
Uncertainty Evaluation Methods for Waste Package Performance Assessment

Prepared by
Y.-T. Wu/SwRI
A. G. Journel/SU
L. R. Abramson/NRC
P. K. Nair/CNWRA

**Center for Nuclear Waste Regulatory Analyses
Southwest Research Institute**

**Applied Earth Sciences Department
Stanford University**

**Prepared for
U.S. Nuclear Regulatory Commission**

AVAILABILITY NOTICE

Availability of Reference Materials Cited in NRC Publications

Most documents cited in NRC publications will be available from one of the following sources:

1. The NRC Public Document Room, 2120 L Street, NW, Lower Level, Washington, DC 20555
2. The Superintendent of Documents, U.S. Government Printing Office, P.O. Box 37082, Washington, DC 20013-7082
3. The National Technical Information Service, Springfield, VA 22161

Although the listing that follows represents the majority of documents cited in NRC publications, it is not intended to be exhaustive.

Referenced documents available for inspection and copying for a fee from the NRC Public Document Room include NRC correspondence and internal NRC memoranda; NRC Office of Inspection and Enforcement bulletins, circulars, information notices, inspection and investigation notices; Licensee Event Reports; vendor reports and correspondence; Commission papers; and applicant and licensee documents and correspondence.

The following documents in the NUREG series are available for purchase from the GPO Sales Program: formal NRC staff and contractor reports, NRC-sponsored conference proceedings, and NRC booklets and brochures. Also available are Regulatory Guides, NRC regulations in the *Code of Federal Regulations*, and *Nuclear Regulatory Commission Issuances*.

Documents available from the National Technical Information Service include NUREG series reports and technical reports prepared by other federal agencies and reports prepared by the Atomic Energy Commission, forerunner agency to the Nuclear Regulatory Commission.

Documents available from public and special technical libraries include all open literature items, such as books, journal and periodical articles, and transactions. *Federal Register* notices, federal and state legislation, and congressional reports can usually be obtained from these libraries.

Documents such as theses, dissertations, foreign reports and translations, and non-NRC conference proceedings are available for purchase from the organization sponsoring the publication cited.

Single copies of NRC draft reports are available free, to the extent of supply, upon written request to the Office of Information Resources Management, Distribution Section, U.S. Nuclear Regulatory Commission, Washington, DC 20555.

Copies of industry codes and standards used in a substantive manner in the NRC regulatory process are maintained at the NRC Library, 7920 Norfolk Avenue, Bethesda, Maryland, and are available there for reference use by the public. Codes and standards are usually copyrighted and may be purchased from the originating organization or, if they are American National Standards, from the American National Standards Institute, 1430 Broadway, New York, NY 10018.

DISCLAIMER NOTICE

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, or any of their employees, makes any warranty, expressed or implied, or assumes any legal liability of responsibility for any third party's use, or the results of such use, of any information, apparatus, product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights.

Uncertainty Evaluation Methods for Waste Package Performance Assessment

Manuscript Completed: November 1990
Date Published: January 1991

Prepared by
Y.-T. Wu, Southwest Research Institute
A. G. Journel, Stanford University
L. R. Abramson, U.S. Nuclear Regulatory Commission
P. K. Nair, Center for Nuclear Waste Regulatory Analyses

Center for Nuclear Waste Regulatory Analyses
Southwest Research Institute
6220 Culebra Road
San Antonio, TX 78228

Applied Earth Sciences Department
Stanford University
Stanford, CA 94305

Prepared for
Division of High-Level Waste Management
Office of Nuclear Material Safety and Safeguards
U.S. Nuclear Regulatory Commission
Washington, DC 20555
NRC FIN D1035

ABSTRACT

This report identifies and investigates methodologies to deal with uncertainties in assessing high-level nuclear waste package performance. Four uncertainty evaluation methods (probability - distribution approach, bounding approach, expert judgment, and sensitivity analysis) are suggested as the elements of a methodology that, without either diminishing or enhancing the input uncertainties, can evaluate performance uncertainty. Such a methodology can also help identify critical inputs as a guide to reducing uncertainty so as to provide reasonable assurance that the risk objectives are met. This report examines the current qualitative waste containment regulation and shows how, in conjunction with the identified uncertainty evaluation methodology, a framework for a quantitative probability-based rule can be developed that takes account of the uncertainties. Current U.S. Nuclear Regulatory Commission (NRC) regulation requires that the waste packages provide "substantially complete containment" (SCC) during the containment period. The term "SCC" is ambiguous and subject to interpretation. This report, together with an accompanying report that describes the technical considerations that must be addressed to satisfy high-level waste containment requirements, provides a basis for a third report to develop recommendations for regulatory uncertainty reduction in the "containment" requirement of 10 CFR Part 60.

TABLE OF CONTENTS

	Page
ABSTRACT	iii
EXECUTIVE SUMMARY	ix
NOMENCLATURE	xiii
1. INTRODUCTION	1
1.1 Background	1
1.2 Objectives and Scope	3
1.3 Report Organization	3
2. WASTE-PACKAGE PERFORMANCE ASSESSMENT	5
2.1 Introduction	5
2.2 Waste-Package Failure Modes	5
2.2.1 General Approach to Assessment of Waste-Package Performance	5
2.2.2 Waste-Package Failure Modes	6
2.2.3 Uncertainties Associated with Failure Modes	7
2.3 Uncertainties in Waste-Package Environments and Materials	7
2.3.1 Uncertainties in Environments	7
2.3.2 Uncertainties in Materials	9
2.4 Probabilistic Performance Assessment and Probabilistic Modeling	9
3. CHARACTERIZATION OF UNCERTAINTIES IN WASTE-PACKAGE PERFORMANCE ASSESSMENT	12
3.1 Introduction	12
3.2 Types of Uncertainty	12
3.2.1 Random Uncertainty	12
3.2.2 Knowledge Uncertainty	13
3.3 Sources of Uncertainty	13
4. WASTE-PACKAGE UNCERTAINTY EVALUATION METHODOLOGY	15
4.1 Uncertainty Evaluation and Uncertainty Reduction	15
4.1.1 Uncertainty Evaluation	15
4.1.2 Uncertainty Reduction	17
4.2 Uncertainty Evaluation Methods	17
4.3 Probability-Distribution Approach	20
4.4 Bounding Approach	21
4.5 Expert Judgment	22
4.5.1 Selection of Issues and Experts	23
4.5.2 Presentation of Issues to the Experts	24
4.5.3 Preparation of Issue Analyses by the Experts	24
4.5.4 Discussion of Issue Analyses	24
4.5.5 Elicitation Training and Elicitation	24
4.5.6 Recomposition and Aggregation of Results	25
4.5.7 Documentation	25
4.6 Sensitivity Analysis	26

5. A PROBABILISTIC FRAMEWORK FOR ASSESSING CONTAINMENT PERFORMANCE	28
5.1 Qualitative vs. Quantitative Criteria	28
5.2 Criteria for a Quantitative Rule	28
5.3 An Example Framework for a Probability-Base Rule for Substantially Complete Containment (SCC)	29
6. SUMMARY	35
7. REFERENCES	36
APPENDICES	
A. UNCERTAINTY MODELING METHODS - A SELECTED REVIEW	A-1
B. PROBABILITY-DISTRIBUTION APPROACH I - GENERAL	B-1
C. PROBABILITY-DISTRIBUTION APPROACH II - FAST PROBABILITY INTEGRATION AND IMPORTANCE SAMPLING UPDATING	C-1
D. THE BOUNDING APPROACH	D-1
E. MODELING THE TRANSFER FUNCTION	E-1
F. BAYESIAN GEOSTATISTICS (A FORMALISM FOR UPDATING PRIOR SOFT INFORMATION)	F-1

FIGURES

	Page
1. Waste-Package Probabilistic Performance Assessment	10
2. Modeling Sequence for Engineered Barrier System (EBS) Performance Assessment	16
3. Waste-Package Number-of-Failures Distribution	31

TABLES

	Page
1. Examples of Uncertainty Characterization in Waste-Package Performance Assessment	14
2. Uncertainty Evaluation Methods for Waste-Package Reliability Analysis	18

EXECUTIVE SUMMARY

Performance-related U.S. Regulatory Commission (NRC) regulations for high-level waste (HLW) include 10 CFR 60.112 on repository system performance, which implements the Environmental Protection Agency (EPA) requirements related to the containment of high-level waste (HLW). The engineered barrier system (EBS) is a subsystem of the overall repository system. The NRC performance requirements for this subsystem are given in 10 CFR 60.113, which requires that containment of HLW be substantially complete during the period when radiation and thermal conditions in the EBS are dominated by fission product decay. The minimum period of containment is specified as 300 to 1,000 years after permanent closure of the geologic repository. The current NRC regulation requires "substantially complete containment" (SCC). However, the term "SCC" is ambiguous and subject to interpretation. The purpose of this report and two other companion reports is to investigate the feasibility of a possible change in the current regulation to reduce the regulatory ambiguity. The first report¹ is a compendium of technical information and considerations that must be addressed to satisfy containment requirements of HLW. As the second report in this series, this report identifies methodologies to deal with uncertainty issues involved in evaluating containment performance. Together, the two reports provide the basis for a third report², which develops recommendations for resolving the regulatory uncertainty in 10 CFR Part 60.113.

As part of the feasibility study, this report has the key objective of identifying and investigating methodologies to deal with uncertainties in waste package performance assessment, without prejudging whether an SCC rule should be qualitative or quantitative. Although the focus of this report is mainly on waste package performance assessment, the methodologies discussed are also broadly applicable to a wide range of other performance assessment issues. Another objective of this report is to critically evaluate the methodologies discussed. By presenting the strengths and limitations of the methodologies, this report aims to guide the analysts in selecting the appropriate approaches for the uncertainty analysis, as well as conveying useful advice for avoiding pitfalls in the application of the selected methodologies.

As a feasibility study and a companion study to the first report, this report has a limited scope. This report presents methodologies, not analyses. As a compilation of methodologies, this report accepts the input data and models and their associated uncertainties as givens and shows how to evaluate the effects of input uncertainties on the output uncertainty, without either attenuating or amplifying the input uncertainties. The input uncertainties stem from the technical considerations discussed in the first report. The output uncertainty relates to the SCC criterion. However, this report does not deal directly with the quality of the inputs; as with any analysis, the quality of the

¹Manaktala, H.K. and C.G. Interrante, "Technical Considerations for Evaluating Substantially Complete Containment of HLW within the Waste Package," CNWRA 90-001, 1990.

²Nair, P.K. and E. Tschoepe, "Substantially Complete Containment Feasibility Study Recommendations Report," CNWRA Letter Report, 1990.

output can be no better than the quality of the inputs. The tasks of specifying the input uncertainties and choosing, developing, and applying the uncertainty methodologies to determine the output uncertainty are the responsibility of the repository designer, i.e., the Department of Energy (DOE).

The containment performance assessment and uncertainty issues are addressed in Sections 2 and 3. Section 2 describes the scope of uncertainty issues involved in waste package performance assessment and suggests a probabilistic analysis framework by which uncertainties can be treated to address the containment issue. It should be noted that a probabilistic approach, if useful for evaluating performance, can be used regardless of whether an SCC rule is qualitative or quantitative. Section 3 characterizes those uncertainties encountered in waste package performance assessment by uncertainty types and sources. Examples of uncertainty characterization are given in the following table. The purpose of the classification is to provide a systematic way to include all sources of uncertainty and to analyze the uncertainties associated with the SCC issues.

Types of Uncertainty	Sources of Uncertainty	Examples	Possible Effects
Random	Fabrication Geologic characteristics	Grain size, defects Rock formation	Stochastic corrosion Random porosity
Knowledge: Parameter	Insufficient data	Geochemistry	Error in specification
Knowledge: Modeling	Lack of scientific models	Water intrusion scenarios Waste-form/canister interaction	Inaccurate time-to-failure prediction
Knowledge: Completeness	Inadequate knowledge	Long-term radiation effects	Not all failure mechanisms considered

If uncertainty could be reduced to zero, it might be possible to design an EBS that provides complete containment with full (100 percent) confidence. In reality, uncertainty cannot be totally reduced because of the difficulties associated with long-term predictions and many practical constraints such as time and cost factors. To minimize risk, one should attempt to build high confidence by designing-in reliability and by evaluating and reducing uncertainty.

To evaluate uncertainty, one can first use probabilistic and statistical tools or expert judgment to characterize uncertainty and then conduct probabilistic and sensitivity analysis to identify critical inputs, i.e., those parameters whose associated uncertainties are the major components of the output uncertainty. Early identification of these critical inputs is essential as a guide to reducing uncertainty so as to provide reasonable assurance that the performance objectives are met. Uncertainty can be reduced by gathering more data (including field data, lab data, and natural analog data), by improving predictive models using advanced research results and accelerated life tests, and by performing validation analyses or tests. Section 4 suggests several uncertainty evaluation methods that can be jointly used to design such a methodology.

Section 4, together with the appendices, constitutes the bulk of the results of this study. The methods discussed are of a very general nature and are not exhaustive. Section 4 discusses four uncertainty evaluation methods and suggests that a methodology may be designed in such a way that the assurance (or confidence) level in the analysis output reflects the degree of uncertainty in data and models, i.e., the assurance level of the output reliability is low if the input quality is poor. Such a methodology would be useful for providing guidance for reducing uncertainty.

A probability-distribution approach that models all the uncertainties as probability-distributions is not feasible because of the large uncertainties in the input parameters and process models. For evaluating containment performance, other approaches must be used to supplement the probability-distribution approach. Four uncertainty evaluation methods are identified as shown in the following table.

Method	Main Role	Discussions
I. Probability-Distribution	<ul style="list-style-type: none"> • Establish waste package performance distributions (reliability analysis) 	Section 4.3 Appendices B, C, E, F.
II. Bounding	<ul style="list-style-type: none"> • Quantify input uncertainty (Bounds, assurance levels) • Provide assurance levels for results of I 	Section 4.4 Appendix D
III. Expert Judgment	<ul style="list-style-type: none"> • Quantify input uncertainty (Distributional input, assurance levels) • Provide assurance levels for results of I 	Section 4.5
IV. Sensitivity Analysis	<ul style="list-style-type: none"> • Identify critical assumptions, parameters 	Section 4.6

The four methods are not mutually exclusive; they should be jointly used. The probability-distribution approach provides a framework for waste package reliability analysis. When there are not enough data to establish a distribution for an uncertainty parameter, the bounding approach can be used to quantify input uncertainty, using bounds and corresponding assurance levels. In addition, in combination with the probability-distribution approach, the bounding approach can provide assurance levels to the outputs. Expert judgment can be used to assist in developing distribution models and process models. In some cases, expert judgment may be the only practical mechanism to justify (or provide reasonable assurance for) the selection of scenarios, process models, and the values of parameters. When there are insufficient data or knowledge to discriminate between several competitive models or assumptions, sensitivity analyses can be performed to identify critical uncertain parameters, so that efforts can be directed to reduce the uncertainty.

The uncertainty evaluation methods discussed in Section 4 provide a basis for exploring probability-based SCC rules. Section 5 examines the current SCC regulation and provides an example that shows how a framework for a quantitative probability-based rule can be developed that takes account of the uncertainties. The basic idea is to evaluate the assurance level of the waste

package reliability, which defines (in a probabilistic sense) the containment requirement in terms of the allowable number (for example, zero) of waste package failures. The assurance level of the reliability can be evaluated using the bounding approach or the reasonable-assurance approach.

A significant part of this study concerns the improvement of the standard Monte Carlo approach, which might be extremely time-consuming for complex problems such as waste-package reliability assessment. A number of methods are proposed in the appendices that provide potentially significant improvement over the standard Monte Carlo method. These more efficient procedures are suitable for providing cost-effective and accurate "what if" answers to different assumptions, to assist in uncertainty and risk analyses.

NOMENCLATURE

cdf	cumulative distribution function
EBS	engineered barrier system
$E(.)$	expected value
$F(.)$	cdf
$F(p;n,t)$	cdf of $P(n,t)$; $Prob[P(n,t) \leq p]$ (see Eq. 4)
K	WP's failing ratio = N/M
K_o	regulatory limit of K ; (see Eq. 1)
M	total number of waste packages in an EBS
$N(t)$	total number of failed waste packages in an EBS at time t
n	an integer associated with random variable $N(t)$
$Prob(.)$	probability of
$P(n,t)$	cdf of N ; $Prob(N(t) \leq n)$
p	reliability
p^*	reliability estimate
p_o	reliability target (regulatory limit)
SCC	substantially complete containment
T_k	time-to-failure of waste package
T_o	regulatory limit of containment time (see Eq. 1)
TF	transfer function
t	time
W_k	number of failure of k^{th} waste package; 0 or 1
WP	waste package
X_i	input parameters to transfer functions
Y_k	output/performance functions
β	risk associated with reliability (see Eq. 8)
β_o	acceptable level of risk (see Eq. 10b)
Φ	standard normal cdf
ϕ_j	transfer functions in waste package performance assessment



1. INTRODUCTION

1.1 Background

The Center for Nuclear Waste Regulatory Analyses (CNWRA) has developed a systematic approach for evaluating and analyzing the regulations affecting the High Level Radioactive Waste program. The systematic approach is described in detail in a CNWRA document (see CNWRA, 1990). The rationale for the systems approach is to enable the analyst to identify, among other things, specific interrelated regulatory requirements, what must be proven to demonstrate compliance with these regulations, and how one might demonstrate compliance with the requirements. As part of this process, uncertainties in the regulations are also evaluated, and potential uncertainty resolution strategies are explored. Preliminary analysis of regulations related to the performance of waste packages within the engineered barrier system indicates that there is uncertainty with respect to these requirements, and the CNWRA is assessing various uncertainty resolution strategies. This report supports an initial effort to identify an acceptable strategy for reducing the regulatory uncertainty in the meaning of the term "substantially complete containment."

Regulations by the Nuclear Regulatory Commission (NRC) with respect to the disposal of high-level radioactive wastes (HLW) in geologic repositories are set out in the Code of Federal Regulations (CFR), Title 10, Part 60. In the current work on the containment of HLW, two paragraphs in 10 CFR Part 60 are the principal focus, that is, 10 CFR 60.112, "Overall System Performance Objective for the Geologic Repository after Permanent Closure," and 10 CFR 60.113, "Performance of Particular Barriers after Permanent Closure." 10 CFR 60.112 provides the performance requirements for the overall geologic repository system and 10 CFR 60.113 includes the performance requirements for the repository subsystems, including the engineered barrier system and the geologic setting.

As noted in the following paragraphs, the performance requirements for the engineered barrier system consist of two parts: (1) a "containment" requirement for HLW waste packages and (2) a radionuclide release rate limit from the engineered barrier system (EBS). Taken together these two parts are intended to control the release of radioactive materials to the geologic setting and to add confidence that the overall system performance objectives for the repository (i.e., 10 CFR 60.112) will be met.

The performance requirements for the EBS, as contained in 10 CFR 60.113, specify that:

(i) "The engineered barrier system shall be designed so that assuming anticipated processes and events: (A) containment of HLW will be substantially complete during the period when radiation and thermal conditions in the engineered barrier system are dominated by fission product decay; and (B) Any release of radionuclides from the engineered barrier system shall be a gradual process which results in small fractional releases to the geologic setting over long times...."

(ii) "In satisfying the preceding requirements, the engineered barrier system shall be designed, assuming anticipated processes and events, so that: (A) containment of HLW within the waste packages will be substantially complete for a period to be determined by the Commission taking into account the factors specified in 60.113(b) provided, that such period shall be not less than 300 years nor more than 1,000 years after permanent closure of the geologic repository; and (B) The release rate of any radionuclide from the engineered barrier system following the containment period shall not exceed one part of 100,000 per year of the inventory of that radionuclide calculated to be present at 1,000 years following permanent closure, or such other fraction of the inventory as may be approved or specified by the Commission; provided, that this requirement does not apply to any radionuclide which is released at a rate less than 0.1% of the calculated total release rate limit. The calculated total release rate limit shall be taken to be one part in 100,000 per year of the inventory of radioactive waste originally emplaced in the underground facility, that remains after 1,000 years of radioactive decay."

Although the requirement in the regulation for limited release from the EBS in the post-containment period is clearly stated in numerical terms, the coupled requirement for "substantially complete containment" (SCC) during the containment period is ambiguous and subject to interpretation. Accordingly, clarification of the meaning of the containment requirement either by rulemaking or some other uncertainty reduction method is needed. One reason why a rapid resolution of the meaning of this regulation has not occurred is that the problem of demonstrating containment for long periods of time is an unprecedented task.

Three factors make the problem different from those solved in the past; and these, listed as follows, will require the judicious application of advanced and relevant technologies.

(1) Length of time

The required service time (over 300 years) specified in the regulation exceeds the times commonly required in engineering design.

(2) Size

The number of waste packages is very large (about 20,000 to 80,000 are expected) which implies that a full-scale laboratory test is practically impossible.

(3) Inaccessibility

In a closed repository that is not subject to active institutional controls, the waste packages will be inaccessible for inspection during the majority of the service life, which is up to thousands of years. Reasonable assurance of the required long-term performance will require the application of conservative assumptions and/or a scientific understanding, with a high level of confidence, of the effects of time and the environment on a repository.

In an effort to clarify and investigate the feasibility of quantifying the containment regulation, the CNWRA at the Southwest Research Institute is examining the feasibility of a potential rulemaking activity related to containment. As part of this work, three background technical reports have been developed to provide a common technical basis from which a rule can be developed, if appropriate. The first report (Report 1, Manaktala and Interrante, 1990) is intended to present and describe the technical topics that must be considered in assessments of the long-term performance of the waste package. As the second report in this series, this report (Report 2) identifies methodologies for assessing how well the various technical considerations can be evaluated in the face of known and, at times, unknown uncertainties. The third companion report, on the basis of the information provided in the first two reports, contains recommendations for the resolution of the issue of SCC of HLW within the waste package.

1.2 Objectives and Scope

As part of the feasibility study, this report has the key objective of identifying and investigating methodologies to deal with uncertainties in waste-package performance assessment, without prejudging whether an SCC rule should be qualitative or quantitative. Although the focus of this report is mainly on waste package performance assessment, the methodologies discussed are also broadly applicable to a wide range of other performance assessment issues. Another objective of this report is to critically evaluate the methodologies discussed. By presenting the strengths and limitations of the methodologies, this report aims to guide the analysts in selecting the appropriate approaches for the uncertainty analysis, as well as conveying useful advice for avoiding pitfalls in the application of the selected methodologies.

As a feasibility study and a companion study to the first report, this report has a limited scope. This report presents methodologies, not analyses. As a compilation of methodologies, this report accepts the input data and models and their associated uncertainties as givens and shows how to evaluate the effects of input uncertainties on the output uncertainty, without either attenuating or amplifying the input uncertainties. The input uncertainties stem from the technical considerations discussed in the first report. The output uncertainty relates to the SCC criterion. However, this report does not deal directly with the quality of the inputs; as with any analysis, the quality of the output can be no better than the quality of the inputs. The tasks of specifying the input uncertainties and choosing, developing, and applying the uncertainty methodologies to determine the output uncertainty are the responsibility of the repository designer, i.e., the Department of Energy (DOE).

1.3 Report Organization

Sections 2 and 3 address containment performance assessment issues and suggest approaches. Section 2 describes the scope of uncertainty issues involved in waste package performance assessment and suggests a probabilistic performance assessment framework by which uncertainties can be treated to address the containment issue. It should be noted that a

probabilistic approach, if useful for evaluating performance, can be used regardless of whether an SCC rule is qualitative or quantitative. Section 3 characterizes those uncertainties encountered in waste-package performance assessment by uncertainty types and sources.

A probability-distribution approach that models all the uncertainties as probability distributions is an idealized approach for probabilistic performance assessment. To use the probability-distribution approach, implementation problems associated with insufficient data and uncertain transfer functions must be resolved. Without adequate data and models, the result of a reliability analysis would be highly uncertain. Three other uncertainty evaluation methods, namely the bounding approach, expert judgment, and sensitivity analysis, were identified to supplement the probability-distribution approach. Section 4, together with the appendices, constitutes the bulk of the results of this study. The methods discussed are of a very general nature and are not exhaustive. This section presents an uncertainty evaluation methodology that combines the above four methods and suggests that a methodology may be designed in such a way that the assurance (or confidence) level in the analysis output reflects the degree of uncertainty in data and models, i.e., the assurance level is low if the input quality is poor. Such a methodology would be useful for providing guidance to reducing uncertainty.

Section 5 examines the current SCC regulation and provides an example that shows how a framework for a quantitative probability-based rule can be developed that takes account of the uncertainties. The uncertainty evaluation methodology of Section 4 would be used to evaluate the uncertainty associated with whatever SCC rule, either qualitative or quantitative, is adopted. A summary of the report is presented in Section 6.

A significant part of this study concerns the improvement of the standard Monte Carlo approach, which might be extremely time-consuming for complex problems such as waste package reliability assessment. A number of methods, discussed in Appendices A-C, provide potentially significant improvement over the standard Monte Carlo method. These more efficient procedures may be suitable for providing cost-effective and accurate "what if" answers to different assumptions to assist in decision-making.

Some of the methodological techniques presented in this report draw heavily on the relatively new field of geostatistics (e.g., see references in Appendix F). Because it arose out of mining and petroleum engineering, geostatistics has been developed largely outside of the mainstream of statistics, and is unfamiliar to many statisticians. However, because it was developed to deal with complex dependency structures where classical statistical methods are inadequate, geostatistics is highly useful for dealing with the spatial and temporal dependencies involved in waste repositories.

2. WASTE-PACKAGE PERFORMANCE ASSESSMENT

2.1 Introduction

In 10 CFR Part 60, the "waste package" is defined as the waste form (radioactive waste materials and any encapsulating or stabilizing matrix) and any containers, shielding, packing, and other absorbent materials immediately surrounding an individual waste container; the EBS is defined as the waste packages plus the underground facility.

The SCC regulation addresses the ability of all the waste packages within the EBS to "contain" the waste during the period when radiation and thermal conditions in the EBS are dominated by fission product decay.

In this report, a failure is defined as loss of containment function caused by the interaction between environments and waste-package degradation processes. A more explicit definition of a waste-package failure can be advanced only when a design is proposed in which the materials, geometries, etc., are well-defined. The consequence of waste-package failure is the availability of nuclear waste for transport to the geologic setting. However, the amount or rate of release is regulated in a separate requirement that is not discussed in this report.

In Report 1 (Manaktala and Interrante, 1990), the various technical considerations required in making a logical argument for satisfying containment requirements were presented. The higher-order classification of the technical considerations includes environment, materials fabrication, degradation processes, and inspections and monitoring. These considerations, when incorporated into a life prediction methodology for waste packages, become the basis of exercising technical judgment on the long-term performance of the containment barriers.

Recognizing the various types of inherent uncertainties in the technical considerations, the remaining discussions in this section attempt to establish a probabilistic analysis framework by which uncertainties in waste-package performance assessment can be treated to address the SCC issue. To achieve this goal, potential failure modes/mechanisms and uncertainties will be described to provide a basis for probabilistic treatment.

2.2 Waste-Package Failure Modes

2.2.1 General Approach to Assessment of Waste-Package Performance

An assessment of waste-package performance may be considered to involve the following steps:

- Identify anticipated processes and events;
- Identify possible failure modes in the context of anticipated processes and events;
- For each failure mode, identify failure paths (i.e., scenarios, environments and mechanisms leading to failures);
- For each failure path, evaluate time-to-failure for each waste package

(Note: assuming each waste package and associated environment have their own characteristics);

- Compute number of waste-package failures as a function of time;
- Use the result from the above step to determine whether the design provides SCC.

A basic performance measure is time-to-failure, which is a function of environment and material characteristics. To evaluate time-to-failure, it is necessary to identify failure modes, defined here as the physical/chemical processes that result in failure. Once the time-to-failure information is available, the number of waste package failures, as a function of time, can be computed. In this report, the two major performance measures, time-to-failure and number of waste-package failures, are treated as random variables to deal with the uncertainties involved in evaluating containment performance.

2.2.2 Waste-Package Failure Modes

In defining general mechanical failure modes, Collins (1981) suggested that a systematic classification might be devised by which all possible failure modes could be predicted. According to Collins, there are:

four manifestations of failure:

- Material change (metallurgical, chemical, nuclear)
- Elastic deformation
- Plastic deformation
- Rupture or fracture

four failure-inducing agents:

- Reactive environment (chemical, nuclear)
- Temperature
- Force
- Time

and two failure locations:

- Body type
- Surface type

Several technical considerations presented in Report 1 fall in the category of "material change." This is an important aspect of waste containment where degradation processes are slow and long-term in nature. Elastic or plastic deformations are applicable to repositories where creep is an important phenomena. The integrity of waste packages can also be affected by transient "overload" conditions within the repository area. This may be due to events such as seismic/tectonics or volcanic activities. In such cases, the manifestations of failure would include elastic or plastic deformation, and rupture or fracture.

The failure-inducing agents in describing containment involve the long-term and time-dependent geochemical environment, the gamma radiation field, decay heat temperatures, and geo-thermo-mechanical conditions.

The degradation of waste-package performance can be body-related (i.e., bulk material property changes) and/or surface-related (i.e., corrosion, erosion and other surface effects).

The above classification, which may neither be complete, nor fully relevant to the waste-package performance assessment problem, nevertheless does provide a way to define possible failure modes and failure paths. As an example, pitting corrosion is a localized attack that leads to the development of pits or holes that may penetrate the container wall. This failure mode may be induced by time (long-term), temperature, reactive environment, and results in material change leading to a pitting failure.

In reviewing a waste-package design, it is necessary to investigate all the possible failure modes, and also to investigate the scenarios (related to failure-inducing agents) associated with each failure mode. Standard risk analysis techniques such as fault-trees and failure-modes-and-effects-analysis can be used in defining failure path and failure consequence.

Examples of the important failure modes that need to be reviewed in waste-package performance assessment are corrosion, radiation damage, and force and/or temperature-induced deformation or fracture (e.g., by seismic force). A more detailed description of the conceivable failure modes can be found in Report 1.

2.2.3 Uncertainties Associated with Failure Modes

As discussed above, there are failure-inducing agents for each failure mode. Uncertainties are anticipated in these failure-inducing agents, as described in Section 2.3. Because failure modes will depend on design parameters such as the selected material, geometry, strength, location, etc., it is difficult to postulate which failure modes and uncertainties will be important. Section 2.3 attempts to describe those uncertainties that are anticipated or that have significant impacts on the overall performance.

2.3 Uncertainties in Waste-Package Environments and Materials

2.3.1 Uncertainties in Environments

Waste-package environments are defined here as time-dependent physical/chemical conditions imposed either externally or internally on a waste package (or any of the barriers) and having the potential of affecting the performance to the extent of causing a waste-package failure. Note that a barrier, such as a packing, has a resistance property, but also provides an environmental condition for other barriers.

Major waste-package environments that may impact the waste-package performance include the following elements:

- Geochemical condition
- Thermal field
- Radiation field
- Fluid flow field
- Stress and seismic fields
- Surrounding medium: rocks and other barriers

The geochemistry of a repository defines important environmental conditions that initiate and/or propagate material degradation processes such as corrosion. The geochemistry characteristics surrounding an EBS could be influenced by the gamma radiation or release of metal ions from other components of the EBS (Manaktala and Interrante, 1990). Thus, the uncertainties in geochemical characteristics are strongly dependent on long-term predictive models.

The thermal field around waste packages is expected to be dependent on the type and amount of nuclear waste. In the case of an unsaturated medium around EBS, the thermal field has a potential effect of keeping liquid water from contacting the waste packages in the first several hundred years after permanent closure. However, once the temperature drops to below boiling point, the water may be in contact with canisters. The thermal field may also enhance the concentration of salts and other corrosive elements in the waste package vicinity. The uncertainties in predicting the thermal field stem mainly from the difficulty in characterizing the thermal properties of the surrounding medium.

The waste package radiation field is determined by the waste package design and the contained nuclear waste. The radiation field may influence the characteristics of the surrounding materials and cause changes in the degradation rates. The characteristics of the radiation field have relatively smaller uncertainty; however, the radiation effect on the materials may have greater uncertainty that can be reduced only by physical understanding of the processes.

The fluid flow field interacts with geochemical conditions and provides a path for potentially hostile chemicals to contact and degrade the waste packages. The locations and amount of water and vapor depend on many factors such as climate, water table, and water flow characteristics of the rock matrix. The uncertainties are dependent on long-term predictive models such as recharge rates and groundwater flow models.

The effects of the stress and seismic fields depend on waste package and EBS design and repository geologic characteristics. The sources of stresses include those due to welding, sealing, and other fabrication processes; and other mechanical forces imposed upon the waste packages by, for example, the collapse of emplacement boreholes due to seismic

shaking or fault motions by an earthquake. The stress field may contribute to corrosion failure and other mechanical failures such as impact failure or buckling failure. To some extent, the fabrication-induced stress can be controlled by proper quality assurance procedures. The stresses induced by seismic shaking and fault motions, on the other hand, are difficult to predict and are, hence, subject to greater uncertainty. A practical solution today is to use a probabilistic model to simulate the occurrence and effects of such events.

The surrounding medium includes rocks, soils, and other barriers that affect the geochemical, thermal and/or fluid flow fields. Uncertainties in material characteristics are anticipated for nonhomogeneous medium such as rock.

In summary, the nature of the waste package environments involves many uncertainties. In general, the uncertainties involve physical variability such as porosity in a rock, and knowledge uncertainty (see Section 3.2) such as tectonic and seismic activities and climate changes on time scales of thousands of years or more.

2.3.2 Uncertainties in Materials

As defined earlier, a waste package consists of the waste form and any containers, shielding, packing, and other absorbent materials immediately surrounding an individual waste container. Detailed definition of uncertainties in waste-package materials will be possible only after materials have been selected.

Material defects may exist in most engineering materials. Although initial (after inspection) defects can be controlled to some extent, "small" defects may be unavoidable because of imperfect manufacturing processes and/or of detection limitations of non-destructive testing methods.

In general, the uncertainties that could impact performance assessment include such random characteristics as material grain size and orientation, location of defects or voids, distribution of residual stresses, and variations in waste form characteristics (amount and type) among waste packages. Scientific uncertainty is expected to exist in the long-term predictive models of material properties. The importance of each mentioned uncertainty depends on waste-package design and potential failure modes.

2.4 Probabilistic Performance Assessment and Probabilistic Modeling

According to Sections 2.2 and 2.3, waste-package performance assessment requires long-term prediction/modeling of environments and material-related processes that are uncertain. This section describes a probabilistic analysis framework and presents several examples of probabilistic modeling that are useful for dealing with the uncertainties.

Figure 1 shows a simplified flow-chart for waste-package probabilistic performance assessment. In this figure, the transfer functions ϕ_j are idealized mathematical models representing real world physical processes that link the inputs to the outputs. The transfer functions are likely to be very complicated and would require sophisticated computer modeling. The input parameters X_i and the transfer functions are generally uncertain.¹ Both the inputs and the transfer functions are in general functions of time.

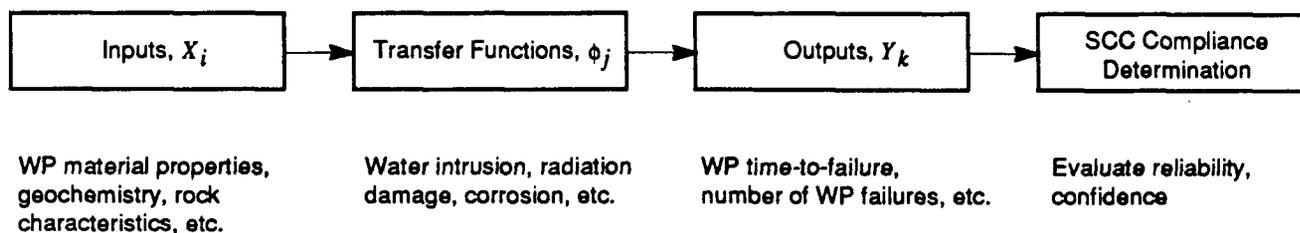


Figure 1. Waste-package (WP) probabilistic performance assessment.

To evaluate uncertainty, the input uncertainties are propagated through the transfer functions to the outputs Y_k . As a result of the input uncertainties, the two major performance measures, time-to-failure and number of waste package failures, are uncertain. Uncertainty evaluation methods that may be used for probabilistic performance assessment are discussed in Section 4.

Probabilistic performance assessment requires proper probabilistic modeling of a variety of uncertainties (see Section 3). For example, the initial defect size might be modeled by a probability distribution that can be developed based on empirical probability detection curves (Woo and Simonen, 1984). Some material degradation processes have been represented as random and modeled probabilistically. For example, the generation of pits on stainless steel can be assumed to occur randomly in time and space, and it may be modeled as a stochastic Markov process (Sato, 1976; Provan and Rodriguez, 1989; Shibata and Takeyama, 1977; Rodriguez and Provan, 1989). In such cases, probabilistic predictive models may be established to predict random pit growth. Similarly, because the deepest pit (rather than the overall corrosion rate or the number of pits) may cause containment failure, an extreme value statistical model is

¹ A possible approach to modeling an uncertain transfer function is to include random variables in a transfer function, to account for scientific or modeling uncertainty. In such cases, these random variables are also treated as inputs.

useful for modeling the time-to-failure (Provan and Rodriguez, 1989). In general, the time-to-failure prediction models should be built based on physical formulations and/or experimental results.

Probabilistic/statistical models may be useful for establishing life prediction models based on accelerated testing and/or natural analogs. For example, an accelerated stress corrosion cracking test method has been proposed by Sato and Murata (1989) to simulate changes in environments, to evaluate material service life. They examined the time-to-failure distributions due to different test conditions, using a Weibull distribution in which the "shape parameter," m , was used to identify the failure type, i.e., "start up" ($m < 1$), "incidental" ($m = 1$) or "wear-out" ($m > 1$). If m can be estimated based on natural analogs or other experimental evidence/models, then m is an useful index for suggesting an accelerated test method for developing a long-term life prediction model.

3. CHARACTERIZATION OF UNCERTAINTIES IN WASTE-PACKAGE PERFORMANCE ASSESSMENT

3.1 Introduction

Section 2 suggests a probabilistic approach to waste-package performance assessment. This section presents a method of characterizing the uncertainties so that the required probability-based uncertainty evaluation methods can be identified.

In this section, uncertainties will be classified into two major types: uncertainty due to random variability and uncertainty due to lack of knowledge about parameters, models, or phenomenologies. Also, examples of uncertainty sources associated with each uncertainty type will be given to demonstrate how to classify waste package uncertainties. Uncertainty evaluation methods that might be suited to treat each type of uncertainty are presented in Section 4.

Uncertainties can be classified in a number of ways. However, for this report, the purpose of classification is to provide a systematic way of analyzing the certainties associated with the SCC issues. Therefore, precise taxonomy is not crucial. Rather, it is more important that all sources of uncertainty can be included so that proper uncertainty evaluation methods can be identified or developed.

3.2 Types of Uncertainty

There are two major types of uncertainty: (1) uncertainty due to random variability, and (2) uncertainty due to lack of knowledge about parameters, models, or underlying physical processes. The essential difference between these two types is that an enlargement of the data base can decrease knowledge uncertainty, but can have no effect on random uncertainty.

The following definitions are drawn mainly from USNRC (1983) and from Vesely and Rasmuson (1984). The former contains a comprehensive treatment of uncertainty, and the latter is a useful overview.

3.2.1 Random Uncertainty

Random uncertainty is uncertainty due to experimental error or inherent variability in some measured physical quantity. This type of uncertainty results when an experiment is repeated under identical conditions, and different outcomes are observed. An example is the failure times for a group of "identical" waste packages subject to "identical" experimental conditions. Random uncertainty is inherent in the physical process involved; it cannot be reduced by enlarging the data base. However, enlarging the data base can provide information about the probability distribution of the random uncertainty, but this information reduces the knowledge uncertainty, not the random uncertainty. For the SCC applications, the random uncertainties will tend to be dominated by the knowledge uncertainties because of the need to predict complicated long-term processes.

In recent years, researchers have become increasingly aware that many deterministic, nonlinear systems can exhibit irregular, unpredictable, and seemingly random behavior. This phenomenon is termed "chaos" (Jensen, 1987). Because chaotic behavior is not affected by increased knowledge, uncertainty stemming from such behavior will be considered random uncertainty.

3.2.2 Knowledge Uncertainty

There are three types of knowledge uncertainty: *parameter*, *modeling*, and *completeness*.

Parameter uncertainties refer to imprecisions and inaccuracies in input parameters, e.g., chloride concentration in the waste-package environments. Parameter uncertainties can be reduced by parameter estimation, using appropriate data. However, the estimation process itself can introduce additional uncertainties if the data are biased, incomplete, or irrelevant, or if an invalid data analysis is used. Additional uncertainties can be introduced if expert judgment is used for parameter estimation. All of these types of uncertainty are considered parameter uncertainties.

Modeling uncertainties refer to uncertainties in a probability model (e.g., a parameterized distribution model) of a random uncertainty, and uncertainties in the applicability and precision of an idealized physical model, stemming from its inadequacies in representing reality, e.g., an empirical corrosion model based on short-term experimental data. Since all models are imperfect representations of reality, all models are subject to a greater or lesser degree to modeling uncertainty. This stems from uncertainty as to whether the model accounts for all the variables that affect the results and from uncertainties in the relations and descriptions used in the model. Modeling uncertainty can also be introduced by applying the model beyond its range of validity.

Completeness uncertainties refer to uncertainties as to whether all the significant phenomena, relationships, and future states have been considered, e.g., whether all scenarios that could significantly affect SCC have been identified. Completeness uncertainties are similar in nature to modeling uncertainties, but occur at the initial stage in the analysis. In addition to inadequate identification of the physical phenomena involved, completeness uncertainties can also result from inadequate consideration of human error, software reliability, or interactions and dependencies among the elements of the processes being modeled.

3.3 Sources of Uncertainty

The sources of the *random* uncertainty involved in a waste-package performance assessment include material fabrication processes. The resulting random variables include grain size, defect size, and defect locations, etc.

The sources of *parameter* uncertainty include insufficient data or knowledge, indirect data (e.g., based on analogs), etc. The sources of *modeling* uncertainty include lack of scientific models and/or experimental data for making long-term predictions. The sources of *completeness* uncertainty include inadequate knowledge to predict long-term events.

Another source of uncertainty is human error. Human error could result in random uncertainty (e.g., crack size due to imperfect inspection or mishandling of equipment) and knowledge uncertainty (e.g., calculation error, modeling error, or lack of completeness).

Table 1 provides some examples of uncertainty sources and their possible effects on waste package performance. In general, it is anticipated that all uncertainties in waste package performance assessment can be classified according to Section 3.2.

Table 1. Examples of Uncertainty Characterization in Waste-Package Performance Assessment

Types of Uncertainty	Sources of Uncertainty	Examples	Possible Effects
Random	Fabrication Geologic characteristics	Grain size, defects Rock formation	Stochastic corrosion Random porosity
Knowledge: Parameter	Insufficient data	Geochemistry	Error in specification
Knowledge: Modeling	Lack of scientific models	Water intrusion scenarios Wasteform/canister interaction	Inaccurate time-to-failure prediction
Knowledge: Completeness	Inadequate knowledge	Long-term radiation effects	Not all failure mecha- nisms considered

4. WASTE-PACKAGE UNCERTAINTY EVALUATION METHODOLOGY

4.1 Uncertainty Evaluation and Uncertainty Reduction

4.1.1 Uncertainty Evaluation

Referring to Figure 1, the input uncertainties X_i are propagated through the transfer functions to the outputs Y_k . The uncertainty evaluation methods discussed in this section accept parameters and transfer functions (and their associated uncertainties) as given. Major Y_k of interest are waste package (WP) time-to-failure (denoted by T_k) and the total number of WP failures in time t (denoted by $N(t)$). Both T_k and $N(t)$ are random variables.

Figure 2 is a flow-chart showing the sequence of process models required to assess compliance with EBS regulations, which separates containment and radionuclide release rate models. For illustration, an EBS surrounded with rock medium is assumed in this flow chart. A transfer function (TF) which could either be deterministic or stochastic replaces one or more process models.

Remarks on the flow-chart:

- (1) "Inward" refers to the WPs; "outward" refers to the biosphere.
- (2) As a simplification, the same groundwater conceptual model is used to assess both the external degradation of WPs and the transport of released radionuclide to the biosphere (TF5).
- (3) The groundwater model is coupled with a rock mechanics model because development of major fracture and/or major collapse could influence flow paths both inward (TF1) and outward (TF5).
- (4) The aggregation model (TF4) accounts for space and time dependence due to the spatial arrangement of WPs in the repository.
- (5) Calculation of the release rate, i.e., compliance to the release rate regulation, is conditioned by the multiple WPs failure model.

The transfer functions are generally coupled (dependent); however, the flow-chart suggests several decoupling possibilities to aid feasibility of the global assessment task:

- (1) Decoupling of the external and internal degradation models (transfer functions TF1 and TF2) until the two processes combine, resulting in actual failure (TF3).
- (2) Decoupling of the single WP and multiple WPs failure models, allowing for consideration of alternative space-time failure-dependence scenarios, without having to evaluate the entire modeling sequence.
- (3) Decoupling of the containment (SCC) and release rate issues.

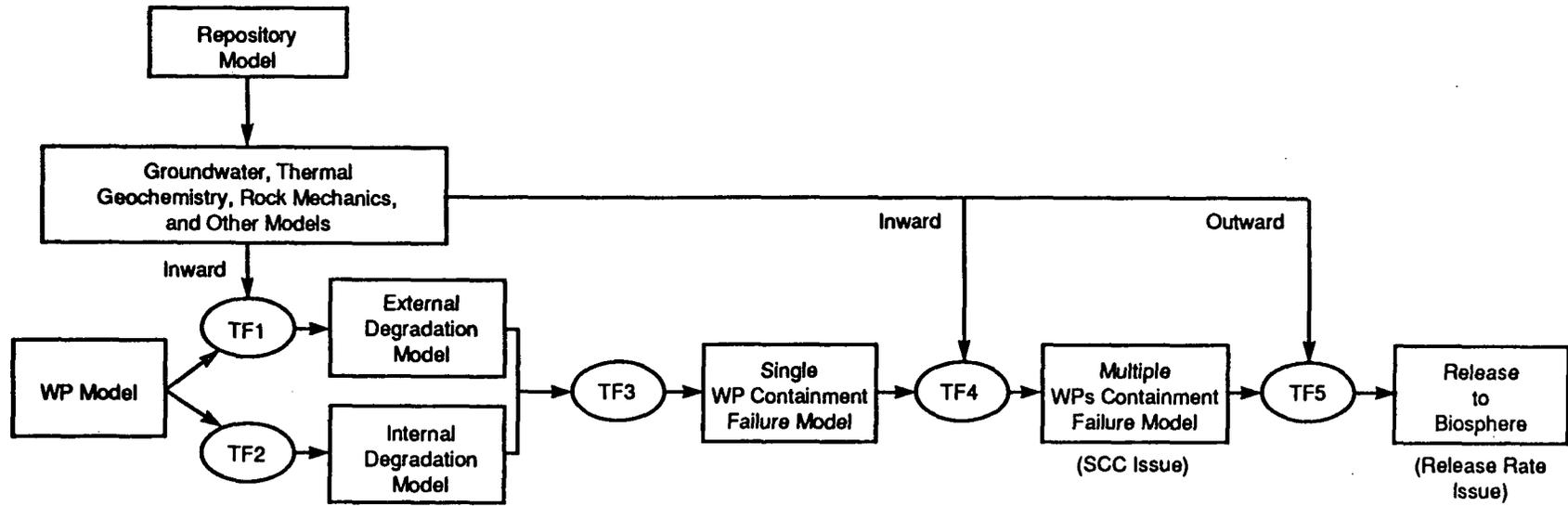


Figure 2. Modeling sequence for EBS performance assessment.

4.1.2 Uncertainty Reduction

If uncertainty could be reduced to zero, it might be possible to design an EBS that provides complete containment, with full (100 percent) confidence. In reality, uncertainty can not be totally reduced because of the difficulties associated with long-term predictions and many practical constraints, such as time and cost factors. To minimize risk, one should attempt to build high confidence by designing-in reliability and by evaluating and reducing uncertainty.

To evaluate the uncertainty, one can first use probabilistic and statistical tools or expert judgment to characterize the uncertainty and then conduct probabilistic and sensitivity analysis to identify critical inputs, i.e., those parameters or models whose associated uncertainties are the major components of the output uncertainty. Early identification of these critical inputs is essential as a guide to the allocation of resources for reducing uncertainty.

Uncertainty can be reduced by gathering more data (including field data, lab data, and natural analog data), by improving predictive models using advanced research results and accelerated life tests, and by performing validation analyses or tests.

Collection of relevant data and development of transfer functions involve many specialized fields and require numerous experts. The scope of this report is not to directly deal with the quality of the input data or transfer functions, nor is it to provide in-depth suggestions for reducing uncertainty. Rather, the goal is to identify or develop uncertainty evaluation methods that are useful for evaluating uncertainty, so as to provide reasonable assurance that the risk objectives are met.

4.2 Uncertainty Evaluation Methods

A selected review of some uncertainty evaluation methods applicable to waste package performance assessment is documented in Appendix A. Several uncertainty methodologies are outlined in Appendices B through F. The applicability of these methodologies to the SCC issue is, as yet, unproven. Although we believe that these methodologies provide a sound basis for an uncertainty evaluation, no claim is made that they are the only appropriate ones.

As with any analysis, the quality of the output can be no better than the quality of the inputs or the models (transfer functions). Without adequate data and models, the result of a reliability analysis would be uncertain. Nevertheless, an uncertainty evaluation methodology can be designed in such a way that the confidence in the output reflects the degree of uncertainty in data and models, i.e., the assurance level is low if the input quality is poor. This section suggests several methods that can be jointly used to design such a methodology.

An idealized approach for probabilistic performance assessment is the probability-distribution approach. This approach models all the input and transfer function uncertainties as probability distributions and calculates the resultant uncertainties of the results. However, there are serious implementation problems associated with insufficient data and uncertain transfer functions, which make the probability-distribution approach impractical for most applications. Therefore, in many engineering problems where safety is a concern, it is common to use the safety factor approach to provide "safety margins" or reliability for a design (Collins, 1981). To account for uncertainties, typically a number of safety factors are applied to the characteristic or nominal values of the major design parameters such as material properties and loads. This approach is simple to implement, but the selection of the safety factors is often experience-based and, more critically, the resulting reliability can not be quantified.

Since an experience base for long-term waste disposal does not exist, the safety factor approach should be used with caution. Unless assurance levels can be established for selected safety factors, this approach is not recommended for addressing the SCC regulation.

The major difficulties associated with the probability-distribution approach include insufficient data or knowledge for establishing distributions and the difficulties in quantifying the uncertainties in transfer functions. To deal with these problems, other approaches must be used to supplement the probability-distribution approach. Four uncertainty evaluation methods are identified in Table 2. The methods suggested are of a very general nature and are not exhaustive. They are potentially useful tools to deal with uncertainty issues in SCC. Most of the detailed technical discussions supporting these methods are included in the appendices.

Table 2. Uncertainty Evaluation Methods for Waste-Package Reliability Analysis

Method	Main Role	Discussions
I. Probability-Distribution	<ul style="list-style-type: none"> Establish WP performance distributions (reliability analysis) 	Section 4.3 Appendices B, C, E, F.
II. Bounding	<ul style="list-style-type: none"> Quantify input uncertainty (Bounds, assurance levels) Provide assurance levels for results of I 	Section 4.4 Appendix D
III. Expert Judgment	<ul style="list-style-type: none"> Quantify input uncertainty (Distributional input, assurance levels) Provide assurance levels for results of I 	Section 4.5
IV. Sensitivity Analysis	<ul style="list-style-type: none"> Identify critical assumptions, parameters 	Section 4.6

The probability-distribution approach provides a framework for waste-package reliability analysis. When there are not enough data to establish a distribution for an uncertainty parameter, the bounding approach can be used to quantify input uncertainty, using bounds and corre-

sponding assurance levels. Expert judgment can be used to assist in developing distribution models and transfer functions. In many cases, expert judgment may be the only practical mechanism to justify (or provide reasonable assurance for) the selection of scenarios, process models, and the values of parameters. When there are insufficient data or knowledge to discriminate between several competitive models or assumptions, sensitivity analyses can be performed to identify critical uncertain parameters, so that efforts can be directed to reduce the uncertainty.

The four approaches are not mutually exclusive. For example, the bounding approach can be used in conjunction with the probability-distribution approach, and expert judgment can be used to provide the necessary input so that the probability-distribution approach can be applied.

The above approaches can be applied according to the types of uncertainties identified in Section 3.

For any input variable or process, the treatment of random uncertainty involves the specification of a probabilistic model for the inherent physical uncertainty in the variable or process. This probabilistic model can be based on theoretical considerations, experiments (including accelerated tests as discussed in Section 2.4) or natural analogs. These can be complemented with expert judgment. If the model can be completely specified, then the treatment of random uncertainty for the input variable or process is complete. However, if there is any significant uncertainty about the choice of model or the values of model parameters, then there is knowledge uncertainty as well as random uncertainty.

Three types of knowledge uncertainty were identified in Section 3.2 - parameter, modeling and completeness. Modeling and completeness uncertainty can best be handled by the formal use of expert judgment. However, if there is significant uncertainty about a critical model (a model whose choice can significantly affect the results), then a sensitivity analysis in conjunction with expert judgment can be used to evaluate different models. Another approach is to bound the uncertainty by using conservative models.

The treatment of parameter uncertainty admits the largest variety of approaches. These approaches can be characterized by the amount of data available for the input variables and parameters. The method requiring the most data is the probability-distribution approach. If only bounding data are available, then the bounding approach can be used. Expert judgment can be used to provide inputs for the probability distribution or bounding approach when insufficient data are available.

Development of a detailed methodology for uncertainty evaluation and its implementation is possible only when a conceptual repository design is available and the waste-package design parameters are well-defined. Actual implementation corresponding to particular data and transfer functions would be the responsibility of the repository designer. As a general guideline,

the SCC problem must be approached in terms of its unique attributes and specific constraints. A collection of approaches such as those reviewed in Appendix A and those discussed in this section must be assembled. New approaches should be developed as needed.

4.3 Probability-Distribution Approach

Referring to Figure 1, a probability-distribution approach would consist of assigning probability distributions to all uncertain inputs X_i and propagating these distributions using the transfer functions ϕ_j to derive the probability distributions of the outputs Y_k . In practice, this approach calls for a global simulation of the entire sequence described in Figure 2, to derive the probability distributions for the key performance measures, including time-to-failure $T_k(k = 1, \dots, M)$ and total number of waste package failures $N(t)$. Here M is the total number of WPs.

Let $W_k(k = 1, \dots, M)$ be the number of failures of the k^{th} waste package. At a given time t , the random variable W_k can take only two values: 0 or 1, with corresponding probabilities determined from the probability distribution of T_k . Thus, the total number of failures is: $N(t) = \sum W_k = W_1 + W_2 + \dots + W_M$, in which every W_k is a random variable that is a function of input random variables X_i .¹

Ideally, given all the transfer function and the joint probability-distribution-function of the input random variables, a standard Monte Carlo simulation can be performed to obtain the distributions of T_k , W_k , and finally $N(t)$.²

The standard Monte Carlo approach would consist of repetitive sampling of the input uncertainty space and transferring of each sample into the output value (or outcome), thus allowing a sampling of the $N(t)$ -output distribution - a task that may be prohibitively costly. To alleviate this problem, appendices B, C, and E suggest some practical approximate computational methods.

Appendix B discusses important limitations of the standard Monte Carlo approach and proposes several alternative Monte Carlo simulation techniques, with emphasis on reducing the sampling effort. In particular, an efficient importance sampling scheme is proposed.

¹ Remark: In general, the random variables $W_k(k = 1, 2, \dots, M)$ are not independent because the realizations of some common input random variables X_i may affect multiple waste packages simultaneously. On the other hand, the random variables W_k are not completely dependent because of the differences in locations, materials, etc., among different WPs.

² In Section 5, the distribution of $N(t)$ is used to develop a framework for a quantitative probability-based rule.

Appendix C proposes a fast probability analysis method that employs a fast probability integration concept combined with an importance sampling scheme. The importance sampling method implemented in this appendix relates to the more general approach discussed in Appendix B. For complex transfer functions, Appendices B and C provide potentially significant improvement over the standard Monte Carlo method.

Appendix E reviews and proposes general principles for decoupling and streamlining sequences of transfer functions without altering their ability to generate "extreme-valued responses" that are important for waste package reliability analysis. Note that "extreme-valued responses" means responses beyond the performance compliance limit, with usually a low probability of occurrence.

Another problem associated with the probability-distribution approach is that the available information might be insufficient to define probability distributions. Section 4.4 and Appendix D discuss the bounding approach that can be used to supplement the probability-distribution approach. Section 4.5 discusses the expert judgment approach that may be useful to assist in defining input probability distributions. Appendix F proposes a methodology for integrating soft¹ information.

4.4 Bounding Approach

In order to implement the probability-distribution approach, it is necessary to specify the joint distribution of all input variables. If this can be done, then, in principle, the joint distribution of the output variables can be calculated (see Appendix B). Although there may be sufficient theoretical foundations or experimental data to specify the distribution of many input variables, it is likely that, for some input variables, no such information exists. Although some information about the distribution of such a variable X may exist based on theory, data, or expert judgment, this information may be too vague to justify assigning a distribution to X . However, in such a case, it may be possible to characterize X as lying in some specified interval (called an uncertainty interval) with high assurance. It is expected that the uncertainty interval and the level of assurance will be specified by expert judgment (see Section 4.5).

Assume that the set of input variables can be divided into two subsets -- \underline{X}' and \underline{X}'' , where $\underline{X}' = \{X_1, \dots, X_m\}$ is a set of input variables with joint distribution $F(x_1, \dots, x_m) = \text{Prob}\{X_1 \leq x_1, \dots, X_m \leq x_m\}$, and $\underline{X}'' = \{X_{m+1}, \dots, X_L\}$ is a set of input variables with uncertainty intervals $[a_{m+1}, b_{m+1}], \dots, [a_L, b_L]$. The bounding approach combines the input distribution F for \underline{X}' with the uncertainty intervals for \underline{X}'' to calculate upper and lower bounds for the joint

¹As opposed to hard data, "soft" data correspond to local geophysical information that needs updating.

distribution of the output variables $\underline{Y} = \{Y_1, \dots, Y_k\}$, without diminishing or enhancing the input uncertainties. A probabilistic inequality is used to relate the assurance level of the output bounds to the assurance levels of the input uncertainty intervals (see Appendix D).

A note of caution is in order. When an input variable X is characterized by an uncertainty interval $[a, b]$, it is assumed that not enough is known about X to justify specifying a distribution for it. Consequently, it would not be appropriate to replace the uncertainty interval with some fitted distribution. For example, some analysts might be tempted to replace an uncertainty interval $[a, b]$ by a distribution for X , e.g., a lognormal, where $Prob\{a \leq X \leq b\}$ is set equal to the assurance associated with the uncertainty interval. This would be inappropriate for several reasons. First, it would be substituting the analyst's judgment for the experts' judgment. Second, the use of any distribution implies an averaging process, which could lead to less uncertainty in the output distribution of Y than is implied by the uncertainty interval for X . Third, the tail behavior of the fitted distribution might possibly dominate the effect of X on the output distribution of Y , and this would violate the implicit judgment of the experts that nothing can be said about the values of X outside of the uncertainty interval. For all of these reasons, only the bounding approach as outlined in Appendix D is recommended to evaluate the effect of input variables in \underline{X} .

4.5 Expert Judgment

The use of expert judgment (often referred to as expert opinion) is an essential aspect of uncertainty modeling for SCC rulemaking. Three areas of particular concern are scenario development, model development, and parameter estimation. Expert judgment is needed in scenario development to identify, classify, and screen events and processes, to formulate and screen scenarios, and to estimate their relative ranking and probabilities of occurrence. Expert judgment is needed in model development to select and interpret data, to develop the conceptual models, and to build confidence in the models and codes. Expert judgment is needed in parameter estimation to identify important parameters and to quantify their uncertainty. Specific techniques for the formal elicitation and use of expert judgment in these areas are discussed in Bonano *et al.* (1990).

The formal use of expert judgment has been extensively applied to a number of recent major studies in the nuclear probabilistic risk assessment area [USNRC (1989), Bernreuter *et al.* (1989), and Risk Engineering, Inc., *et al.* (1989)]. Although scientific inquiry and decision-making have always relied on expert judgment, the formal use of expert judgment as a well-documented systematic process is a relatively new development. It has been necessitated by the need to address questions where alternative sources of information are unavailable, less reliable, or too costly. However, because of the many potential pitfalls in using expert judgment, it is essential that analysts must be familiar with the state of the art and must use the services

of experienced practitioners in order to avoid wasting time and resources. Useful discussions of potential pitfalls and approaches to overcoming them may be found in Meyer and Booker (1990), Mosleh and Bier (1988), and Svenson (1989).

The formal use of expert judgment is most appropriate when extensive, noncontroversial data directly relevant to a problem is lacking, or when the issue studied is complex or is apt to receive extensive review and criticism. A formal expert-judgment process has a predetermined structure for the collection, processing, and documentation of experts' knowledge. The advantages and drawbacks in using such a process as opposed to an informal process are outlined in Bonano *et al.* (1990). The advantages include improved accuracy and reliability of the expert judgments, a reduced likelihood of critical mistakes leading to suspect or biased judgments, enhanced consistency and comparability of procedures, and improved scrutability and documentation for communication and external review. The drawbacks include an increase in the resources and time required to carry out the process, a reduction in the flexibility to make changes in the on-going process, and an enhanced vulnerability to criticism due to the relative transparency provided by a formal documentation of the procedures and findings. Bonano *et al.* warn that, while a formal process often requires more resources and time than an informal process initially requires, a faulty process that fails to withstand criticism or must be redone because of inappropriate design or improper execution may end up failing to satisfy the project's objectives and cost more in both time and resources. The potential for further costs in an informal study should be considered when evaluating the need for a formal process.

The expert judgment process used in NUREG-1150 (USNRC, 1989) is discussed in Ortiz *et al.* (1989). This process for gathering expert judgment was developed in response to criticisms of the previous Reactor Safety Study (USNRC, 1975) and an earlier draft of NUREG-1150. The history of this development underscores the importance of basing the SCC rulemaking analysis on state-of-the-art techniques and of making use of experienced practitioners in this difficult area.

As outlined in Ortiz *et al.* (1989), there were seven steps in the expert judgment process used in NUREG-1150. These steps should form the basis for the use of expert judgment in the uncertainty modeling for SCC rulemaking. Bonano *et al.* (1990) discuss specific techniques for the elicitation, use, and communication of expert judgments, as well as suggestions for the use of expert judgment in HLW disposal.

4.5.1 Selection of Issues and Experts

The selection of issues and experts is closely related. The initial selection of issues should be made by the project staff and used to guide the selections of experts. The experts should then review the list of issues and be invited to propose additions, deletions, or modifications to the list. There are two ways to organize the experts - by panels and by teams. The panel approach was used for NUREG-1150 (one panel for each of six groups

of related issues) and the Lawrence Livermore seismic hazard study (one seismicity panel and one ground motion panel) described in Bernreuter *et al.* (1989). The team approach was used in the Electric Power Research Institute (EPRI) seismic hazard study (six balanced teams, each containing seismicity and ground motion experts) described in Risk Engineering, Inc., *et al.* (1989). Regardless of which approach is used, it is essential that the experts should be knowledgeable about the state of the art and be chosen to represent a diversity of backgrounds, with a wide variety of experience and viewpoints (e.g., academic, consulting, commercial, national laboratory, government).

4.5.2 Presentation of Issues to the Experts

In addition to providing the experts with a clear statement and a good understanding of the issues, this step also provides a mechanism to discuss the state-of-the-art data base for the issues. An essential aspect of issue presentation is issue decomposition, which allows the experts to make a series of simpler assessments rather than one overall assessment of a complex issue. This step should be carried out with great care, as the decomposition of an issue can vary by expert and thereby significantly affect its assessment. Care should also be taken to present the issues so as to minimize potential biases in their assessment.

4.5.3 Preparation of Issue Analyses by the Experts

The experts should be given sufficient time and resources to analyze the issues before the elicitation session. This step may entail support by the project staff, e.g., by performing computer calculations or other requested analyses.

4.5.4 Discussion of Issue Analyses

Before the elicitation session, the experts should be allowed to present the results of their analyses and research. The ensuing discussion can serve to ensure a common understanding of the issues and the data base. The final part of this step is to reach agreement on the exact elicitation variables.

4.5.5 Elicitation Training and Elicitation

The purpose of elicitation training is to help the experts learn how to encode their knowledge and beliefs into probabilistic or other quantitative forms. Elicitation training can significantly improve the quality of the experts' assessments by avoiding psychological pitfalls that can lead to biased and/or overconfident assessments. Whenever the training session takes place, it is important that it not be abbreviated due to time pressure. The training should be carried out by a substantive expert who is knowledgeable about the issues to be assessed and a normative expert who is knowledgeable about decision theory and the practice of probability elicitation.

The elicitation sessions should be held immediately following the discussion of issue analyses and the selection of the elicitation variables. An elicitation team should meet separately with each expert, to avoid pressure to conform and other group dynamics interactions which might occur if the expert judgments were elicited in a group setting. The elicitation team should consist of a substantive expert, a normative expert, and a recorder. It is also useful to add as a fourth member the person who will prepare the final documentation.

4.5.6 Recomposition and Aggregation of Results

Each expert's elicitation should be recomposed by the normative and substantive experts to put them in a form suitable for further analysis. For example, recomposition is necessary to convert the subjective probability distributions provided for each part of the issue decomposition into an assessment for the issue as a whole.

After the recomposition of each expert's elicitation, the results should be aggregated to yield a final assessment for each issue. There are two general classes of aggregation methods - methods that tend to consensus and methods that tend to preserve the variability between the experts. Two informative reviews that discuss many of the aggregation methods that have been proposed are Genest and Zidek (1986) and Uppuluri and Seaver (1986).

Although consensus methods are often easy to implement (e.g., averaging over the experts), they should not be automatically applied without careful consideration. Because one of the primary goals of uncertainty modeling is to reflect the state-of-the-art uncertainty as expressed by the diversity of expert judgments, an aggregation method should not be used if it tends to mask the diversity of expert judgment. For example, consider a case where half the experts judge the probability P of a phenomenon to be close to zero, whereas the other half judge P to be close to one. Averaging over the experts is equivalent to the case where all the experts judge P to be approximately one-half. However, these two cases are quite different, since there is no disagreement among the experts in the second case, whereas there is a great deal of disagreement (and hence uncertainty) in the first case. In the second case, a decision-maker would have high confidence that $P \approx 1/2$, whereas in the first case, he does not know what value to assign to P . If he would want to make one decision if $P = 0$ and another decision if $P = 1$, premature averaging in the first case might deprive the decision-maker of essential information. In general, an aggregation method should be used only if a sensitivity study indicates that it does not destroy information, as expressed by the diversity of expert judgments, that might significantly affect a decision-maker's options.

4.5.7 Documentation

The final step in the expert judgment process is to document the entire process. Documentation has several purposes. First, it can be used by the experts involved to assure

them that their judgments were correctly reflected. Second, it can be used by potential users of the results of the process to enhance their understanding. Third, it can be used by peer reviewers of the process to provide an informed basis for their review. And finally, documentation can be extremely useful to update the analyses, when future research provides additional information.

4.6 Sensitivity Analysis

In engineering design and performance analysis, many sensitivity measures can be defined. A commonly used sensitivity measure in deterministic analysis is performance sensitivity, $\partial Z/\partial X_i$, where Z is a performance function (e.g., time-to-failure) and X_i is a design variable. In some cases, performance sensitivity is defined as $\Delta Z/\Delta X_i$, where a finite change (not necessarily very small) in X_i is taken to measure the change in Z . In general, a reference point (e.g., a design point) must be chosen to perform the analysis. In a deterministic design approach, sensitivity analysis is useful for identifying key design variables and improving/modifying designs. However, the above sensitivity analysis does not take into account the uncertainty or random variability in the input variables.

In probabilistic performance assessment, a more direct sensitivity measure is *probability/reliability sensitivity*, which measures the change in probability/reliability relative to the change in distribution parameters such as mean and standard deviation. Another, perhaps more important, kind of probability/reliability sensitivity analysis is the determination of the relative importance of the random variables. This analysis can be done, for example, by repeated probabilistic analysis in which one random variable at a time is treated as a deterministic variable (i.e., zero variance). The results of the analyses, for example, are a number of cdf curves or reliabilities. Based on the results, the relative importance (i.e., ranking) can be identified.

The concept of sensitivity analysis can be extended further to include all the major assumptions that are uncertain. The uncertainty assumptions could include probability distribution types (Lognormal, Weibull, etc.), process models (matrix flow, fracture flow), and parameters in an empirical life predictive model, etc. In general, probability distribution modeling (with input from expert judgment) should not be automatically applied to all input parameters or models, even when it is possible to do so. In particular, probability distribution modeling should not be applied to "merge" mutually exclusive assumptions/models. Instead, the results based on different models/assumptions should be evaluated and presented to the decision maker, along with their probabilities. In practice, this "sensitivity" analysis should be applied only to critical models/assumptions, so that the number of combinations presented to the decision-maker would not be excessive.

In summary, sensitivity analysis is a powerful tool for supporting probabilistic analysis and identifying key design variables and process models that contribute most to performance uncertainty. Identification of such design variables and process models is important for two reasons. First, it identifies those variables and processes that must be examined with particular care, to ascertain that their contributions to performance uncertainty have been properly assessed. Second, it can help to prioritize research programs that have the greatest potential to decrease performance uncertainty.

5. A PROBABILISTIC FRAMEWORK FOR ASSESSING CONTAINMENT PERFORMANCE

5.1 Qualitative vs. Quantitative Criteria

The Commission has several options: leave the SCC regulation as is, make the regulation less ambiguous but still qualitative, or make the regulation more quantitative. One consequence of leaving the SCC regulation wording qualitative is that the Commission has the flexibility to interpret SCC more precisely as more information becomes available on its feasibility. In this vein, the Commission has postponed the precise definition of the containment period by stating it to be a minimum period of 300 to 1,000 years after permanent closure of the repository. However, using a qualitative regulation increases the chance that the designer (DOE) will over- or under-design the waste packages.

A major motivation for making the SCC regulation more quantitative is to reduce its ambiguity for both DOE and NRC. Of course, any change in the regulation must recognize that, because of the long time periods involved, uncertainty about the geologic setting, and the inaccessibility of the waste packages once the repository is sealed, no design can guarantee complete containment. (The inclusion of the term "substantially" in the current regulation is a reflection of this fact.) The critical regulatory question is: "How much risk is acceptable?"

This section provides an example that shows how a framework for a quantitative probability-based rule can be developed, which takes account of the uncertainties. The example framework for the probability-based rule contains a number of parameters. A key parameter is a proportion K_o of "acceptable" waste package failures during the regulatory period. The value of K_o defines the design goal. Clearly, $K_o = 0$ (complete containment) is the most desirable criterion and appears to correspond most closely to the intent of the current regulation. The term "substantially" is given a quantitative meaning by specifying the probability that $K_o = 0$ is achieved. The purpose of this section is not to suggest a specific SCC rule, but rather to present a class of rules that is sufficiently broad to accommodate a wide range of scientific and regulatory concerns. If the example framework of the rule is adopted by NRC, additional analyses will be required to determine the values of the parameters that specify the precise form of the rule. Furthermore, NRC may have to provide additional guidelines and may have to increase the level of interaction with DOE.

5.2 Criteria for a Quantitative Rule

To maintain flexibility for NRC, while assisting the license applicant to design safe waste packages, it seems reasonable that any adopted quantitative rule should satisfy, as closely as possible, the following six criteria:

- (1) It should be easy to interpret and be unambiguous;
- (2) It should allow for a pass-fail criterion;

- (3) It should reflect the state-of-the-art in both scientific knowledge and uncertainty;
- (4) Its demonstration should be achievable with presently available or easily developed methodology and data, including the use of expert judgment;
- (5) It should allow flexibility for possible later rule modification;
- (6) It should allow flexibility to use data up to licensing hearing time and beyond, up until permanent closure of the repository.

In an attempt to satisfy these criteria, we describe an example framework for a probability-based rule, as outlined below.

5.3 An Example Framework for a Probability-Based Rule for SCC

The basic idea is to control the number of failed waste packages during the SCC performance period. Since the number of emplaced packages will be known, this is equivalent to controlling the proportion K of failed waste packages. However, because of the nature of the physical processes involved and the long time period, K is inherently stochastic and must be treated as a random variable. One possible approach is to control $E(K)$, the expected value of K . This approach is unsatisfactory because the actual value of K might be very far from $E(K)$, if the distribution of K had a large spread. To control K , it makes more sense to control the probability that K exceeds some regulatory limit K_o . Accordingly, a possible rule could be developed based on the following example framework:

"The probability that the proportion K of waste packages failing during the period $[0, T_o]$ does not exceed K_o should be no less than p_o ."

In symbols,

$$Prob\{K \leq K_o\} \geq p_o, \quad (1)$$

where K is the proportion of waste packages failing in time T_o and p_o , called the reliability target, is the minimum acceptable probability that $K \leq K_o$. How to deal with the uncertainty in demonstrating that Eq. (1) holds will be discussed below.

It may be desirable to extend the requirement (1) to several values of K_o and p_o . For example, it may be desirable to specify a value of p_o very close to 1 for a large value of K_o in order to control the risk of a massive failure. However, for simplicity, this report will consider only one value for K_o and p_o . Our results can be easily extended if more than one value of K_o and p_o is used.

As an example, suppose that a total of 10,000 waste packages will be emplaced and NRC determines that no failures are allowed in the first 1,000 years after permanent closure. Rec-

ognizing the inherent unpredictability of the exact number of failures, NRC allows the regulatory limit of 0 failures to be exceeded with no more than 0.1 probability. In this example, $K_o = 0$, $p_o = 0.9$, and the criterion is $Prob\{K \leq 0\} \geq 0.9$.

The specific values of T_o , K_o and p_o will be determined by NRC. In the current SCC regulation, T_o , at a minimum, lies between 300 and 1000 years. The choice of T_o , K_o and p_o will depend on the definition and the consequences of failure and will have to be consistent with the allowable radionuclide release rates subsequent to T_o .

It is convenient to write the requirement (1) in terms of the cumulative distribution function (cdf) of $N(t)$, the random number of failed WPs in $[0, t]$. If M is the total number of emplaced WPs, the cdf of $N(t)$ is defined as:

$$\begin{aligned} P(n, t) &= Prob\{\text{number of WPs failing in } [0, t] \text{ is } \leq n\}, \\ &= Prob\{N(t) \leq n\}, \end{aligned} \quad (2)$$

for all $n = 0, 1, \dots, M$. Only values of $P(n, t)$ for $t \leq T_o$ are relevant for SCC, although values of $P(n, t)$ for $t > T_o$ may be needed for repository evaluation after T_o years.

A sketch of $P(n, t)$ as a function of n is shown in Figure 3 for two values of t , with $t_1 < t_2$. For $n = 0$, $P(0, t)$ is the probability that no WPs fail in $[0, t]$, i.e., the probability of full containment, and is a number between 0 and 1. Since M is the total number of waste packages, $P(M, t) = 1$ for all t . If $t_1 < t_2$, $N(t_1) \leq N(t_2)$, since the number of failed WPs can only increase as time increases. Hence, as Figure 3 indicates, the cdf of $N(t_1)$ is greater than the cdf of $N(t_2)$ for all values of $n < M$. Figure 3 has been drawn under the assumption that $P(n, 0) \approx 1$ for all n , i.e., the number $N(0)$ of failed packages is zero at time $t = 0$. At time $t = \infty$, all packages would certainly have failed; hence $P(n, \infty) = 0$ for all $n < M$.

To write the requirement (1) in terms of $P(n, t)$, note that $K = N(T_o)/M$ in the period $[0, T_o]$. Hence,

$$\begin{aligned} Prob\{K \leq K_o\} &= Prob\{N(T_o) \leq K_o M\} \\ &= P(K_o M, T_o) \end{aligned}$$

and Eq. (1) becomes

$$P(K_o M, T_o) \geq p_o. \quad (3)$$

We see that the requirement (1) will hold if it can be demonstrated that the cdf $P(K_o M, T_o)$ complies with the reliability goal. Assuming that this can be done, a quantitative rule based on Eq. (1) will satisfy the six criteria listed in Section 5.2:

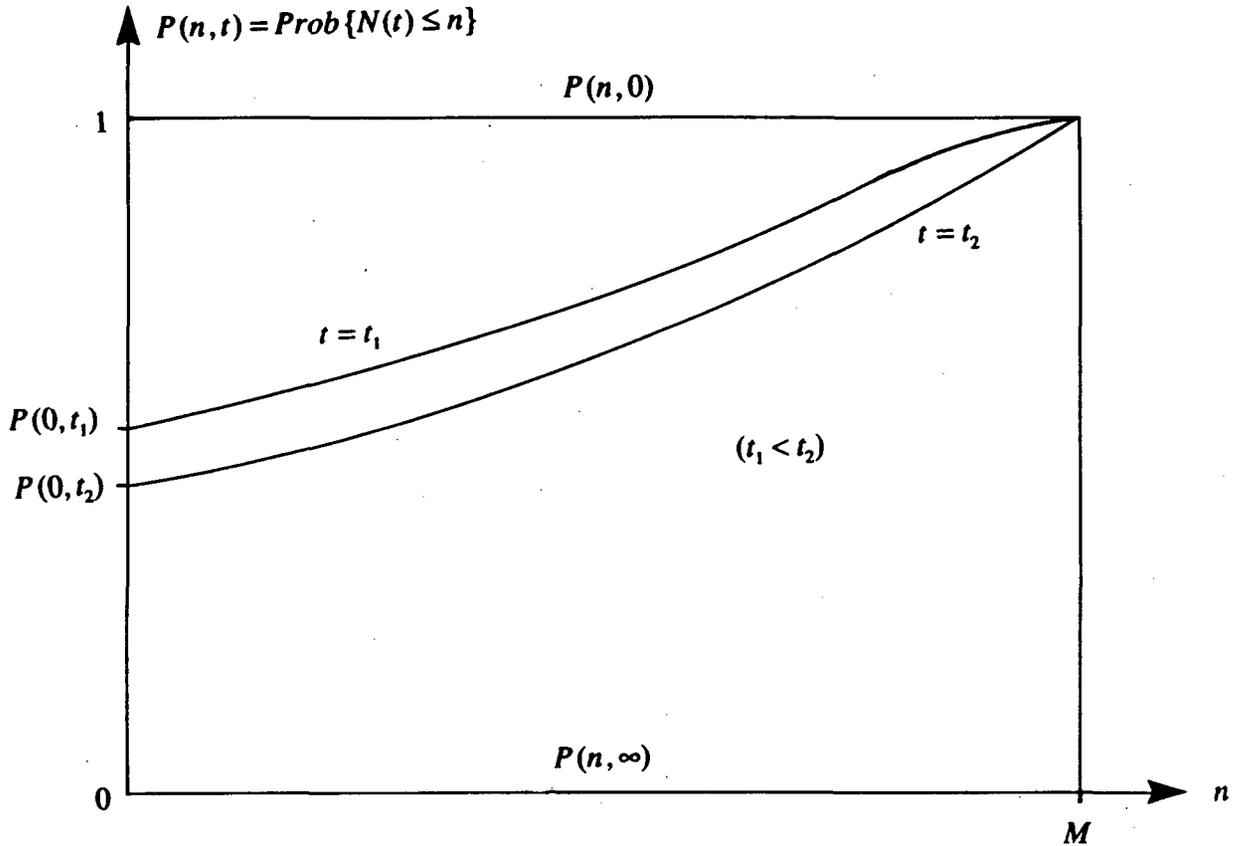


Figure 3. Waste-package number-of-failures distribution.

- (1) Eq. (3) is easy to interpret and is unambiguous.
- (2) The repository will comply with the rule if and only if Eq. (3) holds.
- (3) The demonstration that Eq. (3) holds will very likely require state-of-the-art scientific knowledge.
- (4) The demonstration of compliance should be achievable if the reliability target is not too high.
- (5) Flexibility can be achieved by modifying K_o or p_o . For example, the rule could be made stricter by increasing the reliability target p_o , or by decreasing K_o .
- (6) The rule could allow the use of new data up until permanent closure by requiring Eq. (3) to be evaluated conditional on all information available at time $t = 0$, i.e., permanent closure time.

We have shown that a rule based on Eq. (3) would satisfy the suggested criteria, provided that Eq. (3) can be verified. However, because of incomplete information about the repository and the behavior of the waste packages over the very long time period involved, there will be considerable uncertainty about the value of the probability $P(K_oM, T_o)$. It remains to show how to incorporate this uncertainty into the example framework.

Note that, apart from the uncertainty in the determination of $P(K_oM, T_o)$, there is no ambiguity in the probability-based rule, as there is in the current SCC regulation. Although it is probability-based, the rule boils down to the determination of the value of $P(K_oM, T_o)$. If there were sufficient data and validated models to compute the cdf, then there would be no uncertainty involved in applying the rule. For example, suppose that the waste packages failed independently in $[0, T_o]$, each with a known failure probability equal to q , where q is very small. Then, to an excellent approximation, the number of failures in $[0, T_o]$ would have a Poisson distribution with mean $\lambda = Mq$ and $P(K_oM, T_o)$ could be easily calculated and compared with p_o in Eq. (3). In practice, the calculation of $P(n, t)$ would be much more complicated because failures are correlated due to common uncertain environments affecting multiple waste packages.

Because of the large scientific uncertainty, it is highly unlikely that it will be possible to demonstrate that Eq. (3) holds with certainty for reasonable values of K_o, M, T_o and p_o : at best, it may be possible to demonstrate that Eq. (3) holds with some high assurance, say, 95 percent. Accordingly, we will assume that $P(n, t)$ has a distribution and consider the (second level) cdf defined by

$$F(p; n, t) = \text{Prob}\{P(n, t) \leq p\}. \quad (4)$$

In practice, it will not be necessary to estimate $F(p; n, t)$ for all values of the parameters n, t and p ; it will be sufficient to evaluate Eq. (4) only for the parameter values under consideration. For example, let $M = 10,000$, $K_o = 0$ and $T_o = 1,000$ years. If the reliability target equals 0.9 and 95 percent assurance is assumed to be an acceptable approximation of "certainty," the corresponding SCC rule is that there is at least 95 percent assurance that the probability of zero WPs failing in the first 1000 years is no less than 0.9. Such a rule would be written:

$$\text{Prob}\{P(K_oM, T_o) \geq 0.9\} \geq 0.95 \quad (5)$$

i.e.,

$$\text{Prob}\{P(0 \text{ WP's}, 1000 \text{ yrs}) \geq 0.9\} \geq 0.95.$$

For arbitrary K_o, M and T_o , this SCC rule can be written in terms of $F(0.9; K_oM, T_o)$ as

$$1 - F(0.9; K_oM, T_o) \geq .95. \quad (6)$$

This means there is a high assurance (95 percent) that the reliability is at least 90 percent for any fixed K_o and T_o .

From Eqs. (4) and (6),

$$F(0.9; K_o M, T_o) = Prob\{P(K_o M, T_o) \leq 0.9\} \leq 0.05. \quad (7)$$

This can be interpreted as saying that the probability that the reliability target is not met is no more than 5 percent.

In practice, the situation is somewhat more complicated than is implied by Eq. (7). To demonstrate that the reliability target can be met, the repository designer will have to provide an estimate $p^* = P^*(K_o M, T_o)$ for $P(K_o M, T_o)$ such that $p^* \geq p_o$. In addition, NRC may well require assurance that Eq. (3) is, in fact, satisfied. As part of the rule, NRC may want to control the chance that Eq. (3) is not satisfied, given that $p^* \geq p_o$. This chance is called the risk and is denoted by β , where

$$\beta = Prob\{P(K_o M, T_o) < p_o \mid p^* \geq p_o\}. \quad (8)$$

The next step is to relate β to $F(p; n, t)$. We cannot use Eq. (4) directly, since β is a conditional probability and the events $\{P(K_o M, T_o) < p_o\}$ and $\{p^* \geq p_o\}$ are not independent. Since $p^* = P^*(K_o M, T_o)$ is an estimate of $P(K_o M, T_o)$, it is clear that these events are negatively correlated. Accordingly, if it is known that $p^* \geq p_o$, it is less likely that $P(K_o M, T_o) < p_o$ than if nothing were known about p^* . From Eqs. (4) and (8), this result is written as:

$$\begin{aligned} \beta &= Prob\{P(K_o M, T_o) < p_o \mid p^* \geq p_o\} \\ &< Prob\{P(K_o M, T_o) \leq p_o\} \\ &= F(p_o; K_o M, T_o) \end{aligned}$$

or

$$\beta < F(p_o; K_o M, T_o). \quad (9)$$

From Eq. (9), in order to control the risk β , it is sufficient to control $F(p_o; K_o M, T_o)$. This can be done by showing that $F(p_o; K_o M, T_o) < \beta_o$, where β_o is an acceptable level of risk. From Eq. (9), the risk would be less than β_o .

Combining these results, we see that the rule based on the example framework has two parts. The designer must demonstrate that

$$p^* = P^*(K_o M, T_o) \geq p_o \quad (10a)$$

and

$$F_o = F(p_o; K_o M, T_o) < \beta_o. \quad (10b)$$

The first part of the rule controls the reliability, and the second part controls the assurance that the reliability target is met. An alternative is to omit the second part and use the demonstration of the first part as part of an argument that Eq. (3) holds with reasonable assurance (see "Remarks").

Remarks

Remark 1. Criteria Satisfaction

The rule based on the example framework satisfies all the criteria in Section 5.2 except the first. The first part of the rule does satisfy the first criterion since p^* is the estimate of the reliability target. However, since it seems very unlikely that F_o can be calculated with sufficient precision, there will be unavoidable ambiguity associated with Eq. (10b). This can be dealt with by using either the reasonable-assurance or the bounding approach.

Remark 2. Reasonable-Assurance Approach

It may not be possible to demonstrate that Eq. (10b) holds. The reasonable-assurance approach consists of demonstrating that Eq. (3) holds with "reasonable assurance." For example, one could find an estimate p^* that satisfies Eq. (10a) and develop an argument that, with reasonable assurance, the true reliability $P(K_o M, T_o)$ is greater than the reliability estimate p^* . This would imply that $P(K_o M, T_o) > p^* \geq p_o$, i.e., Eq. (3) holds with reasonable assurance.

The reasonable-assurance approach can also be applied to the demonstration of Eq. (10b). Instead of a model for the second level cdf $F(p; n, t)$ that leads to the demonstration that $F_o < \beta_o$, as required by Eq. (10b), one could settle for a "reasonable assurance" that $F_o < \beta_o$. For example, one could develop a chain of reasoning, including expert judgment, that would culminate in a statement that " $F_o < \beta_o$ with reasonable assurance."

Remark 3. Bounding Approach

It is highly unlikely that it will be possible to characterize the probability structure of all input variables and parameters with enough precision to determine $F(p; n, t)$. If all that is known about some input variables is that they lie in some known intervals with high assurance, then it may be possible to determine an upper bound for $F(p; n, t)$. This could be done by using conservative (bounding) models and/or the bounding approach outlined in Appendix D. If this upper bound is less than β_o , then Eq. (10b) holds.

6. SUMMARY

The technical uncertainty arising from the "substantially complete containment" requirement in 10 CFR 60.113 requires, first, a degree of understanding of the technical elements that must be considered in a systematic approach to address any containment evaluation, and second, a logical approach to define "how well" each of the technical elements can be determined. The first aspect of the requirement is addressed in the technical considerations report (Report 1). The second part is the subject of this report.

This report describes a number of broadly applicable uncertainty evaluation methods based on existing methodologies. To address the uncertainty issues, four uncertainty evaluation methods are identified: the probability-distribution approach, the bounding approach, expert judgment, and sensitivity analysis. The diversity of the various methods identified in this report is a reflection of the diverse nature of the types of technical data or information that can be expected from the technical elements described in Report 1. The associated uncertainties are identified and characterized by uncertainty types and sources. Once the pertinent data for the technical elements are identified, they can be individually or collectively quantified by these methods. The report examines the current SCC regulation and provides an example that shows how a framework for a quantitative probability-based rule can be developed that takes account of the uncertainties.

There is not any "best" methodology that can be recommended. As a general rule, the problem should be approached in terms of its unique characteristics developed from the waste-package design. Several methodologies have been discussed in the appendices. The methodologies discussed are of a very general nature and are not exhaustive. A mix of these and other appropriate methodologies should be explored, taking feasibility and applicability into account. Whatever mixed approach is used, it is essential that the probabilistic framework be built on a solid understanding of the physical, chemical, geological, and other aspects of the waste-package long-term containment problem.

7. REFERENCES

1. U.S. Code of Federal Regulations, "Disposal of High-Level Radioactive Waste in Geologic Repositories," Title 10, Part 60, "Energy," January 1988.
2. Manaktala, H. K. and C. G. Interrante, "Technical Considerations for Evaluating Substantially Complete Containment of HLW within the Waste Package," CNWRA 90-001, 1990.
3. Nair, P. K. and E. Tschoepe, "Substantially Complete Containment Feasibility Study Recommendations Report," CNWRA Letter Report, 1990.
4. The Center for Nuclear Waste Repository Analyses, CNWRA Technical Operating Procedure TOP-001-02, 1990.
5. Woo, H. H. and F. A. Simonen, "The Impact of In-Service Inspection on the Reliability of Nuclear Piping," Advances in Probabilistic Fracture Mechanics, C. Sundararajan (ed.), ASME PVP-Vol. 92, 1984.
6. Sato, N., "Stochastic Process of Chloride Pit Generation in Passive Stainless Steel," Electrochemical Science and Technology, 123 (8), 1976.
7. Provan, J. W. and E. S. Rodriguez, III, "Part I: Development of a Markov Description of Pitting Corrosion," Corrosion Science, 45 (3), pp. 178-192, March 1989.
8. Shibata, T. and T. Takeyama, "Stochastic Theory of Pitting Corrosion," Corrosion-Nace, 33 (7), pp. 243-251, July 1977.
9. Rodriguez, E. S. III and J. W. Provan, "Part II: Development of a General Failure Control System for Estimating the Reliability of Deteriorating Structures," Corrosion Science, 45 (3), pp. 193-206, March 1989.
10. Sato, E., H. Abo, and T. Murata, "Life Time Estimation and Accelerated Stress Corrosion Cracking Test of Stainless Steel in a Neutral Chloride Environment," Corrosion 89 Conference, Paper No. 97, April 17-21, 1989.
11. United States National Regulatory Committee, "PRA Procedures Guide," NUREG/CR-2300, Vols. 1 and 2, 1983. (Washington, D. C.: U.S. Nuclear Regulatory Commission)
12. Vesely, W. E. and D. M. Rasmuson, "Uncertainties in Nuclear Probabilistic Risk Analysis," Risk Analysis, 4, pp. 313-322, 1984.
13. Jensen, R. V., "Classical Chaos," American Scientist, 75, pp. 168-181, 1987.
14. Collins, J. A., Failure of Materials in Mechanical Designs, John Wiley & Sons, Inc., 1981.

15. Bonano, E. J., et al., "Elicitation and Use of Expert Judgment in Performance Assessment for High-Level Radioactive Waste Repositories," NUREG/CR-5411, SAND 89-1821, 1990. (Albuquerque, New Mexico: Sandia National Laboratories.)
16. United States National Regulatory Committee, "Severe Accident Risks: An Assessment for Five U.S. Nuclear Power Plants," NUREG-1150, 1989. (Washington, D. C.: U.S. Nuclear Regulatory Commission)
17. Bernreuter, D. L., J. B. Savy, R. W. Mensing, and J. C. Chen, "Seismic Hazard Characterization of 69 Nuclear Plant Sites East of the Rocky Mountains," NUREG/CR-5250, UCID-21517, 1989. (Livermore, California: Lawrence Livermore National Laboratory)
18. Risk Engineering, Inc., Yankee Atomic Electric Company, and Woodward-Clyde Consultants, "Probabilistic Seismic Hazard Evaluations at Nuclear Plant Sites in the Central and Eastern United States: Resolution of the Charleston Earthquake Issue," NP-6395-D, 1989. (Electric Power Research Institute, Palo Alto, California)
19. Meyer, M. A. and J. A. Booker, "Eliciting and Analyzing Expert Judgment," NUREG/CR-5424, LA-11667-MS, 1990. (Los Alamos National Laboratory, Los Alamos, New Mexico)
20. Mosleh, A. and V. M. Bier, "A Critique of Current Practice for the Use of Expert Opinions in Probabilistic Risk Assessment," Reliability Engineering and System Safety, 20, pp. 63-85, 1988.
21. Svenson, O., "On Expert Judgments in Safety Analyses in the Process Industries," Reliability Engineering and System Safety, 25, pp. 219-256, 1989.
22. Ortiz, N. R., T. A. Wheeler, R. L. Keeney, and M. A. Meyer, "Use of Expert Judgment in NUREG-1150," Proceedings of the International Topical Meeting on Probability, Reliability and Safety Assessment, American Nuclear Society, La Grange Park, Illinois, 1989.
23. United States National Regulatory Committee, "Reactor Safety Study - An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants," WASH-1400 (NUREG-75/014), 1975. (Washington, D.C.: U.S. Nuclear Regulatory Commission)
24. Genest, C. And J. V. Zidek, "Combining Probability Distributions: A Critique and an Annotated Bibliography," Statistical Science, 1 (1), pp. 114-148., 1986.
25. Uppuluri, V. R. R. And D. A. Seaver, "Formal Procedures for Obtaining a Consensus of Opinion: A Critical Evaluation," 1986. (Oak Ridge National Laboratory, Oak Ridge, Tennessee)

APPENDIX A - UNCERTAINTY MODELING METHODS - A SELECTED REVIEW

Principal Author: Andre G. Journal

TABLE OF CONTENTS

	Page
1. INTRODUCTION	A-1
2. TOPICS RELATED TO SUBSTANTIALLY COMPLETE CONTAINMENT (SCC)	A-1
2.1 SCC vs. Release Rate	A-1
2.2 The Scenario Approach	A-2
2.2.1 Scenario Development	A-2
2.2.2 Scenario Probability	A-3
2.2.3 Relative Probability	A-5
2.3 Space and Time Dependence	A-5
2.4 Expected Value and Least Square Criterion	A-7
2.5 Lognormality and Long-Tail Distributions	A-8
2.6 The Multivariate Normal Distribution	A-8
2.6.1 Tail Probabilities	A-8
2.6.2 Maximum Entropy	A-9
2.6.3 Normal-Score Transforms	A-9
3. UNCERTAINTY EVALUATION METHODS	A-10
3.1 Introduction	A-10
3.2 Monte Carlo Simulation	A-10
3.3 Response Surface Methodology	A-13
3.4 Differential Analysis Approach	A-14
3.5 Geostatistical Techniques	A-15
3.6 Propagation of Uncertainties through Multiple Models	A-17
4. OTHER WASTE PACKAGE ISSUES	A-17
5. REFERENCES	A-19
6. BIBLIOGRAPHY	A-21
FIGURE	
A-1. Non-Mutually Exclusive Scenarios	A-4

APPENDIX A - UNCERTAINTY MODELING METHODS - A SELECTED REVIEW

1. INTRODUCTION

To evaluate the appropriateness of a more quantitative regulation, and assess the feasibility of a probabilistic rule framework discussed in Section 5 of the main report, a number of relevant documents were reviewed. This review does not claim exhaustivity; only those documents listed in the references and bibliography of this appendix were reviewed. However, we believe that these documents do reflect present concepts for demonstrating safety compliance, and we wish to comment on some possible shortcomings. In addition, we will suggest possible approaches applicable to Substantially Complete Containment (SCC).

It is clear that the development of a compliance-demonstration methodology accounting for severe shortage of data is an extremely difficult task, and that only the repository designer with access to all available data can hope to develop such methodology. However, an independent review may prove useful for future methodology development.

The literature reviewed for this report can be classified into two categories. The first category focuses on the problem of the release rate to the environment, whereas the second one emphasizes the problem of failure of waste packages (WPs). Since the release rate depends on the WP failure rate that, in turn, depends on critical environment parameters such as water accessing the engineered barrier system (EBS), the two issues cannot be fully decoupled.

2. TOPICS RELATED TO SCC

2.1 Sc vs. Release Rate

Several Sandia-originated reports, as summarized by Davis *et al.* (1989), do not distinguish the two 10 CFR 60.113 rules on SCC and release rate. The Sandia flow charts relate to the modeling of the release of radionuclide to the environment, starting from the initially contained inventory; compliance is limited to the U.S. Nuclear Regulatory Commission (NRC) release-rate regulations and related U.S. Environmental Protection Agency (EPA) regulations. For example, simulation of water flow is done from the WP outward, as a medium of radionuclide release to the environment. Whereas for the sole purpose of SCC performance assessment, one should simulate the inward flow, allowing water and its chemicals to reach the WPs and degrade the containment capability.

Using Sandia terminology, the question is what are the scenarios that could put any number of WPs in contact with substantial water, say more than one gallon per year per container during the containment period (300 to 1,000 years)? Study of the consequences of such scenarios require parallel critical data (or process simulations) about the long-term degradation effect of water, with due account being given to the temperature and failure-chain effect.

Assuming that water is a major source of WP degradation, the scenarios for surface and groundwater reaching the WPs are more critical to containment, whereas the scenarios for radionuclide transport by groundwater flow are important for control of release rate.

Remark: In this regard, the "tentative interpretation of the EBS boundary as the wall of the emplacement hole" (Ramspott, 1988) and the assessment of such interpretation as conservative may not be correct. Indeed, the environment prevailing in the underground facility (in particular humidity) may negatively impact containment. 10 CFR Part 60 defines the EBS as the WPs plus the underground facility. Planning of the underground facility to prevent or retard access of water to the WPs is, thus, part of the SCC requirement.

2.2 The Scenario Approach

As summarized in Campbell and Cranwell (1988), the performance assessment methodology involves three steps:

1. Develop scenarios.
2. Develop models for use in estimating consequences from the scenarios.
3. Perform performance assessment to assess compliance with the regulatory rules.

The corresponding types of uncertainty are (Bonano and Cranwell, 1988):

1. Scenario uncertainty, including completeness of all scenarios, uncertainty about the scenario occurrence and the consequences of each scenario.
2. Modeling uncertainty, including uncertainty about the conceptual model, the mathematical model and potential errors in the computer codes.
3. Data and parameter uncertainty, possibly the easiest to model but also, unfortunately, the least consequential.

As pointed out by the Sandia authors, the major difficulty with this framework is the quantification of uncertainty in scenario completeness and the evaluation of the relative probability of occurrence of each scenario being considered.

2.2.1 Scenario Development

In the Sandia reports, a scenario defines a set of conditions (including events, features, and processes) leading to particular performance measures that may or may not meet compliance. That set of conditions includes all input variables and all the parameters defining the transfer function that processes the input variables into the output values.

The fore-mentioned scenario concept can be generalized if we consider that the parameters of the transfer function are a subset of the input variables. In such cases, a scenario can be interpreted as a particular realization of a multivariate set of usually dependent variables $\{X_1, \dots, X_L\}$ characterizing the state of a system.

Assuming that the "normal" (e.g., unperturbed, base-case) scenario will likely result in meeting the performance goal with high probability, the thrust of scenario development for evaluating SCC compliance should be to select those scenarios that could lead to non-compliance, and to evaluate their probability of occurrence.

For a high-reliability containment design, a scenario of interest would correspond to a "perturbed," possibly "extreme" set of conditions or realization $\{X_1 = x_1, \dots, X_L = x_L\}$, "extreme" defining a realization with low probability of occurrence and leading to non-compliance. When developed by a panel of experts, it thus corresponds to a subjective (and most often unquantifiable) sampling of "extreme corners" of the joint distribution of L input variables $\{X_1, \dots, X_L\}$.

The scenarios in general should be developed as much as possible from data, whether actual or simulated, and expert judgment should be used when quantitative data are not available.

Elements contributing to scenario developments are:

- Relevant data collected on possibly different but analogous (in some sense) sites and time scales.
- Physical process simulations, including simulations leading to extreme output values. Such simulations could provide a ranking and thus some measure of the probability of occurrence.
- Sensitivity analysis indicating which of the L variables $\{X_1, \dots, X_L\}$ are most consequential to, say, WP failure, and pointing at potential harmful (for SCC) couplings of the variables.

A panel of experts may then be convened, with a careful elicitation procedure, to decide on the details and contents of each scenario to be considered, i.e., on the N sets $\{x_1^{(i)}, \dots, x_L^{(i)}\}, i = 1, \dots, N$, if N scenarios are retained (see Section 4.5 of the main report).

2.2.2 Scenario Probability

Ideally a model for the joint L -variate distribution of $\{X_1, \dots, X_L\}$ should be established, from which the probability of any scenario i could be derived, say:

$$\begin{aligned} \text{Prob}\{X_1 \in [x_1^{(i)} \pm dx_1], \dots, X_L \in [x_L^{(i)} \pm dx_L]\} &= P_i \\ &= \text{Probability of scenario } i. \end{aligned}$$

A scenario i (more precisely the class of scenarios i) is defined by the intersection of L intervals, one for each of the L input variable X_i . Note that to any number N of scenarios actually considered should be added an 0^{th} scenario, called the "complement scenario," corresponding to the sum (union) of all joint realizations not accounted for by the L scenarios.

For the scenario probability, $P_i, i = 0, \dots, N$, to be of any use and, in particular, to sum up to 1, the $(N + 1)$ scenarios must not only be complete, but also mutually exclusive, i.e. no two scenarios should share a common set of realizations of all L variables X_i . For example, with $L = 2$, the two scenarios $\{X_1 \in [1, 3], X_2 \in [1, 3]\}$ and $\{X_1 \in [2, 4], X_2 \in [2, 4]\}$ are non-mutually exclusive because they share the common set $\{X_1 \in [2, 3], X_2 \in [2, 3]\}$ defined on all $L = 2$ variables, see Fig. A-1.

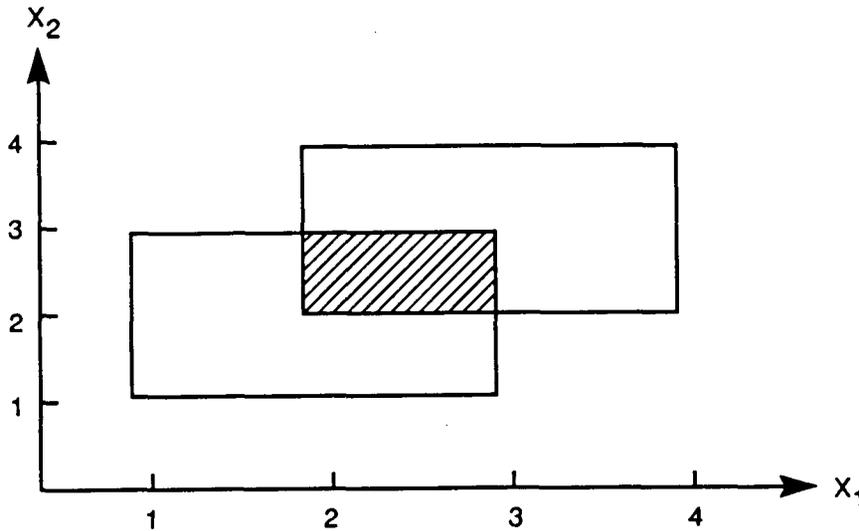


Figure A-1. Non-mutually exclusive scenarios.

Note: The two scenarios $\{X_1 \in [1, 3], X_2 \in [1, 3]\}$ and $\{X_1 \in [2, 4], X_2 \in [2, 4]\}$ share a common set (hatched area) of all defining variables.

If the L variables whose realizations determine a scenario are not specified beforehand, it is possible that scenarios defined by a panel of experts might not be mutually exclusive. In such a case, the scenarios should be transformed to an equivalent set of mutually exclusive scenarios before occurrence probabilities are solicited. (See Appendix B for a discussion of methods to evaluate occurrence probabilities.)

2.2.3 Relative Probability

Rather than seeking the absolute probability of occurrence of each of the N scenarios retained for study, one may limit one's goal to a relative probability ranking of these scenarios. For example scenario i is ranked more probable than scenario i' , without saying how much more probable, if some aggregation of the expert opinion probabilities for i exceed that for i' . Such ranking would not require the condition that the N scenarios be mutually exclusive nor even that they are complete (exhaustivity). But such ranking falls short of providing the probability of occurrence of, say, the impact (response function) of scenario i , thus would not provide per se the required probabilistic assessment of SCC compliance.

2.3 Space and Time Dependence

The interdependence between processes and input parameters pertaining to the risk of failure of any number of waste packages is the major difficulty faced in modeling that risk: interdependent processes cannot be decoupled, and dependent input variables must be sampled jointly to account for their degree of correlation (not necessarily linear). Thus, there is a tendency to make simplifying independence hypotheses where data or, in the lack thereof, geological evidence, would invalidate them. When dependence or couplings of variables are known to result in greater risk of failure, an independence hypothesis must be based on unquestionable data and/or a strong consensus of expert opinions. Alternatively, sensitivity studies may show that the independence assumption does not seriously alter the final conclusions.

In the absence of such justification, the hypothesis of independence should not be made because it may violate the essential intrinsic geological dependence patterns.

Some examples of such dependence are:

- The spatial distribution of conductivity data for the evaluation of water infiltration (an inward process leading to accelerated corrosion of the WPs) and water transport of released radionuclides. For example, there seems to be a developing consensus that major water-flow rates at Yucca Mountain would occur through flow in fractures. A fracture can be seen as a string of high conductivity values very different from the matrix conductivity, i.e., a specific dependence feature of the conductivity field. Moreover, such spatial dependence (clustering) of extreme values is not relevant to linear-correlation analysis and traditional covariance and lognormal modeling (Journel and Alabert, 1989).
- The space and time-dependence of WPs' failure due to common external failure mechanism and to the chain-effect. Water dripping from a fracture intersecting several galleries could simultaneously affect several WPs and jointly shorten their respective containment periods through accelerated corrosion. Similarly, an earthquake may generate rock falls and gallery

collapses throughout part or all of the repository, thus generating dependent WP failures. This geology-induced dependence process can be compounded by a time chain-effect if the heat release of the failed WPs accelerates the corrosion of other nearby WPs.

In general, WP failures caused by external sources (e.g., earthquakes) are most likely to involve statistical dependence in space, whereas WP failures caused by internal sources (e.g., materials) are more likely to be independent from one WP to another. Unfortunately, the external sources are also those most likely to generate a large number of early failures.

- The sequential failure process of the successive barriers of a WP. The internal degradation sources would start acting on the k^{th} outer barrier only when the $(k - 1)^{\text{th}}$ has failed.
- The interdependence of radionuclide release rate and the water-flow process. Indeed water acts simultaneously as an corrosion agent (in fact the major one) and a transport vehicle. Thus, the study of the release rate to the environment under disruptive excess of water cannot be decoupled from that of SCC under the same scenario.

Modeling dependence: When two processes are dependent, they should be studied or simulated jointly. If decoupling is required, say for reasons of computer coding and computation time, then a series of different conditional processes should be considered. Using a probabilistic terminology, consider two dependent variables representing the two processes, X_1 and X_2 .

Either the joint distribution of X_1, X_2 should be considered, i.e.,

$$F(x_1, x_2) = \text{Prob} \{X_1 \leq x_1, X_2 \leq x_2\},$$

or a series of distributions for X_1 conditional on different values (outcomes) of X_2 should be considered, i.e.,

$$F(x_1; x_2^{(i)}) = \text{Prob} \{X_1 \leq x_1 \mid X_2 = x_2^{(i)}\}, i = 1, \dots, N.$$

The number N of classes of conditioning value for X_2 depends on the range of X_2 and the type and strength of the (X_1, X_2) dependence. A word of caution is here relevant:

Heteroscedastic (non-Gaussian) dependence: The function or process $\phi(x_1; x_2^{(i)})$ may be radically different from one conditioning value $x_2^{(i)}$ to another $x_2^{(j)}$, not only different in terms of means or expected values as Gaussian-related models would indicate. Even if both marginal distributions of X_1 and X_2 are lognormally distributed, this does not imply that any of the N conditional distributions $F(x_1; x_2^{(i)})$ are lognormal. The dependence (X_1, X_2) may be limited to extreme values of X_2 , e.g.,

$$\text{Prob} \{X_1 \leq x_1 \mid X_2 = x_2\} = F(x_1), \text{ for all } x_2 \leq b_2,$$

$$\text{Prob} \{X_1 \leq x_1 \mid X_2 = x_2\} = \psi(x_1; x_2), \text{ for } x_2 > b_2 \text{ (large) with } \psi \neq F.$$

For example, the flow process may change drastically when the spatial distribution of conductivity (X_2) shows a large proportion of spatially connected high values (open fractures).

2.4 Expected Value and Least Square Criterion

There is sometimes a misconception that the mean of a distribution, being the "expected" value, should be the preferred value, and that the accuracy of an estimator should be measured in terms of mean square error. Care should be taken in using the mean value as the preferred value: the mean of a distribution may not exist (e.g., Cauchy distribution) or may correspond to a zero-probability outcome (e.g., the mean 0.5 of a 0-1 coin toss); the mean value of an input distribution will not necessarily generate an output value equal to the mean of the output distribution.

It is true that the mean $m = E\{Z\}$ is among all outcomes of a random variable Z the one that minimizes the average squared error, i.e.,

$$m \text{ is such that } E\{[Z - m]^2\} = \min\{E\{[Z - z]^2\}, \text{ all } z\}.$$

(Berger, 1980, p. 60.)

However, one should question the relevance of adopting the mean square criterion $E\{[Z - z]^2\}$ to measure the impact of the random error $[Z - z]$. Had the mean absolute error $E\{|Z - z|\}$ been minimized, the optimal estimate would have been the median. Besides analytical simplicity, which is not a relevant notion for the safety of a repository, the mean square criterion $\min E\{[Z - z]^2\}$ is no better than the mean absolute deviation criterion $\min\{E\{|Z - z|\}\}$ or a particular loss function $L(\cdot)$ criterion: $\min\{E\{L(Z - z)\}\}$ (see Journel, 1989 - Lesson 4 and utility theory references, for further discussion).

For conservative reasons, one should rather choose for input to the analysis the value of z that would yield the most unfavorable (higher failure probability) output response. In a scenario context, choosing the mean of the input variable distribution may be non-conservative. This problem becomes less serious if the distribution of Z is highly positively skewed, for, then, the mean $E\{Z\}$ is equal to a high p -quantile value z_p , with $p \gg 0.5$, which would for most input variables represent a conservative choice.

In fact, and as is explained in Appendix B of the main report, the question is not so much to ensure a conservative response value, i.e., a value that could entail non-compliance, but to attach a probability of occurrence to this response value y seen as a function of the input value(s) $x: y = \phi(x)$. If the transfer function ϕ is monotonic increasing and single-valued, the output value $y_p = \phi(x_p)$ corresponding to the p -quantile x_p of the input distribution is also the p -quantile of the Y -response distribution. Thus, focus should be on mapping p -quantiles from the input space to the output (response) space or, better, conversely.

2.5 Lognormality and Long-tail Distributions

In the Sandia reports and more generally in the hydrological literature, distributions for input variables are often assumed normal or lognormal. If data exist to support a lognormal model for transmissivity averaged over the whole thickness of an aquifer (Hoeksema and Kitanidis, 1985), there are no such data for conductivity distributed in the 3-dimensional space. In fact both data and good judgment indicate that conductivity, in an heterogeneous media consisting of a mixture of multiple rock types and porosity fabrics, features multimodal distributions with modes differing by several orders of magnitude in millidarcies. The fact that core plugs sampled for permeability are not taken in impervious media nor in fractures or fragmented rocks entails a censored sampling resulting in unimodal histograms, but does not refute the existence of extreme modes. Unfortunately, such extreme modes are most consequential for fluid flow (Desbarats, 1987).

The choice of lognormal distributions that are not long-tail compared to, say, hyperbolic distributions, should be carefully documented from either unbiased data (which can sample extreme values) or a large consensus of expert opinions. The argument of analytical convenience does not even hold in favor of the lognormal distribution, since hyperbolic-type distributions are just as convenient (Johnson & Kotz, 1970). Short of data, one should consider sensitivity analysis with distributions other than normal-related.

Since much of the behavior of interest for SCC compliance is in the tail of the output distribution, it is essential to justify any input distribution assumptions that affect the tail of the output distribution.

2.6 The Multivariate Normal Distribution

2.6.1 Tail Probabilities

The multivariate normal distribution, with its exceptional analytical convenience, represents a dangerous attraction to many practitioners. Indeed, an L -variate $\{X_l, l = 1, \dots, L\}$ multivariate distribution is fully characterized by:

~ the means: $\bar{X} = \{E\{X_l\}, l = 1, \dots, L\}$

~ the covariance matrix: $C_{ll'} = \{Cov\{X_l, X_{l'}\}, l, l' = 1, \dots, L\}$.

In addition all marginal and conditional distributions are also normal (Anderson, 1984).

However, this convenience is paid for by very specific properties that can make the multivariate normal model either inappropriate or a non-conservative approximation. In particular, the multivariate normal distribution entails a little known, yet severe, "destruction" (independence) of extreme values in the following sense (see Journel & Alabert, 1989): If $x_p, x_{p'}$ corresponding to high p, p' -quantile threshold values of the marginal distributions of, respectively, X_l and $X_{l'}$, then

$$\begin{aligned} \text{Prob}\{X_i > x_p, X_r > x_p\} &= \text{Prob}\{X_i > x_p\} \cdot \text{Prob}\{X_r > x_p\} \\ &= (1-p)(1-p), \text{ a very small value.} \end{aligned}$$

This result applies even if X_i and X_r are correlated.

In other words, when sampling from a multivariate normal distribution, one disallows the possibility of drawing a L -tuple $\{x_i^{(i)}, i = 1, \dots, L\}$, in which several of the values $x_i^{(i)}$ are simultaneously extreme. This is non-conservative if coincidence of extreme values are the causes of failure. This could happen even if the input variables are dependent.

2.6.2 Maximum Entropy

Some authors use the maximum entropy argument to justify their choice of multivariate normal-related models. Entropy is a measure of disorganization, and its maximization ensures that no structure beyond the imposed constraints is introduced into the model. It is known (Jones, 1979), that the multivariate distribution that maximizes entropy under the constraint of an imposed covariance is the multivariate normal distribution. If the additional constraint of a lognormal marginal distribution is imposed, then the maximum entropy distribution is multivariate lognormal.

Maximum entropy is certainly a quite reasonable criterion. However, the constraints should include all available information, whether quantitative or qualitative. There is clearly much more to, say, patterns of spatial dependence, than a mere histogram and covariance function, in which case the maximum entropy distribution would not necessarily be multivariate normal-related.

2.6.3 Normal-Score Transforms

Another frequent misconception is that one can easily transform any set of variables $\{X_i, i = 1, \dots, L\}$ into a set of multivariate normally-distributed variables $\{u_i, i = 1, \dots, L\}$ by a mere sequence of univariate transforms. This is not the case, unless the initial random variables X_i are mutually independent.

- Each variable X_i can indeed be transformed into a univariate normal u_i by the classical quantile function (normal-score) transform:

$$u_i = \Phi^{-1}(F_i(X_i)) \rightarrow N(0, 1),$$

where $\Phi^{-1}(\cdot)$ is the standard normal quantile function, or inverse standard normal Cumulative Distribution Function (cdf), and $F_i(\cdot)$ is the X_i -marginal cdf. However, the resulting L normal-score transforms u_i are not exactly multivariate normally-distributed, unless the X_i 's are mutually independent or already multivariate normal. This oversight is pervasive in the hydrogeological literature, where the log-transform of, say, transmissivity variables $T_i = T(x_i)$, is assumed to generate a multivariate normal field $\{\text{Ln } T(x_i), i = 1, \dots, L\}$.

- The correct approach (although not the only one) is to consider for normal-score transforms a sequence of conditional functions written as (Rosenblatt, 1952):

$$u_1 = \Phi^{-1}(F_1(X_1)), \text{ with } F_1(\cdot) \text{ being the marginal cdf of } X_1,$$

$$u_2 = \Phi^{-1}(F_2(X_2 | x_1)), \text{ with } F_2(\cdot) \text{ being the conditional cdf of } X_2 \text{ given } X_1 = x_1,$$

.....

$$u_l = \Phi^{-1}(F_L(X_L | x_1, \dots, x_{L-1})), \text{ with } F_L(\cdot | x_1, \dots, x_{L-1}) \text{ being the conditional cdf of } X_l \text{ given the } (L - 1) \text{ conditioning values } X_1 = x_1, \dots, X_{L-1} = x_{L-1}.$$

Clearly, this correct approach is much more demanding than a mere series of marginal normal-score transform $\Phi^{-1}(F_l(X_l)), l = 1, \dots, L$.

The previous discussion is not intended to discard the multivariate normal distribution model, which has a long record of satisfactory results in a large variety of studies. Rather, such a model should not be taken for granted, and its adoption should always be documented. If departure from the multivariate normal assumptions is shown to be inconsequential for the particular study at hand, then the multivariate normal model is a prime candidate because of its well-understood properties.

3. UNCERTAINTY EVALUATION METHODS

3.1 Introduction

The report by Zimmerman *et al.* (1990) addresses the problem of assessing compliance demonstration uncertainty by propagating data and parameter uncertainties through multiple complex transfer functions into a distribution characterizing the uncertainty of the response (output) values. It discusses four uncertainty analysis methods: (1) Monte Carlo simulation, (2) response surface methodology, (3) differential analysis approach, and (4) geostatistical techniques.

Although the four methods listed above should be considered as complementary tools rather than alternative approaches, they will be discussed separately.

3.2 Monte Carlo Simulation

Monte Carlo simulation is a sampling-based technique to map a multivariate input distribution into a possibly multivariate distribution of output values.

If the input variables $\{X_l, l = 1, \dots, L\}$ are in number L , and the transfer function ϕ (see Section 4.3 of the main report) is single-valued, each input L -tuple $\{x_l^{(i)}, l = 1, \dots, L\}$ would result in, say,

one K -tuple of output values $\{y_k^{(i)}, k = 1, \dots, K\}$. The Monte Carlo idea is to run N^1 such input L -tuples $\{x_l^{(i)}, l = 1, \dots, L\}, i = 1, \dots, N$ through ϕ resulting in a distribution of N output K -tuples $\{y_k^{(i)}, k = 1, \dots, K\}, i = 1, \dots, N$. Consider the simpler case, $K = 1$, of one single output value, say the time-to-failure of any particular WP.

The output distribution of the $y^{(i)}, i = 1, \dots, N$, depends of course on the sampling of the input \underline{X} space. The Zimmerman et al. report considers three types of sampling:

- **Random sampling.** A random sample of size N is a set of N independent observations on the vector $\underline{X} = (X_1, \dots, X_L)$ of input random variables. Each observation on \underline{X} is made in accordance with the assumed joint input distribution of X_1, \dots, X_L . Zimmerman et al. point out that random sampling can be prohibitively expensive for many computer models, since N can be very large if the sample is to adequately cover the ranges of the input variables.
- **Stratified sampling.** A stratified sample is a sample from a partition of the \underline{X} input space into n mutually exclusive and exhaustive L -variate cells or strata with known probabilities $p_j \geq 0$, such that $\sum p_j = 1$. (For example, the strata may be of equal probability: $p_j = 1/n$.) A stratified sample is a random sample from \underline{X} of size n_j from stratum j , where $\sum n_j = N$. **Systematic sampling** is stratified sampling where one observation is taken at the center of each stratum ($n_j = 1$). **Importance sampling** is stratified sampling where more observations are drawn in input areas of importance, for example, corresponding to extreme X_l -values. The **marginal sampling** proposed in Appendix B is a particular case of importance sampling where one tries iteratively to zoom toward those input L -tuples that yield a marginal response value, i.e., a response that barely passes compliance.
- **Latin hypercube sampling (LHS).** LHS is a special kind of stratified sampling in which the range of each input variable is divided into intervals of equal probability, and one value is then randomly selected from each interval. For a sample of size N , the range of each input variable $X_l (l = 1, \dots, L)$ is divided into intervals of probability $1/N$, and a value $x_{l,n}$ is chosen at random from the n^{th} interval ($n = 1, \dots, N$). The N values thus obtained for X_l are rearranged according to an independent random permutation of $\{1, 2, \dots, N\}$. The resulting permutations form an $L \times N$ matrix, with the l^{th} row consisting of the permuted values of X_l . The N columns of this matrix constitute the Latin hypercube sample.

LHS has the advantage of covering the range of each of the input variables. As is true for stratified sampling, it can be shown that, under appropriate conditions, estimates based on LHS have smaller variance than estimates based on random sampling. However, LHS has the disadvantages of not being able to accommodate a general dependency structure for the input variables. Even though LHS can be modified to accommodate some types of dependency, there is a danger that use of LHS could lead to biased estimates.

¹In this appendix, N denotes sample size.

In their pioneering paper, Mckay et al. (1979) consider only the case where the input variables X_i are independent. They prove that LHS is unbiased and, provided the transfer function is monotonic in each of its arguments, that LHS is better (i.e., has smaller variance) than random sampling for estimating the mean and distribution of the output Y . Stein (1987) proves that, asymptotically as the sample size $N \rightarrow \infty$, LHS is better than random sampling. While their results are interesting, they are not very useful since stratified sampling is also unbiased and is better than random sampling. Since random sampling is inefficient and is not recommended for problems involving complex transfer functions, LHS is an alternative to stratified sampling and should be compared with it. However, although numerical examples indicate that LHS is better than stratified sampling (e.g., see Mckay et al.), there is no general proof to date. Although Zimmerman et al. claim that LHS is better than either random or stratified sampling (page 5), it is clear from their Appendix A (based on Mckay et al.) that this is only true for the stratified sampling example cited in Mckay et al.

All of the results cited above assume that the input variables are independent. Recognizing that this is often not the case, Iman and Conover (1982) note that: "If a correlation structure exists among the input variables, but the actual sampling takes place as if the input variables were independent, the theoretical properties of the statistics formed from the input may no longer be valid. Estimators intended to be unbiased or consistent may not be." (page 331). In this paper, Iman and Conover present a method for approximately inducing a desired rank correlation structure, on the input variables, that preserves the exact marginal distributions and, with an example, show that it is better than LHS, based on independent inputs. Recognizing that "there is much more to a multivariate input distribution than a mere collection of marginal distributions and a covariance matrix" (page 331), Iman and Conover point out that, in practice, independence is often assumed for simplicity and argue that their method should be used whenever the input random variables are correlated.

Stein (1987) points out that Iman and Conover's procedure does not necessarily yield input variables with even approximately the correct joint distribution. Stein then describes a procedure "...for producing a Latin hypercube sample of size N such that each sample vector has approximately the correct joint distribution when N is large" (page 146). Stein makes no claim that his procedure is better than random sampling, even in the asymptotic case. (However, he does present an example that shows that LHS is a definite improvement over simple random sampling). In fact, he cautions the reader that, "...if N is not large enough, then the joint distribution of Z_j (the j^{th} sample vector) may be substantially distorted, which could lead to bias problems" (page 146).

Recommendation:

Although LHS has been used extensively in probabilistic risk assessments (e.g., USNRC, 1989), the input variables are generally assumed independent or have a simple dependency structure. Since, for the SCC uncertainty analysis, many of the input variables will have a complex dependency structure, it is not clear that LHS can yield unbiased estimators or that it is better than stratified sampling. Unless LHS can be shown to be an accurate and efficient procedure under these circumstances, it is recommended that the more widely used and more well-understood stratified sampling be used. To further reduce computational effort, Appendices B and C suggest more efficient importance sampling methods.

3.3 Response Surface Methodology

The response surface methodology consists of replacing the complex transfer function $Y = \phi(\underline{X})$ by a simple analytical function of a limited number L' of the initial L input variables, with $L' < L$. Most often that function is a linear combination of the $X_l, l = 1, \dots, L'$:

$$\hat{\phi}(\underline{X}) = a_0 + \sum_{l=1}^{L'} a_l X_l.$$

In a more complex case, the fitted transfer function may include some non-linear functions, e.g.,

$$\hat{\phi}(\underline{X}) = a_0 + \sum_{l=1}^{L'} a_l X_l + \sum_{l=1}^{L'} \sum_{\substack{k=1 \\ k \leq l}}^{L'} b_{lk} X_l X_k.$$

The coefficients a_0, a_l, b_{lk} are fitted by least squares from a series of runs of the actual transfer function $Y = \phi(\underline{X})$ yielding, say, N realizations: $\{x_l^{(i)}, l = 1, \dots, L; y^{(i)}\}, i = 1, \dots, N$.

Once the approximation $\hat{\phi}(\underline{X})$ is established, it is used in place of the actual (expensive) transfer function $Y = \phi(\underline{X})$ to propagate uncertainty on \underline{X} into a distribution for Y , in fact for $\hat{Y} = \hat{\phi}(\underline{X})$.

One shortcoming of the response surface methodology not fully recognized in the Zimmerman *et al.* report, is the smoothing effect of the least squares criterion for fit. The variance of the approximation \hat{Y} may be much less than that of the actual response Y . This reduction of variance and consequent reduction of distribution tails may not affect reproduction of the mean response, i.e.,

$$E\{\hat{Y}\} \approx E\{Y\},$$

but it may bias reproduction of the high p -quantile values in a non-conservative (over optimistic) direction:

$$\hat{y}_p \ll y_p, \text{ with } p \text{ large close to } 1,$$

and: \hat{y}_p, y_p being the p -quantiles of the distributions of, respectively, \hat{Y} and Y .

A solution to such smoothing of the tails is to consider a weighted least squares fit slanted toward better reproduction of the important tail, or to consider piece-wise response surfaces fit. However, such solutions may require additional information about the transfer function.

In repository performance assessment, in addition to the mean and median response values, the high p -quantiles may be very important. Therefore, the response surface methodology suffers a potentially severe limitation that makes it a non-recommendable technique except, possibly, for modeling accessory transfer functions with little consequence on the final critical response.

3.4 Differential Analysis Approach

The differential analysis approach is similar in principle to the response surface methodology and suffers from the same limitations. It consists of replacing the complex transfer function $Y = \phi(\underline{X})$ by a Taylor series expansion about some fixed base value, usually the mean input $\underline{\mu} = \{\mu_l, l = 1, \dots, L\}$:

$$Y = \phi(\underline{X}) \approx \phi(\underline{\mu}) + \sum_{i=1}^L \frac{\partial \phi(\underline{\mu})}{\partial X_i} (X_i - \mu_i).$$

From this first-order approximation, the mean and variance of the output Y are estimated as:

$$\mu_Y = E\{Y\} = \phi(\underline{\mu}) + \sum_{i=1}^L \frac{\partial \phi(\underline{\mu})}{\partial X_i} \cdot E\{X_i - \mu_i\} = \phi(\underline{\mu}),$$

which amounts, exactly, to assuming that the mean input $\underline{\mu}$ corresponds to the mean output $E\{Y\}$, and:

$$Var\{Y\} \approx \sum_{i=1}^L \sum_{i'=1}^L \frac{\partial \phi(\underline{\mu})}{\partial X_i} \cdot \frac{\partial \phi(\underline{\mu})}{\partial X_{i'}} \cdot Cov\{X_i, X_{i'}\}.$$

Again, linearization either by Taylor series or by a least squares fit cannot render the complexity of most transfer functions involved in a repository safety assessment. Also the variance approximation requires that each X_i -distribution be of small variance. Last, one should question the Taylor development about the mean $\underline{\mu}$ rather than about any other more appropriate (more extreme) multivariate p -quantile values of \underline{X} .

A second-order Taylor expansion would be better, but would call for second-order partial derivatives of ϕ and moments of order 4 of the input random vector \underline{X} in the calculation of $Var\{Y\}$.

The differential analysis approach yields estimates of the mean and variance of the response variable Y . One would need an additional distributional assumption to make probability statement about that response. When the goal of the analysis is to evaluate the tail distribution, the differential analysis approach may not be sufficient.

A method that may be used to improve the differential analysis approach to establish a response cdf without using higher-order derivatives is discussed in Appendix C.

3.5 Geostatistical Techniques

Geostatistics is a branch of applied statistics that deals with phenomena spread in space and/or time with some pattern of dependence. This pattern of dependence, which may be more complex than mere linear correlation, is inferred from data, modeled by one or a series of covariances, then put to use for space/time interpolation. Kriging, in its common (ordinary) form, is but a generalization of multiple linear regression, in which the independent variables (the data) can be inter-correlated to the same degree that they are correlated to the dependent variable (Journel, 1989 - Lesson 2).

Kriging and its variant cokriging has been widely used, and sometimes misused, in hydrogeological applications to provide input fields of transmissivity values to flow simulators. The same criticisms that were levied against response surface and differential analysis techniques can be applied to kriging: fields of kriging estimates are smooth, and do not reproduce the actual proportions and spatial patterns of connectivity of extreme input values that could lead to poor repository performance (Journel and Alabert, 1989). That smoothing effect is particularly severe when data are sparse, which is, unfortunately, a common situation in repository compliance assessment. Some authors, including Zimmerman *et al.*, argue that kriging provides, together with the least squares-type estimate, a measure of the corresponding estimation uncertainty. This is correct only in the very restrictive case when the spatial distribution under study, say that of transmissivity, can be modeled by a multivariate normal distribution possibly after a normal-score transform such as the logarithm. In all other cases, the kriging variance is but a ranking index of data configuration, a useful index, but yet a far cry from a measure of local estimation accuracy (Journel and Rossi, 1989). In any case, a contour map of local estimation variances cannot be input into a transfer function, such as a flow simulator, to propagate spatial estimation uncertainty.

Notwithstanding the hard sale of geostatistics and the common usage of kriging-cokriging in the hydrogeological literature, the only geostatistical algorithm that is relevant to propagation of uncertainty is that of "conditional simulation," (Journel, 1989 - Lesson 5).

Conditional simulation is an algorithm for generating alternative, equiprobable, spatial fields of dependent input values (say transmissivity) which honor data values at their locations and reproduce some prior measure(s) of spatial dependence, such as a variogram or covariance function. Conditionally simulated input fields have the distinct advantage, over most estimated fields, that they do not smooth actual spatial variability: proportions and spatial connectivity of extreme input values can be correctly reproduced.

Consider, for example, a set of L dependent input variables $\{X_l, l = 1, \dots, L\}$ with L possibly as large as 10^6 . The dependence between any two variables X_l, X_l' can be characterized by either the full bivariate probability distribution:

$$F_{ii'}(x, x') = \text{Prob}\{X_i \leq x, X_{i'} \leq x'\},$$

or any moment thereof such as the covariance:

$$\text{Cov}\{X_i, X_{i'}\} = E\{X_i, X_{i'}\} - E\{X_i\} \cdot E\{X_{i'}\}.$$

In the latter case, only linear correlation between X_i and $X_{i'}$ will be reproduced.

Conditional simulation allows the generation of equiprobable sets of L -tuples $\{x_l^{(i)}, l = 1, \dots, L\}, i = 1, \dots, N$, with N as large as necessary. Each set of L -tuples will match the prior model of spatial dependence, in addition to honoring the data values, that is:

$$x_{l'}^{(i)} = x_{l'}, l' \in (L'), \text{ for all } i = 1, \dots, N.$$

The conditional simulation algorithm allows equiprobable sampling from a, possibly very large, set of spatially dependent input variables. These N equiprobable sets $\{x_l^{(i)}, l = 1, \dots, L\}, i = 1, \dots, N$ could then be processed by the transfer function into a set of N equiprobable response value(s):

$$y^{(i)} = \phi(x_l^{(i)}, l = 1, \dots, L), i = 1, \dots, N.$$

From this latter set, a distribution (histogram) can be built, yielding a measure (more exactly a model) of response uncertainty. For example, the proportion of values $y^{(i)}$ exceeding the regulatory threshold y_{lim} provides an estimate of the probability of exceedance:

$$\text{Prob}\{Y = \phi(\underline{X}) > y_{lim}\}.$$

Geostatistical conditional simulations appear as a versatile convenient (very fast) sampling tool to be used in a Monte Carlo approach. It is misleading to present it as an alternative to such a Monte Carlo approach.

Bootstrap technique: Bootstrap and other resampling techniques (Efron, 1982), can be seen as particular stochastic simulation algorithms, whereby equiprobable sets of L – tuples are obtained by sampling with replacement from an original larger set of data values. Although widely used to assess the sampling distribution of statistics built on such L – tuples, resampling techniques appear limited in their potential applications to repository studies, for two reasons:

- (i) - the resampling algorithm usually requires that the original data be statistically independent, although this limitation can be somewhat removed by orthogonalization of the data (Solow, 1985).
- (ii) - more importantly, the initial data set must be large to very large to allow meaningful (non-redundant) resampling.

3.6 Propagation of Uncertainties through Multiple Models

Section 3 of the Zimmerman *et al.* report suggests using a single technique, the Monte Carlo, to propagate uncertainties through a sequence of models (transfer functions). The Monte Carlo approach is chosen for its versatility, absolute generality of application, and lack of constraining prior hypotheses, such as small uncertainty variances and linearization of the transfer functions. Zimmerman *et al.* insist on the necessity of proper interfacing of the various models, the output distribution of transfer function k being one of the input distributions of transfer functions $k+1$. They insist also on ensuring consistency of any approximation from one model to another. All these recommendations are reasonable and important.

The Monte Carlo approach is a viable technique to propagate uncertainty in a complex performance assessment. However, its implementation requires defining:

- The technique for sampling a high-dimensional multivariate input distribution
- The approximation needed to cut down the number of runs of an extraordinary complex series of transfer functions.

4. OTHER WASTE PACKAGE ISSUES

The report by Stephens *et al.* (1986) provides a review of U. S. Department of Energy (DOE) WP performance assessment methods and its application to the Basalt Waste Isolation Project (BWIP). At the time of that report, application of DOE methodology to the Yucca Mountain site had just started. Comments on some aspects of this report and two related papers by Liebetrau *et al.* (1987) and Ramspott (1988) are given.

These three documents do not propose any global and integrated compliance methodology, but all do recommend a probabilistic approach to compliance assessment. The report by Stephens *et al.* proposes some solutions, based on modularization and decoupling principles, to reduce the computational effort of a "global" Monte Carlo approach.

External failure sources

These three documents concentrated more on the problem of failure sources external to the WPs. Besides rock fall (very unpredictable), the major external source of WP failure is probably accelerated corrosion due to excess water getting in contact with the WP outer barrier.

Stephens *et al.* proposes a convolution - type approach to modeling the probability of sequential failure of concentric barriers going outward; this approach could be applied equally to the inward process of failure due to external water (and also rock fall). Then at a certain point, the two outward and inward failure processes would meet, leading to actual failure of the WP.

The geological processes generating intruding water can be assumed independent of both corrosion and number of failures, and could be simulated independently. The corresponding flow simulation could be inverted from the outward flow models envisioned by Sandia for study of the release rate.

The chain-effect

One important question not addressed in the reports is the modeling of a chain-effect: the influence of any one failed WP in shortening the remainder lifetime of nearby WPs. More generally, what is the space-time dependence between the failure times of any group of WPs? There is little doubt that there would be some clustering effect - in space and time - of failed WPs, if only due to common external sources of failures. Probabilistic evaluation of such clusters would not only help with the compliance assessment issue, but would also help with an engineering design that would minimize the occurrence of such clusters. For example, if one can predict the direction of future fractures (a rock mechanic problem), the original WP locations should be clustered in an orthogonal direction, to minimize the number of WPs affected by any single fracture.

5. REFERENCES

1. Ramspott, L. D., "Assessment of Engineered Barrier System and Design of Waste Packages," UCRL-98029 preprint, 1988.
2. Campbell, J. E. and R. M. Cranwell, "Performance Assessment of Radioactive Waste Repositories," Science, 239, pp. 1389-1392, 1988.
3. Bonano, E. J. and R. M. Cranwell, "Treatment of Uncertainties in the Performance Assessment of Geologic High-level Radioactive Waste Repositories," Math Geology, 20 (5), pp. 543-565, 1988.
4. Journel, A. G. and F. A. Alabert, "Non-Gaussian Data Expansion in the Earth Sciences," Terra Nova 1(2), pp. 123-134, 1989.
5. Berger, J. O., Statistical Decision Theory and Bayesian Analysis, Springer Verlag, p. 617, 1985.
6. Journel, A. G., "Fundamentals of Geostatistics in Five Lessons," Short Course in Geology, 5, Washington, D.C.: AGU publ., p. 40, 1989.
7. Hoeksema, R. J. and P. Kitanidis, "Analysis of the Spatial Structure of Properties of Selected Aquifers," WRR 21(4), pp. 563-572, 1985.
8. Desbarats, A., "Numerical Estimation of Effective Permeability in Sand-shale Formations," WRR 23(2), pp. 273-286, 1987.
9. Johnson, N. L. and S. Kotz, Continuous Univariate Distributions - 1, Wiley & Sons, 1970.
10. Anderson, T. W., Multivariate Statistical Analysis, Wiley & Sons, 1984.
11. Jones, D., Elementary Information Theory, Clarendon Press, 1979.
12. Rosenblatt, M., "Remarks on a Multivariate Transformation," Annals of Math. Stat., 23, No. 3, pp. 470-472, 1952.
13. Zimmerman, D. A. et al., "A Review of Techniques for Propagating Data and Parameter Uncertainties in High-level Radioactive Waste Performance Assessment Models," NUREG/CR-5393, 1990. (U.S. Nuclear Regulatory Commission, Washington, D.C.)
14. Mckay, M. D., W. J. Conover, and R. J. Beckman, "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," Technometrics, 21(2), pp. 239-245, 1979.
15. Stein, Michael, "Large Sample Properties of Simulations Using Latin Hypercube Sampling," Technometrics, 29(2), pp. 143-151, 1987.
16. Iman, R. L. and W. J. Conover, "A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables," Communications in Statistics, B11(3), pp. 311-334, 1982.
17. United States Nuclear Regulatory Commission, "Severe Accident Risks: An Assessment for Five U.S. Nuclear Power Plants," NUREG-1150, 1989. (Washington, D.C.: U.S. Nuclear Regulatory Commission)
18. Journel, A. G. and M. Rossi, "When Do We Need a Trend Model in Kriging?" Math Geology, 21(7), pp. 715-740, 1989.

19. Efron, B., The Jackknife, the Bootstrap, and Other Resampling Plans, Society for Industrial and Applied Mathematics, 1982.
20. Solow, A., "Bootstrapping Correlated Data," Math Geology, 17(7), pp. 769-775, 1985.
21. Liebetrau, A. M. et al., "A Probabilistic Source-term Code for Waste Package Performance Analysis," Waste Management 87, 2, pp. 535-544, 1987.
22. Stephens, K. et al., "Methodologies for Assessing Long-term Performance of High-level Radioactive Packages," NUREG/CR-4477, 1986. (U.S. Nuclear Regulatory Commission, Washington, D.C.)

6. BIBLIOGRAPHY

- Bonano, E. J. and P. A. Davis, "Demonstration of a Performance Assessment and Methodology for Nuclear Waste Isolation in Basalt Formations," SAND 86-2125C, 1986.
- Bonano, E. J. et al., "Demonstration of a Performance Assessment Methodology for High-level Radioactive Waste Disposal in Basalt Formations," NUREG/CR-4759, SAND 86-2325, 1989.
- Borgman, L. E. and Quimby, W. F., "Sampling for Tests of Hypothesis When Data are Correlated in Space and Time," Principles of Environmental Sampling, ed. L. Keith, pub. Amer. Chem. Soc., pp. 25-43, 1988.
- Cranwell, R. M. and J. C. Helton, "Uncertainty Analysis for Geologic Disposal of Radioactive Waste," SAND 81-1718C, 1981.
- Cranwell, R. M. et al., "Risk Methodology for Geologic Disposal of Radioactive Waste: Final Report," NUREG/CR-2452, SAND 81-2573, 1987.
- Davis, P. A. et al., "Components of an Overall Licensing Assessment Methodology," NUREG/CR-5256, SAND 88-3020, 1990.
- Gelhar, L. W. and C. L. Axness, "Three Dimensional Stochastic Analysis of Macrodispersion in Aquifers," WRR 19 (1), pp. 161-180, 1983.
- Gelhar, L. W., A. L. Gutjahr, and R. L. Naff, "Stochastic Analysis of Macrodispersion in a Stratified Aquifer," WRR 15 (6), pp. 1387-1397, 1979.
- Gomez-Hernandez, J. J. and A. G. Journel, "Conditional Simulation of the Wilmington Sand-shale Sequence," Hydraulic Engineering, ed. Ports, M. A., New York: ASCE, pp. 801-806, 1989.
- Hunter, R. L. and Mann, C. J. eds., "Techniques for Determining Probabilities of Events and Processes Affecting the Performance of Geologic Repositories: Literature Review," NUREG/CR-3964, SAND 86-0196, 1, 1989.
- Journel, A. G. and C. J. Huijbregts, Mining Geostatistics, Academic Press, 1978.
- Ripley, B. D., Stochastic Simulation, Wiley & Sons, 1987.

**APPENDIX B - PROBABILITY-DISTRIBUTION APPROACH
I - GENERAL**

Principal Author: Andre G. Journal

TABLE OF CONTENTS

	Page
1. MONTE CARLO APPROACH	B-1
1.1 Introduction	B-1
1.2 Sampling the Input Distribution	B-4
1.3 Determining Input Strata Probabilities	B-7
1.4 Stratified Sampling and Conditional Simulations	B-9
2. FAILURE BOUNDARY DETERMINATION USING IMPORTANCE SAMPLING	B-11
2.1 Introduction	B-11
2.2 Monotonic Transfer Functions	B-13
2.3 Determining Probability of Exceedance	B-14
3. REFERENCES	B-20
FIGURES	
B-1. Standard Monte Carlo Approach	B-2
B-2. Discretizing the Output Intervals	B-3
B-3. ForwardSimulation Approach	B-5
B-4. Determining the Compliance Probability	B-12
B-5. Iterative Search for a Marginal Point	B-16
B-6. Determination of the Output Cumulative Distribution Function (cdf)	B-19

APPENDIX B - PROBABILITY-DISTRIBUTION APPROACH I - GENERAL

1. MONTE CARLO APPROACH

1.1 Introduction

In Section 5 of the main report, a probabilistic Substantially Complete Containment (SCC) rule has been suggested that calls for knowledge of the distribution of the number $N(t)$ of waste packages failing in $[0, t]$. This number can be viewed as the output value Y of a transfer function, see Fig. 1 of the main report and Fig. B-1. Because the multiple input variables to that transfer function are uncertain, they are made (jointly) random, resulting in a random output variable $Y = N(t)$. The distribution of $N(t)$ is the basis for the proposed probabilistic rule.

The problem of determining a methodology for compliance demonstration is now cast as the very general problem of transferring an input uncertainty into an output uncertainty, see Fig. B-1. The standard Monte Carlo approach would consist of repetitive sampling of the input uncertainty space and a transfer of each sample i into the response value (or outcome) $y^{(i)}$, thus allowing a sampling of the Y -output distribution. This appendix discusses important limitations of this standard Monte Carlo approach and, together with Appendix C, proposes some implementation approximations. The implementation approximations proposed here are conceptual in nature. Actual implementations corresponding to a particular transfer function would be the responsibility of the repository designer.

In most applications, the standard Monte Carlo approach is unfeasible, for it would involve a huge number N^1 of runs of an expensive global simulation program denoted by the transfer function ϕ of Fig. B-1. A combinatorial exercise will help elucidate the problem.

Consider the following numbers:

- There are L input variables, some of them categorical, such as rock types, indicators of sub-processes to be used for the global transfer function ϕ (e.g., fracture - controlled flow vs. matrix flow), some of them continuous such as thermal and hydraulic conductivities and material characteristics of the waste packages. The number of uncertain parameters involved in performance assessment may be large.
- Each of these L input variables would require a different discretization level, depending on its nature and level of uncertainty, with a minimum of 2 for a binary variable and up to possibly 10 for a critical parameter with multimodal long-tail distributions. Consider an average of four classes per input variable, defining a total of $N = 4^{100} = 1.6 \times 10^{60}$ classes (!!) for $L = 100$.

¹In this appendix, N denotes sample size.

- There are $K = 5$ output variables whose uncertainty is to be characterized. Considering again an average of four classes for each of these output variables, there are 4^5 output classes. To sample completely a multivariate distribution, a minimum average of 10 outcomes per class would be reasonable, calling for $N = 10 \times 4^5 = 10240 \cong 10,000$, a large number yet much smaller than $N = 4^{100}$. Clearly, running an expensive global transfer function N times to allow exploration of all input classes is not sensible, nor is it needed. What is needed is a determination of the response probability (hyper-) surface constituted of a much smaller number 4^K of classes. Characterizing these 4^K classes requires a much smaller number N of carefully chosen (described below) runs. Yet $N = 10,000$ may still be too large for a complex global transfer function. Thus, there appears to be a need for approximating the standard Monte Carlo approach to make it feasible without biasing the desired answer, i.e., a response probability surface: $Prob\{Y_1 \leq y_1, \dots, Y_K \leq y_K\}$ or the most important parts thereof, see Fig. B-1. All approximations will aim at reducing the number of samples of the joint input distribution.

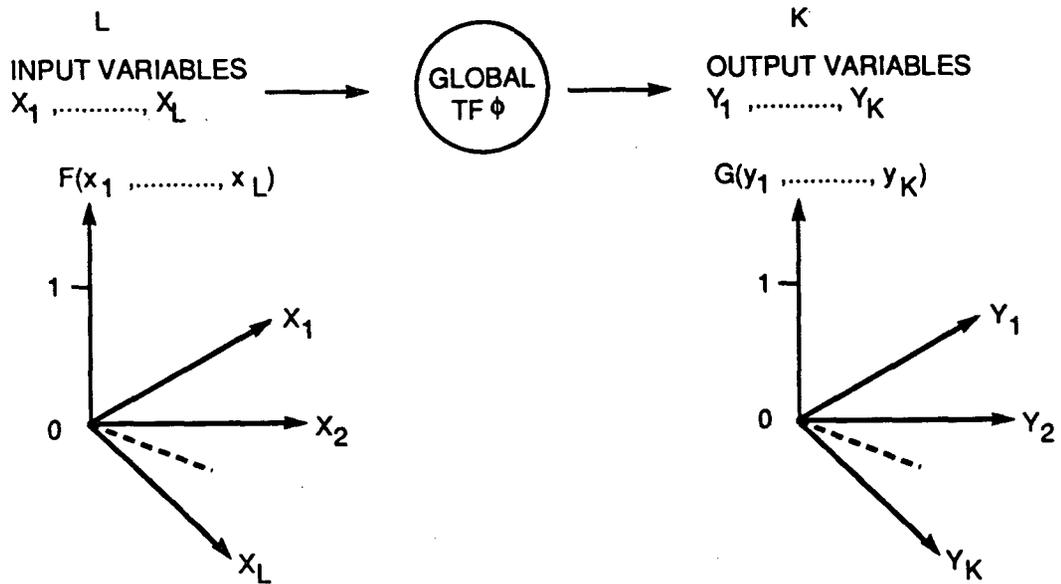


Figure B-1. Standard Monte Carlo approach.

Note: The input L -variate distribution is processed by the global transfer function into an output K -variate distribution (response surface): $F(x_1, \dots, x_L) = Prob\{X_1 \leq r_1, \dots, X_L \leq r_L\}$; $G(y_1, \dots, y_K) = Prob\{Y_1 \leq y_1, \dots, Y_K \leq y_K\}$.

Discretizing the output intervals

Not all 4^K classes of the output distribution are equally important; in general where the output values are more critical one should have a better class-resolution, whether in terms of class amplitude measured in the units of the Y_k 's or in terms of class probability.

In Fig. B-2, the $K = 2$ output variables are each classified into four classes of unequal amplitude.

All output classes should be characterized with a probability value for reason of completeness, even if they are unimportant. Also, output classes should be mutually exclusive to allow unequivocal determination of all class-probabilities, see Fig. A-1 and the related discussion.

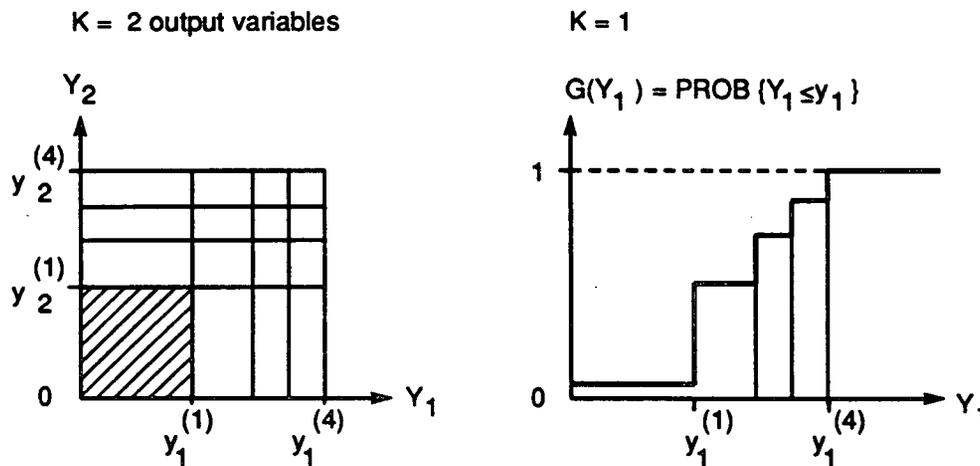


Figure B-2. Discretizing the output intervals.

Note: The critical intervals should be narrower in terms of probability. Non-critical output values can be lumped into a single interval, say for $K = 2: \{Y_1 \leq y_1^{(1)}, Y_2 \leq y_2^{(1)}\}$: hatched area.

Quantile approach:

Very often the whole output distribution is not required for compliance assessment. A quantile approach consists in retaining from the output distribution only one or a few critical probability intervals and/or quantile values.

For example, with $K = 1$ and a given compliance threshold y_{lim} the probability of exceedance may suffice:

$$\text{Prob}\{Y > y_{\text{lim}}\}$$

Alternatively one may only require evaluation of a high p -quantile of the Y -distribution, i.e., the value y_p such that:

$$\text{Prob}\{Y \leq y_p\} = p, \text{ with } p \text{ fixed large, e.g., } p = .95;$$

consequently: $\text{Prob}\{Y > y_p\} = 1 - p$.

With $K = 2$, if compliance requires both variables Y_1 and Y_2 to be below regulatory limits $y_{1,\text{lim}}$ and $y_{2,\text{lim}}$, one may require only evaluation of the probability of compliance:

$$\text{Prob}\{Y_1 > y_{1,\text{lim}}, Y_2 > y_{2,\text{lim}}\}.$$

In such cases there is no need to sample the entire input space so as to build the entire output distribution. Importance sampling scheme may be used to focus on those input areas leading to the relevant bounds of the output distribution. However, for regulatory purposes, it may be desirable to evaluate the y -distribution at values other than the regulatory limits. This is particularly true if the regulatory limits do not have a large built-in safety margin.

1.2 Sampling the Input Distribution

For determining the output class probabilities, the proportion of input sets leading to a response vector \underline{Y} falling into each of the output classes should be evaluated. The problem is thus the inverse mapping of the, say 4^K , classes of output into the corresponding 4^K classes of input values.

One can try either to invert (approximately) the various transfer functions constituting the global transfer function ϕ , or use a forward simulation approach by sampling the input distribution.

In the forward simulation approach, the input space is sampled: each point or sample i corresponds to a particular joint realization of the L input variables, say $\{X_1 = x_1^{(i)}, \dots, X_L = x_L^{(i)}\}$. Then, each point i is transferred (through ϕ) to a particular output class, see Fig. B-3.

At least five sampling schemes deserve consideration:

- **Random sampling.** This scheme is relatively straightforward but can be prohibitively expensive for many computer models, since N can be very large if the sample is to adequately cover the ranges of the input variables.

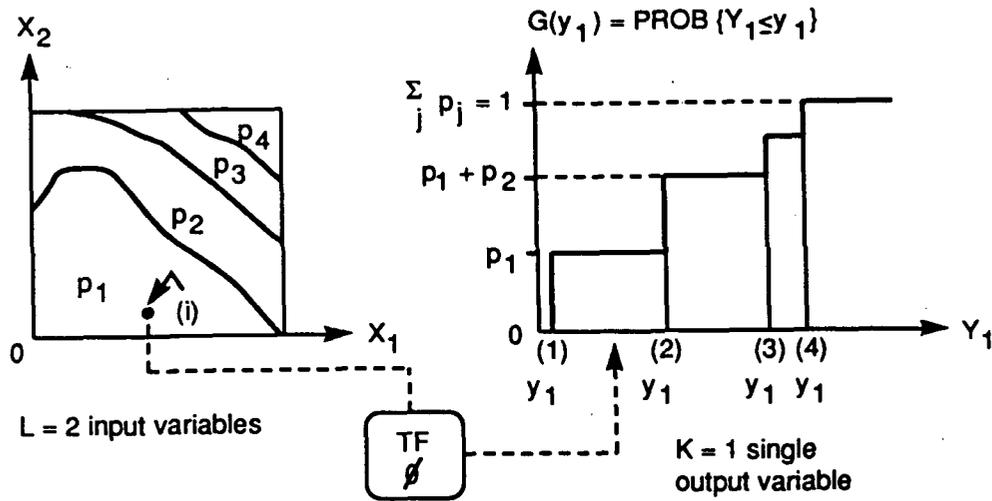


Figure B-3. Forward simulation approach.

Note: Forward transfer of selected points i of the input space (X_1, \dots, X_L) allows a probability mapping of classes of output values. The transfer function ϕ is assumed single-valued, i.e., the response to a given set i of input values is unique.

- **Stratified sampling.** The input space is partitioned into mutually exclusive and exhaustive strata as on the left of Fig. B-2, and each stratum is sampled at random by a single point. Stratified sampling protects from any accidental clustering of the sample points. The strata can be of any "size," as long as their respective probabilities are known. However, to allow mapping the input strata probabilities into output probabilities, there should be more input strata than output classes.
- **Systematic sampling.** Samples the input space at the nodes of a regular grid, either in terms of equal class amplitudes:

$$x_i^{(i+1)} = x_i^{(i)} + \Delta x, \text{ for all class bounds of } X_i,$$

or in terms of equal class probabilities:

$$\text{Prob}\{X_i \in [x_i^{(i)}, x_i^{(i+1)}]\} = p, \text{ constant for all classes } i \text{ of } X_i.$$

Systematic sampling is a special case of stratified sampling, where a unique stratum sample is taken at the "center" of that stratum. Again, the prior probability attached to each stratum must be known. When the L input variables are highly dependent and the number of strata retained is small with regard to L , systematic sampling should be the preferred choice, because it avoids any extreme location of the random sample in each stratum. However, if the input variables are known to be periodic, a regular grid in phase with the input periods could yield severely biased output results.

- Latin hypercube sampling (LHS). Since, for the SCC uncertainty analysis, many of the input variables will have a complex dependency structure, it is not clear that LHS can yield unbiased estimators or that it is better than stratified sampling. Unless LHS can be shown to be an accurate and efficient procedure under these circumstances, it is recommended that the more widely used and more well-understood stratified sampling be used.
- Scenario - type sampling. Define a scenario as a particular realization defining a set of conditions including events, features, processes, and the parameters in the transfer functions (see Appendix A for discussion). When N scenarios are selected, say by a panel of experts, this amounts to picking N points i in some regions in the input space, for example, the upper right-hand corner of the input space (X_1, X_2) of Fig. B-3. The major drawback of this approach is that the N scenarios S_i do not by themselves provide a set of probabilities $p_j, j = 1, \dots, J$, to be assigned to the J output classes as in Fig. B-3. Rather, N subjective probabilities π_i are assigned to the scenarios S_i , such that $\sum \pi_i = 1, i = 1, \dots, N$, and these probabilities π_i are assigned to the N corresponding output responses (Cranwell et al., 1987, p. 53).

Unless the N scenarios S_i represent a mutually exclusive and exhaustive partition of the input space as in Fig. B-2, and the probability elicitation program has been designed to ask the experts for the absolute probability of each stratum, the expert-originated values $\pi_i, i = 1, \dots, N$, are but a relative ranking of the N selected scenarios.

For a high reliability design, it is likely that the selected scenarios are clustered in some corner of the input (hyper-)space (see Appendix A), except for a "base case" isolated in the remainder of that space. Such clustering and the likelihood that the corresponding scenarios are partially redundant (non-mutually exclusive) might lead the experts to overestimate the relative probabilities of the corresponding strata; as a consequence, the probabilities of the corresponding responses would be too high (over-conservative).

Recommendations:

- For a full characterization of the output distribution, implement a stratified sampling with strata of equal joint-probability. The definition of such strata would require a careful modeling of the input space, including:

- ~ definition of input variables ranges and interdependence
- ~ definition of mutually exclusive and exhaustive strata and their respective probabilities
- ~ selection of one point (scenario) per strata

There should be at least as many input strata as there are output classes (two to three times more would be better).

Determination of strata probabilities would call for a careful evaluation of all patterns of dependence (possibly non-linear) between any number of input variables (see Section 1.3).

- For a partial characterization of the output distribution, such as through critical bounds or quantile values, implement an importance sampling focused on those "important" or "marginal" areas of the input space (see Section 2).

These recommendations do not preclude use of carefully elicited expert judgments to make up for lack of data.

1.3 Determining Input Strata Probabilities

Consider, for the sake of simplicity, a parallelepipedic strata defined from L input variables:

$$\{X_1 \in [x_1, x_1 + dx_1], \dots, X_L \in [x_L, x_L + dx_L]\}.$$

Any geometry-complex stratum can be approximated by a series of smaller parallelepipedic strata of the type above.

Evaluating that stratum probability amounts to the general problem of evaluating the joint probability:

$$P_L = Prob\{X_1 \in [x_1, x_1 + dx_1], \dots, X_L \in [x_L, x_L + dx_L]\}.$$

This joint probability can be decomposed into the product of L conditional probabilities:

$$\begin{aligned}
 P_L = & Prob\{X_1 \in [x_1, x_1 + dx_1]\} \\
 & \cdot Prob\{X_2 \in [x_2, x_2 + dx_2] \mid X_1 \in [x_1, x_1 + dx_1]\} \\
 & \cdot Prob\{X_3 \in [x_3, x_3 + dx_3] \mid X_1 \in [x_1, x_1 + dx_1], X_2 \in [x_2, x_2 + dx_2]\} \cdot \\
 & \dots \dots \dots \\
 & \cdot Prob\{X_L \in [x_L, x_L + dx_L] \mid X_1 \in [x_1, x_1 + dx_1], \dots, X_{L-1} \in [x_{L-1}, x_{L-1} + dx_{L-1}]\}. \quad (1)
 \end{aligned}$$

The previous decomposition is exact, but requires calculation of probabilities with increasing conditioning level. The last term, in particular, calls for the calculation of a probability interval of X_L conditioned by intervals on the $(L - 1)$ other variables. With L large, say $L = 100$, such calculation is not practical.

Grouping: The number of L can be substantially decreased by considering grouping of mutually independent input variables. Consider, for example, two groups of variables of respective size L_1 and L_2 , with: $L = L_1 + L_2$, such that the joint realization of any number of variables from the first group L_1 is influenced in no way by knowledge of any joint realization of any number of variables from the second group (L_2).

Then

$$\begin{aligned}
 P_L &= Prob \{X_1 \in [x_1, x_1 + dx_1], \dots, X_{L_1} \in [x_{L_1}, x_{L_1} + dx_{L_1}]\} \\
 &\quad \cdot Prob \{X_{L_1+1} \in [x_{L_1+1}, x_{L_1+1} + dx_{L_1+1}], \dots, X_L \in [x_L, x_L + dx_L]\} \\
 &= P_{L_1} \cdot P_{L_2}.
 \end{aligned} \tag{2}$$

For example, the group of input variables conditioning the internal waste package degradation model can be considered as independent from that conditioning the groundwater/rock mechanics model.

Screen effect: Within each of the previous groups, the number of conditioning variables can again be considerably reduced by using the "screen effect" or Markov effect paradigm commonly used in geostatistics (Journel and Huijbregts (1978) p. 346; Ripley (1987), p. 114). This paradigm amounts to retaining from a list of $(L - 1)$ variables $X_l, l = 1, \dots, L - 1$ conditioning an L^{th} one X_L , only that one which is the "most dependent" on X_L . Let $X_{L_0} \in [1, L - 1]$ be that variable most dependent on X_L , then:

$$Prob \{X_L | X_l, l = 1, \dots, L - 1\} \approx Prob \{X_L | X_{L_0}\}. \tag{3}$$

"Most dependent" can be interpreted in the traditional sense of linear correlation, or better in the sense of rank correlation, or better still in the sense of the conditioning variable that modifies most (for some distributional distance) the prior distribution of X_L .

Assuming that the L variables X_l have been ordered randomly and using the screen-effect approximation (3), the joint probability (1) reduces to a product of one-level conditional probabilities:

$$\begin{aligned}
P_L &= Prob\{X_1 \in [x_1, x_1 + dx_1]\} \\
&\cdot Prob\{X_2 \in [x_2, x_2 + dx_2] \mid X_1 \in [x_1, x_1 + dx_1]\} \\
&\cdot Prob\{X_3 \in [x_3, x_3 + dx_3] \mid X_3' \in [x_3, x_3 + dx_3']\} \\
&\dots\dots\dots \\
&\cdot Prob\{X_l \in [x_l, x_l + dx_l] \mid X_{l'} \in [x_{l'}, x_{l'} + dx_{l'}]\} \\
&\dots\dots\dots \\
&\cdot Prob\{X_L \in [x_L, x_L + dx_L] \mid X_{L'} \in [x_{L'}, x_{L'} + dx_{L'}]\}, \tag{4}
\end{aligned}$$

with: $l' \in [1, l - 1]$, $X_{l'}$ being the variable most dependent on X_l , among all previously considered variables.

One should make sure that the series of approximations involved in relation (4) do not all lead to over/underestimation, so that a compensation-effect can take place.

After appropriate groupings, L could be of the order of 10 for each group. Then, the screen-effect approximation allows calculation of the strata probability P_L from 10 conditioning probabilities involving each no more than one level of conditioning. This process may be tedious, but should be feasible.

Limited (in scope) expert judgments may be elicited for some of the conditional probabilities involved in expression (4). This is preferable to elicitation of a poorly understood scenario probability.

Remarks

The dependence of the approximation (4) on the random ordering of the L variables is a matter of concern, if these L variables are not all related to the same attribute, say porosity at L different locations, as they are in traditional geostatistical applications. Some sensitivity analysis would then be in order.

Because of the approximation (4), the sum of the N strata probabilities may not add exactly to one. A standardization correction would then be in order.

1.4 Stratified Sampling and Conditional Simulations

Stratified sampling can be combined with the geostatistical technique of conditional simulations, to sample some important strata more than others.

Consider the case of an input set $\underline{X} = \{X_l, l = 1, \dots, L\}$ comprised of two subsets:

- A set of spatial variables $X_l, l = 1, \dots, L_0$ e.g., $X_l = T(\underline{u}_l)$ is the transmissivity at node \underline{u}_l of some grid.
- A set of categorical parameters, $X_l, l = L_0 + 1, \dots, L$ defining the transfer function under study.

where L_0 could be of the order of 10^3 to 10^6 , whereas $L - L_0$ may be only of the order of 10, thus:
 $L_0 \gg L - L_0$.

Defining strata on the parameters $X_l, l = L_0 + 1, \dots, L$ is easy, particularly if they are categorical with a discrete distribution. Definition of strata on a very large multivariate spatial distribution is much more difficult: conditional simulations amount to equiprobable random sampling of the spatial input space $X_l, l = 1, \dots, L$, but are not yet a probability - specified stratified sampling (see Appendix A, Section 3.5, "Geostatistical Techniques," and Journel, 1989 - Lesson 5).

The idea is to proceed with stratification whenever possible, i.e., for those variables $X_l, l = L_0 + 1, \dots, L$, with each stratum having a specified joint $(L - L_0)$ - marginal probability, i.e., independently of the L_0 variables $X_l, l = 1, \dots, L_0$. Then, within each of these strata, proceed with random sampling of the remaining L_0 spatial variables using the same (for all strata) N conditional realizations $\{x_l^{(i)}, l = 1, \dots, L_0\}, i = 1, \dots, N$.

Example: Consider two sets of spatially dependent variables, say porosity and transmissivity defined over 10^3 grid blocks, thus: $L_0 = 2 \times 10^3$. Then consider a complementary set of $L - L_0 = 4$ parameters defining a particular flow simulator (transfer function). Assume each of these parameters binary, i.e.,

$$X_l = 1, \text{ with marginal probability } p_l$$

$$0, \text{ with marginal probability } (1 - p_l),$$

where

$$l = L_0 + 1, \dots, L_0 + 4.$$

The four parameters are assumed mutually independent (just for the present accounting), thus define $2^4 = 16$ strata with specified marginal four-variate probabilities $P_j, j = 1, \dots, 16$. For example, the first stratum, corresponding to all four parameters set to 1, has probability $P_1 = p_{L_0+1} \cdot p_{L_0+2} \cdot p_{L_0+3} \cdot p_{L_0+4}$, and $\sum P_j = 1$.

Each of these 16 strata are sampled by the same number, say $N = 100$, of equiprobable conditional realizations of the porosity-transmissivity field:

$$\{\phi^{(i)}(\underline{u}_l), T^{(i)}(\underline{u}_l), l = 1, \dots, 10^3\}, i = 1, \dots, N = 100.$$

Assuming that the L_0 spatial variables are independent of the four process parameters, there are $16 \times 100 = 1600$ input samples available, each of dimension $(2 \times 10^3 + 4)$, and with specified joint L -variate probability. For example, the realization $\{\phi^{(i)}(\underline{u}_l), T^{(i)}(\underline{u}_l), l = 1, \dots, 10^3; X_l = 1, l = L_0 + 1, L\}$ has probability $P_1/100$.

The number 1600 of input samples may still be too large if each sample is to be processed through an expensive, computer-intensive transfer function. One could consider cutting down that number using the following algorithm:

- Select out of the 16 sets of parameter values the one, j_0 , deemed most unfavorable, from the point of view of compliance. Process all 100 conditional realizations under the conditions j_0 , thus obtaining a set of 100 "unfavorable" response values. From a ranking of these response values, retain four conditional realizations, the three leading to the three output distribution quartiles and the one leading, say, to the 95th percentile of that output distribution.

Only these four selected conditional realizations are processed under the remainder 15 sets of process parameters. Thus the total number of runs of the transfer functions is: $100(1) + 4(15) = 160$, a more reasonable number.

The 100 initial realizations are conditional to j_0 , and each would receive 1/25th the weight given to each of the latter 60 realizations. Conditional to each set $j \neq j_0$ of the process parameters, the corresponding four runs can be considered to approximate the three conditional quartiles and the 95th percentile.

2. FAILURE BOUNDARY DETERMINATION USING IMPORTANCE SAMPLING

2.1 Introduction

As indicated in Section 1, the probabilistic approach requires an inverse mapping of the output classes into corresponding input classes (see Fig. B-3). Then, the input class probabilities would be determined and assigned to the relevant output classes.

Short of being able to invert the transfer function ϕ of Fig. B-4, the proposed approach consists of a forward simulation, mapping points (i) of the input space into the output classes. Also, it has been argued that the number of output classes could be quite small, possibly down to only two classes, as in Fig. B-4: the lower class corresponding to, say, compliance, and the complement class. In the case of Fig. B-4, compliance is defined by non-exceedance of the threshold value y_{max} applied to the single output variable Y , i.e., $K = 1$. For example, for SCC compliance, Y could be the random number of waste packages (WPs) having failed by time t .

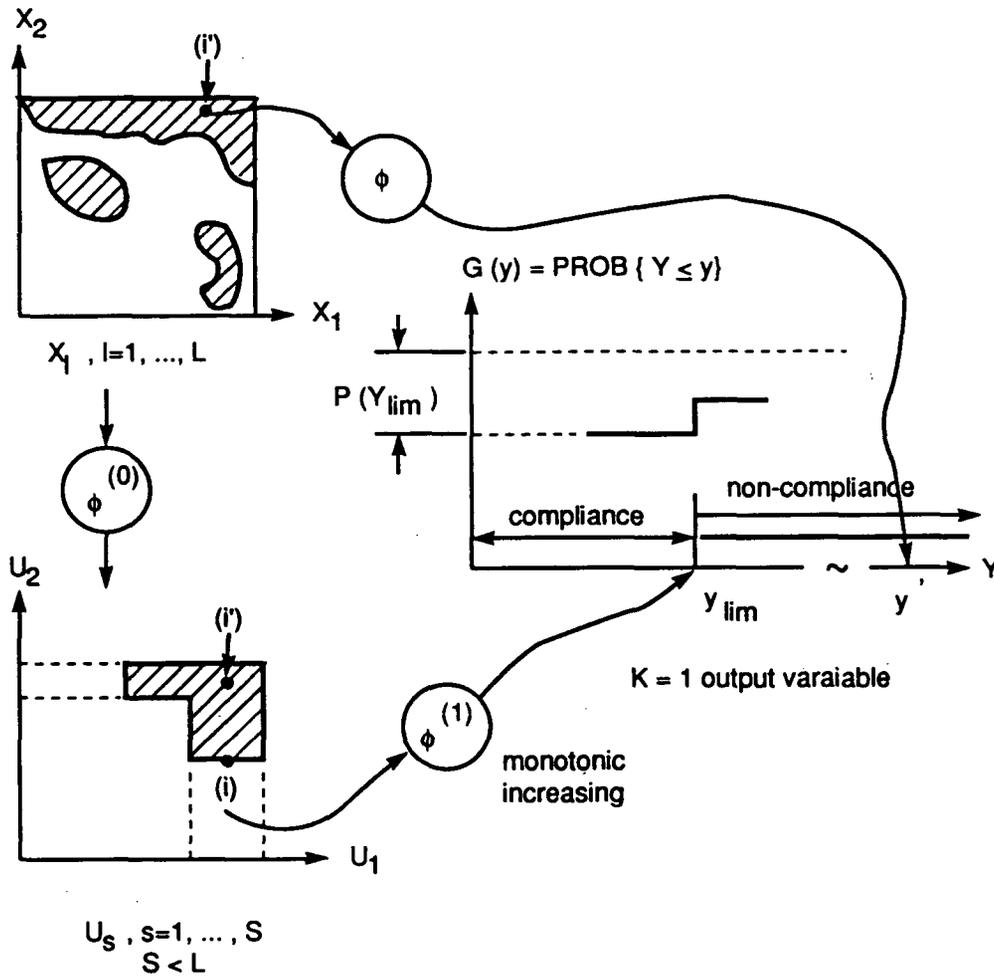


Figure B-4. Determining the compliance probability.

Note: The non-compliance probability $P(y_{lim}) = Prob\{Y > y_{lim}\}$ is made to correspond to a single-body area (hatched) in a space \underline{U} with reduced dimensionality ($S < L$). The boundary of that area is interpolated from a few carefully chosen points (i) corresponding to marginal conditions. Points (i') correspond to "extremely unfavorable" conditions.

In the rest of this section, for ease of notations, only the case ($K = 1$) of Fig. B-4 is considered; however, the arguments presented remain valid for an output surface defined by any number K of variables.

The proposed approach consists in a preferential (importance) sampling of the areas of the input space deemed to lead to "extreme" output classes, see Fig. B-4. "Extreme" is here understood as that (or those) output class(es) exceeding or leading to exceedance of regulatory limits. The

problem is that the extent, thus the probability of those input areas (hatched in Fig. B-4) are not known a priori. Worse, these critical input areas may be disconnected as in the upper left of Fig. B-4; indeed the transfer function $Y = \phi(X_1, \dots, X_L)$ need not be congenially monotonic.

When a transfer functions is well-behaved (e.g., can be approximated by using a number of well-selected polynomial functions - see Appendix C), a fast probability estimation technique as described in Appendix C can be used to identify the regions for importance sampling. However, a scheme is needed for more general transfer functions. In the following section a methodology is proposed to approximate the boundaries of these critical areas in the input space.

2.2 Monotonic Transfer Functions

A single-valued function $Y = \phi(X_1, \dots, X_L)$ of L variables X_L is said to be multivariate monotonically increasing if and only if: for any two input sets $\{x_1^{(i)}, \dots, x_L^{(i)}\}$, $\{x_1^{(j)}, \dots, x_L^{(j)}\}$ such that: $x_l^{(j)} \geq x_l^{(i)}$, $l = 1, \dots, L$, then:

$$Y^{(j)} = \phi(x_1^{(j)}, \dots, x_L^{(j)}) \geq \phi(x_1^{(i)}, \dots, x_L^{(i)}).$$

If a function $Y = \phi(X_1, \dots, X_L)$ is monotonically increasing, then to any set $\{Y > y_{lim}\}$ corresponds a "single-body" set $D(y_{lim})$, which may not be convex, as is the hatched area on the lower left of Fig. B-4.

Determining the boundaries, thus, the probability, of a single-body set is much easier than for a multiple-body set, such as represented by the hatched areas on the upper left of Fig. B-4. Thus, if the transfer function ϕ under consideration is not monotonically increasing, it should be decomposed into a sequence of two transfer functions $\phi^{(0)}$ and $\phi^{(1)}$, with the former ensuring that the latter is monotonic increasing, see Fig. B-4.

Among alternative pre-filters $\phi^{(0)}$ one would prefer the one that also, and in decreasing priority:

- defines transform variables U_s that carry physical significance, so that their probabilities of occurrence are easier to determine, possibly through expert judgment.
- reduces the dimensionality of the input space from L to $S < L$.
- maximizes "decoupling" between the variables U_s , i.e., the S variables U_s should be as much as possible mutually independent, again to allow an easier determination of probability of joint occurrence.

Through inverse-mapping, the output class-probabilities would be determined from single-body areas on the \underline{U} -input space, rather than from multiple-bodies areas on the original \underline{X} -input space, see Fig. B-4.

The pre-filter step $\phi^{(0)}$ amounts to the following recommendation, which may prove critical if the output probabilities are to be inferred from expert opinions in the failure boundary determination approach.

Recommendations:

Any scenario approach must ensure that the input space to be sampled is:

- of low dimensionality (S small), with easily interpretable quasi-independent variables
- such that the transfer function to the input space be single-valued and multivariate monotonic.

According to the prior recommendation given in Section 1, that input space would then be sampled with a stratified sampling scheme.

Remarks:

- If the final transfer function is monotonically decreasing, a very simple transform of its input variables U_s would make it monotonic increasing, e.g., $V_s = 1 - U_s$, or $V_s = 1/U_s$.
- It is yet unclear whether the critical step $\phi^{(0)}$ of pre-filtering a complex function to make it multivariate monotonic is always possible. Additional methodology research is here needed.
- If all S input variables U_s are mutually independent, then input areas corresponding to output intervals of the type $\{Y \in [y_j, y_{j+1}]\}$ would be parallelepipedic of the type $\{U_s \in [u_{s,j}, u_{s,j+1}], s = 1, \dots, S\}$. Thus, the stratified sampling should also be parallelepipedic (see Fig. B-2).

The non-rectangular hatched area on the lower left of Fig. B-4 implies that the two input variables U_1, U_2 to $\phi^{(1)}$ are not independent. If the S mutually independent variables are standardized to a common univariate distribution, say, uniform in $[a, b]$ for all s , or standard Gaussian, then all U -intervals should be of equal size i.e., $U_{s,j+1} - U_{s,j} = \Delta u_{s,j} = \text{constant}$ for all $s = 1, \dots, S$, but possibly: $\Delta u_{s,j} \neq \Delta u_{s',j}$, for $j \neq j'$

2.3 Determining Probability of Exceedance

Consider the case of only two output classes separated by the single threshold value y_{lim} , as in Fig. B-4. The exceedance class $\{Y > y_{lim}\}$ corresponds to the single body "safe" set:

$$D(y_{lim}) = \{U \text{ such that } Y > y_{lim}\}$$

Determination of the failure probability $P(y_{lim}) = P\{Y > y_{lim}\}$ calls for two steps:

- delineation of the boundary of the set $D(y_{lim})$ in the U -space, see hatched area on the lower left of Fig. B-4. In structural reliability this boundary hyper-surface is known as the "limit state surface."
- evaluation of the probability attached to the failure area, i.e.,

$$P(y_{lim}) = \int_{\{D(y_{lim})\}} dF_{U_1 \dots U_s}(u_1, \dots, u_s),$$

with $F_{U_1 \dots U_s}(u_1, \dots, u_s)$ being the joint cdf of the S input variables U_s .

Approximation solutions for evaluation of the probability integral $P(y_{lim})$ defined over a single-body failure set are well known in structural reliability (Madsen et al., 1986; Wu et al., 1989; Bjerager, 1989).

But first the boundary or limit state surface of $D(y_{lim})$ must be delineated, corresponding to step (i). It is proposed to interpolate this boundary, defined in the U space, by a few carefully chosen points (i) , $i = 1, \dots, N$ corresponding to "marginal" points. Marginal points represent those conditions (i.e., a combination of selected input parameters) that would barely pass the compliance criteria $Y \leq y_{lim}$.

To determine the boundary points (i) corresponding to a given threshold value y_{lim} , one could proceed with an iterative process as follows (see Fig. B-5), assuming that the transfer function is continuous.

- (1) Start with two bounding points (i_1) and (i_2) , deemed to result in the inequalities:

$$\phi^{(1)}(u_1^{(i_1)}, \dots, u_s^{(i_1)}) < y_{lim} < \phi^{(1)}(u_1^{(i_2)}, \dots, u_s^{(i_2)}).$$

- (2) Zoom progressively toward the corresponding boundary by considering a third point, for example:

$$(i_3): \{u_1^{(i_3)}, \dots, u_s^{(i_3)}\},$$

with: $u_s^{(i_3)} = (u_s^{(i_2)} + u_s^{(i_1)})/2$, for all $s = 1, \dots, S$.

If, e.g., $\phi^{(1)}(u_1^{(i_3)}, \dots, u_s^{(i_3)}) < y_{lim}$, then consider possibly the fourth point:

with: $u_s^{(i_4)} = (u_s^{(i_2)} + u_s^{(i_3)})/2$, for all $s = 1, \dots, S$.

Retain the π^{th} iteration as an approximation for a marginal point (i) located on the border of the hatched area $D(y_{\text{lim}})$.

- (3) Iterate the process N times, starting each time from step (1) with a different pair $((i_1), (i_2))$ and resulting in N marginal points $(i), i = 1, \dots, N$ defining the boundary of $D(y_{\text{lim}})$. Once the area $D(y_{\text{lim}})$ is defined, evaluate its probability $\text{Prob}\{\underline{U} \in D(y_{\text{lim}})\}$ possibly using the approximations proposed in Section 1. Then, and finally, the sought after probability of exceedance is:

$$P(y_{\text{lim}}) = \text{Prob}\{Y > y_{\text{lim}}\} = \text{Prob}\{\underline{U} \in D(y_{\text{lim}})\}.$$

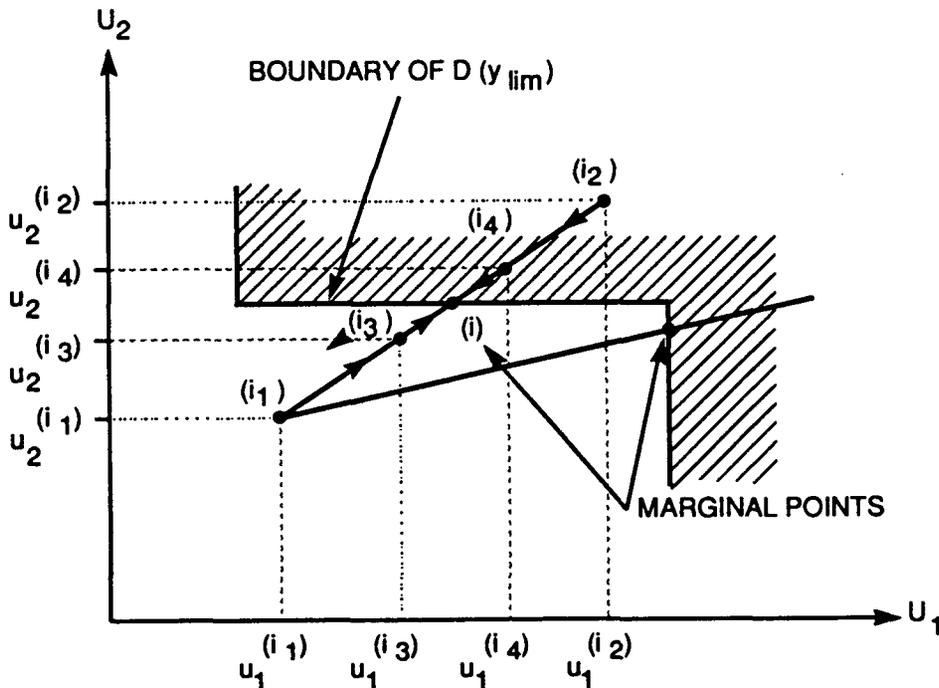


Figure B-5. Iterative search for a marginal point.

Note: The marginal point (i) on the boundary of the set D is approached iteratively by considering the sequence of points $(i_1), (i_2), (i_3), (i_4), \dots, (i_n)$.

One would stop at the π^{th} iteration when the consequence of the condition at (i_π) is judged close enough to the threshold value y_{lim} defining the set D .

Remarks

- If the transfer function $\phi^{(1)}$ is not continuous, the response limit value y_{lim} may not be reached, that is, there may not be any set of input values $(u_s, s = 1, \dots, S)$ such that $\phi^{(1)}(u_1, \dots, u_S) = y_{\text{lim}}$, see Fig. B-4. In such a case, the boundary corresponding to a conservative regulatory limit should be searched instead. In searching for the marginal points, it is important to identify this situation and prevent the iterative search from looping indefinitely.
- If the list of input variables u_s includes a categorical variable characterizing, for example, a model type (e.g., matrix flow vs. fracture flow), the transfer function is likely to involve severe discontinuities whenever that categorical variable changes value. In such a case, the discontinuity problem may be solved by finding the boundaries separately for each realization of the categorical variables, and then all the conditional probabilities should be assembled to evaluate the total probability $P(y_{\text{lim}}) = \text{Prob}\{\underline{u} \in D(y_{\text{lim}})\}$.
- The determination of probabilities in the input space, such as $\text{Prob}\{\underline{U} \in D(y_{\text{lim}})\}$, does not call for knowledge of any of the transfer functions ϕ or $\phi^{(1)}$; it calls for knowledge of the joint multivariate distribution of the S input variables U_s (see Section 1.2).
- During the search process, if a simpler approximation to $\phi^{(1)}$ can be found, such that the essential order relations are preserved (i.e., if (i) leads to a "pass" with the approximation model, it should also lead to a "pass" with the actual transfer function $\phi^{(1)}$, and, similarly for a "fail"), then the full (and computer-intensive) transfer function $\phi^{(1)}$ of Fig. B-4 should be replaced by the simpler function for those intermediate points (such as $(i_1), (i_2)$ on Fig. B-5), in order to reduce computational efforts.

Approximation models could range from a response surface model to an expert judgment about whether point (i) would pass or fail. However, the later iteration points, such as (i_3) and (i_4) on Fig. B-5, should make use of the exact transfer function, to ensure determination of a correct marginal point (i) .

- A parameterization of the input space \underline{u} , in terms of y -isopleth surface as in Fig. B-5, allows the evaluation of safety margins. For example, if the (uncertain) input variable u_s varies over a certain range $u_s \in [u_{s,\text{min}}, u_{s,\text{max}}]$, one can check if that range intersects the boundary of $D(y_{\text{lim}})$.

Some accounting: If an average of two iterations, calling for four points ($\pi = 4$ as in Fig. B-5), is needed to zoom into an approximation for a marginal point (i); and a maximum of 30 marginal points (i) is sufficient to delineate the boundary of $D(y_{lim})$ in a $S = 10$ - dimensional input space, then a maximum² of 120 points would have to be run, a large but not too excessive number.

Obtaining the full output distribution: What has been done for the evaluation of $P(y_{lim}) = Prob\{Y > y_{lim}\}$ can be repeated for several other threshold values y_{lim} , thus allowing us to approximate the output probability distribution $G(y) = Prob\{Y \leq y\}$, as shown at the right of Fig. B-6.

Note that the number N of points does not increase proportionally to the number J of threshold values $y_j, j = 1, \dots, J$ retained to discretize the output distribution, since the iteration processes for say, $G(y_1)$ and $G(y_2)$, can share several points, see Fig. B-6.

In summary, one can visualize the above proposed process for determination of the output probability distribution $G(y)$ as a point-by-point determination of (equal Y -response) contour lines defined on a carefully defined input space \underline{U} , see Fig. B-6.

Remarks

If we define a scenario as a particular realization defining a set of conditions including events, features, processes, and the parameters in the transfer functions, then the above approach can be classified as a scenario-type sampling scheme. The proposed approach provides an alternative means of determining the output probability distribution by focusing on the marginal conditions that would barely pass the compliance criteria. The approach is different from the more traditional scenario approach, in which a small number of points (i') are picked in the input space (usually \underline{X} rather than \underline{U}), yielding output values y' , see Fig. B-4. In practice, the probability attached to such a value y' , $Prob\{Y > y'\}$, cannot be determined easily.

The proposed approach provides a more rigorous determination of the output probability distribution, at the cost of an increased, yet manageable, number of scenarios to evaluate. As opposed to the traditional approach, the emphasis is not on selecting extreme scenarios, but in approximating marginal scenarios that would barely pass the compliance criteria. To reduce computational effort, one should investigate solutions for decoupling and streamlining all transfer functions to be run repetitively (see Appendix E).

²Some of the initial iteration points, say (i_1), can be used as seeds to zoom onto several different marginal points, see Fig. B-5.

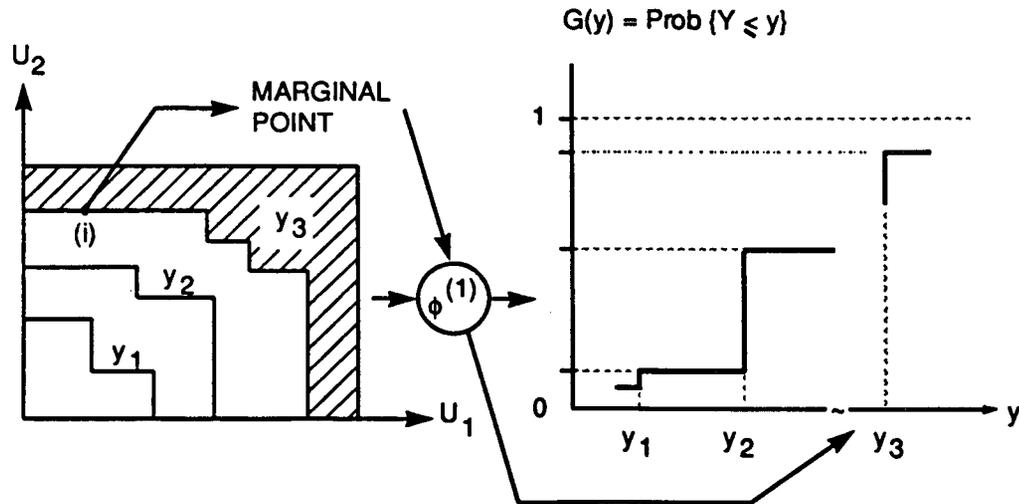


Figure B-6. Determination of the output cdf.

Note: The contour line (y_3) on the \underline{U} -space is the locus of all points (i) leading to a response value close to (y_3). The hatched area corresponds to the probability of exceedance: $1 - G(y_3) = Prob\{Y > y_3\}$. The transfer function ϕ is assumed singled-valued and monotonic increasing.

3. REFERENCES

1. Cranwell, R. M. *et al.*, "Risk Methodology for Geologic Disposal of Radioactive Waste: Final Report," NUREG/CR-2452, SAND 81-2573, 1987.
2. Journel, A. G. and C. J. Huijbregts, Mining Geostatistics, Academic Press, 4th print, 1989.
3. Ripley, B. D., Stochastic Simulation, Wiley & Sons, 1987.
4. Journel, A. G., "Fundamentals of Geostatistics in Five Lessons," Short Course in Geology, Vol. 5, AGU publ. Washington D.C., 1989.
5. Madsen, H. O. Krenk, S. & Lind, N. C., Methods of Structural Safety, Prentice Hall, New York, New York, 1986.
6. Wu, Y.-T. *et al.*, "Probabilistic Methods for Structural Response Analysis," Chapter 7, in Computational Mechanics of Probabilistic and Reliability Analysis, (ed.) Lin and Belytschko, publ. Elmpress, Switzerland, pp. 182-195, 1989.
7. Bjerager, P., "Probability Computation Methods in Structural and Mechanical Reliability," Chapter 2, in Computational Mechanics of Probabilistic and Reliability Analysis, (ed.) Lin and Belytschko, publ. Elmepress, Switzerland, pp. 48-67, 1989.

**APPENDIX C - PROBABILITY-DISTRIBUTION APPROACH
II - FAST PROBABILITY INTEGRATION AND IMPORTANCE
SAMPLING UPDATING**

Principal Author: Y.-T. Wu

TABLE OF CONTENTS

	Page
1. INTRODUCTION	C-1
2. REVIEW OF PROBABILISTIC METHODS	C-1
2.1 Analytical Approaches - General Concept	C-2
2.2 Analytical Approaches - Mean-Based Methods	C-5
2.3 Importance Sampling Methods	C-6
3. PROPOSED IMPORTANCE SAMPLING METHOD	C-7
3.1 General Formulation Using an Approximate Function	C-7
3.2 Example: Importance Sampling Using the AMV-Based Solution	C-8
4. NUMERICAL EXAMPLE	C-10
5. SUMMARY	C-15
5. REFERENCES	C-16
FIGURES	
C-1. Illustration of the Transformation and the Most Probable Point	C-4
C-2. Relationship between the Exact (y) Function and the Approximate (y_b) Function	C-7
C-3. Illustration of the AMV-Base Importance Sampling Method	C-9
C-4. Corrosion Depth Cumulative Distribution Function (cdf) Analysis Using the Advanced Mean Value Method	C-12
C-5. Cdf Analysis Using the Standard Monte Carlo Method	C-12
C-6. Cdf Analysis Using the AMV-Based Importance Sampling Method	C-13
TABLES	
C-1. Data Input for Demonstration Example	C-11
C-2. Importance Sampling Analysis Results	C-14

APPENDIX C - PROBABILITY-DISTRIBUTION APPROACH II - FAST PROBABILITY INTEGRATION AND IMPORTANCE SAMPLING UPDATING

1. INTRODUCTION

In Section 5 of the main report, a probabilistic Substantially Complete Containment (SCC) rule has been discussed that calls for the evaluation of the number $N(t)$ of waste packages failing in $[0, t]$. The distribution of $N(t)$, which is a complex function of the multiple uncertain input variables, is the basis for the quantitative probabilistic rule.

The Cumulative Distribution Function (cdf) of $N(t)$ can be evaluated using the standard Monte Carlo approach. However, as noted in Appendix B, it would involve a large number of repeated runs of a global simulation program - a task that may be prohibitively costly. To alleviate this problem, this appendix, together with Appendices B and E, suggest some practical approximate computational methods. Appendix B proposes a methodology for determining the failure-boundary (or limit-state surface) and computing the corresponding failure probability. Appendix E suggests methods for simplifying transfer functions. This appendix proposes an efficient procedure, based on approximate transfer functions.

When a transfer function is sufficiently "well-behaved" so that it can be approximated by using a number of appropriate polynomial functions, it is possible to quickly and adequately approximate the limit-state surface and the corresponding failure probability. In this appendix, a method is proposed that combines the fast probability integration concept with an importance-sampling scheme. The proposed method uses an approximate, simplified transfer function to define the initial sampling region and gradually increases the sampling region to update the approximate probability.

For complex but well-behaved functions, the method provides a potentially significant improvement over the standard Monte Carlo method. In principle, the method can also be applied to functions that are not well-behaved, provided that reasonably good approximate functions can be established to suggest importance sampling regions. In general, for transfer functions that are not well-behaved, the search procedure as suggested in Appendix B may be used to assist in defining the initial as well as the subsequent sampling regions.

The approximation techniques proposed here are conceptual in nature. Actual implementation corresponding to a particular transfer function would be the responsibility of the repository designer.

2. REVIEW OF PROBABILISTIC ANALYSIS METHODS

This section provides a brief review of probabilistic analysis methods, with emphasis on those requiring minimum function evaluations.

Recent probabilistic analysis methods based on "limit state" (the boundary between "safe" and "fail") formulations and the most probable point concept (described below) have been developed to evaluate structural safety (Ang and Tang, 1984; and Masden *et al.*, 1986). These methods generally employ performance sensitivity calculations and can incorporate efficient sensitivity calculation schemes such as the adjoint method and perturbation algorithms. These methods are analytical in nature and are very efficient relative to simulation methods, and they provide information on the sensitivity or importance (ranking) of random variables. However, there are two problems associated with these methods. (1) If the performance functions or transfer functions are complicated, the solution procedure tends to be cumbersome and possibly too costly. (2) The analyses do not provide error estimates, although "good" accuracies for a wide range of engineering problems have been reported in the literature.

A solution to the first problem described above has been proposed recently, using an advanced mean-based method (Wu *et al.*, 1989; and Wu *et al.*, 1990). In applying this method, the probability solution starts from conventional mean-based approximation. Next, the solution is improved by correcting the predicted performance function values and the performance sensitivities at some properly selected points. This solution procedure may be close to the optimum in requiring a minimum number of function evaluations relative to the accuracy of the performance cdf. However, the advanced mean-based method by itself is an approximate scheme. To confirm or enhance the accuracy, a probability analysis method that combines the fast probability integration technique and an importance sampling method will be proposed in Section 3 of this appendix.

2.1 Analytical Approaches - General Concept

Let a transfer function (or performance function) be defined as:

$$Y = \phi(X_1, X_2, \dots, X_L), \quad (1)$$

where X_i are the input random variables; then the cdf of Y can be computed as

$$P(Y < y_0) = F_Y(y_0) = \int_{Y < y_0} \dots \int f_{\underline{X}}(\underline{x}) d\underline{x}, \quad (2)$$

where y_0 is a specific value of the random variable Y , $f_{\underline{X}}(\underline{x})$ is the input joint probability density function, and $Y < y_0$ is the region of integration. This multiple integral is, in general, very difficult to evaluate. Alternatively, a Monte Carlo solution provides a convenient but usually time-consuming approximation.

For many engineering problems, efficient approximate solutions can be obtained by the recently developed structural reliability analysis methods (Ang and Tang, 1984). To apply these methods, the first step involves the formulation of the limit-state surfaces for a given ϕ -function. Then the ϕ -function is approximated by a polynomial at the most probable point (defined below).

Because the error in the approximated ϕ -function tends to be small around the most significant probability region, fairly accurate solution may be achieved, provided that the ϕ -function is "well-behaved."

The limit state is defined as:

$$Y = y_0 \quad (3)$$

This limit state separates the variable space into two regions: e.g., in structural design, "safe" and "fail." If the cdf of Y is of interest, then a number of y_0 values can be selected to establish $F_y(y_0)$.

The approximation procedure involves the identification of the most probable point, which is defined in the standardized normal (Gaussian) space using the transformation described below.

In general, the non-normal \underline{X} variables can be transformed to the standard normal variables \underline{u} using the following Rosenblatt transformation (Rosenblatt, 1952):

$$\begin{aligned} u_1 &= \Phi^{-1}[F_1(x_1)] \\ u_2 &= \Phi^{-1}[F_2(x_2 | x_1)] \\ &\cdot \\ u_L &= \Phi^{-1}[F_L(x_L | x_1, x_2, \dots, x_{L-1})], \end{aligned} \quad (4)$$

where F_i is the original cdf and Φ^{-1} is the inverse standardized normal cdf. When the variables are mutually independent, the transformation reduces to:

$$u_i = \Phi^{-1}[F_{X_i}(x_i)]. \quad (5)$$

The inverse transformation is:

$$x_i = F_{X_i}^{-1}[\Phi(u_i)]. \quad (6)$$

In many engineering applications, only marginal distributions and correlation coefficients are available. In such cases, Eq. 5 can be used and the corresponding correlation coefficients between u_i variables can be computed (Wu *et al.*, 1989). Afterwards, uncorrelated normal variables can be generated by a standard transformation procedure.

By transforming $\phi(\underline{X})$ to $\phi(\underline{u})$, the most probable point (point on the limit state surface with maximum joint probability density) in the u -space, \underline{u}^* , is the point that defines the minimum distance, β , from the origin to the limit state surface (see Fig. C-1). In this appendix, this point will be used as an approximation point. An approximation point is defined here as a critical point that minimizes the probability error when the original function is replaced by an approximate function.

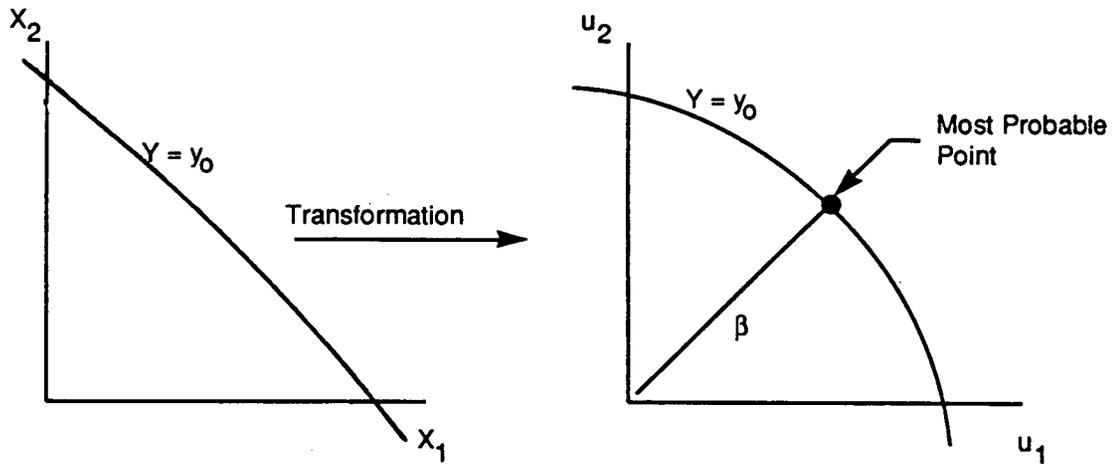


Figure C-1. Illustration of the transformation and the most probable point.

The minimum-distance point may be found by using the optimization or iteration schemes. Afterwards, approximate $\phi(\underline{u})$ function can be established by fitting, in the vicinity of the approximation point \underline{x}^* , the exact function with a simple (e.g., a linear or quadratic polynomial) function. Once the approximate function is obtained, the probability estimate can be easily estimated.

In the transformed u -space, approximate results are available for linear and quadratic functions (Tvedt, 1990). For example, for unique \underline{u}^* , the first-order probability estimate is:

$$P(Y \leq y_0) \approx \Phi(-\beta) \quad (7a)$$

and the asymptotic second-order probability estimate is:

$$P(Y \leq y_0) \approx \Phi(-\beta) \prod_{j=1}^{L-1} (1 + \beta \kappa_j)^{-1/2}, \quad \beta \rightarrow \infty, \quad (7b)$$

where $\kappa_j, j = 1, \dots, L - 1$ are the main curvatures of the limit-state surface at \underline{u}^* .

The key to the above fast probability integration approach is the ability to identify one or more approximation points at the "probability-significant" regions. If good approximation points can be identified, not necessarily the most probable points in the transformed space, then the approximation can be extended to include higher-order (i.e., higher than second-order) polynomials in either the original X -space or a transformed space. Once a polynomial is available, one can apply a standard Monte Carlo approach or other more efficient approximation techniques to estimate the probability. With the current computer technology, it is feasible to apply the standard Monte Carlo approach to problems involving simple functions such as polynomial equations.

It should be pointed out, however, that higher-order approximations are more difficult to construct and require more transfer function calculations. In practice, therefore, some degree of engineering judgment is necessary in order to justify the use of reasonably low-order polynomial approximations. In addition, efficient checking methods, such as the one described below, are needed to confirm or update the solutions.

By repeatedly applying the above approximate methods for a number of different y_0 , the cdf of Y can be constructed. This implies that an approximate function is established for *each* limit state. The fact that the entire Y -function is being represented by more than one approximating function (one for each probability level) is the main reason that low-order approximate functions are "sufficient" for many engineering problems. For the case involving one random variable, this concept is similar to fitting a nonlinear curve with splines.

The above approach has been used quite successfully for structural reliability problems (Ang and Tang, 1984; Madsen *et al.*, 1986; Wu *et al.*, 1989; and Wu *et al.*, 1990). However, when the transfer function Y is implicitly defined and is highly nonlinear, it may be difficult to assess the error due to the low-order approximation. Furthermore, inefficiency and convergence instability may occur due to the limitations in the optimization algorithms and the existence of multiple minimum distance points. A procedure that is robust and fast will be described in the next section. The discussions will be limited to well-behaved functions, namely the Y -function can be approximated by a number of low-order polynomials. For more general functions, the search-procedure for determining the limit-state boundary, as discussed in Appendix B, may be used to define the important regions. Also, other techniques for establishing good approximate functions should be sought.

2.2 Analytical Approaches - Mean-Based Methods

Mean-based analytical methods involve the development of approximate ϕ -function based on the mean values. Assuming that the ϕ -function is "smooth" or can be smoothed, the ϕ -function can be expressed as:

$$\begin{aligned}\phi(\underline{X}) &= \phi(\underline{\mu}) + \sum_{i=1}^L \left(\frac{\Delta\phi}{\Delta X_i} \right) \cdot (X_i - \mu_i) + H(\underline{X}) \\ &= a_0 + \sum a_i X_i + H(\underline{X}) \\ &= \phi_1(\underline{X}) + H(\underline{X}),\end{aligned}\tag{8}$$

where the sensitivities $\Delta\phi/\Delta X_i$ are evaluated at the mean values; $\phi_1(\underline{X})$ is a random variable representing the sum of the linear terms and $H(\underline{X})$ represents the higher-order terms.

For complicated, implicit, performance functions requiring computer algorithms such as finite element or finite difference methods, there are several ways of obtaining a_i , including numerical differentiation methods, the adjoint method, and the least-squares method. When numerical differentiation is applied, the minimum required number of ϕ function evaluations for computing the a_i is $(L+1)$.

Since the ϕ_1 -function is linear and explicit, its cdf can be computed effectively using many methods, including the analytical methods mentioned earlier and the convolution theorem. For brevity, the mean-value first-order (MVFO) solution based on $\phi_1(\underline{X})$ will be called the mean-value (MV) solution. The MV method has also been called differential analysis (see Appendix A).

For nonlinear ϕ -functions, the MV solution is, of course, not accurate. To improve the accuracy, an advanced mean value (AMV) method was developed (Wu *et al.*, 1989; and Wu *et al.*, 1990) to compute the cdf. The AMV method reduces the truncation error by approximating the random function $H(\underline{X})$ as a deterministic function $H(\phi_1)$ dependent on ϕ_1 . The H-function is defined using the most probable point of ϕ_1 . The AMV solution is expressed as:

$$\phi^*(\underline{X}) = \phi_1(\underline{X}) + H(\phi_1). \quad (9)$$

Based on the MV solution, a stepwise AMV procedure can be outlined as follows:

- Obtain the linear approximation $\phi_1(\underline{X})$ using Eq. 8.
- Compute the cdf of $\phi_1(\underline{X})$ at selected y_o points using the analytical methods.
- Select a number of cdf values that cover a sufficiently wide probability range.
- For each cdf value, identify the most probable point \underline{x}^* .
- Re-compute $\phi(\underline{x}^*)$ to replace y_o for the same cdf in the previous step.

The recomputation of $\phi(\underline{x}^*)$ corrects y_o by including the deterministic $H(\phi_1)$ function. For m cdf values, the total number of the ϕ -function evaluations is $(L + 1 + m)$.

The above AMV method is based on approximate most probable point. With more function calculations, the accuracy can be improved based on exact most probable point(s) derived using proper iteration algorithms (Wu *et al.*, 1989; and Wu *et al.*, 1990).

2.3 Importance Sampling Methods

The importance sampling method is based on the idea of sampling only in the critical region. The reason for doing this is that relatively small sample sizes can produce point probability estimates having relatively narrow confidence intervals. Based on this concept, various schemes have been developed particularly for structural safety analysis (Harbitz, 1986). In these methods, the sampling starts from the standardized Gaussian space.

3. PROPOSED IMPORTANCE SAMPLING METHOD

The following Section 3.1 discusses the general concept of using an approximate ϕ -function to define the importance sampling region. Section 3.2 presents a sampling procedure using the AMV solution to define an approximate function. There is no unique way of selecting an approximate function. The reason for selecting the AMV solution is for convenience. In practice, more efficient solutions may be achieved by using more appropriate approximation functions.

3.1 General Formulation Using An Approximate Function

Assume that an approximate, analytical function, $Y^* = \phi^*(\underline{X})$, to the exact function, Y , can be found by some method (e.g., coarse finite element model, mean expansion, response surface, AMV, etc.). Let

$$\phi(\underline{X}) = \phi^*(\underline{X}) + \varepsilon(\underline{X}) \quad (10)$$

where the error function, $\varepsilon(\underline{X})$, is a random variable. Consider a special case where Y^* is a monotonic function of a "base" random variable Y_b , which has easily determined distribution. An example of defining Y_b is included in Section 3.2. Now consider a Monte Carlo simulation. Because of the presence of $\varepsilon(\underline{X})$, the relationship between Y and Y_b is generally random, as illustrated in Fig. C-2.

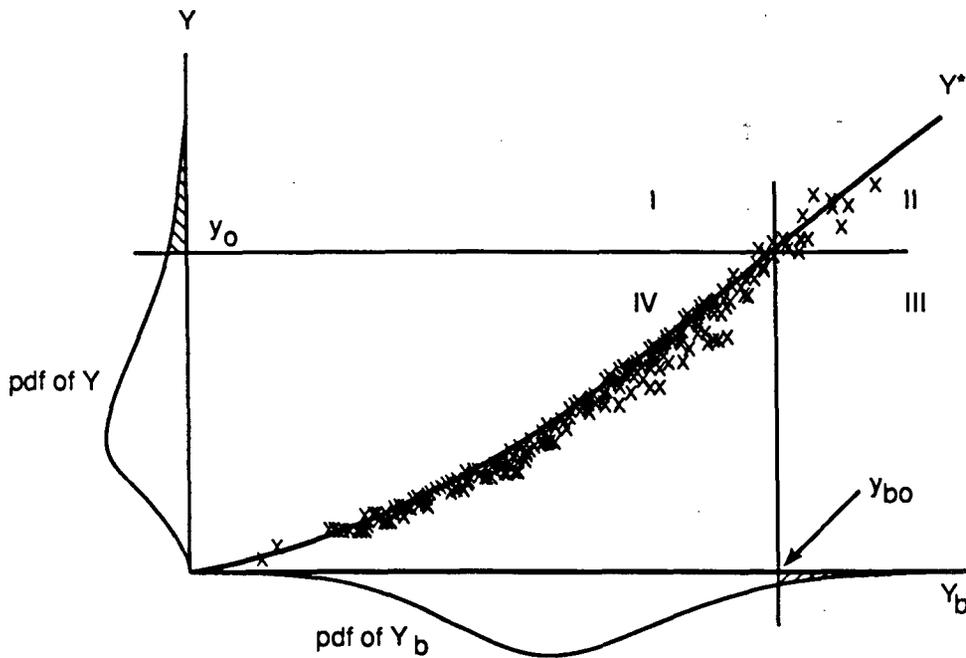


Figure C-2. Relationship between the exact (y) function and the approximate (y_b) function.

In general, it is not feasible to find a Y_b such that $\epsilon(\underline{X})$ is insignificant. However, Fig. C-2 suggests that: if the number of samples in region I (N_I) equals the number of samples in region III (N_{III}), then

$$P(Y > y_0) = P(Y_b > y_{b0}), \quad (11)$$

even if $\epsilon(\underline{X})$ is "large." The two numbers N_I and N_{III} correspond to misclassification of "fail-safe" conditions; letting $N_I = N_{III}$ provides a compensation to the misclassification errors. Thus, given a "sufficient" number of random samples (by increasing samples and sampling region), the point probability of Y can be estimated by finding a point Y_{b0} , which provides the condition $N_I = N_{III}$, and by using the (more easily computed) distribution of Y_b .

Fig. C-2 also suggests that, when $\epsilon(\underline{X})$ is reasonably small, one does not need to concentrate on the entire region of Y_b for computing the point probability of Y . In particular, when the "tail distribution" is of major interest, it is preferred to generate sufficient N_I samples and minimum N_{IV} samples.

The above concept can be extended to non-monotonic functions, where, for each $Y^* = y_0$, there exist multiple Y_{b0} values.

When Y is a complex function requiring extensive computations and Y_b is a simple analytical function (e.g., a polynomial), a simple, efficient procedure can be designed as follows (Wu and Nair, 1988):

- Generate \underline{X} samples and compute Y_b .
- Compute Y only if Y_b is in the important region, e.g., the tail region.

This is a two-stage sampling procedure in which the first stage uses the Y^* values to filter the undesirable random samples. Only the selected samples will go to stage 2, where the "expensive" calculations of the ϕ -function will be performed.

A more general approach to directly generate the desired \underline{X}^* samples is described in the next section, using the μ -space as the starting sampling space.

3.2 Example: Importance Sampling Using the AMV-Based Solution

Applying the previous formulations, the AMV solution can be used, as one of the possible choices, to provide a basis for defining Y_b , Y^* and the importance sampling region.

Let

$$Y_b = Y_1, \quad (12)$$

where Y_1 is a linear function of \underline{X} , and the AMV approximation is:

$$Y^* = Y_1 + H(Y_1), \quad (13)$$

where H is defined in Eq. 9. Note that there are various ways of defining Y_b ; e.g., a second-order polynomial of \underline{X} may be used. Of course, it is better if Y_b closely approximates Y .

Given a limit state $Y = y_0$, the AMV approximation is:

$$Y_1 + H(Y_1) = y_0. \quad (14)$$

In the u -space, the most probable point \underline{u}^* satisfying the limit state is defined by the vector OQ shown in Fig. C-3. The minimum distance is β .

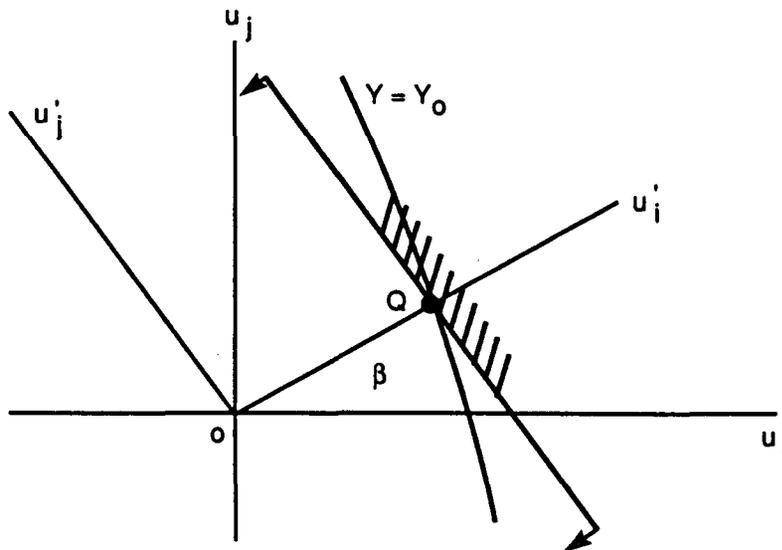


Figure C-3. Illustration of the AMV-based importance sampling method.

Using the rotationally symmetric properties of the Gaussian joint probability density in the u -space, a procedure to generate selected samples is as follows:

- Perform orthogonal transformation of \underline{u} to \underline{u}' such that the vector OQ in Fig. C-3 coincides with u'_1 .
- Select an initial distance $\delta < \beta$.

- Generate a u_1' sample such that $u_1' > \delta$.
- Generate random samples for the remaining u -variables.
- For each u' point, compute the u point.
- Compute X from u .
- Compute $Y_b(X)$ and $Y(X)$.
- Repeat the above five steps as necessary and count numbers of samples: N_I , N_{II} , and N_{III} .
- Reduce distance δ to generate additional samples until δ is sufficiently small.

The convergence criteria for δ include the convergence of y_{b0} (evaluated by letting $N_I = N_{III}$) and the convergence of the sum ($N_I + N_{II}$). An optimum δ is the one that minimizes N_{IV} . Finally, the probability estimate is computed using y_{b0} and Eq. 11, taking advantage of the fact that Y_b is a simple function.

The above procedure may be improved by using the exact most probable point (derived, for example, using the AMV-based iteration procedure) (Wu *et al.*, 1989; and Wu *et al.*, 1990), to define the orthogonal transformation. The reason is that more samples generated in this way will be closer to the most probable point. In terms of the regions defined in Fig. C-2, this means that fewer samples will be generated in Region IV, implying higher efficiency.

4. NUMERICAL EXAMPLE

A waste package corrosion process model (Sutcliffe, 1984) is used to demonstrate the proposed sampling procedure. The performance function is the corrosion depth of a waste package and is expressed as a function of three random variables: chloride concentration (Cl), an empirical constant related to time (d), and a pitting factor (K_p). The present example uses a simple explicit function to facilitate the generation of an "exact" probabilistic solution for comparison purposes. It should be emphasized that the methodology is equally applicable to implicit performance functions (Wu *et al.*, 1990).

Assume that the corrosion depth can be predicted using the following formula:

$$C = K_p K \exp(a/H) O^b Cl^c t^d, \quad (15)$$

where K is the uniform corrosion factor, H is the absolute temperature, t is the exposure time, O is the oxygen concentration, and a , b , and c are dimensionless empirical parameters. The data used for this illustration are defined in Table C-1.

Based on Eq. 15 and Table C-1, a standard Monte Carlo solution was generated. This solution, using 100,000 simulations to ensure "accurate" solutions at the tail of the distribution, will be used as a comparison basis.

Table C-1. Data Input for Demonstration Example

Variable	Mean	Standard Deviation	Distribution
K_p (mm/yr ⁿ)	4.0	1.0	Lognormal
Cl (μg/g;ppm)	6.5	0.65	Normal
d	0.47	0.0329	Lognormal
$H(K)$	373	0	
K (mm/yr ⁿ)	0.1706	0	
O (μg/g;ppm)	7	0	
a	-1402	0	
b	0.2	0	
c	0.543	0	
t (yr)	300	0	

By numerical differentiation at the mean values of the random variables, the following first-order approximation can be obtained:

$$C_1 = A_0 + A_1K_p + A_2Cl + A_3n, \quad (16)$$

where A_i are constants.

The solutions based on the MV and the AMV methods are shown in Fig. C-4, where the cdf is plotted on a normal probability paper. The MV solution is obtained by applying a fast probability integration method (Wu and Wirsching, 1987) to Eq. 16. The AMV solution is obtained using Eq. 9.

By comparing with the Monte Carlo solution, it can be concluded that the AMV solution is good in the cdf range from approximately 0.0003 to 0.99. There is one selected point, at cdf \approx 0.99976, which shows relatively large error. This discrepancy exists because C is non-linear and the most probable point is far away from the mean-values point. To improve the accuracy, iterations can be performed to converge to the exact most probable point. For illustration purposes, only the first iteration has been carried out. The results are plotted in Fig. C-4.

To demonstrate the proposed importance sampling method, the approximate function C_1 will be used as the base random variable. The AMV solution approximates C by C^* . Fig. C-5 shows the random relationship between C and C_1 using 600 random samples. The random error $\varepsilon(X)$ is

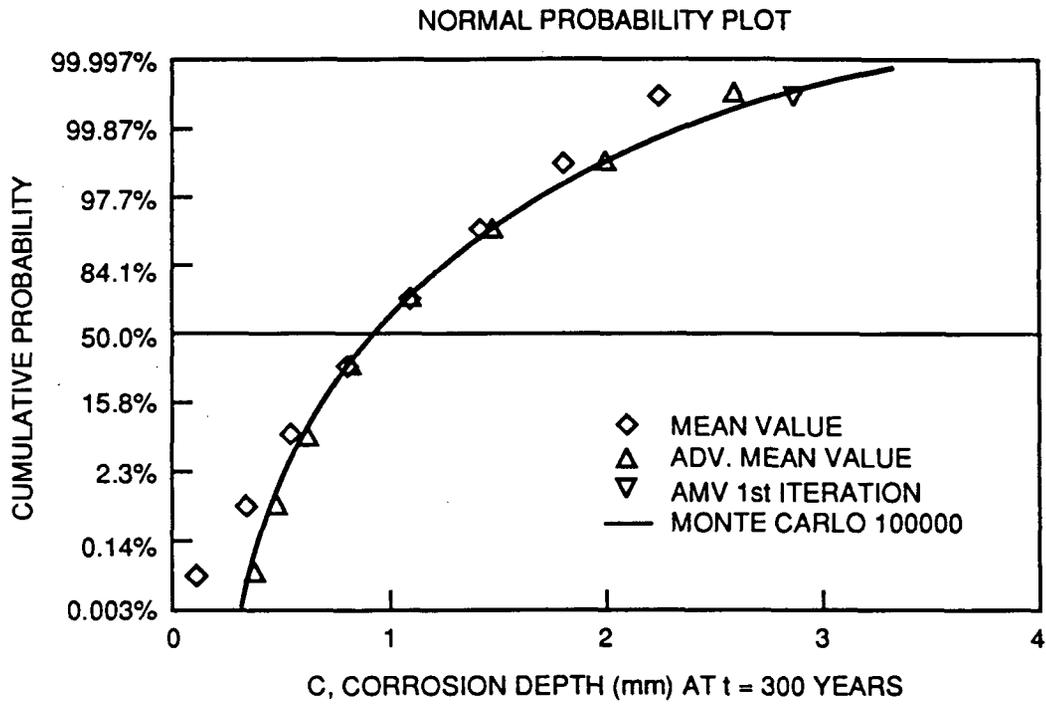


Figure C-4. Corrosion depth cdf analysis using the advanced mean value method.
(Number of performance function evaluations: 12)

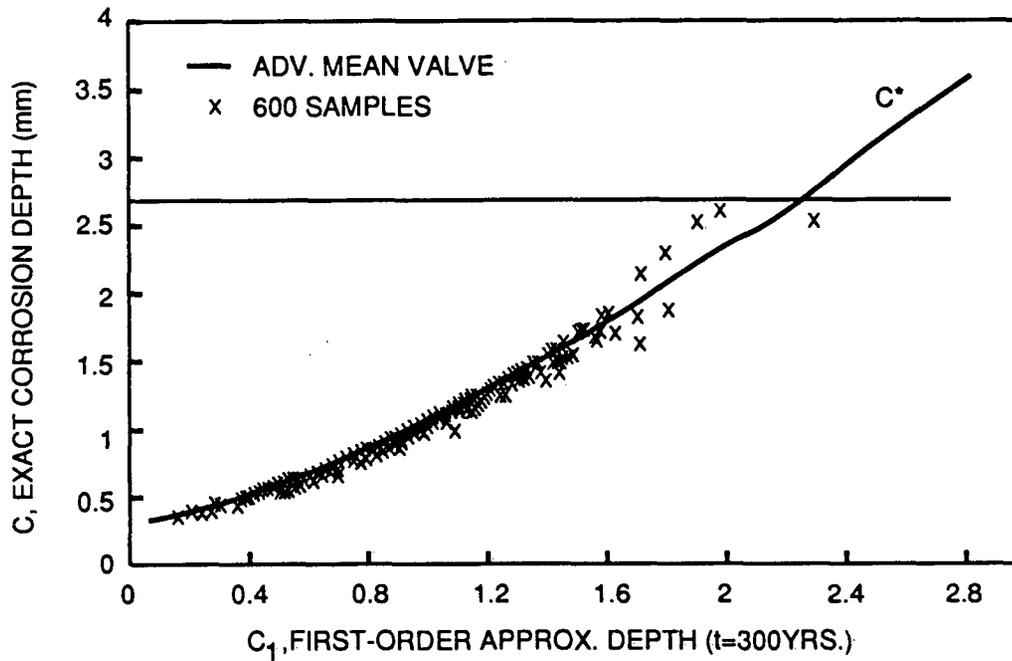


Figure C-5. Cdf analysis using the standard Monte Carlo method.
(Number of performance function evaluations: 600)

higher for larger C_1 . The curve in Fig. C-5 is the AMV solution. It can be seen that C^* and C_1 have a nonlinear relationship, because C is nonlinear in \underline{X} , and C_1 is linear in \underline{X} . Thus, the AMV solution provides, roughly, the "best-fit" curve and explains why the AMV approximation produces reasonable probability results.

To illustrate the sampling procedure, a small probability problem representing a worst case was selected. Assume that the probability of $[C > 2.665]$ is desired. The AMV solution is 2.33×10^{-4} and a standard Monte Carlo solution with 500,000 samples produces 4.41×10^{-4} . Based on Fig. C-4, it appears that the AMV solution is not sufficiently accurate at this probability level.

To improve the AMV solution, the sampling procedure described earlier was applied. The first task is the selection of the initial minimum distance. The AMV minimum distance is 3.53 and the most probable point is $\underline{\mu}^* = (3.28, 0.333, 1.25)$. Several initial distances for importance sampling were selected to find the optimum δ that minimizes N_{IV} . Results are listed in Table C-2, in which N_E is the number of samples, with C exceeding 2.665, and N_T is the total number of samples. The optimum δ was found to be near 2.8. Fig. C-6 shows the results of 100 standard Monte Carlo samples and 500 importance sampling samples with $\delta = 2.8$. The optimum C_{10} (i.e., Y_{b0} in Fig. C-2) is 2.23, whereas the AMV method predicts 2.305.

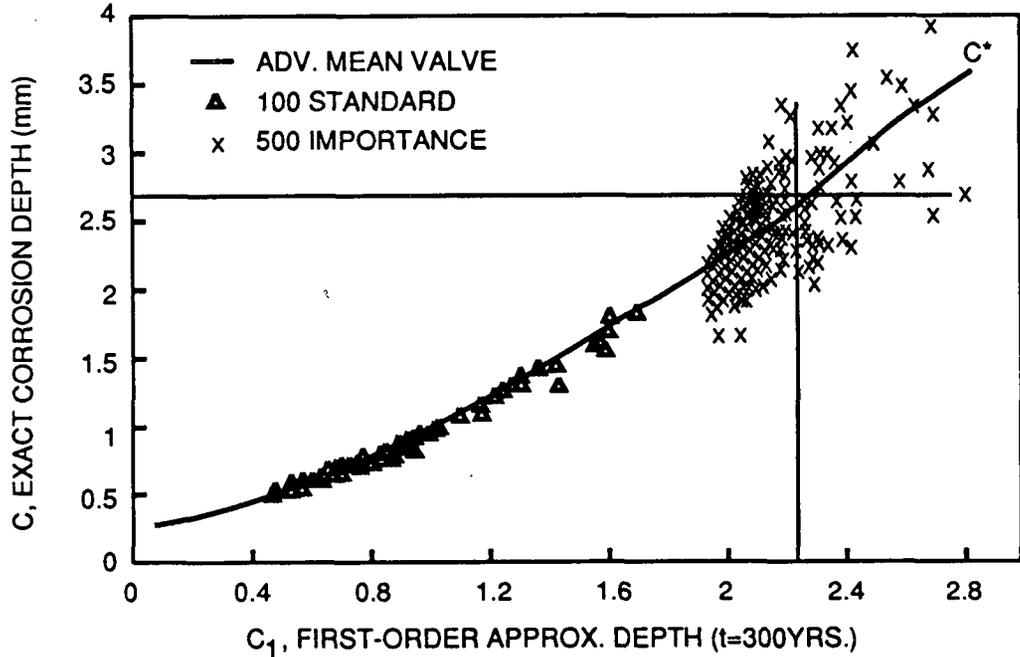


Figure C-6. Cdf analysis using the AMV-based importance sampling method. (Number of performance function evaluations: 600)

The percent error in the probability estimate, with 95 percent confidence, is (Ang and Tang, 1984):

$$\text{percent error} < 200 \sqrt{\frac{1-p}{pN_T}}, \quad (17)$$

where $p = N_E/N_T$.

Assuming a 10 percent error, the required number of samples, N_T , is:

$$N_T = 400 \frac{1-p}{p} \quad (18)$$

For the standard Monte Carlo method, $p \approx 4 \times 10^{-4}$ for $[C > 2.665]$ and $N_T \approx 10^6$. For the importance sampling method, $p \approx 0.15$ (see Table C-2) and $N_T \approx 2266$. Therefore, the importance sampling method improves the efficiency by 440 ($10^6/2266$) times. Two importance samplings with 5000 samples each indicated that the "exact" solution is near 4×10^{-4} . With 500 samples, the solution in Table C-2 is off by less than 10 percent.

Table C-2. Importance Sampling Analysis Results

Distance	Samples	C_{10}	$p = N_E/N_T$	$P[C > 2.665]$
3.0	500	2.247	0.262	0.000353
2.9	500	2.226	0.21	0.000391
2.8	500	2.230	0.148	0.000378
2.8	1000		0.166	0.000424
2.8	2000		0.173	0.000443
2.8	5000		0.160	0.000409
2.8	5000		0.156	0.000399
Std. Monte Carlo	500000		0.000441	0.000441

5. SUMMARY

The efficiency and the analytical nature of the fast probability analysis methods such as the advanced mean-based method provide potential applications to complex computation-intensive problems (such as waste-package performance assessment) that are too costly to analyze using the standard Monte Carlo method. The fast probability analysis result can be checked or enhanced using an importance-sampling updating technique, provided that a reasonable approximate performance function can be determined. The proposed procedure is suitable for providing cost-effective and accurate "what if" answers to different assumptions, to assist in uncertainty and risk analyses.

6. REFERENCES

1. Ang, A. H.-S., and W. H. Tang, Probability Concepts in Engineering Planning and Design, Volume II: Decision, Risk, and Reliability, New York: John Wiley & Sons, Inc., 1984.
2. Madsen, H. O., S. Krenk, and N. C. Lind, Methods of Structural Safety, Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1986.
3. Wu, Y.-T., O. H. Burnside, and T. A. Cruse, "Probabilistic Methods for Structural Response Analysis," Computational Probabilistic Methods, AMD-Vol. 93, ASME; Computational Mechanics of Reliability Analysis, W. K. Liu and T. Belytschko (eds.), Elmepress International, Chapter 7, 1989.
4. Wu, Y.-T., H. R. Millwater and T. A. Cruse, "Advanced Probabilistic Structural Analysis Method for Implicit Performance Functions," AIAA Journal, Vol. 28, No. 9, pp. 1663-1669, September 1990.
5. Rosenblatt, M, "Remarks on a Multivariate Transformation," Annals of Math. Stat., 23, No. 3, pp. 470-472, 1952.
6. Tvedt L., "Distribution of Quadratic Forms in Normal Space - Application to Structural Reliability," Journal of Engineering Mechanics, ASCE, 116 (6), pp. 1183-1197, June 1990.
7. Harbitz, A, "An Efficient Sampling Method for Probability of Failure Calculation," Structural Safety, 3, pp. 109-115, 1986.
8. Wu, Y.-T., and P. K. Nair, "Fast Probabilistic Performance Assessment Methodology Evaluation," Center for Nuclear Waste Regulatory Analyses, Southwest Research Institute, San Antonio, CNWRA88-004, 1988.
9. Sutcliffe, W. G., "Uncertainty Analysis: An Illustration From Nuclear Waste Package Development," Nuclear and Chemical Waste Management, 5, pp. 131-140, 1984.
10. Wu, Y.-T., and P. H. Wirsching, "New Algorithm for Structural Reliability Estimation," Journal of Engineering Mechanics, ASCE, 113 (9), pp. 1319-1336, September 1987.

APPENDIX D - THE BOUNDING APPROACH

Principal Author: Lee R. Abramson

APPENDIX D - THE BOUNDING APPROACH

To implement the probability-distribution approach discussed in Appendix B, it is necessary to specify the joint distribution of all input variables. Although there may be sufficient theoretical foundations or experimental data to specify the distribution of many input variables, it is likely that, for some input variables, no such information exists. Although some information about the distribution of such a variable X may exist based on theory, data, or expert judgment, this information may be too vague to justify assigning a distribution to X . However, in such a case, it may be possible to assign bounds to X . This approach is called the bounding approach, and it is used to propagate uncertainty in a subset $\{X_i\}$ of the input variables when not enough is known about $\{X_i\}$ to use the probability-distribution approach.

Denoting an input variable subject to the bounding approach by X , we assume that X lies in some interval $[a, b]$ with assurance $(1 - \gamma)$. Both the uncertainty interval $[a, b]$ and the assurance are assumed to be supplied by expert judgment. The role of the assurance is to express the experts' assessment of the validity of the assertion that X lies in the stated uncertainty interval. As such, it can be interpreted as a lower bound on the probability that $a \leq X \leq b$. For example, $\gamma = .05$ means that the experts believe there is at least a 95 percent chance that X lies in the stated uncertainty interval. This interpretation will be used for the purpose of applying the Bonferroni inequality to calculate the output assurance as a function of the input assurances (see below).

The assurance is interpreted as a lower bound instead of an equality because only vague information is available about X . In practice, γ will usually be small but positive, so as to reduce the length of the uncertainty interval. If $\gamma = 0$ (assurance = 100 percent), a and b would be absolute bounds on X . However, the length of the uncertainty interval might be too large to yield useful results. In such a case, the length of the uncertainty interval can be reduced by allowing γ to be positive. The trade-off between the assurance and the uncertainty interval will be discussed below.

The bounding approach is recommended when one or more variables can only be characterized by an uncertainty interval. The bounding approach uses the only information available about these variables, i.e., their bounds. Accordingly, let X_{m+1}, \dots, X_L be input variables such that the uncertainty interval of X_i is $[a_i, b_i]$ with assurance $1 - \gamma_i$ ($i = m + 1, \dots, L$). Assume that the input variables X_1, \dots, X_m can be treated by the probability-distribution approach. Using the notation of Appendix B (Fig. B-1), the output variable Y_j can be written as $Y_j = \phi_j(X_1, \dots, X_m, X_{m+1}, \dots, X_L)$, where ϕ_j is the j^{th} component of the global transfer function ϕ , $j = 1, \dots, K$. We write:

$$Y_j = \phi_j(\underline{X}', \underline{X}''), \quad (1)$$

where $\underline{X}' = \{X_1, \dots, X_m\}$ and $\underline{X}'' = \{X_{m+1}, \dots, X_L\}$. If \underline{X}'' had a multivariate distribution, we could write the distribution of the K output variables $\underline{Y} = \{Y_1, \dots, Y_K\}$ as

$$\begin{aligned}
G(\underline{y}) &= G(y_1, \dots, y_k) = \text{Prob}\{Y_1 \leq y_1, \dots, Y_k \leq y_k\} \\
&= \int_{S^*} \text{Prob}\{Y_1 \leq y_1, \dots, Y_k \leq y_k \mid \underline{X}'' = \underline{u}''\} dF(\underline{u}''),
\end{aligned} \tag{2}$$

where $\underline{u}'' = \{u_{m+1}, \dots, u_L\}$ is any fixed set of values of \underline{X}'' , S^* is the domain of \underline{X}'' , and $F(\underline{u}'') = \text{Prob}\{X_{m+1} \leq u_{m+1}, \dots, X_L \leq u_L\}$. Since $\int_{S^*} dF(\underline{u}'') = 1$, bounds on $G(\underline{y})$ can be calculated by bounding the integrand in Eq. 2. This yields:

$$G_1(\underline{y}) \leq G(\underline{y}) \leq G_2(\underline{y}), \tag{3}$$

where

$$\begin{aligned}
G_1(\underline{y}) &= \min_{S^*} \text{Prob}\{Y_1 \leq y_1, \dots, Y_k \leq y_k \mid \underline{X}'' = \underline{u}''\} \\
G_2(\underline{y}) &= \max_{S^*} \text{Prob}\{Y_1 \leq y_1, \dots, Y_k \leq y_k \mid \underline{X}'' = \underline{u}''\}.
\end{aligned}$$

The above bounds hold in general and might be useful even if the probability-distribution approach applied to \underline{X}'' .

In the bounding approach, the only information available about \underline{X}'' is the joint uncertainty interval

$$S^* = \{a_i \leq u_i \leq b_i, i = m + 1, \dots, L\} \tag{4}$$

and the associated assurances. If the domain of \underline{X}'' were S^* with certainty, then Eq. 3 would hold with S^* given by Eq. 4. However, since \underline{X}'' lies in S^* with less than 100 percent assurance, Eq. 3 holds with less than 100 percent assurance. To calculate the assurance of Eq. 3, we interpret the assurance for each component of \underline{X}'' as a lower bound for the probability that it lies in its associated uncertainty interval. We then calculate a lower bound for $\text{Prob}\{S^*\}$ and interpret this lower bound as the assurance of Eq. 3.

It can be shown (see proof on page D-6) that $\text{Prob}\{S^*\} \geq 1 - \gamma^*$, where

$$\gamma^* = \gamma_{m+1} + \dots + \gamma_L. \tag{5}$$

From this result, it can be seen that Eq. 3 holds with assurance $(1 - \gamma^*)$.

From Eq. 4, S^* is an $(L - m)$ -dimensional rectangular parallelepiped. In other words, the minimum and maximum of $G(\underline{y})$ in Eq. 3 are calculated by allowing the components X_i of \underline{X}'' to vary independently over their uncertainty intervals. However, some of the components of \underline{X}'' may have a dependency structure not captured by Eq. 4. For example, it may be known that $X_{m+1} \geq X_{m+2}$.

In such a case, the minimum and maximum of $G(\underline{y})$ in Eq. 3 should be calculated by constraining S^* to take the assumed dependency structure of \underline{X}'' into account. This will have the effect of possibly bringing the bounds on $G(\underline{y})$ closer together. Since the Bonferroni inequality is unaffected by any dependency structure, Eq. 5 will still hold.

Eq. 5 relates the output assurance to the input assurances. Although γ^* should be small (e.g., $\gamma^* = .05$) in order that the bounding approach be useful, the input assurances can vary without changing γ^* , provided only that they satisfy the constraint of Eq. 5. This freedom can be used to adjust the uncertainty intervals so that the bounds for $G(\underline{y})$ are as close as possible. For example, in many applications, only a small number of input variables are critical variables, i.e., significantly affect the output. If these critical variables can be identified with a sensitivity analysis, then the uncertainty intervals for the non-critical input variables can be chosen with high assurance (even 100 percent assurance) in the knowledge that they will not significantly affect the bounds for $G(\underline{y})$. For a fixed γ^* , the uncertainty intervals for the critical input variables can then be adjusted to minimize the distance between $G_1(\underline{y})$ and $G_2(\underline{y})$.

It would appear that the bounding approach has increased the uncertainty by replacing $G(\underline{y})$ with the interval $[G_1(\underline{y}), G_2(\underline{y})]$. However, it must be emphasized that the result is a direct consequence of the uncertainty in the input variables X_{m+1}, \dots, X_L , and hence is unavoidable. For given input uncertainties, uncertainty in the output $G(\underline{y})$ can be reduced only by decreasing the assurance $(1 - \gamma^*)$.

The main result that Eq. 3 holds with assurance $(1 - \gamma^*)$ was derived by first assuming that \underline{X}'' lies in S^* with probability one and then calculating the probability that this assumption is correct. An alternative approach is to assume that $Prob\{\underline{X}'' \in S^*\} = 1 - \gamma^*$ and derive bounds on $G(\underline{y})$. Eq. 2 can be written as

$$G(\underline{y}) = \int_S H(\underline{y}, \underline{u}'') dF(\underline{u}''), \quad (6)$$

where S is the domain of \underline{X}'' and $H(\underline{y}, \underline{u}'') = Prob\{Y_1 \leq y_1, \dots, Y_k \leq y_k \mid \underline{X}'' = \underline{u}''\}$. Now write $S = S^* \cup \bar{S}^*$, where S^* is given by Eq. 4 and \bar{S}^* is the complement of S^* . Then Eq. 6 can be written as

$$G(\underline{y}) = \int_{S^*} H(\underline{y}, \underline{u}'') dF(\underline{u}'') + \int_{\bar{S}^*} H(\underline{y}, \underline{u}'') dF(\underline{u}''). \quad (7)$$

To get bounds on $G(\underline{y})$, note that $Prob\{S^*\} = 1 - \gamma^*$, $Prob\{\bar{S}^*\} = \gamma^*$ and $0 \leq H(\underline{y}, \underline{u}'') \leq 1$ for $\underline{u}'' \in \bar{S}^*$. (The bounds $G_1(\underline{y}) \leq H(\underline{y}, \underline{u}'') \leq G_2(\underline{y})$ hold only when $\underline{u}'' \in S^*$.) It follows from Eq. 7 that

$$G(\underline{y}) \geq \int_{S^*} H(\underline{y}, \underline{u}''') dF(\underline{u}''') \geq (1 - \gamma^*) G_1(\underline{y})$$

and

$$\begin{aligned} G(\underline{y}) &\leq \int_{S^*} H(\underline{y}, \underline{u}''') dF(\underline{u}''') + \int_{S^*} dF(\underline{u}''') \\ &\leq (1 - \gamma^*) G_2(\underline{y}) + \gamma^* \end{aligned}$$

or

$$(1 - \gamma^*) G_1(\underline{y}) \leq G(\underline{y}) \leq G_2(\underline{y}) + \gamma^* [1 - G_2(\underline{y})]. \quad (8)$$

These bounds on $G(\underline{y})$ follow from the assumption that $Prob\{S^*\} = 1 - \gamma^*$ and hold for any distribution function $F(\underline{u}''')$. A comparison of Eq. 3 with Eq. 8 shows that the bounds in Eq. 8 are wider than those in Eq. 3, with equality if and only if $\gamma^* = 0$. The increased spread in the bounds is a reflection of the fact that Eq. 8 is an absolute statement, whereas Eq. 3 holds only with assurance $(1 - \gamma^*)$. Either Eq. 3 or Eq. 8 can be used to bound $G(\underline{y})$.

In practice, the bounds in Eq. 3 may be difficult to calculate. From Eq. 1, the conditional distribution of \underline{Y} given $\underline{X}'' = \underline{u}''$ depends on \underline{u}'' in two ways: directly, through \underline{u}'' in ϕ and indirectly, through the conditional distribution of \underline{X}' given $\underline{X}'' = \underline{u}''$. Although this double dependence may make it difficult to find the minimum and maximum over S^* , it can also be used to find simpler but wider bounds on $G(\underline{y})$.

From Eq. 1, we can write:

$$\begin{aligned} &Prob\{Y_1 \leq y_1, \dots, Y_k \leq y_k \mid \underline{X}'' = \underline{u}''\} \\ &= Prob\{\underline{Y} \leq \underline{y} \mid \underline{X}'' = \underline{u}''\} \\ &= Prob\{\phi(\underline{X}', \underline{X}'') \leq \underline{y} \mid \underline{X}'' = \underline{u}''\}, \end{aligned} \quad (9)$$

where $\phi = \{\phi_1, \dots, \phi_k\}$ is the global transfer function. For any fixed values \underline{u}'' and \underline{v}'' , define

$$H(\underline{y}; \underline{u}'', \underline{v}'') = Prob\{\phi(\underline{X}', \underline{v}'') \leq \underline{y} \mid \underline{X}'' = \underline{u}''\}. \quad (10)$$

From Eqs. 3, 9 and 10, the bounds on $G(\underline{y})$ can be written:

$$G_1(\underline{y}) = \min_{S^*} H(\underline{y}; \underline{u}'', \underline{u}'') \quad (11a)$$

$$G_2(\underline{y}) = \max_{S^*} H(\underline{y}; \underline{u}'', \underline{u}''). \quad (11b)$$

To get a lower bound for $G_1(\underline{y})$, note that:

$$\begin{aligned} \min_{S^*} H(\underline{y}; \underline{u}''', \underline{u}''') &= \min_{\underline{u}'' = \underline{v}''} H(\underline{y}; \underline{u}''', \underline{v}'') \\ &\geq \min_{\underline{v}''} \min_{\underline{u}''} H(\underline{y}; \underline{u}''', \underline{v}''). \end{aligned} \quad (12)$$

The inequality follows because minimizing $H(\underline{y}; \underline{u}''', \underline{v}'')$ when \underline{u}'' and \underline{v}'' can vary independently over S^* is less restrictive than minimizing $H(\underline{y}; \underline{u}''', \underline{v}'')$ when \underline{u}'' is constrained to be equal to \underline{v}'' . Hence:

$$G_1(\underline{y}) \geq G_1^*(\underline{y}), \quad (13)$$

where

$$G_1^*(\underline{y}) = \min_{\underline{v}''} \min_{\underline{u}''} H(\underline{y}; \underline{u}''', \underline{v}'').$$

Similarly,

$$G_2(\underline{y}) \leq G_2^*(\underline{y}), \quad (14)$$

where

$$G_2^*(\underline{y}) = \max_{\underline{v}''} \max_{\underline{u}''} H(\underline{y}; \underline{u}''', \underline{v}'').$$

Finally, from Eqs. 3, 13, and 14,

$$G_1^*(\underline{y}) \leq G(\underline{y}) \leq G_2^*(\underline{y}). \quad (15)$$

The bounds from Eq. 15, although wider than the bounds from Eq. 3, may be simpler to calculate because \underline{u}'' is not constrained to be equal to \underline{v}'' . If the bounds in Eq. 15 are not much wider than those in Eq. 3, then little is lost and much may be gained by using Eq. 15 instead of Eq. 3. In the special case where \underline{X}' and \underline{X}'' are independent, then $G_1(\underline{y}) = G_1^*(\underline{y})$ and $G_2(\underline{y}) = G_2^*(\underline{y})$ and Eq. 15 reduces to Eq. 3. Thus, Eq. 15 may yield useful bounds on $G(\underline{y})$ when \underline{X}' and \underline{X}'' are "almost" independent.

Proof: $Prob\{S^*\} \geq 1 - \gamma^*$.

Consider any set of n statement S_i such that $Prob\{S_i\} = 1 - p_i, i = 1, \dots, n$. We wish to find a bound for the probability of the joint statement

$$S = \{S_1, \dots, S_k\} = \cap S_i.$$

Using standard results from Boolean algebra, and using \bar{S} to denote the complement of S ,

$$\begin{aligned} Prob\{S\} &= 1 - Prob\{\bar{S}\} \\ &= 1 - Prob\{\overline{\cap S_i}\} \\ &= 1 - Prob\{\cup \bar{S}_i\} \\ &\geq 1 - \sum Prob\{\bar{S}_i\} \quad \text{or} \\ Prob\{S\} &\geq 1 - \sum p_i. \end{aligned}$$

This last inequality is called the Bonferroni inequality. It holds regardless of the dependency structure of the S_i . Now let $S_i = \{a_i \leq X_i \leq b_i\}$ be an uncertainty interval for X_i with assurance $1 - \gamma_i$. Then $Prob\{\bar{S}_i\} \leq \gamma_i$. Substituting into the Bonferroni inequality yields $Prob\{S\} \geq 1 - \sum \gamma_i$. If we set $\gamma = \sum \gamma_i$ then $Prob\{S\} \geq 1 - \gamma$. The desired result follows by setting $S = S^*$ and $\gamma = \gamma^*$.

APPENDIX E - MODELING THE TRANSFER FUNCTION

Principal Author: Andre G. Journal

TABLE OF CONTENTS

	Page
1. INTRODUCTION	E-1
2. MODULARIZATION	E-1
3. RESPONSE SURFACE MODELING	E-3
4. IMPORTANT RESPONSE MODELING	E-5
5. EXTRAPOLATION MODELS	E-6
FIGURES	
E-1. Modeling Sequence for Compliance Assessment	E-2
E-2. Response Surface Modeling	E-4

APPENDIX E - MODELING THE TRANSFER FUNCTIONS

1. INTRODUCTION

A key element of the demonstration of the probabilistic rule proposed in Section 5 is the transfer of input uncertainty into a distribution for the critical output variable, the number of waste packages (WPs) failing in the period $[0, t]$. The simpler the transfer function, the easier the task of transfer of uncertainty. Appendix B (Section 2, "Failure Boundary Determination Using Importance Sampling") discusses important implementation shortcuts if the transfer function can be made multivariate monotonic. This appendix reviews some commonly used simplification principles for modeling transfer functions, pointing out cases where such simplifications may result in non-conservative (for the Substantially Complete Containment (SCC) issue) assessment for the output probability distribution.

Modeling the transfer functions is a highly specific technical task that does not pertain to this report on uncertainty evaluation methods. However, in a simulation approach to uncertainty modeling, the transfer functions must be computed repetitively for a large number of alternative sets of input variables. To reduce computational burden, this appendix reviews and proposes general principles for decoupling and streamlining sequences of transfer functions, yet without altering their ability to generate "extreme-valued responses" important for WP reliability analysis. Note that "extreme-valued responses" means responses beyond the compliance limit, with usually a low probability of occurrence.

There are a few general principles that can be put to use to streamline a transfer function before repetitive forward simulation runs:

- Modularization and decoupling
- Response surface modeling, or regression
- Important response modeling
- Extrapolation models

2. MODULARIZATION

The most efficient way to simplify a transfer function, or more generally a complex multivariate input-dependent process, is to decompose it into a series of simpler sub-processes, each calling for a lesser number of input variables. For example, the global transfer function ϕ can be decomposed into a series of five simpler transfer functions TF1 to TF5, as presented in Fig. E-1. These subprocesses can be either in parallel, as for TF1 and TF2 or, better, in series, as for TF3 and TF4.

Ideally, these subprocesses should call, not only each for lesser number of input variables than the global transfer function, but also for independent sets of input variables. Again from Fig. E-1, the two sets of input variables to TF2 and TF4, respectively, can be seen as approximately independent, whereas this is not the case for the two sets of input variables to TF1 and TF4; the

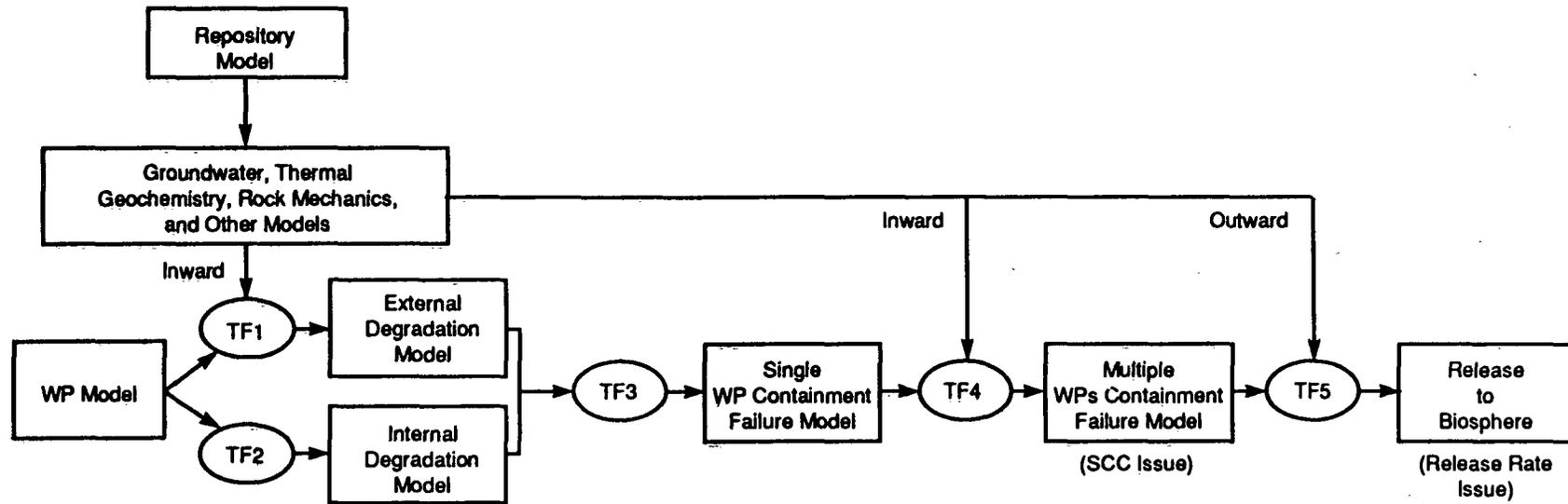


Figure E-1. Modeling sequence for compliance assessment.

latter two share common input variables stemming from a common groundwater/rock-mechanics model. Given any prior single WP failure model, repetitive runs of TF4 can be done independently of the TF2 runs. Similarly, runs of TF3 can be done independently of those of both TF1 and TF2.

The transfer function TF5 conditioning the release rate in Fig. E-1 is most likely quite complex and should itself be made more modular.

3. RESPONSE SURFACE MODELING

Consider any particular subprocess or module $\phi^{(1)}$ that transforms the input set $\underline{X}:\{X_1, \dots, X_L\}$ into the output set $\underline{Y}:\{Y_1, \dots, Y_K\}$, as in Fig. E-2.

Repetitive runs of $\phi^{(1)}$ would yield a correspondence between any N sets of input values $\underline{X}^{(i)}, i = 1, \dots, N$ and N sets of output values $\underline{Y}^{(i)}, i = 1, \dots, N$. From these N pairs of sets, one can determine a regression of any output variable (the dependent variables) on the L (or less) input variables (the independent variables), e.g.,

$$Y_k = \sum_{i=1}^L a_{ki} \cdot T_i(X_i) + \epsilon_k, \text{ for any } k = 1, \dots, K, \quad (1)$$

with $T_i(\cdot)$ being some, possibly non-linear, transform of the input variable X_i .

More generally, instead of considering the regression (1), one could fit some parametric response surface providing a close-form analytical approximation $H(\cdot)$ to those important aspects of the transfer function $\phi^{(1)}$:

$$\underline{Y} = \phi^{(1)}(\underline{X}) \approx H(\underline{X}) + \underline{\epsilon}. \quad (2)$$

Consider, then, the next module $\phi^{(2)}$ that transforms a subset ($K' \leq K$) of the variables Y_k into M new output variables $\underline{Z}:\{Z_1, \dots, Z_M\}$ (see Fig. E-2). One wishes to study the global transfer function $\phi = \phi^{(2)} \cdot \phi^{(1)}$, which transforms \underline{X} into \underline{Z} , i.e.,

$$\underline{Z} = \phi(\underline{X}) = \phi^{(2)}(\phi^{(1)}(\underline{X})).$$

Rather than repetitive, possibly expensive, runs of the global process ϕ , one may consider shortcutting the first module $\phi^{(1)}$ by replacing it by the regression (1) or the response surface (2), e.g.,

$$\underline{Z} \approx \phi^{(2)} \left(\sum_{i=1}^L a_{ki} \cdot T_i(X_i), \quad k = 1, \dots, K' \right). \quad (3)$$

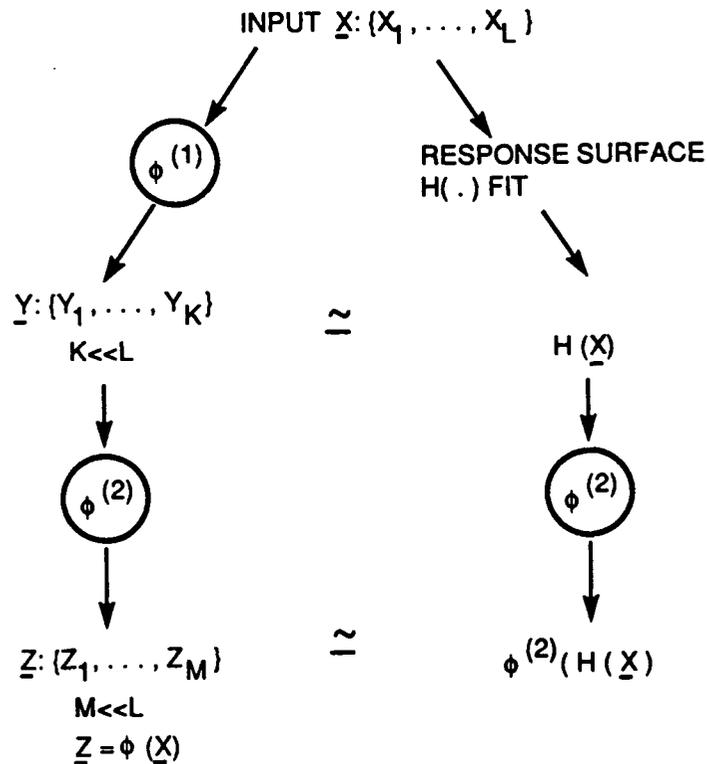


Figure E-2. Response surface modeling.

Note: The first module $\phi^{(1)}$ of the transfer function $\phi = \phi^{(2)} \cdot \phi^{(1)}$ is approximated by a response surface $H(\underline{X})$, with $H(\cdot)$ being some simple close-form analytical expression.
 The second module $\phi^{(2)}$ only considers a limited number of "summary" variables $Y_k, k = 1, \dots, K'$, with $K' \leq K \ll L$.

The approximation (3) is particularly advantageous if the first module $\phi^{(1)}$ is expensive to run and/or if $K' \ll L$, i.e., if the modularization $\phi = \phi^{(2)} \cdot \phi^{(1)}$ has succeeded to cut down the number of input variables from L to K' .

Remarks

- (1) The choice of the regression-type [i.e., the transform $T_i(\cdot)$ or the response surface (2)] is critical for the accuracy of the approximation (3). In repository-type applications, one should be careful in not smoothing out extreme values of the intermediary response

values Y_k , $k = 1, \dots, K'$. If the regression surface (1) or (2) does not reflect any extreme set \underline{Y} , then the outcomes $\underline{Z}^{(i)}$ of the approximation (3) may all be artificially median-valued, even if some of the original input set $\underline{X}^{(i)}$ were extreme.

Any smoothing of the response distribution is usually non-conservative in the sense that it leads to an underestimation of the probability of occurrences of extreme values, a fact not always well-understood, and possibly the major drawback of the regression shortcut for transferring probability distributions (see prior discussion in Appendix A).

(2) The regression (1) can be interpreted as a linearization of the first transfer function $\phi^{(1)}$:

$$\underline{Y} = \phi^{(1)}(\underline{X}) \approx \underline{A} \cdot \underline{T}, \quad (4)$$

with: $\underline{A} = [a_{kl}]$, a $(K \times L)$ matrix

and: $\underline{T} = [T_l(X_l), l = 1, \dots, L]'$, $\underline{Y} = [Y_k, k = 1, \dots, K]'$.

Again, the process of linearization of a transfer function tends to smooth out response distributions resulting possibly in optimistic nonconservative assessments.

4. IMPORTANT RESPONSE MODELING

There are two important aspects to the response surface approximation (1):

- the choice of the particular transform functions $T_l(X_l)$ specific to each input variable X_l ;
- and
- the criterion for determining the model parameters a_{kl} .

Unfortunately, there is no general theory or rule for selecting the transform functions $T_l(\cdot), l = 1, \dots, L$, except for evident parsimony and simplicity.

Sensitivity analysis is used to evaluate how any particular response variable Y_k is dependent on each specific input X_l ; this is done by fixing all other input variables $X_{l'}$ at, usually, their mean or median values. For example, if the variability of $Y_k | X_{l'} = E\{X_{l'}\}$, as a function of $X_l, l \neq l'$, is seen to be correctly fit by the function $T_{l'}(X_{l'})$, then that function $T_{l'}(\cdot)$ is retained for the approximation (1). If that function $T_{l'}(\cdot)$ is not monotonic over the range of $X_{l'}$, then that range can be split into intervals within which monotonic approximations $T_{j_l}(\cdot)$ can be defined:

$$T_l(X_l) = \sum_{j_l=1}^{J_l} T_{j_l}(X_l), \quad (5)$$

with: $T_{j_l}(x) = 0$, for all $x \notin [x_{j_l}, x_{j_l+1}]$.

However, the function $T_l(\cdot)$ may depend on the other input variables $X_{l'}, l' \neq l$; in all rigor, one should denote that function as $T_l(X_l | X_{l'} = x_{l'}, l' \neq l)$, emphasizing that dependence. If one single function $T_l(X_l)$ has to be retained, it is unclear for which set of conditioning values $X_{l'} = x_{l'}, l' \neq l$ it should be determined. The usual choice of mean or median values for the $X_{l'}$ does not necessarily yield a final, overall "best" approximation of type (1); worse, it is most likely not conservative because it does not allow the approximation to reflect a combination of extreme values for different variables $X_l, X_{l'}$. There seems to be a need to appraise what a "good" or "best" approximation of type (1) represents.

As for the criterion for determining the parameters a_{kl} , it is, in the vast majority of applications, least-squares and homoscedastic. That is, the a_{kl} 's are determined to minimize the average square deviations between the model (1) and the observations $\{y_k^{(i)}, k = 1, \dots, K\}$, $i = 1, \dots, N$ independently on the criticality or ranking of the realization (i).

Recommendation:

In the line of importance sampling as developed in Appendix B, it is recommended that good fit of the "marginal" output sets $\{y_k^{(i)}, k = 1, \dots, K\}$ be preferred to good fit of mean or median output sets. By "marginal" it is understood those output sets close to the compliance limit, either barely passing it or failing it.

Consider the case $K = 1$, simpler for notation purposes: marginal output values are those $y^{(i)} = y_{lim} \pm dy$. The fit of a model of type (1), including the choice of the functions $T_l(\cdot)$, should be skewed toward reproducing the proportion of output values greater than y_{lim} . The resulting model would thus be y_{lim} - dependent, and valid only in the vicinity of y_{lim} :

$$Y \approx h(\underline{X}) = \sum_{l=1}^L a_{kl}(y_{lim}) \cdot T_l(X_l; y_{lim}). \quad (6)$$

The approximation (6) must be such that:

$$P\{h(\underline{X}) > y_{lim}\} = P\{Y > y_{lim}\}, \quad (7)$$

in addition to fitting closely all experimental output values $y^{(i)}$.

In the case of an intermediary transfer function for which the notion of marginal output value is not defined, one may replace the value y_{lim} by some high p -quantile value y_p of the output distribution $y^{(i)}, i = 1, \dots, N$. If low values y lead to noncompliance, one would instead retain a low p -quantile value y_p with p close to zero.

5. EXTRAPOLATION MODELS

The two processes of interpolation and extrapolation are distinguished in the sense that the former is performed between two bounds within ranges of dependence, while the latter rests on one single bound and extends beyond any verifiable range of dependence. For example,

- Interpolation is performed within the class $[p_1, p_2]$ defined from two calculated output probabilities $p_1 = P\{Y \leq y_1\}$ and $p_2 = P\{Y \leq y_2\}$, with $y_1 < y_2$, to provide the intra-class distribution $P\{Y \leq y\}, y \in p_1, p_2$.
- Extrapolation is called for the last class $[p_j, 1]$ beyond the last calculated output probability, with y_j equal to or larger than the compliance threshold y_{lim} .

Both interpolation and extrapolation call for models. However, the interpolation model is, or at least can be, data-based, whereas the extrapolation model cannot be, by definition.

The golden rule with extrapolation is to consider many alternative extrapolation models.

A corollary rule is to go as far as possible with data-based or data-charged models, and only then revert to extrapolation-type modeling, i.e., using a model that cannot be refuted by accessible data.

Multivariate normal - related models, including the multivariate lognormal distribution, are examples of models usually called much too early before data-based information has been exhausted. (Refer to previous discussion on multivariate normal and maximum entropy in Appendix A.)

Acceleration principle: Sometimes one can avoid extrapolation models by playing between two variables with similar consequences on the output. This principle is often used to mimic an unfeasible long-duration experiment by increasing (accelerating) another variable. For example, if output data $y(r, t_2)$ are unavailable for very large values of the input time period t_2 , one may consider the substitute data obtained by accelerating, say, the rate of corrosion r and observe the output for smaller time period t_1 , i.e., consider the approximation:

$$y(r_1, t_2) \approx y(r_2, t_1), \text{ with: } r_2 > r_1; \quad t_2 \gg t_1.$$

If based on sound physical grounds and/or experiments relating the y -consequences of the input variables r and t , the acceleration principle allows some data control on extrapolation models.

**APPENDIX F - BAYESIAN GEOSTATISTICS
(A FORMALISM FOR UPDATING PRIOR SOFT INFORMATION)**

Author: A. G. Journel

TABLE OF CONTENTS

	Page
1. INTRODUCTION	F-1
2. BASIC CONCEPTS OF NON-PARAMETRIC GEOSTATISTICS	F-2
2.1 Bayesian Updating	F-2
2.2 Indicator Coding of Information	F-3
2.3 The Projection Theorem	F-4
3. CODING SOFT DATA	F-8
4. PROCESSING SOFT DATA	F-12
5. COKRIGING HARD AND SOFT DATA	F-16
6. REFERENCES	F-23
7. BIBLIOGRAPHY	F-24
FIGURES	
F-1. The Projection Theorem	F-5
F-2. Indicator Coding of Prior Information	F-10
F-3. Calibration of the Information $Y(x)$	F-14
F-4. Covariance Relation Between Soft and Hard Data	F-22

APPENDIX F - BAYESIAN GEOSTATISTICS (A FORMALISM FOR UPDATING PRIOR SOFT INFORMATION)

1. INTRODUCTION

In Appendix B, a combination of stratified sampling for process parameters and conditional simulation sampling for spatial variables has been proposed to characterize the input (multivariate) distributions. Each conditional simulation can be viewed as a random sample of the input space of all spatially variable parameters, such as porosity and permeability. The dimension of that input space can be very large, from 10^3 , if block average (effective) properties are considered, to 10^6 , if core-volume properties are considered. It is therefore essential to reduce as much as possible the spread of that multivariate input distribution by conditioning on as many data as possible, including all sources of information, possibly of varying quality.

The concept and algorithms for conditional simulations are well described elsewhere (Ripley, 1987; Journel, 1989 - lesson IV). The principle of conditional simulation consists of drawing possible outcomes for any yet unsampled (or unknown) input value from a probability distribution model made conditional on the available information. The more information, the smaller the spread (e.g., variance) of that distribution, thus the lesser the uncertainty. This appendix proposes a formalism for generating such probability distribution models conditional on data of different types.

This appendix contains advanced research results that have not yet undergone the test of numerous implementations. However, the author believes these results to be relevant to the Substantially Complete Containment (SCC) issue, which requires the use of soft data.¹

Not all data are equal; they differ by their prior measure of uncertainty and their degree of dependence on the particular unknown being considered. The general Bayesian formalism is ideally suited to update (reduce) uncertainty about any variable by accounting for the correlation between that variable and any relevant piece of information (the data). In the following, a geostatistical twist is given to that general Bayesian formalism, in the sense that all data are assumed to be spatially distributed and statistically inter-dependent, in addition to being related to the unknown. The formalism proposed amounts to modeling local prior distributions as random variables with their specific reliabilities and patterns of spatial dependence. Bayesian updating is then performed yielding posterior (conditional) probability distribution models for the unsampled (or uncertain) values.

This appendix consists of four sections providing:

- 1: a review of the basic principles of non-parametric geostatistics
- 2: a formalism for a common coding of hard and soft data
- 3-4: a formalism for jointly processing hard and soft data

¹ As opposed to hard data, "soft" data correspond to local information that needs updating.

2. BASIC CONCEPTS OF NON-PARAMETRIC GEOSTATISTICS

Non-parametric geostatistics differs from more traditional multinormal-based geostatistics in the sense that it endeavors to characterize uncertainty by direct inference of conditional (posterior) probability values. A parametric approach would consist of the indirect inference of such conditional distributions through their parameters (usually mean and variance), once a distributional type has been assumed.

Non-parametric geostatistics rests on three basic yet very powerful concepts:

- Bayesian updating
- indicator coding of information
- projection theorem

Elements of these three concepts are briefly summarized with the focus being toward the determination of conditional (posterior) probability distributions as measures of uncertainty.

2.1 Bayesian Updating (Berger, 1980)

Consider any two random events A_1 and A_2 , in which outcomes can be either discrete or continuous. (In the following discussion, we will be using discrete notations for convenience.)

The two random variables (RV) are characterized by their joint distribution:

$$P_{12}(a_1, a_2) = P\{A_1 = a_1, A_2 = a_2\}.$$

The conditional or posterior distribution of the event A_1 (the unknown), given information on A_2 (the data), is given by Bayes' relation:

$$P_{1|2}(a_1 | a_2) = P\{A_1 = a_1 | A_2 = a_2\} = \frac{P_{12}(a_1, a_2)}{P_2(a_2)}, \quad (1)$$

with $P_2(a_2) = P\{A_2 = a_2\}$ being the marginal probability of event A_2 .

Bayes' relation can be rewritten as:

$$P_{1|2}(a_1 | a_2) = \frac{P_{2|1}(a_2 | a_1)}{P_2(a_2)} \cdot P_1(a_1) \quad (2)$$

with $P_{2|1}(a_2 | a_1) = P\{A_2 = a_2 | A_1 = a_1\}$ being the conditional probability (likelihood) of observing the data $A_2 = a_2$ if the unknown value $A_1 = a_1$; this probability is called the likelihood function, a function of both a_2 and a_1 .

Relation (2) appears as an updating of the prior (marginal) probability $P_1(a_1) = P\{A_1 = a_1\}$ by the data-value a_2 dependent factor $P_{2|1}(a_2 | a_1)/P_2(a_2)$. This updating calls for prior knowledge of the likelihood function $P_{2|1}(a_2 | a_1)$, i.e., de facto for knowledge of the joint distribution $P_{12}(a_1, a_2)$ since:

$$P_{12}(a_1, a_2) = P_{2|1}(a_2 | a_1) \cdot P_1(a_1).$$

Notwithstanding the success of the formulation (2), this discussion prefers the initial formulation (1), which states unambiguously that Bayesian updating requires knowledge of the joint distribution $P_{12}(a_1, a_2)$, i.e., a prior model of dependence between the unknown A_1 and the data A_2 . That prior model should stem from actual prior information, such as a calibration data set $\{a_1^{(i)}, a_2^{(i)}\}, i = 1, \dots, N$, from which the conditional distributions $P_{2|1}(a_2 | a_1)$ can be inferred for all values a_1 . Indicator covariances are shown hereafter to be exactly such joint distribution models inferred from data.

Unfortunately, often the likelihood function $P_{2|1}(a_2 | a_1)$ is arbitrarily taken from some congenial multinormal-related distribution with no data justification; then, belatedly, Bayesian respectability is gained by using the formulation (2). Bayesian updating is a powerful concept for propagation of information, but cannot make up for the poor quality or arbitrariness of the likelihood functions being used.

2.2 Indicator Coding of Information (Journel, 1989 - Lessons 4-5; Journel and Alabert, 1989)

The probability of any event can be expressed as the expectation of a corresponding indicator (binary) random variable. Define the indicator attached to the random event A as:

$$I_A(a) = \begin{cases} 1, & \text{if } A = a \\ 0, & \text{if not} \end{cases} \quad (3)$$

Then $P\{A = a\} = E\{I_A(a)\}$.

Similarly, if the event A_1 (the unknown) is made conditional on another event A_2 (the data), the corresponding conditional probability can be expressed as the conditional expectation of an indicator RV:

$$P\{A_1 = a_1 | A_2 = a_2\} = E\{I_{A_1}(a_1) | I_{A_2}(a_2) = 1\}. \quad (4)$$

Relation (4) is fundamental, for it allows mapping the problem of estimating conditional (posterior) probabilities into that of estimating conditional expectations, with the projection theorem (point 3 hereafter) providing a general approach to the latter.

Bayes' relation (1) is then rewritten, in all generality:

$$E\{I_{A_1}(a_1) | I_{A_2}(a_2)\} = \frac{E\{I_{A_1}(a_1) \cdot I_{A_2}(a_2)\}}{E\{I_{A_2}(a_2)\}}. \quad (5)$$

It thus appears that updating the prior probability $P_1(a_1) = P\{A_1 = a_1\} = E\{I_{A_1}(a_1)\}$ calls for the indicator covariance $E\{I_{A_1}(a_1) \cdot I_{A_2}(a_2)\} = P\{A_1 = a_1, A_2 = a_2\}$. There is a strict parallel between indicator covariance models and joint probability distributions, between the indicator approach to calculating conditional distributions and Bayesian updating.

2.3 The Projection Theorem (Luenberger, 1969)

Consider any two dependent random events A and B . The conditional expectation $E\{A | B = b\}$ is a function, usually non-linear, of the conditioning value b :

$$E\{A | B = b\} = \psi(b). \quad (6)$$

Randomizing b into the RV B , the conditional expectation (6) $\psi(B)$ appears itself as a random variable. This randomization of b into B is the key conceptual step, for it allows estimation of $\psi(B)$ by projection-type algorithms.

The projection theorem allows defining uniquely the RV $\psi(B)$ as the projection of the RV A onto the Hilbert space H_B of all measurable functions of B (see Figure F-1). That projection is characterized by the orthogonality of the error vector $A - \psi(B)$ to the Hilbert space of projection, leading to the well-known normal equations, belatedly renamed kriging equations:

$$\psi \text{ is s.t. } E\{[A - \psi(B)] \cdot f(B)\} = 0,$$

$$\text{i.e.,} \quad (7)$$

$$E\{\psi(B) \cdot f(B)\} = E\{A \cdot f(B)\}, \text{ for all } f \in H_B.$$

Any data event B can be seen as the intersection of a certain number N , possibly large, of indicator (binary) data $I_{B,\alpha}$, $\alpha = 1, \dots, N$. Thus, the random event B can take up to 2^N outcome values, depending on the state of the N indicator switches $I_{B,\alpha}$. Note that the N indicator RV's $I_{B,\alpha}$ may be dependent, with, for example:

$$E\{I_{B,\alpha} \cdot I_{B,\alpha'}\} = 0, \text{ forbidding the two switches } \alpha \text{ and } \alpha' \text{ to be simultaneously set to one.}$$

The following lemma provides the general expression of any measurable function $f(B)$ of a random event B characterized by N indicator RV's $I_{B,\alpha}$:

Lemma: Any function of n indicator (binary) variables can be written as a linear combination of these indicators taken one by one, then of products of any two, three, ..., up to n indicators:

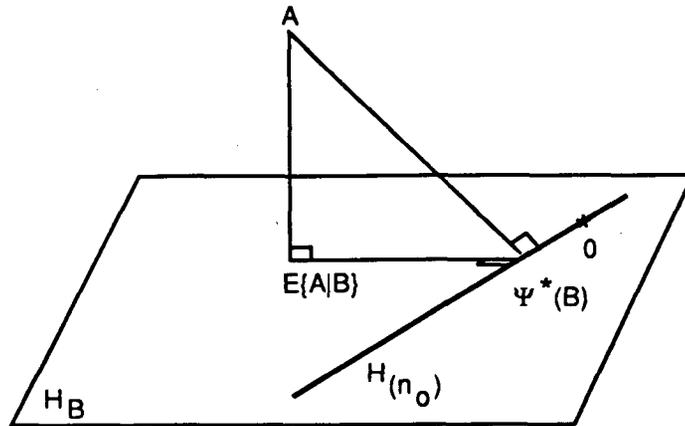


Figure F-1. The projection theorem.

Note: The conditional expectation $E\{A | B\}$ is defined as the unique projection of A onto the Hilbert space generated by all functions of B . That conditional expectation can be approximated by projecting A onto a Subspace $H_{(n_0)} \subset H_B$.

$$\begin{aligned}
 f(B) &= f(I_{B,\alpha} \alpha = 1, \dots, n) \\
 &= a_0 \underline{1} + \sum_{\alpha=1}^n \alpha_0^{(1)} I_{B,\alpha} + \sum_{\alpha_1=1}^n \sum_{\alpha_2 > \alpha_1}^n a_{\alpha_1 \alpha_2}^{(2)} I_{B,\alpha_1} \cdot I_{B,\alpha_2} \\
 &\dots\dots\dots + a^{(n)} \cdot \pi_{\alpha=1}^n I_{B,\alpha} \quad ,
 \end{aligned} \tag{8}$$

with (by convention) the RV $\underline{1}$ almost certainly equal to the unit scalar. There are $1 + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n} = 2^n$ coefficients $a^{(i)}$.

Note that any product of indicators is an indicator (binary) variable; thus, expression (8) is a linear combination of 2^n indicator variables.

The Hilbert space H_B of Fig. F-1 is thus characterized by the basis of the 2^n vectors of expression (8). The conditional expectation $\psi(B) = E\{A | B\}$ appears as a linear combination of 2^n binary events, and is then characterized by 2^n normal equations of type (7):

$$\begin{aligned} \psi(B) = & \lambda_0 \cdot \underline{1} + \sum_{\alpha=1}^n \lambda_{\alpha}^{(1)} I_{B,\alpha} + \sum_{\alpha=1}^n \sum_{\alpha_2 > \alpha_1}^n \lambda_{\alpha_2 \alpha_1}^{(2)} \cdot I_{B,\alpha_1} \cdot I_{B,\alpha_2} \\ & + \dots + \lambda^{(n)} \pi_{\alpha=1}^n I_{B,\alpha} \quad , \end{aligned} \tag{9}$$

with the 2^n coefficients λ determined by the 2^n orthogonality relations:

$$\begin{aligned} E\{\psi(B) \cdot \underline{1}\} &= E\{\psi(B)\} = E\{A\}, \quad \text{ensuring unbiasedness} \\ E\{\psi(B) \cdot I_{B,\alpha}\} &= E\{A \cdot I_{B,\alpha}\}, \quad \alpha = 1, \dots, n \\ E\{\psi(B) \cdot I_{B,\alpha_1} \cdot I_{B,\alpha_2}\} &= E\{A \cdot I_{B,\alpha_1} \cdot I_{B,\alpha_2}\}, \\ & \text{for all pairs } \alpha_2 > \alpha_1, \alpha_1 = 1, \dots, N, \text{ in number } \binom{n}{2} \\ & \dots \dots \dots \\ E\{\psi(B) \cdot \pi_{\alpha=1}^n I_{B,\alpha}\} &= E\{A \cdot \pi_{\alpha=1}^n I_{B,\alpha}\}. \end{aligned} \tag{10}$$

Remarks:

- D. R. Cox (1972) has proposed an expansion for joint probability distributions similar to expression (8).
- The projection theorem coupled with the concept of indicator coding allows expressing any conditional expectation $E\{A | B\}$ as a linear combination of 2^n elementary binary events. The coefficients (λ 's) of this expression are given by a system of linear equations calling for:
 - ~ the covariance matrix between any two elementary binary events
 - ~ the covariance between any elementary binary event and the unknown A
- The full expression (9) is exact inasmuch as Bayes' relation (1) or (5). However, in practice, it may be difficult to infer all $2^n(2^n + 1)/2 + 2^n = 2^n(2^n + 3)$ indicator covariances as required by system (10). One may call for a Markov-type paradigm, and approximate expression (9) by retaining only those $(n_0 + 1)$ elementary events (out of a total of 2^n) deemed most characteristic of event B :

$$\psi^*(B) = [E\{A | B\}]^* = v_0 \cdot \underline{1} + \sum_{\alpha=1}^{n_0} v_{\alpha} \cdot J_{\alpha} \quad , \tag{11}$$

with the n_0 binary events J_{α} taken from the set of 2^n binary events defining B and listed in expressions (8) or (9). The $(n_0 + 1)$ weights v_0, v_{α} of expression (11) are determined by a correspondingly reduced set of normal equations expressing the projection of A onto the vector

subspace $H_{(n_0)} \subset H_B$ generated by the $(n_0 + 1)$ elementary events selected (see Fig. F-1). Note that the projection of A onto H_{n_0} also defines the closest approximation to the conditional expectation $E\{A | B\}$ to be found in the subspace $H_{(n_0)}$ [refer to the theorem of the three perpendiculars in topology (Journal and Huijbregts, 1978, p. 559)].

For example, if the $n_0 + 1 = n + 1$ events first listed in expression (9) are retained, defining the approximations:

$$\psi(B) = E\{A | B\} \approx v_0 + \sum_{\alpha=1}^n v_{\alpha} I_{B,\alpha} \quad , \quad (12)$$

the corresponding $(n + 1)$ normal equations are:

~ first equation equivalent to an unbiasedness condition:

$$v_0 + \sum_{\alpha=1}^n v_{\alpha} E\{I_{B,\alpha}\} = E\{A\},$$

which allows defining the conditional expectation estimate (12) as:

$$[E\{A | B\} - E\{A\}]^* = \sum_{\alpha=1}^n v_{\alpha} [I_{B,\alpha} - p_{\alpha}], \quad (13)$$

with: $p_{\alpha} = E\{I_{B,\alpha}\}, \alpha = 1, \dots, n$.

~ the remainder n normal (linear) equations are:

$$\sum_{\beta=1}^n v_{\beta} Cov\{I_{B,\beta}, I_{B,\alpha}\} = Cov\{I_{B,\omega}, A\}, \alpha = 2, \dots, n \quad (14)$$

with:

$$Cov\{I_{B,\beta}, I_{B,\alpha}\} = E\{[I_{B,\beta} - p_{\beta}][I_{B,\alpha} - p_{\alpha}]\}$$

$$Cov\{I_{B,\alpha}, A\} = E\{[I_{B,\alpha} - p_{\alpha}][A - E\{A\}]\}.$$

Thus, the approximation (13) to the conditional expectation $E\{A | B\}$ only requires knowledge of:

~ the $(n + 1)$ marginal expected values $E\{A\}, p_{\omega}, \alpha = 1, \dots, N$ to ensure unbiasedness, i.e.,

$$E\{[E\{A | B\}]^*\} = E\{A\}$$

- ~ the $n(n + 1)/2$ indicator covariances $Cov\{I_{B,\beta}, I_{B,\alpha}\}, \alpha \neq \beta = 1, \dots, n$ defining the interdependence between the n elementary data events $I_{B,\alpha}$.
- ~ the n indicator covariances $Cov\{I_{B,\omega}, A\}$ characterizing the dependence between each elementary data event $I_{B,\alpha}$ and the unknown A .
- If the unknown event A is itself binary, e.g., $A = I_{A_1}(a_1)$ as defined by relation (3), then the conditional expectation $E\{A | B\}$ is itself a conditional probability value:

$$E\{I_{A_1}(a_1) | B\} = P\{A_1 \leq a_1 | B\},$$

which exact expression is given by the development (9), or which can be approximated by an expression of type (13).

By varying the threshold value a_1 , the entire conditional (posterior) distribution function of A_1 given any outcome of B can be derived. Note from the linear system (14) that changing the (unknown) event A does not require solving a new system, since A affects only the right-hand side covariance values.

In summary, it appears that the non-parametric determination of conditional probability distributions, through the projection theorem and normal equations, is theoretically straightforward. All problems relate to the inference of the covariances characterizing the dependence between the unknown and the data, as in formula (5).

Such inference is made easier if the usually complex assemblage of data (B in Fig. F-1) is decomposed into a series of elementary binary events, the 2^n switches $I_{B,\omega}, I_{B,\alpha_1}, I_{B,\alpha_2}, \dots$, in expression (9). Then inference of a limited number of indicator covariances, characterizing the pattern of dependence of the most important switches, is enough to provide an approximation to the required conditional probability.

The effect of truncation of the exact expression (9) into the more amenable expression (11) or (12) can be evaluated through a convergence analysis of approximations of type (9) with increasing truncation order number.

3. CODING SOFT DATA

The projection theorem approach discussed in the previous section does not impose any limitation on the type of conditioning information. The data event B , in the notations of relations (6) to (14), is constituted of any number N of elementary binary events $I_{B,\omega}, \alpha = 1, \dots, N$. These N , possibly dependent switches generate a potential of 2^N realizations for B .

This remarkable flexibility of the indicator coding of any data event B allows considering almost any type of conditioning data, yet all can be processed through the same normal equations (10) or (14).

Consider the elementary cases of information taking the format of a hard continuous or discrete variable, a constraint interval, and last, a prior probability distribution. Information relates to the spatial distribution of a particular attribute $z(\underline{x})$, say core porosity measured at location \underline{x} .

Indicator Coding:

Discretize the range of Z by K threshold values $z_k, k = 1, \dots, K$, and define from the original RV $Z(\underline{x})$ the K indicator (binary) RV's as:

$$I(\underline{x}; z_k) = 1, \text{ if } Z(\underline{x}) \leq z_k$$

$$0, \text{ if not } \quad k = 1, \dots, K. \quad (15)$$

- **Hard data:** A hard sample value $z(\underline{x}_1)$ would generate a complete indicator data column $i(\underline{x}_1; z_k)$ of zeros followed by ones (see Fig. F-2). What has been lost in the process of indicator coding is the resolution within a class. However, one could consider K as large as necessary and the class intervals can be arbitrary, for example, of equal amplitude, or equal marginal probability, or clustered in some important range of Z -values.
- **Constraint interval:** At some other location \underline{x}_2 the porosity may not have been sampled, but a well log may indicate the rock type prevailing at \underline{x}_2 from which an interval for the porosity may be specified: $z(\underline{x}_2) \in [a_2, b_2]$ (see Fig. F-2). Such a constraint interval would generate an incomplete indicator data column with:

$$i(\underline{x}_2; z_k) = 0, \text{ for } z_k \leq a_2$$

$$\text{undefined, for } z_k \in [a_2, b_2] \quad (16)$$

$$1, \text{ for } z_k > b_2.$$

- **Prior local distribution:** At some third location \underline{x}_3 , the constraint interval $[a_3, b_3]$ corresponds to a rock type for which a porosity distribution is available, for example, from samples taken elsewhere. One may consider for prior information at \underline{x}_3 that distribution, identifying the indicator data to the corresponding cumulative distribution function (cdf) values (see Fig. F-2):

$$i(\underline{x}_3, z_k) = 0, \text{ for } z_k \leq a_3$$

$$P\{Z \leq z_k \mid \text{rock type prevailing at } \underline{x}_3\} \in [0, 1],$$

$$\text{for } z_k \in [a_3, b_3]$$

(17)

$$1, \text{ for } z_k > b_3$$

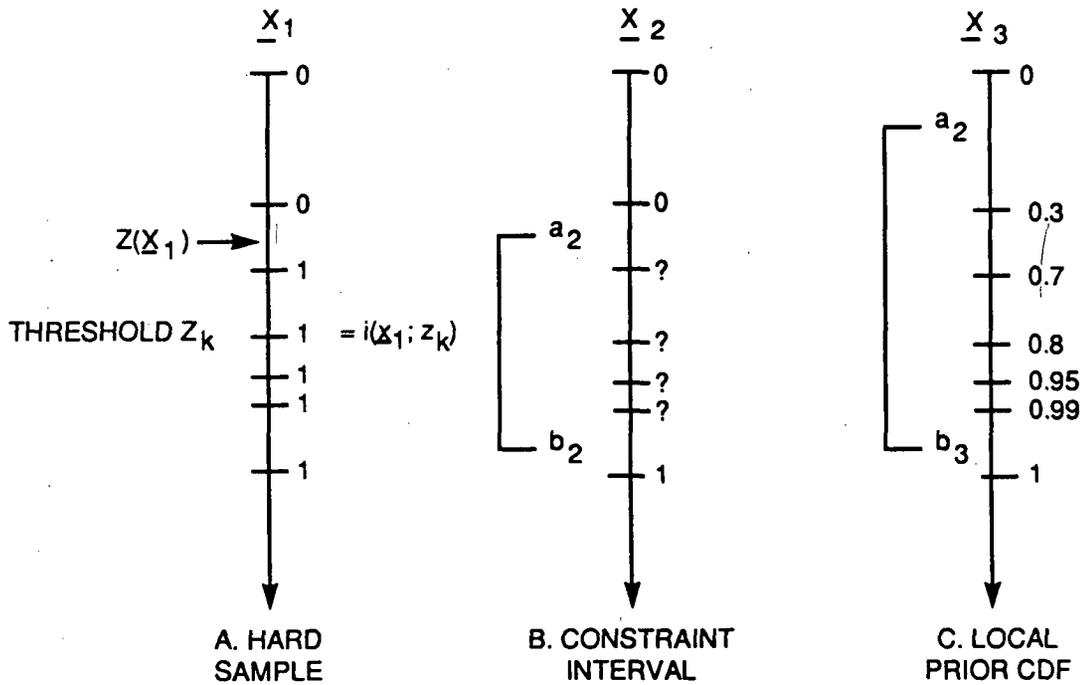


Figure F-2. Indicator coding of prior information.

In this latter case, the indicator data are not any more binary, but are prior probability valued in $[0,1]$. They are said to be "fuzzy" indicator data.

Note that a hard sample of type (15) is but a particular case of prior local distribution with zero variance (no uncertainty), that is, a step cdf:

$$I(\underline{x}_1; z_k) = \Delta(z_k, z(\underline{x}_1)) = 0, \text{ for all } z_k < z(\underline{x}_1)$$

$$1, \text{ for all } z_k \geq z(\underline{x}_1).$$

A constraint interval can also be seen as an incomplete prior probability distribution. Therefore, a general format for input numerical data is a local prior cdf defined as:

$$y(\underline{x}; z_k) = P\{Z(\underline{x}) \leq z_k \mid \underline{x}\} \in [0, 1]. \quad (18)$$

Remarks:

- The notation $y(\cdot; \cdot)$ with a lower case letter y recalls that this information is a deterministic scalar, as opposed to a random variable. In the next section, that information will be made random to allow accounting for its own uncertainty. The notation $\mid \underline{x}$ differentiates this information, specific to location \underline{x} , from the prior histogram of $Z(\underline{x})$ obtained, say, by averaging all hard data $z(\underline{x}_n)$ available at different locations. To avoid confusion, the local prior cdf (18) is sometimes called a pre-posterior distribution. The overall histogram of $Z(\underline{x})$ can be seen as prior information available at all locations \underline{x} , whereas pre-posterior information of type (18) is available only at some selected locations.
- Prior uncertainty at location \underline{x} is characterized by the spread of the local prior Cumulative Distribution Function (cdf) (18). It is maximal if only the overall prior histogram is available at \underline{x} ; it is minimal (zero variance) if (18) reduces to a step cdf corresponding to a hard sample value.

At this point indicator coding, including fuzzy indicator valued between [0,1], has allowed defining a continuum between no local data (maximum prior uncertainty) and hard data (no uncertainty).

Bayesian updating consists of reducing the uncertainty associated to any local prior cdf of type (18) by capitalizing on dependent information available at neighboring locations. The result will be a new distribution at each location \underline{x} , called the posterior distribution, reflecting a reduced uncertainty.

Composite information:

Information available at a given location \underline{x} can include several pre-posterior, or local prior, distributions related to different attributes. For example, in addition to a local porosity permeability distribution, there could be available a local transmissivity distribution, and/or a distribution of rock types, if the well log analysis could not determine with certainty which rock type prevails at \underline{x} :

- ~ for porosity $\phi: y_1(\underline{x}; \phi_k) = P\{\Phi(\underline{x}) \leq \phi_k \mid \underline{x}\}$
- ~ for transmissivity $T: y_2(\underline{x}; t_k) = P\{T(\underline{x}) \leq t_k \mid \underline{x}\}$
- ~ for rock type (assuming three rock types $r = 1, 2, 3$):

$$y_3(\underline{x};r_1) = P\{\underline{x} \in r_1 | \underline{x}\} \quad (19)$$

$$y_3(\underline{x};r_2) = P\{\underline{x} \in r_2 | \underline{x}\}$$

$$y_3(\underline{x};r_3) = P\{\underline{x} \in r_3 | \underline{x}\} = 1 - y_3(\underline{x};r_1) - y_3(\underline{x};r_2).$$

All these data at \underline{x} may be interdependent and, provided this dependence has been modeled, can be used to update each other: for example, the local prior rock type distribution can (must) be used to update both local prior transmissivity and porosity distributions.

In addition, thanks to spatial cross-correlation, the rock type information $y_3(\underline{x};r)$ at location \underline{x} can be used to update the transmissivity distribution at a neighboring location \underline{x}' where no pre-posterior rock distribution is available.

4. PROCESSING SOFT DATA

Although the local prior distribution $y(\underline{x};z_k)$ as defined by relation (18) already carries a measure of uncertainty of that information, that measure may be itself uncertain.

To allow for a modeling of that second level uncertainty, the idea is to randomize $y(\underline{x};z_k)$ itself into a random function (RF) $Y(\underline{x};z_k)$. Of course that RF is likely to present spatial auto-correlation, since local prior information at two neighboring locations \underline{x} , \underline{x}' are likely to be dependent. Also, and hopefully, $Y(\underline{x};z_k)$ is related to the binary indicator RF $I(\underline{x};z_k)$ as defined in (15) and whose conditional expectation is precisely the posterior cdf:

$$E\{I(\underline{x};z) | (n)\} = P\{Z(x) \leq z | n\} = F(x;z | (n)) \quad (20)$$

with the notation (n) representing all information available at all locations, including, but not limited to location \underline{x} .

The projection theorem indicates that an estimate of the conditional cdf (20) is obtained by projecting the (unknown) RV $I(\underline{x};z)$ onto the linear vector space generated by all data $Y(\underline{x}_\alpha; z_k)$, $\alpha = 1, \dots, n$ available at all neighboring locations \underline{x}_α including \underline{x} itself. That estimator is written, similarly to expression (13):

$$[F(x;z | (N)) - F(z)]^* = \sum_{\alpha=1}^n v_\alpha \cdot [Y(\underline{x}_\alpha; z) - E\{Y(\underline{x}_\alpha; z)\}] \quad (21)$$

with $F(z) = P\{Z(x) \leq z\}$ being the prior cdf (not a local prior!) of the stationary RF $Z(x)$ inferred, say, from the histogram of all z -data available over the stationary field.

Determination of the estimator (21) requires knowledge of:

~ the two mean values $F(z)$ and $E\{Y(\underline{x};z)\}$

- ~ the $(n \times n)$ covariance matrix, $[Cov\{Y(\underline{x}_\alpha; z), Y(\underline{x}_\beta; z)\}]$, $\alpha, \beta = 1, \dots, n$, characterizing the interdependence (redundancy) between the n data $Y(\underline{x}_\alpha; z)$
- ~ the n covariances $Cov\{Y(\underline{x}_\alpha; z), I(\underline{x}; z)\}$, $\alpha = 1, \dots, n$, characterizing the relation of each datum $Y(\underline{x}_\alpha; z)$ with the unknown $I(\underline{x}; z)$.

Covariance determination:

In the following discussion, to lighten notations the parameter-threshold value z will be dropped:

$Y(\underline{x}; z)$ is short denoted as $Y(\underline{x})$

$I(\underline{x}; z)$ is short denoted as $I(\underline{x})$.

However, it will be remembered that all results are threshold z -dependent.

The following constitutive hypothesis is key to the proposed covariance determination:

"Hard information $I(\underline{x})$ always prevails over any soft collocated information $Y(\underline{x})$, that is:

$$\begin{aligned}
 E\{I(\underline{x}') | I(\underline{x})=1, Y(\underline{x})=y\} &= E\{I(\underline{x}') | I(\underline{x})=1\} \\
 E\{I(\underline{x}') | I(\underline{x})=0, Y(\underline{x})=y\} &= E\{I(\underline{x}') | I(\underline{x})=0\} \\
 &\text{for all } y \in [0, 1], \text{ and all } \underline{x}', \underline{x}.
 \end{aligned} \tag{22}$$

The tedious but otherwise not particularly difficult calculations are given hereafter. The results are strikingly simple.

Let:

$$\begin{aligned}
 C_I(h) &= Cov\{I(\underline{x} + \underline{h}), I(\underline{x})\} = E\{I(\underline{x} + \underline{h}) \cdot I(\underline{x})\} - F^2, \\
 &\text{with } F = F(z) = E\{I(\underline{x}; z)\},
 \end{aligned}$$

be the covariance of the indicator RF $I(\underline{x})$, as inferred from all indicator data $i(\underline{x}_\alpha; z)$ available at all sample locations throughout the stationary field of $Z(\underline{x})$. Then:

$$\begin{aligned}
 Cov\{I(\underline{x} + \underline{h}), Y(\underline{x})\} &= B \cdot C_I(h) \\
 Cov\{Y(\underline{x} + \underline{h}), Y(\underline{x})\} &= B^2 \cdot C_I(h) \\
 E\{Y(\underline{x})\} &= F \cdot m^{(1)} + (1 - F)m^{(0)} = BF + m^{(0)} \\
 Var\{Y(\underline{x})\} &= F(1 - F)B^2 + [F\sigma^{2(1)} + (1 - F)\sigma^{2(0)}]
 \end{aligned} \tag{23}$$

with:

$$m^{(1)} = E\{Y(\underline{x}) | I(\underline{x}) = 1\} \in [0, 1]$$

$$m^{(0)} = E\{Y(\underline{x}) | I(\underline{x}) = 0\} \in [0, 1]$$

$$B = m^{(1)} - m^{(0)} \in [-1, +1]$$

$$\sigma^{2(1)} = \text{Var}\{Y(\underline{x}) | I(\underline{x}) = 1\}$$

$$\sigma^{2(0)} = \text{Var}\{Y(\underline{x}) | I(\underline{x}) = 0\}$$

$m^{(1)}$ is the expected value of the pre-posterior information $Y(\underline{x})$ given that the actual value is $I(\underline{x}) = 1$; it can be inferred from a calibration data set where pre-posterior values $y(\underline{x}_\alpha)$ and $i(\underline{x}_\alpha)$ co-exist (see Fig. F-3). Similarly, the variance $\sigma^{2(1)}$ can be inferred from the variance of the data $y(\underline{x}_\alpha)$ for which the collocated $i(\underline{x}_\alpha) = 1$.

Similar definitions hold for $m^{(0)}$ and $\sigma^{2(0)}$.

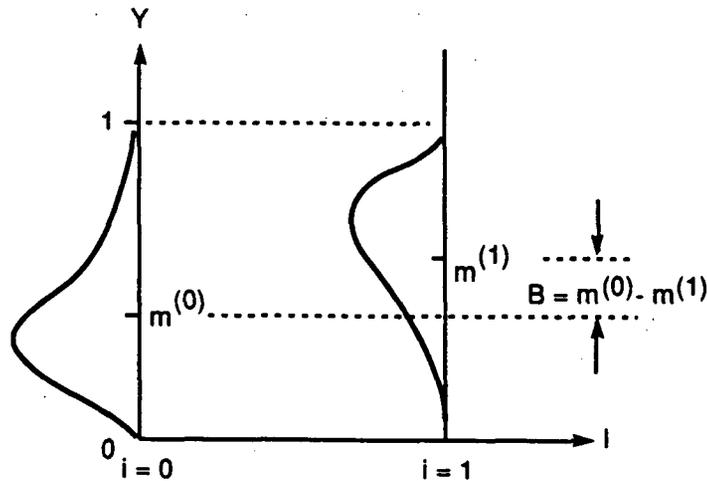


Figure F-3. Calibration of the information $Y(\underline{x})$.

Note: Using a data set for which pre-posterior information $y(\underline{x}_\alpha)$ and actual values $i(\underline{x}_\alpha)$ co-exist, the two conditional means $m^{(i)} = E\{Y(\underline{x}) | I(\underline{x}) = i\}$ are determined.

Information value:

Maximum conditional consistency for the information $y(\underline{x})$ would call for $m^{(1)} = 1$ and $m^{(0)} = 0$, i.e., $B = 1$.

Maximum inconsistency, such that $Y(\underline{x}) = 1 - I(\underline{x})$, would yield: $m^{(1)} = 0$ and $m^{(0)} = 1$, i.e., $B = -1$.

The case $m^{(1)} = m^{(0)}$, i.e. $B = 0$, corresponds to an average non-informative pre-posterior $Y(\underline{x})$.

- When $B=1$:

$Cov\{Y(\underline{x} + \underline{h}) \cdot Y(\underline{x})\} = Cov\{I(\underline{x} + \underline{h}) \cdot Y(\underline{x})\} = C_I(\underline{h})$, thus, as far as the projection theorem, the pre-posterior information $Y(\underline{x})$ is treated equal to the hard local information $I(\underline{x})$.

- When $B=-1$, similarly $Y(\underline{x})$ is treated equal to $1 - I(\underline{x})$.
- When $B=0$, all covariances involving $Y(\underline{x})$ vanish; therefore, the information $Y(\underline{x})$ is ignored in the normal equations.

Thus, it appears that the average value of the information $Y(\underline{x})$ is measured by B^2 , not B , just as the value of information in a traditional regression is measured by ρ^2 , not ρ the coefficient of correlation.

Updating: Unless $B = 1$, $Y(\underline{x})$ is not considered equal to $I(\underline{x})$; thus, any pre-posterior cdf $y(\underline{x})$ will be updated (i.e., modified) by information of type $y(\underline{x}_\alpha)$ or $i(\underline{x}_\alpha)$ available at surrounding locations \underline{x}_α .

Composite information:

The previous development and the relations (23) only consider one single type of pre-posterior information $Y(\underline{x})$. In case of composite information of type (19), where several types of pre-posterior information $Y_l(\underline{x}), l = 1, \dots, L$ co-exist, one would need to repeat the previous calibration exercise L times defining L consistency indices $B_l, l = 1, \dots, L$.

A more difficult problem will be that of evaluating the cross-covariances of the type $Cov\{Y_l(\underline{x} + \underline{h}), Y_{l'}(\underline{x})\}, l \neq l'$, measuring the degree of redundancy between these different sources of information. Such evaluation may require setting a hierarchy of data screening similar to the constitutive hypotheses (22). Additional research is warranted here.

5. COKRIGING HARD AND SOFT DATA

As developed in Sections 3 and 4,

- a hard data $Z(x)$ generates a full indicator data column:

$$[I(\underline{x}; z_k), k = 1, \dots, K] \text{ 'y, cf. relation (15)}$$

- soft local information can be coded as a pre-posterior distribution:

$$y(\underline{x}; z_k) = P\{Z(\underline{x}) \leq z_k | \underline{x}\} \text{ [cf. relation (18)], then randomized into } Y(\underline{x}; z_k).$$

The goal is to estimate (actually model) the posterior distribution: $E\{I(\underline{x}; z) | (n)\} = P\{Z(\underline{x}) \leq z | (n)\}$. This conditional expectation is estimated, e.g., by a linear combination of hard indicator data of the type $I(\underline{x}_{\alpha_1}, z_k)$ and fuzzy indicator data of the type $Y(\underline{x}_{\alpha_2}; z_k)$ $\alpha_1 = 1, \dots, n_1; \alpha_2 = 1, \dots, n_2, n_1, n_2 \in (n)$:

$$\begin{aligned} [P\{Z(x) \leq z_k | (n)\} - F(z_k)]^m &= \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1} \cdot [I(\underline{x}_{\alpha_1}; z_k) - F(z_k)] \\ &+ \sum_{\alpha_2=1}^{n_2} \nu_{\alpha_2} [Y(\underline{x}_{\alpha_2}; z_k) - E\{Y(\underline{x}_{\alpha_2}; z_k)\}] \end{aligned} \quad (24)$$

with: $E\{I(\underline{x}; z)\} = F(z) = P\{Z(x) \leq z\}$, for all \underline{x}

and: $E\{Y(\underline{x}; z)\}$ as given by later relation (31).

The $(n_1 + n_2)$ weights λ_{α_1} and ν_{α_2} are given by a system of $(n_1 + n_2)$ normal equations of type (7), also called the cokriging system. Dropping from the notations the threshold parameter z_k , this system is written:

$$\begin{aligned} \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1} \text{Cov}\{I(\underline{x}_{\beta_1}), I(\underline{x}_{\alpha_1})\} + \sum_{\beta_2=1}^{n_2} \nu_{\beta_2} \text{Cov}\{Y(\underline{x}_{\beta_2}), I(\underline{x}_{\alpha_1})\} \\ = \text{Cov}\{I(\underline{x}), I(\underline{x}_{\alpha_1})\}, \alpha_1 = 1, \dots, n_1 \\ \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1} \text{Cov}\{I(\underline{x}_{\beta_1}), Y(\underline{x}_{\alpha_2})\} + \sum_{\beta_2=1}^{n_2} \nu_{\beta_2} \text{Cov}\{Y(\underline{x}_{\beta_2}), I(\underline{x}_{\alpha_2})\} \\ = \text{Cov}\{I(\underline{x}), Y(\underline{x}_{\alpha_2})\}, \alpha_2 = 1, \dots, n_2. \end{aligned}$$

Solution of this system calls for knowledge of three covariance functions, hereafter assumed stationary:

- hard indicator covariance:

$$C_I(h) = \text{Cov}\{I(\underline{x} + h), I(\underline{x})\} \quad (25)$$

- fuzzy indicator covariance:

$$C_Y(\underline{h}) = Cov\{Y(\underline{x} + \underline{h}), Y(\underline{x})\}$$

• hard-fuzzy indicator cross variance:

$$C_{IY}(\underline{h}) = Cov\{I(\underline{x} + \underline{h}), Y(\underline{x})\}.$$

The hard indicator covariance $C_I(\underline{h}; z_k)$ is inferred directly from the hard indicator data $i(\underline{x}_{\alpha_1}; z_k)$, as deduced through relation (15) from the hard original data $z(\underline{x}_{\alpha_1})$. As for the two other covariances $C_Y(\underline{h})$, and $C_{IY}(\underline{h})$, they could also be inferred from actual data if there are enough collocated Y and I -data. Otherwise, building on a Markov hypothesis of type (22), they can be derived directly from the hard indicator covariance $C_I(\underline{h})$.

Covariance determination [proof of relations (23)]:

Let

$$\Phi^{(1)}(\underline{x}; p) = P\{Y(\underline{x}) \leq p \mid I(\underline{x}) = 1\} \in [0, 1] \quad (26)$$

be the conditional cumulative distribution function (cdf) of $Y(\underline{x})$. Recall that $Y(\underline{x})$ is itself a random cdf.

Similarly:

$$\Phi^{(0)}(\underline{x}; p) = P\{Y(\underline{x}) \leq p \mid I(\underline{x}) = 0\} \in [0, 1].$$

To allow inference of the covariances $C_Y(\underline{h})$ and $C_{IY}(\underline{h})$, we must assume (rather a decision) some form of stationarity for the random functions $Y(\underline{x}), I(\underline{x})$. Strict stationarity of $Y(\underline{x})$ and $I(\underline{x})$ entails stationarity of the previous cdf's (26) which can then be denoted without the location parameter \underline{x} as:

$$\Phi^{(1)}(p), \Phi^{(0)}(p) \text{ for the conditional cdf's} \quad (27)$$

$$\phi^{(1)}(p), \phi^{(0)}(p) \text{ for the corresponding pdf's}$$

with

$$\phi^{(i)}(p) = \frac{\partial \Phi^{(i)}(p)}{\partial p}, i = 1, 2.$$

* Stationarity for $Y(\underline{x}; z)$ is strictly a model decision that allows inference of the covariance $C_Y(\underline{h}; z)$. It still leaves the possibility for realizations $y(\underline{x}; z), y(\underline{x} + \underline{h}; z)$ to be different, that is, allows for different pre-posterior distributions at different locations.

Hypothesis: "Hard information $I(\underline{x})$ prevails over any other information when conditioning at location \underline{x} , and prevails over any other collocated information," that is:

$$\begin{aligned} E\{I(\underline{x} + \underline{h}) | I(\underline{x}) = 1, Y(\underline{x}) = p\} &= E\{I(\underline{x} + \underline{h}) | I(\underline{x}) = 1\} \\ &= K_I(\underline{h})/F, \quad \text{for all } p \in [0, 1] \end{aligned} \quad (28)$$

with:

$$\begin{aligned} E\{I(\underline{x})\} &= E\{I^2(\underline{x})\} = K_I(0) = F \\ K_I(\underline{h}) &= E\{I(\underline{x})I(\underline{x} + \underline{h})\} = C_I(\underline{h}) + F^2 \end{aligned}$$

and similarly:

$$\begin{aligned} E\{I(\underline{x} + \underline{h}) | I(\underline{x}) = 0, Y(\underline{x}) = p\} &= E\{I(\underline{x} + \underline{h}) | I(\underline{x}) = 0\} \\ &= \{F - K_I(\underline{h})\}/(1 - F), \quad \text{for all } p \in [0, 1]. \end{aligned}$$

Also:

$$\begin{aligned} P\{Y(\underline{x}) = p | I(\underline{x}) = i, Y(\underline{x} + \underline{h}) = p'\} \\ = P\{Y(\underline{x}) = p | I(\underline{x}) = i\}, \quad \text{for all } \underline{h}, p', i = 0, 1. \end{aligned}$$

• Using Bayes' relation, develop the joint probability:

$$\begin{aligned} P\{I(\underline{x} + \underline{h}) = 1 \cdot Y(\underline{x}) = p\} &= \\ P\{I(\underline{x} + \underline{h}) = 1 \cdot Y(\underline{x}) = p | I(\underline{x}) = 1\} \cdot F &+ P\{I(\underline{x} + \underline{h}) = 1 \cdot Y(\underline{x}) = p | I(\underline{x}) = 0\} \cdot (1 - F). \end{aligned}$$

Since:

$$\begin{aligned} P\{I(\underline{x} + \underline{h}) = 1, Y(\underline{x}) = p | I(\underline{x}) = 1\} \\ = P\{I(\underline{x} + \underline{h}) = 1 | Y(\underline{x}) = p, I(\underline{x}) = 1\} \cdot P\{Y(\underline{x}) = p | I(\underline{x}) = 1\} \\ = K_I(\underline{h})/F \cdot \phi^{(1)}(p). \end{aligned}$$

Similarly:

$$P\{I(\underline{x} + \underline{h}) = 1, Y(\underline{x}) = p | I(\underline{x}) = 0\} = \frac{F - K_I(\underline{h})}{1 - F} \cdot \phi^{(0)}(p).$$

Thus:

$$P\{I(\underline{x} + \underline{h}) = 1, Y(\underline{x}) = p\} = K_I(\underline{h}) \cdot \phi^{(1)}(p) + [F - K_I(\underline{h})] \cdot \phi^{(0)}(p).$$

The non-centered cross-covariance between $I(\underline{x} + \underline{h})$ and $Y(\underline{x})$ is obtained by integrating the previous density:

$$\begin{aligned} E\{I(\underline{x} + \underline{h}) \cdot Y(\underline{x})\} &= K_I(\underline{h}) \int_0^1 p \phi^{(1)}(p) dp + [F - K_I(\underline{h})] \int_0^1 p d\phi^{(0)}(p) \\ &= K_I(\underline{h}) \cdot m^{(1)} + \{F - K_I(\underline{h})\} \cdot m^{(0)} \\ &= K_I(\underline{h}) [m^{(1)} - m^{(0)}] + F m^{(0)} \end{aligned} \quad (29)$$

with:

$$m^{(1)} = E\{Y(\underline{x}) \mid I(\underline{x}) = 1\}: \text{mean of the pdf } \phi^{(1)}(p)$$

$$m^{(0)} = E\{Y(\underline{x}) \mid I(\underline{x}) = 0\}: \text{mean of the pdf } \phi^{(0)}(p).$$

The marginal pdf of $Y(\underline{x})$ appears as a mixture of the two conditional pdf's $\phi^{(1)}(p)$ and $\phi^{(0)}(p)$, indeed:

$$\begin{aligned} P\{Y(\underline{x}) = p\} &= P\{Y(\underline{x}) = p \mid I(\underline{x}) = 1\} \cdot F + P\{Y(\underline{x}) = p \mid I(\underline{x}) = 0\} \cdot (1 - F) \\ &= F \phi^{(1)}(p) + \{1 - F\} \phi^{(0)}(p). \end{aligned} \quad (30)$$

Thus, the mean of $Y(\underline{x})$ is:

$$E\{Y(\underline{x})\} = F m^{(1)} + (1 - F) m^{(0)}. \quad (31)$$

Recall: $E\{I(\underline{x})\} = F$.

Finally, the centered cross-covariance between hard and soft data is written from relations (29) to (31):

$$\begin{aligned} C_N(\underline{h}) &= \text{Cov}\{I(\underline{x} + \underline{h}), Y(\underline{x})\} \\ &= E\{I(\underline{x} + \underline{h}) \cdot Y(\underline{x})\} - E\{Y(\underline{x})\} \cdot E\{I(\underline{x})\} \\ &= \{m^{(1)} - m^{(0)}\} \cdot \{K_I(\underline{h}) - F^2\} = B \cdot C_I(\underline{h}) \end{aligned} \quad (32)$$

with $B = m^{(1)} - m^{(0)} \in [-1, +1]$

- Similarly, using Bayes' relation, develop the joint probability:

$$P\{Y(\underline{x}) = p, Y(\underline{x} + \underline{h}) = p'\} = P_1 + P_2 + P_3 + P_4,$$

with:

$$P_1 = P\{Y(\underline{x}) = p, Y(\underline{x} + \underline{h}) = p', I(\underline{x}) = 1, I(\underline{x} + \underline{h}) = 1\}$$

$$P_2 = P\{Y(\underline{x}) = p, Y(\underline{x} + \underline{h}) = p', I(\underline{x}) = 1, I(\underline{x} + \underline{h}) = 0\}$$

$$P_3 = P\{Y(\underline{x}) = p, Y(\underline{x} + \underline{h}) = p', I(\underline{x}) = 0, I(\underline{x} + \underline{h}) = 1\}$$

$$P_4 = P\{Y(\underline{x}) = p, Y(\underline{x} + \underline{h}) = p', I(\underline{x}) = 0, I(\underline{x} + \underline{h}) = 0\}$$

and:

$$\begin{aligned} P_1 &= P\{Y(\underline{x}) = p, Y(\underline{x} + \underline{h}) = p' \mid I(\underline{x}) = 1, I(\underline{x} + \underline{h}) = 1\} \cdot K_I(h) \\ &= K_I(h) \cdot [P\{Y(\underline{x}) = p \mid Y(\underline{x} + \underline{h}) = p', I(\underline{x}) = 1, I(\underline{x} + \underline{h}) = 1\}] \\ &\quad \cdot [P\{Y(\underline{x} + \underline{h}) = p' \mid I(\underline{x}) = 1, I(\underline{x} + \underline{h}) = 1\}]. \end{aligned}$$

Since the hard information $I(\underline{x})$ prevails over any other soft information at \underline{x} , the second term in square brackets is equal to $P\{Y(\underline{x}) = p \mid I(\underline{x}) = 1\}$. Similarly, the third term is equal to $P\{Y(\underline{x} + \underline{h}) = p' \mid I(\underline{x} + \underline{h}) = 1\}$. Thus:

$$P_1 = K_I(h) \cdot \phi^{(1)}(p) \cdot \phi^{(1)}(p'). \quad (33)$$

Similarly:

$$P_2 = [F - K_I(h)] \cdot \phi^{(1)}(p) \cdot \phi^{(0)}(p')$$

$$P_3 = [F - K_I(h)] \cdot \phi^{(0)}(p) \cdot \phi^{(1)}(p')$$

$$P_4 = [1 - 2F + K_I(h)] \cdot \phi^{(0)}(p) \cdot \phi^{(1)}(p').$$

And finally:

$$\begin{aligned} E\{Y(\underline{x}) \cdot Y(\underline{x} + \underline{h})\} &= \sum_{i=1}^4 \int_0^1 \int_0^1 pp' \cdot P_i dp dp' \\ &= B^2 K_I(h) + m^{(0)}[m^{(0)} + 2FB]. \end{aligned}$$

Recalling the expression (31) for the marginal mean of $Y(\underline{x})$, the centered covariance is:

$$\text{Cov}\{Y(\underline{x}), Y(\underline{x} + \underline{h})\} = B^2 \cdot C_I(h), \quad \text{for } h > 0 \quad (34)$$

with: $B = m^{(1)} - m^{(0)}$.

The hypothesis calling for dominance of hard data on soft data allows the direct determination of $C_Y(h)$ and $C_N(h)$ from the indicator covariance $C_I(h)$ and the consistency index B. That index measures the consistency of soft information with hard indicator information, and should be inferred from calibration data of the type of Fig. F-3.

Remarks:

$$C_Y(0) = \text{Var}\{Y(x)\} = E\{Y^2(x)\} - [E\{Y(x)\}]^2$$

with $E\{Y^2(x)\} = F[\sigma^{2(1)} + m^{2(1)}] + (1-F)[\sigma^{2(0)} + m^{2(0)}]$,

since the pdf of $Y(x)$ is a mixture of $\phi^{(1)}(p)$ and $\phi^{(0)}(p)$, and $\sigma^{2(1)}$, $\sigma^{2(0)}$ are the variances of respectively $\phi^{(1)}(p)$ and $\phi^{(0)}(p)$.

Accounting for the expression (31) for the mean, it comes:

$$\begin{aligned} C_Y(0) &= [F(1-F) \cdot B^2] + [F\sigma^{2(1)} + (1-F)\sigma^{2(0)}] \\ &= V_c^2 + V_f^2. \end{aligned}$$

V_c^2 can be read as a "consistency term," best if B^2 is close to one. V_f^2 is a "fluctuation term," increasing with the conditional variances $\sigma^{2(1)}$ and $\sigma^{2(0)}$ of the pre-posterior distribution information $Y(x)$. The smaller this term, the better the soft information $Y(x)$.

Now, from relation (34) it comes:

$$\lim_{h \rightarrow 0} C_Y(h) = B^2 C_I(0) = V_c^2 \leq .25$$

$$\text{since } C_I(0) = \text{Var}\{I(x)\} = F(1-F) \leq .25.$$

Thus, the covariance $C_Y(h) = B^2 C_I(h)$ is proportional to the indicator covariance $C_I(h)$, except for an additional discontinuity at the origin of amplitude V_f^2 (see Fig. F-4).

If $V_f^2 = 0$, i.e., if the soft information is precise no matter if not accurate ($B \neq 1$), as long as lack of consistency can be corrected through knowledge of B, the soft information $Y(x)$ is as good as a hard indicator data $I(x)$.

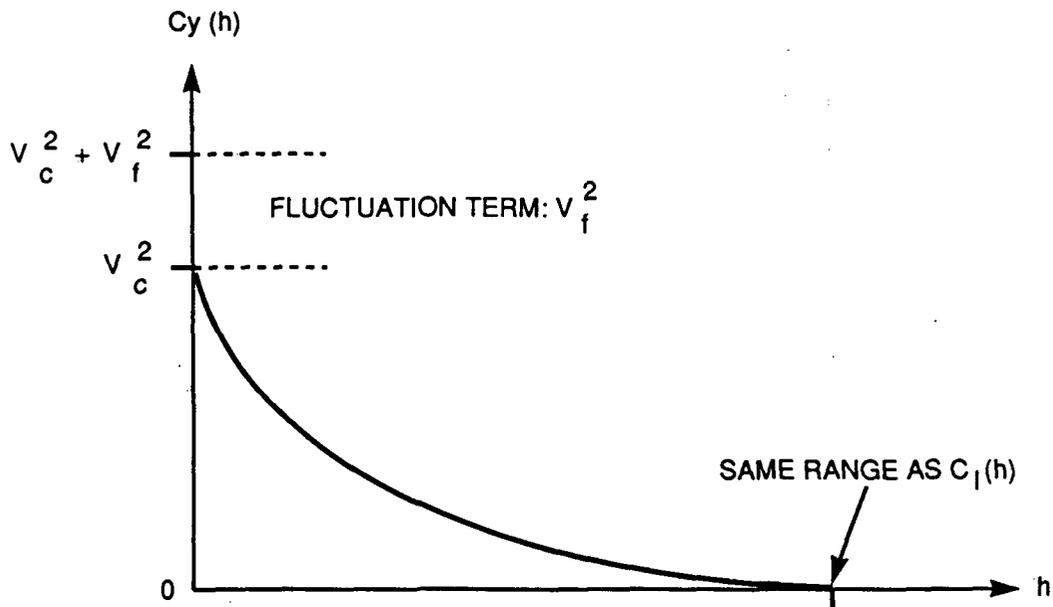


Figure F-4. Covariance relation between soft and hard data, (under dominance of hard data).

Note: Spatial autocorrelation of the soft data decreases as the fluctuation term V_f^2 (nugget effect) increases.

6. REFERENCES

1. Ripley, B. D., Stochastic Simulation, Wiley & Sons, 1987.
2. Journel, A. G., Fundamentals of Geostatistics in Five Lessons, Short Course in Geology, 5, Washington, D.C.: AGU, 1989.
3. Berger, J. O., Statistical Decision Theory and Bayesian Analysis, Springer Verlag, 1980.
4. Journel, A. G. and F. A. Alabert, "Non-Gaussian Data Expansion in the Earth Sciences," Terra Nova 1 (2), pp. 123-134, 1989.
5. Luenberger, V., Optimization by Vector Space Methods, Wiley & Sons, 1969.
6. Cox, D. R., "The Analysis of Multivariate Binary Data," Applied Statistics, 21 (2), pp. 113-120, 1972.
7. Journel, A. G. and C. J. Huijbregts, Mining Geostatistics, Academic Press, 1978.

7. BIBLIOGRAPHY

- Anderson, T. W. (1984). Multivariate Statistical Analysis, Wiley & Sons.
- Benjamin, J. R. and A. Cornell, Probability, Statistics and Decision for Civil Engineers, McGraw-Hill, 1970.
- Borgman, L. E. and W. F. Quimby, "Sampling for Tests of Hypothesis When Data Are Correlated in Space and Time," in Principles of Environmental Sampling, L. Keith (ed.), American Chemical Society, pp. 25-43, 1988.
- Desbarats, A., "Numerical Estimation of Effective Permeability in Sand-Shale Formations," WRR 23 (2), pp. 273-286, 1987.
- Efron, B., The Jackknife, the Bootstrap and Other Resampling Plans, SIAM, 1982.
- Gomez-Hernandez, J. J. and A. G. Journel, "Conditional Simulation of the Wilmington Sand-Shale Sequence," Hydraulic Engineering, Ports, M. A. (ed.), New York: ASCE, pp. 801-806, 1989.
- Johnson N. L. And Kotz S., Continuous Univariate Distributions - 1, Wiley & Sons, 1970.
- Journel, A. G. And M. Rossi, "When Do We Need a Trend Model in Kriging?," Math Geology, 21 (7), pp. 715-740, 1989.
- Hoeksema R. J. And P. Kitanidis, "Analysis of the Spatial Structure of Properties of Selected Aquifers," WRR 21 (4), pp. 563-572, 1985.
- Gelhar, L. W. And C. L. Axness, "Three Dimensional Stochastic Analysis of Macrodispersion in Aquifers," WRR 19(1), pp. 161-180, 1983.
- Gelhar, L. W., A. L. Gutjahr, and R. L. Naff, "Stochastic Analysis of Macrodispersion in a Stratified Aquifer," WRR 15 (6), pp. 1387-1397, 1979.
- Solow, A., "Bootstrapping Correlated Data," Math. Geology 17 (7), pp. 769-775, 1985.

BIBLIOGRAPHIC DATA SHEET

(See instructions on the reverse)

1. REPORT NUMBER
(Assigned by NRC, Add Vol., Supp., Rev.,
and Addendum Numbers, if any.)

NUREG/CR-5639

2. TITLE AND SUBTITLE

Uncertainty Evaluation Methods for Waste Package Performance Assessment

3. DATE REPORT PUBLISHED

MONTH	YEAR
January	1991

4. FIN OR GRANT NUMBER

FIN D1035

5. AUTHOR(S) Y.-T. Wu¹, A. G. Journel², L. R. Abramson³, P. K. Nair¹

1. Center for Nuclear Waste Regulatory Analyses, Southwest Research Institute
2. Applied Earth Sciences Department, Stanford University
3. Probabilistic Risk Analysis Branch, Office of Nuclear Regulatory Research, U.S. Nuclear Regulatory Commission

6. TYPE OF REPORT

Technical

7. PERIOD COVERED (Inclusive Dates)

8. PERFORMING ORGANIZATION - NAME AND ADDRESS (If NRC, provide Division, Office or Region, U.S. Nuclear Regulatory Commission, and mailing address; if contractor, provide name and mailing address.)

Center for Nuclear Waste Regulatory Analyses
Southwest Research Institute
6220 Culebra Road
San Antonio, TX 78228

9. SPONSORING ORGANIZATION - NAME AND ADDRESS (If NRC, type "Same as above"; if contractor, provide NRC Division, Office or Region, U.S. Nuclear Regulatory Commission, and mailing address.)

Division of High-Level Waste Management
Office of Nuclear Material Safety and Safeguards
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

10. SUPPLEMENTARY NOTES

11. ABSTRACT (200 words or less)

This report identifies and investigates methodologies to deal with uncertainties in assessing high-level nuclear waste package performance. Four uncertainty evaluation methods (probability-distribution approach, bounding approach, expert judgment, and sensitivity analysis) are suggested as the elements of a methodology that, without either diminishing or enhancing the input uncertainties, can evaluate performance uncertainty. Such a methodology can also help identify critical inputs as a guide to reducing uncertainty so as to provide reasonable assurance that the risk objectives are met. This report examines the current qualitative waste containment regulation and shows how, in conjunction with the identified uncertainty evaluation methodology, a framework for a quantitative probability-based rule can be developed, which takes account of the uncertainties. Current NRC regulation requires that the waste packages provide "substantially complete containment" (SCC) during the containment period. The term "SCC" is ambiguous and subject to interpretation. This report, together with an accompanying report which describes the technical considerations that must be addressed to satisfy high-level waste containment requirements, provides a basis for a third report to develop recommendations for regulatory uncertainty reduction in the "containment" requirement of 10 CFR Part 60.

12. KEY WORDS/DESCRIPTORS (List words or phrases that will assist researchers in locating the report.)

waste package; engineered barrier system; waste containment; high-level radionuclide waste management; waste management; performance assessment; probabilistic performance assessment; failure mode; time-to-failure; uncertainty evaluation; uncertainty propagation; uncertainty modeling; uncertainty characterization; uncertainty reduction; risk assessment; reliability analysis; failure probability; safety margin; safety factor; probability-distribution approach; bounding approach; expert judgment; sensitivity analysis; Monte Carlo simulation; stratified sampling; Latin hypercube sampling; importance sampling; geostatistics; fast probability integration; Bayesian geostatistics

13. AVAILABILITY STATEMENT

Unlimited

14. SECURITY CLASSIFICATION

(This Page)

Unclassified

(This Report)

Unclassified

15. NUMBER OF PAGES

16. PRICE

THIS DOCUMENT WAS PRINTED USING RECYCLED PAPER.



**UNITED STATES
NUCLEAR REGULATORY COMMISSION
WASHINGTON, D.C. 20555**

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, \$300

SPECIAL FOURTH-CLASS RATE
POSTAGE & FEES PAID
USNRC
PERMIT No. G-67