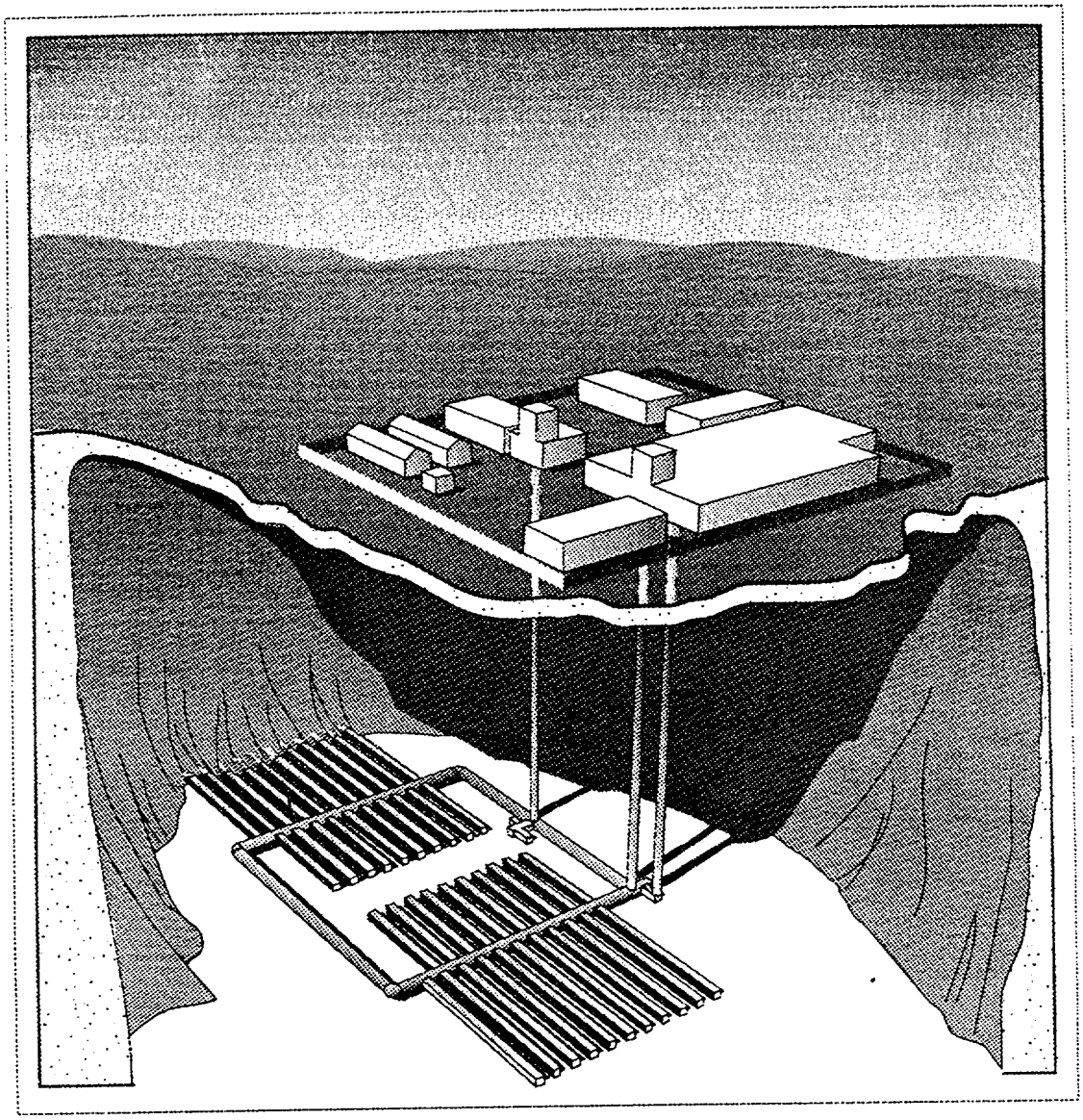
rec'd w/ap letter
dtd. 7/13/92

Repository Integration Program



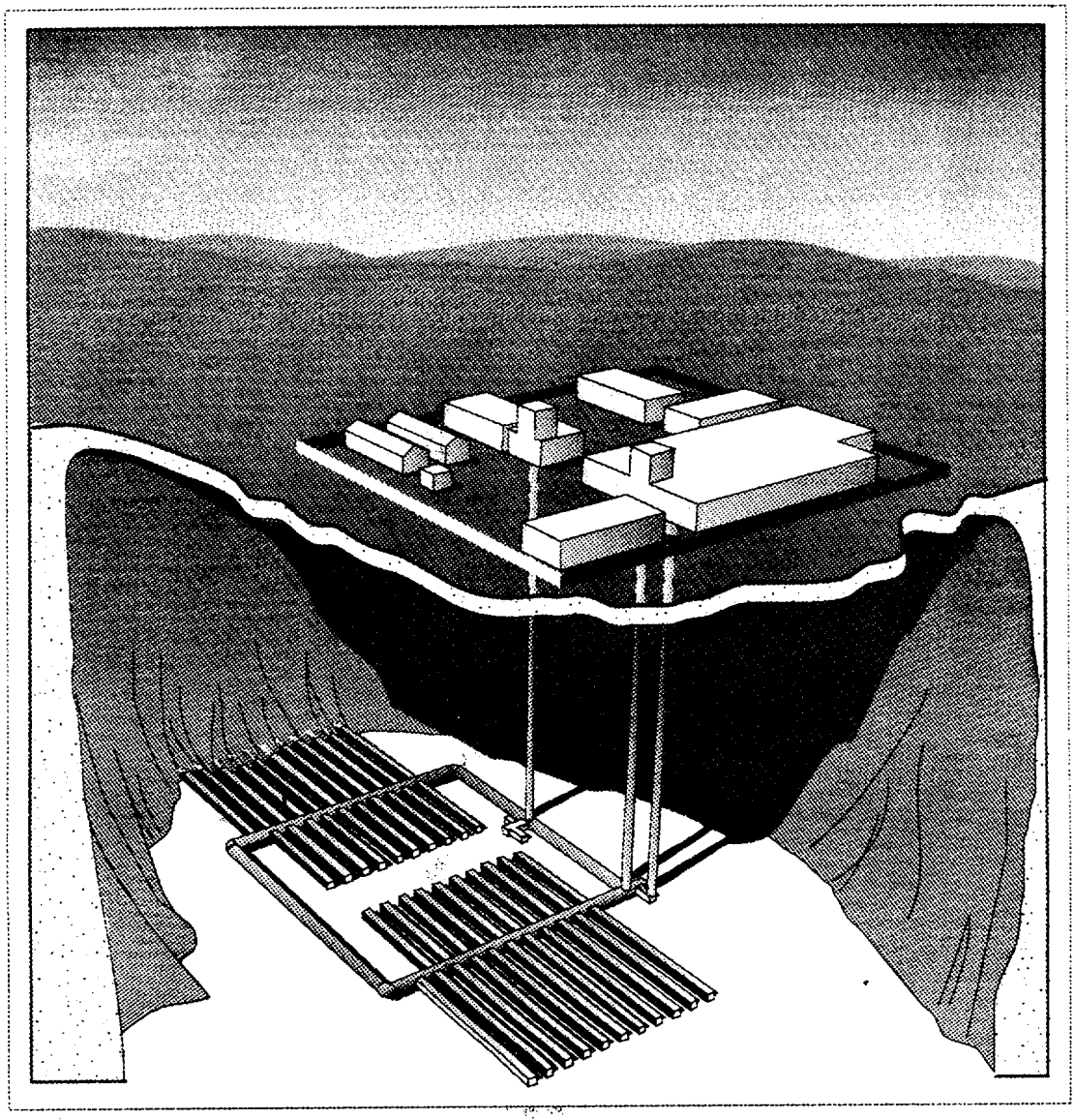
The RIP Repository Performance Assessment and Strategy Evaluation Model: Theory and Capabilities

Golder Associates

102

rec'd w/ap letter dtd. 7/13/92

Repository Integration Program



The RIP Repository Performance Assessment and Strategy Evaluation Model: Theory and Capabilities

Golder Associates

102

Golder Associates Inc.

4104-148th Avenue, NE
Redmond, WA USA 98052
Telephone (206) 883-0777
Fax (206) 882-5498

RIP REPOSITORY PERFORMANCE ASSESSMENT AND STRATEGY EVALUATION MODEL: THEORY AND CAPABILITIES

Version 1.00

I. Miller
R. Kossik
M. Cunnane

Golder Associates Inc.

Prepared for the U.S. DOE Office of Civilian Radioactive Waste Management

Work Performed Under Argonne National Laboratory Contract Number 02352401

March 31, 1992

903-1371.203

ABSTRACT

This report describes the theory and capabilities of RIP (Repository Integration Program). RIP is a powerful and flexible computational tool for carrying out probabilistic integrated total system performance assessments for geologic repositories. It embodies probabilistic decision analysis tools that allow it to:

- examine parameter sensitivity;
- evaluate alternative conceptual designs; and
- evaluate alternative site characterization strategies.

The primary purpose of RIP is to provide a management tool for guiding system design and site characterization. In addition, the performance assessment model (and the process of eliciting model input) can act as a mechanism for integrating the large amount of available information into a meaningful whole (in a sense, allowing one to keep the "big picture" and the ultimate aims of the project clearly in focus). Such an integration would be useful both for project managers and project scientists.

RIP is based on a "top down" approach to performance assessment which concentrates on the integration of the entire system, and utilizes relatively high-level descriptive models and parameters. The key point in the application of such a "top down" approach is that the simplified models and associated high-level parameters must incorporate an accurate representation of their uncertainty. RIP is designed in a very flexible manner such that details can be readily added to various components of the model without modifying the computer code. Uncertainty is also handled in a very flexible manner, and both parameter and model (process) uncertainty can be explicitly considered. Uncertainty is propogated through the integrated PA model using an enhanced Monte Carlo method.

RIP (and any practical PA model) must rely heavily on subjective assessment (expert opinion) for much of its input. The process of eliciting the high-level input parameters required for RIP is critical to its successful application. As a result, in order for any project to successfully apply a tool such as RIP, an enormous amount of communication and cooperation must exist between the data collectors, the process modelers, and the performance assessment modelers.

ACKNOWLEDGEMENTS

The authors of this document are not experts in technical areas of waste package behavior, hydrology of the Yucca Mountain site, or disruptive events. The model that is presented here incorporates concepts from a number of individuals who generously gave of their time to attend a series of workshops held by Golder Associates over the past year. These workshops were attended by numerous scientists and engineers, both from within and outside of the Yucca Mountain project. The workshops were attended by representatives from Lawrence Livermore National Laboratory (which hosted one of the workshops), Sandia National Laboratory, Los Alamos National Laboratory, the U.S. Geological Survey (USGS), Lawrence Berkeley Laboratory, RE/SPEC, Battelle Pacific Northwest Laboratory, Science Application International Company (SAIC), Roy F. Weston Inc., Electric Power Research Institute (EPRI), Yucca Mountain Site Characterization Project Office (YMP), Mifflin and Associates, Geomatrix Consultants, Intera Sciences, University of Minnesota, and the Georgia Institute of Technology. We would like to thank these organizations for their input, without which this model could not have been developed.

In addition, we would like to thank Russ Dyer and Jerry Boak of the YMP, who were instrumental in helping to bring the various parties together for the workshops.

Although input from the organizations listed above was utilized in the design of RIP, Golder Associates formulated the computational structure of the model independently and is solely responsible for the contents of this document.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont.)

TABLE OF CONTENTS (Cont.)

LIST OF FIGURES

TABLE OF CONTENTS (Cont.)

LIST OF APPENDICES

EXECUTIVE SUMMARY

Introduction

The Office of Civilian Radioactive Waste Management (OCRWM) is responsible for providing oversight and guiding the implementation of specific technical studies conducted by the Yucca Mountain Site Characterization Project Office (YMP). As part of this effort, the Director of the OCRWM has requested Golder Associates Inc. (GAI) to develop a methodology and recommend a strategic plan of action for evaluating the suitability of the Yucca Mountain site for the development of a repository.

The approach selected by GAI is based on two cornerstones:

- Development and continual updating of a preliminary *total system performance assessment model*; and
- Integration of the performance assessment model with various proposed design and site characterization activities using a *probabilistic decision analysis* approach in order to evaluate alternative strategies and guide and prioritize the site characterization process.

The integrated repository performance assessment and strategy evaluation software developed by GAI is known as RIP (Repository Integration Program). The major portion of the software is the total system performance assessment model, which consists primarily of a series of inter-connected, coupled component models with input/output relationships for radionuclide transfer. The performance model itself is embedded within a probabilistic decision analysis model which allows the user to evaluate alternative site characterization strategies.

Report Contents

This report provides a detailed description of the theory and capabilities of the RIP performance assessment and strategy evaluation model (which is briefly summarized below). It defines the overall methodology on which the program is based and describes the actual algorithms embodied within the software. It does not specifically describe the application of the software to the Yucca Mountain (or any other) site. A user's guide for the software and a description of an application of RIP to the Yucca Mountain site are available as separate documents.

Overview of Methodology

The Role of Performance Assessment

Traditionally, performance assessment has been considered as an essential method to be implemented at the later stages in a repository siting program, after the site has been characterized. However, performance assessment can also be used as a tool to aid in the

development and implementation of the entire siting procedure. In fact, GAI and others have long maintained that the only effective way to make valid site characterization decisions (what should be studied, when, and in what detail) and to improve conceptual designs is by using preliminary integrated performance assessments. This view has recently been reiterated by both the National Research Council (1990) and the Nuclear Waste Technical Review Board (1990).

This view of an ongoing performance assessment-driven characterization process is based on the premise that by evaluating system performance using *currently available knowledge* and the *anticipated level of knowledge that can be obtained following completion of a characterization activity*, it will be possible to identify which activities will be most critical to rapid and accurate determination of site suitability. That is, integrated performance assessment is intended to be used in an ongoing review process to continually reevaluate which data needs are most critical, and to reallocate program resources appropriately. Figure ES-1 illustrates the intended process graphically.

Development of the Performance Assessment Model

At the heart of the GAI methodology is the integrated total system performance assessment model. To assist in the conceptual development of the model, GAI held a number of workshops, which were attended by scientists and engineers from both within and outside of the Yucca Mountain Project. Based on these workshops, it became apparent that the level of understanding of the site (and designs) is such that a practical performance assessment model based completely on low-level physically-based parameters is not currently feasible. That is, the uncertainty in quantifying the basic physical processes controlling waste release, groundwater flow, and radionuclide transport at the site is such that GAI concluded that a practical performance assessment model must rely on relatively high-level "lumped" descriptive parameters (e.g., radionuclide travel times, container failure rates), which in turn are based on lower-level "process" models and on subjective assessments from qualified experts.

Moreover, it is GAI's view that due to computational constraints and the inherent complexity of such systems, it is unlikely that this situation will change in the foreseeable future. Although better simulation tools, more powerful computers, and additional field data will undoubtedly shed considerable light on the controlling processes, integrated total system performance models will continue to rely primarily on high-level parameters. Additional detailed process modelling and field studies will simply act to increase the accuracy and decrease the uncertainty in the experts' subjective assessments of these parameters.

As a result, the total system performance assessment model incorporated into RIP *is not* a detailed deterministic model based completely on low-level physically-based parameters (e.g., it does not explicitly simulate radionuclide transport by solving a three-dimensional advection-diffusion equation). Rather, it is a *descriptive probabilistic model* based on a *"top down" approach* to performance assessment that *describes* rather than *explains* the system behavior and is intended to directly represent the uncertainties in processes and events and their controlling parameters. This is not to say that detailed models of the controlling

the activity. This information is required in order to actually *prioritize* activities into an efficient and effective strategy. For example, suppose two proposed activities are both concerned with studying a particular parameter which has been shown to have a strong influence on system performance. While we can state that these activities both merit further consideration, we can not determine whether both activities are necessary, or which would be most effective, without quantifying the extent to which each activity will reduce the uncertainty in system performance, and incorporating cost and duration considerations.

Moreover, because of the complexity of the system and the large number of interconnected activities, actual prioritization of activities can not be based simply on an evaluation of individual activities with respect to cost, duration, and reduction in performance uncertainty. Rather, the entire site characterization plan must be considered as a whole. The proposed methodology accomplishes this prioritization by defining and evaluating *alternative characterization strategies* (a characterization strategy consisting of a specific set or sequence of characterization activities). The strategy portion of RIP provides the second cornerstone of the GAI approach by quantitatively integrating the performance assessment model with the characterization activities. The strategy model is essentially a decision analysis shell around the performance assessment model which allows the user to evaluate the alternative site characterization strategies.

The performance assessment model and the strategy evaluation model are discussed in more detail below.

Performance Assessment Model Overview

Basic Concepts

The integrated total system performance model consists of three coupled components which address:

- waste package behavior and radionuclide release;
- radionuclide transport pathways to the accessible environment; and
- disruptive events (such as volcanism and human intrusion) which can affect system parameters.

Figure ES-2 is a schematic of the performance assessment model structure.

As pointed out previously, RIP is a *descriptive* model which relies heavily on subjective assessments of relatively high-level descriptive parameters. It is also a *probabilistic* model intended to represent the uncertainties in processes and events. That is, uncertainty in both the model parameters and the component models themselves can be explicitly represented by RIP. Due to the inherent uncertainties resulting from our lack of knowledge, many model parameters will be represented by probability density functions (pdfs). The integrated performance assessment model uses a *simulation approach*, utilizing

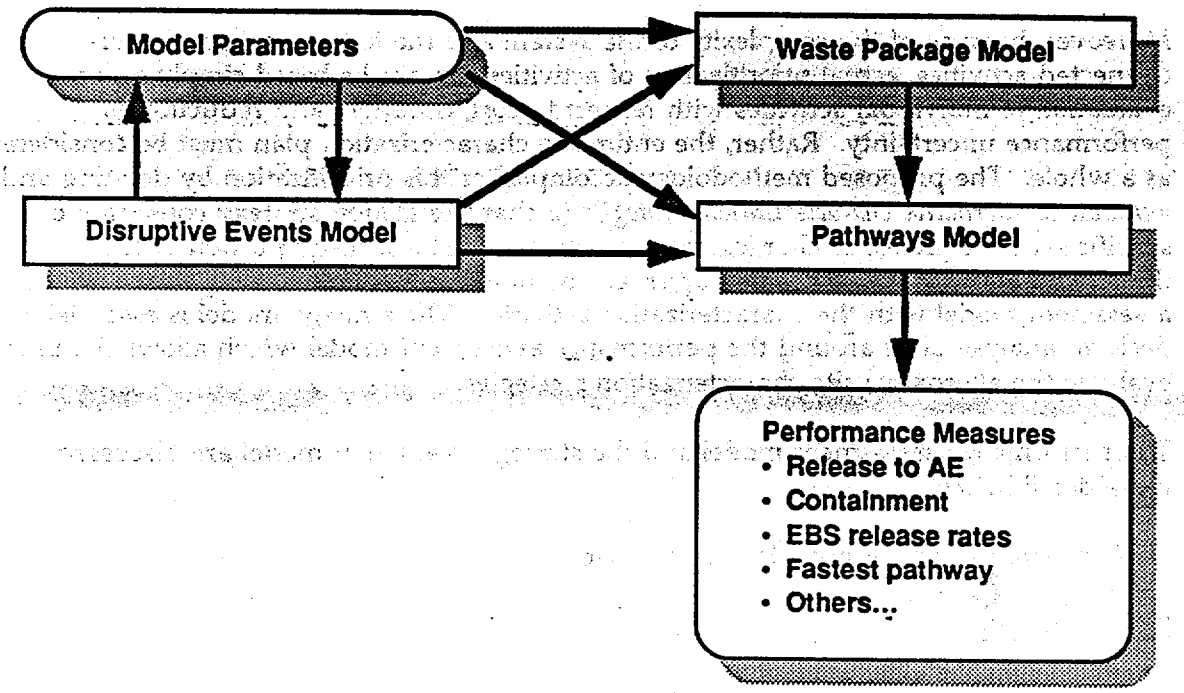


FIGURE ES

RIP PERFORMANCE ASSESSMENT MODEL COMPONENTS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DWG. NO. 27822 DATE 2/19/92 DRAWN TB

Golder Associates

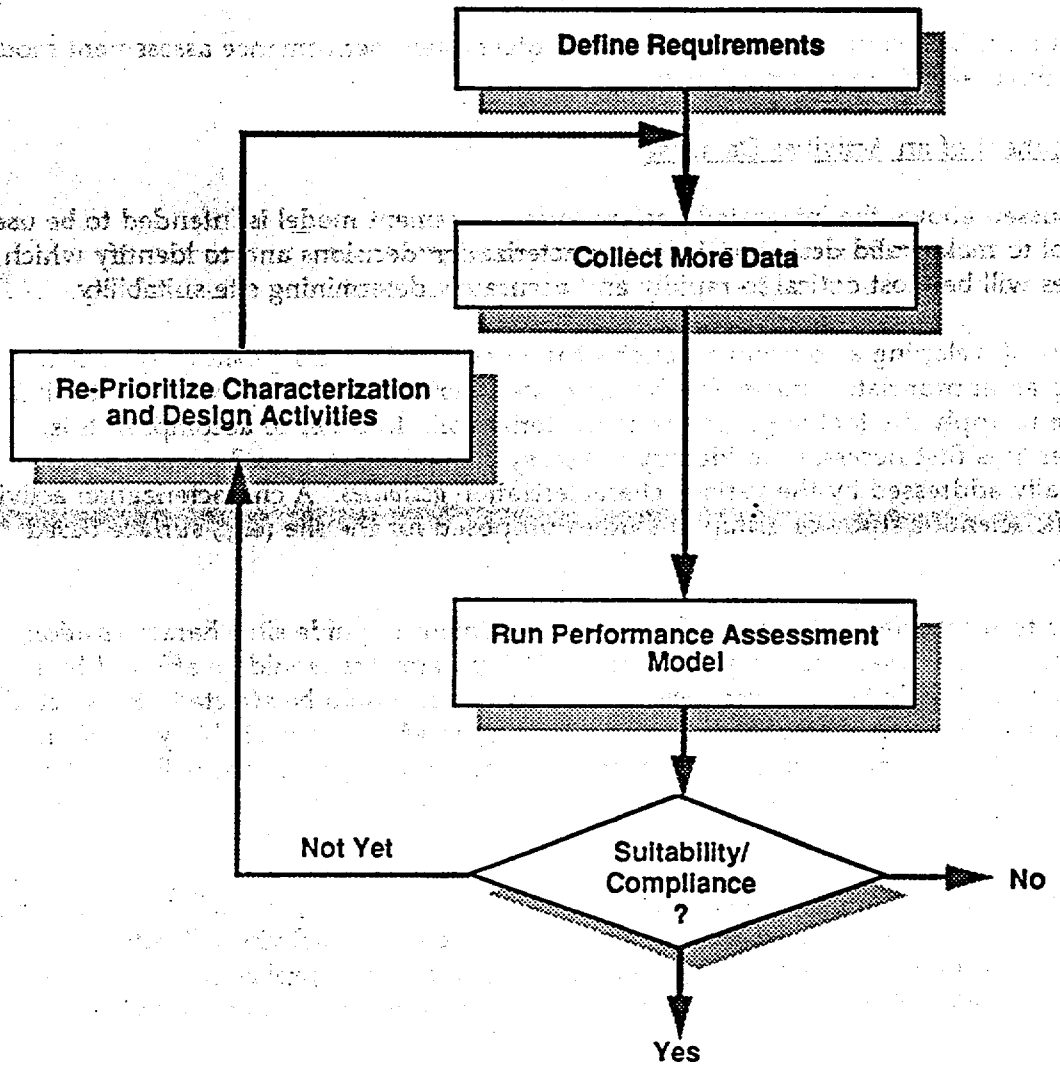


FIGURE ES-1
METHODOLOGY OVERVIEW
ANALYSIS/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DWG. NO. 27821 DATE 3/30/92 DRAWN TB

Golder Associates

processes are ignored. Quite to the contrary, detailed process models form the foundation for a systems model such as RIP and are required to generate the appropriate input parameters (e.g., in the form of response surfaces or analytical expressions).

The structure and concepts of the integrated total system performance assessment model will be discussed in greater detail below.

Development of an Activities Database

As discussed above, the integrated performance assessment model is intended to be used as a tool to make valid design and site characterization decisions and to identify which activities will be most critical to rapidly and accurately determining site suitability.

After first developing a conceptualization of the behavior of the repository system and creating an appropriate data set for the integrated performance assessment model, it is possible to apply the tool to guide site characterization. In order to accomplish this, however, it is first necessary to identify which system parameters and processes are specifically addressed by the various characterization *activities*. A characterization activity is a specific scientific study or group of studies proposed for the site (e.g., surface-based testing).

In order to most effectively use performance assessment to guide site characterization, however, it is necessary not only to identify *which* parameters would be affected by a particular activity, but to quantify *how* those parameters would be affected. In effect, this entails quantifying how a proposed activity is likely to affect the probability distribution describing the uncertainty in a given system parameter (i.e., how much will we learn?). This requires evaluations by experts of the nature of the uncertainty in a parameter, and the extent to which it is likely to be resolved by a particular activity.

An *activities database* must be developed which includes the name of each proposed activity, a brief description, identification of the system parameters the activity will affect, a quantitative estimate of how those parameters will be affected, probabilistic estimates of cost and duration of the activity, and precedence requirements for other activities.

Evaluating Characterization Activities and Strategies

Having developed an activities database, the simplest way to use RIP to evaluate characterization activities is to post-process the results of a performance assessment such that the sensitivities of various system performance measures (e.g., cumulative radionuclide release to the accessible environment) to individual model parameters are identified. In simple terms, activities which provide information on model parameters to which system performance is sensitive should be given the highest priority.

Although such an approach is useful in providing a preliminary screening and evaluation of site activities, it is limited because while it qualitatively considers which parameters are affected by a particular activity, it does not quantitatively incorporate the extent to which parameter uncertainty might be reduced by a particular activity, nor the cost or duration of

the Monte Carlo method to sample the probability distributions for the uncertain parameters (describing both processes and events) and simulate a large number of system realizations in order to determine probability distributions of site performance (e.g., cumulative release). That is, RIP creates a time history of disruptive events and other system parameters for each system realization, simulates the behavior of the system under those conditions, and then combines the results of all the realizations in an appropriate manner to determine probability distributions of site performance. RIP utilizes both *importance sampling* and *stratified (Latin-Hypercube) sampling* to increase the efficiency of the Monte Carlo sampling process.

In general terms, the output for the performance assessment model consists of performance measures for the repository system. A variety of performance measures can be considered (e.g., cumulative radionuclide release to the accessible environment, maximum annual release from the waste packages). These performance measures are probabilistic in nature. That is, output is not a single value, but a distribution which specifies the probability of exceedence for any particular value of a performance measure.

Flexibility of the Software

A fundamental feature of the RIP methodology, which it is critical to understand in order to appreciate the power and flexibility of the software, is the concept that a given model parameter can be specified by the user as a *constant*, a *stochastic* (i.e., represented by a probability distribution), or as a *function of time or of other parameters* (which themselves can be constants, stochastics, or functions).

Representing system parameters as stochastics allows the RIP user to directly specify the degree of uncertainty in a particular parameter. The RIP user interface (which consists of a series of interactive menus and pop-up input windows) allows the user to choose from a wide variety of probability distributions. For each Monte Carlo realization of the repository system, all of the stochastic parameters are sampled from their specified distributions. The probability distributions used for stochastic parameters can themselves be stochastic, or functions of other parameters.

The ability to represent system parameters as functions of other parameters imparts to the user the ability to readily add detail to any given system parameter or process represented by RIP. The user can even create and specify processes and parameters which are not explicitly incorporated in RIP. In a sense, the RIP program is similar to a spreadsheet: while it contains a large amount of built-in logic and calculational capabilities, the problem that is solved is entirely defined by the user. That is, RIP has been developed such that it is relatively free of assumptions regarding the details of waste package behavior and radionuclide transport processes. Hence, the model consists of a basic computational framework representing the controlling processes, but is intentionally flexible such that the user can represent the processes with as much detail as desired.

The user interface of the software is designed such that for function-type parameters the user simply types in the desired function. In addition to standard mathematical operators (e.g., SIN, COS, MAX, ERF, LOG), the user can define functions using relational operators

(e.g., >,<, =) and IF,THEN logic. Such flexibility allows the user to easily modify and add detail to the conceptual and computational model *without having to make changes to the software* (i.e., without modifying and compiling the source code). This allows the conceptual model to be continuously and easily modified as more information becomes available.

Components of the Performance Assessment Model

The major features of the three component models which comprise the performance assessment model are summarized briefly here.

The *waste package behavior and radionuclide release component* requires as inputs descriptions of the radionuclide inventories in the waste packages, a description of near field environmental conditions (which may be defined as temporally and spatially variable), and subjective assessments of high-level parameters describing container failure, matrix alteration/dissolution, and radionuclide mass transfer. The waste package component model can simulate two layers of containment (e.g., outer package and zircaloy cladding). Waste package failure rates, along with matrix alteration/dissolution rates, are used to compute the rate at which radionuclides are *exposed*. Once exposed, RIP computes the rate of *mass transfer* out of and away from the waste package. Parameters describing waste package failure and radionuclide exposure and mass transfer can be described as a function of near-field environmental conditions.

The output from this component (for each system realization) consists of time histories of release for each radionuclide from the waste packages, and acts as the input for the transport pathways component.

The task of the *radionuclide transport pathways component* is to probabilistically simulate radionuclide transport through the near and the far field. Workshops held by GAI indicated that existing continuum-based models, while representing portions of the system well, *fail to capture the essence of the overall hydrologic system*. While this may change as more data is collected, the current model must be capable of representing what is known about the site *at the present time*. Thus, the RIP model uses a phenomenological approach which attempts to *describe* rather than *explain* the system.

The resulting transport algorithm is based on a network of user defined *pathways*. The pathways reflect the major features of the hydrologic system and are conduits through which transport occurs. The pathways may be used for both flow balance and radionuclide transport purposes, and may account for either gas or liquid phase transport. The purpose of a pathway is to represent large scale heterogeneity of the hydrologic system, such as geologic structures and formation scale stratigraphy. In highly complex systems such as Yucca Mountain, up to 30 pathways may be required.

The pathways may be subdivided into *flow modes*, which address heterogeneity at the local scale (e.g., flow in rock matrix, flow in fractures). The flow modes are primarily distinguished from one another based on flow velocity in the mode, although retardation parameters may also differ between flow modes.

The actual transport of radionuclides along a pathway is based on a *breakthrough curve,* which is developed as a cumulative probability distribution for radionuclide travel times along a pathway. The breakthrough curve combines the effects of all flow modes and retardation on the radionuclide travel time, and determines the expected proportion of mass which has traversed the pathway by any specified time. The breakthrough curve is computed based on a Markov process algorithm for exchange between different flow modes.

The third performance assessment component model represents *disruptive events.* Disruptive events are defined as discrete occurrences which have some quantifiable effect on the processes described by the other two component models. Examples of disruptive events include volcanism, faulting, and human intrusion. The user first identifies all *significant* events (i.e., events that are both credible and consequential). Having done so, each event is assigned a rate of occurrence and, if desired, one or more *descriptor parameters,* which define the characteristics and magnitude of the event (e.g., length of a volcanic dike). Descriptor parameters may be described stochastically. Event occurrences are simulated as Poisson processes.

The user defines probability distributions for the event *consequences* (which may be functions of event descriptors). A consequence may take the form of a number of *discrete responses* (e.g., disrupting a number of waste packages, moving radionuclides from some waste packages directly to the accessible environment). It is also possible for an event to directly modify parameters defined in the other two component models, and this capability can be used to specify *long-term consequences* (e.g., raising the water table or opening a new pathway).

Strategy Evaluation Model Overview

The performance assessment model is embedded within a decision analysis model which allows the user to evaluate alternative site characterization strategies, where a strategy is defined as a group of activities. Figure ES-3 provides a schematic of the RIP strategy evaluation model.

For any user-specified characterization strategy, RIP provides three outputs by which alternative strategies can be ranked and compared:

- a probabilistic estimate of cost;
- a probabilistic estimate of duration; and
- a probabilistic evaluation of the predicted site performance resulting from implementation of the strategy (i.e., "what will our performance predictions be <u>after</u> we carry out this strategy?").

Computation of the first two outputs is straightforward, and consists of simply integrating within a Monte Carlo framework the cost and duration estimates for the individual

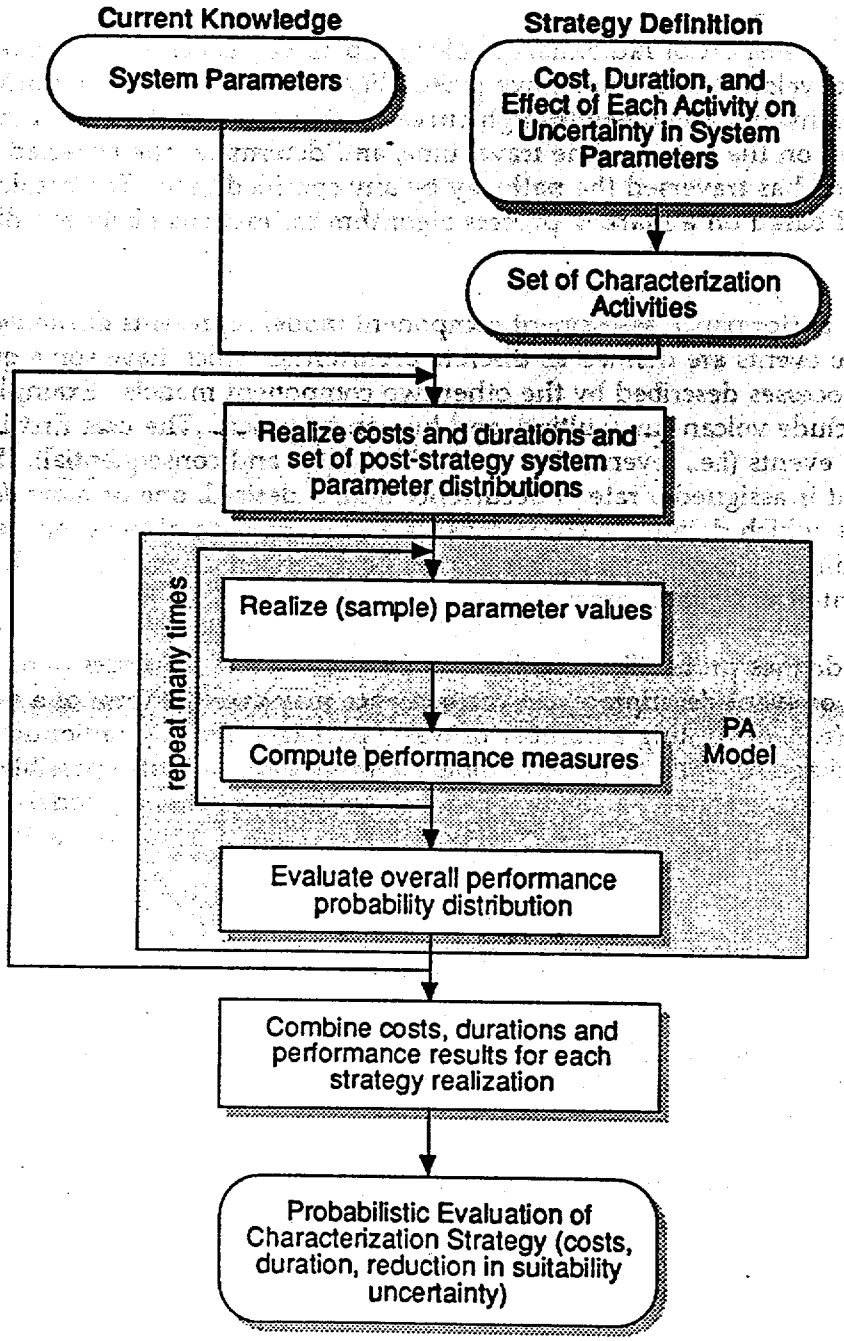


FIGURE ES
OVERVIEW OF RIP
STRATEGY EVALUATION
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 27827 DATE 2/19/92 DRAWN BY TB

Golder Associates

activities, taking into account any precedence requirements. The third output relies on subjective assessments by experts of the extent to which model parameter uncertainty will be reduced by a particular activity. Given these assessments, along with the current state of knowledge, RIP uses a Bayesian computational algorithm to simulate how probability distributions representing parameter uncertainty will change as a result of a particular characterization strategy, and develops a probabilistic evaluation of the repository performance that will be predicted after carrying out the strategy. Based on these outputs, it is possible to evaluate that effectiveness of alternative site characterization strategies.

Model Application

As discussed above, the primary objective of RIP is to provide a management tool for guiding design and site characterization. A related and equally important use for the model has become apparent to GAI in the process of eliciting information regarding models and input data from scientists and engineers investigating the Yucca Mountain site. Namely, the performance assessment model (and the process of eliciting model input) can act as a mechanism for integrating the large amount of available information generated by different groups into a meaningful whole (in a sense, allowing one to keep the "big picture" and the ultimate aims of the project clearly in focus). Such an integration would not only be useful for project managers (and is ultimately required in order to carry out a meaningful assessment of site suitability), but also to individual project scientists to help them formulate and present the results of their research in a manner such that it can be *readily applied* to the ultimate goal of the Project, *predicting the performance and determining the suitability of the site.*

RIP (and any other practical PA model) must rely heavily on subjective assessment (expert opinion) for much of its input. The process of eliciting the high-level input parameters required for RIP is critical to its successful application. As a result, in order for any project to successfully apply a tool such as RIP, an enormous amount of communication and cooperation must exist between the data collectors, the process modelers, and the performance assessment modelers. This is because such a tool must constantly evolve and is only valid for decision making when it actually incorporates the *current state of knowledge.* This is only possible if project scientists *think in terms of performance assessment,* at least to the extent that their results can be readily incorporated into a total system model. In effect, they must be familiar with that portion of the total system performance assessment model which represents the particular process or parameter that they are studying, such that they can recommend modifications in the data (or the model itself if necessary) as more information becomes available.

Without this type of integration between the performance assessment modelers and the project scientists, a tool such as RIP can not be validly applied.

1. INTRODUCTION

1.1 Background

The Office of Civilian Radioactive Waste Management (OCRWM) within the U.S. Department of Energy (DOE) has the task of siting, designing, constructing, obtaining a license for, operating, and decommissioning the nation's first high-level nuclear waste repository. The Yucca Mountain Site Characterization Project Office (YMP) is responsible for evaluating the suitability of the proposed Yucca Mountain, Nevada site to host a repository.

As part of this effort, OCRWM contracted Golder Associates Inc. (GAI) to independently develop a performance-based repository site-suitability methodology, and, based on the methodology, recommend a strategic plan of action for guiding site characterization and evaluating the suitability of the Yucca Mountain site for the development of a repository. The work is intended to complement the studies currently underway within the Project. This report presents a summary of the initial step in GAI's project, which was to develop appropriate performance assessment and strategy evaluation software.

The methodology presented herein, and any opinions expressed, are solely those of Golder Associates Inc. and do not necessarily represent positions of the Department of Energy.

1.2 Objective and Scope of Work

The primary objective of this study was to independently develop a methodology and recommend a strategic plan for guiding site characterization and evaluation of site suitability. The approach selected by GAI was based on two cornerstones:

- Development and continual updating of a preliminary *total system performance assessment model*; and

- Integration of the performance assessment model with various proposed design and site characterization activities using a *probabilistic decision analysis* approach in order to guide and prioritize the site characterization process.

The integrated repository performance assessment and strategy evaluation software developed by GAI is known as RIP (Repository Integration Program). The major portion of the software is the total system performance assessment (PA) model, which consists primarily of a series of inter-connected, coupled component models with input/output relationships for radionuclide transfer. The RIP PA model is intended to integrate existing knowledge and component-level performance models that have already been developed within the YMP or elsewhere. The performance assessment model itself is embedded within a probabilistic decision analysis model which allows the user to evaluate alternative site characterization strategies.

This report focusses only on the *theory and capabilities* of the RIP computer program. It is not a user's manual, nor does it discuss the application of the model to Yucca Mountain (or any other site). A user's manual has been prepared as a separate companion document. The development of a conceptual model of the proposed Yucca Mountain repository and the application of the RIP program to this conceptual model is presented in a third document. Both of these documents will be discussed further in Section 1.5.

In the following section, a brief overview of the RIP methodology is presented. Section 1.4 then discusses the applicability of the software. Section 1.5 summarizes related RIP documentation. Finally, Section 1.6 discusses the organization of the remainder of the report.

1.3 Overview of RIP Methodology

1.3.1 Problem Definition

Programs for selecting and evaluating the suitability of geologic sites for disposal of high level radioactive wastes all include three general components: site characterization, system design, and safety (performance) assessment. *Site characterization* activities are primarily concerned with determining the ability of the natural barrier (the geologic system itself) to provide containment and control migration of radionuclides from the repository. *System design* is concerned with the ability of the engineered barrier system (the waste package itself and the associated man-made systems) to provide containment and control migration. Obviously, site characterization and system design are not independent of each other. That is, the performance of the engineered barrier system is dependent on characteristics of the geologic system (e.g., near-field conditions), and the characteristics of the geologic system can be altered and affected by the engineered barrier system. As a result, site characterization includes activities which are intended to provide input for the system design.

In simple terms, *performance (or safety) assessment* involves the evaluation of the anticipated performance of a proposed repository with respect to radionuclide release to the environment. It ultimately depends upon both the available data (obtained during site characterization) and the proposed engineered barrier system design. (Depending on policy and regulatory constraints, however, the actual contribution of the engineered barrier system may or may not be fully included in a safety assessment.)

The task of performance assessment is complicated by the large uncertainties involved. Due to the long times scale of interest and the complexity of most geologic systems, a large degree of uncertainty exists with respect to the parameters and processes controlling such a system. As a result of these uncertainties, deterministic analysis alone is inappropriate and a probabilistic approach to performance assessment must be used. As will be described in detail in subsequent sections, RIP is a probabilistic model intended to represent these uncertainties.

1.3.2 The Role of Performance Assessment

Traditionally, performance assessment has been considered as an essential method to be implemented at the later stages in a repository siting program in order to demonstrate that the candidate site is in compliance with the licensing criteria. However, performance assessment can also be used as a tool to aid in the development and implementation of the entire siting procedure. In fact, GAI and others have long maintained that the only effective way to make valid site characterization decisions (what should be studied, when, and in what detail) is by using preliminary integrated performance assessments (e.g., GAI, 1977; NRC, 1983; GAI, 1986). This view has recently been reiterated by both the National Research Council (1990) and the Nuclear Waste Technical Review Board (1990).

This view of an ongoing performance-assessment driven characterization process is based on the premise that by evaluating system performance using *currently-available knowledge* and the *anticipated level of knowledge that can be obtained following completion of a characterization activity*, it will be possible to identify which activities will be most critical to rapid and accurate determination of site suitability. Figure 1-1 illustrates the intended process graphically.

Therefore, the RIP methodology is intended to not only provide an initial evaluation of the current YMP program, but to also provide a management tool to help assure continuing program refinement and an early resolution of site suitability. That is, RIP is intended to be used in an ongoing review process to continually reevaluate which data needs are most critical, and to reallocate program resources appropriately.

A related and equally important use of the model has become apparent to GAI in the process of eliciting information regarding models and input data from scientists and engineers investigating the Yucca Mountain site. Namely, the performance assessment model (and the process of eliciting model input) can act as a mechanism for integrating the large amount of available information generated by different groups into a meaningful whole (in a sense, allowing one to keep the "big picture" and the ultimate aims of the project clearly in focus). Such an integration would not only be useful for project managers (and is ultimately required in order to carry out a meaningful assessment of site suitability), but also to individual project scientists to help them formulate and present the results of their research in a manner such that it can be *readily applied* to the ultimate goal of the project, *predicting the performance and determining the suitability of the site.*

1.3.3 Performance Assessment Modelling Approach

At the heart of the RIP methodology is the integrated total system performance assessment model. In general terms, performance assessments can be broadly divided into two categories: those based on a "top down" approach, and those based on a "bottom up" approach. RIP is based on a "top down" approach".

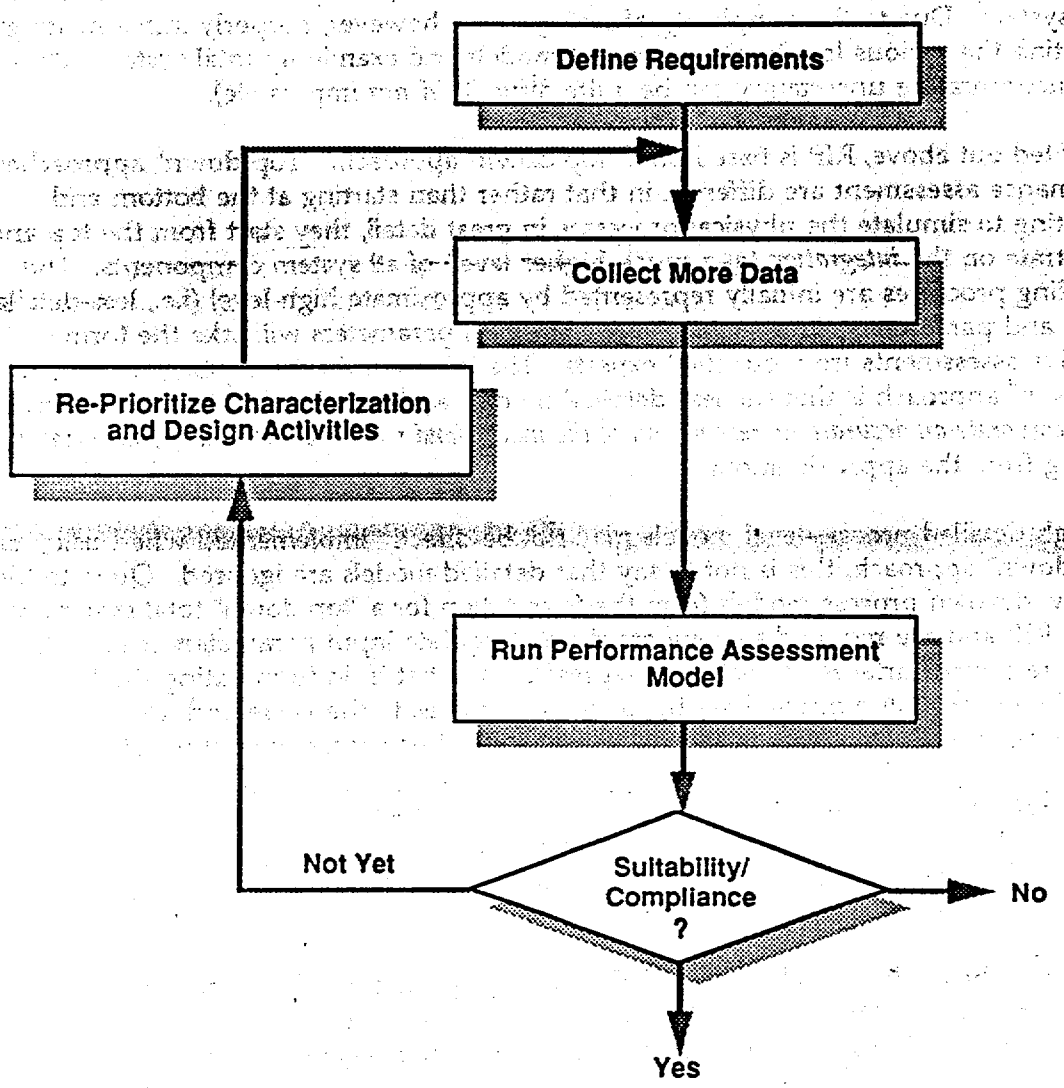


FIGURE 1-1
METHODOLOGY OVERVIEW
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DWG. NO. 26609 DATE 3/30/92 DRAWN CB

Golder Associates

"Bottom up" approaches to performance assessment attempt from the outset to model the various controlling processes in detail, and typically make use of *complex physically-based models* for the various system components. The emphasis is on understanding and explaining the lower-level processes in order to eventually describe the behavior of the entire system. Due to the complexity of real systems, however, properly implementing and integrating the various low-level component models and examining total system behavior while incorporating uncertainty can be quite difficult (if not impossible).

As pointed out above, RIP is based on a "top down" approach. "Top down" approaches to performance assessment are different in that rather than starting at the bottom and attempting to simulate the physical processes in great detail, they start from the top and concentrate on the *integration* (at a much higher level) of all system components. The controlling processes are initially represented by approximate high-level (i.e., less-detailed) models and parameters. In general, these high-level parameters will take the form of subjective assessments from qualified experts. The key point in the application of such a "top down" approach is that the less-detailed models and associated high-level parameters *must incorporate an accurate representation of the model (and associated parameter) uncertainty* resulting from the approximations.

Although detailed process-level models may not be directly implemented when using such a "top down" approach, this is not to say that detailed models are ignored. Quite to the contrary, detailed process models form the foundation for a "top down" total system model such as RIP and are required to generate the appropriate input parameters (e.g., in the form of response surfaces or analytical expressions). That is, in formulating the high level input, the expert will typically base his or her opinion as to the value and associated uncertainty of a parameter or model on available detailed process model results.

It is important to understand that a "top down" model does not have to be "simple". Whereas a "simple" model might completely *ignore* a key process, a well designed "top down" model *approximates* the process while *explicitly incorporating the resulting uncertainty* which is introducted.

A review of the current levels of knowledge regarding the controlling processes and parameters at Yucca Mountain (or for that matter, any proposed repository site anywhere in the world) readily indicates that a practical performance assessment model based completely on low-level physically based models and parameters (a "bottom up" model) is not currently feasible. That is, the uncertainty in quantifying the basic physical processes controlling waste release, groundwater flow, and radionuclide transport over the large time scales of interest at *any site* is such that a practical performance assessment model must incorporate a "top down" approach and rely on relatively high-level "lumped" descriptive parameters (e.g., radionuclide travel times, container failure rates).

Moreover, it is our view that due to computational constraints and the inherent complexity of such systems, it is unlikely that this situation will change in the foreseeable future. Although better simulation tools, more powerful computers, and additional field data will undoubtedly shed considerable light on the controlling processes, integrated total system performance models will continue to rely primarily on high-level parameters. Additional

detailed process modelling and field studies will simply act to increase the accuracy and decrease the uncertainty in the expert's subjective assessments.

As a result, the total system performance assessment model incorporated into RIP *is not* a detailed deterministic model based completely on low-level physically-based parameters (e.g., it does not explicitly simulate radionuclide transport by solving a three-dimensional advection-diffusion equation). Rather, it is a *descriptive probabilistic model* that *describes* rather than *explains* the system behavior and is intended to directly represent the uncertainties in processes and events and their controlling parameters.

As opposed to representing all processes with great detail from the outset (whether or not it is justified), the RIP methodology is based on a "top down" approach in which each model component can initially be represented at a relatively high level. RIP is designed such that the model can evolve (by adding detail to specific components) as further information becomes available. A key feature of the "top down" approach is that details are only added when it is warranted (e.g., if results indicate that performance is sensitive to a process which is currently represented in a simplified manner with a correspondingly large degree of uncertainty). That is, details are added only to those processes which are identified as being important with respect to total system performance. Such an approach can help to keep a project focussed on *total system performance* without getting lost in what may prove to be unnecessary details.

As discussed above, high-level "top down" models such as RIP rely heavily on subjective assessment (expert opinion). The process of eliciting the high-level parameters required for RIP is critical to its successful application. As will be seen below, expert elicitation also plays a key role in the RIP strategy model. Appendix A contains a summary paper outlining the methods for obtaining defensible subjective probability assessments.

In order to define the capabilities needed in a general performance assessment model such as RIP, it has been necessary to evaluate the validity of the current conceptual models (for Yucca Mountain) of subsystem performance and the current knowledge of the necessary physical parameters for these subsystem models. The knowledge base required for this evaluation lies primarily with the principal investigators involved in the YMP in addition to a number of knowledgeable third parties. Therefore, a significant part of this study has involved acquiring from these individuals the data and concepts necessary to construct the integrated model. To assist in the conceptual development of the model, GAI held a number of workshops, which were attended by scientists and engineers from both within and outside of the Yucca Mountain Project. The structure and components of the integrated total system performance assessment model incorporated into RIP are described in detail in subsequent chapters.

1.3.4 Development of an Activities Database

As discussed above, the integrated performance assessment model is intended to be used as a tool to make valid site characterization decisions and to identify which activities will be most critical to rapidly and accurately determining site suitability.

After first developing a conceptualization of the behavior of the repository system and creating an appropriate data set for the integrated performance assessment model, it is possible to apply the tool to guide site characterization. In order to accomplish this, however, it is first necessary to identify which system parameters and processes are specifically addressed by the various characterization *activities*. A characterization activity is a specific scientific study or group of studies proposed for the site (e.g., surface-based testing). A characterization *strategy* is a specified set of characterization activities.

In order to most effectively use performance assessment to guide site characterization, however, it is necessary not only to identify *which* parameters would be affected by a particular activity, but to quantify *how* those parameters would be affected. In effect, this entails quantifying how a proposed activity is likely to affect the probability distribution describing the uncertainty in a given system parameter (i.e., how much will we learn?). This requires evaluations by experts of the nature of the uncertainty in a parameter, and the extent to which it is likely to be resolved by a particular activity.

An *activities database* must be developed which includes the name of each proposed activity, a brief description, identification of the system parameters the activity will affect, a quantitative estimate of how those parameters will be affected, probabilistic estimates of cost and duration, and precedence requirements for other activities.

1.3.5 Evaluating Characterization Activities and Strategies

Having developed an activities database, the simplest way to use RIP to evaluate characterization activities is to post-process the results of a performance assessment such that the sensitivities of various system performance measures (e.g., cumulative radionuclide release to the accessible environment) to individual model parameters are identified. In simple terms, activities which provide information on model parameters to which system performance is sensitive should be given the highest priority.

Although such an approach is useful in providing a preliminary screening and evaluation of site activities, it is limited because while it qualitatively considers which parameters are affected by a particular activity, it does not quantitatively incorporate the extent to which parameter uncertainty might be reduced by a particular activity, nor the cost or duration of the activity. This information is required in order to actually *prioritize* activities into an efficient and effective strategy. For example, suppose two proposed activities are both concerned with studying a particular parameter which has been shown to have a strong influence on system performance. While we can state that these activities both merit further consideration, we can not determine whether both activities are necessary, or which would be most effective, without quantifying the extent to which each activity will reduce the uncertainty in system performance, and incorporating cost and duration considerations.

Moreover, because of the complexity of the system and the large number of inter-connected activities, actual prioritization of activities can not be based simply on an evaluation of individual activities with respect to cost, duration, and reduction in performance uncertainty. Rather, the entire site characterization plan must be considered as a whole. The proposed methodology accomplishes this by defining and evaluating

alternative characterization strategies (a characterization strategy consisting of a specific set or sequence of characterization activities). The strategy portion of RIP provides the second cornerstone of the GAI approach by quantitatively integrating the performance assessment model with the characterization activities. The strategy model is essentially a decision analysis shell around the performance assessment model which allows the user to evaluate the alternative site characterization strategies.

For any user-specified characterization strategy, RIP provides three outputs by which alternative strategies can be ranked and compared:

- a probabilistic estimate of cost;
- a probabilistic estimate of duration; and
- a probabilistic evaluation of the predicted site performance resulting from implementation of the strategy.

Computation of the first two outputs is straightforward, and consists of simply integrating within a Monte Carlo framework the cost and duration estimates for the individual activities, taking into account any precedence requirements. The third output relies on subjective assessments by experts of the extent to which model parameter uncertainty will be reduced by a particular activity. Given these assessments, along with the current state of knowledge, RIP uses a Bayesian computational algorithm to simulate how probability distributions representing parameter uncertainty will change as a result of a particular characterization strategy, and develops a probabilistic evaluation of the anticipated repository performance (i.e., probabilities of future repository performance results).

It should be noted that the degree to which such a methodology can be used to prioritize site characterization activities is inherently limited by the level of detail included in the performance assessment model. For example, it may be possible to determine, based on model results, that with respect to determining site suitability, further information regarding process A would be more beneficial than further information regarding process B. This would imply that activities which obtain information on process A should be given higher priority than those related to process B, and would be a useful finding. However, in order to further prioritize the various activities which specifically investigate the details of process A, it may be necessary to add additional detail to that particular component of the PA model.

Details of both the performance assessment model and the strategy evaluation model are presented in subsequent chapters.

1.4 Model Applicability

Before describing RIP in detail, it is useful to briefly mention the applicability of this simulation tool. As has been pointed out above, RIP was specifically developed for

application to the proposed high-level radioactive waste repository at Yucca Mountain, Nevada. *Its use, however, is not necessarily limited to this particular site.* This is because, for the most part, RIP is designed in a very flexible manner (as will be discussed in the next chapter). In simple terms, *RIP is a simulation tool designed to probabilistically model the release of radionuclides from buried waste packages and the subsequent transport of those radionuclides through the environment.* Because the physical processes which are built into the RIP algorithms are relatively fundamental and not site specific, the software could be applied to a variety of sites. As will be discussed below, the ability for the user to add detail and represent various conceptual models is considerably enhanced by the flexibility of the software. Nevertheless, in some cases, it may prove more effective to modify the RIP software directly to better represent a specific application.

1.5 Related RIP Documentation

As pointed out above, this report deals specifically with describing the theory and capabilities of the integrated performance assessment and strategy evaluation model RIP. That is, it describes the algorithms contained within the computer program. It does not specifically describe the application of RIP to the Yucca Mountain site, nor is it a user's manual. Related RIP documentation is summarized below:

- *RIP Repository Performance Assessment and Strategy Evaluation Model: User's Guide* (Golder Associates, 1992). This document describes in detail the manner in which data is entered into the computer program, as well as the form of the output. That is, it explains in detail how to apply the model to an actual problem. The RIP User's Guide is the companion document to the Theory and Capabilities manual and is referenced throughout the present document.

- *RIP Verification Report.* This document describes the verification test problems which were carried out in order to demonstrate that the major algorithms of the software (discussed in the Theory and Capabilities document and the User's Guide) are properly implemented.

- *Analysis of Radionuclide Release to the Accessible Environment from a Nuclear Waste Geologic Repository at Yucca Mountain, Nevada.* As mentioned previously, during development of the software, it was necessary to concurrently develop a conceptual model for the behavior of the Yucca Mountain system in order to ensure that RIP included the important processes and events which may be active at the site. This report presents this conceptual model, identifies (and supports) the appropriate model input parameters, and discusses the application of the RIP performance assessment model to the Yucca Mountain site. A series of sensitivity analyses were also carried out to provide a preliminary identification of sensitive parameters.

- *Evaluation of Alternative Site Characterization Strategies at Yucca Mountain, Nevada: Demonstration of the RIP Methodology.* The formal methodology for

incorporating RIP into a procedure for evaluating alternative strategies and guiding site characterization at Yucca Mountain is demonstrated in this document.

1.6 Report Organization

Chapter 2 discusses the basic concepts of the RIP performance assessment and strategy evaluation model. This provides a summary description of the various features and components of RIP. Readers interested only in the basic features and concepts of the model may wish to read only Chapter 2. Chapters 3, 4 and 5 discuss the details of the three major components of the RIP performance assessment model: waste package behavior and release, radionuclide transport pathways, and disruptive events. Chapter 6 describes the details of the strategy evaluation model. References are listed in Chapter 7.

The report also includes a number of appendices. These appendices describe additional details pertaining to the RIP model and are referenced in the main text.

2. BASIC CONCEPTS OF THE REPOSITORY INTEGRATION PROGRAM (RIP)

Having introduced in very general terms the overall philosophy and methodology upon which RIP is based, in this chapter a more detailed description of the basic concepts of the RIP software are presented. Section 2.1 first discusses some terminology which is used throughout the report. Section 2.2 then describes the structure and fundamentals of the performance assessment model, including an overview of each of the three major model components. The structure and fundamentals of the strategy evaluation model are discussed in Section 2.3. Section 2.4 discusses the flexibility of the RIP software, which imparts tremendous power to the model. Finally, Section 2.5 provides a summary of the key features of RIP.

2.1 Terminology

The experienced reader will note that the term "model" is used somewhat loosely in this report. Because modelling and simulation terminology is generally not used uniformly throughout the engineering and scientific community, it is important to clearly define at the outset our set of terminology. In particular, it is important to differentiate between what we refer to as the *computational model*, the *conceptual model*, and the *simulation model*.

As used here, the *computational model* is the tool which provides the mathematical and computational framework for modeling a system of processes. It is more accurately referred to as the *computational algorithm* or *simulator*. It does not specifically describe an actual physical system. Rather, it simply consists of a set of flexible tools (i.e., a group of algorithms) for doing so. The computational model is embodied within software (i.e., a computer code such as RIP). The user of the software must provide the proper input data in order to simulate an actual physical system.

A *conceptual model* is a representation of the physical and chemical properties and processes controlling behavior of a physical system. That is, a conceptual model is essentially a body of ideas, based on available data, which summarizes the current understanding of a system. The conceptual model not only includes a description and quantification of the controlling processes and parameters, but also quantifies the uncertainties involved. (In fact, due to large uncertainties involved in geological systems, a number of *alternative conceptual models* are likely to be developed.)

The *simulation model* is the actual implementation of the computational model for a given conceptual model. That is, the simulation model couples the tools within the computational model (the actual computer code) with a particular conceptual model (represented by a distinct set of input data for the computer code).

Summarizing in general terms, the *computational model* is the mathematical framework embodied in the actual computer code. A *conceptual model* represents the current understanding of the physical system to be modeled. A *simulation model* is the application of the computer code using a specific data set in order to simulate the behavior of the system described by the conceptual model.

In this document, the term "model" used by itself refers only to the *computational model.* Hence, when we speak of the "radionuclide transport component model" or the "waste package component model", we are referring to the computational algorithms describing radionuclide transport pathways or waste package behavior, respectively (a component model being a subset or part of the total computational model). The complete terms "*conceptual model*" and "*simulation model*", on the other hand, are always used explicitly when referring to these two concepts.

2.2 Total System Performance Assessment Model Overview

The major portion of the RIP software is the *total system performance assessment (PA) model.* The PA model is essentially a radionuclide transport model, and consists of a series of interconnected, fully-coupled component models. The component models consist, in general, of simple functions relating various system parameters which control processes affecting radionuclide transport. The three primary component models address:

- waste package behavior and radionuclide release;
- transport pathways to the accessible environment; and
- ancillary processes (disruptive events) which may affect system parameters, such as human intrusion, tectonics and volcanism.

A schematic of the main components of the RIP PA model is shown in Figure 2-1.

Before describing in greater detail what the RIP PA model *is*, it is important to reemphasize what it *is not*. RIP *is not* a detailed deterministic model based completely on low-level physically-based parameters. (For example, it does not explicitly model radionuclide transport through the unsaturated and saturated zone by solving a three-dimensional advection-diffusion equation.) As pointed out in the previous chapter, the RIP PA model is a *descriptive probabilistic model* that *describes* rather than *explains* the system behavior and is intended to directly represent the uncertainties in processes and events and their controlling parameters. As a result, RIP relies heavily on subjective assessments of relatively high-level descriptive parameters. It is also important to note that RIP PA model is a *preliminary* performance assessment model. It is intended to be used as a management tool to help guide site characterization, as well as a tool for project scientists who wish to carry out sensitivity analyses and/or test alternative hypotheses with respect to system performance. In its present form, it is not intended to be a licensing tool. It should be noted, however, that the structure of RIP is such that it can continually evolve, and, as more details are incorporated into the model, RIP could potentially evolve into a licensing tool.

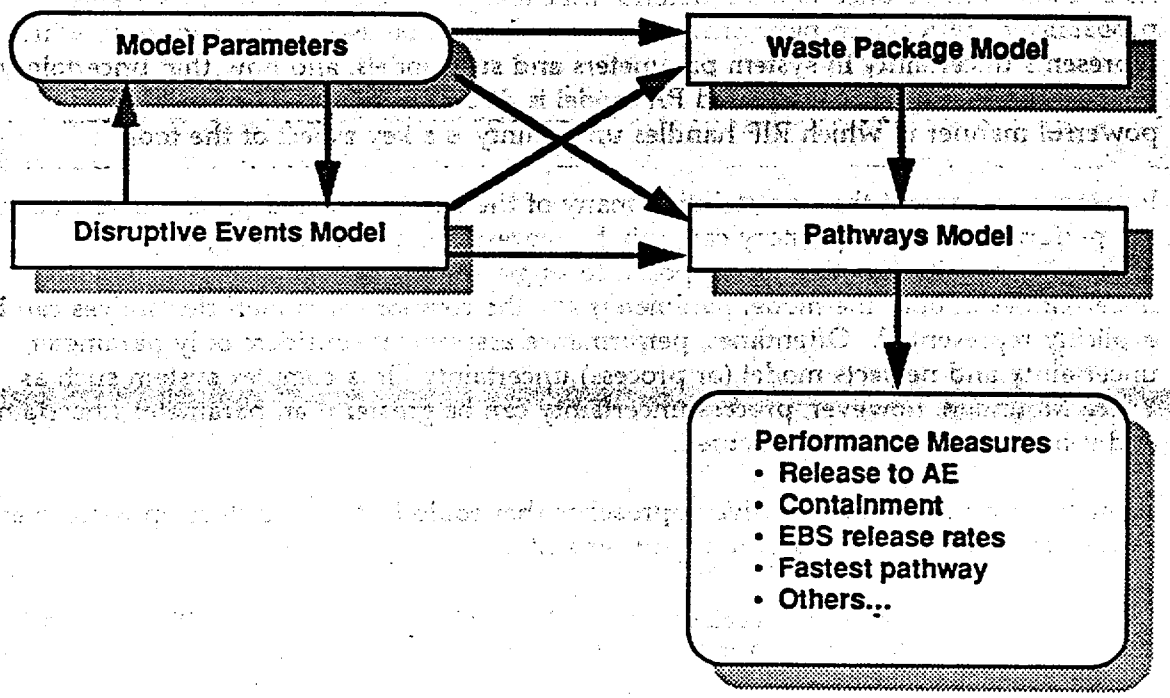


FIGURE 2-1
RIP PERFORMANCE ASSESSMENT MODEL COMPONENTS
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DWG. NO. 27823 DATE 2/19/92 DRAWN TB

Golder Associates

2.2.1 Treatment and Propogation of Uncertainty

The purpose of the RIP PA model is to predict the future, allowing for uncertainties in our knowledge about the system being modelled, including uncertainties about the processes and future events that may occur, and uncertainties in the parameters controlling those processes and events.

As a result of these uncertainties, deterministic analysis alone is inappropriate and a probabilistic approach to performance assessment must be used. The manner in which RIP represents uncertainty in system parameters and sub-models, and how that uncertainty is propogated through the integrated PA model is discussed below. The flexible and powerful manner in which RIP handles uncertainty is a key aspect of the tool.

In order to represent the uncertainties, many of the parameters and processes related to the performance of a repository can only be represented probabilistically (e.g., by probability density functions, as opposed to single values). RIP is designed such that uncertainties in both the model parameters *and* the component models themselves can be explicitly represented. Oftentimes, performance assessment considers only parameter uncertainty and neglects model (or process) uncertainty. In a complex system such as Yucca Mountain, however, process uncertainty can be greater than parameter uncertainty and it must be properly represented.

There are a number of alternative approaches that could be taken to develop performance assessment predictions incorporating uncertainty:

- evaluation of the base case (a representation of the expected or most likely future for the repository system), followed by sensitivity evaluation of the ways in which the base case results would vary as a function of the uncertain parameters (eg., "if the climate gets wetter by X%, what will happen to the system?");

- evaluation of the base case plus a series of scenarios (i.e., alternative futures), which represent significant variants from the base case due to specific events or processes that may occur (eg., different degrees of water-table rise, alternative types of human intrusion, alternative climates, etc.);

- sampling and simulation of all possible futures randomly and repeatedly using the Monte Carlo method (i.e., repeatedly simulating the time history of repository system behavior, each time sampling the various parameter distributions describing system processes and events).

There are many variants and hybrids of the above approaches, and Appendix B compares and contrasts scenario-based approaches and simulation-based approaches.

The RIP model uses the *simulation approach*, utilizing the *Monte Carlo method* to sample the probability distributions for the uncertain parameters (describing both processes and events) and simulate a large number of system realizations in order to determine

probability distributions of site performance (e.g., cumulative release, risk, transport time). That is, RIP creates a time history of disruptive events and other system parameters for each system realization, simulates the behavior of the system under those conditions, and then combines the results of all the realizations in an appropriate manner to determine probability distributions of site performance. Because model parameters are described stochastically, each realization produces a different time history of events and processes. The integrated model is designed such that it can simulate all combinations of model parameters and time histories which might be realized.

An enhanced sampling scheme using both *importance-sampling* and *Latin Hypercube sampling* was developed in order make the Monte Carlo method more efficient by improving its ability to resolve low-probability, high-consequence areas of the site performance pdf's. This scheme is discussed again in Chapter 5 and is described in some detail in Appendix C.

In general terms, the outputs from the performance assessment model consist of probabilistic performance measures for the repository system (e.g., cumulative radionuclide release to the accessible environment, maximum annual release from the waste packages). That is, outputs are not single values, but distributions which specify the probability of exceedence for any particular value of the performance measures.

Figure 2-2 shows the overall logic of the RIP program, illustrating the Monte Carlo simulation approach.

2.2.2 RIP Performance Assessment Component Models

The integrated model consists of a series of fully-coupled component models. These component models consist, in general, of simple functions directly relating various system parameters in order to describe processes affecting radionuclide transport. Overall consistency is maintained by ensuring that in a given realization a parameter has a single value for all of the component models that depend on it. Correlations between sets of parameters are also incorporated. As discussed above, the integrated model consists, in simple terms, of three major component models:

- a model that defines and describes the performance of the waste package system (failure and radionuclide release);
- a model that defines and describes the various radionuclide transport pathways from the waste package to the accessible environment; and
- a model which describes disruptive events which can directly or indirectly affect waste package performance and/or transport pathways (e.g., volcanism, human intrusion, seismic activity).

A schematic of the major components of the RIP PA model was presented in Figure 2-1.

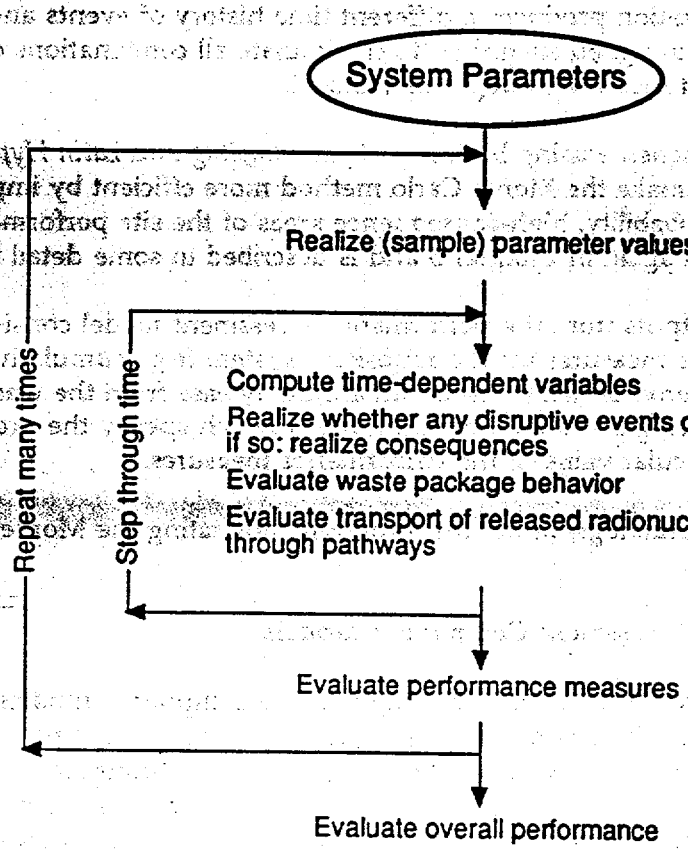


FIGURE 2
RIP PERFORMANCE ASSESSMENT
MODEL SIMULATION LOGIC
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 27825 DATE 2/19/92 DRAWN BY TB

Golder Associates

The PA model incorporates radioactive decay and production of daughter products. The algorithm used to implement this is discussed in Appendix D. Additional details regarding this process are presented in the RIP User's Guide.

The *waste package behavior and radionuclide release component model* considers: 1) breaching of the waste containment system (including both container failure and dissolution and/or alteration of the waste matrix) which exposes the waste; and 2) the subsequent mass transfer of the radionuclides present within the containers to the geologic environment.

The output for the waste package component model provides input for the *transport pathways component model*, which considers transport of the radionuclides through the geologic environment and eventual discharge to the accessible environment (e.g., the ground surface and/or the saturated zone at some specified distance from the repository).

The *disruptive events component model* considers the effects of discrete pertubations (e.g, earthquakes, volcanism, human intrusion) on the behavior of the other two PA model components. That is, disruptive events could potentially change the behavior of the containment system (e.g., waste packages could be instantaneously disrupted or moved), or modify the behavior of the geologic transport pathways (e.g, a new transport pathway could be created or the properties of an existing pathway could be modified).

The RIP structure for each component model was designed from the top down, starting with a broad description of processes of interest. The software is designed such that the user can increase the amount of detail and complexity in nearly any portion of the component models. That is, through a method of encapsulation, the component models can themselves be made up of sub-components.

For example, the waste package model contains sub-models for corrosion of the outer and inner barriers, dissolution of the waste matrix, and radionuclide transport in the near field. The user might elect to define subsidiary lower-level models to support the sub-models. For example, the user could define a model of the chemical evolution of the near-field environment. The encapsulation process may proceed to an arbitrary depth, but is limited in complexity by the necessity for the calculations to be extremely rapid (to facilitate efficient use of the Monte Carlo method).

Sub-models range from simple analytical expressions to more complex numerical subroutines. Some sub-models (and parameters) are time-dependent. That is, as shown in Figure 2-2, the integrated model essentially "time-steps" through the simulations, taking into account the time-dependent parameters affecting the waste package source term and other system parameters, as well as radionuclide decay chains. For example, corrosion processes could have a sensitivity to the time-varying thermal, chemical and moisture conditions near the waste packages.

In the interest of computational efficiency (which is necessary for probabilistic analysis), the component models built into RIP were greatly simplified compared to state-of-the-art process models. Where appropriate, the increased uncertainty caused by such simplification can be reflected in parameters describing *model uncertainty*. This explicit

representation of model uncertainty is important for two reasons: 1) it provides a more accurate representation of the true overall uncertainty; and 2) it allows evaluation of the potential benefits of development of better models.

It is important to emphasize that although it is currently infeasible to use detailed low level models directly for probabilistic analysis, it is possible to use these models indirectly. That is, it is intended that the high-level parameters for the component models within RIP should be based, when possible, on the results of these detailed models. For example, a detailed model could be run externally a number of times to create a response surface or an analytical expression which could then be used to describe the dependencies of a high-level parameter within RIP.

The basic concepts of the three major components models (waste package behavior and radionuclide release, radionuclide transport pathways, and disruptive events) are summarized in the following sections. These components are discussed in detail in Chapters 3, 4, and 5.

2.2.2.1 Waste Package Behavior and Radionuclide Release Component Model

The waste package model is described in detail in Chapter 3. A brief summary of its capabilities is provided in outline form below.

General Model Structure

- The waste package component model is part of the total system performance model, and is directly coupled to component models describing disruptive events and radionuclide transport pathways.
- The model is based on subjective assessments of relatively high level phenomenological parameters, such as container and cladding/pour canister failure distributions, and generalized alteration and mass transfer parameters.
- The model simulates groups of waste packages (rather than tracking the behavior of each of the individual waste packages in the repository).
- Waste package parameters can be described by the user as a function of temporally varying environmental conditions.
- The model incorporates radioactive decay and production of daughter products.

Environmental Conditions

- Temperature and moisture conditions are explicitly included in the model. The temperature at the edge of the waste package varies with time. Moisture conditions (intended primarily for unsaturated repositories) refer to the mode

of water contact at a waste package (e.g., "wet-drip"). Both temperature and moisture conditions can be specified as variable across the repository.

- Other environmental conditions (e.g., chemistry, stress) can be added by the user. Environmental conditions can be described as being spatially and temporally variable.

- A simple rewetting model is used which assumes that waste packages return to specified moisture conditions upon reaching a specified temperature. At temperatures above the specified temperature, the waste package is considered to be dry.

Waste Package Failure Rates

- Two levels of waste package containment (and failure) are explicitly simulated: the primary container (the waste package itself); and the secondary container (e.g., cladding for spent fuel, pour canisters for high level defense waste).

- Waste package failures are described in terms of density functions of failure frequency.

- Containers and cladding/pour canisters can fail by one or more failure modes. The model combines failure modes by assuming that the failures can be treated independently (they do not act with synergism).

- Container failure modes can be affected by temporally varying environmental conditions.

Mass Exposure

- Exposure of radionuclides is brought about by container and cladding/pour canister failure, as well as matrix alteration/dissolution processes.

- The radionuclide inventory is made up of three additive components: the free inventory, which is exposed immediately upon primary container failure, the gap inventory, which is exposed immediately upon secondary container (e.g., cladding) failure in a failed container, and the bound inventory, whose exposure is controlled by alteration/dissolution of the waste matrix.

- Exposure of the bound inventory is controlled by two types of processes: dissolution of the matrix, and air alteration of the matrix. Dissolution is described in terms of a matrix dissolution rate and an effective wetted surface area. Dissolution is zero prior to rewetting of the waste package. Air alteration rates are specified directly by the user.

- The effective wetted surface area of matrix can be described as a function of container failure mode. It is assumed that for a waste package that fails by more than one mode, the effects on wetted surface area are additive.

Mass Transfer

- For an aqueous radionuclide, mass transfer can be described as an advectively or a diffusively controlled process, each of which may be limited by radionuclide solubilities.
- Aqueous mass transfer is set equal to zero prior to rewetting of the waste package.
- Once mass transfer commences, it can not be described as a function of the time of container failure.
- Several mass transfer parameters can be described as functions of container failure mode. It is assumed that for a waste package that fails by more than one mode, the effects on mass transfer are additive.
- The mass transfer rate for gaseous species is directly specified by the user.

Effect of Disruptive Events

- Disruptive events can affect the behavior of the waste package. Disruptive events are simulated in a separate model component. Their consequences can manifest themselves in the waste package model in three ways: 1) a portion of the waste packages can be disrupted in place; 2) a portion of the waste packages (and their inventory) can be moved directly to the accessible environment or some other location; and 3) the parameters describing waste package behavior and/or environmental conditions may be changed.

Linkage to Transport Pathways Model

- The waste package model is coupled with the transport pathways model in two ways: 1) it relies upon the same large scale hydrologic parameters utilized and/or defined within the transport pathways component (e.g. repository level infiltration rate); and 2) each waste package type discharges its mass to a specified transport pathway (or pathways) defined by the user.

2.2.2.2 Radionuclide Transport Pathways Component Model

The pathways model is described in detail in Chapter 4. A brief summary of its capabilities is presented below.

The task faced by the RIP pathways model is to probabilistically simulate radionuclide transport through the geosphere. Knowing that continuum-based models cannot be

practically used for this task, the RIP model algorithm was designed in a different way. The near and far field radionuclide transport algorithm for the RIP model is based on a network of user defined pathways. The pathways of RIP reflect major features of the hydrologic system and are conduits through which transport occurs. In a RIP application, the user sets up a pathway "network" relevant to what is known about the hydrologic system. The pathways may be used for both flow balance and radionuclide transport purposes. They may account for either gas or liquid phase transport from the repository level to the accessible environment.

The purpose of a pathway is to represent a homogeneous region within the large scale heterogeneity of the hydrologic system, such as geologic structures, and formation scale stratigraphy. In a simple homogeneous system, only a single pathway may be necessary. In highly complex systems, such as Yucca Mountain, up to 30 pathways are more likely required in order to account for the large scale heterogeneity.

The RIP pathways are characterized by flow modes, which address heterogeneity at the local scale. For example, flow in rock matrix and flow in fractures may be two interacting flow modes within a single pathway. The flow modes are primarily distinguished from one another based on flow velocity in the mode. However, retardation parameters and the proportion of the total pathway flow may also be different from one flow mode to another.

The actual transport of radionuclides along a pathway is based on a breakthrough curve. In RIP, the breakthrough curve is developed as a cumulative probability distribution for radionuclide travel times along the pathway. The probability gives the expected proportion of radionuclides which would have traversed the pathway by a specified time, which is equivalent to a breakthrough curve. The breakthrough curve for a pathway combines the effects of all flow modes and retardation on the radionuclide travel time. It is developed based on a Markov process algorithm, which is similar to a random walk through the flow modes.

The RIP pathways model incorporates a number of simplifying approximations made in the interest of reducing computer run times. Nevertheless, it is a very flexible and adequately accurate approach for stochastic modelling.

As can be seen even from the brief description provided above, the RIP pathways model is significantly different from other models that have been applied to groundwater flow and radionuclide transport at Yucca Mountain. The reason for this is the authors' observation that available models, while representing portions of the overall system well, *fail to capture the essence of the overall hydrogeologic system.* One-dimensional models cannot capture effects such as fingering, perching, and diversion to faults. Even three-dimensional models do not capture all the important aspects of the fracture system, and do not have adequate spatial resolution to accurately simulate small-scale behavior such as fracture termination and matrix-block imbibition. None of the available models adequately represents coupled water/vapor/air flow within the mountain. The absence of adequate models is revealed most tellingly in the following observation: *there is no valid water-balance model which explains the flow of water at the Yucca Mountain site, and predicts results consistent with observed field data.*

Most of the deficiencies noted above are due to two factors: 1) the unsaturated zone hydrogeology at Yucca Mountain is complex, and 2) DOE has been unable to collect much data. While the second of these factors should be resolved in due course, the RIP authors have to develop a model capable of representing what is known about the site *at the present time*. Thus, the RIP model is based on a phenomenological approach which attempts to *describe* rather than *explain* the system.

While the YMP scientists may not yet have adequate data and models to develop a full mathematical model of the site, they do have a good understanding of the kinds of processes that can occur there, both now and in the future. The RIP pathways model was developed to account for these processes. While many of the necessary input data will currently be little more than educated guesses based on limited data and physical bounds, the results of the Site Characterization Program and more advanced models should allow the input data to be significantly improved in the future. As has been pointed out, the primary purpose for developing RIP is to evaluate what types of data and models will be most valuable.

2.2.2.3 Disruptive Events Component Model

The disruptive events model is described in detail in Chapter 5. A brief summary of its capabilities is presented below.

Disruptive events are defined here as discrete perturbations of the repository system. That is, disruptive events are discrete occurrences which have some quantifiable effect on the processes described by the other two component models. Note that discrete is a relative term, and does not necessarily imply instantaneous. Given the long time scales of interest (10,000 years or more), something taking place over a period of 100 years could be considered a discrete event. Examples of disruptive events under this definition include volcanism, faulting, and human intrusion. In general, the disruptive events component model is intended to represent relatively rare occurrences. Events that occur continuously throughout the time period of interest are more efficiently modeled as processes within the other two component models.

Climate change is not treated as a disruptive event, since this is a process which occurs gradually, and some sort of climate change is expected in the future (although the actual nature of the change is uncertain). Climate change (and the corresponding changes in environmental conditions such as water table elevation and infiltration rate) can be treated explicitly in the transport pathways and waste package behavior components.

The *occurrence* of disruptive events is described as follows:

- RIP requires input by the user describing all *significant events*: events that are both credible and consequential.

- Each significant event must be assigned an annual rate of occurrence, λ. It is also necessary to specify whether an event can reoccur. In addition, each disruptive event can be assigned one or more *descriptor parameters*. Descriptor

parameters define the characteristics and magnitude of the event. For example, the descriptor parameters for a human intrusion/drilling disruptive event might be the number of boreholes drilled, and the deepest pathway (i.e., formation) intersected.

- Event occurrences are simulated as Poisson processes.

For each disruptive event, there are four types of internally-defined disruptive event *consequences* which describe possible discrete responses of the system:

- The radionuclides in a number of waste packages are moved directly to the accessible environment. It is assumed that the inventory released from each package is equal to the inventory of an unfailed container at the time of the event.

- The radionuclides in a number of waste packages are moved directly to a specified pathway (e.g., the saturated zone). It is assumed that the inventory released from each package is equal to the inventory of an unfailed container at the time of the event. It is also assumed that all of the waste package's inventory is immediately released to the pathway (i.e., the waste package is completely disrupted during the movement and the contents are not limited by any alteration, dissolution, or mass transfer processes at the waste package).

- A number of waste packages are disrupted in place. It is assumed that the cladding (or pour canister) instantaneously fails with the container.

- A portion of the mass (previously released from the waste packages) contained within a pathway is immediately discharged to the accessible environment.

In addition to the four *discrete consequences* described above, which are explicitly included in the model, it is also possible to directly influence parameters defined in the waste package and transport pathway modules, and this capability can be used to specify *long-term consequences*.

2.2.3 Performance Assessment Model Input and Output

RIP is designed with an interactive menu-driven user-interface that allows the user to easily enter input parameters. The specific types of input parameters required by each component model are discussed in detail in the following chapters. The RIP User's Guide describes the user-interface in detail. Section 2.4 discusses in general terms the features of the user-interface which allow enormous flexibility in specifying input parameters.

The RIP user-interface is also designed to allow the user to specify and manipulate the output and present it in a variety of forms. The specific forms of the output are discussed in general terms in subsequent sections of this document and in detail in the RIP User's Guide.

As discussed above, in general terms, the output for RIP consists of performance measures for the repository system. These performance measures are probabilistic in nature. That is, output is not a single value, but a distribution which specifies the probability of exceedence for any particular value for the performance measure.

It is worthwhile to briefly discuss here the ways in which the probabilistic results generated by RIP are presented. As an example, we will consider one of the major probabilistic performance measures output by RIP: the (normalized) cumulative release of radionuclides to the accessible environment (e.g., over a 10,000 year period).

Recall that the Monte Carlo method utilized by RIP will essentially simulate a large number of system realizations. Each realization will produce a single value for cumulative release. Each individual result will be weighted in a manner which reflects how the realizations were sampled. For example, if realizations are sampled in a totally random manner, each result will have an equal weight; if realizations were sampled in a biased manner (as described in Appendix C), results would have different weights.

The simplest way in which to display these results would be in the form of a probability density function (pdf). In simple terms, this plots the frequency of the various cumulative release results, and is illustrated schematically in Figure 2-3a. A probability density function is seldom suitable for presenting Monte Carlo results due to the discrete form of the results, which produces "jagged" probability densities.

An alternative manner of presenting the same information is the cumulative distribution function (cdf). This is formed by simply integrating over the pdf and is illustrated schematically in Figure 2-3b. By definition, the total area under the pdf must integrate to 1.0, and the cdf therefore ranges from 0.0 to 1.0. As shown in the figure, a particular point, say $[r_i, p_1]$, on the cdf is interpreted as follows: p_1 = the probability that the cumulative release is *less* than r_i.

The final (and most common) manner of presenting this information is the complementary cumulative distribution function (ccdf). The ccdf is illustrated schematically in Figure 2-3c. As shown in the figure, a particular point, say $[r_i, p_2]$, on the ccdf is interpreted as follows: p_2 = the probability that the cumulative release is *greater* than r_i. Note that the ccdf is the complement of the cdf. That is, $p_2 = 1 - p_1$.

In addition to presenting probabilistic results in the three manners outlined above, RIP can also manipulate the output data in order to carry out detailed *sensitivity analyses*. Note that for each individual realization, RIP not only saves the performance result, but also saves all of the input parameters. This allows the user to carry out a variety of post-processing exercises to examine the sensitivity of the performance results to specific model parameters or groups of parameters.

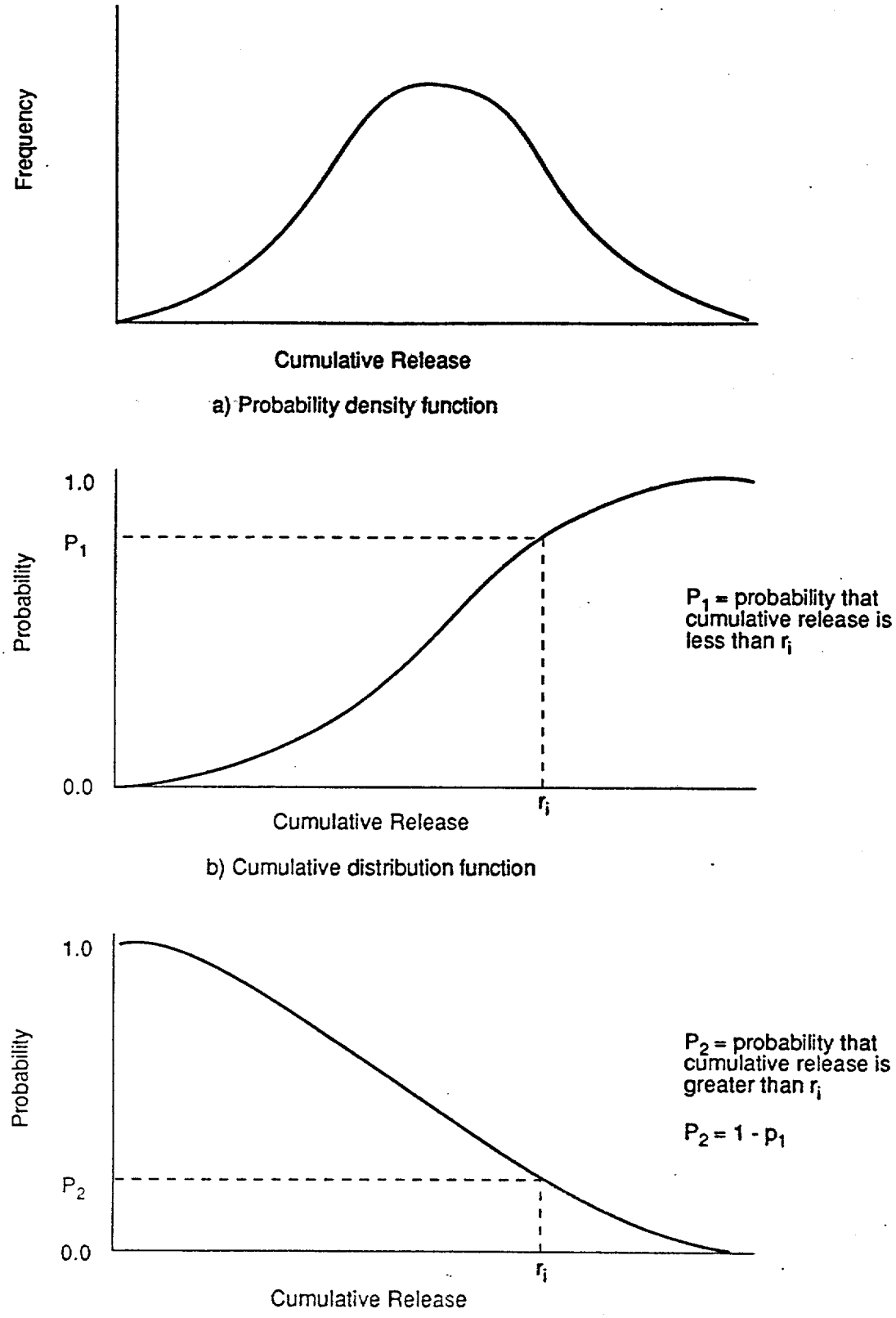


FIGURE **2-3**
PRESENTATION OF PROBABILISTIC RESULTS
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 26694 DATE 10/8/91 DRAWN BY EA

Golder Associates

2.2.4 Results Post-Processing

A RIP analysis can typically involve hundreds of independent stochastic parameters, and numerous performance measures may be calculated. The results of several thousand realizations of such an analysis would be very difficult to interpret without good analytical tools. Therefore, RIP has a post-processing module which assists the user in making sense of the results, and identifying the most significant parameters and the way in which they affect the system performance.

The result post-processing module in RIP allows the user to review and analyze the results of the Monte Carlo simulations. The post-processing is primarily graphical or statistical analyses of how the system results are affected by different input parameters.

The user takes the following three steps in post-processing:

- he selects a subset of the results (e.g., all results, or the worst 5% of results, etc.)
- he selects which result he is interested in (e.g., a time-history of radionuclide release, or a waste package failure rate, etc),
- he selects the type of analysis to perform, and the parameters which he wishes to analyze.

There are a number of options for each of the above steps. For example, the user can select a subset of results based on how the system performed (e.g., the 'bad' results), or on the value of some input parameter (e.g., all results involving human intrusion), or both. Similarly, a variety of analyses may be performed, including the following:

- direct display of the inputs and results for the Monte Carlo realizations;
- a display of correlations between results and input parameters;
- a scatter plot of the result versus an input parameter;
- a 3-d plot of the result as a function of two input parameters;
- a display of time-histories of selected results from particular 'bad' realizations.

These post-processing functions are intended to assist the user in developing an understanding of which combinations or ranges of input parameters result in better or worse performance of the repository system.

An additional method of post-processing analysis, strategy modeling, is described in Chapter 6. Strategy analysis allows the RIP user to evaluate the likely results of new information (from site characterization activities) in terms of changes in the confidence of performance predictions.

The results post-processor for RIP is discussed in detail in the RIP User's Guide.

2.3 Strategy Evaluation Model Overview

As described previously, the performance assessment model is embedded within a decision analysis model which allows the user to evaluate alternative site characterization strategies, where a strategy is defined as a group of activities. Figure 2-4 provides a schematic of the RIP strategy evaluation model. As discussed in Chapter 1, the RIP strategy model provides three outputs by which alternative strategies can be evaluated:

- a probabilistic estimate of cost;
- a probabilistic estimate of duration; and
- a probabilistic evaluation of the predicted site performance resulting from implementation of the strategy.

Computation of the first two (probabilistic cost and duration) is straightforward. Computation of the third, the probabilistic reduction in performance assessment uncertainty, is relatively novel and complex and is discussed briefly here.

Given a current data set of system parameters, and information describing how parameter uncertainty will be reduced by a given characterization strategy, RIP realizes a set of *post-strategy system parameter distributions*. That is, the current data set (in particular, the uncertainty in specified system parameters) is modified to reflect the effects of carrying out the activities within a characterization strategy. This modification is carried out using a Bayesian updating algorithm. Because we can not predict precisely how a parameter's distribution will change due to an activity, but can only estimate how the *magnitude of the uncertainty* will be affected (i.e., the distribution shape may narrow, but we can't predict whether it will shift up or down), the results of a particular strategy can only be described probabilistically. Figure 2-5 shows how the probability distribution for one particular parameter might change as a result of carrying out a strategy. (In this example, it has been assumed that the characterization strategy will reduce the parameter uncertainty by a factor of three.) In effect, a particular strategy defines a *probability distribution* of *sets of resulting system parameter distributions*. The RIP strategy model samples this distribution of sets many times.

Based on the parameter distributions within a given realized set, RIP uses the performance model to produce a probabilistic evaluation of repository performance measures. As illustrated in Figure 2-6, each realized set of parameter distributions produces unique probabilistic performance results (e.g., in the form of a CCDF of cumulative release). By realizing many different sets of parameter distributions for a particular strategy, producing a probabilistic performance result for each, and combining this suite of results in an appropriate manner, it is possible to determine probabilistically how that particular strategy is likely to affect the evaluation of site suitability.

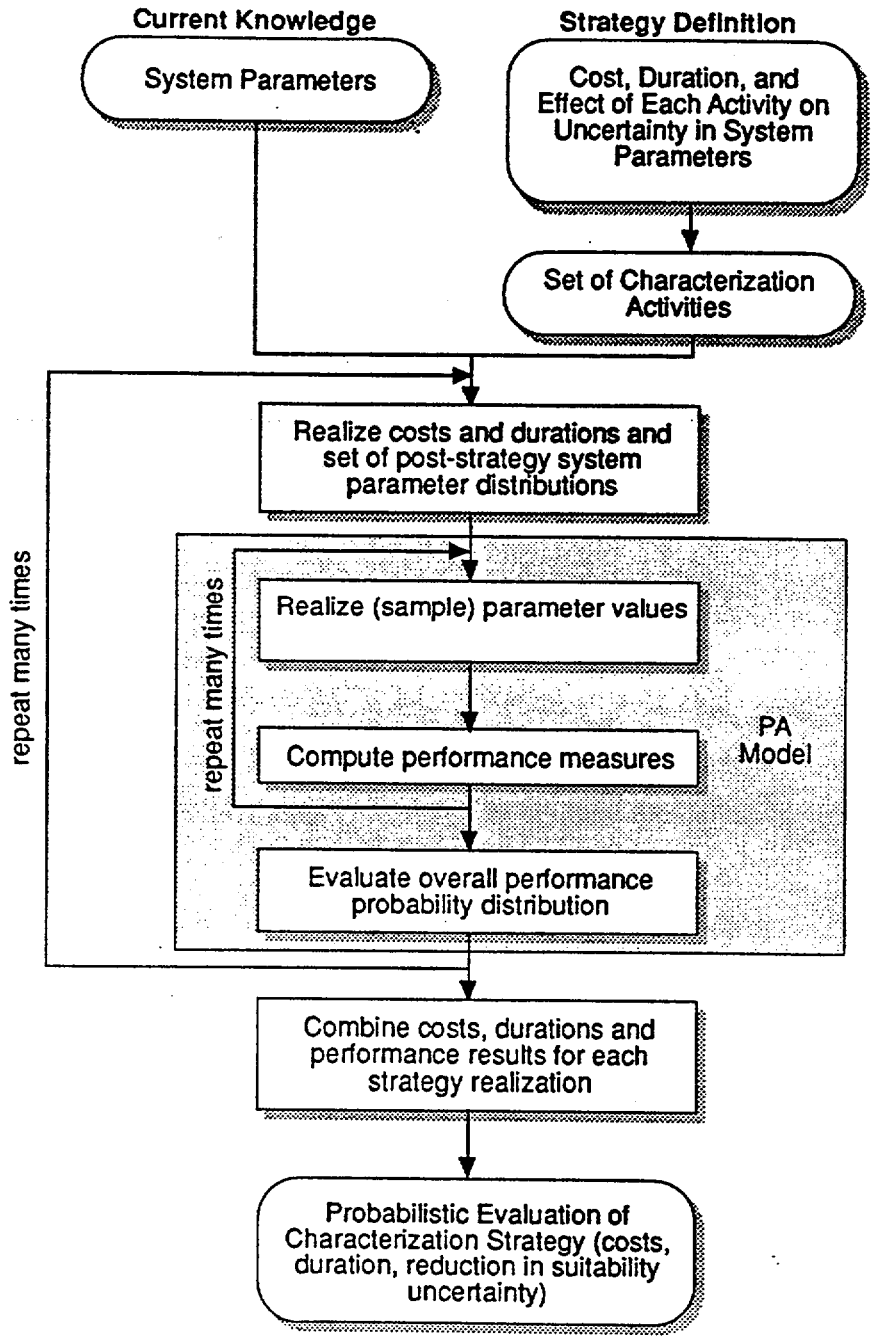


FIGURE 2
OVERVIEW OF RIP STRATEGY EVALUATION
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 27828 DATE 2/19/92 DRAWN BY TB

Golder Associates

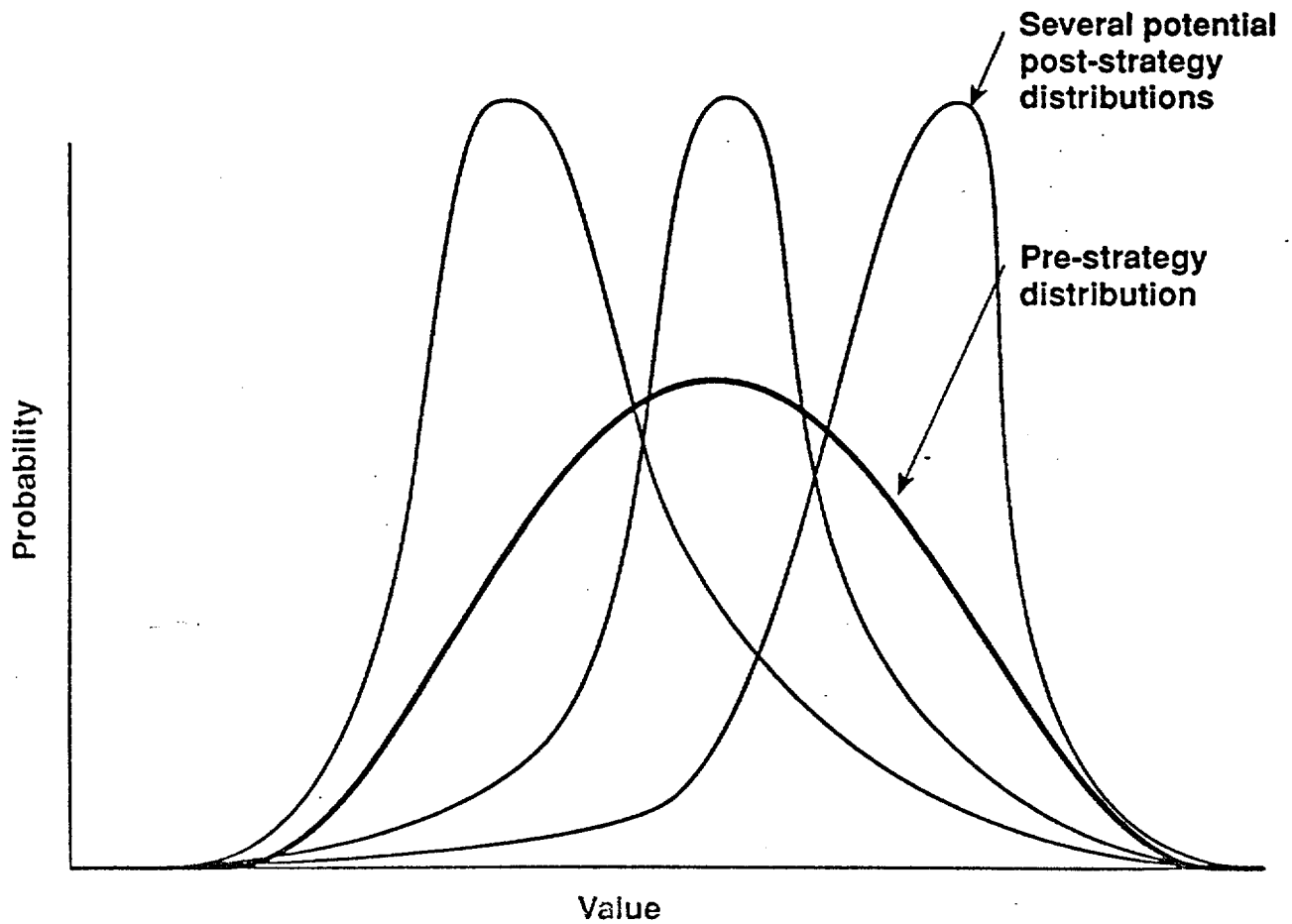


FIGURE 2-5
UPDATING THE PROBABILITY DISTRIBUTION FOR A SINGLE SYSTEM PARAMETER
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 27832 DATE 2/19/92 DRAWN BY TB

Golder Associates

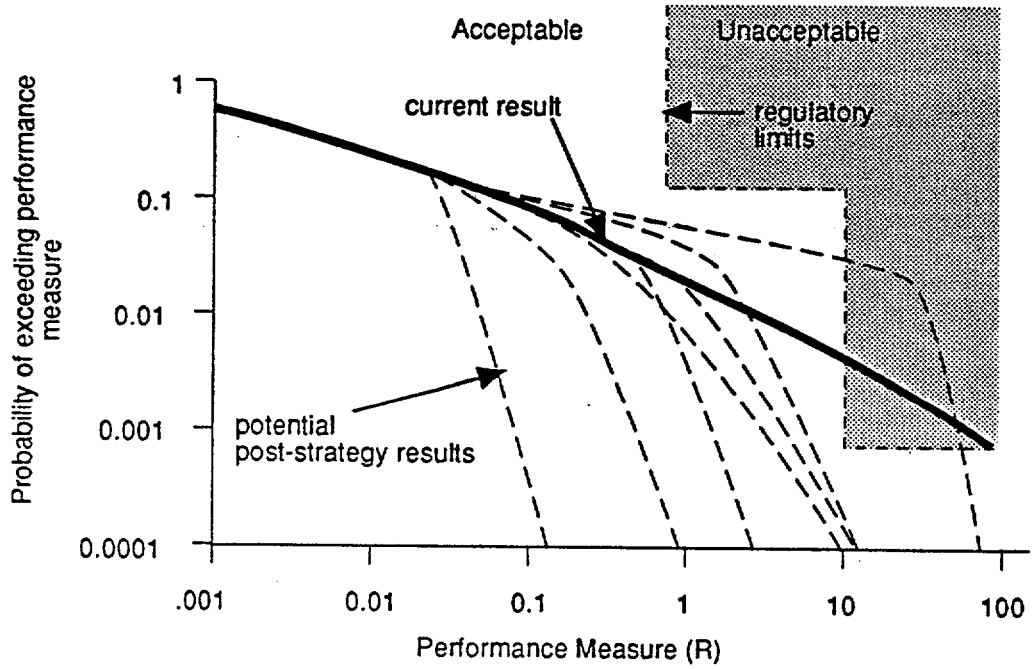


FIGURE 2
SUITE OF PERFORMAN
RESULTS FOR STRATEGY A
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 27830 DATE 4/6/92 DRAWN BY TB

Golder Associates

In practice, the algorithm used in RIP is somewhat simpler than that indicated in Figure 2-4. It is not actually necessary to re-run the performance model each time the updated probability distributions are produced by simulating a strategy. In fact, it is only necessary to run one full set of performance model analyses. Subsequently, the relative probabilities of each realization can be modified based on the updated probabilities for the parameters to create the revised performance probability distributions (such as those shown in Figure 2-6).

An example of how this can be carried out is schematically illustrated in Figure 2-7. In this example, the post-strategy probability distribution of a performance result for a particular value of a performance measure (R = 10) is plotted (i.e., a vertical slice through Figure 2-6 at R = 10). Strategy A is the same strategy shown in Figure 2-6. Strategy B is an alternative strategy. This figure indicates that very little knowledge regarding system performance will be gained by carrying out strategy B, while strategy A is likely to greatly change the current performance prediction. Note the asymmetric nature of the post-strategy distributions, indicating that the provision of additional information is most likely to move the distribution towards the expected value. That is, the expected result of a strategy is to bring the tails of the distribution inwards toward the prior mean. While there is always the possibility of a bad surprise (i.e., predicted performance is worse), this is what reducing uncertainty means.

The strategy evaluation model is described in detail in Chapter 6.

2.4 Parameter Definition and the Flexibility of the RIP Software

Throughout this report reference will be made to 'input parameters' for RIP. While certain types of input parameters are explicitly required by RIP, the software contains logic that allows the user to define additional parameters and equations for very general kinds of system components. That is, the model consists of a basic computational framework representing the controlling processes, but is intentionally flexible such that the user can represent the processes with as much detail as desired.

The user can even create and specify processes and parameters which are not explicitly incorporated into RIP. In a sense, the RIP program is similar to a spreadsheet: while it contains a large amount of built-in logic and calculational capabilities, the problem that is solved is entirely defined by the user. That is, RIP has been developed such that it is relatively free of assumptions regarding the details of waste package behavior and radionuclide transport processes. Similarly to a spreadsheet user, the RIP user can define a very simple model, or a very complex one. RIP can be run on a personal computer, for simple problems, or can be linked to a powerful workstation for solving more complex ones. The reader of this report who may be daunted by some of the theory described in later chapters should not abandon hope: it is possible to use RIP in a relatively simple manner to address relatively simple, though perhaps very important, issues.

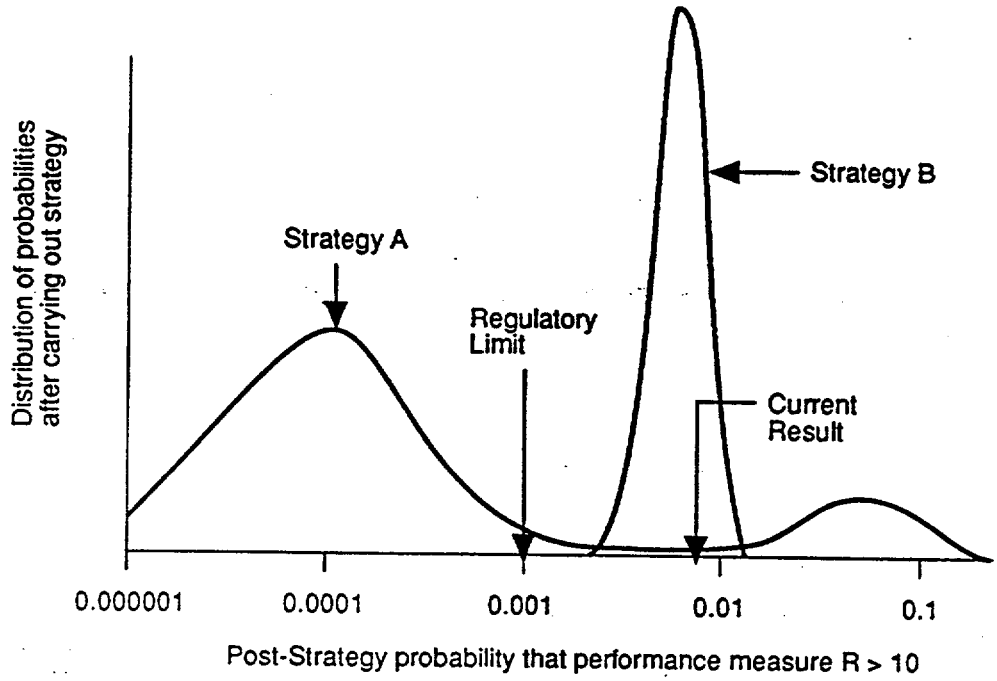


FIGURE 2
PROBABILITY OF POST-STRATE
PERFORMANCE RESULTS
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 27832 DATE 4/6/92 DRAWN BY TB

Golder Associates

In fact, the cornerstone of the RIP methodology, which it is critical to understand in order to take full advantage of the software, is the concept that model input parameters can be defined by the user with a great deal of flexibility. In particular, *a given system parameter can be specified as a constant, a stochastic (i.e., represented as a probability distribution), or as a function of other parameters* (which themselves can be constants, stochastics, or functions).

Representing system parameters as *stochastics* allows the RIP user to specify the degree of *uncertainty* in a particular parameter. For example, if the current level of knowledge regarding a particular system parameter (such as an elemental solubility) is such that it is only possible to specify its value within certain limits (e.g., greater than 10^{-4} g/m^3 but less than 10^{-1} g/m^3), it would be most appropriate to specify this parameter as a probability distribution (e.g., a log-uniform distribution ranging between 10^{-4} and 10^{-1}). RIP allows the user to choose from a wide variety of probability distributions (e.g., normal, log-normal, triangular, beta, gamma, discrete, etc.) for a stochastic parameter. For each Monte Carlo realization of the repository system, all of the stochastic parameters are sampled from their specified distributions.

The ability to represent system parameters as *functions* of other parameters imparts to the user the ability to readily add detail to any given system parameter or process represented by RIP. The user interface of the software is designed such that the user literally types in the desired function. In addition to standard mathematical operators (e.g., SIN, COS, MAX, ERF, LOG), the user can define functions using relational operators (e.g., >,<, =) and IF,THEN logic. Such flexibility allows the user to easily modify and add detail to the conceptual and computational model *without having to make changes to the software* (i.e., without modifying and compiling the source code).

As an example, suppose that instead of simply representing an elemental solubility as a stochastic parameter (as illustrated above), the user wished to incorporate a solubility model which explicitly accounted for the effects of temporally or spatially variable chemistry (e.g., pH) conditions. To accomplish this, the user could define a new parameter, called *pH*. This parameter could be defined as a constant, a stochastic, or a function (of other parameters or of time). It could also be assigned a random spatial variability throughout the repository. The elemental solubility could then be described as a function of this parameter. For example:

$$SOL_n = (A \times pH) + B \qquad (2.1)$$

where SOL_n is the solubility of element n and A and B are constants or additional user-defined parameters.

At a more complex level, the concepts of stochastics and functions can be combined. For example, the solubility of a species could be defined as a log-normal distribution whose mean and standard deviation were functions of pH.

Use of the *Parameters Module*, the software component which implements the concepts described above, is discussed in detail in the RIP User's Guide.

2.5 Summary of the Repository Integration Program

RIP is a powerful and flexible computational tool for carrying out probabilistic integrated total system performance assessments for geologic repositories. It embodies probabilistic decision analysis tools that allow it to:

- examine parameter sensitivity;
- evaluate alternative conceptual designs; and
- evaluate alternative site characterization strategies.

The primary purpose of RIP is to provide a management tool for guiding system design and site characterization. In addition, the performance assessment model (and the process of eliciting model input) can act as a mechanism for integrating the large amount of available information about a repository into a meaningful whole (in a sense, allowing one to keep the "big picture" and the ultimate aims of the project clearly in focus). Such an integration would be useful both for project managers and project scientists.

RIP is based on a "top down" approach to performance assessment which concentrates on the integration of the entire system, and utilizes relatively high-level descriptive models and parameters. The key point in the application of such a "top down" approach is that the simplified models and associated high-level parameters must incorporate an accurate representation of their uncertainty.

RIP is designed in a very flexible manner such that details can be readily added to various components of the model without modifying the computer code. Uncertainty is also handled in a very flexible manner, and both parameter and model (process) uncertainty can be explicitly considered. Uncertainty is propogated through the integrated PA model using an enhanced Monte Carlo method.

RIP (and any practical PA model) must rely heavily on subjective assessment (expert opinion) for much of its input. The process of eliciting the high-level input parameters required for RIP is critical to its successful application. As a result, in order for any project to successfully apply a tool such as RIP, an enormous amount of communication and cooperation must exist between the data collectors, the process modelers, and the performance assessment modelers. This is because such a tool must constantly evolve and is only valid for decision making when it actually incorporates the *current state of knowledge*. This is only possible if project scientists *think in terms of performance assessment*, at least to the extent that their results can be readily incorporated into a total system model. In effect, they must be familiar with that portion of the total system performance assessment model which represents the particular process or parameter that they are studying, such that they can recommend modifications in the data (or the model itself if necessary) as more information becomes available. Without this type of integration between the performance assessment modelers and the project scientists, a tool such as RIP can not be validly applied.

The details of the RIP algorithms are described in the remaining chapters. Chapters 3, 4 and 5 describe the three components of the performance assessment model, and Chapter 6 discusses the strategy evaluation model in which the PA model can be embedded.

3. WASTE PACKAGE BEHAVIOR AND RADIONUCLIDE RELEASE MODEL

3.1 Introduction

As discussed in Chapter 2, the RIP integrated repository performance assessment model is a complex radionuclide transport model, and consists primarily of a series of inter-connected, fully coupled component models with input/output relationships for radionuclide transfer. The three major component models address waste package behavior and radionuclide release, transport pathways to the accessible environment, and ancillary processes such as disruptive events.

The purpose of this chapter is to describe the structure of the integrated performance assessment model's waste package behavior and radionuclide release component model. The general methodology and assumptions incorporated into the software are presented in detail, and application of the model is discussed. The waste package component model discussed below includes both the waste package itself and the engineered barrier system (EBS).

Considerable effort and progress has been made at developing performance assessment tools for the waste package and the EBS over the last several years (more so than in other areas of repository performance modeling, such as unsaturated zone radionuclide transport), and several models have been developed and are continuing to be revised (e.g., Liebetrau et al., 1987; O'Connell, 1990; Robinson and Worgan, 1991). The waste package model within RIP relies upon some of the same basic concepts employed by these existing codes.

RIP is primarily distinguished by its unique "parameters module" (described briefly in Chapter 2 and in detail in the RIP User's Guide) and user-interface which make the software enormously flexible as well as powerful, allowing the user to easily define new model parameters and create alternative conceptual models of waste package behavior. Uncertainty and variability can easily be incorporated by the user into any model parameter which has been defined. Furthermore, because RIP is an integrated total system performance model, of which the waste package model described here is but one component, waste package behavior is directly coupled with component models describing disruptive events and radionuclide transport through the geological environment. This allows waste package parameters and processes to be directly analyzed with respect to *total system performance.*

These features are important, because in addition to acting as a management tool, it is intended that RIP can become a valuable tool to project scientists who wish to carry out sensitivity analyses and/or test alternative hypotheses with respect to total system performance. The software was specifically designed to facilitate such analyses.

It is intended that existing lower-level (i.e., more detailed) waste package models eventually be used to provide input for RIP and/or as benchmarks during model validation exercises.

3.1.1 Chapter Organization

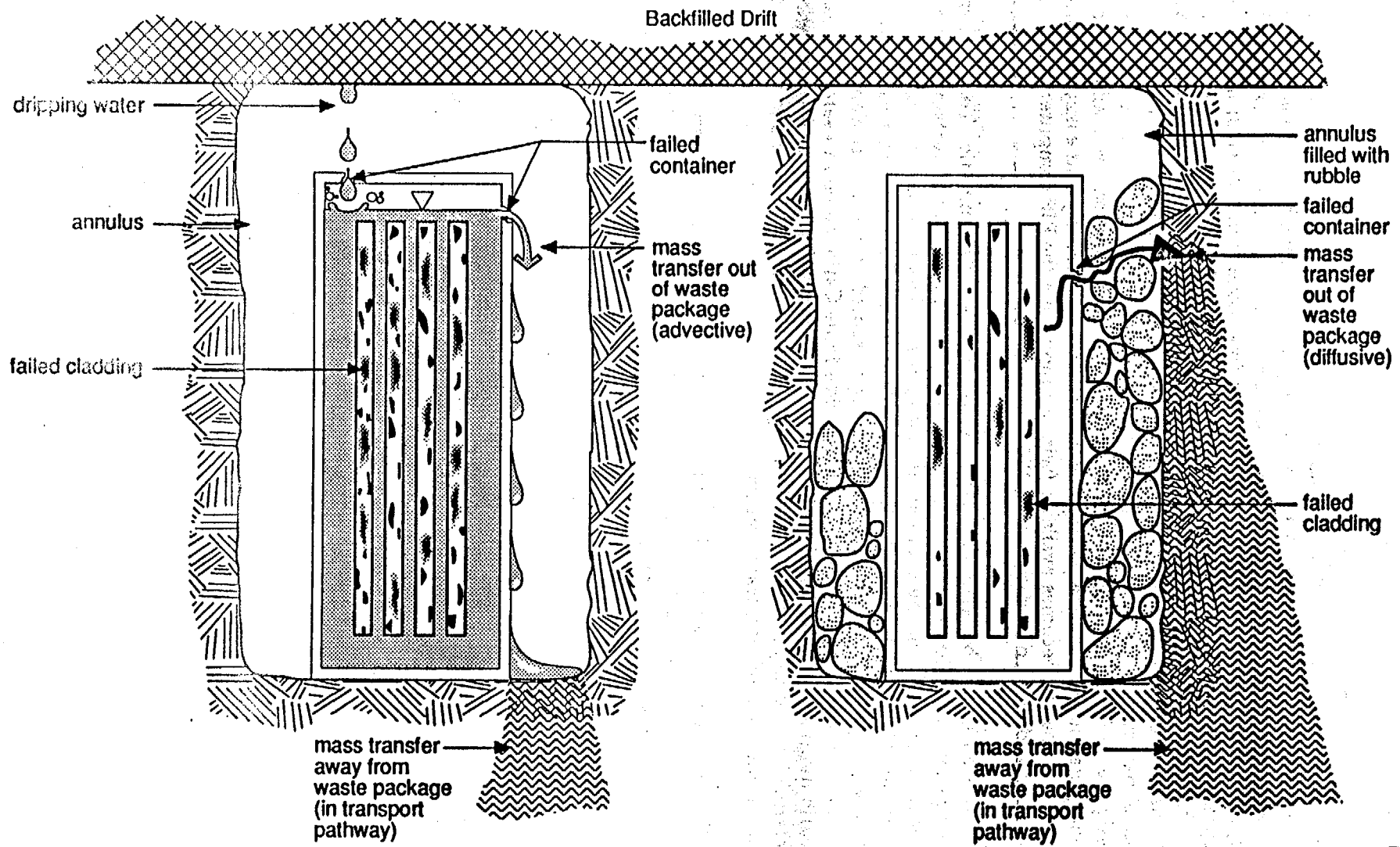
The RIP waste package behavior and radionuclide transport model is described in detail in the following sections. Section 3.2 presents the general methodology for computing radionuclide release from waste packages. Section 3.3 then describes how near-field environmental conditions (and variability in these conditions) are superimposed on this methodology. Section 3.4 presents a general discussion of the required input parameters and the form of the model output. Section 3.5 provides a general schematic summary of the model, and also reiterates the inherent assumptions on which the computational structure is based. This is critical as model results must always be viewed with a full understanding of the assumptions and limitations of the computational algorithm and the input data. To assist the reader, a list of symbols (and their definitions) which are used is included in Section 3.6.

3.2 General Computational Algorithm for Waste Package Behavior and Radionuclide Release

The purpose of the waste package behavior and radionuclide release model is to generate a time history of release from the emplaced waste packages for each radionuclide. To accomplish this, the waste package behavior model must consider two types of processes: 1) breaching of the waste containment system itself (in the case of spent fuel waste, container and cladding) which exposes the waste; and 2) the subsequent mass transfer of the radionuclides present within the containers to the geological environment. These processes are dependent on environmental conditions. Figure 3-1 is a schematic of the waste package release processes. As shown in this figure, both exposure of the waste, as well as mass transfer of the radionuclides, are dependent on environmental conditions. The model explicitly considers the influence of environmental conditions, such as moisture and temperature (including modeling a "thermal period" during which waste packages may be dried out upon heating and subsequently rewet upon cooling.)

Due to the complexity of the processes controlling waste package behavior, it is generally agreed that it is not presently possible to build a practical waste package behavior model based on low level physical parameters and first principles (Golder Associates, 1991a). As pointed out in Chapter 1, a more realistic approach is to build a model based on subjective assessments of relatively high-level phenomenological parameters. In particular, the waste package model described below is built primarily upon two types of high-level parameters: those describing container and cladding failure distributions (which determine when and what portion of the waste inventory is exposed), and those describing radionuclide exposure and mass transfer.

mass release rate = f(mass exposure, mass transfer)
mass exposure, mass transfer = f(environmental conditions)
environmental conditions ≡ moisture, temperature, chemistry...



(after Apted et al. 1991)

FIGURE 3-1
WASTE PACKAGE SCHEMATIC
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 26458 DATE 2/20/92 DRAWN BY TB

Golder Associates

Consistent with this approach, the model has been developed such that it is relatively free of assumptions regarding the details of waste package failure, mass exposure, and mass transfer processes. That is, the model consists of a basic computational framework representing these controlling processes, but is intentionally flexible such that the user can represent the processes with as much detail as desired. For example, a model parameter (such as the matrix alteration rate), can be described by the user not only as a constant or a stochastic (i.e., uncertain) variable, but if necessary, as a complex function of other defined model parameters. As described in Chapter 2, the user interface of the software is designed such that the user literally types in the desired function. This allows the conceptual model of waste package behavior to be continuously and easily modified as more information becomes available (without modifying the RIP computer code). The user interface is described in detail in the RIP User's Guide.

The waste package behavior algorithm described here is deterministic. That is, it simulates a single system realization. However, as pointed out previously, uncertainty in both the model parameters and the component models themselves is explicitly included in the integrated stochastic model. Due to the inherent uncertainties resulting from our lack of knowledge, many of the parameters will be represented by probability density functions (pdf's). The integrated model uses a modified Monte Carlo method to sample these distributions and simulate a large number of random system realizations (using the deterministic model) in order to determine probability distributions of site performance (e.g., cumulative release, transport time).

Both parameter *uncertainty* and parameter *variability* can be explicitly represented in the waste package model. *Uncertainty* in a model parameter implies a lack of knowledge regarding the actual value of that parameter. This is represented in the model by describing the uncertain parameter as being stochastic (i.e., represented as a distribution). The parameter's distribution is randomly sampled each realization. That single parameter value is then used throughout the realization.

Variability in a model parameter implies that for a given realization, a distribution of parameter values exists. An example of a variable parameter would be one which describes water contact modes (moisture conditions) at a waste package (in an unsaturated repository). Due to hydrogeological variability, one would expect the mode of water contact to vary from waste package to waste package throughout the repository. In the model, variability is represented by discretizing the repository into groups of waste packages, and summing the contributions from these groups. The waste packages within a given group are assumed to be under similar environmental conditions, which vary from group to group. These groups are defined based on a number of environmental factors (such as water contact mode) whose variability is considered to be significant. The use of groups has important implications with respect to the capabilities and limitations of the model and will be discussed further in subsequent sections.

It should be noted that although the waste package computational algorithm discussed below is based on high-level phenomenological parameters, the reader will find that it is still fairly complex, and includes a variety of options which enable more detailed representation of various processes. This complexity was unavoidable in order to ensure

that the software is able to realistically represent critical aspects of waste package behavior and radionuclide release. Nevertheless, it is important for the reader to realize that the complexity of the waste package model is completely controlled by the user. That is, due to the flexibility of the user-interface, the user can quickly define a very simple (and approximate) waste package model, or, alternatively, can take more time and effort and define a very detailed waste package model. This flexibility may be difficult to fully appreciate prior to reading the user's guide and actually using the software directly.

The general methodology for developing the failure distributions and using them in combination with radionuclide inventories, dissolution rates, and mass transfer rates to generate waste package release rates for each radionuclide is discussed in detail below. Representation and incorporation of near-field environmental conditions is discussed in Section 3.3, and a summary of model input parameters is provided in Section 3.4. Those readers who simply want to obtain an overview of the waste package model are encouraged to skip directly to Section 3.5, which provides a schematic summary and outlines the major model assumptions.

3.2.1 Overview of Radionuclide Release Calculation

In order for radionuclides to be released from a waste package, they must first be exposed to the near field environment. This is brought about as a result of waste package failure and subsequent waste matrix alteration/dissolution. Even after a given mass of radionuclide is exposed, however, it is not considered to be released until the mass is physically transferred away from the immediate vicinity of the waste package (e.g., a low solubility constituent may be exposed but never released). This transfer occurs via advective and diffusive processes, and is described in the RIP model by parameters defining the rate of mass transfer. Figure 3-2 summarizes the radionuclide release calculation.

The instantaneous release rate for a given radionuclide is dependent on both the rate at which mass is exposed to the near field environment and the maximum rate at which mass can be transferred out of and away from the waste package. In general, the slower of these two rates is the rate-limiting process defining the actual release rate. Because these two rates may vary temporally, however, this generalized rule is not strictly correct.

Consider, for example, a situation in which the mass transfer rate is initially much smaller than the exposure rate. Under these conditions, the release rate is controlled by the mass transfer rate. As a result, exposed mass "accumulates" at the waste package (since mass is being exposed faster than it is being transferred away). Imagine now that the exposure rate eventually drops below the maximum possible rate of mass transfer (e.g., due to a change in environmental conditions or exhaustion of the supply of unexposed mass). Under these conditions, it is incorrect to assume that the release rate is controlled by the slower of the two rates (the exposure rate), since a quantity of mass (exposed previously) is still available to be released.

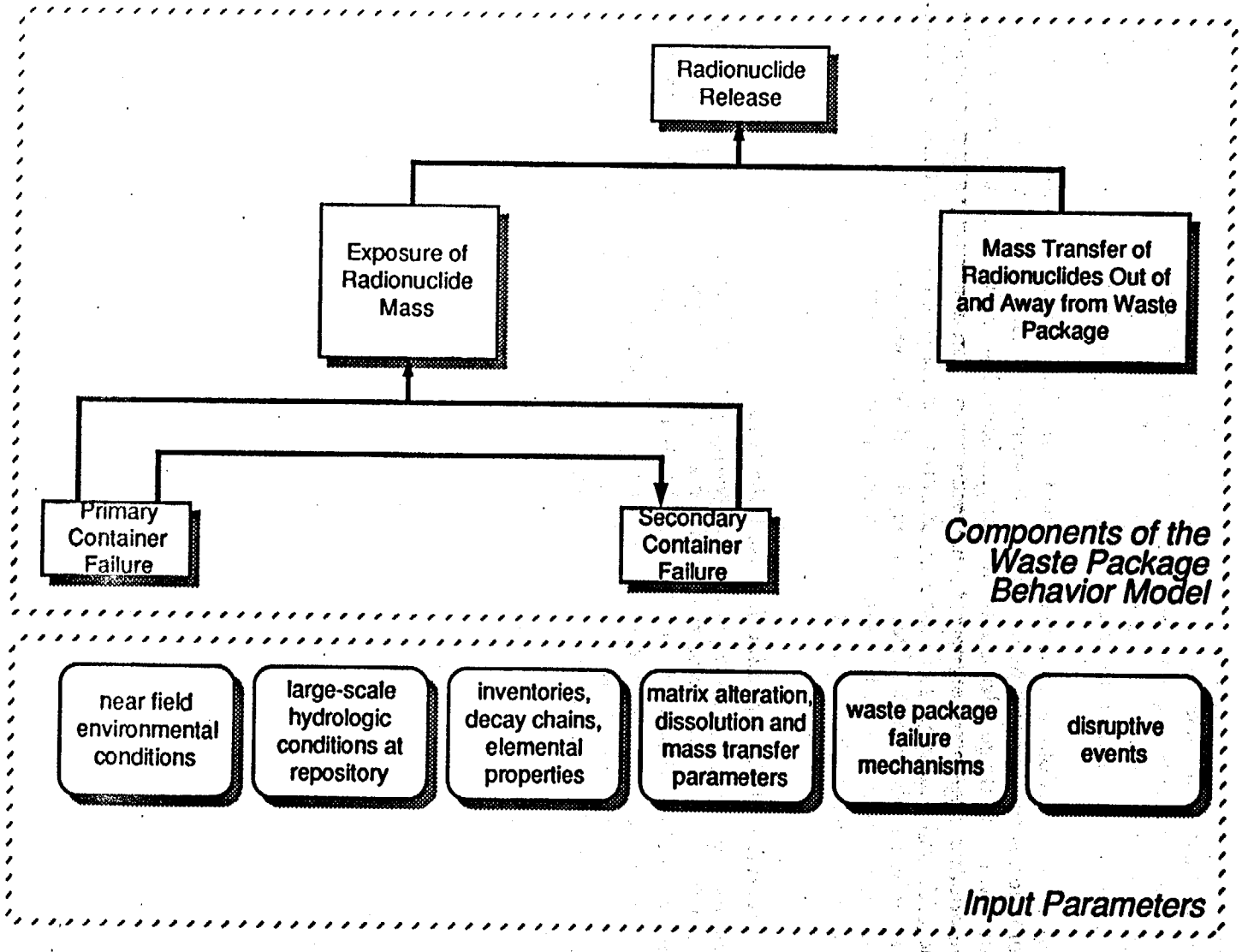


FIGURE 3-2
OVERVIEW OF WASTE PACKAGE BEHAVIOR MODEL
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DWG. NO. 24935 DATE 2/20/92 DRAWN EA

Golder Associates

It is therefore apparent that the instantaneous release rate is not only dependent on the rate of exposure and rate of mass transfer, but is also dependent on the amount of available (i.e., previously exposed) mass still present in the immediate vicinity of the waste package. The actual calculation of the release rate as a function of these three parameters is shown below.

Let $r(n,t)$ = the release rate of radionuclide n from the waste package at time t [M/t];

$e(n,t)$ = the rate at which radionuclide n within the waste package is being exposed to the near field environment at time t (a function of the waste package failure and matrix alteration/dissolution rates) [M/t];

$k_t(n,t)$ = the maximum possible mass transfer rate of radionuclide n out of and away from the waste package at time t [M/t]; and

$M(n,t)$ = the amount of exposed (available) mass of radionuclide n at the waste package at time t [M].

Given these definitions,

$$M(n,t) = \int_0^t [e(n,\tau) - r(n,\tau)]\, d\tau \tag{3.1}$$

The release rate, $r(n,t)$, is determined as follows:

$$\begin{aligned} &\text{if } M(n,t) > 0, \text{ then } r(n,t) = k_t(n,t) \\ &\text{if } M(n,t) = 0, \text{ then } r(n,t) = \text{the minimum of } k_t(n,t) \text{ and } e(n,t). \end{aligned} \tag{3.2}$$

The first case represents a situation in which there is an excess of exposed (available) mass in the immediate vicinity of the waste package, and the release rate is therefore controlled only by the rate of mass transfer. The latter case represents a situation where no excess mass is exposed. Under these conditions, the release rate is controlled by whichever of the two rates is smaller.

Others (e.g., Sadeghi et al., 1989) have typically considered three types of releases: 1) solubility-limited releases, 2) congruent releases from the waste matrix, and 3) releases of readily soluble species. After describing the calculation of $e(n,t)$ and $k_t(n,t)$ in detail in the following sections, the manner in which the above methodology accommodates each of these types of releases will be summarized.

Note that in the discussion above, $e(n,t)$, $k_t(n,t)$, $r(n,t)$, and $M(n,t)$ referred to a single waste package only. Hence, the calculations would need to be carried out for each individual waste package and summed over all waste packages to obtain the total repository release. In order to accurately represent uncertainty and waste package-to-waste package variability, however, such an approach would be prohibitive in terms of computational

expense (i.e., a large number of individual waste packages would need to be simulated every realization). To avoid this difficulty, RIP takes a different approach and *does not simulate individual waste packages.* Instead, RIP divides the population of waste packages in the repository into a discrete number of *waste package groups.* These groups are used to represent repository-wide random variability in near-field environmental conditions. How these groups are defined will be discussed in detail in Section 3.3. For the present purposes, however, it is sufficient to note that e(n,t), k_t(n,t), and r(n,t) represent the mass exposure rate, mass transfer rate, and mass release rate *not for a single waste package, but for a group of waste packages.* Release from the entire repository is calculated by summing the contributions from each group. This simplification has important implications for the interpretation of k_t(n,t) and the resulting release rate, and will be discussed again in subsequent sections.

Finally, it is important to remember that the release rate represents the total rate of release *from* the waste packages. It says nothing about where the radionuclides are released *to.* The RIP integrated model, however, as noted in Chapter 2, consists of components describing transport pathways and disruptive events, in addition to the waste package behavior component. The transport pathways component model defines the transport pathways along which radionuclides can migrate, linking the waste packages to the accessible environment. When the radionuclides are released *from* the waste packages (as defined by r(n,t)), they will be released *to* one or more transport pathways. The available pathways will be defined within the transport pathways component model. The user must specify how a particular nuclide released from a particular type of waste package is partitioned into the various transport pathways. This partitioning may, in some cases, be controlled by environmental conditions. Where the transport pathways are defined to begin is entirely up to the model user. For example, consider a waste package emplaced in a borehole with a backfilled annulus. The waste package model could be used to simulate mass transfer out of the container and through the annulus, with one or more transport pathways *beginning at the edge of the borehole.* Alternatively, the waste package model could simulate mass transfer out of the container, with one or more transport pathways *beginning at the edge of the container,* simulating mass transfer through the annulus and subsequently into the rock. These two conceptual approaches would be differentiated by the specification of mass transfer parameters, as discussed in Section 3.2.4. The transport pathways model is described in detail in Chapter 4.

Details of waste package failure rates and the subsequent radionuclide exposure and transfer rates are presented in the following sections.

3.2.2 Waste Package Failure Rates

The overall radionuclide exposure rate is a function of the rate at which the waste package containers are failing. RIP explicitly considers two layers of waste package containment: the *primary container* (the outside waste package itself), and the *secondary container* (e.g., the cladding for spent fuel or the pour canister for high level defense waste). Note that whenever the word "container" is used alone, the primary container is implied.

As discussed above, rather than explicitly modeling failure mechanisms in detail, waste package failure is represented in terms of high level phenomenological parameters: distributions of container lifetimes. Similar approaches have been used by others (e.g., Bullen, 1990). The manner in which these failure distributions are represented and computed is described below.

3.2.2.1 Primary Container Failure Rate

The general shape of a primary container failure distribution is shown in Figure 3-3. This is a density function of failure frequency. Failure is defined as the initial breaching of the container. In this figure, a waste package has only two states: unfailed or completely failed. The y-axis, c(t), represents the fraction of containers failing at a given time, and the x-axis represents time of failure. This distribution can be thought of equivalently as a container failure rate plotted versus time after repository closure. The integral over time of c(t) would produce the actual fraction of containers which had been breached by a given time.

The shape of c(t) depends on the actual failure modes (such as uniform corrosion, pitting, etc.) which are of importance. The model assumes that it is possible to develop a separate failure distribution for each failure mode. The failure modes are then combined in an appropriate manner to obtain c(t). Appendix E describes in detail the manner in which the separate failure modes are mathematically combined. It is assumed that the different modes operate independently and without synergism. For the purposes of illustration, Figure 3-3 represents the combination of three failure modes: one due to flawed packages or emplacement, and two that represent different corrosive mechanisms.

Mode 1 in the figure represents the failure of containers due to flawed packages and/or flawed emplacement. This process might be represented by an exponential distribution. Modes 2 and 3 represent the failure of containers due to two different corrosion mechanisms. A number of distributions could be used to describe these corrosive failure modes. One such distribution is the Weibull distribution, which is used in many engineering applications to predict component failure. As shown in Figure 3-3, the three distributions describing the different failure modes are combined to produce the total failure distribution c(t).

Each of the independent failure mode distributions must be described by one or more parameters (e.g., the Weibull distribution requires three parameters). Two other pieces of information are also required to define a container failure mode distribution: 1) the probability that the mode is active at any given waste package (e.g., only a small fraction of the waste packages will be susceptible to failure due to flawed packages or emplacement); and 2) whether the mode can start immediately upon closure or cannot start until the waste package is rewet (after drying out due to the thermal pulse). The rewetting time is a function of the thermal and moisture conditions in the vicinity of the waste package and will be discussed in detail in Section 3.3.

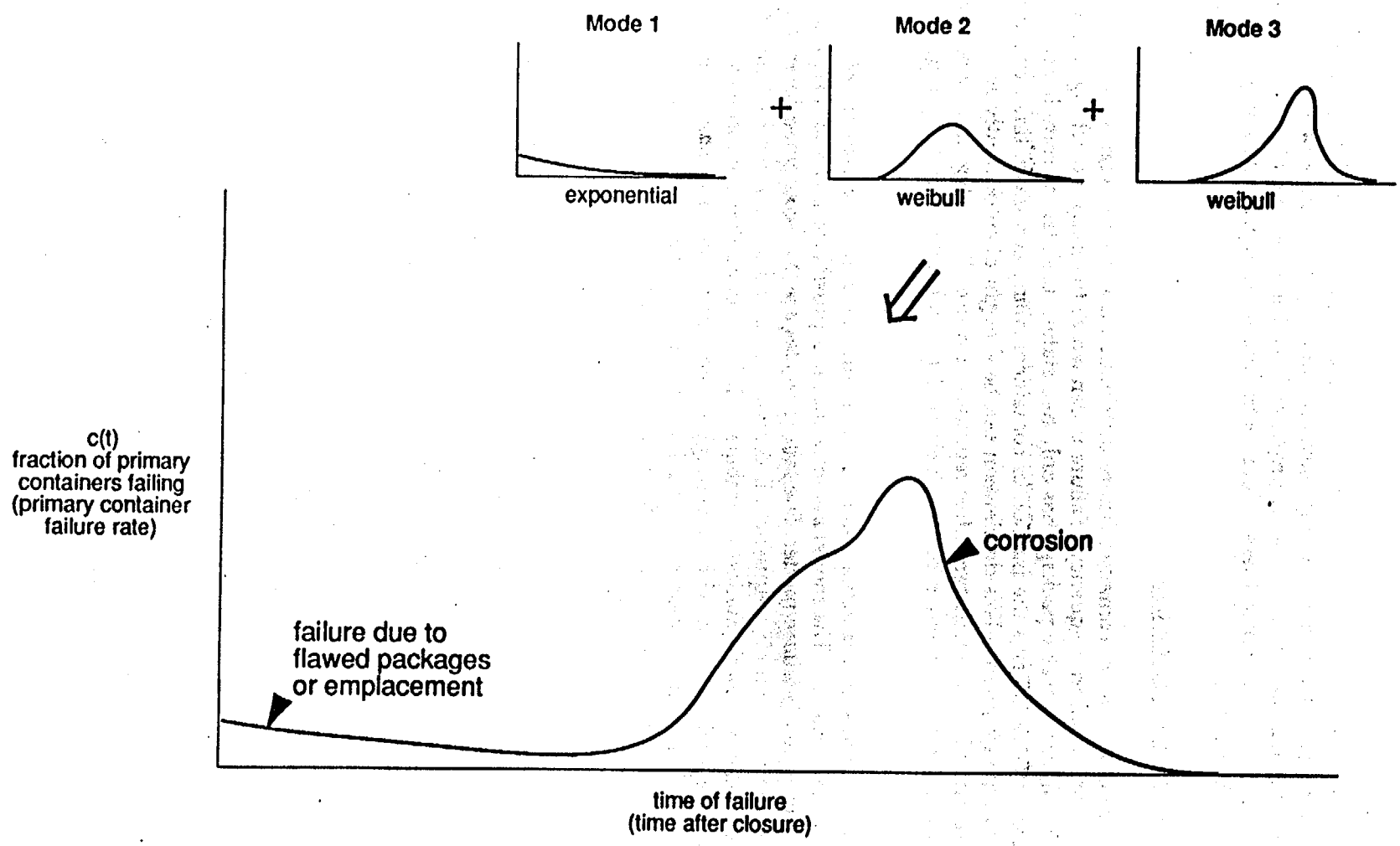


FIGURE 3-3

PRIMARY CONTAINER FAILURE DISTRIBUTION

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 24936 DATE 2/27/92 DRAWN BY CB

Golder Associates

The mathematical details of the most common distributions which are likely to be used to represent container and cladding failure (the exponential, Weibull, uniform, and degenerate) are discussed in Appendix F.

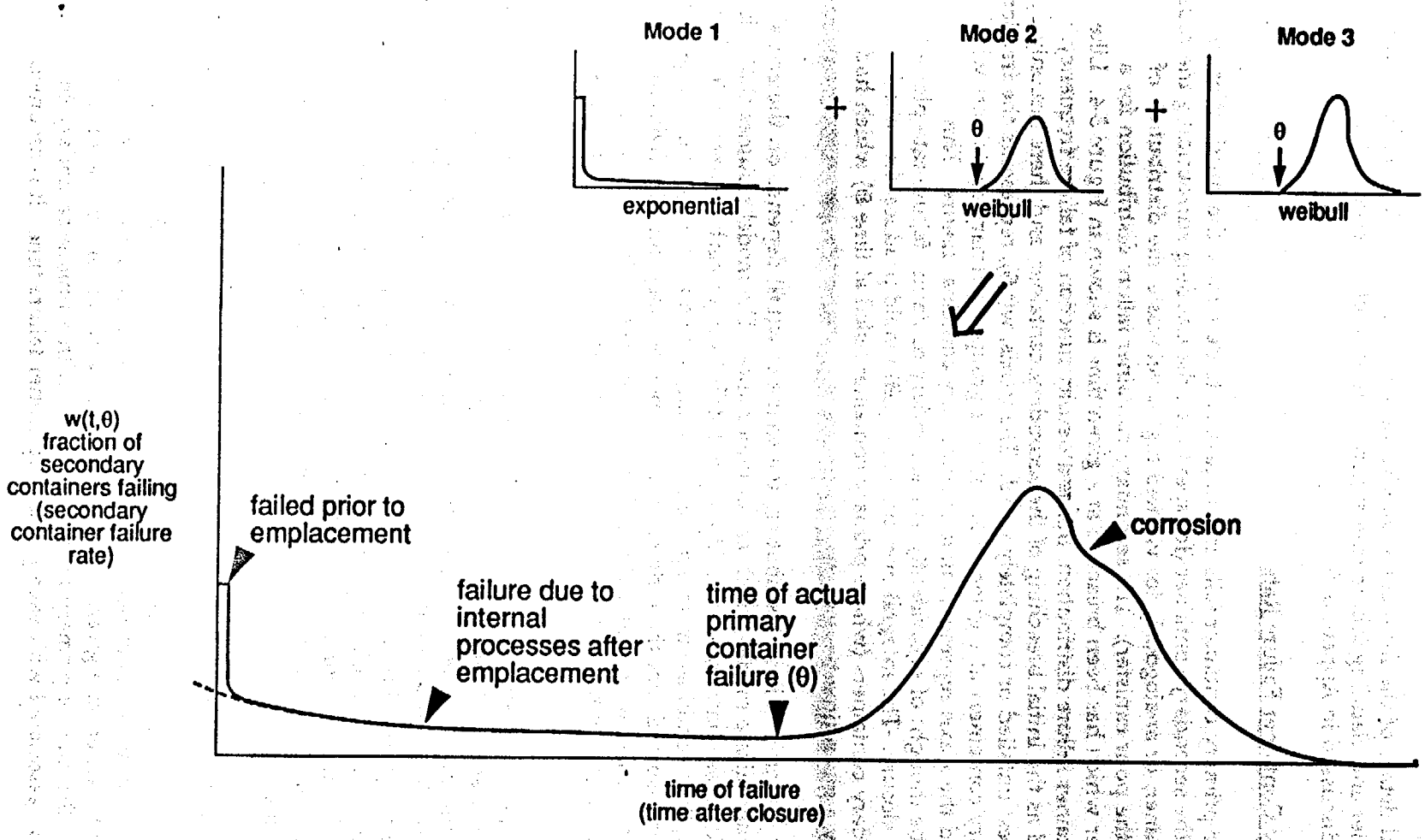
3.2.2.2 Secondary Container Failure Rate

Once a particular primary container fails, it is necessary to determine the distribution of failure times for the secondary container (for spent fuel, the cladding surrounding the rods within that container; or analogously for vitrified high-level waste, the distribution of failure times for the pour canister). The secondary container failure distribution for a primary container which has been breached at a given time is shown in Figure 3-4. Like the primary container failure distribution, this is a density function of failure frequency. Failure is defined as the initial breaching of the secondary container, and there are only two possible states: unfailed or completely failed. The y-axis, $w(t,\Theta)$, represents the fraction of failed secondary containers at a given time within a group of primary containers which fail at time Θ, and the x-axis represents time of secondary container failure. This distribution can be thought of equivalently as a secondary container failure rate plotted versus time after closure. The integral over time of $w(t,\Theta)$ would produce the actual fraction of secondary containers (whose primary container failed at time Θ) which had been breached by a given time t.

Like the primary container failure distribution, the shape of $w(t,\Theta)$ depends on the actual secondary container failure modes which are of importance. The model assumes that it is possible to develop a separate independent failure distribution for each failure mode. The failure modes are then combined in an appropriate manner to obtain $w(t,\Theta)$ (as detailed in Appendix E). It is assumed that the different modes operate independently (without synergism). For the purposes of illustration, Figure 3-4 represents the combination of three failure modes: one due to internal processes, and two that represent different corrosive mechanisms.

Mode 1 in Figure 3-4 represents failure due to internal processes. A spike exists at time = 0, representing the fraction of secondary containers (e.g., cladding) which has failed prior to repository closure due to other processes (transportation damage, internal processes, etc.). The tail of mode 1 represents the failure of the secondary container after closure but prior to actual primary container failure due to additional internal processes. This might be represented by an exponential distribution. Modes 2 and 3 represent the secondary container failure due to corrosion mechanisms following primary container failure. Like the primary container failure distribution, each of these mechanisms could be represented by a Weibull distribution.

In addition to the parameters describing the various independent failure modes, Θ, the time of actual primary container failure, is also required to fully define the shape of the total secondary container failure distribution. It is important to remember that Figure 3-4 is effectively the secondary container failure distribution for a group of primary containers which have been breached at a given time Θ. Hence, the actual shape of the distribution is directly related to the time of primary container failure, with a unique secondary container failure distribution associated with each primary container failure time. It is assumed in



Note: Secondary container is cladding for spent fuel, canister for HLDW.

FIGURE 3-4

SECONDARY CONTAINER FAILURE DISTRIBUTION

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 24937 DATE 2/27/92 DRAWN BY CB

Golder Associates

this model that the time of primary container failure does not affect the values of the other parameters describing the distribution. That is, the only difference between a secondary container failure distribution with primary container failure occurring at time Θ and a second distribution with primary container failure occurring at time $\Theta+\Delta t$ is that in the latter distribution, certain components of the total failure distribution (in this example, the Weibull portions of the distribution) are correspondingly shifted to the right (i.e., increasing time).

Two other pieces of information are needed to define a secondary container failure mode distribution: 1) the probability that the mode is active at any given waste package; and 2) whether the mode can start immediately upon closure or cannot start until the container fails.

3.2.2.3 Rate of Aging for Containers

The choice of using Weibull (or other distributions) to define the failure rates as a function of time is convenient but has a serious shortcoming. Because the parameters of the distributions defining primary and secondary container failure rates (e.g., the three parameters describing the Weibull distribution) cannot vary with time (i.e., a parameter which defines a distribution in time cannot vary with time), it is not possible for the failure rates to respond directly to changing environmental conditions. That is, the failure distribution for a given mode must be defined assuming a given set of conditions. Any temporal changes in environmental factors (e.g., temperature, saturation) which may affect the failure distribution must be known prior to its specification and indirectly incorporated into the form of the distribution.

Because such an approach limits the manner in which time-varying effects can be simulated, an additional feature is incorporated into the RIP model to increase flexibility. This feature allows the failure rates to respond directly to time-varying environmental factors. As described above, all failure mode distributions must be defined in terms of a standard set of constant environmental conditions. Each mode, m, however, has associated with it an additional parameter, known as the rate of aging, $R_{age}(m)$. The rate of aging is a function of environmental conditions (which, as will be described in Section 3.3, may be temporally variable). For example, the aging rate could be specified to increase with increases in temperature or pH.

It is easiest to illustrate the use and physical significance of this parameter by considering an example. Suppose that $R_{age}(m)$ for a certain failure mode was described as a function of temperature such that at a certain elevated temperature the container aged (i.e., failed) at twice the rate as under the standard temperature under which the failure distribution for the mode was defined. If the container was subjected to this elevated temperature for, say the first 1000 years, the *effective* age of the container (with respect to that failure mode) would then be computed as 2000 years, twice the actual age. Hence, if the original failure distribution was defined such that 25% of the containers failed within 1000 years, and 60% of the containers failed within 2000 years, the model would compute that under the elevated temperature conditions, 60% of the containers would fail within 1000 years.

This approach has the effect of stretching and/or shrinking the time axis of the defined failure mode distributions. The effective age of the containers with respect to each mode is continuously tracked. For example, if a mode was not active under certain temperature conditions, $R_{age}(m)$ would be zero, and the effective age of the containers with respect to that mode would not change with time. Note that if $R_{age}(m)$ is always equal to one, the effective age of a container is equal to the actual age.

Due to computational considerations, aging rates are only applied to primary container failure rate calculations and are not currently incorporated into secondary failure rate calculations. As a result, secondary container failure rates cannot respond directly to temporally changing environmental factors.

3.2.3 Mass Exposure Rates

As described in Section 3.2.1, the radionuclide release rate for radionuclide n at time t, r(n,t), is a function of the rate at which mass is exposed, e(n,t), the maximum rate of mass transfer out and away from the waste package, $k_t(n,t)$, and the total amount of available (previously exposed) mass, M(n,t). The manner in which the exposure rates are computed is described below.

Mass exposure is brought about by failure of the primary and secondary containers and exposure of the radionuclide mass to the surrounding environment. In other words, *exposed* implies that the mass is *made available* for mass transfer away from the waste package. Nevertheless, exposure does not imply that mass transfer can immediately take place. Mass transfer itself requires certain conditions. For example, if a waste package fails under dry conditions, although the mass is exposed, a non-gaseous radionuclide cannot be transferred in the absence of water. As will be seen in the next section, this dependence is built directly into the mass transfer rates.

The model assumes that the exposure rate for a given radionuclide is made up of three additive components:

- A portion of the inventory is located between the secondary container (e.g., cladding) and the primary container wall and is assumed to be exposed instantaneously upon primary container failure. This is referred to as the "free" inventory.

- A portion of the inventory is located in the gap between the secondary container (e.g., cladding) and the waste matrix, and is assumed to be exposed instantaneously upon secondary container failure. This is referred to as the "gap" inventory. For spent fuel, some species in this inventory will tend to be present at grain boundaries and exposed surfaces of the fuel. It is assumed that these species are exposed instantaneously along with the gap inventory upon cladding failure.

- The remaining portion of the inventory is bound within the waste matrix and is assumed to be exposed only upon alteration and/or dissolution of the matrix itself. This is referred to as the "bound" inventory.

Note that these three categories, although slightly different, are similar to those considered by others (e.g., Johnson et al., 1985; O'Connell and Drach, 1986; Apted et al., 1987).

Since each of these three inventories is exposed in a different manner, it is necessary to compute the exposure rate for radionuclide n at time t for each of the three inventories and sum these values to obtain the total exposure rate:

$$e(n,t) = e_f(n,t) + e_g(n,t) + e_b(n,t) \quad (3.3)$$

where:

$e(n,t)$ = the total exposure rate for radionuclide n at time t [M/t];

$e_f(n,t)$ = the exposure rate for radionuclide n in the free inventory at time t [M/t];

$e_g(n,t)$ = the exposure rate for radionuclide n in the gap inventory at time t [M/t]; and

$e_b(n,t)$ = the exposure rate for radionuclide n in the bound inventory at time t [M/t].

These three components can be computed as functions of the following parameters:

$c(t)$ = primary container failure rate at time t [1/t];

$w(t,\Theta)$ = secondary container failure rate at time t for a primary container which has failed at time Θ [1/t]

$I_f(n,t)$ = free inventory of radionuclide n *per unfailed container* at time t (e.g., that portion located between the cladding and the container wall) [M];

$I_g(n,t)$ = gap inventory of radionuclide n *per unfailed container* at time t (e.g., that portion located in the gap between the fuel and the cladding) [M];

$I_b(n,t)$ = bound inventory of radionuclide n *per unfailed container* at time t (e.g., that portion bound in matrix) [M];

$k_{a/d}(t)$ = fractional alteration/dissolution rate of waste matrix [1/t];

N_c = number of containers; and

M_c = mass of waste matrix per container [M].

c(t) and w(t,Θ) have been discussed in Section 3.2.2. $I_f(n,t)$, $I_g(n,t)$, and $I_b(n,t)$ define the radionuclide inventory in an unfailed container. These inventories change with time due to decay processes (some nuclides decreasing and others increasing due to the production of daughter products). $k_{a/d}(t)$ is the alteration/dissolution rate of the waste matrix. $k_{a/d}(t)$ can be a function of environmental parameters (e.g., temperature and moisture), as will be discussed in Section 3.3.

Recall that the mass exposure rate, e(n,t), is computed for each waste package group.

3.2.3.1 Free Exposure Rate

The exposure rate for that portion of the waste which is exposed instantaneously upon primary container failure is directly proportional to the container failure rate distribution:

$$e_f(n,t) = N_c \times (\text{free inventory}) \times (\text{container failure rate})$$

$$e_f(n,t) = N_c \times I_f(n,t) \times c(t) \tag{3.4}$$

3.2.3.2 Gap Exposure Rate

The exposure rate for that portion of the waste which is exposed instantaneously upon secondary container failure is somewhat more complex, and is a function of both the primary and secondary container failure distributions:

$$e_g(n,t) = N_c \times (\text{gap inventory}) \times (\text{rate of cladding breaching in failed containers})$$

$$e_g(n,t) = N_c \times I_g(n,t) \times g(t) \tag{3.5}$$

where

$$g(t) = \int_0^t [c(t) \times w(\tau,t) + c(\tau) \times w(t,\tau)]\, d\tau \tag{3.6}$$

The first term in the integral represents the contribution from those primary containers which are instantaneously failing at time t whose secondary container had previously failed due to internal processes. The second term represents the contribution from primary containers which had failed prior to time t whose secondary container is instantaneously failing at time t.

3.2.3.3 Bound Exposure Rate

The exposure rate for that portion of the waste which is exposed upon alteration/dissolution of the matrix is a function of the mass of radionuclide which is unprotected (i.e, primary and secondary container have failed) but is still bound in an unaltered/undissolved matrix, and the alteration/dissolution rate of the matrix:

$e_b(n,t)$ = (mass of unprotected, unaltered, undissolved matrix) x (matrix alteration/dissolution rate) x (bound inventory) / (mass of matrix per container)

$$e_b(n,t) = M_{uu}(t) \times k_{a/d}(t) \times I_b(n,t) / M_c \tag{3.7}$$

Note that this equation assumes that *exposure* of bound nuclides is controlled by dissolution/alteration of the matrix. (As will be seen in the following section, however, *release* of bound nuclides may be controlled by individual radionuclide solubility considerations.)

$M_{uu}(t)$, the mass of unprotected, unaltered, undissolved matrix, can be determined by solving the following differential equation:

$dM_{uu}(t)/dt$ = (rate at which matrix is being unprotected) - (rate at which unprotected matrix is being altered/dissolved)

$$dM_{uu}(t)/dt = N_c \times M_c \times g(t) - M_{uu}(t) \times k_{a/d}(t) \tag{3.8}$$

Since $k_{a/d}$ may be described as a function of time, Equation 3.8 cannot be solved analytically and is solved numerically using a first-order finite difference approximation:

$$M_{uu}(t) = (M_{uu}(t-\Delta t) + N_c \times M_c \times g(t)\Delta t)/(1 + k_{a/d}(t)\Delta t) \tag{3.9}$$

Note that $k_{a/d}$ in this equation represents any process that can act to expose nuclides which are bound in the matrix. It is assumed that this can come about by two types of processes: 1) aqueous dissolution of the matrix; and 2) air alteration of the matrix. Air alteration is included because it may greatly increase the surface area of the matrix, effectively exposing bound nuclides. Hence, $k_{a/d}$ is computed as a sum:

$$k_{a/d} = k_{wat} + k_{air} \tag{3.10}$$

where:

k_{wat} = dissolution rate of matrix in water [1/t]; and

k_{air} = alteration rate of matrix in air [1/t].

Note that both k_{wat} and k_{air} may be described as functions of environmental conditions, and, as a result, it is likely that they will not occur simultaneously (e.g., k_{air} may only occur at high temperatures before the waste package rewets).

k_{wat} is calculated in the model as follows (Stout and Bourcier, 1991):

$$k_{wat} = R_{dis} \times S \times f_w \tag{3.11}$$

where:

R_{dis} = matrix dissolution rate [$M/L^2/t$];

S = effective surface area of waste matrix in a failed container per unit mass [L^2/M]; and

f_w = fraction of waste matrix surface area in a failed container which is wet.

Note that R_{dis}, the matrix dissolution rate, like other system parameters, can be described by the user as a function of local environmental conditions (e.g., temperature, chemistry).

The fraction of the waste matrix surface area which is wet, f_w, can also be described by the user as a function of environmental conditions. In addition, however, the user has the option of describing f_w as a function of the *mode of primary container failure.* This allows one to take into account the fact that the size and nature of container perforations vary depending on the failure mode, and as a result, the nature and characteristics of water contact with the waste matrix can be affected.

As described in detail in RIP User's Guide, if the user chooses to describe f_w as a function of failure mode, he/she must enter the fraction of waste matrix surface area which is wetted, $f_{w,m}$, for each primary container failure mode m. These are then combined to form the value of f_w used in Equation 3.11 as follows:

$$f_w = \frac{\sum_{m=1}^{NM} N_m \times f_{w,m}}{N_T} \tag{3.12}$$

where:

$$N_m = N_c \int_0^t c_m(\tau)\, d\tau \tag{3.13}$$

$$N_T = N_c \int_0^t c(\tau)\, d\tau \tag{3.14}$$

and

N_m = cumulative number of containers which have failed by mode m;

N_T = cumulative number of containers which have failed by any mode;

N_c = total number of containers;

NM = total number of container failure modes;

c(t) = total container failure rate at time t [1/t]; and

$c_m(t)$ = container failure rate by mode m at time t [1/t].

Note that for containers which fail due to disruptive events, the value of $f_{w,m}$ used is the highest of the values specified for the regular failure modes.

By definition,

$$\sum_{m=1}^{NM} N_m \geq N_T \tag{3.15}$$

As a result, Equation 3.12 implies that the effects of failure modes are additive. That is, if a container fails by two modes, one of which wets 10% of the surface, and another which wets 5% of the surface, the model assumes that 15% of the surface is wetted. In reality, the failures may not be independent (i.e., the failures may affect the same part of the waste package). Therefore assuming that the failures are additive is somewhat conservative. Note the that model automatically ensures that f_w is less than or equal to one.

As pointed out above, use of failure modes to describe f_w is entirely optional. That is, a single value of f_w, independent of failure mode, may be used directly in Equation 3.11 if desired.

k_{wat} is automatically set to zero under dry conditions (i.e., at temperatures above the rewetting temperature, discussed below in Section 3.3.4). The air alteration rate, k_{air}, is directly specified by the user.

The radionuclide exposure rate distributions with time are illustrated schematically in Figure 3-5.

3.2.4 Mass Transfer Rate

Even after a given mass of radionuclide is exposed, it is not considered to be released until the mass is physically transferred away from the immediate vicinity of the waste package. This transfer occurs via advective and/or diffusive processes, and is described by a parameter defining the maximum possible rate of mass transfer for a given radionuclide, $k_t(n,t)$.

It is important to differentiate the mass transfer rate out of the waste package from the mass transfer rate through transport pathways (discussed in Chapter 4), and to describe how these are related. The RIP waste package component model computes a maximum possible mass transfer rate *out of the waste packages*. As discussed previously, this is equivalent to the maximum possible mass transfer rate into the beginning of a transport *pathway*. Theoretically, if transport through the pathway is rapid, it will have no effect on

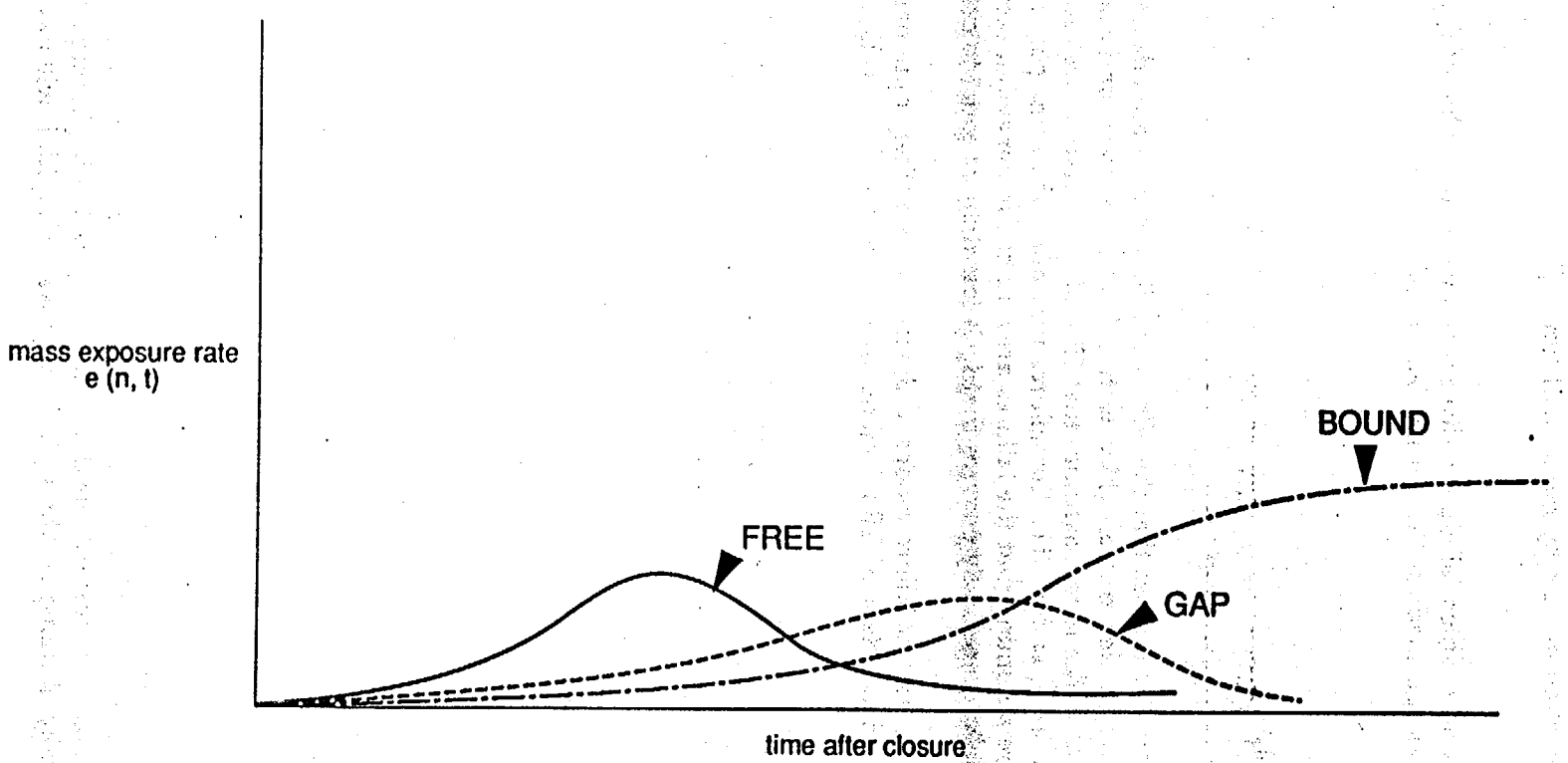


FIGURE 3-5
RADIONUCLIDE EXPOSURE DISTRIBUTION
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 24938 DATE 9/6/91 DRAWN BY CB

Golder Associates

the mass transfer rate out of the waste package (i.e., mass is diffused or advected away as it is released into the pathway). If a "bottleneck" exists at the beginning of the pathway, however, this would influence the mass transfer rate out of the waste package (essentially causing exposed mass to accumulate at the waste package). In the RIP algorithm, however, this is not accounted for within the waste package model and *mass transfer out of the waste package is independent of mass transfer through the pathway to which the mass is released.*

Instead, this is accounted for within the transport pathways component model. Algorithmically, a bottleneck in the first pathway will simply act to accumulate mass in the beginning of the pathway. As a result, simulated waste package releases may be overestimated if pathway bottlenecks are present. Nevertheless, accumulating mass at the beginning of the first pathway (as opposed to at the waste package itself) produces equivalent results in terms of mass transport *through the entire system.*

As pointed out in section 3.2.1, e(n,t) and $k_t(n,t)$ represent the mass exposure rates and the mass transfer rate, not for an individual waste package, but for a group of waste packages. Hence, $k_t(n,t)$ represents the total amount of mass that can be transferred from a defined group of waste packages. It is computed as the product of the mass transfer rate from a single failed container, $k_{tp}(n)$, and the total number of failed containers in that group. As will be discussed in Section 3.3, such an approach allows the user to more easily represent waste package-to-waste package variability in near-field conditions throughout the repository. However, this approach also imposes some limitations on the manner in which mass transfer away from the waste package can be represented, and the user must have an understanding of these limitations.

First, this approach implicitly assumes that all of the exposed mass at individual waste packages is evenly distributed between the failed packages within a given group. In reality, the exposed mass may be concentrated in a fraction of the failed containers (e.g., those that have failed recently). Assuming that the exposed mass is evenly distributed throughout the group results in an overestimate for $k_t(n,t)$. This is because mass transfer from a single waste package can be limited by solubility considerations and/or limited water flux. Spreading the same mass over a group of waste packages reduces both of these limitations, increasing the maximum possible mass transfer rate. This limitation can result in conservative (overestimated) mass transfer rates.

The second limitation is that mass transfer rates can not be described as a function of the time of container failure. This is because the mass transfer rate is represented by a single value for a particular group, but within RIP individual waste packages within a group do not necessarily fail at the same time. As a result, only a *steady-state* mass transfer rate can be used. This steady-state value can change temporally (e.g., due to changing temperature conditions), but because it represents the mass transfer rate for the entire group, it must be independent of the time at which the container was breached and mass transfer away from the package commenced. In effect, transient solutions for individual waste packages can not be represented. Although investigations using analytical solutions (e.g., Pigford et al., (1990)) have indicated that the mass transfer rate (e.g., through a rubble filled annulus) does indeed vary with time and, depending on the conditions (e.g., retardation factors,

porosity), may take many years to reach steady state, we believe that, in many cases, the uncertainty in the parameters and boundary conditions controlling mass transfer are large enough such that the assumption of steady-state rates is appropriate.

Nevertheless, in situations where it is important to represent such transient effects, RIP allows the model user to do so by defining transport pathways appropriately (within the transport pathways component model). As discussed in Section 3.2.1, the user is free to define exactly where the waste package model ends and the transport pathways model begins. For example, if sorption of radionuclides within a rubble-filled annulus was important to consider, the user could define an "annulus" pathway (or pathways) beginning at the edge of the container. Since the transport pathways component model explicitly simulates transient mass transfer (including the effects of retardation), the transient behavior within the backfill would be represented.

The computation of k_{tp}, for both aqueous and gaseous species, is detailed below.

3.2.4.1 Mass Transfer of Aqueous Radionuclides

For an aqueous radionuclide, n, the maximum possible mass transfer rate from a single waste package is computed as a sum of advectively-controlled mechanisms and diffusively-controlled mechanisms:

$$k_{tp}(n) = k_{tp,a}(n) + k_{tp,d}(n) \tag{3.16}$$

where:

$k_{tp,a}(n)$ = maximum possible advective mass transfer rate for radionuclide n [M/t]; and

$k_{tp,d}(n)$ = maximum possible diffusive mass transfer rate for radionuclide n [M/t].

Note that both $k_{tp,a}(n)$ and $k_{tp,d}(n)$ are automatically set to zero prior to rewetting of the waste package.

The maximum possible *advective* mass transfer rate is computed as follows:

$$k_{tp,a}(n) = \text{INFIL} \times \text{CATCH} \times C(n) \tag{3.17}$$

where:

INFIL = repository level groundwater infiltration rate [L/t];

CATCH = effective cross-sectional catchment area associated with a single waste package; it is multiplied by the repository level infiltration rate to compute the effective volume rate of flow of water which is at a concentration of C(n) [L^2];

C(n) = effective concentration of radionuclide n in water at the waste package [M/L^3].

Note that INFIL x CATCH essentially quantifies the volume of water which reaches the effective concentration C(n). Hence, for purely advective transport (e.g., such as the situation illustrated in the left side of Figure 3-1), the mass transfer rate is assumed to be proportional to the volume rate of flow of water coming into contact with the waste. It is further assumed that this water is at a given effective concentration, which as will be detailed below, is either controlled by the solubility of the constituent or by the radionuclide's inventory.

The maximum possible *diffusive* mass transfer rate is computed as follows:

$$k_{tp,d}(n) = D_{eff}(n) \times \omega \times C(n) \tag{3.18}$$

where:

$$D_{eff}(n) = \tau_D \times D_m(n) \tag{3.19}$$

and

$D_{eff}(n)$ = the effective diffusion coefficient for radionuclide n (taking into account tortuosity) [L^2/t];

$D_m(n)$ = the molecular diffusion coefficient for radionuclide n in water [L^2/t];

τ_D = tortuoisty for diffusion through a porous medium; and

ω = geometric factor for diffusive mass transfer [L].

The form of ω is left to the user. A large number of formulations are available in the literature for describing diffusive mass transfer (e.g., Kerrisk, 1984; Aidun et al., 1988; Sadeghi et al., 1989; Pigford et al., 1990). Recall that only steady-state formulations can be used in the RIP model. As discussed previously, the form of $D_{eff}(n)$ and ω will depend on how the subsequent transport pathways into which the waste packages release mass are defined. As illustrative examples, two simple steady state formulations for ω are presented below.

Steady state diffusion from a bare spherical waste form of radius R through a medium with effective porosity η_e (Chambre et al., 1985):

$$\omega = 4\pi R \eta_e \tag{3.20}$$

Steady state diffusion through circular perforations in container wall (Aidun et al., 1988):

$$\omega = \frac{N_p \pi \eta_e a^2}{[\pi \frac{a}{4} + \frac{D_{eff} \eta_e L}{D_{eff,h} \eta_{e,h}}]} \tag{3.21}$$

where:

a	=	radius of circular perforations [L];
N_p	=	number of perforations;
$D_{eff,h}$	=	effective diffusion coefficient in perforation [L^2/t];
L	=	thickness of container wall [L];
η_e	=	effective porosity in backfill; and
$\eta_{e,h}$	=	effective porosity in perforation.

Note that these are simply examples, and no specific formulation for ω is hard-coded into the program; the user must specify ω explicitly, and is free to use any form desired.

Both $k_{tp,a}(n)$ and $k_{tp,d}(n)$ require a value for C(n), the effective concentration of the radionuclide in the water in contact with the matrix. This is calculated as follows:

$$C(n) = MIN[\frac{M_p(n)}{V}, C_s(n)] \tag{3.22}$$

where:

$$M_p(n) = \frac{M(n,t)}{N_T} \tag{3.23}$$

and

$M(n,t)$	=	total exposed mass of radionuclide n at time t [M];
N_T	=	cumulative number of primary containers which have failed by any mode;
$M_p(n)$	=	average exposed mass of radionuclide n per failed waste package [M];
V	=	water volume in contact with matrix in failed waste package [L^3]; and
$C_s(n)$	=	saturation concentration of radionuclide n [M/L^3].

Note that $C_s(n)$ is computed by the model as a function of the elemental solubility and the isotopic mass fraction. The user should be aware, however, that RIP only considers those isotopes which are specifically identified by the user. Isotopes which are not explicitly identified in the input data set (e.g., non-radioactive isotopes neglected by the user) are not considered in this calculation. Not including a significant isotope can result in computed saturation concentrations for other isotopes of the same element which are overestimated.

The formulation for C(n) in Equation 3.22 is an attempt to account for highly soluble species whose effective concentrations may not be controlled by solubility considerations (i.e., the solubility is high and the concentration is determined by how much mass has been exposed and the volume of water into which it has dissolved). Note, however, that V is a difficult parameter to estimate. Specification of a very small value for V is conservative in that it ensures that $C(n) = C_s(n)$.

Note that $k_{tp,a}(n)$ and $k_{tp,d}(n)$ can be described as functions of environmental conditions such that one of the two mechanisms will dominate under a given set of conditions. For example, in an unsaturated repository, under the "wet-drip" water contact mode, mass transfer is likely to be predominantly an advective process, while under the "moist-continuous" water contact mode, mass transfer may be predominantly a diffusive process (Apted et al., 1991). Note that if advection and diffusion are of similar importance, it may not be mathematically appropriate to add these two contributions independently. A more appropriate representation may be to specify ω such that one equation ($k_{tp,d}$) represents both advection and diffusion (e.g., Sadeghi et al., 1989), $k_{tp,a}$ being set to zero.

As discussed in the beginning of this section, k_{tp} represents the mass transfer rate from a single waste package. The model, however, does not simulate waste packages individually, but carries out calculations in terms of groups of waste packages. That is, k_t in Equation 3.2 is the mass transfer rate for a group of waste packages. Therefore, k_t is calculated as the product of the mass transfer rate from a single waste package and the total number of failed waste packages in the group:

$$k_t(n,t) = k_{tp}(n) \times N_T \tag{3.24}$$

As will be discussed in detail in Section 3.3, all of the parameters which define k_{tp} can be described in terms of environmental conditions. In addition, two of these parameters, CATCH and ω, can also be described as functions of the *mode of primary container failure* (completely analogous to f_w described in Section 3.2.3.3). This allows one to take into account the fact that the size and nature of container perforations vary depending on the failure mode, and as a result, the nature and characteristics of mass transfer can be affected.

If the user chooses to describe these parameters as functions of failure mode, a different value of the mass transfer coefficient results for each failure mode. That is, k_{tp} becomes subscripted by m ($k_{tp,m}$), where m is a subscript denoting failure mode. Rather than using Equation 3.24, $k_t(n,t)$ is then computed as follows:

$$k_t(n,t) = \sum_{m=1}^{NM} k_{tp,m}(n) \times N_m \tag{3.25}$$

Note that for containers which fail due to disruptive events, the value of $k_{tp,m}$ used is the highest of the values specified for the regular failure modes.

This equation is completely analogous to the calculation of f_w in Equation 3.12. That is, it essentially assumes that the effects of failure modes are additive. Hence, if a container fails by two modes, one of which produces a mass transfer rate of X, and another which produces a mass transfer rate of Y, the model assumes that the effective mass transfer rate for such a package is X + Y. In reality, the failures may not be independent (i.e., the failures may affect the same area of the waste package). Therefore assuming that the failures are additive is somewhat conservative. However, if most waste packages fail by a mode which produces only restricted mass transfer, the model will be able to more realistically simulate the effects.

As was the case for the parameter f_w discussed in Section 3.2.3.3, use of failure modes to describe CATCH and ω is entirely optional. That is, single values for these parameters, independent of failure mode, may be used directly in Equations 3.17 and 3.18 if desired. To ensure consistency, however, the RIP user interface requires that if f_w is specified as a function of failure mode, CATCH and ω must also be specified as a function of failure mode. That is, either *all three parameters* (f_w, CATCH, and ω) are dependent on failure mode, or *all three* are independent of failure mode.

3.2.4.2 Mass Transfer of Gaseous Radionuclides

For gaseous radionuclides (e.g., ^{14}C), it is not appropriate to compute $k_{tp}(n)$ as a function of solubilities and fluxes. It is quite likely that mass transfer of gaseous radionuclides under unsaturated conditions will be relatively rapid (i.e., k_{tp} will be large). Under these conditions, the rate of exposure would always determine the release rate. However, because it is conceivable that physical situations may arise which limit the rate of gaseous mass transfer (e.g., localized saturation conditions, failed containers whose openings are "clogged"), the model allows the user to directly specify a mass transfer rate for gaseous constituents, $k_{tp,g}(n)$. Note that if gaseous mass transfer is rapid, $k_{tp,g}(n)$ is simply specified as a large number.

3.2.5 Summary of Radionuclide Release Calculation

As described in Section 3.2.1 and summarized in Figure 3-2, the release rate for a given radionuclide at a given time, r(n,t), is computed as a function of the rate of exposure, e(n,t), the maximum rate of mass transfer, $k_t(n,t)$, and the amount of previously exposed (but unreleased) mass at the waste package, M(n,t). Recall that M(n,t) is simply a function of the time histories of e(n,t) and $k_t(n,t)$. The manner in which e(n,t) and $k_t(n,t)$ are computed by RIP has been outlined in the previous sections.

As pointed out previously, others (e.g., Sadeghi et al., 1989) have typically considered three types of releases: 1) solubility-limited releases, 2) congruent releases from the waste matrix, and 3) releases of readily soluble species. Given the information presented in the previous sections, it is instructive to summarize how the methodology described above accommodates these three releases. The category that a particular radionuclide falls into is dependent on the specified input parameters.

Solubility-limited species, by definition, are controlled by the mass transfer rate, $k_t(n,t)$, which is dependent on solubility considerations (Equations 3.17 and 3.18). That is, for solubility-limited species, the exposure rate is faster than the rate at which mass can be transferred away (i.e., $e(n,t) > k_t(n,t)$), and the release rate is therefore equal to the rate of solubility-limited mass transfer (i.e., $r(n,t) = k_t(n,t)$). Note, however, that a species does not have to be solubility limited in order to be controlled by the mass transfer rate (i.e., the saturation concentration is just one parameter which affects the value of $k_t(n,t)$).

Similarly, for species which are controlled by alteration/dissolution of the waste matrix (congruent release), the exposure rate becomes the rate limiting step. That is, the release rate is equal to the exposure rate (i.e., $r(n,t) = e(n,t)$).

For readily soluble species (e.g., released from the fuel-cladding gap and grain boundaries), it can not be stated a priori whether exposure or mass transfer will control release. The important point to note is that the mass transfer rate is not calculated based on solubility considerations (which could be overly conservative since these species may be present at concentrations which are below solubility limits), but based on an estimate of the actual concentration at the waste package (Equation 3.22).

3.2.6 Simulating Different Waste Package Types

The previous sections have often referred to waste packages in terms of spent fuel waste, which will make up the majority of the waste packages. In addition to spent fuel waste, however, high level defense waste will also be emplaced in the repository. Unlike spent fuel waste packages, which consist of fuel rods surrounded by cladding which is then encapsulated within the waste container, defense waste consists of borosilicate glass wastes contained in a single pour canister which in turn is encapsulated in a disposal container. A number of different designs for the waste packages are currently under consideration.

The methodology outlined above, however, is applicable to practically any kind of waste package. Obviously, the radionuclide inventories, alteration/dissolution rates, and failure distributions will differ for different waste types or designs. However, the computational structure and user-interface flexibility is such that RIP can accommodate almost any type of waste package which the user wishes to define.

The standard reference designs for spent fuel and defense waste packages essentially consist of dual containment: a "secondary" container (or containers) within a "primary" container. In the case of spent fuel, the "secondary" containment is the cladding surrounding the individual fuel rods, while for the defense waste, it is the single pour

canister. In both cases, the "primary" container is the outer waste package container itself. Hence, in order to represent defense waste packages, the "cladding failure distribution" becomes the "pour canister failure distribution". Obviously, the actual form of the failure, exposure and mass transfer parameters will differ for the two types of waste packages, but the same methodology can be applied to both. The most important limitation of the methodology is that it is currently limited to explicitly considering only two layers of containment.

Hence, if more than one type of waste package needs to be considered within a simulation, it is simply necessary to specify a separate set of input parameters for each waste package type (e.g., one set pertaining to spent fuel waste packages, and one set pertaining to defense waste packages). That is, a number of the parameters previously introduced become subscripted by waste package type (e.g., $k_{air}(i)$, where i is waste package type). As described inthe RIP User's Guide, the model can accommodate any number of different waste package types.

3.2.7 Incorporation of Disruptive Events

Disruptive events (such as earthquakes, volcanic activity, and human intrusion) may also affect the behavior of the waste package. Disruptive events are simulated in a separate RIP model component. The user defines the types of disruptive events which are possible, their probabilities of occurrence, and their consequences. The manner in which disruptive events are realized and simulated is discussed in detail in Chapter 5. Suffice it to say that their consequences will manifest themselves in the waste package model in three ways:

- A portion of the waste packages can be disrupted in place (e.g., due to mechanical failure induced by some disruptive event).
- A portion of the waste packages (and their inventory) can be moved directly to the accessible environment (AE) or some other location (e.g., the saturated zone).
- The parameters describing waste package behavior and/or environmental conditions for all or a portion of the waste packages may be changed (e.g., flooding of the repository, enhanced corrosion rates).

Note that disruptive events may also have other consequences which affect transport pathways to the accessible environment but do not directly affect the waste packages themselves (and are therefore not discussed here).

The first consequence, in which some portion of the waste packages are disrupted in place, is incorporated into the waste package model by automatically adding an additional primary container failure mode to the failure distributions. Such a failure mode is represented by a spike (mathematically represented by a Dirac delta function) at the time the event occurs. It is assumed that the secondary container instantaneously fails with the primary container. Hence, in addition to the container failure distribution, the cladding

failure distribution $w(t,t_d)$ would also be affected (where t_d is the time of the disruptive event).

The second consequence listed above involves physically moving a portion of the waste packages with their inventory to the accessible environment or some other location (e.g., the saturated zone). In the model, this is represented by instantaneously releasing the corresponding radionuclide inventory to the specified location. It is assumed that the inventory released is equal to the inventory of the specified number of *unfailed* containers at the time of the event (i.e., even if the container failed prior to the event, it is conservatively assumed that most of the inventory would still be present within the waste package). It is also assumed that if the waste package is moved, all of its inventory is immediately released (the constituents are not limited by alteration or mass transfer processes).

The third consequence listed above is concerned with long-term effects on the repository system (e.g., raising the water table, opening a new transport pathway, changing the hydraulic gradient in the saturated zone). For example, one might specify an event which has a consequence of directly changing the value of mass transfer and/or corrosion parameters (perhaps as a result of a change in saturation or a temperature change).

Details of how disruptive event occurrences and consequences are specified are discussed in Chapter 5.

3.3 Representation of Near-Field Environmental Conditions

As pointed out in Section 3.2, the parameters controlling waste package behavior (i.e., the parameters describing the primary and secondary container failure distributions, exposure rates, and mass transfer rates) may be dependent on near-field environmental conditions. This section describes how near-field environmental conditions are represented and used in the waste package model.

Near-field environmental conditions are used by the waste package model in two ways:

- They are automatically incorporated into some of the algorithms which control waste package failure and radionuclide release described in the previous sections (e.g., the infiltration rate at repository level controls advective mass transfer, rewetting behavior can influence both failure rates and mass transfer rates).

- Parameters describing near-field environmental conditions (referred to here as *environmental factors*) can be used directly in the definition of any user-defined model parameter. That is, waste package model input parameters (e.g., matrix dissolution rate, radionuclide solubilities) can be defined by the user as direct functions of environmental factors (e.g., temperature, pH).

As described in Section 2.4, the RIP software is structured such that the user can define input parameters as functions of other user-defined parameters. What sets environmental factors apart from regular user-defined parameters is that they can be defined as having *variability* across the repository.

Recalling the discussion of uncertainty and variability in Section 3.2, it is possible for model parameters representing environmental factors to be both uncertain and variable. Variability should not be confused with uncertainty. Uncertainty in a model parameter implies a lack of knowledge regarding the actual value of the parameter. This is represented in RIP by describing the uncertain parameter as being stochastic (i.e., represented by a distribution). The parameter's distribution is sampled each realization. That single parameter value is then used throughout the realization. Variability in a model parameter, on the other hand, implies that for a given realization, a distribution of parameter values exist. That is, the parameter can vary from waste package to waste package in a given system realization. The mode of water contact at a waste package in an unsaturated repository is an example of a variable parameter (two different contact modes are schematically illustrated in Figure 3-1). Note that the variability that is represented here is a *random variability*. There are no spatial trends implied. The variability in a particular parameter occurs randomly throughout the repository.

Section 3.3.1 describes how variability of near-field environmental conditions is explicitly included in the model. Sections 3.3.2 and 3.3.3 discuss how two major variable environmental factors, temperature and moisture conditions, are represented in the model. As pointed out above, this does not imply that other factors (e.g., chemistry) must be excluded. Because moisture and temperature are clearly critical factors, however, they are explicitly "hard-wired" into the code and are discussed in detail below. The structure of the model is such that other variable environmental factors can easily be defined by the user if necessary, and this is discussed in Section 3.3.4.

Section 3.3.5 briefly discusses the repository level groundwater infiltration rate, a parameter which is defined within the transport pathways component (Chapter 4), but is critical to the waste package component also because it controls the advective mass transfer rate of radionuclides away from a waste package. Finally, Section 3.3.6 describes how the drying and rewetting behavior of waste packages in an unsaturated repository can be represented.

3.3.1 Representation of Variability Between Waste Packages

Waste package-to-waste package variability is incorporated in the model by internally discretizing the waste package population into a discrete number of *waste package groups*. Groups are defined as a function of both near-field conditions and waste package type (e.g., spent fuel, defense waste). That is, a particular waste package group includes waste packages of a particular type which are subject to a specific set of near-field conditions. As was discussed in Section 3.2.6, the user specifies the number of waste packages of each type which are emplaced in the repository. We describe here how waste package groups represent variable near-field conditions.

Near-field conditions are defined in terms of specified variable *environmental factors* (water contact mode, temperature, pH, redox conditions, stress, etc.). The user must determine which environmental factors he/she wishes to include in the model. Two environmental factors, water contact mode (i.e., moisture conditions) and temperature, are explicitly incorporated in the computational algorithm. As will be shown, the structure of RIP is such that other user-defined environmental factors can readily be added.

For each environmental factor which the user wishes to represent as being variable, it is necessary to define a distribution describing its variability. Based on additional user input (described in detail in the RIP User's Guide), RIP then converts the distribution describing the parameter's variability into an *equivalent discrete distribution*. (As will be discussed below, water contact mode is an exception to this rule: it's variability is directly specified in a discrete manner.)

As a simple example, consider a case in which we assume that the environmental conditions can be defined in terms of only two controlling variable environmental factors: the mode of water contact and the temperature. To further simplify the example, assume that the variability distributions of these controlling parameters are already discretized. Since these parameters are described in terms of discrete distributions, they will automatically define a discrete number of *waste package groups*. For example, if we define water contact mode in terms of three discrete categories (e.g., "wet-drip", "moist-continuous" and "dry"), and temperature conditions in terms of three discrete categories (e.g., "hot", "warm", "cool"), the waste package population would be divided into 3 x 3 = 9 waste package groups. If more than one waste package type is included in a simulation (e.g., spent fuel and defense waste), the population would also be distinguished by waste package type, resulting effectively in 3 x 3 x 2 = 18 groups. This is illustrated schematically in Figure 3-6.

A certain fraction of the entire waste package population will be located in each of these groups. For example, if 30% of the waste packages are "wet-drip", 20% of the waste packages are "warm", and 80% of the waste packages are of the type "spent fuel", then the "wet-drip", "warm", "spent fuel" waste package group would contain (0.3 x 0.2 x 0.8 = 0.048) 4.8% of the waste packages. (When computing the number of waste packages in each group, RIP rounds numbers to the nearest integer. Any deviation in the total number of waste packages is then corrected by adding or subtracting an appropriate number of waste packages to the largest group.)

Note that this representation of variability requires that the different environmental factors defined by the user be independent of each other (e.g., the distribution of water contact modes must be independent of the distribution of temperature).

Each environmental factor has a name by which it can be referenced (e.g., CONTAC, TEMP, PH). System parameters can be made dependent on environmental factors by referencing these names in a function (this is described in greater detail in the RIP User's Guide). The value of these parameters would therefore vary from group to group. Within the computational algorithm, at every timestep within a realization RIP computes the release from each group, and then sums the results over all the groups to compute the

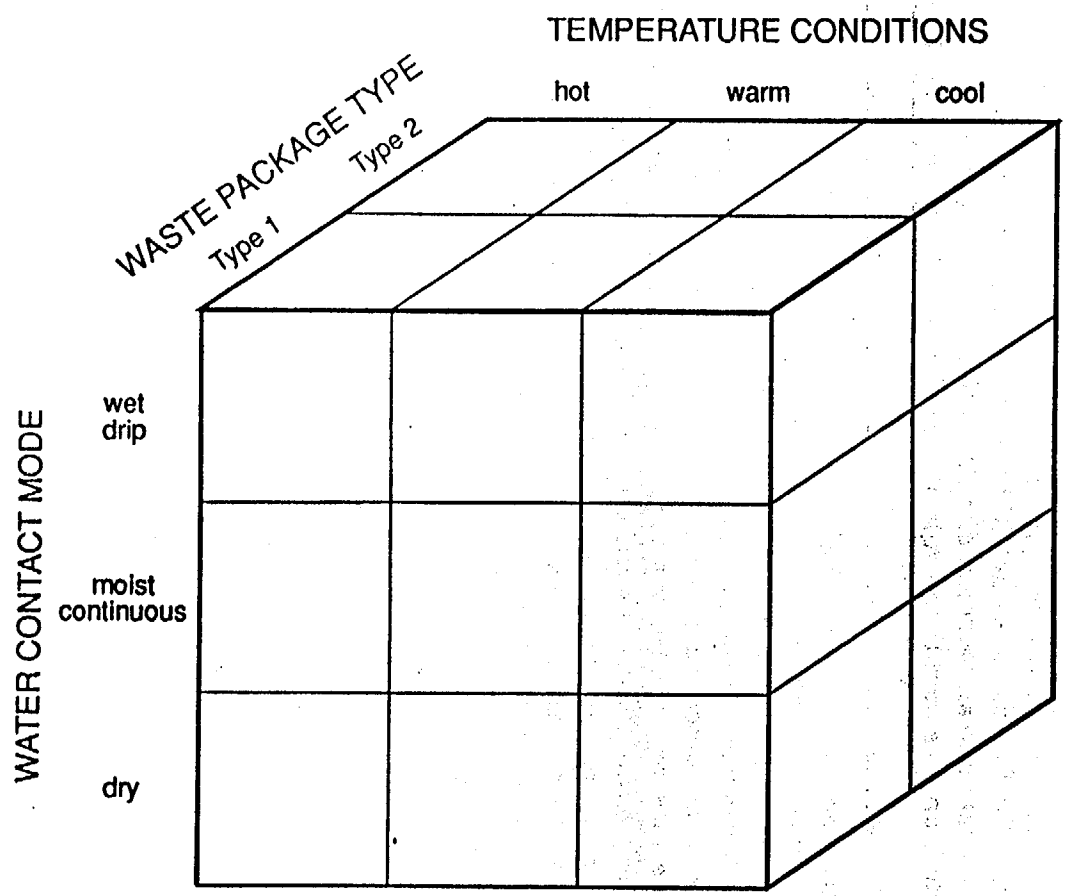


FIGURE **3-6**

EXAMPLE OF DISCRETIZATION OF REPOSITORY INTO ENVIRONMENTAL GROUPS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 24939 DATE 9/6/91 DRAWN BY CB

Golder Associates

total release from the repository. Note that although the parameters controlling waste package release are described in terms of the entire group (e.g., exposure rate for the group, mass transfer rate for the group), it is important to understand that *the model does not assume that all waste packages within a group behave identically.* This is because within a group waste packages fail according to a defined failure distribution, with some waste packages failing earlier than others. As discussed in Section 3.2, this distribution of failure times is explicitly incorporated into the release calculation for a given group.

As discussed in Section 3.2, this method of incorporating variability results in some limitations, particularly regarding the representation of mass transfer. An alternative approach would involve simulating a large number of individual waste packages, the environmental conditions, failure times, and parameters describing mass exposure and mass transfer at each package being sampled from specified distributions. For each realization, however, a large number of individual waste packages would need to be simulated in order to adequately represent the full range of uncertainty in container failure times and the full range of variability in near-field environmental conditions. By simulating waste package groups (as opposed to individual waste packages) we are better able to represent the effect of this variability (since the failure rate distributions are explicitly included in the calculations analytically). Given the computational advantages of this approach, and the fact that the resulting limitations are not severe (see discussion in Section 3.2.4), the authors believe that the waste package group approach is appropriate.

The actual manner in which variable environmental factors are specified is presented in the following three sections.

3.3.2 Representation of Moisture Conditions - Water Contact Modes

The nature of the geological materials at actual repository sites is likely to be quite heterogeneous, producing a wide variety of moisture conditions throughout unsaturated repositories (such as at the proposed Yucca Mountain site). As a result, treating all waste containers as if they will be subject to the same mode of water contact may be a very poor assumption. Various containers will be subject to different moisture conditions as a result of hydrogeological variability. This variability can be represented by a discrete distribution of water contact modes in the immediate vicinity of the waste packages. Figure 3-7 shows an example schematic of such a distribution. In this simple example, it consists of only three categories : "wet drip", "moist-continuous" and "dry" (e.g., O'Connell and Drach, 1986; Apted et al., 1991). Note that the specification and definition of the various categories and the variability throughout the repository (i.e., what fraction of the containers are subjected to each mode) must be determined by the user.

The water contact mode is referenced by the user by assigning an integer value to the parameter CONTAC, each integer value indicating (i.e., acting as an identifier for) a different water contact mode. For each value of CONTAC (i.e., for each water contact mode) the user must specify the fraction of waste packages subjected to those conditions. Each waste package group would then have a particular value for CONTAC (in the example shown in Figure 3-6, "wet-drip", "moist-continuous", and "dry" refer to different

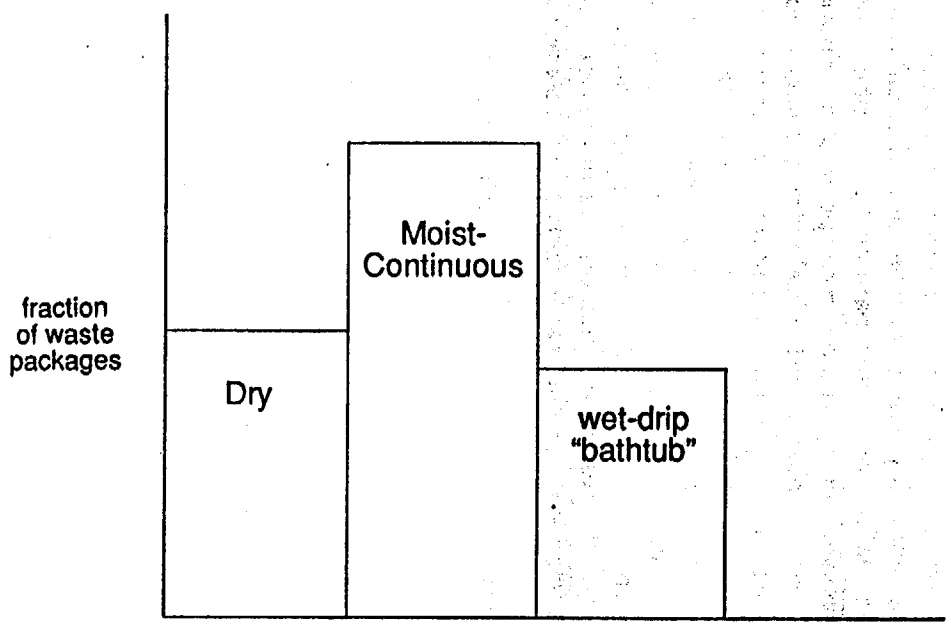


FIGURE **3-7**

EXAMPLE OF WASTE PACKAGE WATER CONTENT MODE DISTRIBUTION

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 24944 DATE 9/6/91 DRAWN BY CB

Golder Associates

values of CONTAC). Waste package parameters may then be specified in terms of the water contact mode by referencing the parameter CONTAC. The specification and use of variable water contact modes, including the manner in which the distributions can be made to change in response to other environmental conditions (e.g., increased infiltration) is discussed in greater detail in the RIP User's Guide.

Note that the water contact mode variability is assumed to be due to local random hydrogeologic conditions at the scale of individual waste packages: it is not meant to represent large-scale spatial trends. In addition, note that the water contact mode represents the long-term distribution expected during the post-thermal period. That is, it is independent of the transient effect of heating in the vicinity of the waste package due to radioactive decay processes. Due to the thermal pulse, some waste packages may remain dry for long periods (and, if included in the contact mode distribution shown above, this would be manifested as a larger fraction in the "dry" category). Although this phenomenon can be represented in the model (as described below in Section 3.3.6), it is not represented by the water contact mode distribution, which only represents post-thermal conditions.

Water contact modes will typically only be applicable for unsaturated repositories. When simulating saturated repositories, a single mode of water contact (i.e., no variability) is likely to be appropriate.

3.3.3 Temperature Conditions at Waste Packages

The second major variable environmental factor which influences waste package behavior is the temperature in the repository and at the waste packages. Temperature behavior is essentially a function of the inventory of the various waste packages, their age, the total number of waste packages, the repository design (e.g., spacing of the waste packages), and the thermal characteristics of the surrounding geologic materials. The moisture content of the repository will also affect thermal behavior (i.e., energy is used in heating and boiling water).

The temperature in the repository will change with time as the waste decays. The thermal behavior of the repository is represented in the RIP model by a parameter defined as the average temperature at the edge of a waste package as a function of time. This parameter is a function of waste package type:

$$TEMPAV_i(t) = T_m(t) + T_i(t) \tag{3.26}$$

where:

$TEMPAV_i(t)$ = the average temperature at the edge of waste package type i at time t;

$T_m(t)$ = the mean temperature (in the rock) throughout the entire repository at time t; and

$T_i(t)$ = the incremental temperature above the mean repository temperature at the edge of waste package type i at time t.

The user must directly specify the forms of $T_m(t)$ and $T_i(t)$. Figure 3-8 illustrates the anticipated form of $T_m(t)$ and $T_i(t)$. As described in the RIP User's Guide, the time histories $T_m(t)$ and $T_i(t)$ can be described stochastically to account for uncertainty.

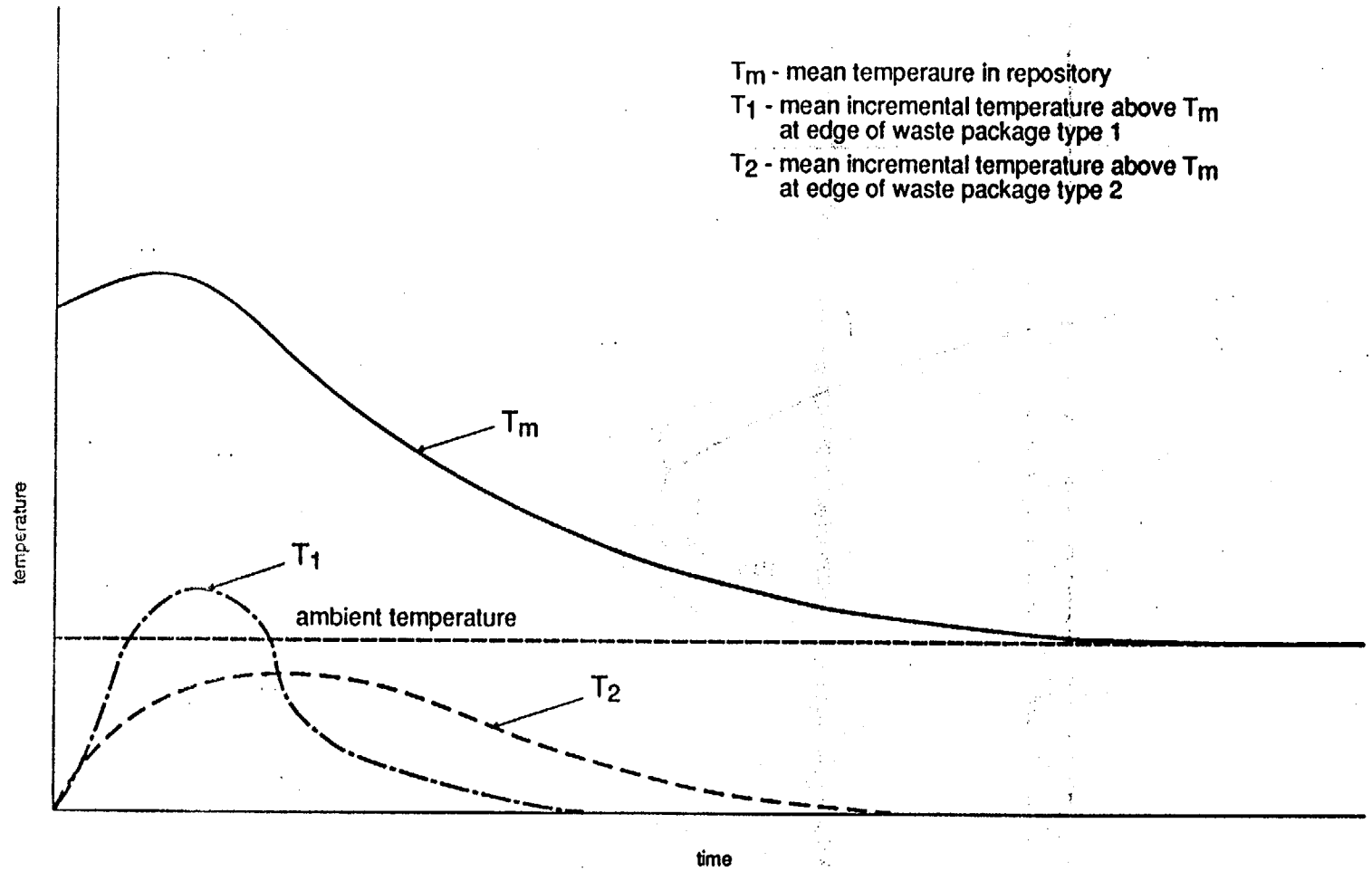
Because some of the controlling parameters are likely to be variable (e.g., geologic materials in the repository are expected to be somewhat heterogeneous; waste packages will have different heat production rates depending on their type, their age, and where they are located within the repository), the actual temperature history at a waste package may vary significantly throughout the repository. This is represented in the model by defining the *variability* around the average temperature $TEMPAV_i(t)$ and creating a *discrete distribution* of temperature variability.

Specification of temperature variability is carried out in two steps. First the user specifies the form of a variability distribution which is used to modify the average temperature TEMPAV. We refer to the variability parameter as TEMPV. In the current version of RIP, the distribution of this parameter can be either triangular or uniform, and must be symmetric about the value 1 (e.g., a uniform distribution between 0.85 and 1.15). Having specified the distribution, the user then specifies how this distribution is to be discretized by defining several (up to six) cumulative probability levels (e.g., 0.33, 0.66, 1.0). RIP then discretizes the distribution by using the expected value of the ranges defined by the specified probability levels as the discrete values. For example, specifying a uniform distribution with a variability of 15% (0.85 to 1.15), and discretization at the 0.33, 0.66, and 1.0 cumulative levels, defines the following discrete distribution:

TEMPV	PROBABILITY
0.9	0.33
1.0	0.33
1.1	0.34

The distribution has been discretized into three discrete values corresponding to the three ranges defined by the specified probability levels: 0 to 0.33, 0.33 to 0.66, and 0.66 to 1.0. The discrete values for the distribution are computed as the expected value of each range (e.g., the 0.333 cumulative probability level corresponds to a value of 0.95; the expected value of a uniform distribution between 0.85 and 0.95 is 0.9). Note that if a triangular distribution had been specified, the values would have been different, since the expected value in each range will be shifted toward 1.0.

The three discrete values described above would define three temperature categories which define variability across the repository: 33% of the packages would have a temperature of 0.9 x $TEMPAV_i(t)$, 33% of the packages would have a temperature of 1.0 x $TEMPAV_i(t)$, and 34% of the packages would have a temperature of 1.1 x $TEMPAV_i(t)$. This is illustrated in Figure 3-9.



Note: temperature at edge of waste package type 1 = $T_m + T_1$

FIGURE **3-8**
MEAN TEMPERATURE AS A FUNCTION OF TIME
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 24940 DATE 9/6/91 DRAWN BY CW

Golder Associates

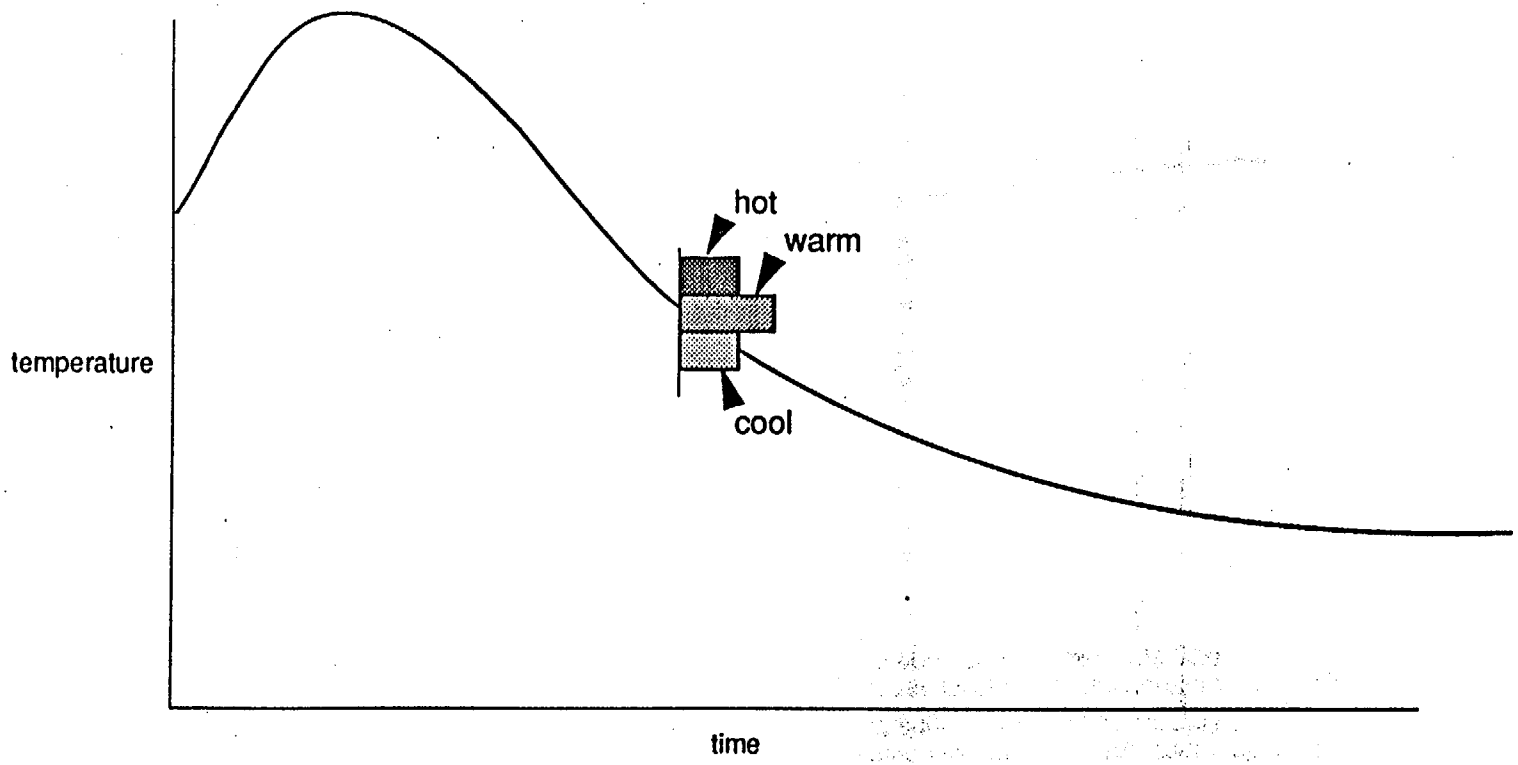


FIGURE 3-9
EXAMPLE OF TEMPERATURE DISTRIBUTION AT WASTEPACKAGE EDGE
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 24941 DATE 2/20/92 DRAWN BY TB

Golder Associates

A single value of TEMPV of 1.0 (implying no variability in temperature) is the default for RIP.

The actual temperature at a particular waste package is referenced by the internal environmental factor TEMP (which is completely analogous to the environmental factor CONTAC discussed previously). Other model parameters can subsequently be defined as a function of temperature, by referencing the name TEMP. At any given time, TEMP, as shown above, is computed internally by RIP as follows:

$$TEMP = TEMPV \times TEMPAV_i(t) \tag{3.27}$$

In the example illustrated in Figure 3-6, "hot", "warm" and "cool" refer to different values of TEMPV (e.g., 1.1, 1.0, 0.9). Hence, in this simple example, TEMPV and CONTAC are the environmental factors which define the waste package groups. Since TEMPV varies between waste package groups, TEMP also varies between waste package groups.

3.3.4 Defining Other Variable Environmental Factors

As discussed above, although only two environmental factors are explicitly incorporated into RIP, the software is structured such that the user can define other environmental factors (e.g., pH, stress, redox conditions) whose variability the user wishes to include in the model. Hence, the user can add variability to any parameter he or she wishes to define.

The user first defines the name of the environmental factor (e.g., PH) and the average value about which the environmental factor's variability is to be centered (e.g, 7). Note that this average value need not be a constant, and may be described stochastically or as a function. A variability distribution is then defined in the same manner as described for temperature (e.g., 0.9, 1.0, 1.1). We refer to each of these discrete values in the distribution as the *variability factor* for each environmental factor category. In this example, we would refer to this factor as PH_V (that is, the environmental factor name with "_V" appended on to the end). PH_V is analogous to TEMPV discussed in Section 3.3.3. This would define a number of discrete PH categories (e.g., 0.9 x 7 = 6.3, 1.0 x 7 = 7.0, 1.1 x 7 = 7.7).

The user-defined environmental factor (in this case PH) is completely analogous to the environmental factors CONTAC and TEMP discussed previously. That is, other model parameters could subsequently be defined as a function of the pH, which would be variable through the repository, by referencing the environmental factor name PH. In the example illustrated in Figure 3-6, adding the three pH categories (e.g.,"acidic" corresponding to 6.3, "neutral" corresponding to 7.0, and "basic" corresponding to 7.7) would add a fourth dimension to the diagram, resulting in 3 x 3 x 2 x 3 = 54 waste package groups. As shown above, the actual value of the environmental factor (in this case, the pH) at any time in any particular group is computed as the product of the average value of the factor and the variability factor. In this example,

$$PH = (\text{average value of pH}) \times PH\_V \tag{3.28}$$

Because PH_V varies from group to group, PH (which can be referenced by the user) varies from group to group.

The definition and use of environmental factors is discussed in more detail in the RIP User's Guide.

3.3.5 Repository Level Groundwater Infiltration Rate

As discussed in Section 3.2.4, the advective mass transfer rate is described as a function of the repository level infiltration rate. The amount of water reaching the repository is ultimately controlled by long-term climatic factors (which control infiltration rates at the repository level), as well as the hydrogeological characteristics of the site (e.g., hydraulic conductivity, fracture connectivity). Climate (and hence infiltration rate at repository level) generally changes rather slowly (typically thousands of years for major changes). Nevertheless, the effective infiltration rate at the repository could conceivably change significantly within time scales of interest. This long term trend can be explicitly represented by the model.

The amount of water reaching the repository is represented by a vertical infiltration rate at the repository level. As will be described in Chapter 4, this can be defined as a function of the infiltration at the ground surface (which in turn will respond to long-term climate changes). Furthermore, it may be described as having a specified spatial trend (as opposed to random spatial variability), some portions of the repository having a higher infiltration rate than others. (As shown in the RIP User's Guide, the model allows the infiltration rate to be specified as a function of waste package type. This permits the user to specify spatial trends in the infiltration rate by simply defining a different waste package *type* for each section of the repository experiencing different hydrologic conditions.)

3.3.6 Rewetting Behavior for Waste Packages

For unsaturated repositories, after the waste is emplaced, radioactive decay of the waste can produce a thermal pulse which will evaporate water in the immediate vicinity of the packages. This pulse may be sufficient to keep the waste packages initially dry. As the waste packages subsequently cool (as the rate of radioactive decay decreases), the material in the vicinity of the waste packages will rewet. As discussed in Section 3.2, several waste package input parameters are sensitive to the rewetting behavior (e.g., contain failure, matrix dissolution, and/or aqueous mass transfer may not proceed prior to rewetting).

Rewetting is a complex process dependent upon near field heat and moisture transport phenomena. Questions remain as to precisely how the repository will behave thermally. Furthermore, the effect of the rewetting process can be quite complex. For example, even if the waste packages do dry out and rewet, mass transfer may be initially limited due to an inward advective gradient. Nevertheless, a simple rewetting model is given here. It is understood, however, that it may be necessary to expand on this simple model in the future to better simulate this complex process.

It is assumed here that a waste package is initially dry and remains dry for a time period t_{wet}, after which water is immediately allowed to contact the waste package (according to the moisture contact modes described above). In reality, this rewetting process will occur gradually, but for simplicity it is treated as an immediate process here (see Figure 3-10). The rewetting time, t_{wet}, is defined in terms of the temperature at which a waste package will become rewet, referred to here as T_r. Note that T_r may be described as a function of other environmental factors (e.g., water contact mode) if desired. Given a value for T_r, the temperature history for a given group of waste packages can then be used to directly determine at what time the waste package temperature reaches the rewetting temperature, thereby defining the rewetting time. This is illustrated graphically in Figure 3-11.

3.4 Model Input Parameters and Output Results

The type of input required by the RIP waste package model can be summarized into the following general categories:

1) *large-scale hydrologic conditions* (effective repository level infiltration) reflecting both hydrogeological influences and climatic change over the period of the simulation) and *near field environmental conditions* (incorporating both waste package design features and environmental influences in the vicinity of the repository);

3) waste package *inventories, elemental properties* (e.g., solubilities), and *decay chains;*

4) waste package (primary and secondary container) *failure mechanisms;*

5) *mass exposure* (matrix dissolution and alteration) and *mass transfer parameters;* and

6) *disruptive events*:
   a) definition and probabilities of disruptive event occurrences;
   b) definition and probabilities of disruptive event consequences.

Note that the six general categories listed are not strictly independent of each other in that they are based on some of the same lower level information. For example, both when defining and generating the waste package failure mechanisms (item 4) or the near field environmental conditions (item 1), one must implicitly assume some specifics of waste package design and a general description of the range of potential environmental conditions. The reason that the required input parameters cannot be considered to be completely independent is that, for the most part, they are not low-level physically based parameters. As pointed out in Section 3.2, it is not practical at this time to build a model based on low-level physical parameters and first principles. Instead, the current model is based on subjective assessments of relatively high-level phenomenological parameters. In the list above, only items 3 and 6a could be considered to be low-level parameters. The rest can be considered high-level phenomenological parameters and must be based on expert interpretation of available data and more detailed modeling of processes.

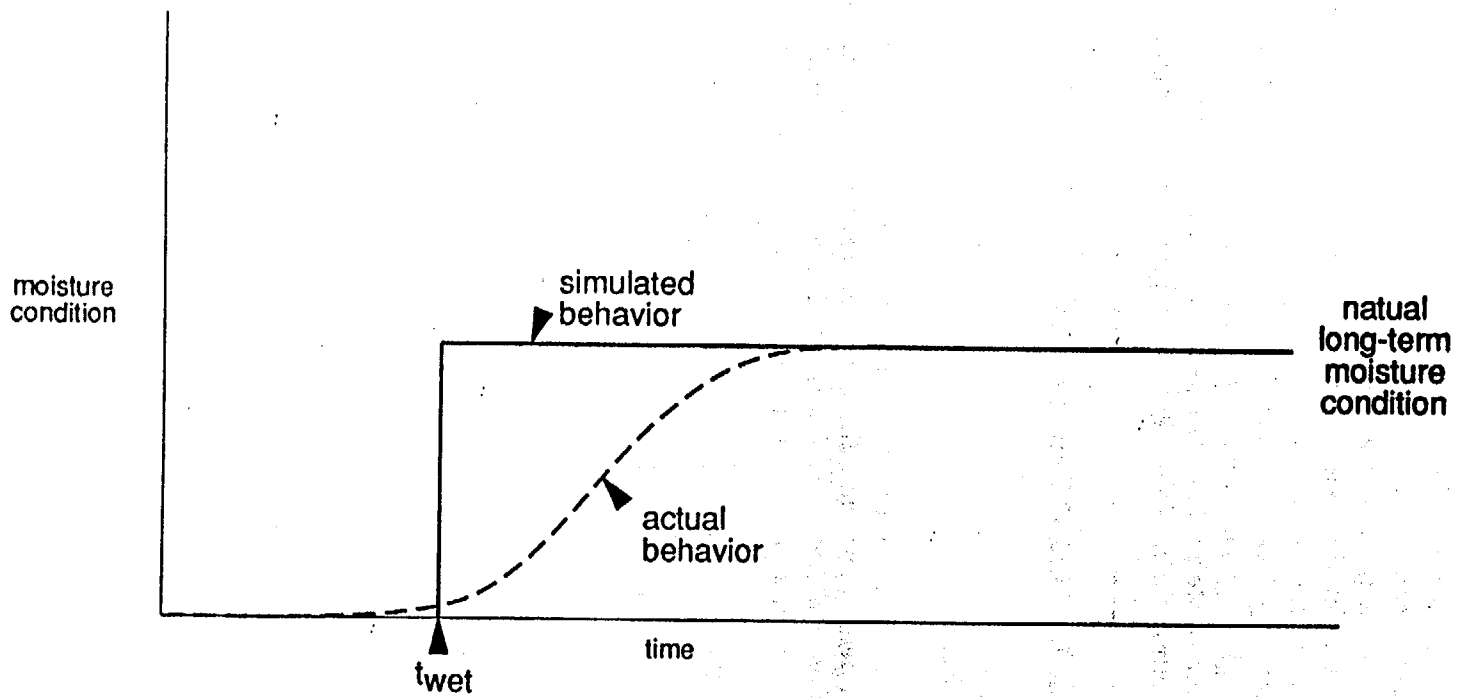


FIGURE 3-10
SCHEMATIC OF WASTE PACKAGE REW[illegible]TING
ARGONNE/MODEL DE[illegible]PMENT

PROJECT NO. 903-1371.203 DRAWING NO. 24942 DATE 9/6/91 DRAWN BY CB

Golder Associates

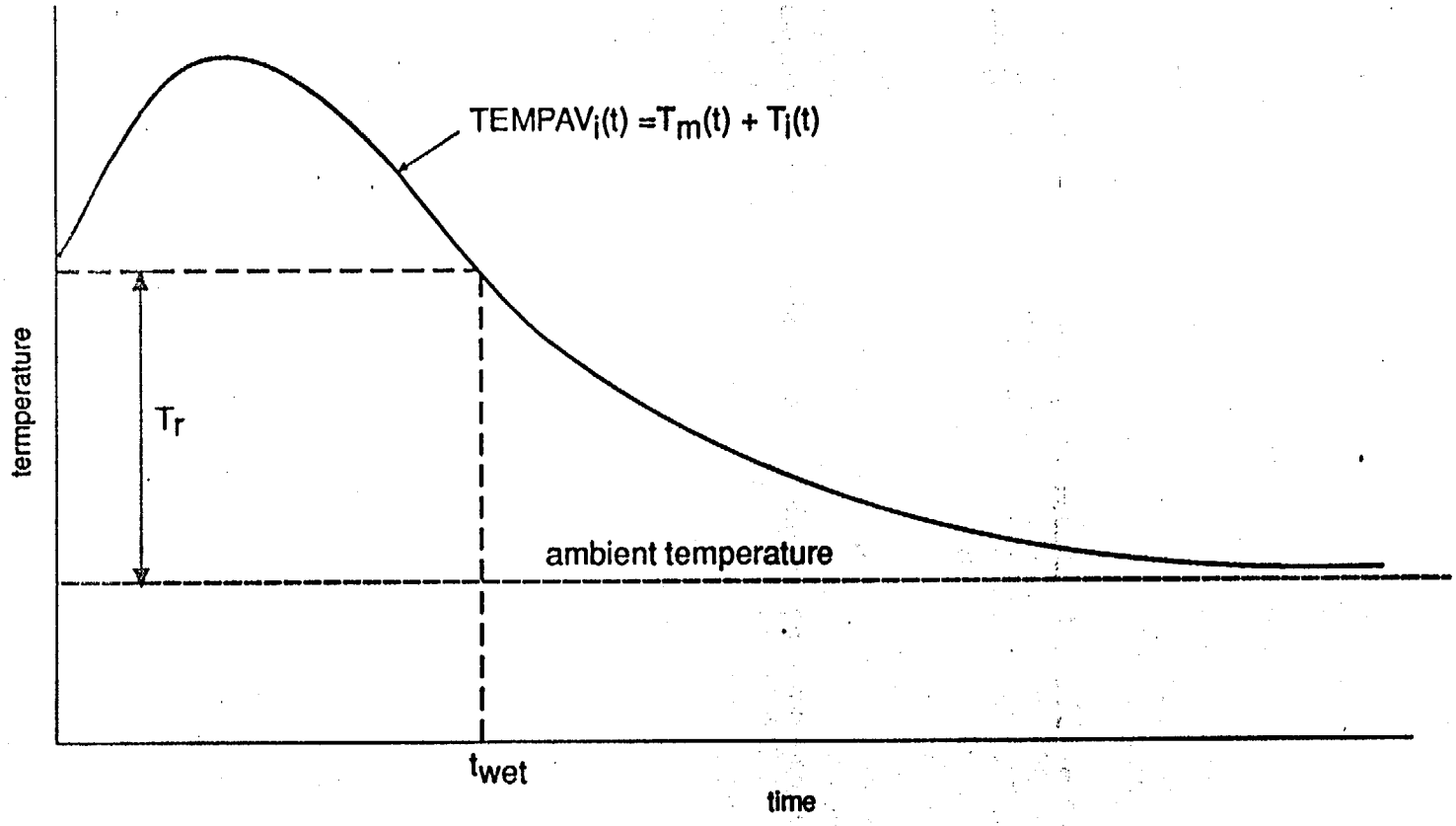


FIGURE **3-11**

CALCULATION OF REWETTING TIME

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 24943 DATE 2/20/92 DRAWN BY TB

Golder Associates

This is illustrated graphically in Figure 3-12. An ideal waste package model based on low level parameters and first principles would model each waste package individually and would require the following input:

- waste package inventories, elemental properties (e.g., solubilities) and decay chains;
- EBS and WP design characteristics;
- geological and hydrogeological characteristics;
- climatic variables; and
- definition and probabilities of disruptive event occurrences.

These parameters would then be input into and manipulated by physically based sub-models which would output higher level parameters (e.g., near field moisture and temperature characteristics). In the present model, however, the low-level physically based sub-models are not included (both because many of the necessary conceptual models and data do not currently exist, and because of the excessive computational time which would be required) and these sub-models are therefore "skipped". The higher level parameters therefore become required input parameters. As can be seen in Figure 3-12, some of these higher level parameters are dependent on some of the same lower level parameters. Although this dependency is not explicitly incorporated in the waste package model (since the low-level sub-models are not included), it must be implicitly incorporated during the process of developing subjective assessments of the high-level input parameters.

Conceptually, as more information becomes available, it may be possible in the future to replace some of the high-level input parameters with lower level parameters and their associated models which will output the higher level parameters directly.

3.4.1 Required Input Parameters

The required input parameters for the waste package model discussed above are summarized in greater detail below. These input parameters will be based on subjective assessments of experts, which, in turn, will be based on experimental data and more detailed modeling of some of the "lower-level" processes discussed above.

Recall that parameters can be specified as constants, stochastics, or functions. A constant parameter is simply assessed as a single value. A stochastic parameter is assessed not as a single value but as a distribution in order to represent the associated uncertainties. This distribution is sampled every system realization. Functions (i.e., dependent parameters) are described as a function of other parameters (which themselves may be either constant, stochastic or dependent).

The user may want to include a stochastic model error term as part of the definition of a dependent input parameter. For example, if k_{air} is described as a function of temperature, the following equation might be used:

$$k_{air} = f(TEMP) \times err_1 \tag{3.29}$$

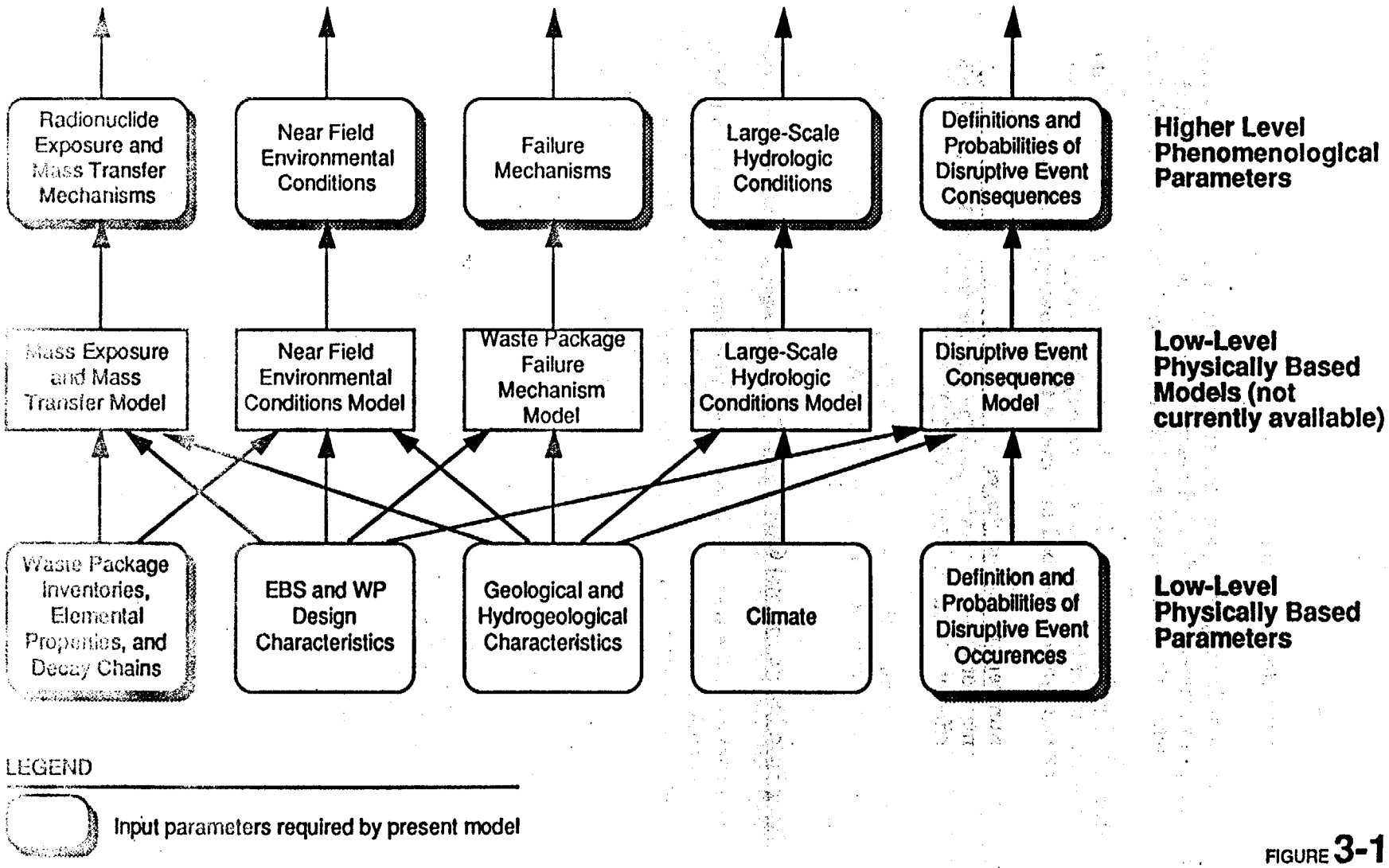


FIGURE 3-12
WASTE PACKAGE INPUT PARAMETERS
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DWG. NO. 24945 DATE 2/20/92 DRAWN TB

Golder Associates

where err_1 is the stochastic model error term.

Note that the following listing supplies only a general description of the waste package model input parameters. A detailed description of the user interface and the actual manner in which these parameters are input is provided in the RIP User's Guide.

Waste Package Description. This data provides a general description of the characteristics of the waste package type(s) which are placed in the repository. *Each waste package type* is described in terms of the following parameters:

- number of waste packages
- mass of waste per waste package [both in units of kg and MTIHM]
- effective waste burnup [MWd/MTIHM]

Note that the mass of waste in terms of MTIHM and the effective waste burnup are required in order to normalize results to EPA limits in the manner specified by 40 CFR Part 191.

Near Field Environmental Conditions.

The behavior of the waste package is dependent on environmental conditions. Some of this dependency is explicitly included in the RIP computational algorithm (e.g., rewetting of waste packages after a thermal period). In addition, other parameters in the model can be described by the user as a function of specified environmental factors. Two environmental factors are automatically incorporated into the model: temperature and moisture conditions. They require the following input:

- the mean repository temperature history (in the rock)
- the incremental temperature history (above the mean repository temperature) at the edge of each different waste package type
- the uncertainty in the temperature histories specified above
- a description of the waste package-to-waste package variability about the mean temperature
- the temperature at which a waste package will rewet (i.e., revert to its long-term moisture condition)
- the fraction of waste packages in each of a number of user-defined *water contact modes* (after rewetting)

Note that the user can, if necessary, define other environmental parameters. These user-defined environmental parameters can also have a specified variability throughout the repository.

Radionuclide Inventory, Decay Chains, and Elemental Properties. This information must be provided for *every radionuclide in every type of waste package:*

- radionuclide name
- radionuclide specific activity (Ci/g)

- first order decay constant (yr^{-1})
- major daughter product
- radionuclide inventory [Ci/package].
- fraction of inventory which is classified as "free", "gap" and "bound"
- elemental solubility [g/m^3]
- effective diffusion coefficient [m^2/yr]
- is the element gaseous?

Waste Package Failure (Degradation).

Each waste package type is assigned by the user one or more failure modes for the primary containment layer (i.e., the outer waste package container). For each failure mode, the following information is required:

- a container failure distribution (failure rate vs. time) for the failure mode
- the probability of the failure mode being active in any given waste package
- the rate of aging (an acceleration/deceleration factor) for the failure distribution, which is typically described as a function of environmental factors

Each waste package is also assigned one or more failure modes for its secondary containment layer (i.e., the cladding for spent fuel; the pour canister for defense waste). For each of these failure modes, the following information is required:

- a container failure distribution (failure rate vs. time) for the failure mode
- the probability of the failure mode being active in any given waste package

Radionuclide Exposure Parameters.

Each waste package type requires specification of several exposure parameters. These parameters describe exposure of bound radionuclides (i.e., those radionuclides bound within the waste matrix) after failure of containment.

- the waste matrix dissolution rate [$g/m^2/yr$]
- the effective surface area of waste matrix (per unit mass) in a failed container [m^2/g]
- the effective fraction of surface area of waste matrix in a failed container which is wet (if desired, this parameter can be described as a function of container failure mode)
- the effective fractional air alteration rate of waste matrix [yr^{-1}]

Radionuclide Mass Transfer Parameters.

Each waste package type requires specification of several parameters which control mass transfer of radionuclides out of and away from failed containers.

- the infiltration rate at repository level for each waste package type [m/yr]

- the effective catchment area for infiltration which determines the amount of water which actually comes into contact with waste matrix in a failed container (if desired, this parameter can be described as a function of container failure mode) [m^2]
- geometric factor for diffusive (aqueous) mass transport (if desired, this parameter can be described as a function of container failure mode) [m]
- volume of water contacting waste matrix in a failed container [m^3]
- effective mass transfer rate for a gaseous radionuclide (must be specified for *each gaseous radionuclide*) [g/yr]

The last two categories of input couple the waste package component model to the radionuclide transport pathways component model and the disruptive events component model:

Waste Package Discharge Distribution.

Each radionuclide in each waste package type is assigned one or more discharge pathways. Mass released from the waste packages will be discharged to these radionuclide transport pathways. The data is entered as follows *for each discharge pathway*:

- name of the discharge pathway
- fraction of waste package release which is discharged to that pathway

Definition of radionuclide transport pathways is discussed in detail in Chapter 4.

Disruptive Events

- disruptive event definitions, probabilities, and consequences (discussed in detail in Chapter 5).

3.4.2 Model Output

The output of the waste package component model essentially consists of probabilistic waste package performance measures (e.g., cumulative release over time period of interest, maximum annual release). These measures can be examined for individual radionuclides, individual waste package groups, or summed over all radionuclides and all groups. Time histories of radionuclide release from the waste packages can also be produced.

The RIP User's Guide discusses the details of the various output options and how the data can be processed and graphically displayed (e.g., as a PDF, CDF or CCDF). As discussed in that document, the values of *all the input parameters for every realization* are also saved. This allows extensive sensitivity analyses to be carried out in order to identify relationships between parameter values and performance.

The structure of the RIP model is such that the waste package component model can be run independently of the radionuclide transport pathways component model, producing

the results outlined above. When the entire integrated model is run, additional outputs pertaining to the pathways component model are also produced (e.g., cumulative release to the accessible environment). These outputs are discussed in Chapter 4.

3.5 Summary of the Waste Package Model

In order to properly apply this (or any) complex simulation tool, it is necessary to have a thorough understanding of the major assumptions on which the computational algorithm is based. The general outline of the waste package behavior model (presented in Chapter 2), highlighting the major model assumptions is reproduced below.

General Model Structure

- The waste package component model is part of the total system performance model, and is directly coupled to component models describing disruptive events and radionuclide transport pathways.
- The model is based on subjective assessments of relatively high level phenomenological parameters, such as container and cladding/pour canister failure distributions, and generalized alteration and mass transfer parameters.
- The model simulates groups of waste packages (rather than tracking the behavior of all the individual waste packages in the repository).
- Waste package parameters can be described by the user as a function of temporally varying environmental conditions.
- The model incorporates radioactive decay and production of daughter products.

Environmental Conditions

- Temperature and moisture conditions are explicitly included in the model. The temperature at the edge of the waste package varies with time. Moisture conditions (intended primarily for unsaturated repositories) refer to the mode of water contact at a waste package (e.g., "wet-drip"). Both temperature and moisture conditions can be specified as variable across the repository.
- Other environmental conditions (e.g., chemistry, stress) can be added by the user. Environmental conditions can be described as being spatially and temporally variable.
- A simple rewetting model is used which assumes that waste packages return to specified moisture conditions upon reaching a specified temperature. At temperatures above the specified temperature, the waste package is considered to be dry.

Waste Package Failure Rates

- Two levels of waste package containment (and failure) are explicitly simulated: the primary container (the waste package itself); and the secondary container (e.g., cladding for spent fuel, pour canisters for high level defense waste).

- Waste package failures are described in terms of density functions of failure frequency.

- Containers and cladding/pour canisters can fail by one or more failure modes. The model combines failure modes by assuming that the failures can be treated independently (they do not act with synergism).

- Container failure modes can be affected by temporally varying environmental conditions.

Mass Exposure

- Exposure of radionuclides is brought about by container and cladding/pour canister failure, as well as matrix alteration/dissolution processes.

- The radionuclide inventory is made up of three additive components: the free inventory, which is exposed immediately upon primary container failure, the gap inventory, which is exposed immediately upon secondary container (e.g., cladding) failure in a failed container, and the bound inventory, whose exposure is controlled by alteration/dissolution of the waste matrix.

- Exposure of the bound inventory is controlled by two types of processes: dissolution of the matrix, and air alteration of the matrix. Dissolution is described in terms of a matrix dissolution rate and an effective wetted surface area. Dissolution is zero prior to rewetting of the waste package. Air alteration rates are specified directly by the user.

- The effective wetted surface area of matrix can be described as a function of container failure mode. It is assumed that for a waste package that fails by more than one mode, the effects on wetted surface area are additive.

Mass Transfer

- For an aqueous radionuclide, mass transfer can be described as an advectively or a diffusively controlled process, each of which may be limited by radionuclide solubilities.

- Aqueous mass transfer is set equal to zero prior to rewetting of the waste package.

- Once mass transfer commences, it can not be described as a function of the time of container failure.

- Several mass transfer parameters can be described as functions of container failure mode. It is assumed that for a waste package that fails by more than one mode, the effects on mass transfer are additive.

- The mass transfer rate for gaseous species is directly specified by the user.

Effect of Disruptive Events

- Disruptive events can affect the behavior of the waste package. Disruptive events are simulated in a separate model component. Their consequences can manifest themselves in the waste package model in three ways: 1) a portion of the waste packages can be disrupted in place; 2) a portion of the waste packages (and their inventory) can be moved directly to the accessible environment or some other location; and 3) the parameters describing waste package behavior and/or environmental conditions may be changed.

Linkage to Transport Pathways Model

- The waste package model is coupled with the transport pathways model in two ways: 1) it relies upon the same large scale hydrologic parameters utilized and/or defined within the transport pathways component (e.g. repository level infiltration rate); and 2) each waste package type discharges its mass to a specified transport pathway (or pathways) defined by the user.

As noted previously, the fact that RIP simulates groups of waste package (as opposed to individual waste packages) is an approximation which has two effects: 1) it tends to overestimate mass transfer rates; and 2) it requires that mass transfer be independent of the time of container failure (although, as discussed previously, this can be compensated for by appropriately defining transport pathways). Note that *conceptually*, modifying the model to simulate individual waste packages would not be difficult. It would, however, create a *computational* problem. That is, in order for such an approach to accurately represent variability throughout the repository, a large number of individual waste packages would need to be simulated, and this would be computationally intensive, severely limiting the practical use of the model. Due to the large amount of uncertainty in required input parameters, the effects mentioned above which are introduced by not modeling waste packages individually were deemed to be of secondary importance. This situation may change, however, as more experience is gained with waste package performance assessment, and future versions of RIP may include explicitly sampling and simulating of individual waste packages.

Figure 3-13 expands on Figure 3-2 and represents a detailed waste package influence/information flow schematic. This diagram illustrates how the various input parameters and component models are linked together. It shows in detail how the various input and output parameters are moved between smaller component models to generate the waste package release distribution with time, graphically summarizing the structure of the waste package model.

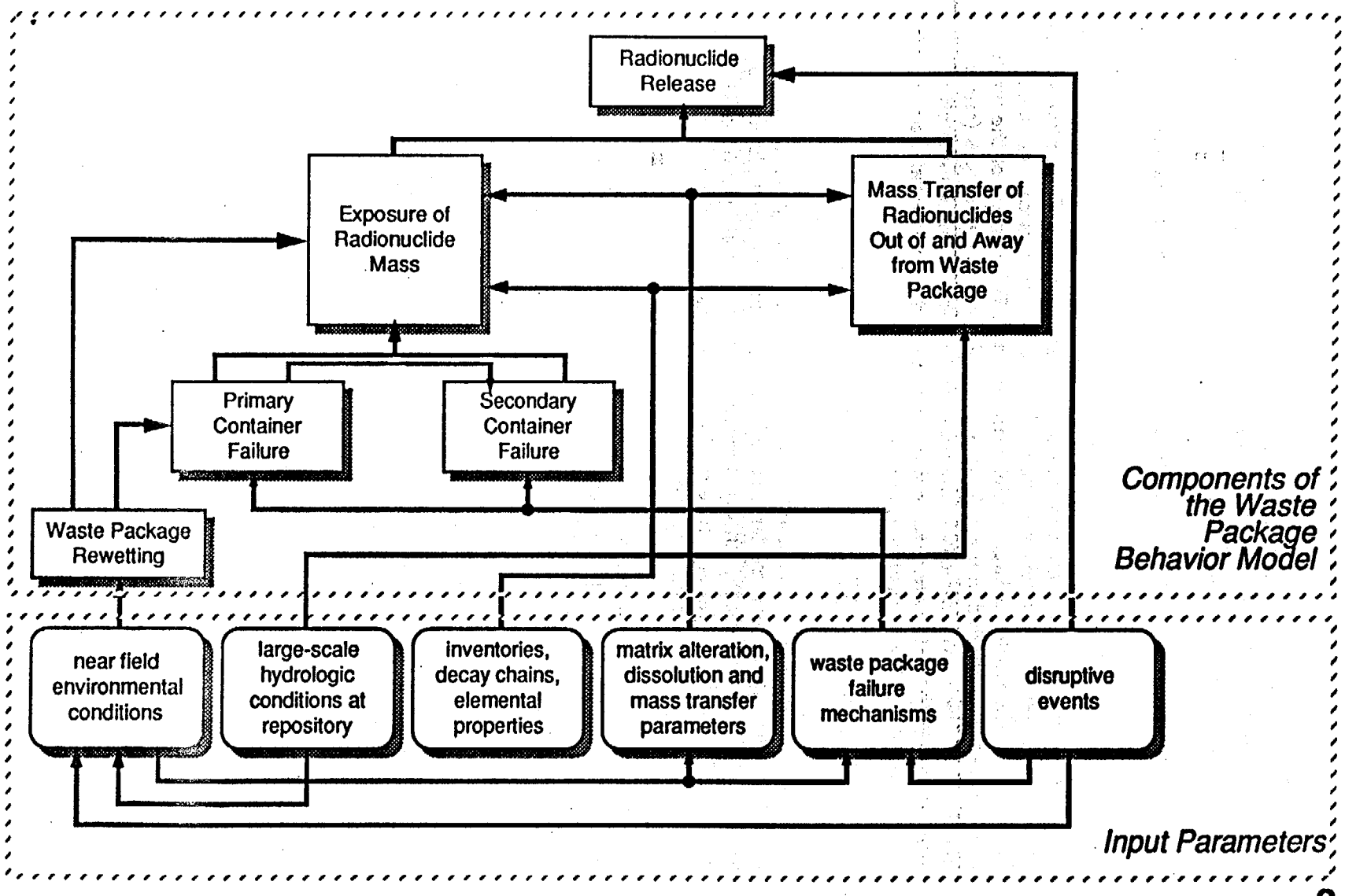


FIGURE 3-13

DETAILED WASTE PACKAGE INFLUENCE/INFORMATION FLOW SCHEMATIC

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DWG. NO. 24959 DATE 2/20/91 DRAWN TB

Golder Associates

Note that arrows connecting input parameters indicate that these parameters may be described as functions of other input parameters. For example, waste package failure mechanisms can be defined in terms of near field conditions. Likewise, disruptive events can affect near field conditions (which in turn may affect other input parameters).

3.5.1 Summary of Waste Package Computational Algorithm

A simplified summary of the computational algorithm for RIP, emphasizing the waste package portion of the model, is presented below in Figure 3-14.

3.6 List of Symbols

c(t) -	total primary container failure modes at time t [1/t]
$c_m(t)$ -	container failure rate for failure mode m at time t [1/t]
C(n) -	effective concentration at waste package for radionuclide n [M/L^3]
$C_s(n)$ -	saturation concentration of radionuclide n in water [M/L^3]
CATCH -	effective catchment area for a waste package [L^2]
CONTAC -	integer which is used to reference the water contact mode, each mode having been assigned a different value
$D_{eff}(n)$ -	effective diffusion coefficient for radionuclide n [L^2/t]
$D_m(n)$ -	moleculare diffusion coefficient for radionuclide n in water [L^2/t]
e(n,t) -	total exposure rate of radionuclide n for a group of waste packages at time t [M/t]
$e_f(n,t)$ -	exposure rate of radionuclide n in the free inventory for a group of waste packages at time t [M/t]
$e_g(n,t)$ -	exposure rate of radionuclide n in the gap inventory for a group of waste packages at time t [M/t]
$e_b(n,t)$ -	exposure rate of radionuclide n in the bound inventory for a group of waste packages at time t [M/t]
f_w -	fraction of waste matrix surface area in a failed container which is wet
$f_{w,m}$ -	fraction of waste matrix surface area which is wet due to failure mode m

Begin nr = 1, number of Monte Carlo realizations

realize all non-temporal system parameters
realize all disruptive events over time period of interest

Begin t = 1, number of time steps

increment elapsed time
compute time-dependent variables
if disruptive event occurs, realize consequences

Begin j = 1, number of environmental groups

evaluate all parameters that depend on environmental factors

Begin n = 1, number of radionuclides

compute exposure rate
compute maximum mass transfer rate
compute release rate
release mass to specified pathways
update amount of exposed, unreleased mass
decay exposed, unreleased mass

end n loop

end j loop

decay radionuclides in waste packages
simulate radionuclide transport through pathways for this time step

end t loop

compute performance parameters for current realization

end nr loop

FIGURE 3-
WASTE PACKAGE MOD
COMPUTATIONAL ALGORITHM
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 26630 DATE 10/1/91 DRAWN BY CB

Golder Associates

$f_{free}(n)$ - fraction of mass of radionuclide n in the "free" inventory

$f_{gap}(n)$ - fraction of mass of radionuclide n in the "gap" inventory

$g(t)$ - rate of secondary container failure in failed primary containers [1/t]

i - index referring to waste package type

$I_f(n,t)$ - free inventory of radionuclide n *per unfailed container* at time t (that portion located between the cladding and the container wall) [M]

$I_g(n,t)$ - gap inventory of radionuclide n *per unfailed container* at time t (that portion located in the gap between the fuel and the cladding) [M]

$I_b(n,t)$ - bound inventory of radionuclide n *per unfailed container* at time t (that portion bound in matrix) [M]

INFIL - repository level groundwater infiltration rate [L/t]

$k_{a/d}$ - combined fractional alteration/dissolution rate of waste matrix [1/t]

k_{air} - fractional air alteration rate of waste matrix [1/t]

k_{wat} - fractional aqueous alteration/dissolution rate of waste matrix [1/t]

$k_t(n,t)$ - maximum possible mass transfer rate of radionuclide n out of and away from a group of waste packages at time t [M/t]

$k_{tp}(n)$ - maximum possible mass transfer rate of aqueous radionuclide n out of and away from a single waste package [M/t]

$k_{tp,a}(n)$ - maximum possible advective mass transfer rate of aqueous radionuclide n out of and away from a single waste package [M/t]

$k_{tp,d}(n)$ - maximum possible diffusive mass transfer rate of aqueous radionuclide n out of and away from a single waste package [M/t]

$k_{tp,g}(n)$ - maximum possible mass transfer rate of gaseous radionuclide n out of and away from a single waste package [M/t]

$k_{tp,m}(n)$ - maximum possible mass transfer rate of aqueous radionuclide n out of and away from a single waste package due to container failure mode m [M/t]

m - index referring to container and/or cladding failure mode

$M(n,t)$ -	amount of exposed (available) mass of radionuclide n in a group of waste packages at time t [M]
$M_p(n)$ -	amount of exposed (available) mass of radionuclide n per waste package [M]
M_c -	mass of waste matrix per waste package [M]
$M_{uu}(t)$ -	mass of unprotected, unaltered undissolved matrix around a group of waste packages at time t [M]
NM -	total number of primary container failure modes
n -	index referring to radionuclide type
N_c -	number of waste packages
N_m -	cumulative number of waste packages which have failed by mode m
N_T -	total number of waste packages which have failed by any mode
$r(n,t)$ -	release rate of radionuclide n from a group of waste packages at time t [M/t]
$R_{age}(m)$ -	aging rate for container failure mode m
R_{dis} -	matrix aqueous alteration/dissolution rate [$M/L^2/t$]
S -	effective surface area of waste matrix in a failed container per unit mass [L^2/M]
t -	time [t]
T_r -	temperature at which a waste package rewets, returning to its long term moisture condition [T]
t_{wet}	time at which a group of waste packages rewets [t]
$TEMPAV_i(t)$ -	the average temperature at the edge of waste package type i at time t [T]
TEMPV -	parameter describing waste package to waste package temperature variability
$T_m(t)$ -	the mean temperature throughout the entire repository at time t [T]
$T_i(t)$ -	the incremental temperature above the mean repository temperature at the edge of waste package type i at time t [T]

V -	water volume in contact with matrix in failed waste package [L^3]
$w(t,\Theta)$ -	secondary container failure rate at time t for a primary container which has failed at time Θ [1/t]
Θ -	time of primary container failure [t]
τ -	dummy integration variable for time [t]
τ_D -	tortuosity for diffusive mass transfer
ω -	geometric factor for diffusive mass transfer [L]
η_e -	effective porosity

4. NEAR AND FAR FIELD RADIONUCLIDE TRANSPORT MODEL

4.1 Introduction

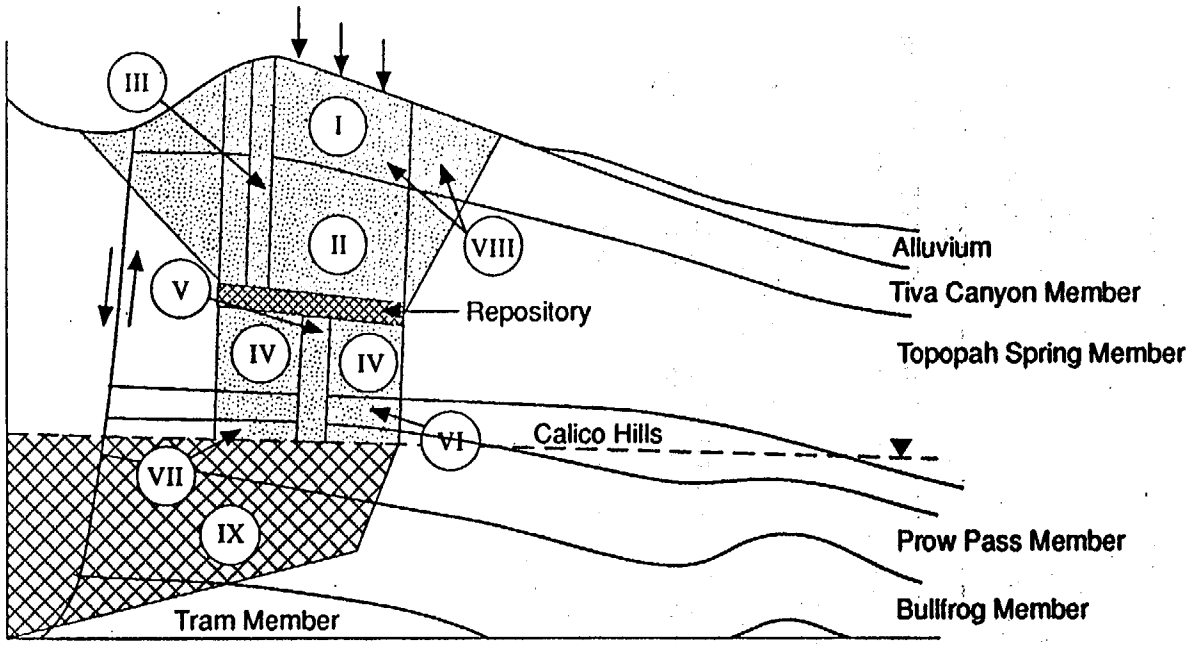
As discussed in Chapter 2, the integrated repository performance assessment model RIP is a complex radionuclide transport model, and consists primarily of a series of interconnected, fully coupled component models with input/output relationships for radionuclide transfer. The three major component models address waste package behavior and radionuclide release, transport pathways to the accessible environment, and ancillary processes such as disruptive events.

The component of the RIP model discussed in this chapter is responsible for calculations pertaining to the transport of radionuclides through the near and far-field to the accessible environment. The output from the waste package component model serves as input for the transport pathways component model. The general methodology and assumptions incorporated into the software are presented in detail. The RIP pathways component model is significantly different from other transport models which have been applied to performance assessment in that it attempts to describe rather than explain the controlling processes. Some of the rationale for choosing such an approach is discussed in Chapter 2.

The transport algorithm described here propagates radionuclide mass along a network of interconnected *pathways*, where a pathway is a distinct hydrologic feature in the physical domain. Radionuclide mass input to the pathways is obtained from the waste package release model discussed in Chapter 3.

The pathways and their linkages are defined by the RIP user, and the model is structured such that a tremendous amount of flexibility as to the complexity of the pathway network and specifications is allowed. The RIP pathways may be utilized to represent radionuclide transport and fluid flow in a variety of ways. Pathways may coexist in space, and can be used for fluid balance purposes, for transporting radionuclide mass (in the aqueous or the gaseous phase), or both.

Figure 4-1 shows a *hypothetical* pathway configuration applied to Yucca Mountain. This is not intended to represent the preferred conceptual model for the site, but is presented here simply to illustrate some of the basic features of the pathway algorithm. For example, pathway IV represents liquid phase transport in the Topopah Spring member, perhaps through the matrix and fractured portions of the strata, while pathway V represents a "short circuit" transport feature (e.g., saturated fracture flow), which traverses all strata between the repository and the water table. Both pathways IV and V coexist in space and represent liquid phase transport. Likewise, pathways III and VIII coexist in space over the repository. In this case, pathway III is only used for liquid flow balance purposes, while pathway VIII is used for gaseous phase mass transport. Note that the actual geometry and locations of the pathways are not rigidly defined in space, the *connectivities* between pathways being the important features.



Unsaturated Groundwater Flow

- (I) Tiva Canyon Member
- (II) Topopah Spring Member
- (III) Episodic Flow in Fractures

Gaseous Phase Transport

- (VIII) Region Above Repository

Flow and Transport

- (IV) Topopah Spring Member (Unsaturated Zone)
- (V) "Short Circuit Pathway" (Unsaturated Zone)
- (VI) Calico Hills (Unsaturated Zone)
- (VII) Prow Pass Member (Unsaturated Zone)
- (IX) Prow Pass and Bullfrog Member (Saturated Zone)

LEGEND

Water Table

FIGURE **4-1**

EXAMPLE PATHWAY CONFIGURATION FOR YUCCA MOUNTAIN

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1571.203 DRAWING NO. 26686 DATE 10/9/91 DRAWN BY EA

Golder Associates

Each pathway is defined by a number of parameters. In general, each of the parameters may be defined stochastically (as a probability distribution) by the user as a means to account for variability and uncertainty. RIP accounts for variability and uncertainty by probabilistic simulation (e.g., by Monte Carlo sampling of pathway lengths, volumetric flow rates, and cross-sectional areas). Parameters may be dependent on disruptive events and climate change. All dependencies are created by the user when defining the parameters.

The primary algorithm used in RIP to move radionuclide mass along a pathway is not based on balance laws of fluid flow. This departs from the normal case in hydrologic modeling of advective-dispersive transport. Although conventional continuum-based models have theoretical advantages, their ability to capture the heterogeneity of a physical domain is generally limited, and the computational work required for the solutions is a tremendous burden to a probabilistic simulation tool such as the RIP model. Given these limitations, the implementation of an adequately valid continuum-based model in the RIP code was precluded. For simple idealized systems, however, the RIP model does offer a continuum-based solution as an alternative transport algorithm (based on the one-dimensional advection-dispersion equation).

The primary pathways transport algorithm in RIP is based on a multi-state Markov process (Cox and Miller, 1965). This algorithm is similar to a random walk algorithm for particles, which has been implemented for diffusion models (e.g., Fischer et al., 1979). In the multi-state Markov process, the particles travel along a pathway in the direction of the pathway fluid flux, moving over randomly generated distances in the various states of the flow system. The states are referred to as *flow modes* and interpreted as the representation of local variation in the fluid flow field due to heterogeneity of the physical domain. The variation in the flow field is primarily represented by markedly different flow speeds among the flow modes. For example, flow in fractures and flow in the rock matrix may be two flow modes within a single pathway. The flow speed distribution for this pathway may resemble the distribution shown on Figure 4-2.

There are significant benefits to using the multi-state Markov process algorithm for radionuclide transport. One is the opportunity to represent multiple flow modes with a pathway, thus addressing local heterogeneity. This is very difficult to accomplish using continuum-based simulators. Another is the reduction in computer time which is achieved by the Markov algorithm in comparison to continuum models of transport. As stated elsewhere in this document, the limitations of computer speed are important to probabilistic simulation and reductions in complexity must be sought wherever reasonably possible.

Section 4.2 provides an overview of the computational algorithm used by the pathways model. The mathematical details of the algorithm are then presented in Section 4.3. Section 4.4 discusses the required input parameters and Section 4.5 discusses the general form of the output. A summary is provided in Section 4.6. Section 4.7 contains a list of symbols referred to in this chapter.

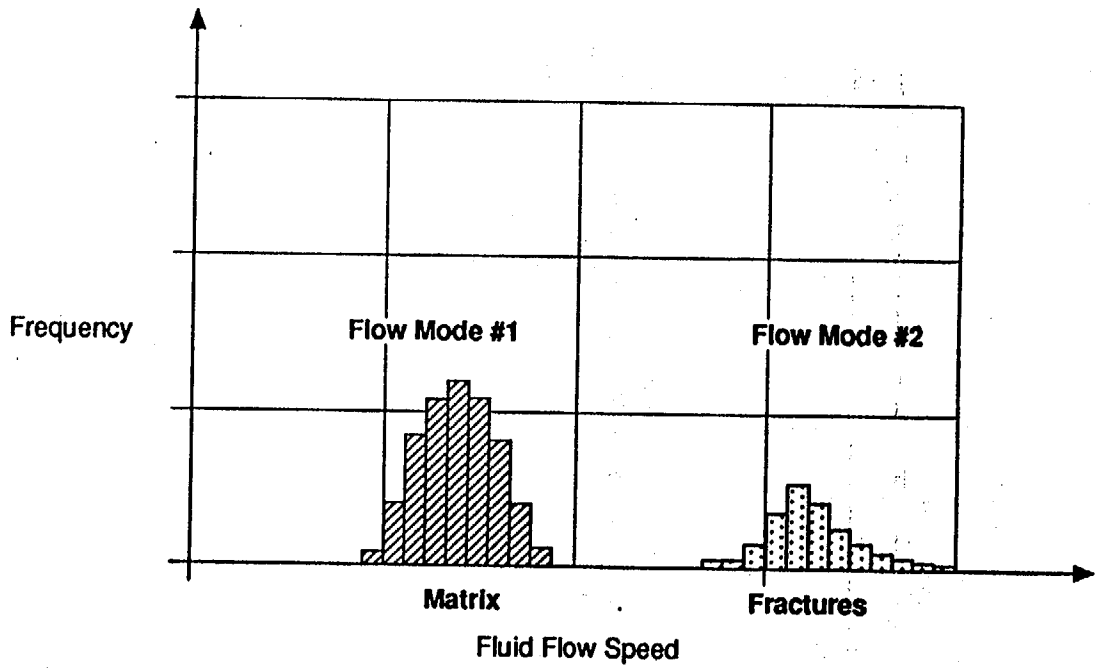


FIGURE **4-2**

HYPOTHETICAL FLUID FLOW FREQUENCY DISTRIBUTION FOR A PATHWAY

ARGONNE/MODEL D OPMENT

PROJECT NO. 371.203 DRAWING NO. 26605 DATE 9/27/91 DRAWN BY CB

Golder Associates

4.2 Overview of Transport Pathways Computational Algorithm

4.2.1 General Concepts of Radionuclide Transport in Pathways

The central feature of the transport model is the *pathway*. Pathways are linked successively to one another in a manner which allows for radionuclide transport from the waste packages to the accessible environment. The pathways may be arranged in series or parallel throughout the physical domain. Pathways may be used for describing flow balances as well as radionuclide transport (although, as will be discussed below in Section 4.4.2, RIP does not explicitly compute a water balance).

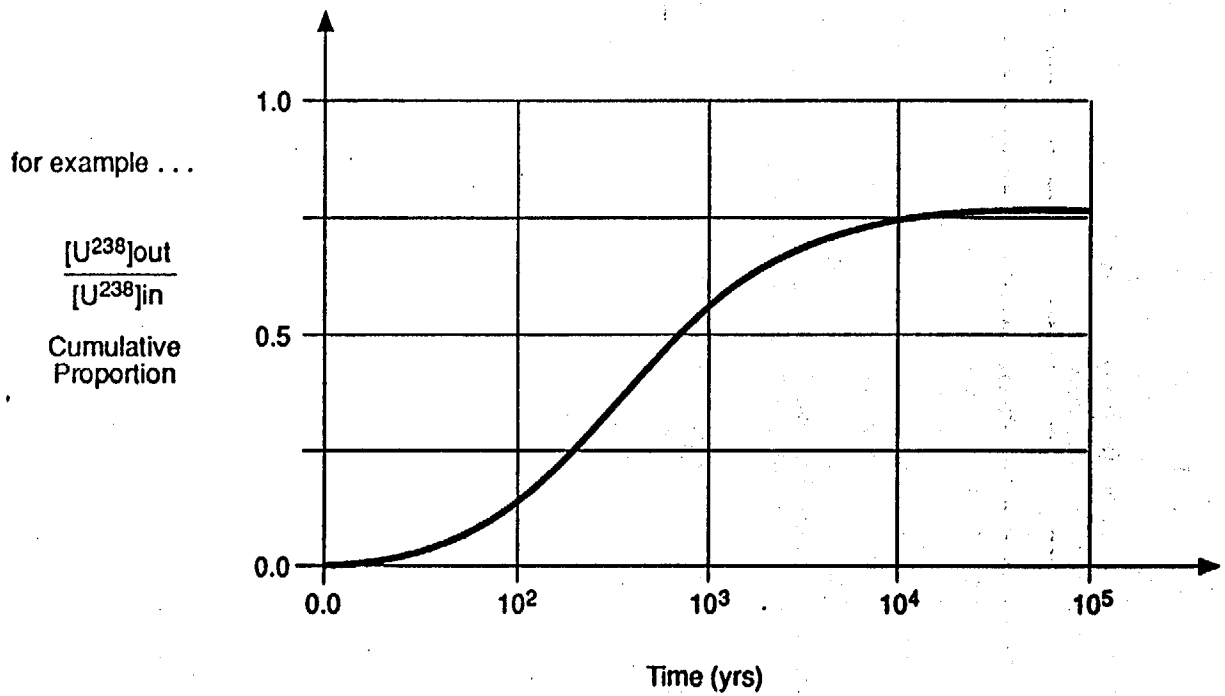
The use of more than one pathway to simulate transport in the physical domain is a means of addressing large scale heterogeneity in the hydrologic system (e.g. among different geologic media and geologic structures). Distinct pathways may be used to represent "short circuit structures" as well as stratigraphic layering. Within a pathway, the hydrologic heterogeneity may be further addressed by the use of *flow modes*. A flow mode is intended to model local heterogeneity within a pathway, such as flow in fractures versus flow in matrix portions of the rock mass. Each flow mode within a pathway has a defined set of hydrologic parameters.

The user may define pathways which transport radionuclide mass in either the gas or liquid phase by specifying the appropriate properties. Radionuclide exchange between phases is not possible within a single pathway, although the pathways may coexist in space. For the purposes of constructing a flow balance in the pathway system, it is not necessary that pathways transport radionuclides. Pathways may be constructed which do not receive mass, but interact with flow boundary conditions (e.g. precipitation), providing subsequent pathways with consistent volumetric flow rates.

Within a pathway, radionuclide transport is based on a *breakthrough curve*, as shown on Figure 4-3. The breakthrough curve indicates the proportion of mass released from the pathway to subsequent pathways as a function of time, based on a unit mass input at time zero (i.e, the time when mass first enters the pathway). For pathways consisting of more than one flow mode, the breakthrough curve is computed by a *multi-state Markov process algorithm*. Alternatively, in a single mode pathway, the breakthrough curve can be computed based on a one-dimensional advection-dispersion solution.

4.2.2 Major Features of the Breakthrough Curve Algorithm

The radionuclide travel times obtained from the breakthrough curve include the effects of multiple flow modes and retardation within the various flow modes. Breakthrough curves may be computed in two alternative ways: 1) based on the multi-state Markov process algorithm, or 2) based on a solution to the one-dimensional advection-dispersion equation for solute transport in groundwater. The latter solution method is an alternative for



Note: Cumulative proportion is less than one due to radioactive decay

FIGURE **4-3**

EXAMPLE BREAKTHROUGH CURVE FOR A RADIONUCLIDE

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 28542 DATE 2/20/92 DRAWN BY TB

Golder Associates

pathways in which the user defines only one flow mode. It requires input of a linear average velocity, the medium's dispersivity, and radionuclide retardation parameters.

The multi-state Markov process algorithm produces a probability distribution for particle travel time over a pathway. By definition, this distribution is the breakthrough curve for a slug input to the pathway. Hence, we use the terminology breakthrough curve and travel time probability distribution interchangeably elsewhere in the report.

In the mechanics of the Markov algorithm, particles travel a certain length in a flow mode and then transition to another flow mode. In the new flow mode they again travel a certain length prior to transition. This process is continued until the particle exits the pathway, as shown on Figure 4-4. The likelihood, or probability, for a particle to be in a specific flow mode is directly related to the proportion of the total flow for the pathway occurring in the flow mode. For example, if most of the flow in a pathway occurred through the rock matrix, there would be greater chance at any given time that a particle would be in the matrix than in other flow modes of the pathway.

The distribution of the lengths traveled by particles within a flow mode are modeled using a Poisson distribution. The Poisson distribution is widely used to describe natural and industrial phenomena. It is often utilized in Markov process analyses (Cox and Miller, 1965). The Poisson distribution is based on a rate parameter, referred to as the rate of occurrence. The Poisson distribution is used to describe the probability distribution for the number of occurrences of an event over an interval of either space or time. Applied to radionuclide transport, the rate parameter is the transition rate from one flow mode to another, per unit length of the pathway. It is a geometric property of the hydrologic system, accounting for connectivity and proximity of the various flow modes in a pathway. Under multiphase conditions, the Poisson rate parameter is also a function of the fluid saturation.

An equivalent way to represent a process which follows a Poisson distribution is to determine the length intervals (i.e., distance traveled) between transitions. This is actually what is done in a random simulation of transport along a pathway. These intervals are exponentially distributed with the exponential rate equal to the Poisson rate parameter. The inverse of the exponential distribution is used to randomly generate the length intervals between transitions. When used in the exponential distribution, the Poisson rate parameter is analogous to the decay constant for radioactive decay.

The Poisson rate parameter, although necessary input to the Markov process algorithm, can not be directly measured in field testing programs. However, information concerning the parameter may be obtained by considering the actual flow modes of a pathway and the natural limitations on how far a particle can travel in a mode before a transition must occur. For example, in cooling fractures of a volcanic tuff, the particle must transition from the fractures within a distance equal to the formation thickness. More complex analysis methods involving random simulation may also be used to evaluate the rate parameters. Dating of groundwater samples collected in various flow modes also may aid in determining the Poisson transition rates. In the RIP model, it is necessary to provide a rate parameter for only one mode in each pathway. Rate parameters for the remaining flow

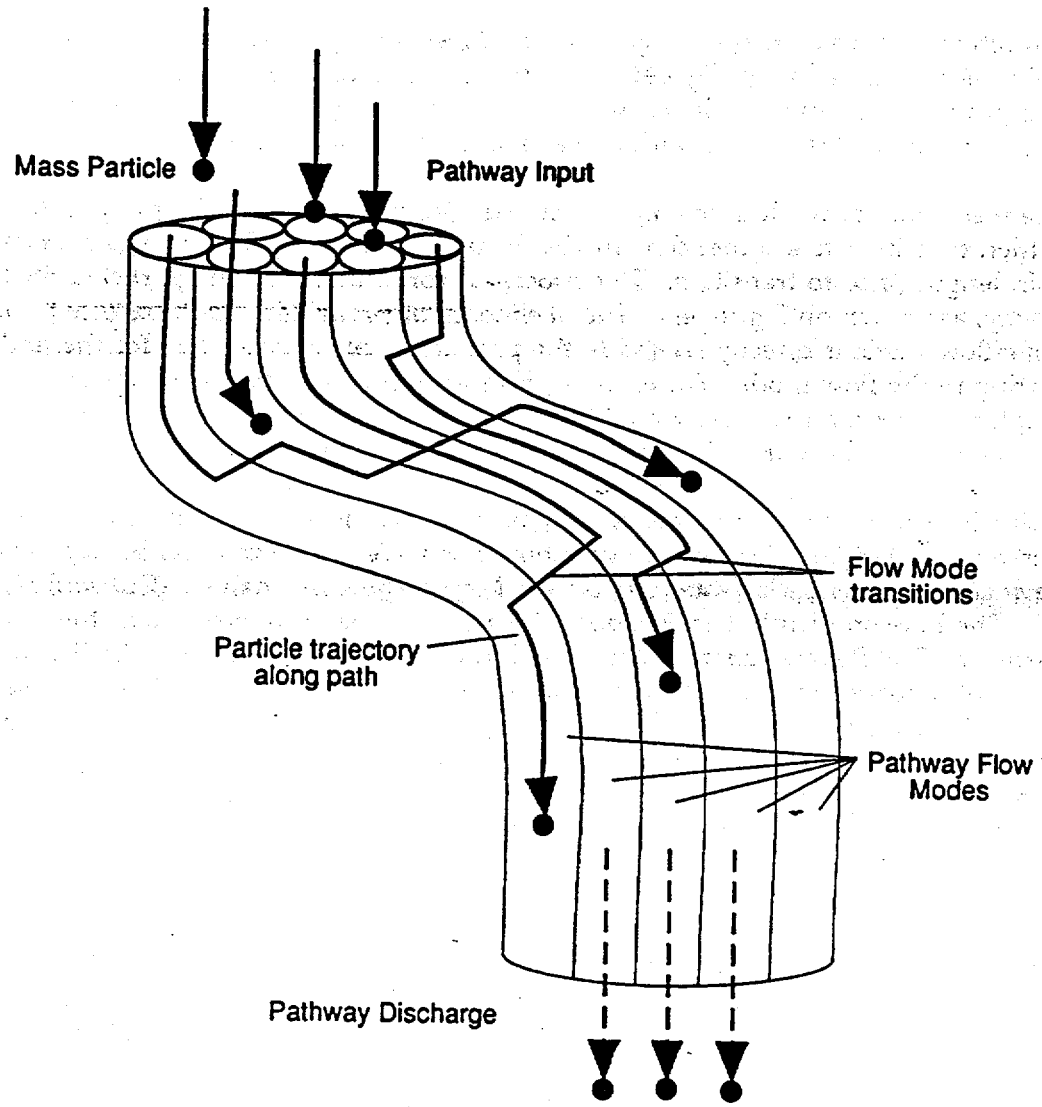


FIGURE 4-4
MARKOV PROCESS ALGORIT
FOR PARTICLE TRANSPO
ALONG A PATHWAY
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 26543 DATE 9/26/91 DRAWN BY EA

Golder Associates

modes are then internally computed based on a limited number of assumptions, which are discussed in Appendix G. Input parameters are discussed in more detail in Section 4.4.

4.2.3 Consequences of Disruptive Events for Radionuclide Transport

Disruptive events include anomalous events which may disturb waste packages at the repository level and/or radionuclide masses distributed in the near and far field regions of Yucca Mountain. In the RIP model, this list of events is reduced to those considered to be "significant", which means they actually may occur and if they do occur their impacts could affect the natural or engineered barriers of the repository. These events include seismic and volcanic activity, and human intrusion. The effects of these events on waste packages are discussed in Section 3.2.7. A detailed discussion concerning simulation of the events is provided in Chapter 5.

With respect to the pathways, there are three possible consequences of the disruptive events:

1) radionuclide mass may be moved from formerly intact waste packages into a pathway (e.g. to pathways in the saturated zone);

2) radionuclide mass may be moved from a pathway directly to the accessible environment; and

3) pathway parameter values may change (e.g. pathway length or flow mode hydraulic properties)

Consequences 1 and 2 are implemented by moving mass proportions from the sources to the targets. Data input determining how much mass is transferred from the sources to the targets is described in Chapter 5.

Consequence 3 applies to the following pathway parameters:

- pathway length,
- flow mode velocity or porosity,
- flow mode proportion of total fluid flow in the pathway, and
- flow mode Poisson transition rate.

The user must define the dependencies of these parameters on the disruptive events. These definitions may vary in complexity and may include dependencies which reflect ordinary climate change as well. For example, flow mode velocity may be defined as a function of a disruptive event and also of climatic parameters. The pathway length may be defined as a function of the water table elevation, which in turn may depend on disruptive events and climatic parameters.

Because changes in the pathway parameter values may cause additional calculations to be performed in the RIP model (due to dependencies), minimum, or threshold, changes are

required before the change is actually implemented in the simulation. These requirements reduce the computational work performed for a simulation. The minimum required change is defined by the user in terms of the fractional difference between a new value and the old value. In the RIP model this fractional difference is computed as follows:

$$\Delta p = \left|\frac{p^{new}}{p^{old}} - 1\right| \tag{4.1}$$

where:

Δp	is the fractional difference for parameter p,
p^{new}	is the new value for parameter p, and
p^{old}	is the old value for parameter p.

The fractional difference, Δp, is compared to the user specified tolerance value to identify if the change should be implemented in the simulation. If Δp does not exceed the tolerance value, p^{old} is not updated and the change does not take place.

In determining the tolerance values, the user must consider the importance of changes in the parameters. For example, because travel time is inversely proportional to velocity, a 1% change in the velocity of a flow mode will map approximately to a 1% change in the expected travel time along the pathway. In consideration of the uncertainty in the travel time for the pathway, the user must decide if the 1% change is significant.

The handling of consequence 3 in the pathway algorithm varies depending on the parameter value which undergoes a change. For any parameters which undergo significant change (i.e. the tolerance value for the parameter is exceeded by the change) the breakthrough curve for each radionuclide is recalculated for the pathway. A significant change in the pathway length parameter has additional ramifications depending on whether the length was reduced or increased. If the length is increased, only the breakthrough curve is recalculated. However, should the length be decreased, a quantity of mass for each radionuclide present in the pathway is released from the pathway. The mass is released to the defined discharge pathways and is released in fractions equal to the fractional change in pathway length. For example, if the pathway was shortened by 20%, then 20% of the mass would be released to the "downstream" pathway(s). Subsequent to this mass release, all mass currently resident in the pathway is lumped and redistributed according to the newly calculated breakthrough curve.

4.2.4 Summary of Computational Algorithm

Figure 4-5 provides a schematic of the computational algorithm for transport through a single pathway. Recall that the individual pathways are linked together into a network, so that output from one pathway is input for one or more additional pathways. Transport

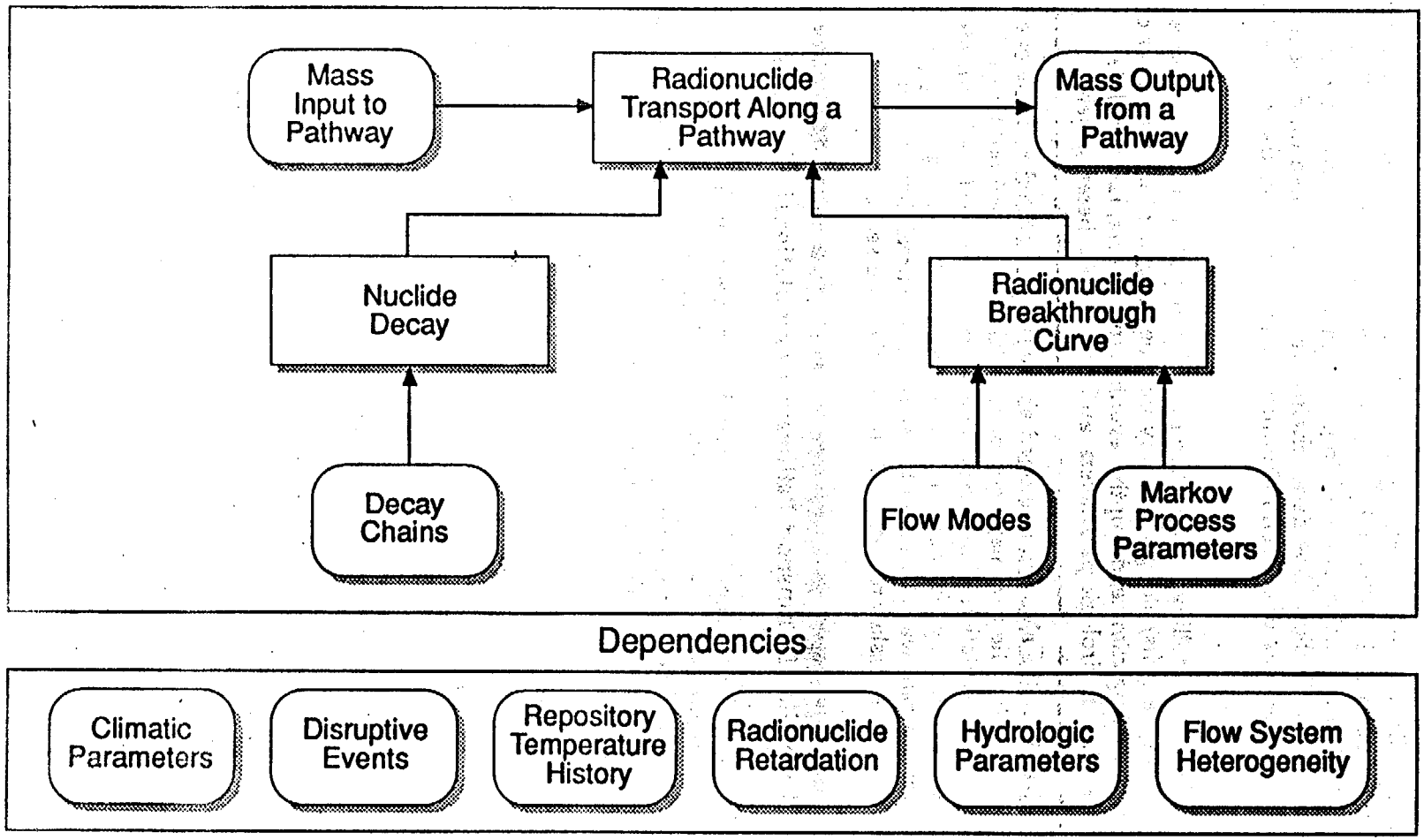


FIGURE **4-5**

OVERVIEW OF RADIONUCLIDE TRANSPORT PATHWAYS MODEL

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 26688 DATE 10/8/91 DRAWN BY EA

Golder Associates

times for individual radionuclides through a given pathway are controlled by a *breakthrough curve*. Each radionuclide in each pathway has a different breakthrough curve. The form of the breakthrough curve is a function of the various flow modes and is computed using a Markov process algorithm. Radionuclide decay chains are also accounted for directly, so that within a pathway, the mass of a given radionuclide may increase or decrease with time.

Note that the parameters controlling radionuclide transport (e.g., flow mode velocities, transition rates, pathway lengths) along a pathway may be dependent on a number of system parameters, including climate, disruptive events, temperature, radionuclide retardation characteristics, hydrologic parameters, and flow system heterogeneity.

Figure 4-6 summarizes how the algorithm outlined above is incorporated into the overall Monte Carlo scheme of RIP. The two major computational tasks carried out by the pathways model are 1) computation of the breakthrough curves for each radionuclide in each pathway (based on pathway and flow mode characteristics); and 2) propogation of mass through the various pathways based on the shape of the breakthrough curves. The mathematical details of these tasks are described in the following section.

4.3 Mathematical Details of Transport Pathways Computational Algorithm

The mathematical details of the computational algorithm are described in this section. This section describes the mechanics of how the algorithm described above is actually numerically implemented in the computer program. Although these details are of use to the serious RIP user, they are not critical to achieving a broad understanding of the algorithm and the reader may therefore wish to skip directly to Section 4.4.

Section 4.3.1 first discusses the manner in which mass is propagated through a pathway (given a radionuclide breakthrough curve). Section 4.3.2 then describes in detail how an actual breakthrough curve is computed (given pathway characteristics and a set of flow modes).

4.3.1 Radionuclide Mass Propagation Algorithm for Pathways Transport

Radionuclide mass transport within a pathway is based on simply evaluating (based on the breakthrough curve) what fractions of the total input to the pathway exit over different time periods. Recall that there is one breakthrough curve for each radionuclide. Hence, what is described below is actually carried out separately for each radionuclide.

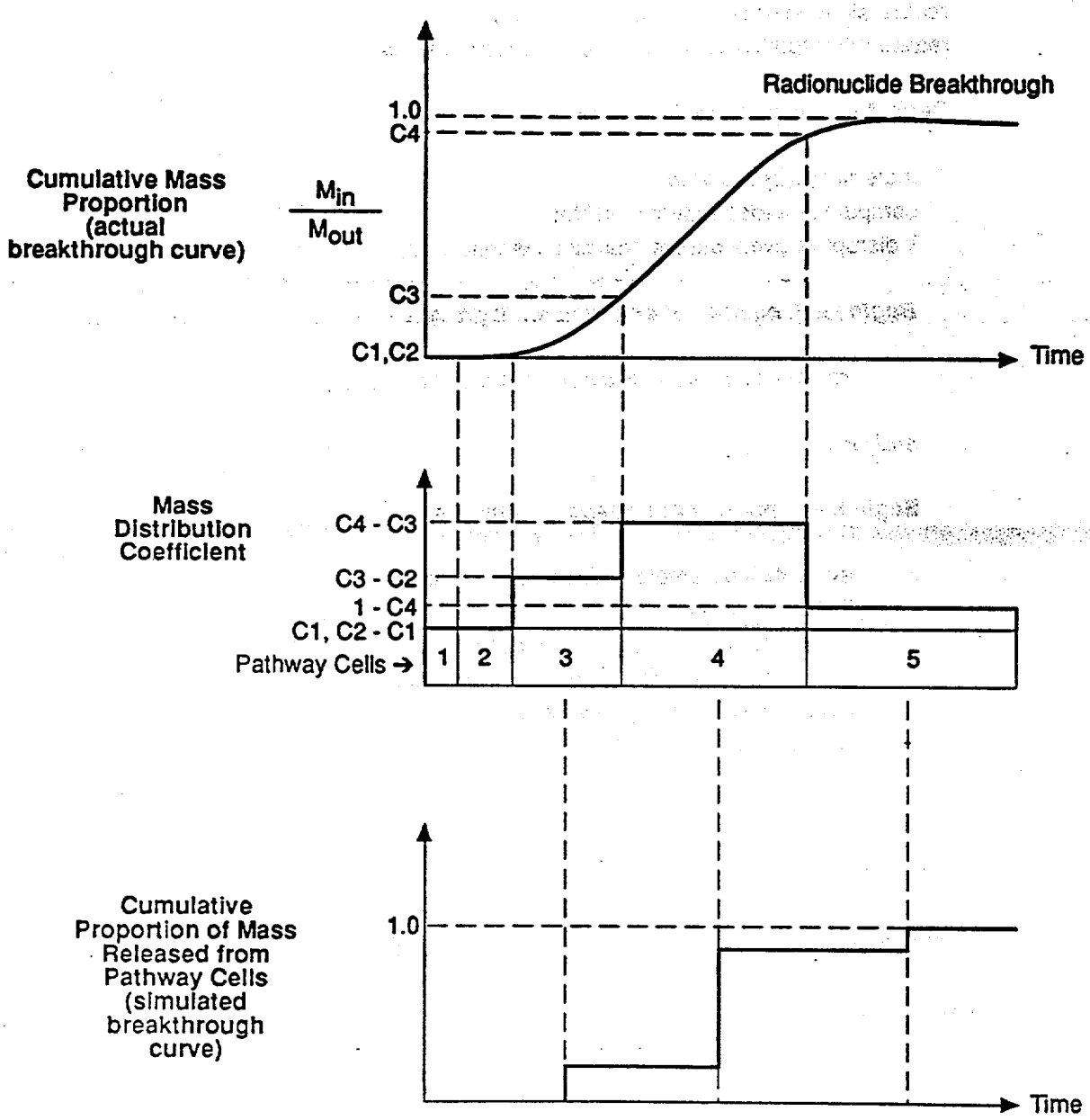
Algorithmically, the pathway is a single row of "cells" to which incoming mass is distributed in parallel. The mass remains in the cells for time periods based on the breakthrough curve and then is discharged, starting with the fastest cells. The mapping of the breakthrough curve onto the cells for a pathway is shown on Figure 4-7 and described below.

```
Begin nr = 1, number of Monte Carlo realizations

    realize all non-temporal system parameters
    realize all disruptive events over time period of interest

    Begin t = 1, number of time steps

        increment elapsed time
        compute time-dependent variables
        if disruptive event occurs, realize consequences

        Begin j = 1, number of environmental groups

            release mass from waste packages to pathways

        end j loop

        Begin k = 1, number of transport pathways

            evaluate if parameter values have changed
            if changed or t = 1 then
                compute breakthrough curve
                compute mass distribution coefficients
            decay mass in each pathway cell
            release mass to "downstream" pathway

        end k loop

    end t loop

    compute performance parameters for current realization

end nr loop
```

FIGURE 4-6
RADIONUCLIDE TRANSPORT MODEL
COMPUTATIONAL ALGORITHM
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 26690 DATE 10/8/91 DRAWN BY EA

Golder Associates



FIGURE

MAPPING OF BREAKTHROUGH CURVE ONTO PATHWAY CELLS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 26544 DATE 2/20/92 DRAWN BY TB

Golder Associates

To determine when mass should be discharged from a cell, each cell has a computed *residence time* and a *release time*. For a sequence of cells beginning with cell 1, the residence times are given by the following:

$$\delta t_i = f^{i-1} \Delta t \tag{4.2}$$

where:

δt_i is the residence time for cell i,

Δt is the RIP time step interval (e.g. 100 years), and

f is a factor determining the geometric rate of increase of δt_i.

Thus, cell 1 has a residence time of Δt, cell 2 has a residence time of $f\Delta t$, cell 3 has a residence time of $f^2\Delta t$, and so on. In the RIP model, the last cell always has a residence time which exceeds the total simulation time. In our preliminary analysis of this discretization method, we have assigned f a value of 2, which produced generally accurate results. These results are discussed in Appendix H. That is, Equation 4.2 becomes

$$\delta t_i = 2^{i-1} \Delta t \tag{4.3}$$

The release time for a cell is computed as a function of the residence time. The release time for a cell is the time at which mass in the cell will be released from the pathway. The release time is initialized the first time mass enters the pathway. The first release time for a cell is given by:

$$t_i = t_{initial} + \delta t_i \tag{4.4}$$

where:

t_i is the release time for cell i,

$t_{initial}$ is the simulation time when mass first enters the cell, and

δt_i is the residence time for cell i.

When the simulation time exceeds the release time for a cell, mass is released from the cell and the release time is incremented by δt_i. Thus, the mass in a cell is transported through the pathway in a time less than or equal to the residence time for the cell.

This will become clearer by considering an example. If the time step interval was 100 years (and assuming a value of f in Equation 4.2 of 2), the residence time would be 100 years for cell 1, 200 years for cell 2, 400 years for cell 3, and so on. Let's further assume that mass first enters the pathway at time = 100 years, and enters continuously (i.e., every time step

thereafter). Given this information, we will examine the behavior of, say, cell 3. The initial release time for cell 3 would be 100 + 400 = 500 years. Between time 100 and 500 mass will be distributed every timestep to cell 3, where it will accumulate (the fraction of the total mass distributed to a particular cell each step will be discussed subsequently). At time = 500, all the mass in cell 3 will be discharged to subsequent pathways, and the release time for cell 3 will be reset to 500 + 400 = 900 years. Between time 500 and 900 mass will again be input every timestep to cell 3 and will accumulate. At time = 900, all the mass in cell 3 will be discharged, and the release time for cell 3 will be reset to 900 + 400 = 1300. The process will repeat itself (for all cells in all pathways) until the end of the simulation (e.g., at 10,000 years).

The fraction of the total mass distributed to a particular cell is specific to each radionuclide and is based on the mapping of the breakthrough curve onto the time interval spanned by the cell. The time interval spanned by a cell is centered about the cell residence time. The right endpoint for the time interval of cell j is given by the following:

$$t_{r,j} = 2\,\delta t_j - t_{l,j} \tag{4.5}$$

where:

$t_{r,j}$ is the right endpoint for the time interval about cell j,

δt_j is the residence time for cell j, and

$t_{l,j}$ is the left endpoint for the time interval about cell j.

Note that $t_{r,j}$ is equal to $t_{l,j+1}$ which allows us to compute $t_{r,j+1}$ and so on, from Equation 4.5. Equation 4.5 cannot used for the first cell, and the value $t_{r,1}$ is set equal to Δt as a starting point for computing the remaining interval endpoints.

Referring again to the example discussed previously, the left and right endpoints for the time interval about the first three cells would be as follows:

j	$t_{l,j}$	$t_{r,j}$
1	0	100
2	100	300
3	300	500

The mapping of the breakthrough curve onto the cell produces a mass distribution coefficient which is the proportion of mass for a specific radionuclide allocated to the cell (see Figure 4-7). The mass distribution coefficient for the ith cell and nth radionuclide is given by the following:

$$w_{in} = F_n(t_{r,i}) - F_n(t_{l,i}) \tag{4.6}$$

where:

w_{in} is the mass distribution coefficient for cell i and radionuclide n,

$F_n(t)$ is the breakthrough curve value at time t for radionuclide n,

$t_{l,i}$ is the left endpoint for the time interval about cell i, and

$t_{r,i}$ is the right endpoint for the time interval about cell i.

While radionuclides reside in a pathway cell, they undergo radioactive decay and daughter-product generation. This process is simulated at every time step. The radioactive decay algorithm is discussed in detail in Appendix D.

In the RIP model, the radionuclide masses in each cell are simultaneously decayed and decay series are accounted for. However, the daughter product generation for a time step is *not redistributed* to the pathway cells using the mass distribution coefficients until they are released from the cell of the parent. That is, daughters products remain in the cell of the parent until the release time and are not distributed. This incurs some error for those daughter products which have significantly different retardation characteristics than the parent.

If during a simulation transient events should change the shape of the breakthrough curve for a radionuclide, the mass distribution coefficients are recomputed. The mass existing in the pathway cells at this time is generally not adjusted for the new mass distribution coefficients. All subsequent mass entering the pathway and the mass generated by radioactive decay, however, is distributed using the newly computed coefficients. In the case where the pathway length is decreased (eg., due to a rising water table), mass contained in the cells is adjusted accordingly and redistributed to the pathway. This redistribution is discussed in Section 4.3.3.

4.3.2 Computing the Radionuclide Breakthrough Curve for a Pathway

The two methods available in the RIP model for computing breakthrough curves are 1) the Markov process algorithm, and 2) the solution for the advection-dispersion equation. The latter solution method can only be used for pathways in which only a single flow mode is defined and is discussed in detail in Bear (1979, page 266). It is derived for one-dimensional transport in an infinite column with uniform flow of a tracer slug.

The remainder of this section describes how breakthrough curves are computed using the Markov process algorithm. The RIP model actually implements an approximation to this algorithm. We explain this approximation by first describing the Markov process and then explaining the approximation implemented in RIP.

In the Markov process algorithm, the breakthrough curve for a radionuclide is a function of pathway and flow mode parameters. This is also true for the approximation

implemented in RIP. These parameters may be functions of time, causing the breakthrough curve to also change with time during a simulation. The specific parameters on which the breakthrough curves depend are the pathway length, the proportions of flow occurring in each flow mode, the fluid velocities of the flow modes, the Poisson rate parameters, and the radionuclide retardation parameters. Currently, with the exception of the retardation parameters, all of these parameters may be transient in the RIP model.

The Markov process algorithm has two steps carried out repeatedly to simulate transport of a particle along a pathway. Step 1 consists of assigning the particle to a new flow mode. This assignment is a function of the proportions of flow occurring in the modes. Step 2 consists of propagating the particle a random length in the flow mode based on the Poisson rate parameter. In a random simulation, the outcome of step 1 is obtained by inverting the following conditional probability distribution:

$$\Pr(\text{ flow mode } j \mid \text{flow mode } i\ ) = f_{Q_j}/(1-f_{Q_i}) \tag{4.7}$$

where:

Pr(flow mode j | flow mode i)	is the probability of transition to flow mode j if previously in flow mode i, and
f_{Qi}	is the proportion of the total fluid flow occurring in the ith flow mode.

The relation between probability and flow mode is an assumption, which is only valid if the proportions of flow in the flow modes remain constant over all infinitesimal lengths along the pathway (i.e., anywhere along a pathway, the same flow proportions would be obtained for the various flow modes).

Once assigned to a flow mode, the random length traveled by a particle is given by:

$$l = -\frac{1}{\lambda_j}\ln[r(0,1)] \tag{4.8}$$

where:

l	is the random length interval,
λ_j	is the Poisson transition rate for flow mode j, and
$r(0,1)$	is a random number between 0 and 1.

Equation 4.8 is based on the inverse of the exponential probability distribution (Benjamin and Cornell, 1970).

Given average linear fluid velocities for the flow modes, the random length traveled by a particle in a flow mode is converted to a travel time. The total travel time over the pathway is the sum of all travel times along the individual flow mode intervals. By repeating the particle simulation procedure a large number of times, we can numerically approximate the probability distribution for travel time through the pathway. This distribution may be directly equated to the breakthrough curve.

Figure 4-8 illustrates the form of the breakthrough curve for a three-mode system. Note that no particles can travel faster than the fastest mode, nor slower than the slowest. Also, some particles may cover the entire path length in a single mode, resulting in "steps" in the breakthrough curve.

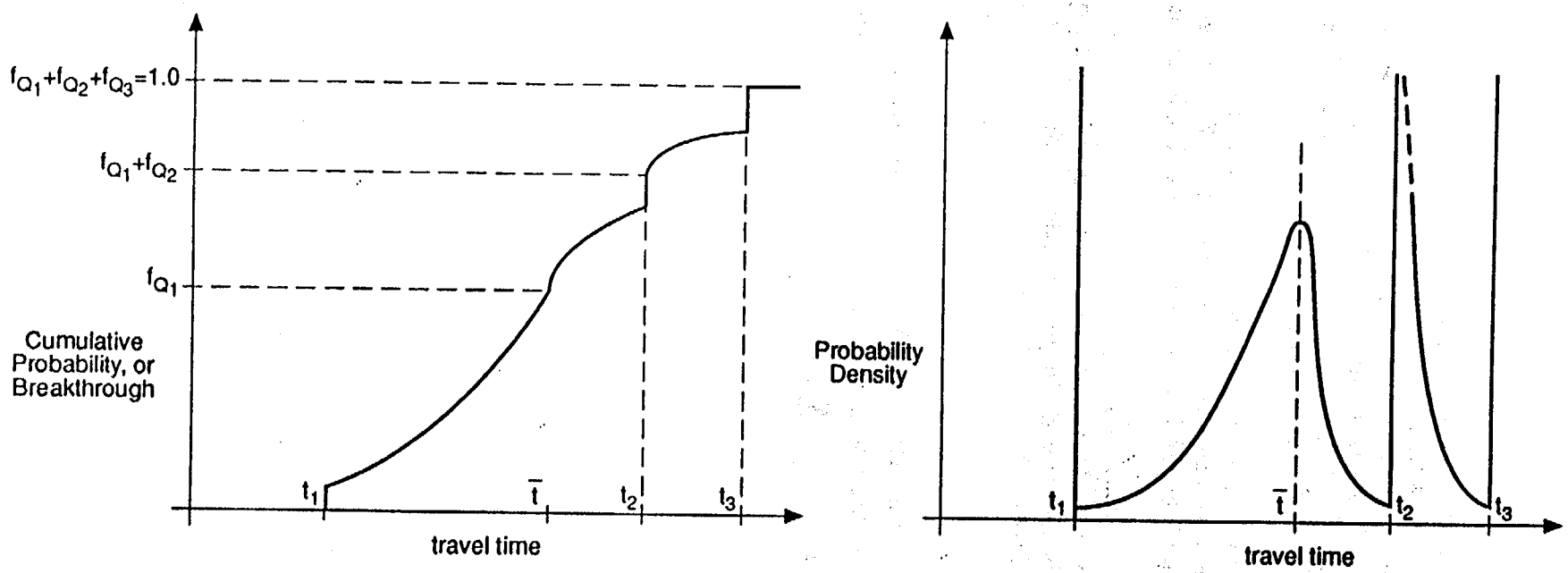
In the RIP model, the approximation to the Markov process algorithm reduces the quantity of computational work which must be performed. In the approximate Markov process algorithm, particles travel only their first random length in an individual flow mode. Subsequently, any remaining length of the pathway is traveled at the expected (average) travel time over all modes.

Based on the approximation, the travel time for a particle is given by the following:

$$t_n = \left[l_i / V_i + (L_p - l_i) \sum_{j}^{N} f_{Q_j} / V_j \right] R_n \tag{4.9}$$

where:

t_n	is the travel time for radionuclide n,
l_i	is the initial length interval traveled in flow mode i,
V_i	is the average linear fluid velocity for flow mode i,
L_p	is the pathway length,
N	is the number of flow modes in the pathway,
f_{Qj}	is the proportion of the total fluid flow occurring in pathway j,
V_j	is the average linear fluid velocity for flow mode j, and
R_n	is the retardation parameter for radionuclide n.



3-Flow Modes

t_i = flow mode travel time (i=1,2,3)

$\bar{t}$ = expected travel time

FIGURE **4-8**

APPROXIMATE MARKOV PROCESS ALGORITHM EXAMPLE BREAKTHROUGH AND DENSITY PLOTS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 [illegible].203 DRAWING NO. 26631 DATE 10/1/91 DRAWN BY CB

Golder Associates

The first term on the right side of Equation 4.9 is the travel time along the initial length interval of flow mode i. The second term on the right side of Equation 4.9 is the travel time over the remaining length at the expected travel time.

In the approximate Markov process algorithm, some fraction of particles may travel the entire path length in each individual flow mode (as illustrated by the "spikes" in the probability density function in Figure 4-8). Equation 4.9 is formulated such that it can directly account for the fraction of particles which may traverse the pathway relatively quickly, in comparison to the average travel time.

The benefit of the approximate Markov process algorithm is that the breakthrough curve has a closed form solution, thus, direct calculation of breakthrough fractions for each cell can be achieved.

This direct calculation for the breakthrough curve is as follows:

$$F_n(t) = \sum_{j}^{N} f_{Q_j} J_{jn}(t) \tag{4.10}$$

where:

$F_n(t)$	is the probability for radionuclide n to have a travel time less than t,
N	is the number of flow modes in the pathway,
f_{Qj}	is the proportion of flow in the jth flow mode, and
$J_{jn}(t)$	is the probability for radionuclide n after starting in flow mode j to have a travel time less than t.

Equation 4.10 is simply the sum over the pathway flow modes of the intersection of two events. The first event is that a particle started traveling in mode j and has probability f_{Qj}. The second event is that the particle's travel time was less than t and has probability $J_{jn}(t)$.

The mode probability distribution, $J_{jn}(t)$, is discontinuous, requiring several decision steps dependent upon the flow mode fluid velocity relative to the pathway average fluid velocity. For a given travel time, t, the mode probability distribution is given by:

$$J_{jn}(t) = \begin{bmatrix} 0, & \text{for} & t \le \frac{L_p}{V_j} R_n < \bar{t} \\ 1-F_j(l), & \text{for} & \frac{L_p}{V_j} R_n < t < \bar{t} \\ 1, & \text{for} & \frac{L_p}{V_j} R_n < \bar{t} \le t \\ 0, & \text{for} & t \le \bar{t} < \frac{L_p}{V_j} R_n \\ F_j(l), & \text{for} & \bar{t} < t < \frac{L_p}{V_j} R_n \\ 1, & \text{for} & \bar{t} < \frac{L_p}{V_j} R_n \le t \end{bmatrix} \tag{4.11}$$

where:

L_p	is the pathway length,
V_j	is the fluid velocity for the jth flow mode,
R_n	is the retardation parameter for radionuclide n,
$\bar{t}$	is the average travel time for the pathway, and
$F_j(l)$	is the exponential probability distribution for length intervals traveled in flow mode j

The average travel time for the pathway is computed from the following:

$$\bar{t} = L_p \sum_{j}^{N} \frac{f_{Q_j}}{V_j} \tag{4.12}$$

The length parameter, l, which is an argument to $F_j(l)$ in Eq. 4.11, is obtained by solving Eq. 4.9 for l_i. This solution is given by:

$$l = L_p \frac{t-\bar{t}}{\left[\frac{L_p}{V_j} R_n - \bar{t}\right]} \tag{4.13}$$

where the parameters in both Eq. 4.12 and Eq. 4.13 are defined in Eq. 4.11. The first three conditions in Eq. 4.11 apply to flow modes in which the travel time for a particle remaining

in the flow mode for the entire pathway ($L_p/V_j R_n$) is less than the pathway average travel time ($\bar{t}$). The last three conditions apply to flow modes for which the individual mode travel time is greater than the average travel time.

Comparisons of the approximate and exact Markov process algorithms are presented in Appendix I. The two methods were visually compared for selected sets of parameters. The comparisons were generally favorable, although, in certain cases considerable error occurred. The range of the error was generally within a factor of 1 to 3 and included both under- and over-estimation of the travel time by the approximate method. If the pathway and flow mode parameters have high uncertainty, this error will not add significantly to the total error and will cause no harm to the analysis. If uncertainty in the pathway and flow mode parameters is low (which will typically be very unlikely), use of the approximate method will have greater impact on the analysis and may be inadequate for the simulations.

4.4 Defining Input Parameters for the Transport Algorithm

4.4.1 General Description of Required Input Parameters

This section provides general documentation on the meaning of input parameters used to define radionuclide transport. The definition of stochastic and dependent parameters using the Parameters Module was briefly discussed in Chapter 1. Additional documentation is provided in the RIP User's Guide.

There are two categories for input to the pathways component of RIP:

1) tolerance values for transient parameters, and
2) pathway parameters.

Within the pathway parameters data category, there are three sub-categories. These are: 1) bulk pathway parameters, 2) pathway discharge parameters, and 3) pathway flow mode parameters. These data categories are discussed below.

Tolerance Values

The use of tolerance values in the RIP model is discussed in Section 4.3.3. Tolerance values apply to the minimum fractional change required for a parameter value before dependencies of the parameter are recalculated. Thus, although a parameter may change, other dependent parameters will not be recalculated until the magnitude of the change exceeds the tolerance value. This concept is used to reduce computational work during a simulation. The tolerance values specified by users remain constant for each simulation.

There are five parameters for which tolerance values must be specified. The tolerance values are applied to these parameters in all pathways and all flow modes. The tolerance values requested by the RIP model are the following:

- pathway length tolerance
- flow proportion tolerance
- fluid velocity tolerance
- porosity tolerance
- Poisson transition rate tolerance

Only the first parameter, pathway length, applies to the pathway as a whole. The remaining four parameters are specific to the flow modes of a pathway. These latter parameters may be dependent on higher level parameters, such as the pathway volumetric flow rate and the pathway area. Thus, tolerance values specified for these lower level parameters could implicitly specify tolerance values for the higher level parameters.

Pathway Parameters

Each pathway is defined with general parameters, discharge parameters, and flow mode parameters. There is considerable flexibility for the user in defining the parameters. The user may define dependencies upon parameters outside the pathway (e.g. in other pathways) and among the parameters within the pathway. For example, the pathway total volumetric flow may be a function of the volumetric flow in another pathway. As another example, the proportion of flow in a flow mode may be a function of the pathway's total volumetric flow.

Pathway Bulk Parameters

The following pathway parameters must be entered for each pathway defined:

- pathway identification
- pathway description
- pathway length (L)
- pathway total volumetric flow (L^3/t)
- pathway total area (L^2)

The pathway identification and description are constant values. The description is for user reference. The identification may be used elsewhere, such as when defining the discharge of a subsequent pathway. The pathway length, total volumetric flow, and total area all have the potential to be either constant, stochastic (i.e. random variables), or functions of other defined parameters, including the simulation time.

Pathway Discharge Parameters

Each pathway may discharge radionuclide mass to other pathways or to the accessible environment. If a pathway discharges to the accessible environment, it cannot discharge to

other pathways as well. There are two parameters per discharge pathway which must be entered. These parameters are the following:

- discharge pathway identification
- mass fraction of remaining mass discharged to the pathway

The pathway identification is the user-defined identification for the target pathway. The mass fraction of remaining mass determines the proportion of mass which will be discharged from the current pathway to the target pathway. Because this proportion is determined based on the remaining (unallocated) mass, the order in which discharge pathways are defined is important. This method of allocating mass to a pathway facilitates defining the parameter as a stochastic variable. The mass fraction parameter may be either constant, stochastic, or a function of other parameters.

Flow Mode Parameters

Flow modes are used to define local heterogeneity in the fluid flow field for a pathway. There may be an unlimited number of flow modes for a pathway. Each flow mode is defined by several parameters. At this level of parameter definition, it is likely the user will also need to use additional user-defined parameters which are not directly solicited by RIP. For example, the fluid velocity may be a function of saturation, hydraulic conductivity, and volumetric flow rate. Both saturation and hydraulic conductivity are not mandatory input for the RIP model and would have to be defined by the user in order to construct the dependency. Chapter 2 presents more detailed information concerning the definition of "sub-models" such as these.

For each flow mode defined in a pathway, the RIP model requests the following input parameters:

- flow mode description
- fraction of pathway total volumetric flow rate,
- velocity or porosity (user preference) (L/t or [])
- Poisson transition rate (see below) (L^{-1})
- medium dispersivity (L)
- radionuclide retardation parameters

The flow mode description is a text field for user definition of the flow mode. The remaining parameters, except for retardation parameters, may be entered as constant, stochastic, or functions of parameters. It is only possible to enter either the velocity or porosity because the two parameters are dependent on one another as follows:

$$v_i = \frac{f_{Q_i} Q}{A n_i} \tag{4.14}$$

where:

v_i	is the average linear fluid velocity for the ith flow mode,
f_{Qi}	is the fraction of the pathway volumetric flow rate occurring in the flow mode i,
Q	is the total volumetric flow rate for the kth pathway,
A	is the total area normal to the flow direction for the kth pathway, and
n_i	is the effective porosity for the ith flow mode.

The Poisson transition rate can be entered for only one flow mode in a pathway. The remaining transition rates are computed as discussed in Appendix G based on the given transition rate and the fraction of flow in each mode. The medium dispersivity for a flow mode is not used by the approximate Markov process algorithm. However, if a single flow mode is defined, this parameter is used in the hydrodynamic dispersion term of the one-dimensional advection-dispersion equation (Bear, 1979, page 268). Neglecting molecular diffusion, this relation is given by:

$$D_{H_i} = |V_i| \alpha_i \tag{4.15}$$

where:

D_{Hi}	is the hydrodynamic dispersion coefficient for flow mode i,
V_i	is the average linear fluid velocity of the ith flow mode, and
α_i	is the longitudinal dispersivity for the ith flow mode.

Radionuclide retardation parameters are solicited on a group basis for each flow mode of a pathway. The radionuclide groups consist of suites of radionuclides which are chemically similar with regard to retardation. The group classification simplifies the user input and is defined in the Waste Package module of the RIP model. Each radionuclide group may have a sorption and matrix diffusion retardation parameter. These parameters are combined into a single retardation value as shown in Appendix J. The combined retardation value remains constant for the simulation.

4.4.2 Managing the Total Flux and Repository Flux

Flow of groundwater through Yucca Mountain is a variable of major inquiry among project scientists. The long-term release of radionuclides to the accessible environment will be strongly dependent on this flow. The actual quantity of flow contacting waste at the repository level is governed by the quantity of flow available and the hydraulic properties of the intervening geologic media. Likewise, the flow available to transfer mass from the repository to the accessible environment is governed by these same parameters.

The RIP model does not explicitly compute a water balance for the simulation domain. Rather, the user must define this balance through the specification of pathway total volumetric flow rates. These rates must be specified with dependencies on one another such that flow at the repository level and elsewhere in the simulation domain is consistent with the flow volumes entering the domain at ground surface (presumably this is the only source of liquid water to the unsaturated zone of Yucca Mountain). A similar approach must be taken for pathways representing the saturated zone below the repository.

In constructing the flow balance, users may wish to define pathways which do not transport radionuclide masses (e.g., pathway III in Figure 4-1). Pathways of this nature may be defined between ground surface and the repository level. Such pathways could include a precipitation model and account for the hydraulic properties of the intervening strata. At the repository level, the volumetric flow could be used to define the repository level groundwater infiltration rate, which is used to determine the advective radionuclide mass transfer rate away from a failed waste package, as discussed in Chapter 3.

4.5 Pathways Model Output

The output of the pathways component model essentially consists of probabilistic waste package performance measures (e.g., cumulative release *to* the AE or *from* a specific pathway over time period of interest). These measures can be examined for individual radionuclides or summed over all radionuclides and all groups. Time histories of radionuclide release can also be produced.

The RIP User's Guide discusses the details of the various output options and how the data can be processed and graphically displayed (e.g., as a PDF, CDF or CCDF). As discussed in that document, the values of *all the input parameters for every realization* are also saved. This allows extensive sensitivity analyses to be carried out in order to identify relationships between parameter values and performance.

The structure of the RIP model is such that the pathways component model can be run independently of the waste package component model (by assuming a specified waste package release), producing the type of results outlined above. When the entire integrated model is run, additional outputs pertaining to the waste package component model are also produced (e.g., cumulative release from the waste packages). These outputs are discussed in Chapter 3.

4.6 Summary

The near and far field radionuclide transport algorithm for the RIP model is based on a network of user defined pathways. The pathways of RIP reflect major features of the hydrologic system. Pathways may represent large scale geologic structures, such as faults, and formation scale stratigraphy, as is necessary to account for the large scale heterogeneity of the system. Within a pathway, local heterogeneity may be addressed through user defined flow modes. The flow modes are primarily distinguished from one another based on flow velocity in the mode. However, flow modes may also conduct varying proportions of the total pathway flux, and realize different radionuclide retardation parameters.

Within a single pathway, radionuclide transport is based on a breakthrough curve (i.e., the cumulative proportion of particles to traverse the pathway within a given time). The breakthrough curve for a pathway combines the effects of all flow modes and retardation parameters on the radionuclide travel time. It is developed based on the Markov process, which is very similar to a random walk. The Markov process was favored for the RIP model over continuum based models of transport because it is able to capture the effects of local heterogeneity in the fluid flow field at a reasonable level of computational work.

4.7 List of Symbols

D_{Hi}	Hydrodynamic dispersion for flow mode i in a pathway [L^2/t].
f	geometric factor for pathway cell residence times.
$F_n(t)$	Breakthrough curve value for radionuclide n at time t.
f_{Qi}	fraction of pathway fluid flow occurring in the ith flow mode.
α_i	Dispersivity for flow mode i in a pathway [L].
λ_j	Poisson transition rate for the jth flow mode of a pathway [L^{-1}].
$J_{jn}(t)$	Cumulative probability distribution for travel time of radionuclide n in flow mode j.
l_i	initial length traveled by a radionuclide in flow mode i [L].
L_p	pathway length [L].
Δp	fractional change in parameter p.
p^{new}	New value for parameter p.

p^{old}	Old value for parameter p.
r(0,1)	Uniformly distributed random number between 0 and 1.
R_n	retardation parameter for radionuclide n.
$\bar{t}$	Average travel time for a particle in a pathway [t].
Δt	RIP time step interval (e.g., 100 years) [t].
δt_i	residence time for pathway cell i [t].
$t_{l,j}$	left endpoint for time interval about pathway cell j [t].
t_n	travel time for radionuclide n along a pathway.
$t_{r,j}$	right endpoint for time interval about pathway cell j [t].
μ	mean value
V_i	Average linear fluid velocity for flow mode i [L/t].
w_{in}	Mass distribution coefficient for pathway cell i and radionuclide n.

5. DISRUPTIVE EVENTS COMPONENT MODEL

5.1 Introduction

As pointed out in Chapter 2, the RIP integrated performance assessment model is essentially a radionuclide transport model, consisting of a series of inter-connected, coupled component models with input/output relationships for radionuclide transfer. The two major component models describe 1) waste package behavior and radionuclide release; and 2) transport pathways to the accessible environment. The third component model, disruptive events, acts to modify the behavior of the other two components.

The purpose of this chapter is to describe the structure of the disruptive event component model. The general methodology and assumptions incorporated into the software are presented in detail, and application of the computer program is discussed.

As described in detail in Chapters 3 and 4, RIP is structured such that waste package behavior or transport pathways can be simulated independently. Each of these two modules produces its own set of output results. The disruptive events module, however, only acts to modify the waste package behavior and/or the transport pathways models, and therefore is not structured to be run separately. Because the disruptive events module acts to modify the behavior of the other two modules, it is recommended (although not essential) that the reader acquire an understanding of the manner in which waste package behavior and transport pathways are simulated in RIP prior to reading this chapter.

The basic RIP computational algorithm described in Chapters 3 and 4 is deterministic. That is, the computational algorithm simulates a single *system realization* (i.e., the behavior of the repository system given specified values or time histories of model parameters). However, uncertainty in both the model parameters and the component models themselves is explicitly included in the integrated stochastic model. Due to the inherent uncertainties resulting from our lack of knowledge, many of the parameters will be represented by probability density functions (pdf's). The integrated model uses the Monte Carlo method to sample these distributions and simulate a large number of random system realizations (using the deterministic model) in order to determine probability distributions of measures of site performance (e.g., cumulative release, transport time). This chapter will describe the manner in which disruptive events are incorporated into this scheme.

Section 5.2 provides an overview of the disruptive events modeling approach. Section 5.3 describes how disruptive event *occurrences* are selected and defined. Section 5.4 then specifically describes how disruptive event *consequences* are defined. Section 5.5 summarizes the disruptive events computational algorithm. Section 5.6 provides a summary discussion of the required disruptive event input parameters. Finally, Section 5.7 contains a glossary of terms used throughout Chapter 5. This glossary is included as an aid to the reader to avoid confusion with regard to terminology.

5.2 Overview of Disruptive Events Methodology

The RIP integrated performance assessment model explicitly incorporates two types of uncertainty:

- uncertainty regarding the parameters describing the behavior of the repository system under a specified set of expected present and future conditions.
- uncertainty regarding the occurrence and consequences of "disruptive events" which change conditions and influence system parameters.

The first type of uncertainty is explicitly included in the waste package behavior and transport pathways models by treating the model parameters representing these processes as uncertain (i.e., represented by distributions which are randomly sampled each Monte Carlo system realization). That is, *processes* (such as container corrosion, heating and cooling of the waste packages, failure of borehole walls, and migration of radionuclides through fractures or faults) are treated explicitly in the waste package and transport pathways models.

The second type of uncertainty is represented by the disruptive events component model described in this chapter. *Disruptive events* are defined here as discrete perturbations of the repository system. That is, disruptive events are discrete occurrences which have some quantifiable effect on the processes described by the other two component models. Note that discrete is a relative term, and does not necessarily imply instantaneous. Given the long time scales of interest (10,000 years or more), something taking place over a period of 100 years could be considered a discrete event. Examples of disruptive events under this definition include volcanism, faulting, and human intrusion. In general, the disruptive events component model is intended to represent relatively rare occurrences. Events that occur continuously throughout the time period of interest are more efficiently modeled as processes within the other two component models.

Note that climate change is not treated as a disruptive event, since this is a process which occurs gradually, and some sort of climate change is expected in the future (although the actual nature of the change is uncertain). Climate change (and the corresponding changes in environmental conditions such as water table elevation and infiltration rate) can be treated explicitly in the transport pathways and waste package behavior components.

As discussed by NRC (1990), there are two general approaches for analysis of the uncertainties listed above. The first approach (referred to here as "simulation") consists of incorporating all uncertainties directly into the models and data bases describing the repository system. The second consists of developing and separately simulating "scenarios", which explicitly represent alternative ways in which the repository environment might change in the future. Each scenario may require a different conceptual and computational model for simulating the performance of the repository system. Most analyses use a combination of the two approaches (e.g., Bertram-Howery et al., 1990). The differences between the approaches are not severe, and both simulation approaches and approaches

which explicitly incorporate scenarios can be designed to produce results that are essentially equivalent. A more detailed discussion of the two approaches is presented in Appendix B.

The RIP model takes the first approach (simulation) and explicitly incorporates all uncertainties directly into the component models and parameters, using a Monte Carlo method to sample parameters describing both processes and events. That is, RIP creates a time history of disruptive events (and other system parameters) for each Monte Carlo system realization of repository performance. Because model parameters are described stochastically, each realization produces a different time history of events and processes. The integrated model is designed such that it can simulate all combinations of model parameters and time histories which might be realized.

As will be described below, a given disruptive event will have associated with it one or more consequences. That is, a disruptive event has specific effects on waste package behavior and/or radionuclide transport in pathways. These consequences can be treated stochastically. For example, if a certain disruptive event has the effect of disrupting a number of waste packages (i.e., instantly failing container and cladding), the actual number of waste packages which are affected could be a stochastic parameter.

Figure 5-1 schematically summarizes the relationships between the three component models and the model parameters within RIP. The first type of uncertainty discussed above (parameter uncertainty) is represented by the stochastic model parameters which control the waste package behavior and transport pathways models. The second type of uncertainty discussed above (events) is represented by the disruptive events model, which can directly affect the other two component models and/or their parameters. Note that the disruptive events model also has stochastic input parameters.

The following three sections describe the selection and description of the disruptive events, specification of disruptive event consequences, and a summary of the computational algorithm by which the occurrences and consequences are incorporated into the integrated model.

5.3 Selection and Description of Disruptive Events

Figure 5-2 illustrates the set of all conceivable disruptive events that could occur at a proposed geologic repository such as Yucca Mountain. Only a small portion of these events can be considered credible. A *credible event* is defined as an event possessing a significant probability of occurrence at the site over the time period of interest. According to the current U.S. regulations (40 CFR, Part 191, App. B), performance assessments need not consider events or processes that are estimated to have less than one chance in 10,000 of occurring over 10,000 yrs (i.e., an occurrence rate of 10^{-8} yr^{-1}).

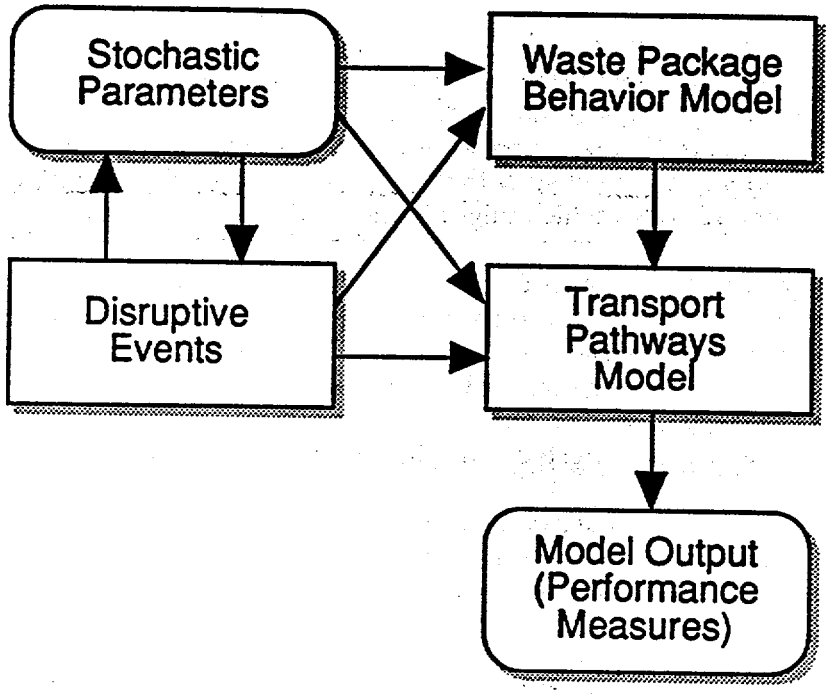


FIGURE 5-1
REPRESENTATION OF UNCERTAINTY IN RIP
ARGONNE/MODEL DEVELOPMENT/WA

PROJECT NO. 903 1371.203 DRAWING NO. 26471 DATE 9/10/91 DRAWN BY EA

Golder Associates

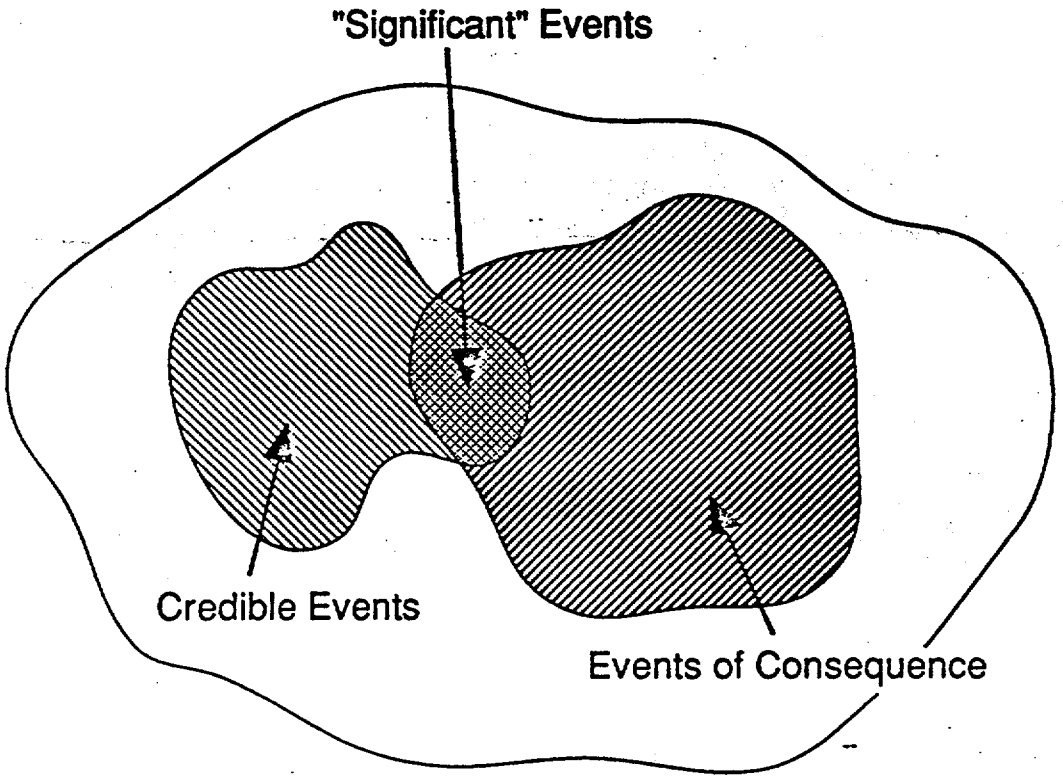


FIGURE 5-2
SET OF ALL CONCEIVABLE DISRUPTIVE EVENTS
ARGONNE/MODEL DEVELOPMENT/WA

PROJECT NO. 903 1371.203 DRAWING NO. 26472 DATE 9/10/91 DRAWN BY EA

Golder Associates

Likewise, only a portion of the conceivable events at Yucca Mountain can be considered to be *events of consequence* with respect to the performance of the repository. As illustrated in Figure 5-2, the disruptive events of interest are formed by the intersection of these two groups: those events that are both credible *and* of consequence. These are referred to here as *significant events*. It is the significant events that are explicitly considered by the disruptive events model. Identification of a complete set of significant events is a critical data input requirement for the model. The process by which a complete set of significant events can be identified is discussed in detail by others (e.g., Cranwell et al., 1990; Guzowski et al., 1990). A preliminary list of anticipated disruptive events for the Yucca Mountain site consists of the following three items: 1) seismic activity/faulting; 2) volcanism; and 3) human intrusion. This short list is based on the results of a preliminary workshop attended by a number of project scientists in August 1991 in GAI's Redmond, Washington office, and is presented here only for purposes of illustration (Golder Associates, 1991b).

RIP assumes that all disruptive events can be simulated as Poisson processes. Events described using a Poisson distribution are assumed to occur singly and independently, and the probability that an event will occur in a short time interval is proportional to the length of the interval (Cox and Miller, 1965).

Note that the assumption of independence of events is not strictly accurate in all cases. The occurrence of some events (e.g., faulting) may be related to the time period since a faulting event has previously occurred. Although the model could be modified to handle such a dependence, given the various uncertainties involved in the input parameters, a simple Poisson representation was considered to be adequate for the present purposes.

The Poisson process is described by a single parameter, λ, the rate of occurrence, which has units of time^{-1}. λ is the expected number of occurrences per unit time.

The probability that x events will occur during a time interval Δt is given by the following expression:

$$\frac{e^{-\mu}\ \mu^{x}}{x!} \tag{5.1}$$

where

$$\mu = \lambda \Delta t \tag{5.2}$$

The Poisson distribution also has the property that the intervals between events are independently distributed with the exponential probability density function:

$$\lambda\ e^{-\lambda t} \tag{5.3}$$

In the RIP model, each event is defined by an *occurrence parameter* (i.e., an identifier for that particular event). The value of the occurrence parameter is internally computed by the model during each system realization as a function of the rate of occurrence. In particular, the occurrence parameter is computed every time step during a given system realization,

taking on a value of 1 if the event has occurred that time step and a value of 0 if it has not occurred (based on Equations 5.1 through 5.3).

A particular event can be specified such that it can, by definition, only occur once (i.e., it may not make sense physically for the event to occur more than once). The event is still represented by a Poisson process, but only two states are defined in any given realization at a given time: either the event has occurred, or it hasn't occurred. Once the event occurs, it is not allowed to reoccur.

In addition to the rate of occurrence, λ, and specification of whether the event can reoccur, each disruptive event can also be assigned one or more *descriptor parameters*. Descriptor parameters are stochastic parameters which define the characteristics and magnitude of the event. Descriptor parameters are defined by the user for each disruptive event and are realized (i.e., a specific parameter value is randomly chosen from its distribution) whenever a disruptive event occurs. They are intended to be simple descriptive parameters which define the event, and can be used to quantify event consequences. For example, the descriptor parameters for a human intrusion/drilling disruptive event might be the number of boreholes drilled, and the deepest pathway (i.e., formation) intersected. For a volcanic event, the descriptor parameters might be the length and width of the dike, and where it is located with respect to the repository.

5.4 Specification of Disruptive Event Consequences

Consequences of disruptive events are specified in two ways:

- by *consequence parameters* which describe the magnitude of a specific set of internally-defined (i.e., hard-wired into the model) discrete responses; and
- by explicitly identifying model parameters (describing waste package behavior and/or radionuclide transport through pathways) which are functions of disruptive event occurrence and descriptor parameters.

The first type of consequence represents discrete responses to the event, such as the disruption of some waste packages. The second type of consequence represents *long-term effects* on the repository system (e.g., changing the hydraulic gradient in the saturated zone, raising the water table, opening a new transport pathway).

Each of these two types of consequences is described in detail below.

5.4.1 Discrete Response Consequences

For each disruptive event, there are four types of internally-defined disruptive event consequences which describe discrete responses:

1) *The radionuclides in a number of waste packages are moved directly to the accessible environment.* It is assumed that the inventory released from each package is equal to the inventory of an unfailed container at the time of the event. The consequence parameter is the number of waste packages affected.

2) *The radionuclides in a number of waste packages are moved directly to a specified pathway (e.g., the saturated zone).* It is assumed that the inventory released from each package is equal to the inventory of an unfailed container at the time of the event. It is also assumed that all of the waste package's inventory is immediately released to the pathway (i.e., the waste package is completely disrupted during the movement and the contents are not limited by any alteration, dissolution, or mass transfer processes at the waste package). The consequence parameters are the number of waste packages affected, and the pathway to which they are discharged.

3) *A number of waste packages are disrupted in place.* It is assumed that the cladding (or pour canister) instantaneously fails with the container. The consequence parameter is the number of waste packages disrupted. The exact manner in which this consequence is implemented in the RIP algorithm, and the resulting approximations involved, are discussed in detail in the RIP User's Guide.

4) *A portion of the mass (previously released from the waste packages) contained within a path is immediately discharged to the accessible environment.* The consequence parameters are the mass fraction discharged from each selected pathway.

Note that the consequence parameters mentioned above can be represented stochastically. Furthermore, they can be described as functions of the descriptor parameters. For example, the number of waste packages disrupted by a drilling event could be described as a function of the number of boreholes. Of course, if a particular event does not produce one of the consequences outlined above, the appropriate consequence parameter is simply set to zero.

5.4.2 Consequences Which Modify Waste Package and Transport Pathways Parameters

In addition to the four consequences described above, which are explicitly included in the model, it is also possible to directly influence parameters defined in the waste package and transport pathway modules, and this capability can be used to specify long-term consequences. A general discussion of the manner in which model parameters can be directly modified by disruptive events is summarized below. The precise manner in which the user can specify this dependence requires an understanding of the user-interface software, and is described in detail in the RIP User's Guide.

The user interface was designed to specifically incorporate user-defined sensitivities to disruptive events when defining parameter values. In particular, for any independent parameter (i.e., a parameter which is not described as a function; it must either be a stochastic or a constant), the user can directly specify the parameter's sensitivities to specific disruptive events. The user simply defines how the current value of the parameter is to be

modified should a particular disruptive event occur. The parameter of interest is modified by a user-defined *influence parameter*. An influence parameter can either *replace, multiply*, or be *added* to the original parameter value. This influence parameter may, in turn, be defined as a function of event descriptor parameters or occurrence parameters.

This capability is best illustrated by example. Suppose a magmatic intrusion disruptive event (represented by the *occurrence parameter* MAGMA) can have the effect of creating a dike which increases the hydraulic gradient in a saturated zone transport pathway (represented by the transport pathways model parameter GRADSZ). Assume that the user wants to specify that the magnitude of the gradient change is proportional to SIZE, which is a *descriptor parameter* for the event, describing the size of the magmatic intrusion (the constant of proportionality being 0.001).

The user could simply define the sensitivity of GRADSZ as follows:

> if disruptive event MAGMA occurs, modify GRADSZ such that
> GRADSZ = GRADSZ + MAGMOD, where MAGMOD = 0.001 x SIZE

MAGMOD is the user-defined *influence parameter*, described here as a function of the descriptor parameter SIZE. GRADSZ would retain the modified value for the remainder of the realization. If a second magmatic intrusion (or any other event which affected GRADSZ) occurred at a later time in the realization, the value of GRADSZ would once again be updated. Note that SIZE is realized every time a magmatic event occurs.

Since the influence parameter can be defined as any complex function, this representation can be quite powerful. A detailed description of the parameters module of the user's interface is provided in the RIP User's Guide).

5.5 Disruptive Events Computational Algorithm

A simplified summary of the computational algorithm for RIP, emphasizing the disruptive events portion of the model is presented below in Figure 5-3.

Within the loop of Monte Carlo realizations, the algorithm steps through time simulating waste package behavior and radionuclide transport pathways. If an event occurs during a given time step, event descriptor parameters, consequence parameters, and influence parameters are realized, and affected system parameters are modified prior to simulating waste package behavior and transport through pathways for the time step.

At the end of the simulation time, the performance parameters for the current realization (e.g., cumulative release to accessible environment) are computed and saved. The algorithm then returns to the top and simulates another system realization. As discussed in Chapter 2, after an appropriate number of system realizations are simulated, the results are combined (e.g., into a CCDF) and analyzed.

Begin nr = 1, number of Monte Carlo realizations

realize all non-temporal system parameters
realize all disruptive events over the time period of interest

Begin t = 1, number of time steps

increment elapsed time
compute time-dependent variables

if one or more events occur during this time step, then for each event:

realize event descriptor parameters
realize descrete response consequences
modify affected system parameters

simulate waste package behavior and transport through pathways, and decay mass for this time step

end t loop

compute performance parameters for current realization

end nr loop

FIGURE
DISRUPTIVE EVENTS
COMPUTATIONAL ALGORITHM
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 26693 DATE 10/8/91 DRAWN BY EA

Golder Associates

5.5.1 Monte Carlo Sampling Algorithm

Completely random Monte Carlo sampling could require a very large number of system realizations to statistically represent the range of system parameters and event occurrences, and this could prove to be computationally prohibitive. The RIP model, therefore, is structured to carry out the Monte Carlo sampling in a intelligent manner to increase efficiency. As pointed out in Section 5.1, RIP explicitly incorporates two types of uncertainty: 1) uncertainty regarding the parameters describing the behavior of the repository system under a specified set of expected present and future conditions; and 2) uncertainty regarding the occurrence and consequences of disruptive events which change conditions and influence system parameters. Both of these uncertainties can be sampled in a biased manner (using importance sampling) to reduce the number of realizations required.

This is illustrated schematically in Figure 5-4, which shows the set of all possible futures for the repository system. The RIP Monte Carlo sampling algorithm can be intentionally biased towards realizations in which disruptive events occur, and realizations at the extremes of the input distributions. The results of the realizations sampled in this manner are subsequently weighted in an appropriate manner before they are combined. This produces better resolution of the high-consequence, low probability "tail" of the results distributions. This can be combined with a stratified (Latin-Hypercube) sampling approach to further improve the efficiency. The importance (and stratified) sampling algorithms are discussed in detail in Appendix C.

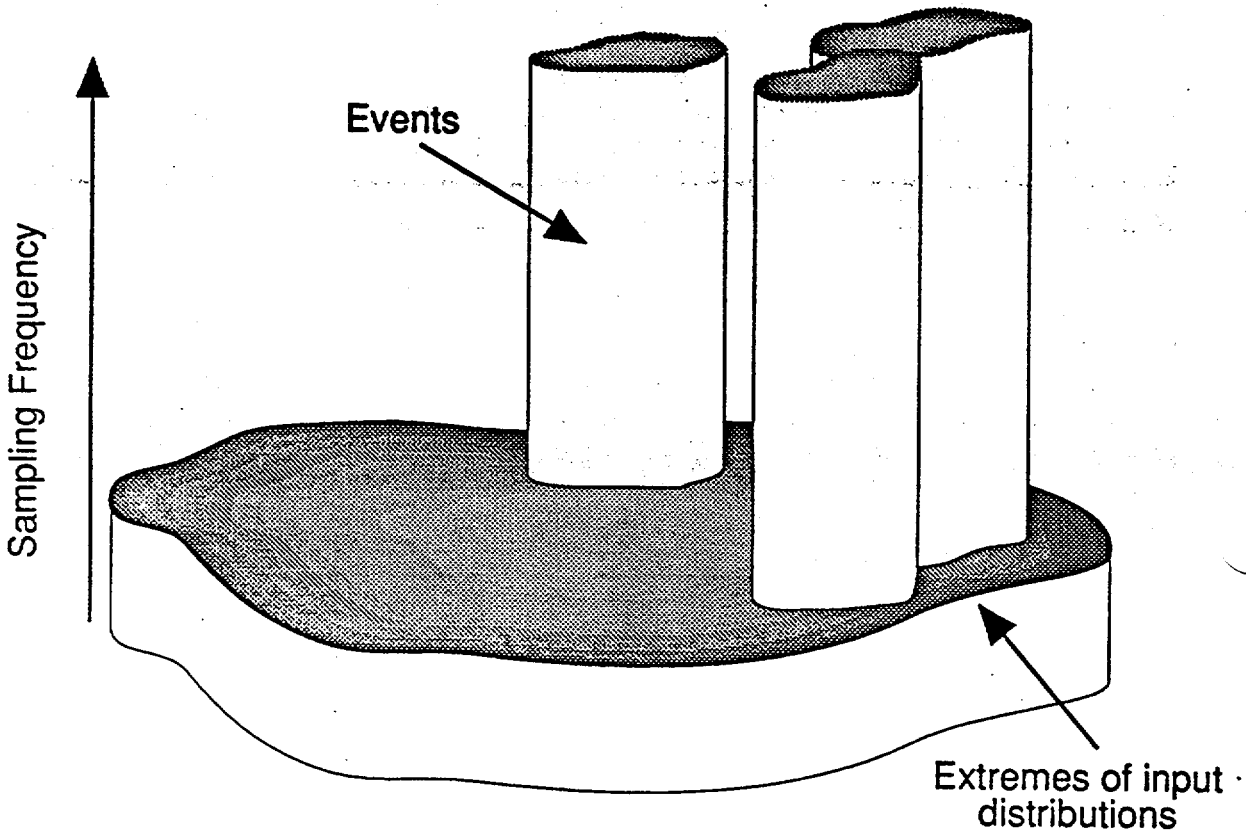
5.6 Summary of Required Input Parameters

The first, and perhaps most difficult task in attempting to simulate disruptive events at the repository is to develop a list of events which are both credible and of consequence with respect to repository performance. The selection procedure must be designed such that the completeness of the set can be demonstrated (in as much as that is possible). As pointed out in Section 5.3, the procedures recommended for carrying out such an exercise have been documented by others.

Once this list has been developed, the RIP model requires specific input for each defined event. The required input for describing the disruptive events is straightforward and is summarized below.

- *Rate of occurrence for disruptive event.* This represents the expected number of occurrences per unit time.

- *Can the event reoccur?* A particular event can be specified such that it can, by definition, only occur once.



Bias towards realizations in which events occur, and realizations at the extremes of the input distribution

FIGURE 5
IMPORTANCE SAMPLING OF PARAMETERS AND EVENTS
ARGONNE/MODEL DEVELOPMENT/WA

PROJECT NO. 903 1371.203 DRAWING NO. 26473 DATE 10/7/91 DRAWN BY EA

Golder Associates

- *Event descriptors.* These are user-defined parameters which are intended to be simple descriptive parameters which define the event, and can be used to quantify event consequences. For example, the descriptor parameters for a human intrusion/drilling disruptive event might be the number of boreholes drilled, and the deepest pathway (i.e., formation) intersected.

- *Discrete response consequences.* There are four discrete response consequences defined by the model. These describe instantaneous, discrete consequences of the disruptive event:

  1) the number of waste packages whose contents are immediately moved directly to the accessible environment.

  2) the number of waste packages whose contents are immediately moved directly to a specified pathway (i.e., the saturated zone).

  3) the number of waste packages which are immediately disrupted in place.

  4) the fraction of the mass in any specified pathway (e.g., the unsaturated zone) which is immediately discharged to the accessible environment.

  Note that these responses can be defined as functions of event descriptors.

- *Long-term consequences.* Long-term consequences of the event can also be defined. These are specified by defining the sensitivity of selected model parameters to disruptive events in terms of an influence parameter. If the particular event occurs, the user-defined influence parameter can replace, multiply or be added to the original parameter value. Model parameters which could be affected by disruptive events might include the hydraulic gradient in saturated zone, the water table elevation, or the fraction of mass released from waste packages which is partitioned into a given transport pathway. Like the discrete response consequences, the long-term consequences can also be defined as a function of event descriptors.

Specifying input for the disruptive events component model is discussed in more detail in the RIP User's Guide.

5.7 Disruptive Events Glossary

consequence parameter - a parameter which describes the magnitude of one of a specific set of four internally-defined discrete instantaneous responses to a disruptive event.

credible event - an event possessing a significant probability of occurrence over the time period of interest. This document uses as the definition of "credible" that provided by 40 CFR, Part 191, App B, which states that only events with an occurrence rate greater than 10^{-8} yr^{-1} need be considered.

descriptor parameter - a user-defined parameter which provides a simple description (quantification, characterization) of a disruptive event, and can be used to quantify event consequences.

disruptive event - discrete perturbations of the repository system which have some quantifiable effect on waste package behavior and/or transport of radionuclides.

influence parameter - a user-defined parameter which replaces, is added to, or multiplies a specified parameter value when a disruptive event occurs. This provides a means for representing long-term effects of disruptive events.

occurrence parameter - the identifier for a disruptive event which automatically takes on the value of the number of event occurrences during the present model time step.

system realization - simulation of the repository system given a value (or time history) for each parameter and event occurrence.

scenario - an explicit representation of an alternative way in which the repository might perform in the future.

significant event - an event which is both credible and of consequence with respect to repository behavior. It is the significant events which are explicitly considered by the disruptive events model.

6. THE RIP STRATEGY EVALUATION MODEL

6.1 Basic Concepts

The previous chapters have described in some detail the RIP performance model. However, RIP is also capable of a higher order of modelling, strategy modelling, which is described in this chapter.

The strategy evaluation model was discussed in general terms in Sections 1.3 and 2.4. Additional details are presented in this chapter.

As discussed in Section 2.4, a 'strategy', as used in RIP, refers to a set of activities or elements which are intended to be carried out in order to develop a better understanding of a proposed repository system. The elements of a strategy represent activities such as the following:

- A test or set of tests which are intended to provide additional information about one or more of the parameters that define the performance model. Note that such tests do not always measure the affected RIP parameter directly, and may often measure it indirectly.

- Development of an improved model which will reduce overall model uncertainty. For example, such an activity could consist of developing and making a number of runs of an improved lower-level model of a subsystem. The result might be a modified component model in RIP, or improved RIP parameter precision, or a reduced model-error level.

- Construction of an infrastructure element (eg. a borehole, or a road), or completion of a required procedural activity such as acquiring a license.

RIP's strategy module allows the user to define strategies, and to evaluate their effectiveness by examining probability distributions of three specific results:

- the *cost of the strategy*

- the *duration of the strategy*

- the strategy's likely *effect on the performance model's predictions.*

Each of these three types of result is evaluated by RIP as a stochastic variable. Thus, RIP provides the user with probability distributions for the overall cost and duration of a strategy. Similarly, RIP provides the user with probability distributions for what the performance-measure distributions may look like after the strategy is carried out. This last item can be a confusing one, because it represents our probabilistic estimate now of what our probabilistic performance estimates will be in the future.

Figures 6-1 and 6-2 show the form of these results. Figure 6-1 shows typical cost and duration probability distributions. Figure 6-2 shows an example of the current probability distribution for system performance, plus confidence bounds for the distribution after the characterization strategy is executed.

The issue of uncertainty in the cost and duration of a strategy is an important one. While it would be convenient to be able to plan with confidence about the necessary costs and time, in reality neither science, politics, nor regulatory bureaucracy is subject to control by the agency attempting to develop the repository. The RIP strategy module contains what is, in effect, a version of the Critical Path Method which incorporates uncertainty.

The general concepts discussed above are best explained by considering a simple example.

6.1.1 Strategy Evaluation Example

Consider the following simplified example: suppose our current RIP model showed a 90% likelihood of acceptable performance by the repository system. The performance hinged on a single critical issue, which was the possible existence of an undetected fault near to the repository. There was a 10% likelihood that a fault existed, and its existence was necessary and sufficient to cause unacceptable performance of the repository.

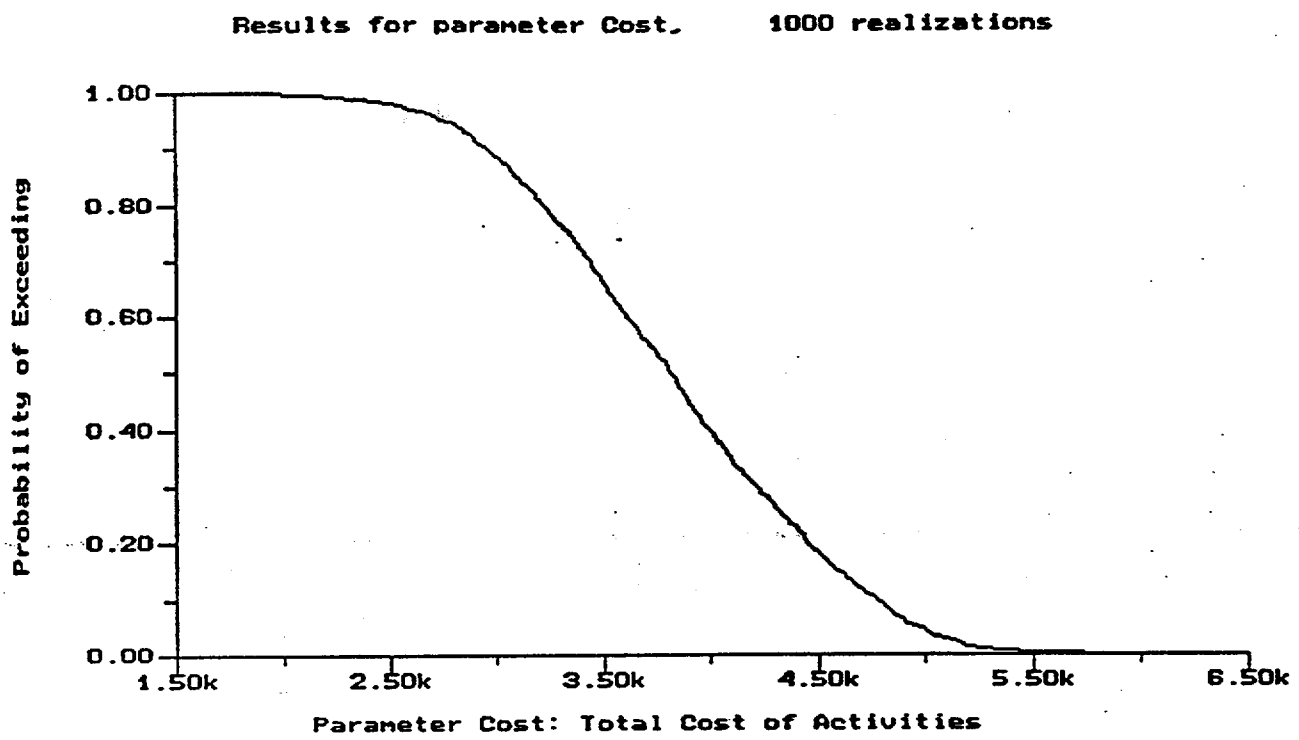
A strategy of extensive test drilling was defined for this problem. The drilling would be definitive, so that if the fault existed it would definitely be found. However, the drilling and data evaluation program would be costly and protracted.

The RIP strategy module for this strategy would provide appropriate estimates of the cost and time distributions to carry out the drilling, and would show a likelihood of 90% that subsequently the performance model would show the site was definitely suitable, and a likelihood of 10% that it would be definitely unsuitable.

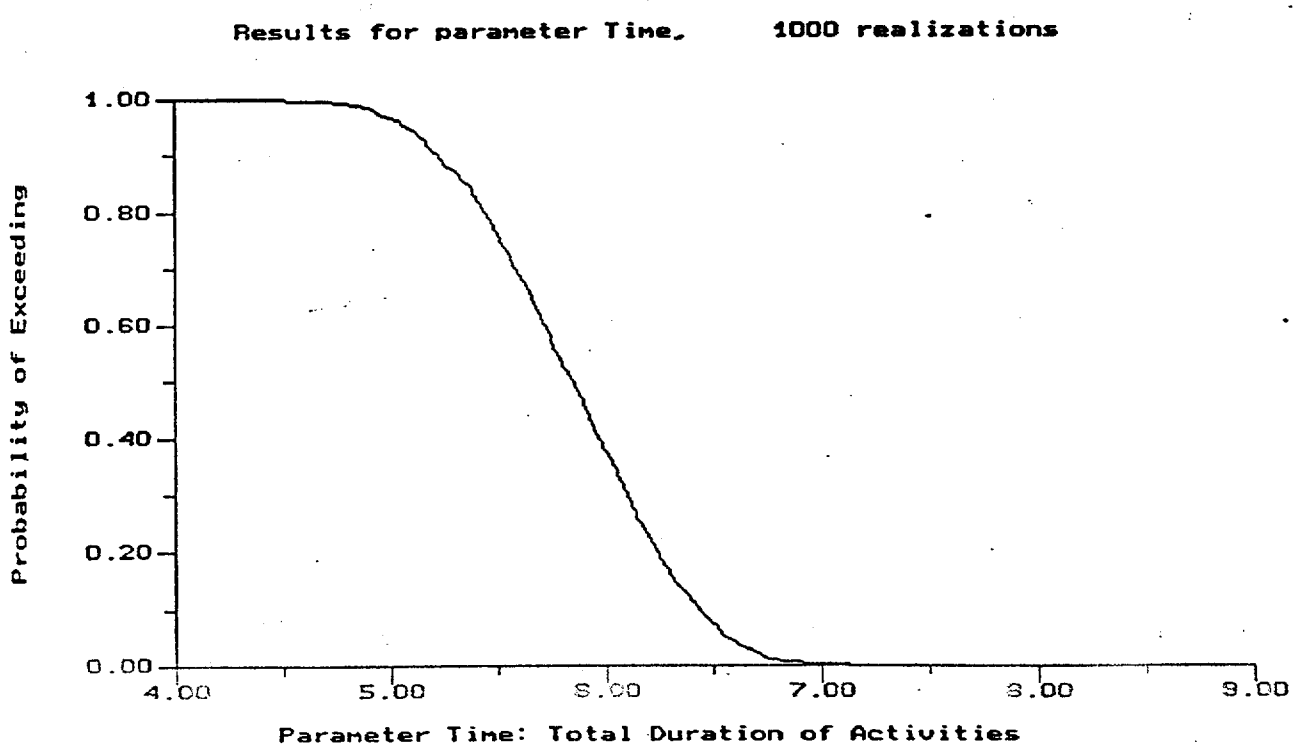
An alternative program might be proposed, however, which while being significantly quicker and cheaper would be less definitive: it would reduce the uncertainty about the fault by one order of magnitude, so that there would be a 99% confidence in the test result. For this strategy, RIP would show a 90% likelihood that the end product would be a 99% confidence in the site, and a 10% likelihood that the end product would be a 1% confidence in the site.

Depending on the regulatory level of confidence required ('compliance criteria'), and the available cost and time, one of the above two strategies would be preferable. For the second strategy, it might be optimal to plan to reduce the residual 1% uncertainty during a 'performance confirmation' phase subsequent to construction of the repository.

If the real world were as simple as the above example, RIP would not be needed. However, in more realistic situations where there are hundreds or thousands of parameters, which have greater or lesser impact on performance, and where few tests are definitive, the situation can be much more complex. It is here that RIP's strategy module



6-1a Cost



6-1b Duration

FIGURES 6-1

PROBABILISTIC ESTIMATE OF COST AND DURATION

ARGONNE/MODEL DEVELOPMENT

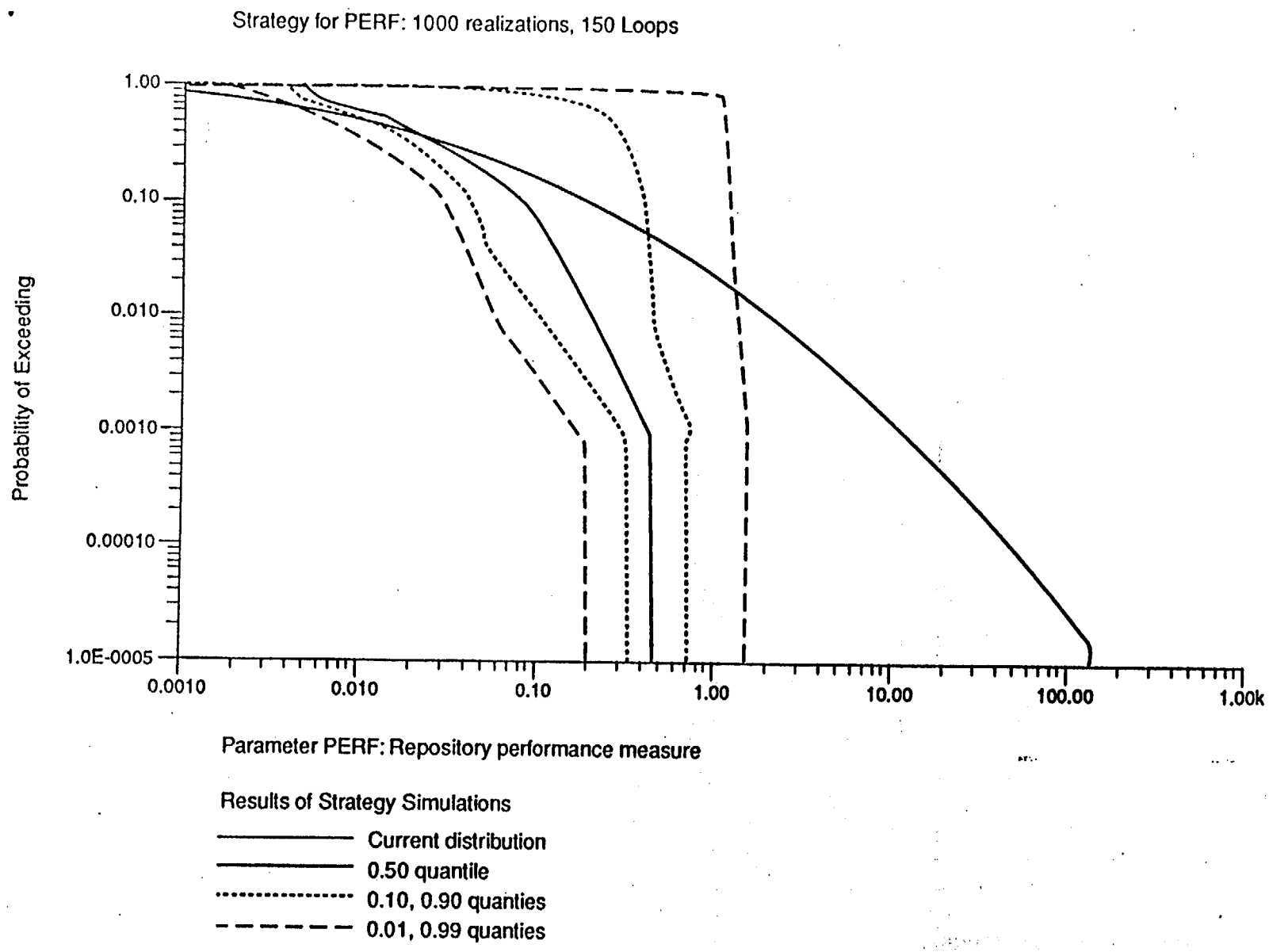


FIGURE 6-2

EFFECT OF STRATEGY ON PERFORMANCE ASSESSMENT RESULT

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903 1371.203 DRAWING NO. 28167 DATE 4/2/92 DRAWN BY TB

Golder Associates

can be valuable, allowing the user to compare and contrast alternative strategies, which may themselves be quite complex, in order to develop an optimal approach.

6.2 Development of an Activities Database

As discussed in Section 1.3.4, the RIP strategy module requires the user to construct an activities database. Each entry in the database contains the following information:

- An identifier for the activity.
- A description of the activity.
- The duration of the activity. The duration may be entered as a constant, as a stochastic value, or as a function of other parameters.
- The cost of the activity. The cost may be entered as a constant, as a stochastic value, or as a function of other parameters. In particular, the cost may be a function of the activity's duration.
- A list of any required precedent activities. For each precedent activity, either the start or finish of the current activity can be constrained by the start or finish of the precedent activity, with a user-entered time-lag. For example, an activity could be constrained such that it could not start until three months after its precedent activity finished.
- A list of all performance-model parameters which would be affected by the activity, and the amount of new information about the parameter that would be generated by the activity. The definition of the measure of new information is discussed below in Section 6.3.

6.3 Updating - Simulating the Effect of New Information

The process of updating prior knowledge based on new test results is simulated in RIP, using an algorithm that captures in a simple way what is in fact a very complex process. In RIP, the prior knowledge is expressed in each stochastic parameter's original probability distribution. The prior knowledge input will normally have come from elicitations of one or more experts in the area, based on reviews of experimental data and available modelling results.

RIP's algorithm for updating is based on Bayesian updating of a parameter. Bayes' theorem states that, for some particular value x of parameter X:

$$p'(x) = \frac{p(x)\ p(x_t|x)}{p(x_t)}$$

where:

$p'(x)$ = updated probability of x;
$p(x)$ = current probability of x;
$p(x_t|x)$ = probability of test result x_t given that x was the true value; and
$p(x_t)$ = current probability of test result.

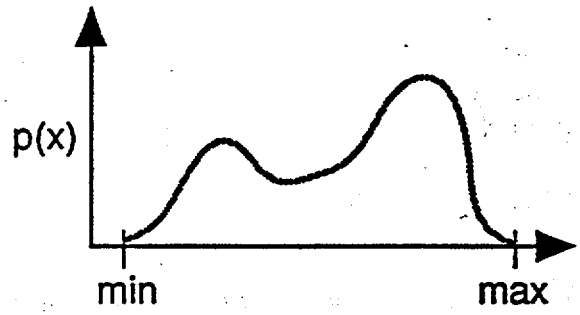
The updated distribution is often referred to as the 'posterior' distribution. Note that the test result does not have to be a measurement of X itself: the test need only measure some value that is affected by the value of X. For example, the parameter X might be the annual precipitation at a site, whereas the test might measure the moisture content in the soil. A considerable amount of additional information, plus some modelling, would be needed to derive the relationship between soil moisture content and precipitation.

Figure 6-3 shows graphically the process that RIP uses to simulate updating. It is essentially a two-step approach to updating: 1) first RIP simulates (by Monte Carlo sampling) the value of x implied by the test result, x_t; and 2) then it computes an updated ('posterior') distribution for X.

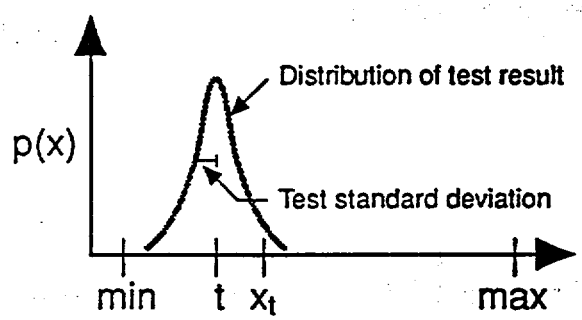
The user is required to define a term called the *test standard deviation*, which represents the level of uncertainty in the test-generated estimate x_t. The test standard deviation is input in a normalized form as a multiple of the prior standard deviation of X. For example, if the prior probability distribution of a rock property was based on 16 samples, and a test was planned which would evaluate 9 more samples, the test standard deviation would be 1.33 of the prior (because the standard deviation of a sample mean varies inversely with the square root of the number of samples). If the new test was to evaluate 100 samples, the test standard deviation would be 0.4 of the prior's.

Hence, a test standard deviation greater than one implies that the test provided a smaller amount of information than the prior, a test standard deviation equal to one implies that the test provided an amount of information equal to that currently suppied by the prior, and a test standard deviation of less than one implies that the test provided a larger amount of information than the prior. A test standard deviation of 0 implies that the test was definitive: there would be no remaining uncertainty after the test.

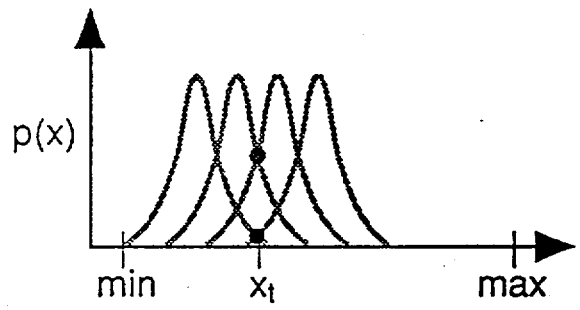
RIP generates a value for x_t by randomly realizing the 'true' value of X (based on the prior distribution), and then randomly realizing the test result x_t based on the true value and the test standard deviation. The process of Bayesian updating of the prior distribution for X is then carried out.



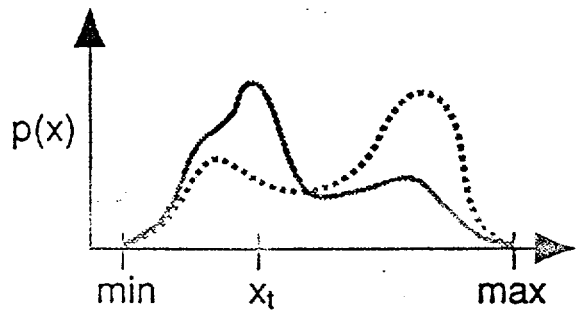
1. Prior distribution of x



2. Realize 'true' value t
3. Construct distribution of test result based on "true" value and test standard deviation
4. Realize test result x_t



5. Bayesian update:
   $p'(x) \propto p(x)\, p(x_t|x)$



6. Normalize the posterior p'

FIGURE 6-3
SIMULATING EFFECT OF NEW INFORMATION
ARGONNE/MODEL DEVELOPMENT

The probability distribution of x_t uses a normal distribution. The updating process also uses a normal distribution, in order to compute the 'probability of the test result if x was the true value'. This distribution is evaluated for a series of values x_i, using logic similar to that described above, and is used to evaluate the likelihood of having seen x_t if the real result was x_i. Regardless of the form of the prior distribution of X, the output is a 10-point 'cdf' cumulative distribution. The range of the output distribution is the same as the prior, and the internal points are arranged so as to best represent the true shape of the distribution's curve.

In order to gain computational efficiency, RIP short-cuts calculating the denominator in Bayes' theorem by simply normalizing the posterior distribution to have a total probability of unity. This is a valid approach which does not affect the result.

6.4 Strategy Evaluation Results

6.4.1 Strategy Cost and Time Distributions

Probability distributions for cost and elapsed time can be displayed for any stage of the strategy. These distributions are developed by Monte Carlo realizations of the entire strategy, using the methods described elsewhere in this report. These displays are available from the RIP module in which the user defines the strategy activities.

6.4.2 Effect of a Strategy on a Parameter

Within the parameter-editing portion of RIP it is possible to evaluate the sensitivity of any stochastic parameter or function of stochastic parameters to the selected strategy. The result is displays of probability distributions of different quantiles (eg, the 0.9 exceedance level) for the parameter subsequent to carrying out the strategy (in a form similar to that of Figure 6-2).

The results are developed by repeatedly executing the strategy and doing Bayesian updates of all the affected parameters from their prior probability distributions. The probability distribution of the parameter being studied is evaluated using the updated distributions, and the different quantiles (0.01, 0.1, 0.5, 0.9, 0.99) are evaluated and saved. After a number of repetitions of this process, it is possible to develop the distributions of each quantile subsequent to carrying out the strategy.

6.4.3 Effect of a Strategy on Performance Assessment Results

Within the RIP post-processing module (see Section 2.4.1), the user can evaluate the effect of the proposed strategy on any of the performance results calculated by RIP. This evaluation is made not by re-running the back end, but by simply re-evaluating the relative

likelihood of each previously-evaluated realization. In this way it is possible to rapidly evaluate alternative strategies without the time-consuming necessity of re-running the back end. The result of such an evaluation has a form similar to that shown in Figure 6-2.

6.5 Summary of the Strategy Evaluation Model

The strategy portion of RIP provides the second cornerstone of the GAI approach by quantitatively integrating the performance assessment model with the characterization activities. The strategy model is essentially a decision analysis shell around the performance assessment model which allows the user to evaluate alternative site characterization strategies.

For any user-specified characterization strategy, RIP provides three outputs by which alternative strategies can be ranked and compared:

- a probabilistic estimate of cost;
- a probabilistic estimate of duration; and
- a probabilistic evaluation of the predicted site performance resulting from implementation of the strategy.

Computation of the first two outputs is straightforward, and consists of simply integrating within a Monte Carlo framework the cost and duration estimates for the individual activities, taking into account any precedence requirements. The third output relies on subjective assessments by experts of the extent to which model parameter uncertainty will be reduced by a particular activity. Given these assessments, along with the current state of knowledge, RIP uses a Bayesian computational algorithm to simulate how probability distributions representing parameter uncertainty will change as a result of a particular characterization strategy, and develops a probabilistic evaluation of the anticipated repository performance.

7. REFERENCES

Aidun, C.K, S.G. Bloom and G.E. Raines, 1988, *Radionuclide Transport Through Perforations in Nuclear Waste Containers*, Mat. Res. Soc. Symp. Proc., Vol 112, pp. 261-272.

Apted, M.J., A.M. Liebetrau and D.W. Engel, 1987, *Spent Fuel as a Waste Form: Analysis with AREST Performance Assessment Code*, Waste Management '87, Vol. 2, pp. 545-554.

Apted, M.J., W.J. O'Connell, K.H. Lee, A.T. MacIntyre, T.-S. Ueng, T.H. Pigford, and W.W.-L. Lee, 1991, *Preliminary Calculation of Release Rates from Spent Fuel in a Tuff Repository*, High Level Radioactive Waste Management, Proceedings of the 2nd Annual International Conference, Las Vegas, NV, April 28-May 3, 1991.

Bear, J., 1979, Hydraulics of Groundwater, McGraw-Hill Book Co., New York.

Benjamin, J.R. and C.A. Cornell, 1970, Probability, Statistics, and Decision for Civil Engineers, McGraw-Hill Book Company, New York.

Bertram-Howery, S. G. et al., 1990, *Preliminary Comparison with 40 CFR Part 191, Subpart B for the Waste Isolation Pilot Plant, December 1990*, SAND90-2347, prepared by Sandia National Laboratories for the U.S. DOE.

Bullen, D.B., 1990, *Engineered Barrier System Failure Model*, Demonstration of a Risk-Based Approach to High-Level Waste Repository Evaluation, prepared by Risk Engineering, Inc., Golden, CO, EPRI NP-7057, Electric Power Research Institute, Palo Alto, CA.

Chambre, P.L., T.H. Pigford, W.W.-L. Lee, J. Ahn, S. Kajwara, C.L. Kim, H. Kimura, H. Lung, W.J. Williams and S.J. Zavoshy, 1985, *Mass Transfer and Transport in a Geologic Environment*, LBL-19430, Lawrence Berkeley Laboratory.

Cox, D.R. and H.D. Miller, 1965, *The Theory of Stochastic Processes*, Chapman and Hall, New York.

Cranwell, R.M., R.W. Guzowski, J.E. Campbell, and N.R. Ortiz, 1990, *Risk Methodology for Geologic Disposal of Radioactive Waste*, NUREG/CR-1667, SAND80-1429, prepared by Sandia National Laboratories for the U.S. Nuclear Regulatory Commission.

Golder Associates, 1977. "Development of Site Suitability Criteria for the High Level Waste Repository, prepared for Lawrence Livermore Laboratory, Livermore, CA.

Golder Associates, 1986. "Application of Performance-Based Decision Process in Selection of Crystalline Rock Sites for Deep Geologic Disposal of Nuclear Waste", prepared for Battelle Memorial Institute, Office of Crystalline Repository Development.

Golder Associates Inc., 1991a, *Performance Assessment Modeling to Support Site Characterization and Suitability Evaluation at Yucca Mountain: Report on Initial Workshop*, 903-1104, January 30, 1991, Redmond, WA.

Golder Associates Inc., 1991b, *Memorandum on Disruptive Events Performance Assessment Workshop, August 22-23, 1991*, 903-1371, September 2, 1991, Redmond, WA.

Golder Associates Inc., 1992, *RIP Repository Performance Assessment and Strategy Evaluation Model: User's Guide*, 903-1371, prepared for OCRWM under contract to Argonne National Laboratory.

Guzowski, R.V., 1990, *Preliminary Identification of Scenarios That May Affect the Escape and Transport of Radionuclides From the Waste Isolation Pilot Plant, Southeastern New Mexico*, SAND89-7149, prepared by Sandia National Laboratories for the U.S. DOE.

Hogg, R.V. and A.T. Craig, 1978, Introduction to Mathematical Statistics, 4th Edition, Macmillan Publishing Co., Inc., New York.

Johnson, L.H., N.C. Garisto and S. Stroes-Gascoyne, 1985, *Used Fuel Dissolution Studies in Canada*, Waste Management '85, Vol. 1, pp. 479-482.

Kerrisk, J.F., 1984, *Solubility Limits on Radionuclide Dissolution at a Yucca Mountain Repository*, LA-9995-MS, Los Alamos National Laboratory.

Liebetrau, A.M., M.J. Apted, D.W. Engel, M.K. Altenhofen, C.R. Reid, D.M. Strachan, R.L. Erikson, and D.H. Alexander, *AREST: A Probabilistic Source-Term Code for Waste Package Performance Analysis*, Waste Management '87, Vol. 2, pp. 535-544.

National Research Council, 1990. "Rethinking High-Level Radioactive Waste Disposal - A Position Paper of the Board on Radioactive Waste Management", National Academy Press, Washington, D.C.

Nuclear Regulatory Commission, 1983. "In Situ Test Programs Related to Design and Construction of High-Level Nuclear Waste (HLW) Deep Geologic Repositories", NUREG/CR-3065, prepared by Golder Associates for the Office of Nuclear Material Safety and Safeguards, U.S. Nuclear Regulatory Commission.

Nuclear Regulatory Commission, 1990, *Phase I Demonstration of the Nuclear Regulatory Commission's Capability to Conduct a Performance Assessment for a HLW Repository*, Final Draft, April 20, 1990.

Nuclear Waste Technical Review Board, 1990. "First Report to the U.S. Congress and the U.S. Secretary of Energy", U.S. Government Printing Office, Washington, D.C.

O'Connell, W.J. and R.S. Drach, 1986, *Waste Package Performance Assessment: Deterministic System Model Program Scope and Specification*, UCRL-53761, Lawrence Livermore National Laboratory.

O'Connell, W.J., 1990, *Status of Integrated Performance Assessment of the Waste Packages and Engineered Barrier System*, High Level Radioactive Waste Management, Proceedings of the International Topical Meeting, Las Vegas, NV, April 8-12, 1990.

Pigford, T.H., P.L. Chambre and W.W.-L. Lee, 1990, *A Review of Near-Field Mass Transfer in Geologic Disposal Systems*, LBL-27045, Lawrence Berkeley Laboratory.

Robinson, P. and K. Worgan, 1991, *The CALIBRE Source-Term Code, Technical Documentation for Project-90*, SKI TR 91:20, Swedish Nuclear Power Inspectorate, Stockholm.

Sadeghi, M.M., T.H. Pigford, P.L. Chambre and W.W.-L. Lee, 1989, *Equations for Predicting Release Rates for Waste Packages in Unsaturated Tuff*, Lawrence Berkeley Laboratory.

Stout, R.B. and W.L. Bourcier, 1991, Personal Communication, June 28, 1991, Lawrence Livermore National Laboratory.

APPENDIX A

METHODS FOR DEVELOPING DEFENSIBLE SUBJECTIVE PROBABILITY ASSESSMENT

METHODS FOR DEVELOPING DEFENSIBLE SUBJECTIVE PROBABILITY ASSESSMENTS

William J. Roberds, ScD
Golder Associates Inc.
4101 148th Ave NE
Redmond (Seattle), WA 98052
(206) 883-0777

ABSTRACT

Typically, some degree of uncertainty exists in the senarios and value of parameters at any site (e.g., due to insufficient data, natural spatial variability, or possible changes with time). Often, this uncertainty must be quantified (e.g., in terms of probability distributions that express the relative likelihood of any value). Because of inevitable data base deficiencies, those probability distributions must be based to some degree on subjective assessments, reflecting personal opinions and judgement, consistent with all available information (site-specific and generic) and recognizing the entire range of possible values. Subjectively derived probability distributions can represent the opinions of individuals or of groups. There are problems associated with either, which, if uncorrected, render the results suspect and difficult to defend. Various techniques have been developed to conduct subjective probability assessments with varying effort and success in mitigating such problems. Thus, the appropriate technique is that which provides the desired level of defensibility at least cost.

INTRODUCTION

Various scenarios can occur at a site, and the associated parameters are often complex, varying spatially and, in some cases, with time, as a function of scale and possibly other factors. The scenarios and values of parameters must often be estimated (e.g., for analysis and design). However, data bases regarding these scenarios/parameters, in many cases, will contain a small number of samples and inexact representation of the true conditions, because of the cost involved in gathering representative data. Scenarios/parameter values cannot be determined accurately in such cases where the data base is statistically insufficient. Instead, the scenarios/parameter values must be estimated based on whatever information is available, including generic as well as site-specific data. Those estimates, depending on their application, may represent conservative assumptions (e.g., to demonstrate compliance with some criteria) or best guesses (e.g., to predict actual performance of alternatives as input to decision making). Such estimates must necessarily incorporate interpretations and judgements regarding the data base, which are subjective and in many cases non-unique, and thus may be open to controversy. Controversy can significantly delay a project and cause unnecessary expense, especially if an ultra-conservative assumption results. Hence, the objective is to cost effectively produce appropriately defensible estimates of scenarios/parameter values where significant uncertainty exists and must be subjectively assessed.

Many examples of such uncertainty analyses exist[1]. This paper is derived from a detailed manual developed by the author specifically for conducting subjective probability assessments[2].

UNCERTAINTY/PROBABILITY CONCEPTS

Variables can represent 1) the state at a particular place and time, or 2) where the state may vary spatially and/or temporally, a group statistic (e.g., mean or variance of the population). The variables, in either case, would have a unique value. Moreover, such variables may be "continuous" (i.e., each may have an infinite number of possible states) or may be "discrete" (i.e., each may have a finite number of possible states).

Often, the state of a variable has not been directly and accurately observed, and there will generally be some uncertainty as to what the state of that variable actually is, was, or will be. The possible sources of this uncertainty can be summarized as follows[2]:

- **Statistically insufficient data:** In direct observations of a variable state, measurement errors (random or systematic) and accuracy limitations may exist. Where the state has not been directly observed, it must be inferred (e.g., interpolated, extrapolated, or analytically derived) from other information. In analytically deriving a variable state from other site- and time-specific measurements, there may be imperfect understanding regarding the processes involved and approximations and simplifications in the analytical procedure. The applicability of indirect observations in the inference of the variable state must be considered. In assessing group statistics, there may not be enough data to be statistically significant or the data may not accurately represent the population (i.e., biased sampling).

- **Natural spatial and/or temporal variability:** The variable state may vary spatially or temporally (i.e., change with time) or both. For example, the space may not be homogeneous and uniform, and may instead have heterogeneities, or the variable state may be affected by future events that cannot be predicted with certainty. In interpolating or extrapolating from observations (direct or indirect) elsewhere and/or at other times, this spatial or temporal variability and the effects of heterogeneities and of events and processes (both past and future) must be considered.

The uncertainty in the actual state of a variable can be quantitatively expressed in various related ways (e.g., ranges, accuracy measures, confidence levels, or probability distributions)[3]. As illustrated in Figure 1, probability distributions can be defined for:

- Discrete variables (Figure 1a), in terms of a probability mass function (pmf), which expresses the probability of each possible variable state.

- Continuous variables (Figure 1b), in terms of:

- probability density function (pdf), which expresses the relative likelihood of each possible variable state; and/or

- cumulative distribution function (cdf), which expresses the probability that the variable state will be less than or equal to each possible variable state.

- **Group statistics (Figure 1c), in terms of pdfs and/or cdfs for:**

  - mean (μ_x), first moment about zero;
  - variance (σ_x^2), second moment about the mean;
  - standard deviation (σ_x), square root of the variance; and/or
  - other higher moments of a distribution.

- **Multiple variables (Figure 1d), in terms of:**

  - joint pmf/pdf, which expresses the probability or relative likelihood of each possible combination of discrete or continuous variable states actually occurring;

  - marginal pmf/pdf, which expresses the probability or relative likelihood of each possible state of one variable actually occurring, regardless of the state of the other variable;

  - conditional pmf/pdf, which expresses the probability or relative likelihood of each possible state of one variable actually occurring given the state of another variable; and/or

  - covariance function/correlation coefficient, which expresses the relationship of the state of one variable to the state of another variable (including spatial and temporal correlation).

INDIVIDUAL PROBABILITY ASSESSMENT TECHNIQUES

The potential problems associated with a single individual subjectively developing probability assessments include the following[2]:

- **Poor quantification of uncertainty:** The assessor might not express uncertainty in a self-consistent or proper fashion. It has been shown[4] that people not trained in probabilistic analysis typically have problems in accurately quantifying their uncertainty. For example, if someone expresses a 90% probability or level of confidence that something will happen, it should happen nine out of 10 times on the average under similar circumstances. However, typically when verified, it has been shown that the event happens much less than nine (more like five) out of 10 times.

- **Poor problem definition:** The parameter for which the value is to be assessed might have been ambiguously defined so that the basis of the assessment

might not be correct. For example, in assessing the value for hydraulic conductivity with respect to groundwater flow through the site, the scale (large-scale averages versus small-scale laboratory values) might not have been specified.

- **Unspecified assumptions:** The assessor might not specify (or even be aware of) the assumptions which underlie his/her assessment so that the conditional nature of the assessment might not be apparent. For example, the assessor may have assumed porous flow through the rock mass rather than considering flow through intersecting fractures in assessing the value for average hydraulic conductivity with respect to groundwater flow through the site.

- **Uncorrected biases:** The assessor might not specify (or even be aware of) biases which underlie his/her assessment, so that the assessment does not accurately reflect the assessor's knowledge. Biases fall within various categories:

  - "Motivational", where the assessor's statements and conscious beliefs are inconsistent. Motivational biases, in turn, can be categorized as follows:

    * **"Management" bias** refers to the assessor's possible view of an uncertain variable (e.g., as an objective rather than an uncertainty). For example, if the objective is to achieve a low groundwater flow through the site, then the average hydraulic conductivity may be understated.

    * **"Expert" bias** refers to a possible reaction that the assessor may have to being considered as an expert. The assessor may feel that experts are expected to be certain of things. This bias tends to promote central bias (i.e., a tendency for the assessor to understate uncertainty). For example, the assessor may understate the range in the average hydraulic conductivity of a specific site.

    * **"Conflict" bias** refers to a reward structure that might encourage the assessor to bias the estimates high or low. For example, an unethical assessor might understate the value of a significant parameter (e.g., average hydraulic conductivity), if it was personally beneficial (e.g., to make a project appear feasible).

    * "Conservative" bias refers to the assessor's desire to err on the safe side. For example, if an event has an adverse impact, then the assessor may want to avoid underestimating the probability of that event (e.g., by consciously overstating its probability), thereby bounding the assessment rather than truthfully estimating it.

  - "Cognitive", in which the assessor's conscious beliefs do not reflect the available information. Cognitive biases, in turn, can be categorized as follows:

* **"Anchoring"** refers to the tendency of individuals to produce estimates by starting with an initial value (suggested perhaps by the formulation of the problem) and then adjusting the initial value to yield the final answer. The adjustment is typically insufficient. For example, the assessor might estimate the most likely value first and then the range in possible values, where this estimated range would probably be larger if assessed first.

* **"Availability"** (or incompleteness) bias refers to the fact that if it is easy to recall instances of an event's occurrence (e.g., the event had some personal significance to the subject), then that event tends to be incorrectly assigned a higher probability. For example, if the assessor had been involved previously with a high groundwater flows, then the resulting assessment of the average hydraulic conductivity would tend to be higher than without this experience.

* **"Base rate"** bias (or lack of moderation, law of small numbers) refers to the tendency of the assessor to focus only on specific information. Empirical evidence shows that assessors often tend to attach less importance to general information. For example, if the specific information is some recent data (e.g., the results of recent field tests), then the importance of that information might be overrated in the assessor's mind.

* **"Coherence and conjunctive distortions"** refers to the tendency of an assessor to not properly account for and combine all of the components of a problem. For example, in assessing groundwater flow where various parameters (e.g., average hydraulic conductivity, gradient) must all be within specific bounds for the flow to be acceptable, people seem especially prone to overestimating the probability that the flow will be acceptable.

* **"Representativeness"** refers to the tendency of an assessor to treat all information equally, even though it may not be statistically representative. For example, intact rock (with low hydraulic conductivity) may be more easily sampled than highly fractured rock (with high hydraulic conductiviry) so that there is a larger percentage of low hydraulic conductivities in the laboratory data base than there is in reality. If this sampling bias was not recognized, the average hydraulic conductivity might be underestimated.

* **"Overconfidence"** refers to the tendency of an assessor to underestimate the uncertainty about the value of a parameter. For example, the assessor might not recognize and properly account for other possible values of the parameter.

- **Imprecision:** The assessor may be indifferent over a specific range of values, so that there is some "fuzziness" in the assessments. For example, an assessment of 20-30% probability that something will happen should be able to be refined further with additional consideration.

- **Lack of credibility:** If the assessor cannot be considered an expert in the technical field, the assessment (regardless of the other limitations) may lack credibility. Such an assessment would not be defensible to other experts or, often more importantly, to the public. For example, a recent graduate engineer with little experience should not be making critical assessments alone.

As summarized in Table 1, the techniques available for eliminating or mitigating the potential problems associated with developing individual subjective probability assessments, include the following[2]:

- **Self-assessment:** The simplest approach to developing an individual subjective probability assessment is "self-assessment"[5,6], where the analyst interprets the available information and quantifies an assessment of the likely value and its uncertainty. The rationale behind the assessment should be well-documented, including a description of the available information and an evaluation of that information, to enhance defensibility of such subjective probability assessments. Although attractive because of its obvious simplicity, this method has significant limitations:

  - poor quantification of uncertainty;

  - uncorrected biases and/or unspecified assumptions, possibly in spite of documentation;

  - imprecision; and

  - lack of credibility, if the analyst cannot be considered an expert in the technical field.

- **Informal solicitation of expert opinion:** One of the most common methods of developing an individual subjective probability assessment consists of "informal solicitation of expert opinion"[7,8], where the analyst asks an "expert" to interpret the available information and quantify an assessment of the likely scenario or value of a parameter, and its uncertainty. The defensibility of such assessments is increased over self-assessment techniques due primarily to the increased credibility of the expert involved. As for self-assessment, the expert's rationale for the assessment should be well-documented, including a description of the information available to the expert as well as the expert's evaluation of that information, to further enhance defensibility of subjective assessments. Although generally an improvement over self-assessment techniques, due to increased credibility, informal solicitation of expert opinion has similar significant limitations, as well as increased cost and potentially poor problem definition.

TECHNIQUE	POTENTIAL PROBLEMS						
	Poor Quantification of Uncertainty	Poor Problem Definition	Uncorrected Biases/ Unspecified Assumptions	Imprecision	Lack of Credibility	Group Dynamics	Expense
INDIVIDUAL							
Self Assessment	●	○	●	●	●	NA	○
Informal Solicitation of Expert Opinion	●	●	●	●	◑	NA	◑
Calibrated Assessment	◑	●	◑	●	◑	NA	◑
Probability Encoding	○	○	○	◑	◑	NA	◑
GROUP (BEHAVIORAL)							
Open Forum	●	◑	●	●	◑	●	●
Delphi Panel	●	◑	◑	●	◑	○	●
Group Probability Encoding	○	○	○	◑	◑	◑	●
Formal Group Evaluation	○	○	○	◑	◑	◑	●

● Technique does not significantly mitigate potential problem

◑ Technique partially mitigates potential problem

○ Technique effectively mitigates potential problem

Table A-1. Evaluation of Subjective Assessment Techniques

- **Calibrated assessment:** A systematic approach to developing an individual subjective probability assessment is through the use of "calibrated assessments"[9-11], where the assessor's biases are identified and calibrated, and the assessments are adjusted to correct for such biases. Hence, two sets of assessments are required:

  - the assessor's assessment (e.g., through the informal solicitation of expert opinion); and

  - an assessment of the assessor's biases.

  The assessment of the assessor's biases can be done either: subjectively by peers (i.e., in the same way as other subjective assessments), or objectively through a set of experiments or questionnaires. The objective approach typically consists of asking the assessor a series of questions for which the true answer is available but unknown to the assessor. For example, the assessor's identified biases can be corrected in the following way:

  1. the assessor may be given a set of relevant data which does not include the direct measurement of the parameter of interest even though such a measurement exists;

  2. the assessor estimates the parameter value based on the available data;

  3. the assessor's estimate is compared with the true value, as given by the measurement; and

  4. a correction or calibration factor is determined for the assessor, which when applied to the assessor's estimate results in the true value.

  Although, a general improvement over self-assessment or informal solicitation of expert opinion techniques, due to the mitigation of some biases, calibrated assessments entail similar significant limitations (even after calibration) as well as increased costs and inherent difficulties in objectively determining calibration factors for many of the scenarios/parameters of interest, since direct measurements might never be available for verification and the calibration factor may not be constant in any case.

- Probability encoding: The most systematic and defensible approach to developing individual subjective probability assessments, but also the most expensive, is "probability encoding"[1,12-15]. In probability encoding, analysts trained in probability theory elicit in a proper and self-consistent manner a technical expert's assessment of the pdf of scenarios or a parameter value, which expresses that expert's uncertainty in the value in quantified terms. This is done in a formalized way[2] in five stages:

  1. motivating;

2. structuring;

3. conditioning;

4. encoding; and

5. verifying.

During this process the analyst attempts to:

- train the subject to properly quantify uncertainty;
- identify and minimize the subject's bias tendencies;
- define (and document) the item to be assessed in an unambiguous manner;
- elicit and document the subject's rationale, including the available information, for assessment;
- elicit (directly or indirectly) and document the subject's quantitative assessment of uncertainty and check for self-consistency; and
- verify the assessment with the subject, repeating the process if necessary.

As illustrated by the example given in Figure 2, the subject's quantitative assessment of uncertainty can be elicited indirectly by determining the probability of various states through:

- comparison with familiar reference events (e.g., poker hands); or
- choosing between two lotteries (e.g., probability wheel or intervals, Figure 2a and b), until indifference is achieved.

A cdf can then be defined, consistent with the various assessments (Figure 2c). Although a general improvement over other available methods, due to mitigation of most of the potential problems, some imprecision may remain and probability encoding is relatively costly because it is labor intensive.

CONSENSUS PROBABILITY ASSESSMENT TECHNIQUES

Differences may exist in the assessment of individuals comprising a group, which may arise from a number of sources, including[2]:

- **Disagreement on the assumptions or definitions that underlie assessments:** Individual assessments are based on specific assumptions and definitions. If

these assumptions and/or definitions differ between individuals, then the individual assessments may differ. For example, one individual may have ruled out a specific case that another individual assumes likely, or one individual may have defined the parameter of interest at a different scale than another individual.

- **Failure to overcome assessment errors and biases:** In conducting the individual assessments, a key objective is to eliminate anchoring, availability, overconfidence, and other common distortions. Training individuals and allowing them to practice making probability judgments prior to the individual assessment help to overcome biases, but such errors may persist. For example, overconfidence may have been mitigated to a large extent in one individual's assessment but not in another's, so that although the means of their probability distributions may be similar the variances may be significantly different.

- **Judgments based on differing information sources:** Both specific data and general knowledge are relevant to the encoding process. Such knowledge varies even among highly specialized experts. Specific information may vary in quantity and quality, while general information may vary due to differences in training and experience. For example, one individual may have based an assessment on a specific data set in conjunction with his/her personal experience, whereas another individual may have used a different specific data set in conjunction with different personal experience.

- **Disagreement on how to interpret available information:** The available information must be interpreted by the individuals. In this interpretation, individuals may disagree, for example, on the methods used to obtain data, the relevance of such data to the quantity being assessed, or on the appropriateness of a particular theory or model. For example, individuals may disagree on how to interpret well stem tests and their validity to assessing large-scale hydraulic conductivities at a site with respect to groundwater flow.

- **Different opinions or beliefs about the quantity of concern:** Even after agreeing on the basis for the assessment, the information available, and how to interpret this information, individuals may still have a difference of opinion. For example, individuals may arrive at different pdf's for average hydraulic conductivity, even after agreeing on all the preliminary aspects.

It is typically desirable to attempt to resolve these differences of opinion, with the following outcomes possible:

- **Convergence:** A single assessment is determined that expresses the common belief of all individuals in the group, as expressly agreed to by the group members.

- **Consensus:** A single assessment is determined, although the assessment may not reflect the beliefs of each individual; the consensus assessment may be derived from the individual assessments without the express agreement of the

individuals (forced) or it may be expressly agreed to by the group for a particular purpose (agreed).

- **Disagreement** Multiple assessments are determined where convergence or consensus on a single assessment is not possible (e.g., owing to major differences of opinion).

In general, convergence is generally most desirable, as it is most defensible, but may be difficult to achieve. Agreed consensus, i.e., with the concurrence of the group, is slightly less defensible but also less difficult to achieve. Forced consensus, without concurrence of the group, may be difficult to defend but is very simple. Disagreement may be difficult to use, as it is non-unique, but is defensible.

Techniques available for resolving differences of opinion amongst a group of individual assessors can be categorized in terms of "mechanical aggregation" and "behavioral procedures":

- **Mechanical aggregation** of individual assessments is a relatively simple approach to achieving at least forced consensus, and involves applying a mathematical formula or procedure to combine the various individual probability distributions[16,17]. If the individuals in the group agree to the resulting distribution, then agreed consensus (and possibly convergence) can be achieved. In general, mechanical aggregation techniques are most useful when the means, rather than the variances, of the individual probability distributions differ. Also, mechanical aggregation techniques can be used when a single distribution is required, but the scenario/parameter in question is not significant enough to warrant large amounts of effort to achieve convergence or agreed consensus.

  The various forms of mechanical aggregation include the following[2]:

  - Averaging, which is the simplest mechanical aggregation technique, involves simply averaging the individuals' probabilities for each possible value. Several empirical studies[18,19] have shown that averaged probabilities are often superior to individual assessments. As an example, if one individual assessed an 80% probability of the average large-scale hydraulic conductivity (log) being less than -3.5 and the other individual in the group assessed a 60% probability, then the group average would be a 70% probability.

  - Group statistics, which is a somewhat more rigorous treatment, involves determining the group's distribution of opinions regarding the probability for each value, thereby developing a "fuzzy" assessment or an assessment which corresponds to a given level of conservatism for the group. More complete statistical methods are available that incorporate dependence among variables and experts[9] and least squares or partitioning methods[20]. As an example, the statistics of the group

members' opinions of the probability of the average large-scale hydraulic conductivity (log) being less than -3.5 could be determined and used.

- **Weighting methods**, which are elaborations on either averaging or group statistics, involve the weighting of individual assessments by an external procedure to incorporate biases or differing levels of expertise among the individual assessors, similar to individual calibrated assessments. There are essentially two weighting procedures:

  * "calibration exercise", in which the natural biases and tendencies of the individual assessors are evaluated and mitigated through 1) the administration of a series of general questions to determine each assessor's ability to make correct assessments, and 2) the determination and application of weighting factors for each assessor to reflect that assessor's ability (relative to the other assessors) to make correct assessments; and

  * "peer ratings", in which each of the individual assessor's relative ability to make correct assessments is assessed subjectively by peers, although such a subjective assessment may itself introduce additional biases.

- **Behavioral procedures** can be used to attempt to develop convergence or at least agreed consensus, and involve interaction among the individuals in the group, which allows for the explicit identification and resolution of differences of opinion. Although there is evidence that such interaction results in better assessments[21,22] and that the results are generally more defensible, because the group agrees on a given distribution, behavioral procedures tend to entail significantly more effort, because the various individual assessors must be involved. Such behavioral procedures are necessary when at least agreed consensus (or disagreement) is required (i.e., for significant parameters), and are especially useful when the differences between the individual assessments are large.

  As summarized in Table 1, the various forms of behavioral procedures include the following[2]:

  - **Open forum** is a very informal means of achieving consensus and does not require prior individual assessments. The group attempts to achieve convergence or agreed consensus by open discussion of whatever each individual deems important to resolving the problem. A major limitation of this method is that the result can be distorted by the dynamics of the group, such as domination by an individual because of status or personality[23]. For example, the persuasiveness of a vocal individual or the desire of some individuals to avoid dissension may distort the results. Other potential limitations to this method are the same as for the development of individual assessments through the informal solicitation of expert opinion, i.e., poor quantification of uncertainty, uncorrected

biases, unspecified assumptions, and poor problem definition. The method is also limited by the credibility of the group members.

- **Delphi panel** is a systematic and iterative approach to achieving consensus, and has been shown to generally produce results which are reasonably reproducible across independent groups[22,24-27]. Each individual in a well-defined group is provided with the same set of background information, and is asked to conduct and document (in writing) a self-assessment. These assessments are then provided anonymously to each of the other assessors, who are encouraged to adjust their assessments in light of their peers' assessments. Typically, the individual assessments tend to converge. Such iterations are continued until either consensus is achieved or the results stabilize otherwise (i.e., disagreement). Because the Delphi technique maintains anonymity and independence of thought through physical separation of the panelists, it precludes the possibility that any one member of the panel may unduly influence the others due to actual or perceived personality dominance. Otherwise, it tends to have limitations similar to those for open forum.

- **Group probability encoding** is a formal process in which a single probability distribution is assessed directly from a group of individuals, such as for the development of individual assessments by probability encoding[1]. However, this requires the group to reach agreement on each question posed during the encoding process, which would be a difficult and tiresome procedure. As for the open forum, face-to-face interaction among participants can create destructive pressures within the group and distort the results.

- **Formal group evaluation** is a formal process of resolving differences between previously developed individual assessments[1]. This process is similar to probability encoding in that it is a joint undertaking between a trained analyst and, in this case, a group that has completed individual assessments. It consists of six steps:

  1) motivating;
  2) identifying differences in the individual assessments;
  3) discussing the basis for each individual assessment;
  4) discussing information sources and interpretations;
  5) re-encoding (if warranted); and
  6) reconciling differences).

  In this process the analyst fulfills an essential role in questioning and probing the group, helping them to understand the differences, and guiding them through the resolution process, often conducting group re-assessments. This sharing of knowledge tends to produce a commonality (i.e., in definitions, assumptions, information bases, and interpretations) that is a key step in reducing the differences between individual

assessments. As for open forums, face-to-face interaction among participants can create destructive pressures within the group and distort the results. However, the analyst can be alert to such pressures and mitigate their effects to a large extent.

RECOMMENDED PROCEDURES

As summarized in Figure 3, the recommended procedure for selecting the appropriate subjective probability assessment technique consists of the following steps[2]:

1. Prior to conducting subjective probability assessments:

   - develop the model(s) for the system of interest, e.g., a model would be needed to determine groundwater flow;

   - conduct sensitivity studies to determine the relative significance of each of the various scenarios and model parameters, e.g., sensitivity studies on the model might show that groundwater flow is very sensitive to the average large-scale hydraulic conductivity at the site; and

   - obtain the available data regarding the various scenarios/parameters, where the relative significance of each scenario/parameter will determine the appropriate level of effort in gathering data, e.g., the data on hydraulic conductivity might be limited to inference from measured physical properties, as well as generic information.

2. Each scenario/parameter to be assessed, on the basis of the model, must be defined unambiguously, e.g., considering temporal and spatial variability, as well as conditional factors (such as scale). Also, it may be useful to decompose a scenario or parameter into more elemental variables for assessment. For example, hydraulic conductivity might be defined as being large scale (i.e., averaged over 10's of meters), recognizing that the value may vary spatially within one geologic unit, as well as within the time frame of interest (i.e., 1000's of years). Hydraulic conductivity could be defined separately for the rock mass (e.g., for equivalent porous flow analyses) or for fractures (e.g., for fracture flow analyses). Hydraulic conductivity could be decomposed into permeability and viscosity.

3. The appropriate level of assessment must be determined on the basis of relative significance of each scenario/parameter to be assessed (from sensitivity studies). For example, if a parameter is relatively insignificant (e.g., density), a low level assessment (with corresponding low costs and low defensibility) would be appropriate. However, if a parameter is relatively significant (e.g., hydraulic conductivity), a high level assessment (with corresponding high costs and high defensibility) would be appropriate. For cost-efficiency, high level assessments should only be used for the most significant scenarios/parameters where high defensibility is required, thus justifying their high costs.

4. The most cost-effective assessment technique is chosen (Figure 3) on the basis of the necessary level of assessment and, in conjunction with the data base, implemented for each scenario//parameter (e.g., using specific procedures[2]).

SUMMARY AND CONCLUSIONS

Subjective probability assessments must often be made (e.g., to accurately predict performance and/or to make decisions among alternatives) wherever the data are not statistically sufficient to make objective assessments. Such subjective probability assessments must be defensible enough to adequately resolve potential controversies. The required defensibility of such assessments is proportional to the significance of each parameter being assessed (e.g., as determined by sensitivity studies).

Potential problems have been identified that are associated with developing individual subjective probability assessments and with developing consensus subjective probability assessments amongst a group that, if uncorrected, can affect defensibility of the results. The available techniques for addressing these potential problems, with varying success and effort, have been presented. Procedures for cost-effectively conducting appropriately defensible subjective probability assessments have been developed.

ACKNOWLEDGEMENT

The author is indebted to Dr. L. Merkhofer and Dr. R. Schwartz of Applied Decision Analysis, Menlo Park, California for their significant contributions on past projects[1], from which some of the concepts presented in this paper have evolved. This paper has been adapted from a paper with the same title and by the same author published in the Transportation Research Record, No. 1288, dated January 1990, pp. 183-190.

REFERENCES

1. Golder, 1987, *Groundwater and Methane Inflow Study for the BWIP ESF at Hanford, Washington,* SD-BWI-DER-001, Rev. 0, by Golder Associates, Inc., in conjunction with Applied Decision Analysis, Inc., to Basalt Waste Isolation Project, Westinghouse Hanford Company, Richland, WA.

2. Roberds, W.J., 1988, *Manual for Conducting Subjective Probability Assessments,* Draft Report, Golder Associates Inc., Redmond, WA, April.

3. Ang, A., and W.H. Tang, 1984, *Probability Concepts in Engineering Planning and Design, Vol. II: Decision, Risk, and Reliability,* John Wiley & Sons, New York, NY.

4. Capen, E.C., 1976, "The Difficulty of Assessing Uncertainty," *Journal of Petroleum Technology,* Vol. 28, pp. 843-849, August

5. Good, I.J., 1965, *The Estimation of Probabilities,* Research Monograph No. 30, The M.I.T. Press, Cambridge, MA.

6. von Holstein, C.-A.S., 1970, *Assessment and Evaluation of Subjective Probability Distributions,* Economic Research Institute, Stockholm, Sweden.

7. Morgan, M.G., M. Hennion, and S.C. Morris, 1979, *Expert Judgments for Policy Analysis,* BNL 51358, Brookhaven National Laboratory, Upton, NY.

8. Bernreuter, D.L., 1980, *Seismic Hazard Analysis: Solicitation of Expert Opinion,* NUREG/CR-1582, Vol. 3, TERA Corporation, Berkeley, CA.

9. Agnew, C.E., 1985, "Multiple Probability Assessments by Dependent Experts," *Journal of the American Statistical Society,* Vol. 80, No. 390, pp. 343-347.

10. Dawid, A.P., 1982, "The Well Calibrated Bayesian," *Journal of the American Statistical Association,* Vol. 77, No. 379, pp. 605-613, September.

11. Winkler, R.L., 1969, "Scoring Rules and the Evaluation of Probability Assessors," *Journal of the American Statistical Association,* Vol. 64, pp. 1071-1078.

12. Spetzler, C.S., and C.-A.S. von Holstein, 1972, "Probability Encoding in Decision Analysis," in Attachment F, *Example of a Probability Encoding Technique,* Stanford Research Institute, K-258802, ORSA-TIMS-AIEEE Joint National Meeting, Atlantic City, NJ.

13. Zamora, R.M., 1975, *ONR Probability Encoding Task,* working paper, SRI Project 4030, prepared for Office of Naval Research, Operational Decision Aids Project, Stanford Research Institute, Menlo Park, CA.

14. von Holstein, C.-A.S., and J.E. Matheson, 1979, *A Manual for Encoding Probability Distributions,* final report, SRI International, Menlo Park, CA.

15. Merkhofer, M.W., and P. McNamee, 1982, *The SRI Probability Encoding Process: Experiences and Insights,* technical report, SRI International, Menlo Park, CA.

16. Harman, A.J. and S.J. Press, 1975, *Collecting and Analyzing Expert Group Judgement Data,* The Rand Corporation, Santa Monica, CA.

17. Ashton, H.A., and R.H. Ashton, 1985, "Aggregating Subjective Forecasts: Some Empirical Results," *Management Science,* Vol. 31, No. 12, pp. 1499-1508.

18. von Holstein, C.-A.S, 1972, "Probabilistic Forecasting: An Experiment Related to the Stock Market," *Organizational Behavior and Human Performance,* Vol. 8, pp. 139-158.

19. Brown, R.V., 1973, *An Experiment in Probabilistic Forecasting,* R-944-ARPA, The Rand Corporation, Santa Monica, CA.

20. Lindley, D.V., A. Tversky, and R.V. Brown, 1979, "On the Reconciliation of Probability Assessments," *Journal of the Royal Statistical Society A*, Vol. 142, Part 2, pp. 146-180.

21. Seaver, D.A., 1976, *Assessment of Group Preferences and Group Uncertainty for Decision Making*, SSRI 76-4, Social Science Research Institute, University of Southern California, Los Angeles, CA.

22. Rohrbaugh, J., 1979, "Improving the Quality of Group Judgement: Social Judgment Analysis and the Delphi Technique," *Organizational Behavior and Human Performance*, Vol. 24, pp. 73-92.

23. Delbecq, A., A. Van de Ven, and D. Gustafson, 1975, *Group Techniques for Program Planning*, Scott-Foresman, Glenview, IL.

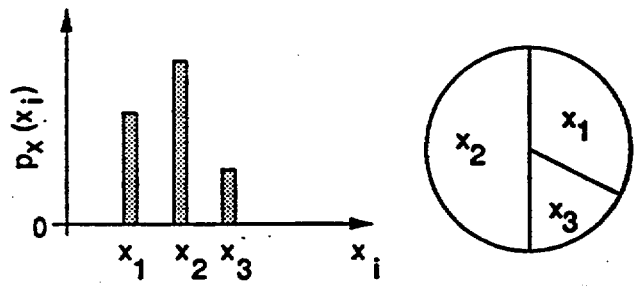
24. Dalkey, N. and O. Helmer, 1962, *An Experimental Application of the Delphi Method to the Use of Experts*, AD A081 351, Memorandum RM-727-PR (Abridged), The Rand Corporation, Santa Monica, CA.

25. Helmer, O., 1966, *The Delphi Method for Systematizing Judgments about the Future*, MR-61, University of California at Los Angeles, Los Angeles, CA.

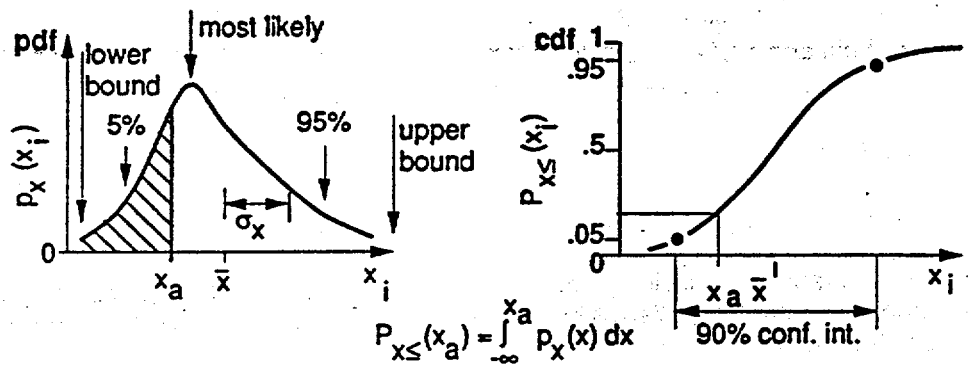
26. Helmer, O., 1967, *Systematic Use of Expert Opinions*, technical paper, The Rand Corporation, Santa Monica, CA.

27. Linstone, H.A. and M. Truoff, eds. 1975, *The Delphi Method: Techniques and Applications*, Addison-Wesley, Reading, MA.

a) Discrete Variable (e.g., a scenario)

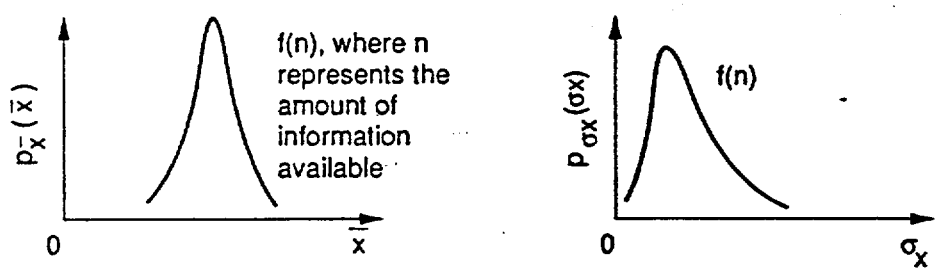


b) Continuous Variable (e.g., a parameter with a unique value)



$$P_{X\leq}(x_a) = \int_{-\infty}^{x_a} p_X(x)\,dx$$

c) Group Statistics (e.g., a parameter with a population of values)



d) Joint, Marginal, and Conditional pdf's

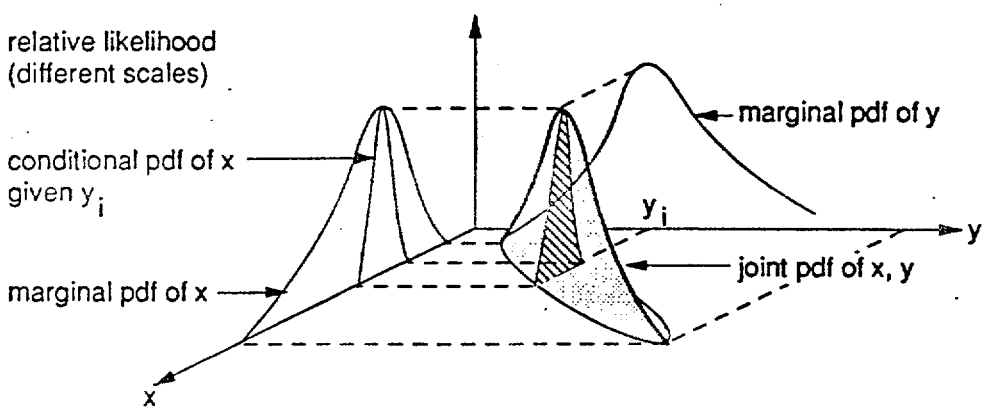
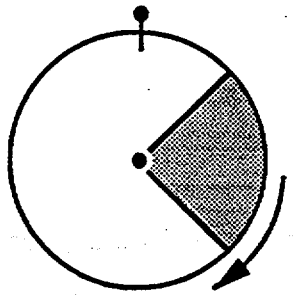


Figure A-1. Probability Distributions

a) Probability Wheel

"Which would you rather pick?"
(Change size of black area until indifferent)



"Spin and land in black area"

vs.

"The average large-scale hydraulic conductivity (log) at a site will be less than -3.1."

b) Interval Technique

"Which interval would you rather pick?" (Change threshold value until indifferent)

"The average large-scale hydraulic conductivity (log) at a site will be less than -3.5."

vs.

"The average large-scale hydraulic conductivity (log) at a site will be greater than -3.5."

c) Cumulative Distribution Function

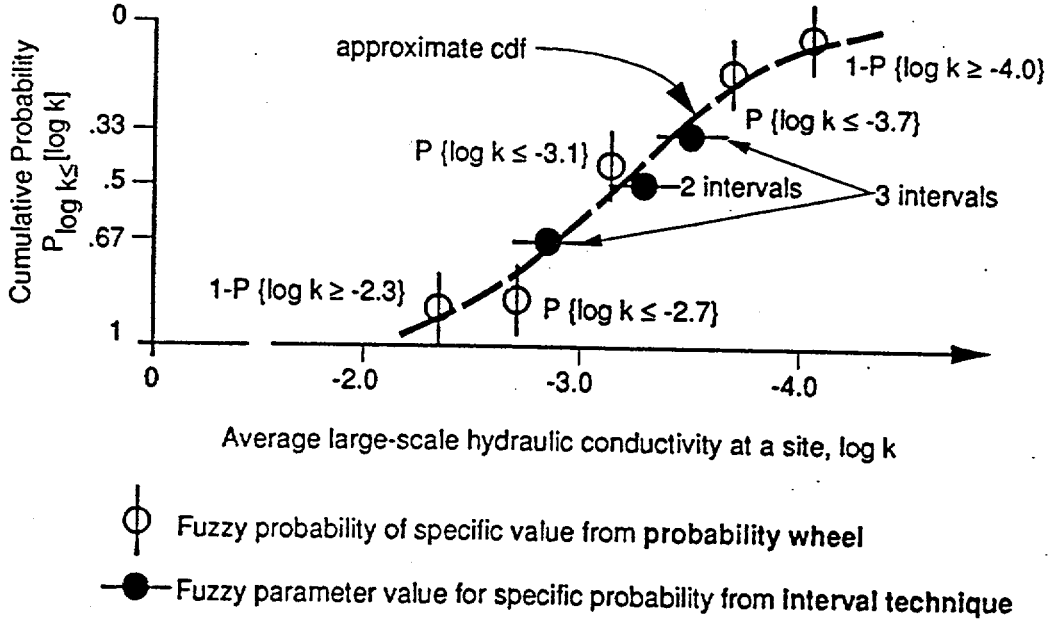


Figure A-2. Probability Encoding Examples

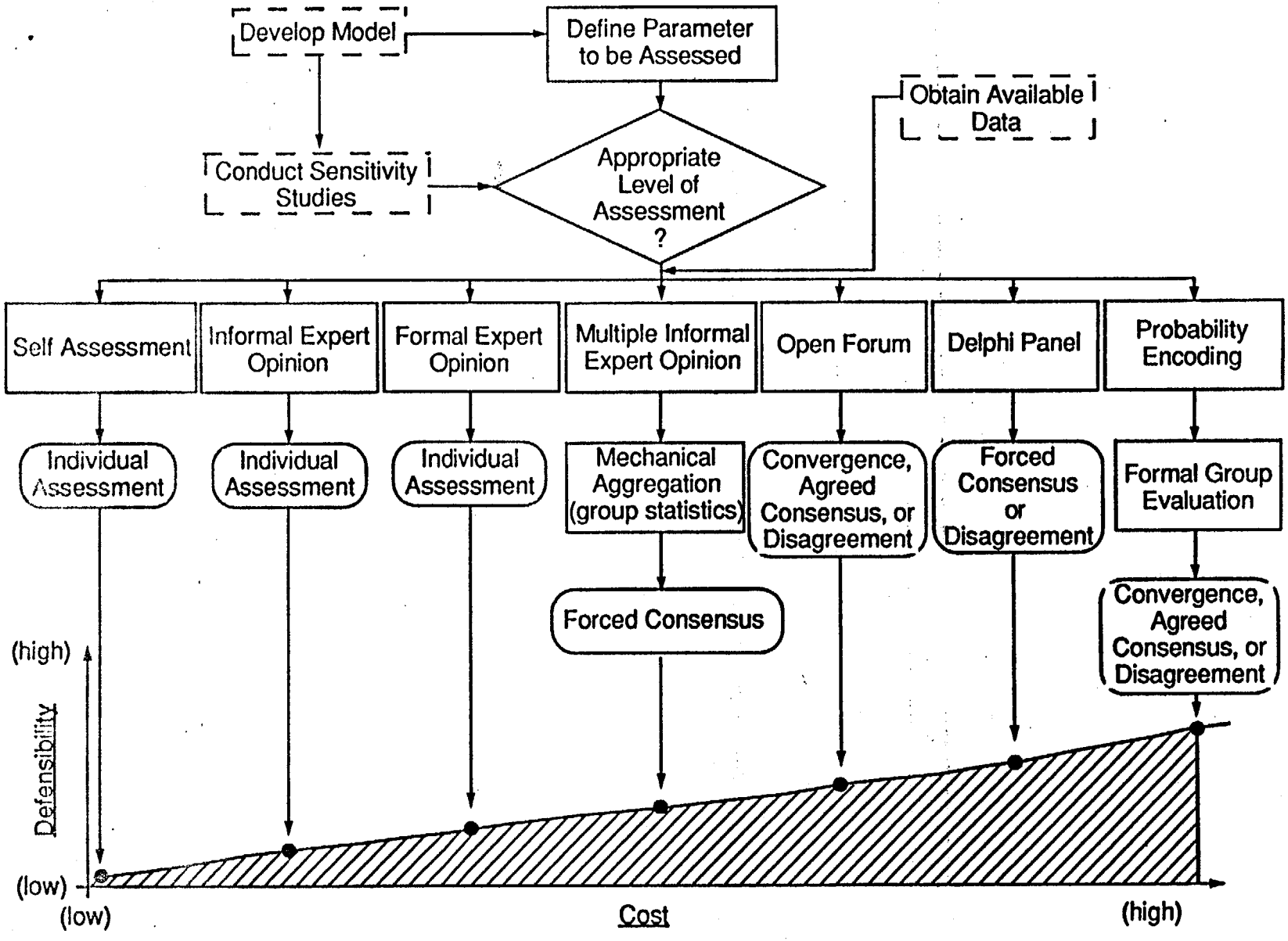


Figure A-3. Procedure for Selecting Technique

903 1371.203/3-20-92

APPENDIX B

DISCUSSION OF SCENARIO AND SIMULATION APPROACHES TO PERFORMANCE ASSESSMENT MODELING

The reader will note that the RIP model does not directly incorporate the concept of a 'scenario'. RIP treats all possible states of the repository system, and all possible future states of nature, as alternative realizations of a single 'universal' scenario. This approach differs from that taken by many other organizations, and is compared to other approaches in the following discussion.

There is a general consensus within the international community involved with long-term disposal of radioactive wastes that safety-assessment should be founded on a "scenario-consequence" modeling approach (NEA,1991). That is, a number of scenarios should be identified which represent possible futures for the repository system, and the likelihood and consequences of each scenario should be evaluated. The overall safety of the repository system would then be evaluated by combining the scenario consequences appropriately weighted by their likelihoods.

However, the term 'scenario' is used in a number of different ways, and there has been some confusion and debate as to the 'proper' definition of a scenario. There appear to be at least five points at issue:

1) Should a scenario just define the behavior of the external environment as it affects the repository, or should a scenario define the condition of both the external environment and the repository system itself?

2) Should a scenario be developed for each possible significantly different behavior of the system, or should 'archetype' scenarios be developed which each are considered to be representative of a set of similar scenarios? (For example, a 'drilling' scenario could be developed which would be considered to be representative of (or a worst-case of) all drilling intrusion cases).

3) Should scenarios be used to refer only to future changes which might affect the repository system, or should they also incorporate possible different behaviors due to uncertainty in the as-built system? (For example, there might be some possibility of an undetected fault at a site: should that be a scenario?)

4) Should scenarios or scenario-classes be defined externally to the modeling system, or generated within the stochastic model(s) itself?

5) Where there is an unresolved dispute as to which one of several alternative conceptual models is the most appropriate, is it valid to assign likelihoods to each model, and weigh the models into the combined probability distribution of system performance?

For mathematical validity, it is essential that whatever definitions are used, the set of scenarios must be comprehensive (ie. represent all significant behaviors) and mutually exclusive (ie. do not overlap).

The choice of how to interpret the above points at issue is essentially one of convenience: for each interpretation of what a 'scenario' is, it is possible to develop a completely valid system model. (It is also quite easy to develop a less than completely valid model, if sufficient care is not taken). The issues of convenience relate to such factors as ease of defining the necessary scenarios, computational expedience and efficiency, ease of interpreting the output of the models, ease of looking at 'what if' questions, etc.

Guzowski (1990) provides a synopsis of different approaches to scenario development that have been taken, and summarizes the approach used for the Waste Isolation Pilot Plant (WIPP). NRC (1990) describes an approach quite similar to that used at WIPP. The Yucca Mountain Site Characterization Plan (DOE, 1988) proposes a quite different approach, as set out in the following table.

TABLE B.1

Issue	WIPP	NRC	DOE-SCP	RIP
External/Internal Processes?	Both*	External only	Both	Both
Archetype Scenarios?	No*	?	No	No
Future Changes Only?	No	Yes	No	No
Scenarios Built-In to the Model?	No	No	No?	Yes[1]
Apply Probabilities to Alternate Conceptual Models?	No	No	?	?

Note: the above table represents the authors' interpretation of the cited references.
* no final decision made yet
? means not discussed in the reference
1 a user option in RIP

There is a general uniformity by the authors of all of the above references, in that the particular method they selected, while perhaps imperfect, is seen as avoiding the pitfalls that alternative methods will fall into. The authors of RIP feel no differently. The approach taken in RIP, direct simulation, has been suggested to have several drawbacks. The following notes address these perceived drawbacks, and the ways in which RIP attempts to minimize them.

- *Combining 'internal' with 'external' processes and events makes it difficult for the regulators and the public to directly address 'what if' questions.* This is a valid issue, as simply producing a CCDF probability distribution from a 'black box' model will not provide the degree of understanding of the overall repository system's behavior that will be needed to produce confidence. RIP's approach, however, allows this issue to be addressed in two different ways:

1) The defining probabilities can be altered by the user so that a particular event always occurs. The resulting system behavior then represents a scenario model, and can be studied and interpreted appropriately.

2) When a massive RIP analysis, involving possibly thousands of simulations is performed, the individual system simulations are not discarded. The inputs to and the response of the system for each realization are retained for further analysis. The RIP user can selectively display subsets of the results in order to further evaluate the system's performance. For example, the user could select and review the results of all simulations involving volcanic events, or all simulations involving increased precipitation. This capability should provide a very powerful ability to develop an understanding of the system's behavior.

- *Excessive numbers of simulations will be required.* This perceived drawback is not valid- if anything, RIP's approach will require the least number of simulations. In any stochastic analysis based on sampling, it is necessary to analyze a large number of samples if the 'tails' of the system probability distribution are to be evaluated. However, the number of samples required is essentially independent of the degree of complexity of the system, and depends only on how precise the probability distribution needs to be. Approaches based on separate stochastic analyses of a number of scenarios will require each scenario to be sampled repeatedly, and the results aggregated, so the overall number of simulations will be higher than with RIP's approach. Approaches based on deterministic analysis of a number of scenarios could not produce defensible probability distributions unless the number of scenarios is very large. Approaches based on deterministic, worst-case scenarios may be overly conservative, in which case a good repository would be discarded, or may founder in debate over whether they are truly a worst case.

It is important to note that a repository system does not have a finite number of discrete states: it entails a large number of continuously-varying stochastic parameters. Approaches based on 'tree'-type logic diagrams are very useful in developing an understanding of the kinds of processes that may occur, but unfortunately will not lead directly to a valid probability-tree approach to scenarios.

RIP minimizes the number of simulations required by using two separate 'importance sampling' techniques, as discussed in Appendix C. RIP does not sample all possible systems equally: it can sample certain realizations more frequently than others. The resulting probability distributions of performance are corrected to compensate for the bias in sampling. The result is that higher accuracy in the system probabilities is achieved with fewer simulations. RIP's two importance-sampling techniques are 1) selectively enhanced sampling of system histories where disruptive events occur; and 2) selectively enhanced sampling nearer to the extremes (tails) of parameter distributions. RIP also

incorporates a stratified (Latin-Hypercube) sampling scheme to further enhance the sampling efficiency.

- *The model will become impossibly complex.* There is a trade-off here: because RIP attempts to simulate all possible system behaviors, it is a sort of 'jack of all trades and master of none'. RIP has to use simplified sub-models, and the price is in an increased level of model error. In principle, the data input to RIP will be based on 'detailed' external models of components of the system, but within RIP these models may be represented by interpolation tables or curve-fits to response surfaces. It is conceivable that the performance assessment model eventually required for a license application will be more like an orchestra leader, where a driving program will invoke one or more of a suite of specialized simulators as required by the specific realization.

RIP's rather direct approach has the benefit of avoiding the problems associated with defining 'mutually exclusive' scenarios. Since RIP encompasses all futures within what is essentially a single scenario, the issue does not arise. Also, RIP avoids the issue of aggregating a set of scenarios (a 'scenario group') into an archetypical scenario. Any set of archetypical scenarios may be criticized as not being adequately refined: perhaps a slightly different combination of timing or magnitudes would produce different results. With RIP, refinement (discretization) is an automatic process, and is based only on computing an adequate number of simulations. Finally, RIP will automatically create 'combined' events or scenarios if they are credible, simply by its random sampling logic, avoiding another difficulty with scenario-driven methods.

References

DOE 1988, Site Characterization Plan, Section 8.3.5.13, U.S. Department of Energy, Office of Civilian Radioactive Waste Management, December, 1988

Guzowski 1990, *Preliminary Identification of Scenarios That May Affect the Escape and Transport of Radionuclide from the Waste Isolation Pilot Plant, Southeastern New Mexico*, Sandia report SAND89-7149, prepared by Sandia National Laboratories for the U.S. DOE.

NEA 1991, Can Long-Term Safety Be Evaluated?, Nuclear Energy Agency, Organization for Economic Co-operation and Development, Paris, 1991

Nuclear Regulatory Commission, 1990, *Phase I Demonstration of the Nuclear Regulatory Commission's Capability to Conduct a Performance Assessment for a HLW Repository*, Final Draft, April, 1990

APPENDIX C

ENHANCED SAMPLING TECHNIQUES USED IN RIP

The EPA's 40CFR191 regulation stipulates that a high-level nuclear waste repository must be demonstrated to "Have a likelihood of less than one chance in 1,000 of exceeding ten times the quantities calculated according to Table 1". For a Monte Carlo-based performance-assessment model such as RIP, this implies runs involving 10,000 or more realizations will be required. In order to get a reasonably accurate estimate of the 0.999 quantile of a result, as required by 40CFR191, as many as 100,000 realizations may be required. Even with a very fast, simple model which took just one second to evaluate a realization, 100,000 realizations would take over a day to perform. Thus, it is important to seek ways to decrease the number of realizations required to reach a certain desire accuracy.

RIP minimizes the number of Monte Carlo simulations required to construct a CCDF by using two separate 'importance sampling' techniques. These techniques work so that RIP does not sample all possible systems equally: it samples certain realizations more frequently than others. The resulting probability distributions of performance are corrected to compensate for the bias in sampling. The result is that higher accuracy in the system probabilities is achieved with fewer simulations. RIP's two importance-sampling techniques are 1) selectively enhanced sampling of system histories where disruptive events occur, and 2) selectively enhanced sampling nearer to the extremes (tails) of parameter distributions. This appendix discusses the mathematical basis of each of the two methods, and presents numerical results demonstrating their effectiveness.

In addition to importance sampling, RIP offers the user an option of conventional or Latin-Hypercube sampling of stochastic parameters. The Latin-Hypercube option, described in Section C.3, offers some additional advantages over conventional Monte Carlo sampling.

C.1 Importance-Sampling of Events

Because RIP represents disruptive events as Poisson processes, it is not appropriate to do performance-sampling of events by simply increasing the rate of occurrence of the events. For example, having volcanoes erupt every year, but with a low 'weight', is physically unreasonable. Instead, RIP uses a 'pruning' process to selectively discard a fraction of the less-important realizations. As a simple example, the user could choose to model every realization which involved a volcanic event, but to discard nine out of ten non-volcano realizations. RIP automatically carries out this pruning, and corrects the resulting CCDFs by weighing the resulting realizations inversely by their pruning factor. That is to say, if a particular event-class was pruned nine times out of ten, then for each realization that was retained the result would be processed as if it had occurred ten times.

The following table shows an example of this approach.

Sampling Class	Fraction Kept (un-pruned)	Class Criteria	Weight Applied (Pruning Factor)
I	1	Volcano or Earthquake	1
II	0.4	Human Intrusion Events	2.5
III	0.1	All others	10

Table C.1 Example Event-Pruning Table

The benefit of this approach is that unusual events will be sampled much more frequently than under the normal Monte Carlo approach, and so their likelihood and consequences will be represented much more precisely in the resulting CCDFs.

For example, suppose that volcanic events normally occur at a rate of one every 10,000,000 years. For 10,000-year simulations, a volcanic event would normally occur only once every 1,000 realizations. A Monte Carlo simulation using 1,000 realizations would have only a 0.63 likelihood of realizing one or more volcanic events, and therefore would have an extremely poor representation of volcanic events in its CCDF. Importance sampling where volcanic events were selected ten times more frequently than other cases would on average result in ten volcanic events over a thousand realizations: enough for a rough approximation of their effects. More realizations, or a higher selectivity of volcanic events, would produce an even better result.

It should be noted that there can be too much of a good thing, and excessive pruning of the base case would result in degradation of the base-case portion of the CCDF, so that the computed expected value or first moment of the distribution would be poor. For example, a pruning-factor of 1000 in the above example for "all others" would have resulted in only one in a thousand realizations of the base case being retained. This clearly would be too few.

The RIP user defines event importance-sampling in a simple manner. Similarly to the preceding table, the user creates event-classes, which jointly span all possible histories. Each event class is given a pruning factor, and a list of which events define each class. The pruning and weighing operations are performed automatically by RIP.

As a simple demonstration of event importance-sampling, a system was defined whose 'base-case' performance was represented by a normal distribution with mean 0 and standard deviation 1. Two events were defined, as follows:

- Event 1 had an expected number of occurrences equal to 0.01, with the consequence having a normal distribution with a mean of 4 and a standard deviation of 0.4,

In each case, if the event occurred then the event consequence was added to the base-case consequence. This system was sampled 1,000 times using pruning-factors of 1, 10, and 100 for the base case, with the resulting density functions shown in the following figure. For comparison, the result of a conventional Monte Carlo analysis using 10,000 realizations is shown. Note the low probability occurrences of multiple events apparent in Figure C-1. As can be seen in Figure C-1, the importance-sampling approach radically improved the ability of the Monte Carlo method to represent the consequences of low-likelihood events.

C.2 Importance-Sampling of Tails of Distributions

C.2.1 Theory

In the Monte Carlo simulation process, each random variable is normally sampled purely randomly or by Latin-hypercube sampling. However, in performance assessment the area of interest is usually the extreme tail of the results distribution. In order to increase the resolution of the tail for a given number of system realizations, RIP uses an importance-sampling technique which can sample the extremes of the random variables at an enhanced rate. To counteract the increased sampling rate, a sampling weight is calculated for each variable, where the weight is inversely proportional to the degree of sampling enhancement. Each realization of the entire system has a weight equal to the product of the weights of the individual, independent stochastic parameters.

For a random variable x with probability density f(x), normal Monte Carlo sampling will yield a sample within the range (x,x+dx) about f(x)dx fraction of the time, for a large number of samples and a small dx. In the importance sampling scheme, a sample in the range (x,x+dx) is generated b(x)f(x)dx fraction of the time, with an associated weight of 1/b(x). b(x) is termed the 'bias function' in RIP.

In the conventional Monte Carlo process, a random number 'r' is selected from the uniform distribution (0,1), and is used as the probability level in the cumulative distribution function F(x). Thus, the normal Monte-Carlo selection is $x = F^{-1}(r)$.

RIP transforms r to 'u', the 'used' random number, as follows:

$$u = 0.5\,(2r)^n$$

where n (n greater than or equal to 1) is a factor that controls the extent of the biassing towards the tails. The weight associated with $F^{-1}(u)$ is simply du/dr:

$$w = du/dr = n(2r)^{n-1}$$

Note that this approach amplifies both tails of the probability distribution equally. This is because we do not know ad hoc which tail is of the most interest. The form selected for u is not very strong, and there may be cases where a transformation which is stronger, or which amplifies just the upper (or lower) end of the distribution would be preferred.

Graphically, we are simply transforming r into u, as shown in Figure C-2.

The following table of u and w shows how the amount of biassing changes with different values of n:

	r=0.01		r=0.1		r=0.5	
n	u	w	u	w	u	w
1	0.01	1	0.1	1	0.5	1
1.33	0.0028	0.3658	0.0588	0.7820	0.5	1.33
1.67	0.0007	0.1215	0.0340	0.5681	0.5	1.67
2.0	0.0002	0.0400	0.0200	0.4000	0.5	2.0

Table C.2 Values of the 'used' random number u and its weight w.

n is the bias exponent
r is the original random number which was selected
u is the 'used' random number
w is the weight
Note: the table does not show values of r>0.5, as the weighing is symmetrical about r=0.5.

The net effect of the importance-sampling technique is to develop relatively more realizations at the ends of the probability distributions. Where the goal is to develop an accurate measure of the cumulative probability at a very high consequence level, this technique should reduce the number of realizations required.

In RIP, any stochastic variable can be assigned an importance-sampling bias factor, using the terminology 'some bias' for n=1.33, 'moderate bias' for n=1.67, and 'maximum bias' for n=2.

C.2.2 Test Problem

A test was carried out involving realizing the function $y = x_1 + x_2e^{x3}$ a thousand times. The input parameters were defined as follows: x_1 had a normal distribution with mean 5 and standard deviation 2, x_2 had a normal distribution with mean 1 and standard deviation 0.5, and x_3 had a triangular distribution with lower bound 0, upper bound 3, and most likely value 1.

The resulting distribution of y was sampled at a number of probability levels.. The entire process was repeated twenty times, and the statistics of the resulting distributions of the sampled quantiles were calculated. This was done for two different values of the biassing parameter, n: n = 1 (no biasing), and n = 2 (maximum biasing). The results, as shown in Figure C-3, demonstrate a dramatic improvement in the tails at the higher value of n. For comparison, the result using Latin-Hypercube sampling (without any importance sampling) are also shown.

C.3 Latin-Hypercube Sampling (LHS)

RIP also presents an option to implement a Latin-Hypercube sampling (LHS) scheme. The LHS option results in forced sampling from each "stratum" of a parameter. The parameter's distribution is divided into up to 250 equally likely strata or slices. The strata are then "shuffled" into a random sequence, and a random value is then picked from each stratum in turn. This approach ensures that a uniform spanning sampling is achieved.

Up to 250 strata are used, depending on the number of realizations to be done. Where more than 250 realizations are specified, sets of strata are defined as shown in examples in the following table:

Number of LHS Strata		
# of realizations	# of sets	# of strata per set
100	1	100
200	1	200
250	1	250
251	2	126
1000	4	250
3333	14	239

Table C.3 Latin-Hypercube Strata

When more than one set of strata are used, each parameter is randomly re-positioned in its sampling sequence after each set is completed.

LHS appears to have a significant benefit for problems involving only a few independent stochastic parameters, and with moderate numbers of realizations. In no case does it perform worse than true random sampling, and accordingly LHS sampling is the default for RIP.

Note that Latin-Hypercube sampling is not meant to be an alternative to importance sampling. Rather, importance sampling can be implemented simultaneously with Latin-Hypercube sampling to further augment the sampling scheme. In general, Latin-Hypercube sampling is effective at delineating the base-case portion of a stochastic result (i.e., the expected value or first moment). It is not efficient at sampling the tails of distribution. Importance sampling, however, is designed to effectively sample the low probability tails. Hence, a combined Latin-Hypercube/importance sampling scheme is likely to be the most efficient sampling approach. Sampling schemes are discussed in more detail in the RIP User's Guide.

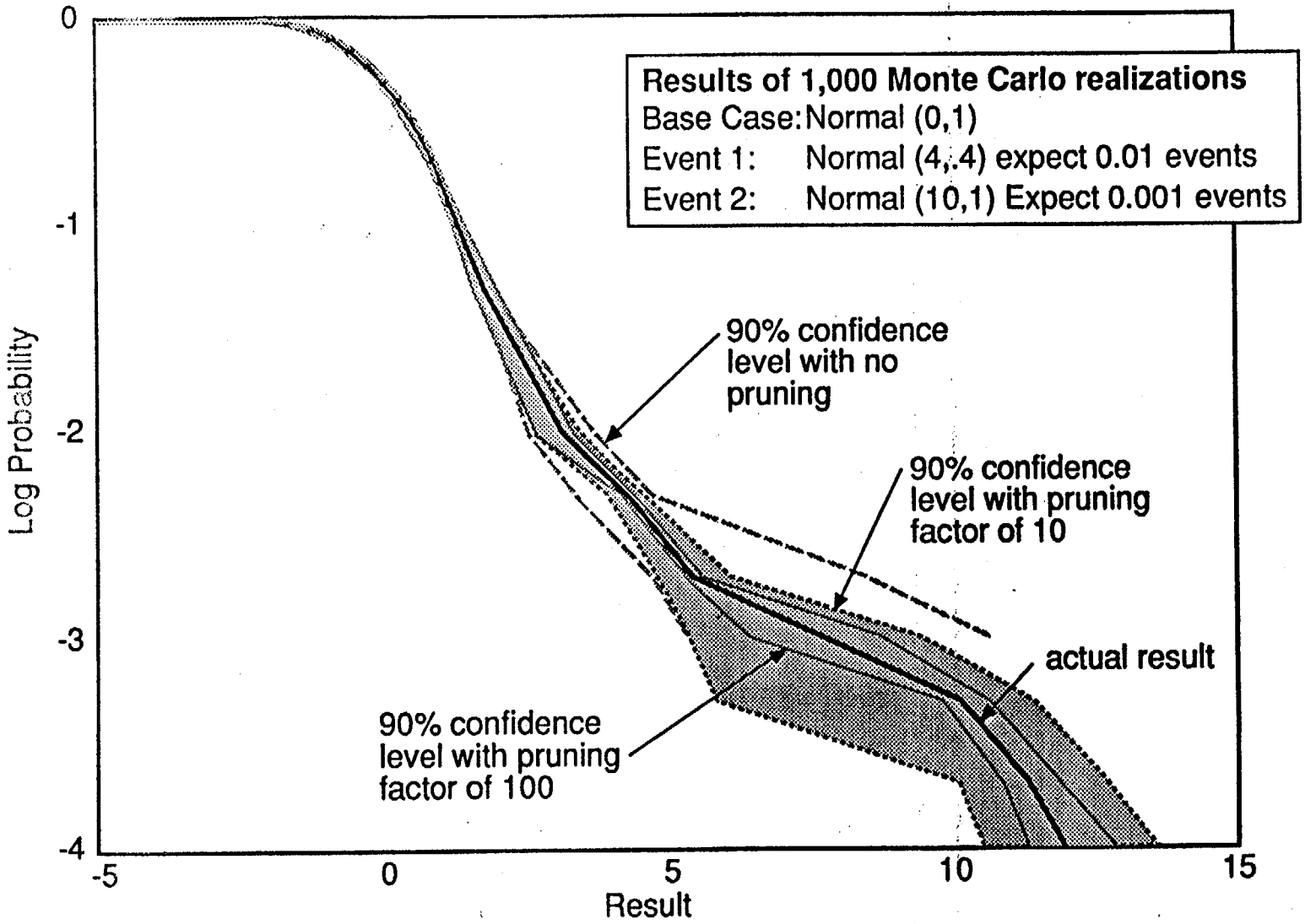


FIGURE **C-1**
EVENT IMPORTANCE SAMPLING - RESULTS
ARGONNE/MODEL DEVELOPMENT/WA

PROJECT NO. 903 1371.203 DRAWING NO. 26741 DATE 12/18/91 DRAWN BY EA

Golder Associates

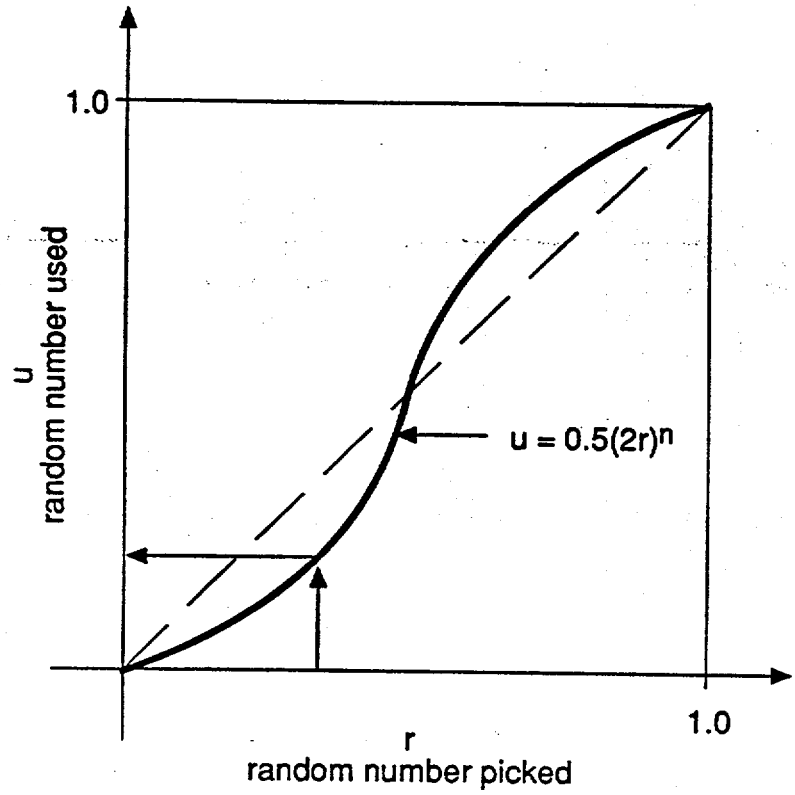


FIGURE C

PARAMETER IMPORTANCE - SAMPLING ALGORITHM

ARGONNE/MODEL DEVELOPMENT/WA

PROJECT NO. 903 1371.203 DRAWING NO. 26725 DATE 12/18/91 DRAWN BY EA

Golder Associates

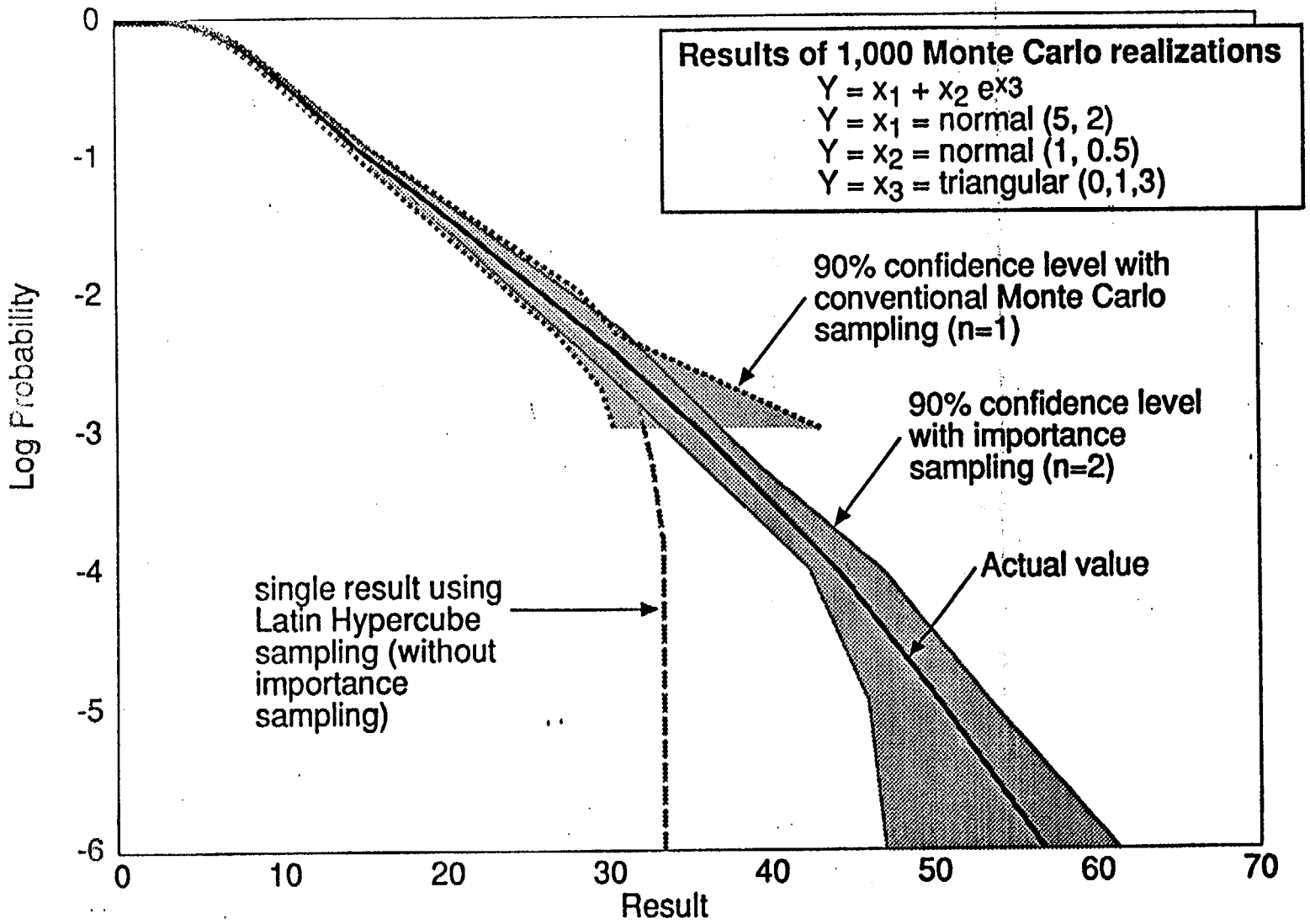


FIGURE **C-3**

PARAMETER IMPORTANCE SAMPLING - RESULTS

ARGONNE/MODEL DEVELOPMENT/WA

PROJECT NO. 903 1371.203 DRAWING NO. 26732 DATE 2/21/92 DRAWN BY EA

Golder Associates

APPENDIX D

RADIOACTIVE DECAY ALGORITHM

This appendix explains the general algorithm for radioactive decay used in the RIP model. We begin by considering the mass balance for a single radionuclide in a closed system. This equation is given as follows:

$$\frac{dM_i}{dt} = -\lambda_i M_i + \sum_j^K \lambda_j M_j \qquad \text{(D-1)}$$

where

- M_i is the mass of the ith radionuclide,
- M_j is the mass of the jth parent to radionuclide i,
- $\lambda_{[i,j]}$ is the radioactive decay constant for the ith or jth radionuclide, and
- t is time.

In the RIP model, we solve Equation D-1 for M_i at time t using a backward difference approximation for the derivative. Replacing dt with Δt, equation K-1 is solved for $M_i(t)$ as follows:

$$M_i(t) \approx (1-\lambda_i \Delta t) M_i(t-\Delta t) + \Delta t \sum_j^K \lambda_j M_j(t-\Delta t) \qquad \text{(D-2)}$$

where

- M_i (t) is the mass of the ith radionuclide at time t, and
- M_i (t-Δt) is the mass of the ith radionuclide at time t - Δt.

The approximate solution for $M_i(t)$ given by equation D-2 has error of order Δt. The actual error is also influenced by the radioactive decay constants. For small decay constants (long half-lives), a larger value of Δt may be used than for more rapidly decaying radionuclides.

For a large inventory of radionuclides, including decay chains of several members in length, it is convenient to write the individual equations D-2 as a linear system. The linear system is given by the following:

$$[A]\ M(t-\Delta t) = M(t) \qquad \text{(D-3)}$$

where

- [A] is the coefficient matrix,
- M(t-Δt) is the radionuclide mass vector at time t-Δt, and
- M(t) is the radionuclide mass vector at time t.

Obtaining the values of M(t) is the goal of equation D-3.

The matrix elements of [A] are given as follows:

$a_{ij} = 1-\lambda_i\Delta t$, if i = j
$a_{ij} = \lambda_j\Delta t$, if i <> j and j is a parent of i, and
$a_{ij} = 0$ if i <> j and j is not a parent of i.

In executing the RIP model, radionuclide decay calculations are made on every time step. The time step size is commonly 100 years, however, this is not mandatory, and it could vary by orders of magnitude. Consequently, we do not use the time step size in the decay calculation. Rather, we use a potentially much smaller time increment, repeating the calculation until the time step decay is computed. Thus, for a time step of size Δt, we choose a smaller time increment δt, and repeat the calculation $N = \Delta t/\delta t$ times for each time step. The calculation by this method may be expressed as follows:

$$[D]M(t-\Delta t) = M(t) \tag{D-4}$$

where
[D] is equal to $[A]^N$

A preprocessing step in the RIP algorithm computes the matrix [D]. Decay is then computed at each later time step by carrying out the matrix multiplication $[D]M(t-\Delta t)$ to obtain M(t).

As a matter of efficiency, N multiplications of [A] are not actually completed to obtain [D]. Rather, we select δt, such that $\Delta t/\delta t = N = 2^L$. Subsequently we solve $[A][A]=[A]^2$, $[A]^2[A]^2=[A]^4$, etc. The selection of δt is made as follows:

step 1: assume a value $\delta t = 1$ year,
step 2: solve for L by $\ln(\Delta t)/\ln(2) = L$,
step 3: round L upward to the nearest integer, and
step 4: compute $\delta t = \Delta t/2^L$.

For a time step size of 100 years, $\delta t = 0.78$ years and L = 7. Thus, only 7 matrix multiplications are completed to compute $[A]^{128}$.

Note that the current algorithm does not allow a parent to split into more than one daughter. That is, a daughter can have more than one parent, but a parent can have no more than one daughter.

APPENDIX E

COMBINING WASTE PACKAGE FAILURE MODE DISTRIBUTIONS

In Chapter 3, waste package failure models of the RIP model are discussed in detail. Generally, the failure models for either canisters or cladding include more than one failure mechanism, or mode. This appendix presents the derivation of the failure model used in RIP, which consists of a combination of one or more modes.

Let

N	=	number of failure modes;
$f_i(t)$	=	frequency of failure by mode i at time t;
$F_i(t)$	=	probability of failure by mode i by time t;
p_i	=	fraction of waste packages which can fail by mode i;
f(t)	=	frequency of failure by all modes combined at time t; and
F(t)	=	probability of failure by all modes combined by time t.

By definition,

$$F_i(t) = \int_0^t f_i(\theta)d\theta \tag{E-1}$$

and

$$F(t) = \int_0^t f(\theta)d\theta \tag{E-2}$$

We would like to develop an expression such that the combined distributions f(t) and F(t) can be expressed in terms of the distributions for the modes, $f_i(t)$ and $F_i(t)$. The derivation is straight forward.

$$1 - p_i\, F_i(t) = \textit{probability of surviving mode } i \textit{ by time } t \tag{E-3}$$

Assuming that the failure modes are uncorrelated, it follows that

$$\prod_{i=1}^{N} [1 - p_i F_i(t)] = \textit{probability of surviving all modes by time } t \qquad \text{(E-4)}$$

and

$$F(t) = 1 - \prod_{i=1}^{N} [1 - p_i F_i(t)] = \textit{probability to not survive all modes by time } t \qquad \text{(E-5)}$$

In the RIP model, only F(t) is used in the calculations. Where needed, f(t) is approximated by $[F(t) - F(t-\Delta t)]/\Delta t$. This approximation ensures the pdf integrates to 1, as required. This approximation eliminates balancing errors which may occur due to a poor approximation of f(t) over the interval $(t, t + \Delta t)$.

APPENDIX F

USING THE EXPONENTIAL, WEIBULL, AND OTHER DISTRIBUTIONS TO REPRESENT WASTE PACKAGE FAILURE

Container and cladding failure rates will be obtained from probability density functions. The failure rates of either containers or cladding will be based on estimated probability density functions for the lifetime of containers and the lifetime of cladding. The failure rates at any time during a simulation are equal to the lifetime probability density function value at that time.

The RIP model currently accommodates four distribution types for describing waste package failure: exponential, weibull, uniform, and degenerate. The use of these distributions is discussed below.

Exponential Distribution

The exponential distribution is fully described by a single parameter. This parameter, λ, is the inverse of the expected lifetime and appears in the exponential probability density function (pdf) as follows:

$$f(t)=\lambda e^{-\lambda t} \qquad t \geq 0, \quad \lambda > 0 \tag{F-1}$$

where the random variable t is the lifetime (Hogg and Craig, 1978). The mode of the exponential distribution always occurs at x = 0, and the variance is equal to the square of the expected value. Hence, the expected value and standard deviation are also equal.

The value of f(t) is the frequency of lifetimes equal to t. Thus, in a population of canisters of size N, we expect Nf(t)dt of the containers to be breached in a time interval from t to t+dt. Hence, f(t) is a failure rate for time interval dt. Likewise, for a container containing Nr spent fuel rods we expect Nr x f(t)dt of the rods to be breached (ie. undergo cladding failure) in the time interval from t to t+dt. Figure F-1 shows Equation F-1 for various values of λ.

The determination of λ may be simplified by considering the cumulative probability density function (cdf), which is given as follows:

$$F(t)=1-e^{-\lambda t} \qquad t \geq 0, \quad \lambda > 0 \tag{F-2}$$

The cdf relates a total fraction to a time given a value of λ. For example, if a project scientist can estimate that 50% of the containers will be breached in 1,000 years due to engineered and/or handling related defects, λ can be solved for directly from Equation F-2. For this example,

$$\lambda=\frac{\ln(50\%/100)}{1{,}000} = -0.000693 \ yr^{-1} \tag{F-3}$$

To aid this elicitation procedure, it is also worth noting that 63% of the canisters will be breached at the expected lifetime, 86% at a time equal to twice the expected lifetime, and 95% at a time equal to three times the expected lifetime. This procedure may be used to obtain bounds on the range of λ.

Weibull Distribution

The Weibull distribution is more complex than the exponential distribution, requiring the input of three parameters rather than one. The pdf for the Weibull distribution is given as follows:

$$f(t)=\frac{\alpha}{\beta-\epsilon}\left(\frac{t-\epsilon}{\beta-\epsilon}\right)^{\alpha-1} e^{-\left(\frac{t-\epsilon}{\beta-\epsilon}\right)^{\alpha}} \tag{F-4}$$

$$t \geq \epsilon,\ \beta > \epsilon,\ \alpha > 0$$

where the random variable t is the lifetime (Benjamin and Cornell, 1970). The parameter ϵ is a minimum lifetime value at which f(t) = 0. The expected and modal values of t are proportional to β, while the variance of t is proportional to β^2. The parameter α relates to the shape of f(t) and influences the expected value, variance and mode.

The expected value and variance of t in the Weibull distribution are given as follows:

$$E(t)=\epsilon+(\beta-\epsilon)\Gamma(1+\frac{1}{\alpha})$$

$$Var(t)=(\beta-\epsilon)^2[\Gamma(1+\frac{2}{\alpha})-\Gamma^2(1+\frac{1}{\alpha})] \tag{F-5}$$

The most probable value of t, or distribution mode, is given by:

$$\tilde{t}=\epsilon+(\beta-\epsilon)\left(\frac{\alpha-1}{\alpha}\right)^{\frac{1}{\alpha}} \qquad \alpha > 1$$

$$\tilde{t}=0 \qquad \alpha \leq 1 \tag{F-6}$$

(Benjamin and Cornell, 1970).

Figure F-2 shows Equation F-4, for various values of α, when β = 5,000 years and ϵ = 1,000 years. As α increases from 0.5 to 7 the distribution form shifts from that similar to the exponential function to a more bell-shaped curve. In fact, for α = 1., the Weibull pdf reduces to the exponential pdf discussed above with $\lambda = 1/(\beta-\epsilon)$. As alpha continues to

increase to $+\infty$, f(t) becomes a spike centered very close to β. For large values of α, the variance of t asymptotically approaches 0, causing the distribution to become essentially degenerate.

Figure F-3 shows the behavior of the mean (expected value), variance, and mode of lifetimes following the Weibull distribution, as functions of the parameter α. The ordinate axis, labeled parameter weightings, provides the values of the right-hand sides of Equations F-5 and F-6 when they are arranged to contain only the parameter α.

It is important to note from these curves, that the mode and mean are equal only when $\alpha = 3.345$. When $\alpha = 3.345$ the mean value is equal to $\epsilon + (\beta-\epsilon)\cdot 0.898$. To the left of $\alpha = 3.345$, the mode occurs to the left of the mean, whereas to the right of $\alpha = 3.345$, the mode occurs to the right of the mean. For values of $\alpha > 1$, the lifetime variance rapidly approaches 0.

The elicitation of the Weibull distribution parameters may likely be pursued in many ways. One method for their elicitation is the following:

1) Determine ϵ independently from α and β. ϵ represents the first possible time of failure due to conditions in the repository. Consequently, it can be interpreted as independent of α and β.

2) Determine the distribution of α with the aid of Figure F-3. Based on the desired population of distribution <u>shapes</u>, a distribution for α may be estimated.

3) Evaluate the distribution of <u>mean</u> lifetime assuming $\epsilon = 0$. A mean lifetime distribution for $\epsilon > 0$ may be obtained by adding ϵ to the mean when $\epsilon = 0$. The difference, $\beta-\epsilon$, which fully determines the Weibull distribution, is obtained from the mean lifetime using Equation F-7 below.

$$\beta-\epsilon=\frac{E(t,\epsilon=0)}{\Gamma(1+\frac{1}{\alpha})} \qquad \text{(F-7)}$$

Uniform Distribution

The uniform distribution is usually described by two parameters, a and b, which define the interval of the random variable over which the distribution has a density greater than 0. The probability density for any value of the random variable within the closed interval [a,b] is equal to 1/(b-a), as shown on Figure F-4. Although the uniform distribution is conventionally not described in terms of its mean and variance, these statistics are equal to (b+a)/2 and $(b+a)^2/12$, respectively.

With respect to container and cladding lifetime, we have chosen to define the uniform distribution in a slightly different but equivalent manner. Rather than specifying the interval endpoints, we specify the first possible lifetime, equivalent to the parameter a, and then the duration of positive probability density. This duration is equivalent to the difference b-a. These two parameters fully define the uniform distribution.

Degenerate Distribution

The degenerate distribution has all its probability density concentrated at a single value of the random variable, ie. there is only one possible value of the random variable which can occur (Figure F-5). The value taken on by the random variable is the mean value, or expected value, and the variance is 0.

This distribution is used to cause total failure of a fraction (0-1) of containers and/or cladding at any time during the simulation. The impact of disruptive events on waste packages is simulated by using this distribution with an adjusted density at the time of the event. The adjusted density reflects the magnitude of the event with respect to container and cladding failure.

References

Benjamin, J.R. and C.A. Cornell, 1970, Probability, Statistics, and Decision for Civil Engineers, McGraw-Hill Book Company, New York.

Hogg, R.V. and A.T. Craig, 1978, Introduction to Mathematical Statistics, 4th Edition, Macmillan Publishing Co., Inc., New York.

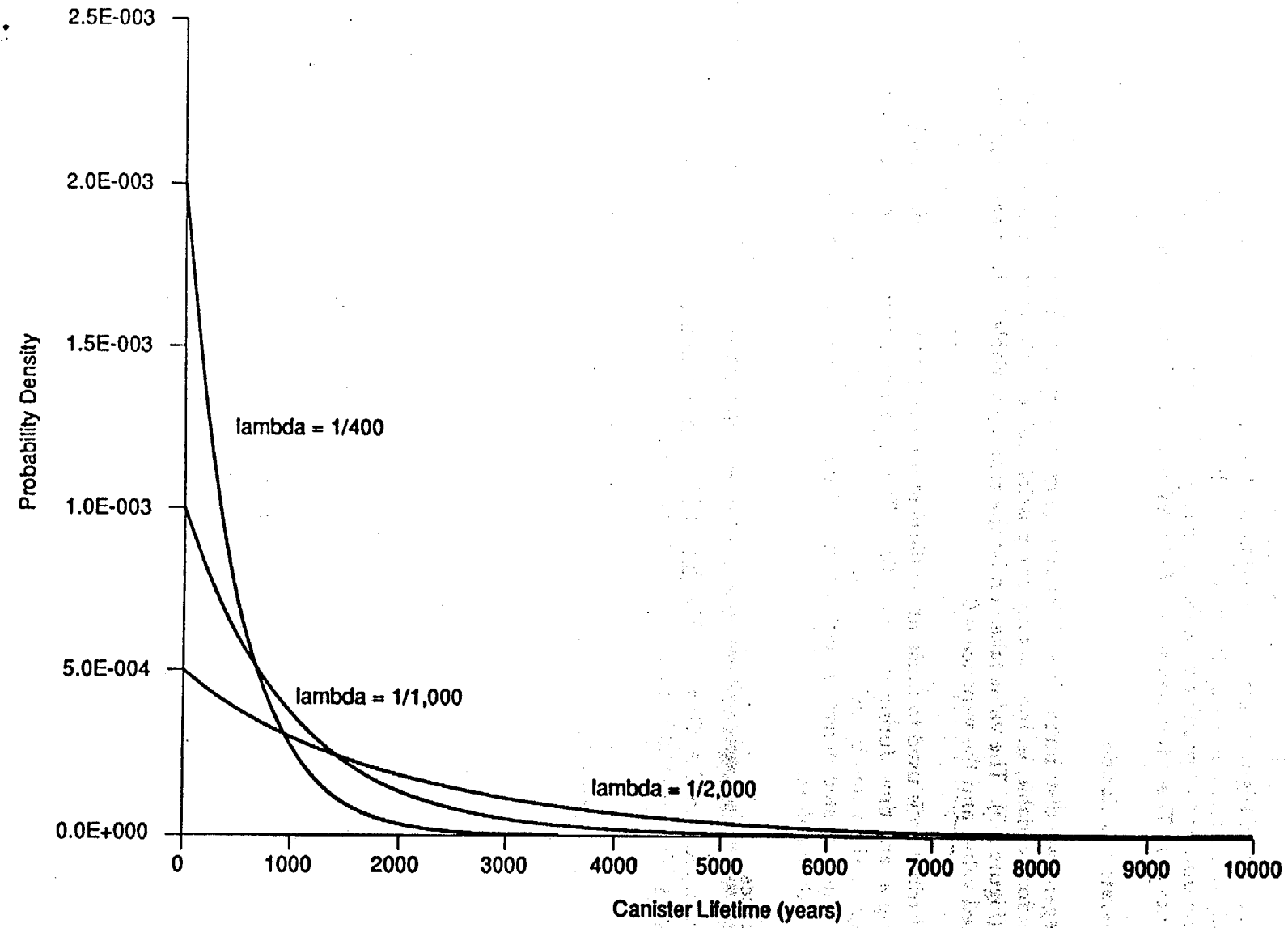


FIGURE **F-1**

EXPONENTIAL PROBABILITY DENSITY FUNCTIONS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 25550 DATE 6/20/91 DRAWN BY CB

Golder Associates

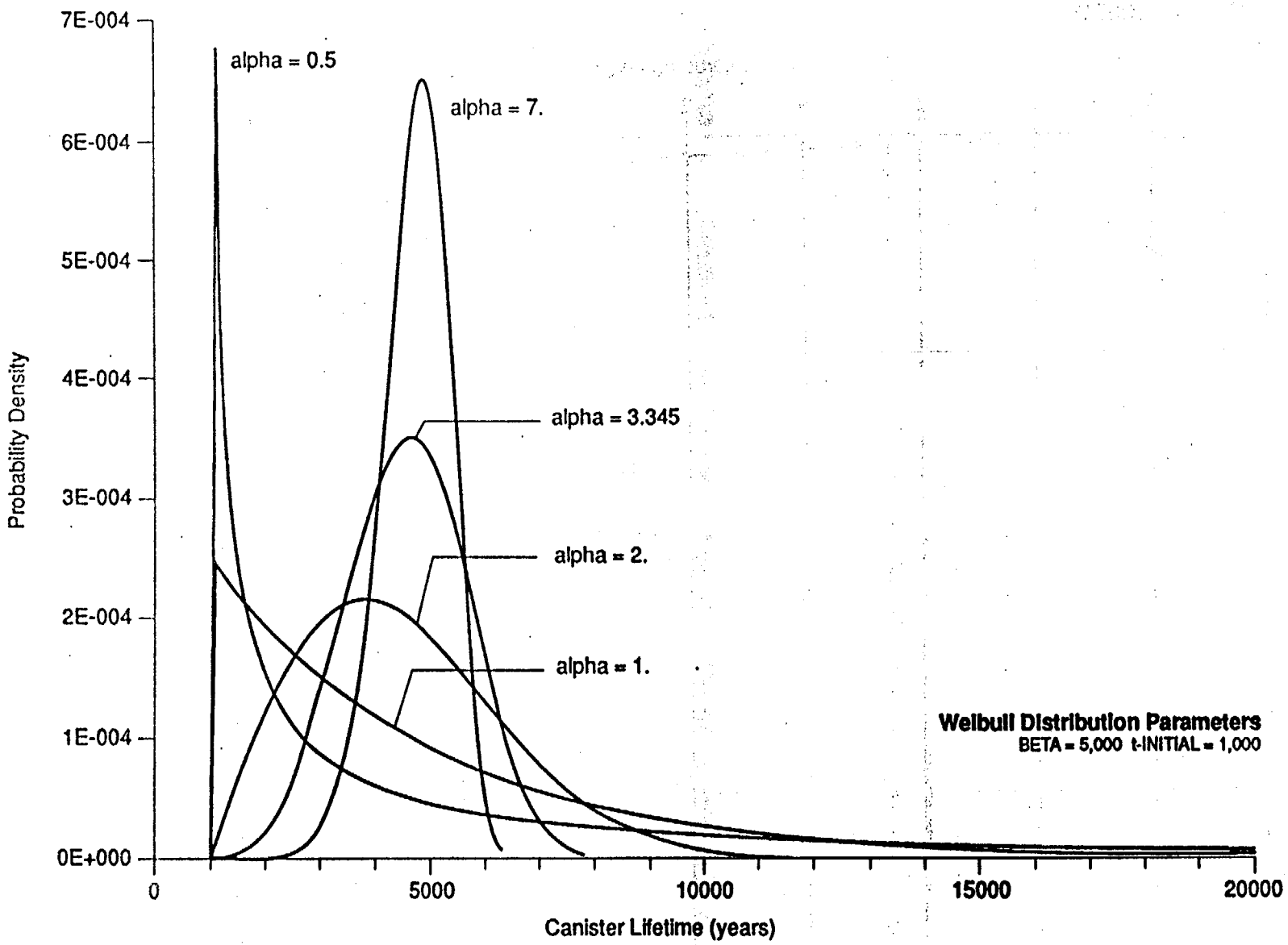


FIGURE **F-2**
WEIBULL PROBABILITY DENSITY FUNCTIONS
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 25549 DATE 6/20/91 DRAWN BY CB

Golder Associates

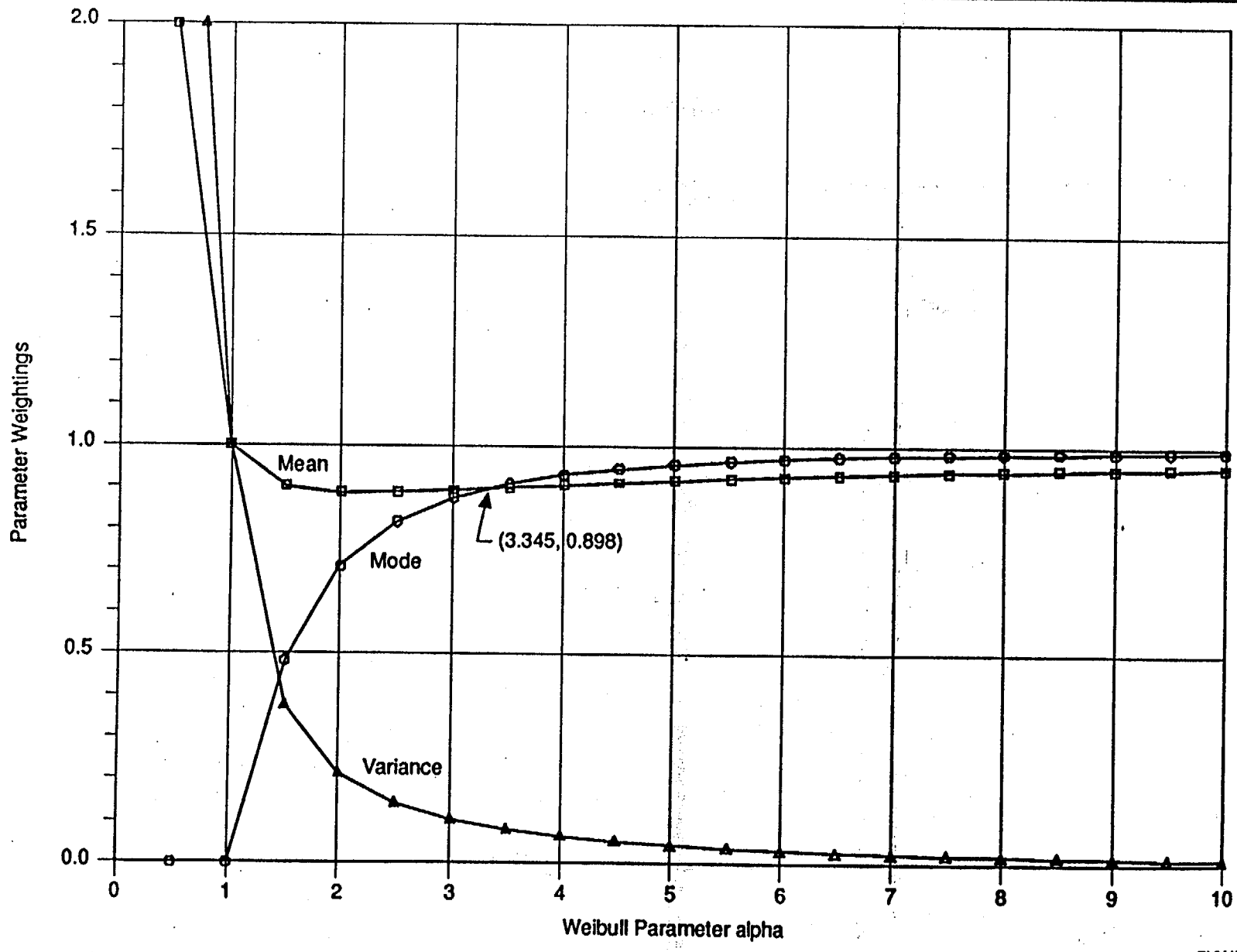


FIGURE **F-3**

WEIBULL DISTRIBUTION PARAMETER WEIGHTINGS

ARGONNE/MODEL DEVE MENT

PROJECT NO. 903-13/1.203 DRAWING NO. 25551 DATE 6/20/91 DRAWN BY CB

Golder Associates

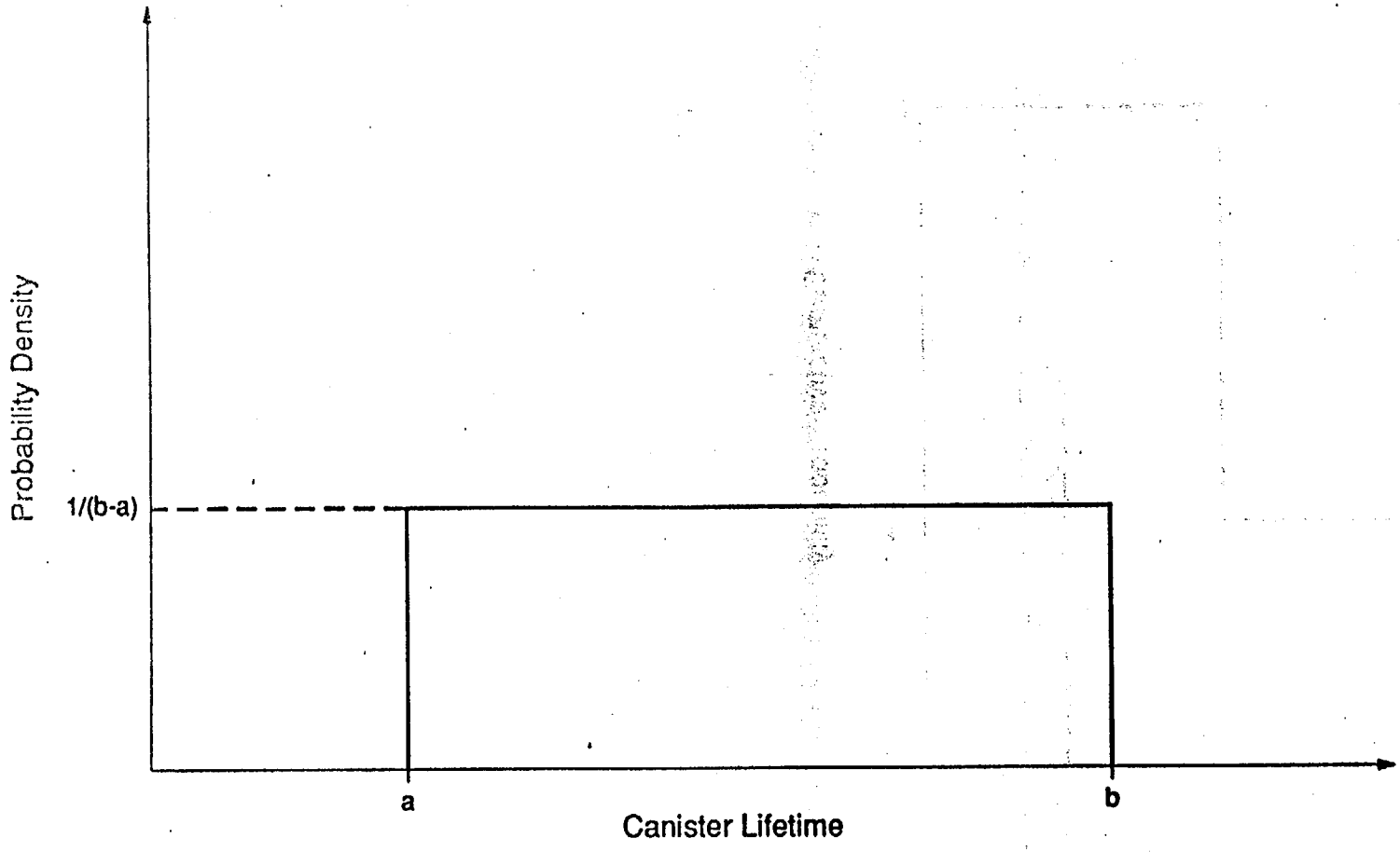


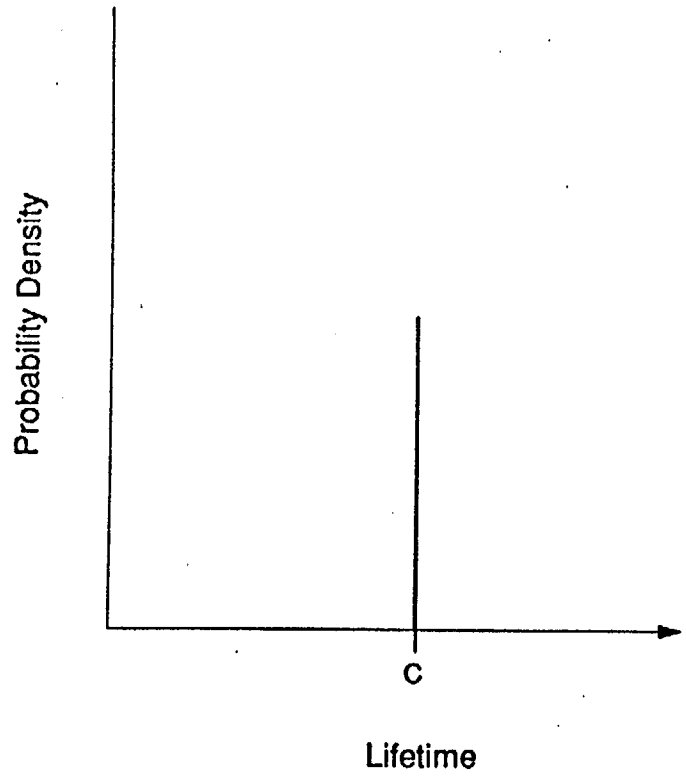
FIGURE **F-4**

UNIFORM PROBABILITY DENSITY FUNCTION

ARGONNE/3/W DEVELOPMENT/WA

PROJECT NO. 903-1371.203 DRAWING NO. 25506 DATE 6/19/91 DRAWN BY EA

Golder Associates



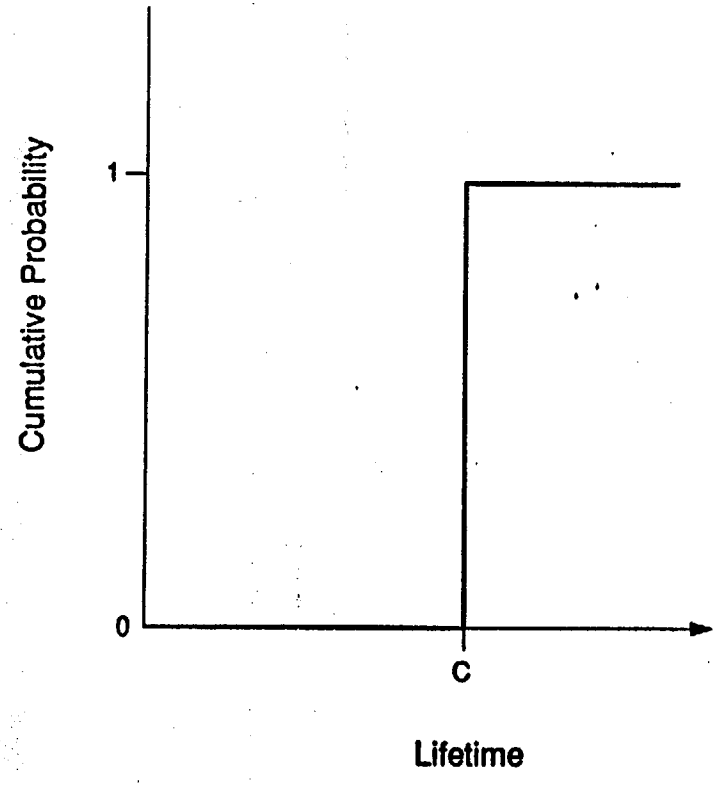


FIGURE **F-5**

DEGENERATE PROBABILITY DISTRIBUTION

ARGONNE/3/W DEVELO T/WA

PROJECT NO. 903-1371.203 DRAWING NO. 25507 DATE 6/19/91 DRAWN BY EA

Golder Associates

APPENDIX G

COMPUTING POISSON TRANSITION RATES FOR PATHWAY FLOW MODES

Poisson transition rates and pathway flow modes are discussed in Chapter 4 of the report. This appendix presents the solution method used to solve for the values of Poisson transition rates in a pathway. As discussed in Chapter 4, although each flow mode of a pathway is partially characterized by a Poisson transition rate, this rate is specified by the user for only one of the flow modes. The remaining transition rates are computed using the given transition rate and the flow proportions for the flow modes, f_{Qi}.

The solution for flow mode transition rates is derived by combining the Markov and Poisson processes. The fundamental property of the m-state Markov process is that the state probability distribution at time n+1 depends only on the state probability distribution at time n (Cox and Miller, 1965). In the RIP model, the m-states are m-flow modes of a pathway. The state probability distributions indicate the likelihood for a particle to be in a particular flow mode during transit along a pathway. The m-state Markov process may be written in terms of matrix and vector quantities as the following:

$$p^{(n+1)} = [P]\ p^{(n)} \tag{G-1}$$

where:

$p^{(n+1)}$ is the m-state probability distribution at time n+1,

[P] is the transition probability matrix (m by m), and

$p^{(n)}$ is the m-state probability distribution at time n.

The elements of the matrix [P] are transition probabilities, indicating the likelihood to transition from one state to another. Thus, element p_{ij} is the probability to go from state j to state i. The elements of the vector $p^{(n)}$ are the probabilities for a given nuclide to be in each of the misstates. Thus, vector element p_j is the probability to be in state j. Consequently, the product $p_{ij}p_j^{(n)}$ is the probability to go from state j to state i at time n+1.

For large values of n, the state probability distribution is argued to be asymptotic to an equilibrium probability distribution, thus, in Eq. 1, $p^{(n)} = p^{(n+1)}$ as n increases. For large n, Eq. 1 may be written as:

$$p = [P]\ p \tag{G-2}$$

where p is the equilibrium state probability distribution vector.

The Poisson probability distribution relates the likelihood for a transition to occur to a rate parameter (e.g. Hogg and Craig, 1978). Applied to particle transitions between flow modes of a pathway, the probability for 0 and 1 transition to occur in a length dl is given by the following:

$$P(N = 0, l, l+dl) = 1 - rdl - o(dl)$$
$$P(N = 1, l, l+dl) = rdl + o(dl) \qquad \text{(G-3)}$$

where:

N	is the number of transitions occurring in dl,
l,l+dl	is the length interval,
r	is the Poisson transition rate, and
o(dl)	is a function such that as dl-->0, o(dl)/dl --> 0, and o(dl) + o(dl) = o(dl).

The diagonal terms of the transition probability matrix [P] may be equated to the probability for 0 transitions (1-rdl-o(dl)) on an interval dl, as they are the transition probabilities to go from state i to state i, i.e. the probability not to transition from the current state.

In the matrix [P], the remaining off-diagonal transition probabilities are unknown, thus, the linear system in Equation G-2 has far more unknowns than knowns and cannot be used to obtain a meaningful solution for the Poisson transition rates. In order to solve for the Poisson transition rates, we make an assumption which results in the transition probabilities of [P] becoming functions of the transition rates, r_i. Our assumption is that the transition probabilities represent the intersection of two independent events. One event is that a particle leaves the current state, and has probability rdl + o(dl). The other event is that the particle enters a new state, conditional on not remaining in the current state. This latter event has probability $p_i/(1-p_j)$ for a particle leaving mode j and entering mode i. Recall the probability, p_i, is the proportion of flow occurring in state i, i.e. $f_{Qi} = p_i$. The off-diagonal transition probabilities of [P] are given by the product of the two event probabilities as follows:

$$p_{ij} = \frac{p_i}{1-p_j}\left[r_j dl + o(dl)\right] \qquad \text{(G-4)}$$

Equation 2 may now be solved for the Poisson transition rates as follows:

Step 1: Note that [P] = [I] + dl[T] where [I] is the m by m identity matrix and [T] is a new matrix with diagonal terms equal to $-r_i$ - o(dl)/dl and off diagonal terms equal to $p_i/(1-p_j)(r_j + o(dl)/dl)$; dl is a scalar;

Step 2: Substitute the results of step 1 into Equation G-2 to obtain the linear system [T]p = 0;

Step 3: Take the limit as dl --> 0 of [T] , which results in all terms o(dl)/dl -->0

Step 4: by inspection rewrite $[T]\mathbf{p} = \mathbf{0}$ such that $[T^*]\mathbf{r} = \mathbf{0}$ where $\mathbf{r}$ is the vector of transition rates and $[T^*]$ is a new matrix. The diagonal terms of $[T^*]$ are equal to $-p_j$ and the off-diagonal terms are equal to $p_i p_j/(1-p_j)$.

The linear system given by $[T^*]\mathbf{r} = \mathbf{0}$ has an infinite number of possible solutions (i.e. in reduced row echelon form, the last row of $[T^*]\mathbf{r} = \mathbf{0}$ contains all zeros). If one value of $\mathbf{r}$ is specified, say r_k, the linear system may be solved uniquely for the remaining r_i. This linear system may be written as $[A]\mathbf{x} = \mathbf{b}$ where the terms in $\mathbf{b}$ are equal to $-r_k p_i p_k/(1-p_k)$ and $[A]$ is an (m-1) by (m-1) matrix. The coefficients of $[A]$ are identical to $[T^*]$, although, the kth row and column from $[T^*]$ is not included in A. In the RIP model, the solution to this system is computed by performing LU decomposition of $[A]$, and then solving for $\mathbf{r}$ by backsubstitution.

APPENDIX H

ERROR ANALYSIS OF PATHWAY DISCRETIZATION OF RADIONUCLIDE BREAKTHROUGH

The use of radionuclide breakthrough curves in the RIP model is discussed in Chapter 4. In summary, the RIP model discretizes the breakthrough curve into pathway cells. This appendix presents results which compare the RIP model algorithm to an analytical solution of breakthrough based on a solution for the one-dimensional advection-dispersion equation. The analytical solution applies to uniform flow in an infinite column with a fixed concentration Co at $x = 0$, and $t > 0$ (Bear, 1979, page 268).

Four tests were run to evaluate the error between the analytical solution output and the RIP model output. The error was visually analyzed by plotting the cumulative mass releases for a variety of conditions based on flow rate and dispersivity. The release was observed at a length of 100 m from the source location and retardation was set to 1, i.e. no retardation. These results are shown on Figures H-1 through H-4. Table H-1 presents the flow rate and dispersivity parameters for each test.

TABLE H-1

CALCULATION PARAMETERS FOR RIP BREAKTHROUGH ERROR ANALYSIS

	Flow rate (m/yr)	
Dispersivity (m)	4×10^{-2}	1.33×10^{-1}
10 m	Test 1	Test 3
33 m	Test 2	Test 4

The test results compare with significant accuracy. The accuracy improves at higher flow velocities. The RIP model tends to overestimate the breakthrough under low flow velocities (e.g. Figure H-1), whereas, at higher flow velocities the results are closer (Figure H-6). Thus, in general the RIP model is conservative.

The RIP algorithm will generally overestimate mass release because it transfers mass in a quasi-parallel manner, rather than sequentially, i.e., mass entering a pathway at time t may be discharged with mass which entered at a previous time, $t-\Delta t$, where Δt can be as large as the pathway cell residence time.

While the RIP results have significant error, they are considered adequate considering the generally high level of uncertainty about flow system at Yucca Mountain.

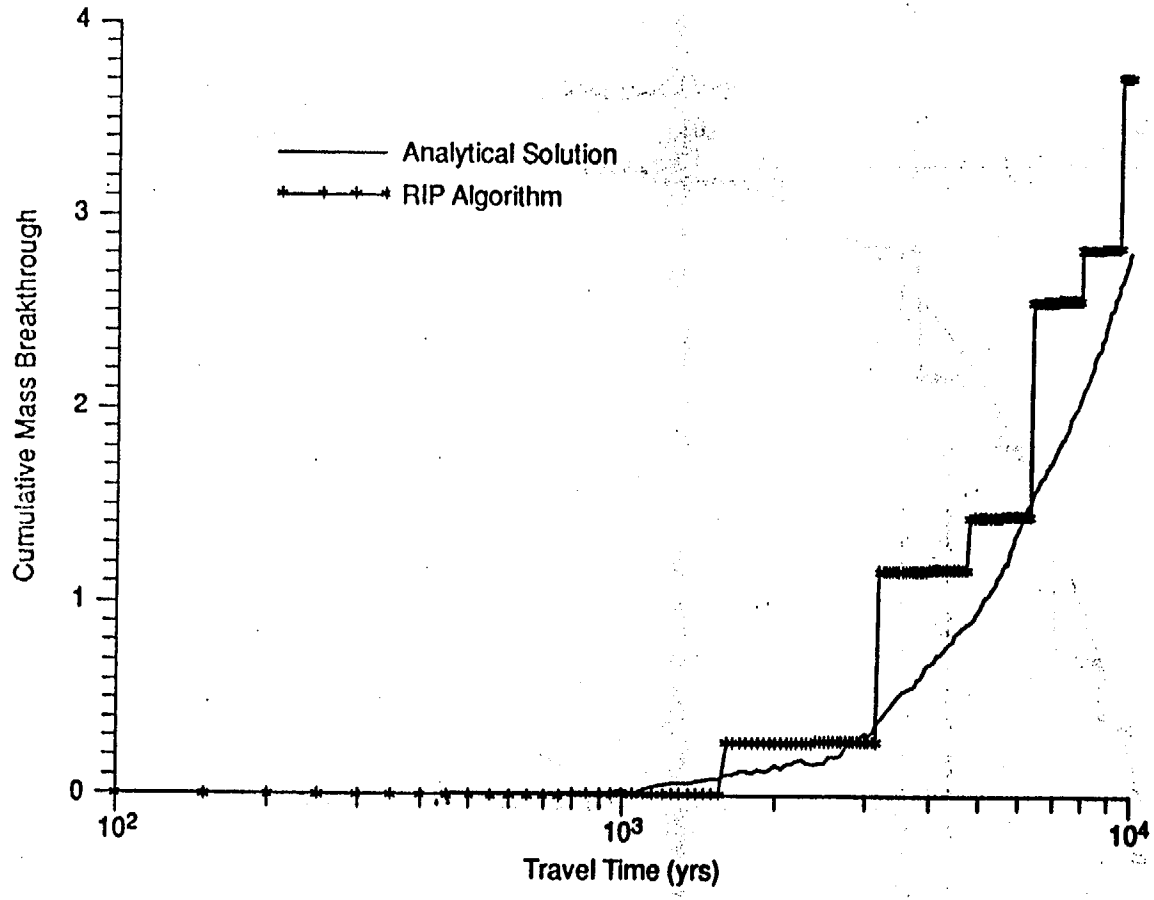


FIGURE **H-1**

TEST 1 - RIP BREAKTHROUGH ALGORITHM ERROR ANALYSIS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 26584 DATE 9/26/91 DRAWN BY CB

Golder Associates

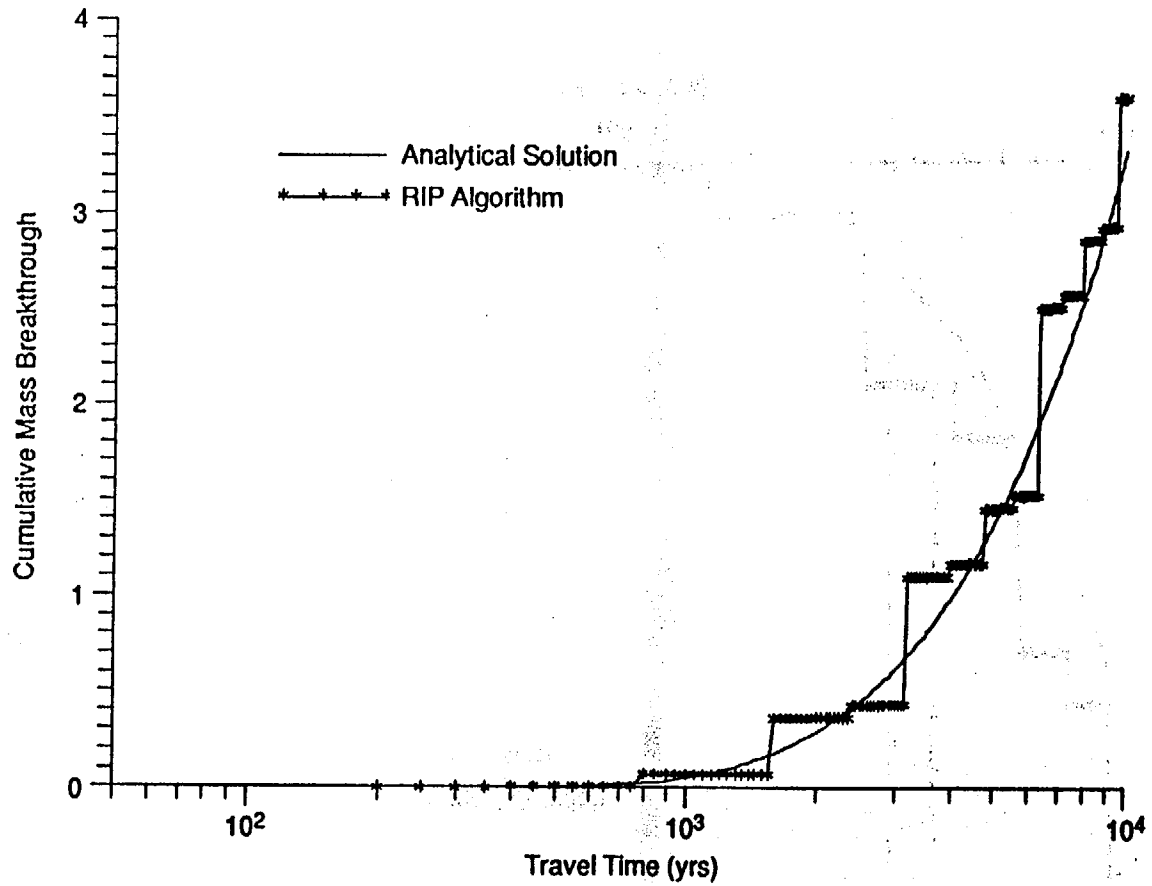


FIGURE **H-2**

TEST 2 - RIP BREAKTHROUGH ALGORITHM ERROR ANALYSIS

ARGONNE/MODEL D[illegible]PMENT

PROJECT NO. 903-1371.203 DRAWING NO. 26586 DATE 9/26/91 DRAWN BY CB

Golder Associates

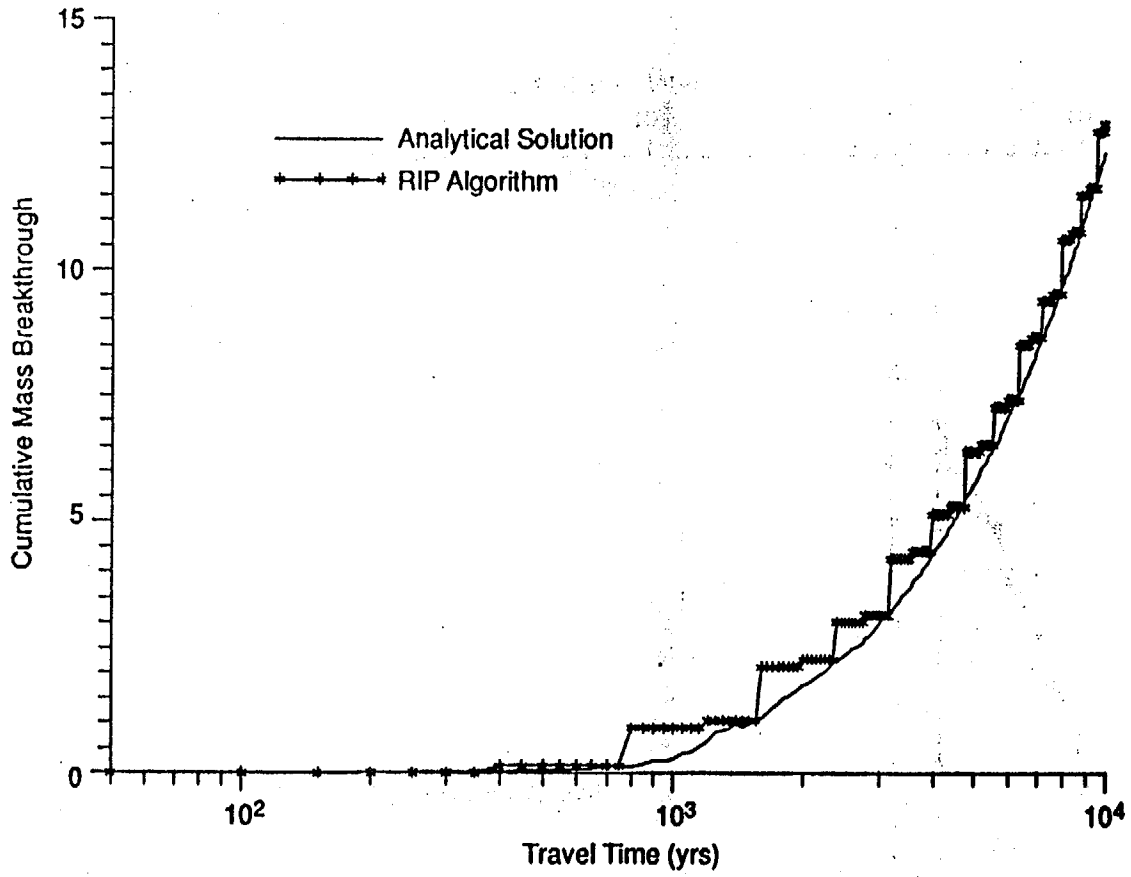


FIGURE **H-3**

TEST 3 - RIP BREAKTHROUGH ALGORITHM ERROR ANALYSIS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 26587 DATE 9/26/91 DRAWN BY CB

Golder Associates

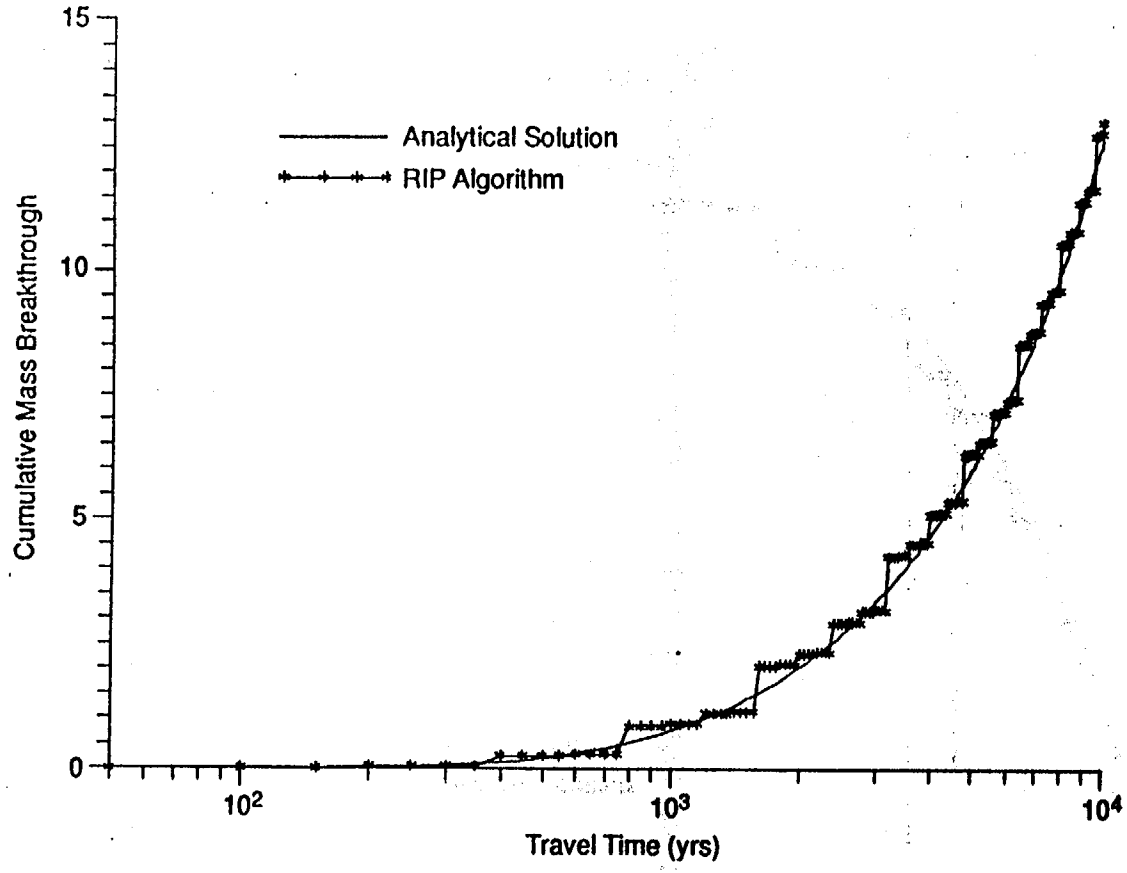


FIGURE **H-4**

TEST 4 - RIP BREAKTHROUGH ALGORITHM ERROR ANALYSIS

ARGONNE/MODEL DEV[illegible]MENT

PROJECT NO. 903-[illegible].203 DRAWING NO. 26588 DATE 9/26/91 DRAWN BY CB

Golder Associates

APPENDIX I

ERROR ANALYSIS OF APPROXIMATE MARKOV PROCESS ALGORITHM

The approximate Markov process algorithm is used to compute breakthrough curves for the RIP model. This algorithm is discussed in detail in Chapter 4. This appendix presents an error analysis conducted to compare the approximate Markov process algorithm to the true solution computed by random simulation. The random simulation results are based on particle tracking over the length of the pathway.

A total of 10 tests were run for a unit-length pathway to compare the approximate Markov process algorithm to the true solution obtained by random simulation. The tests were set up to be generally similar to the anticipated data sets for the RIP model. The tests were different from one another in their flow distributions and flow mode transition rates. In general, while holding the flow distribution constant, tests were run at low, intermediate, and high transition rates. The low transition rates were on the order of 10^{-2} m^{-1}. The intermediate transition rates were on the order of $10^{-1}m^{-1}$. The high transition rates were on the order of $10^{1}m^{-1}$. Three to four variations in the flow distributions were used in the tests.

The test results are shown on Figures I-1 through I-10. Tests 1 through 3 are examples of a two-flow mode pathway. One of the flow modes has a high velocity (1 m/yr) and a low proportion of the flow (0.1). The other mode has substantially lower velocity (10^{-4} m/yr) and the majority of the flow (0.9). This pathway may be similar to a pathway in which unsaturated groundwater flow occurs. The transition rates in tests 1, 2, and 3 were low, intermediate, and high, respectively. The results show that for low transition rates the approximation matches the true solution very well. As the transition rates increase, however, the approximation tends to overestimate the travel time. The largest overestimation is about 1.5 times greater than the true solution. As discussed in Chapter 4, the significance of this error depends on the amount of uncertainty in the flow mode parameters.

The next sequence of tests, 4, 5, and 6, are basically identical to tests 1, 2, and 3, except the flow distribution was changed from (0.1, 0.9) to (0.5, 0.5). The results for these tests are shown on Figures I-4, I-5, and I-6. The results follow a similar pattern to those of the first three tests. In test 6, however, the late travel times are underestimated by the approximation, rather than overestimated. The true solution shows greater dispersion than the approximation. Again, the margin of error is relatively small.

In tests 7, 8, and 9, a two-flow mode system is again used for the comparison. The velocities have not been changed, however, the high velocity flow mode receives the majority of the flow (0.9). This flow distribution may be representative of a saturated groundwater system in fractured rock. The tests include results for low, intermediate, and high transition rates, which are shown on Figures I-7 through I-9. The results show that under low transition rates little error occurs. At greater transition rates, however, the error increases and the approximate method tends to underestimate travel time at the intermediate transition rates. At the high transition rate, the approximate solution compares well to the true solution, although it contains less dispersion in the travel time distribution causing error to occur for the later travel times. Again, this error is within a factor of about 1.5 of the true solution.

Test 10 includes 5 flow modes with a flow velocity variation on the order of 10^4. The results of test 10 are shown on Figure I-10. The flow is distributed to the flow modes in approximately equal portions. The transition rates are at the intermediate level relative to the previous tests. The test results show that in the region of intermediate travel times, i.e. 10 to 1500 years, the approximation overestimates the true solution. After about 1500 years, the approximation underestimates the true solution. In general, a close match is only achieved for the very early and late travel times over the pathway.

APPENDIX J

COMPUTING A RETARDATION PARAMETER FOR MULTIPLE RETARDATION MECHANISMS

This appendix presents the calculation method used to combine retardation parameters for use in the RIP algorithm. Retardation parameters are used to adjust flow mode travel times for individual radionuclides. Chapter 4 of the report discusses pathway flow modes and radionuclide retardation.

The retardation parameter based on the advection-dispersion equation for solute transport in porous media is given by the following:

$$R = 1 + \frac{M_s}{M_d} \tag{J-1}$$

where:

R is the total retardation parameter,

M_s is the total sorbed mass, and

M_d is the total dissolved mass.

For more than one retardation mechanism, the quantity of sorbed mass for the ith mechanism is given by the following:

$$M_{s_i} = M_d (R_i - 1) \tag{J-2}$$

where:

M_{si} is the sorbed mass resulting from the ith retardation mechanism,

R_i is the retardation parameter for the ith mechanism, and

M_d is the total dissolved mass.

The total sorbed mass is the sum of all sorbed mass given by Eq. J-2, and is given by the following:

$$M_s = M_d \sum (R_i - 1) \tag{J-3}$$

where the parameters are as defined in Eq. J-1 and Eq. J-2. The total retardation parameter representing all retardation mechanisms is obtained by substitution of J-3 into J-1 and is given by:

$$R = 1 + \sum (R_i - 1) \tag{J-4}$$

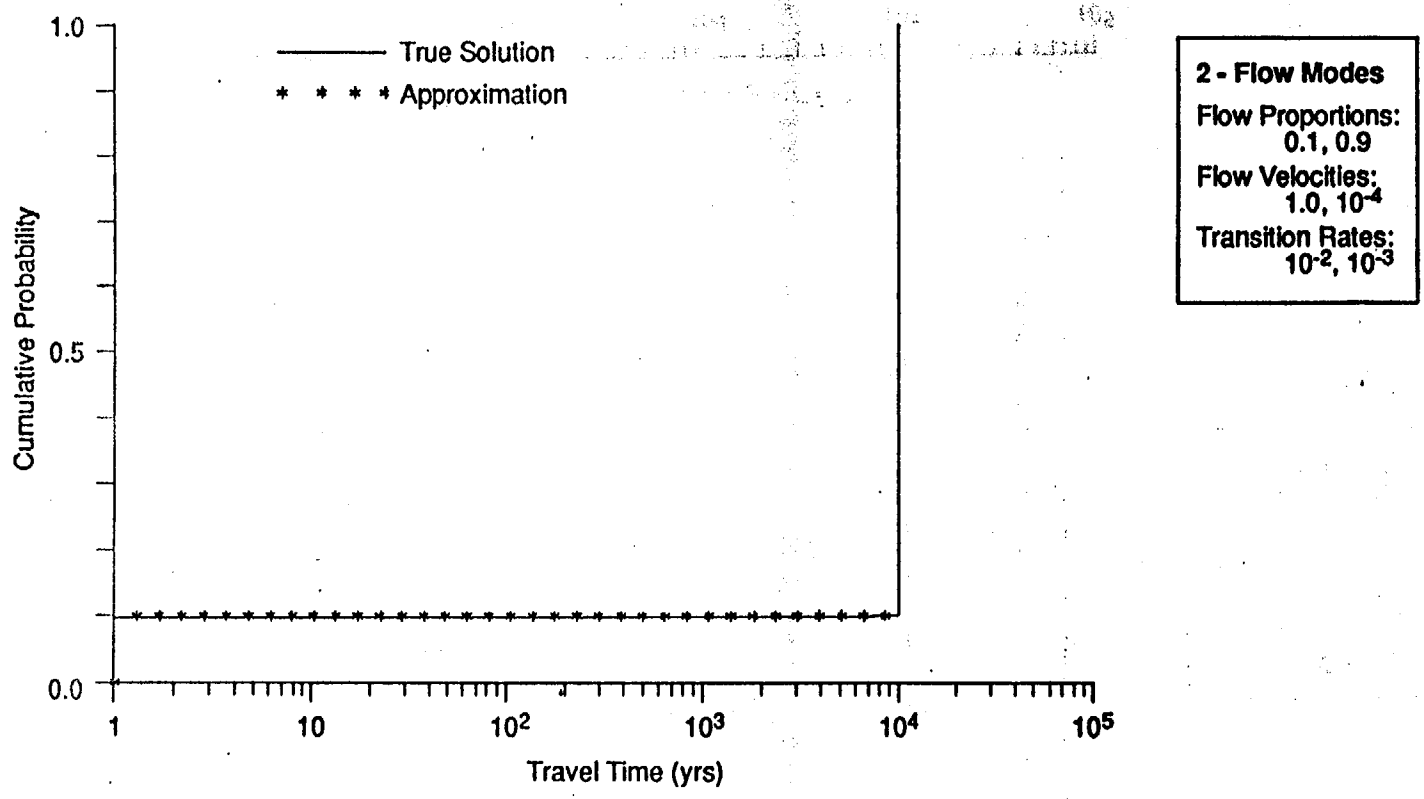


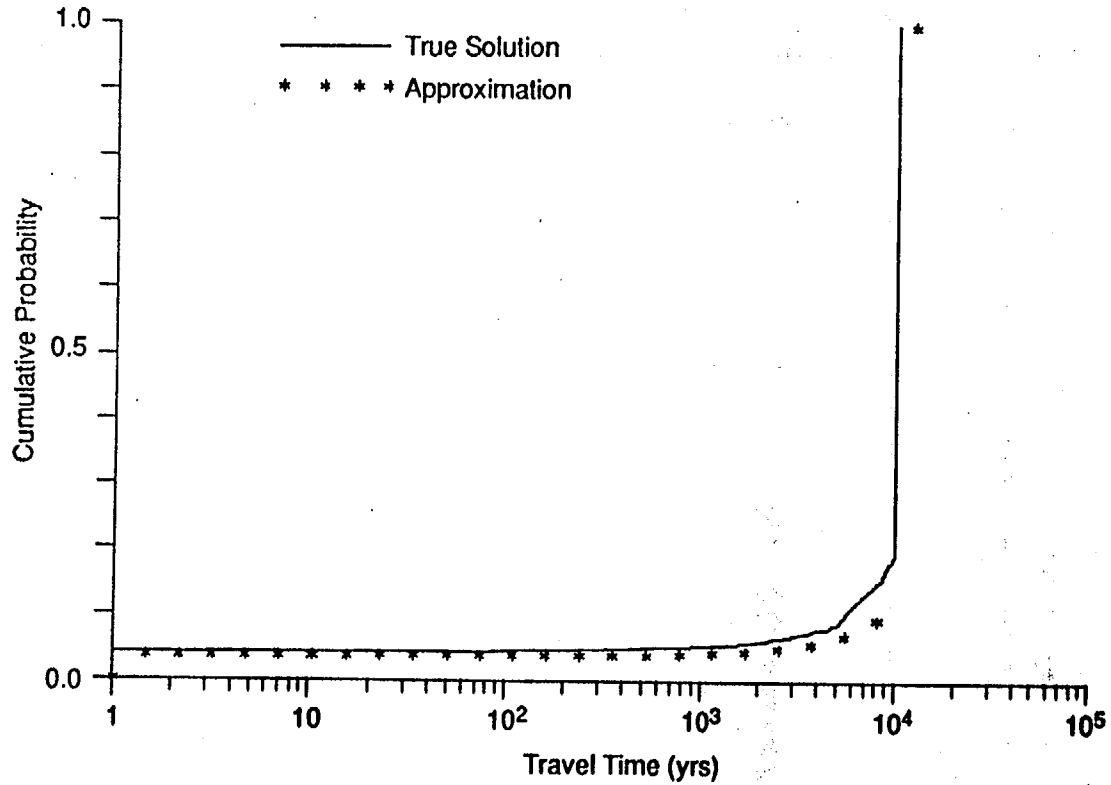
FIGURE **I-1**

TEST 1 - APPROXIMATE MARKOV PROCESS ALGORITHM ERROR ANALYSIS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 26589 DATE 9/27/91 DRAWN BY CB

Golder Associates



2 - Flow Modes

Flow Proportions: 0.1, 0.9

Flow Velocities: 1.0, 10^{-4}

Transition Rates: 1.0, 10^{-1}

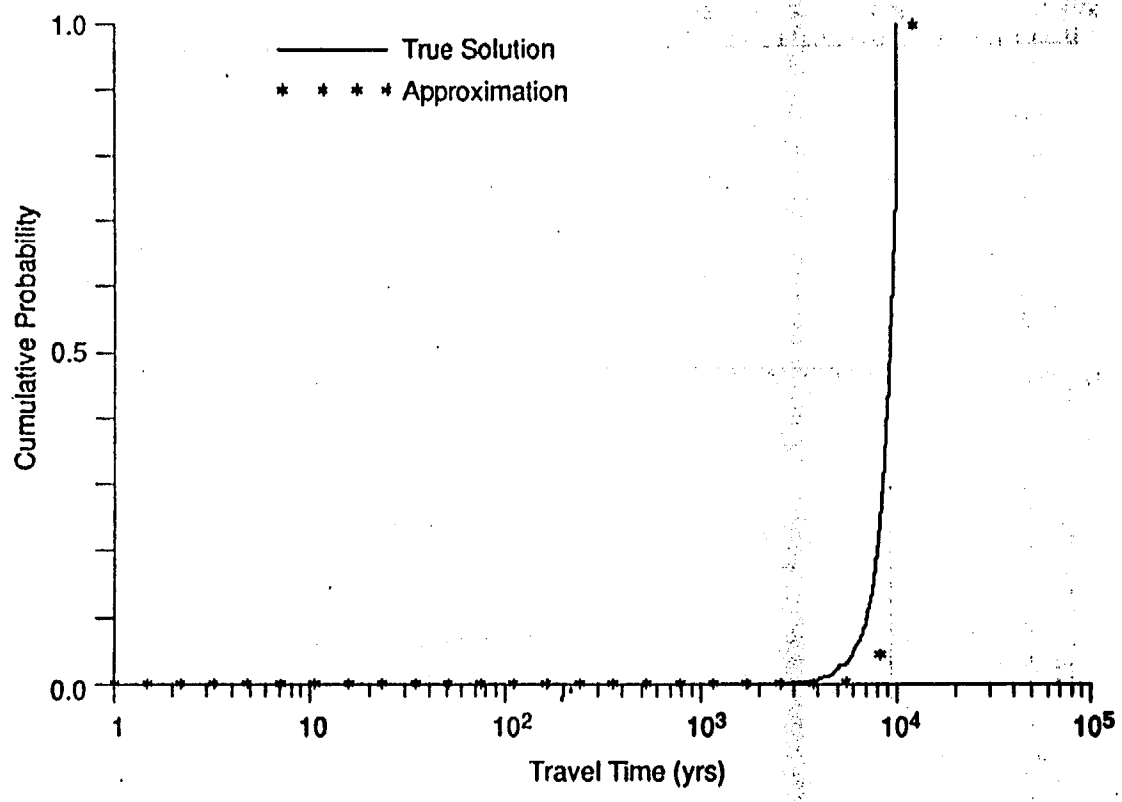
FIGURE **I-2**

TEST 2 - APPROXIMATE MARKOV PROCESS ALGORITHM ERROR ANALYSIS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 26590 DATE 9/27/91 DRAWN BY CB

Golder Associates



2 - Flow Modes
Flow Proportions: 0.1, 0.9
Flow Velocities: 1.0, 10^{-4}
Transition Rates: 10.0, 1

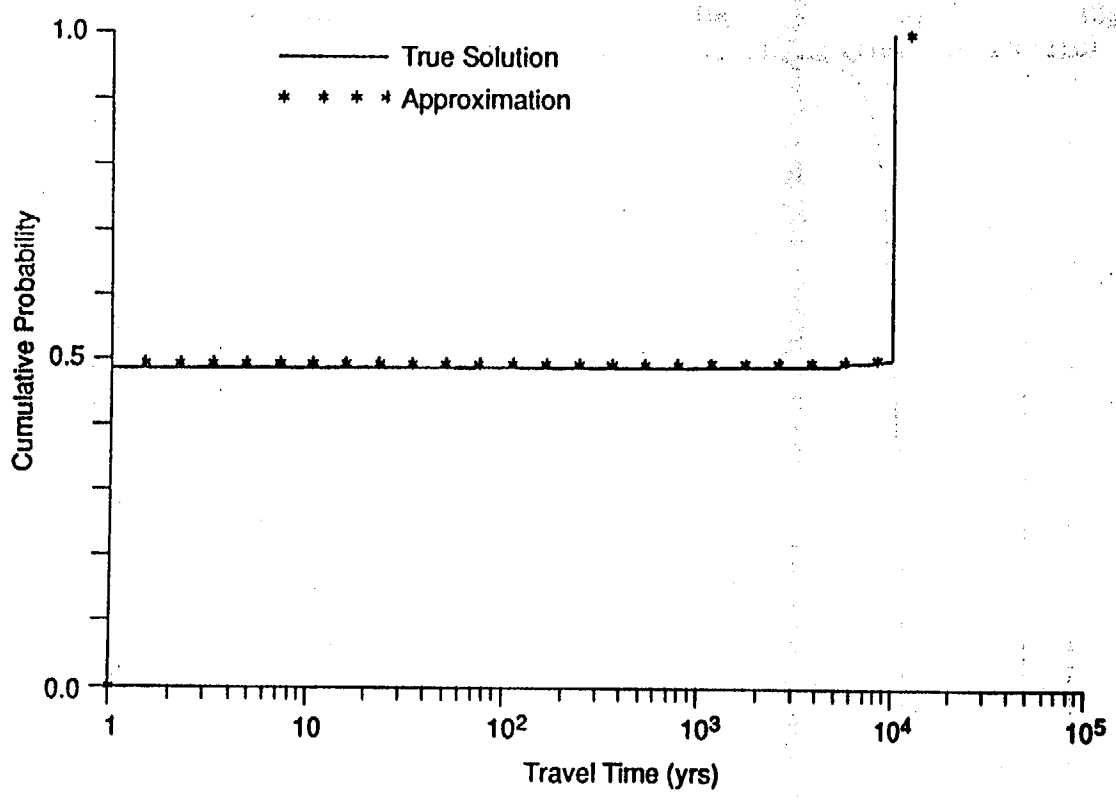
FIGURE **I-3**

TEST 3 - APPROXIMATE MARKOV PROCESS ALGORITHM ERROR ANALYSIS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 26591 DATE 9/27/91 DRAWN BY CB

Golder Associates



2 - Flow Modes

Flow Proportions: 0.5, 0.5

Flow Velocities: 1.0, 10^{-4}

Transition Rates: 10^{-2}, 10^{-2}

FIGURE **I-4**

TEST 4 - APPROXIMATE MARKOV PROCESS ALGORITHM ERROR ANALYSIS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 26592 DATE 9/27/91 DRAWN BY CB

Golder Associates

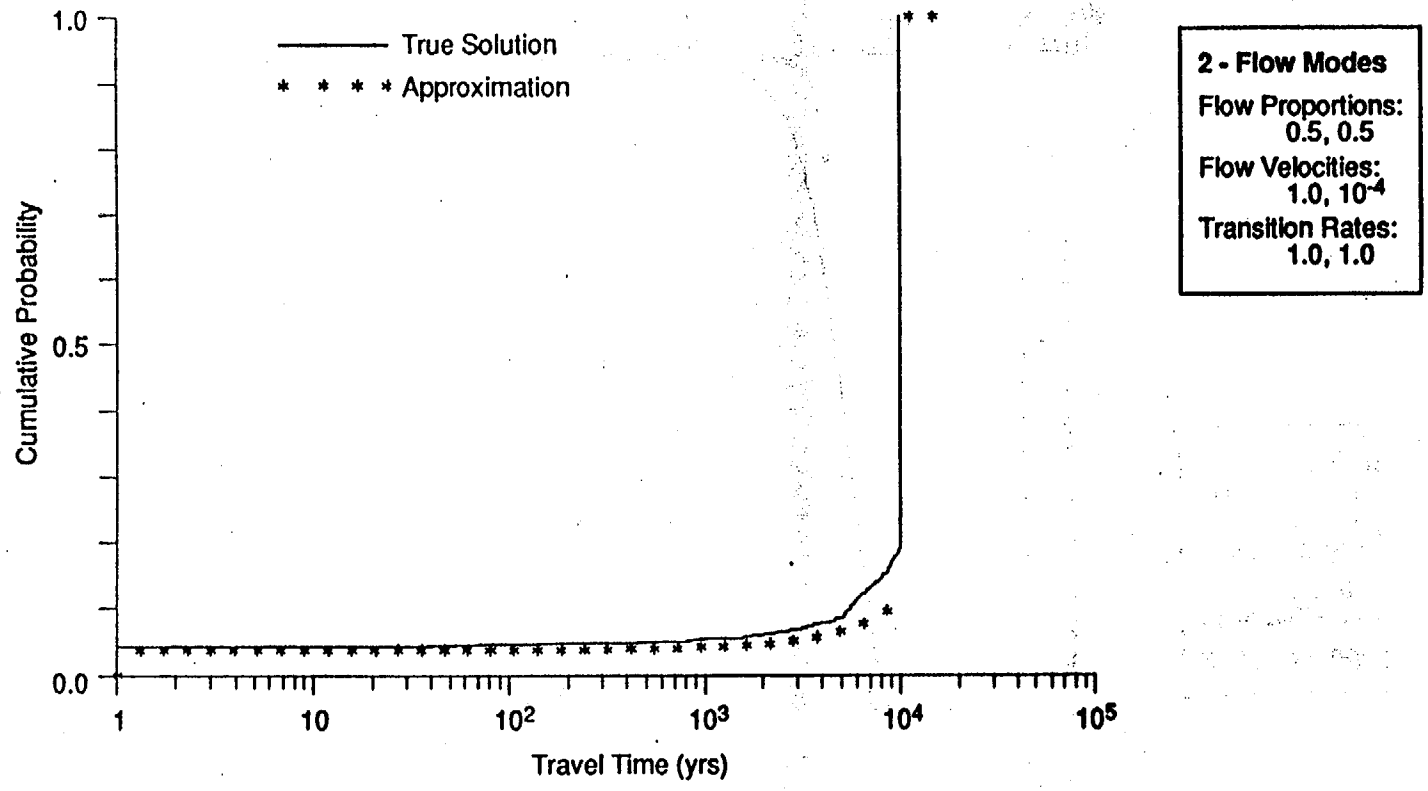


FIGURE **I-5**

TEST 5 - APPROXIMATE MARKOV PROCESS ALGORITHM ERROR ANALYSIS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 26593 DATE 9/27/91 DRAWN BY CB

Golder Associates

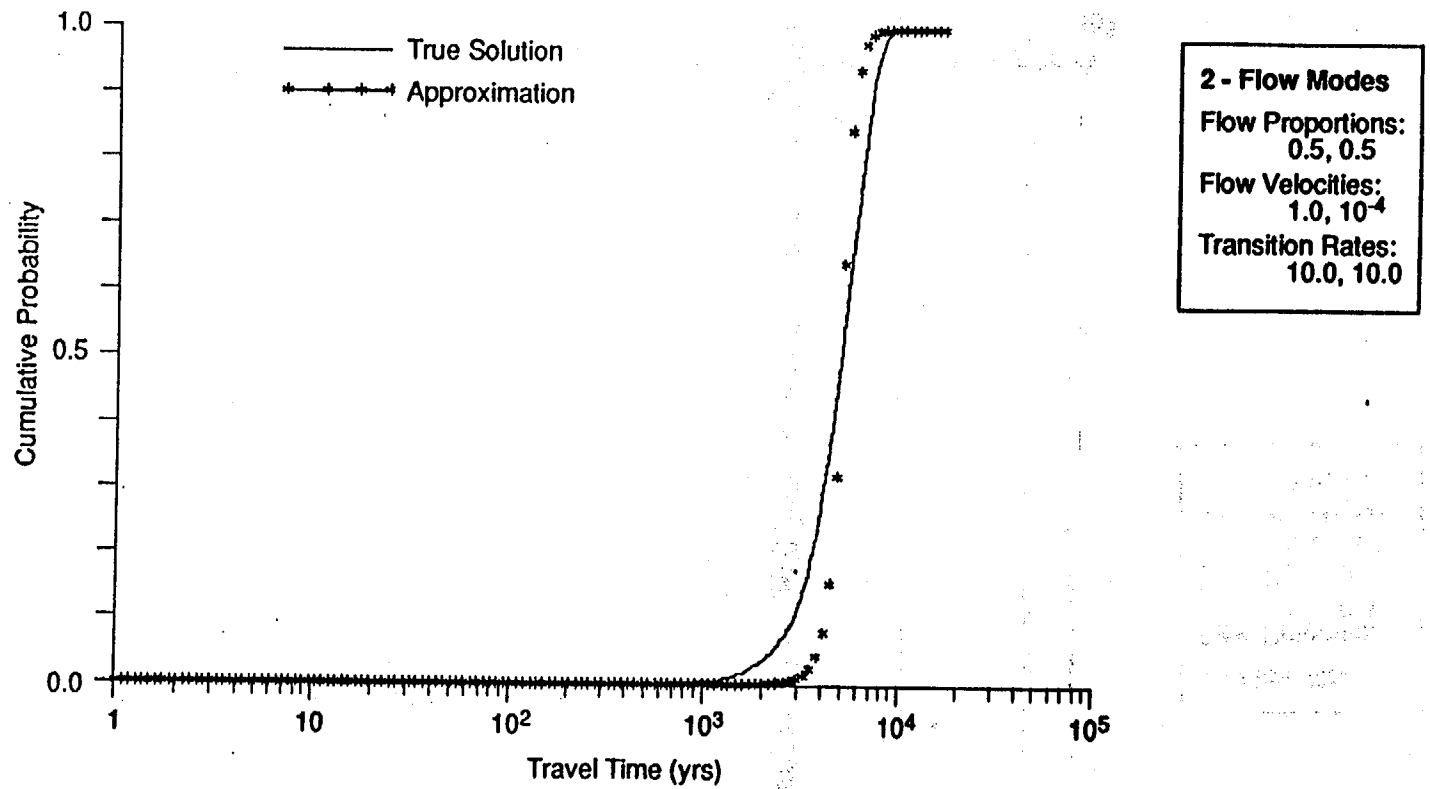


FIGURE I-6
TEST 6 - APPROXIMATE MARKOV PROCESS ALGORITHM ERROR ANALYSIS
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1203 DRAWING NO. 26594 DATE 9/27/91 DRAWN BY CB

Golder Associates

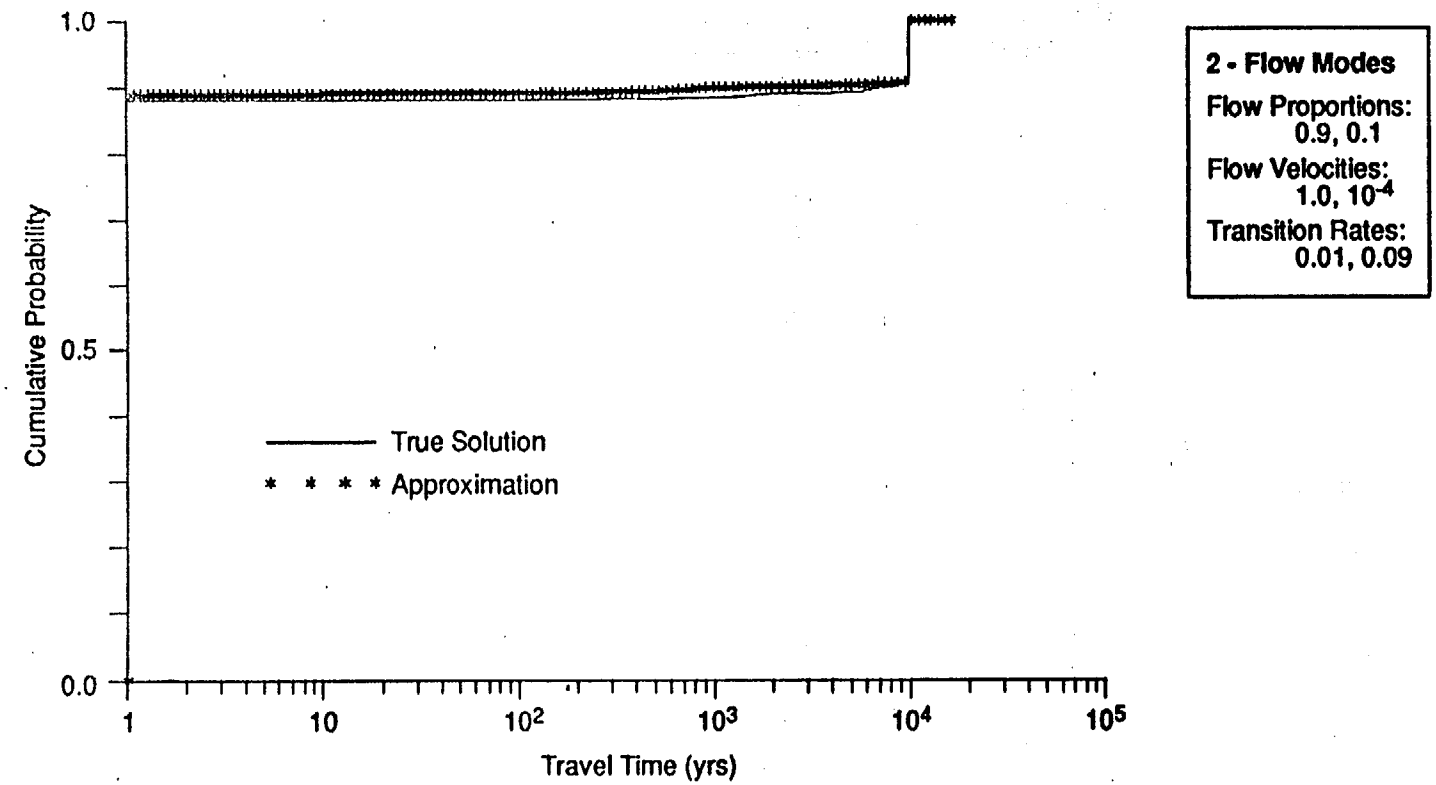


FIGURE **I-7**

TEST 7 - APPROXIMATE MARKOV PROCESS ALGORITHM ERROR ANALYSIS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 28595 DATE 9/27/91 DRAWN BY CB

Golder Associates

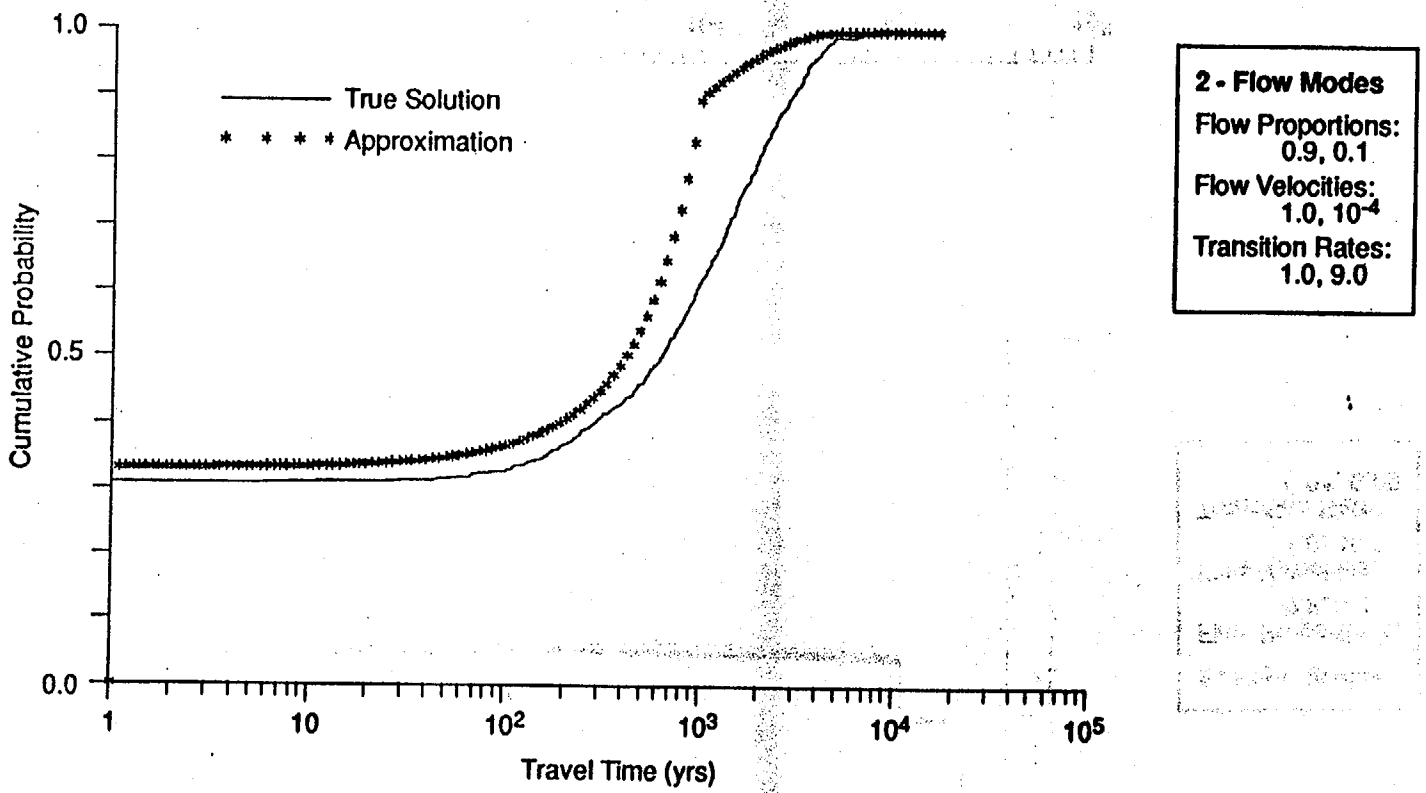


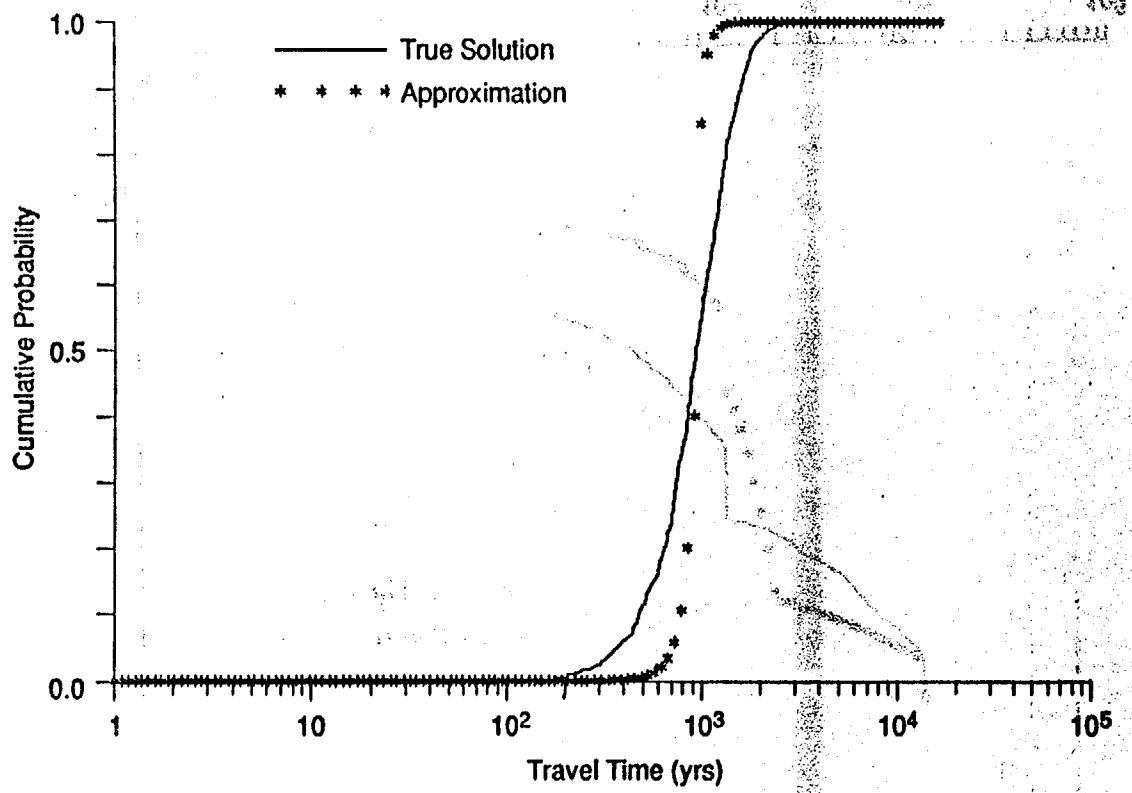
FIGURE I-8

TEST 8 - APPROXIMATE MARKOV PROCESS ALGORITHM ERROR ANALYSIS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 9[illegible]371.203 DRAWING NO. 26596 DATE 9/27/91 DRAWN BY CB

Golder Associates

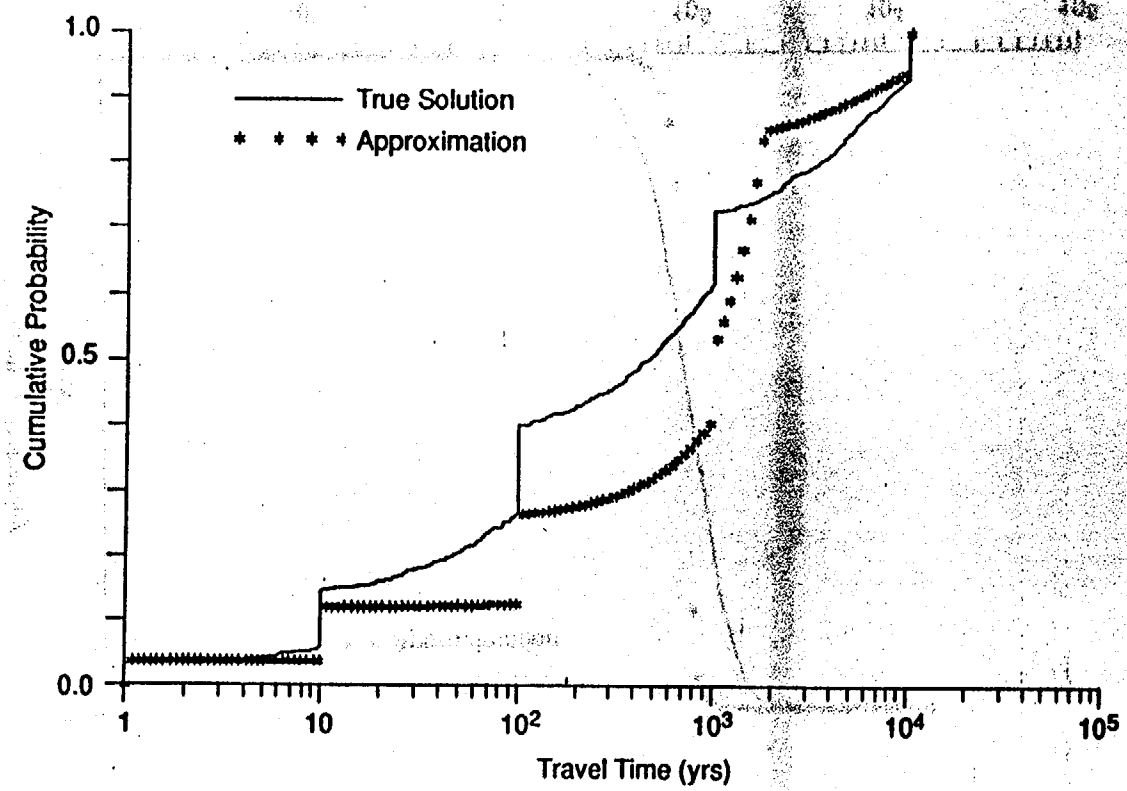


2 - Flow Modes
Flow Proportions: 0.9, 0.1
Flow Velocities: 1.0, 10^{-4}
Transition Rates: 10.0, 90.0

FIGURE **I-9**
TEST 9 - APPROXIMATE MARKOV PROCESS ALGORITHM ERROR ANALYSIS
ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 26603 DATE 9/27/91 DRAWN BY CB

Golder Associates



5 - Flow Modes

Flow Proportions:
0.1,0.2,0.3,0.25,0.1

Flow Velocities:
$1.0, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$

Transition Rates:
1.0,0.89,0.78,0.83,0.94

FIGURE **I-10**

TEST 10 - APPROXIMATE MARKOV PROCESS ALGORITHM ERROR ANALYSIS

ARGONNE/MODEL DEVELOPMENT

PROJECT NO. 903-1371.203 DRAWING NO. 26597 DATE 9/27/91 DRAWN BY CB

Golder Associates