I-50609

NNA.891204.0186

A VALIDATION METHODOLOGY FOR PERFORMANCE-ASSESSMENT MODELS

Prepared by the Validation Oversight Group

for the U.S. Department of Energy

Office of Civilian Radioactive Waste Management

Draft Version 0.0
October 1989

SUMMARY

This report contains a proposed methodology for validating performance-assessment models. The methodology was developed by the Validation Oversight Group for the Department of Energy's Office of Civilian Radioactive Waste Management. The early development and implementation of a model validation methodology is important because performance assessment models will support the license application to construct, operate, and close a mined geologic repository for high-level nuclear waste. Thus, the DOE must prepare technical justification to document and support the evidence provided by the models.

In addition to describing the methodology, the report contains an analysis of the validation process. This was done to clarify the objectives of the activities in the methodology and explain why they are important.

The validation methodology consists of activities that have been broadly divided into three components: 1) a record of model development; 2) a description of the laboratory and field investigations and the resulting data supporting the development of the model; and 3) technical reviews. The methodology attempts to establish that a model is appropriate and adequate for the problem being addressed, that it was logically developed using the best available technology, that it can be supported by experimental and observational data, that the quality of the data is high, and that the limitations of the model are well understood.

A technical review committee will decide when the information provided in the model development record is adequate for demonstrating the validity of the model. Noted deficiencies will be resolved by the modeler(s) and experimenters in this iterative process. The validation decision is conditional in that the review committee will periodically review the model during the construction and operation of the repository as new data become available from performance confirmation activities and other sources.

Additional work is needed to 1) develop qualitative and quantitative criteria for deciding when a model is valid and 2) evaluate the likelihood the criteria will be met. The proposed methodology addresses qualitative technical criteria, e.g., providing the technical basis for the hypotheses included in the model. The quantitative criteria, such as the acceptable variance between a model's predicted behavior and that measured in an experiment or natural analog, have not been addressed. Furthermore, the issue of how to account for the uncertainty in natural systems that cannot be quantified must be dealt with. While the specific criteria to be used are the responsibility of each technical review committee, a reasonable set should be developed at this time to verify that the expectations being placed on the models are achievable.

TABLE OF CONTENTS

A VALIDATION METHODOLOGY FOR PERFORMANCE-ASSESSMENT MODELS

Draft Version 0.0
October 1989

1.0 INTRODUCTION

The Nuclear Regulatory Commission (NRC) and the Environmental Protection Agency (EPA) have issued federal regulations stipulating the performance requirements for the permanent storage of high-level nuclear waste in a mined geologic repository. In order to demonstrate compliance with these requirements, the Department of Energy (DOE) will need to present evidence from several sources including laboratory and field experiments, observation of natural analog systems, and performance assessment analyses. The performance assessment analyses will be based on the construction of numerical models designed to simulate the state and response over time of the physical repository system, its component systems, and its environment. These performance-assessment models and evidence will be used to support a license application to construct, operate, and close the nuclear waste repository.

The evidence provided in the license application is likely to be subject to legal challenge regarding the correctness of the data and its interpretation. Thus, in addition to collecting and interpreting data, the DOE must also prepare technical justification to document and support the evidence provided.

Current NRC/EPA regulations specify the license approval decisions will be made and the standards that are required for the evidence presented. The only statement in 10 CFR Part 60 that appears to directly address the issue of validating performance-assessment models is contained in Chapter 21 (c) (ii):

> "Analyses and models that will be used to predict future conditions and changes in the geologic setting shall be supported by using an appropriate combination of such methods as field tests, in situ tests, laboratory tests which are representative of field conditions, monitoring data, and natural analog studies."

Some additional information vaguely related to the validation of performance-assessment models is provided in the 10 CFR Part 60 Proposed Rule Making for Conforming Amendments:

> "The applicant will be required to submit a systematic and thorough analysis of potential releases and the Commission will issue a license only if it finds a substantial, though unquantified, level of confidence

that compliance with release limits will be achieved. As we have stated previously (48 FR 28201), in order to make a finding with 'reasonable assurance,' the performance assessment which has been performed in the course of licensing review must indicate that the likelihood of exceeding the EPA standard is low and, further, the Commission must be satisfied that the performance assessment is sufficiently conservative, and its limitations are sufficiently well understood, that the actual performance of the geologic repository will be within predicted limits."

In the absence of specific model validation requirements in the regulations and in consideration of the importance performance assessment data will have in the licensing process, DOE management directed that a draft model validation methodology be developed for use within the Civilian Radioactive Waste Management (CRWM) Program.

A Validation Oversight Group (VOG) made up of modelers, experimenters, and quality assurance personnel from within DOE and its contractors was established to prepare the methodology. This paper discusses the approach developed by the VOG. The paper is being widely distributed to obtain a thorough review by the scientific community and the content of the methodology may change based on the comments received.

The validation methodology for performance-assessment models consists of activities that have been broadly divided into three components: 1) a record of model development; 2) a description of the laboratory and field investigations and the resulting data supporting the development of the model; and 3) technical reviews (see Figure 1). Stated briefly, the methodology attempts to establish that the model is appropriate and adequate for the problem being addressed, that it was logically developed using the best available technology, that it can be supported by experimental and observational data, that the quality of the data is high, and that the limitations of the model are well understood. Any deficiencies identified in these areas are to be documented in subsequent technical reviews and corrected by the modelers and/or experimenters.

The high level of detail implicit in the model validation methodology recognizes the exceptional circumstances under which many of the performance assessment models are being used. Traditionally, model validation has entailed comparing predictions made by the model with experimental results over times and distances similar to those for which the model eventually will be used. Clearly, this level of validation is not possible over the long time periods (up to 10,000-years) over which regulatory performance needs to be evaluated. Because of this, the model validation methodology is focused on those aspects of modeling that are especially important for demonstrating compliance with performance requirements and building confidence in the efficacy of the model.

A number of pre- and postclosure models are being used to support the siting, design, and operation of the proposed geologic repository at Yucca Mountain and to assess the probable performance of the repository system during the next 10,000 years. Because of the diversity of models required, the model validation methodology must be both flexible and comprehensive. The VOG attempted to avoid making the methodology unnecessarily complex or overly prescriptive. The methodology is essentially an attempt to document the scientific method. However, the use of the output from the models in the license application requires that a higher level review be made of the model development process receive a higher level review and that the various stages of the model-development process be carefully documented.

To a large extent, the emphasis of the validation methodology is on documenting the validation process rather than on the specific activities involved. While a large number of activities are considered in the methodology (e.g. hypothesis testing, sensitivity analyses, field experiments, etc.). To maintain the requisite flexibility, however, the methodology does not prescribe that they all be used. Ultimately, the modelers have the responsibility for deciding the mix of activities that will be used to support the validation of the performance-assessment models. A technical review committee will decide whether these activities and the accompanying documentation are adequate for demonstrating the validity of a model. All deficiencies identified by the review committee must be addressed by the modelers and experimenters, and they must provide additional information for resolving the deficiencies cited by the review committee.

In the proposed validation methodology, a model can be submitted for formal technical review at any stage of its development; it is not necessary that the model be implemented in a fully operational computer code. The benefit of early review is that a preliminary assessment of the conceptual model and the strategy for addressing the problem being considered can be obtained before investing the large amounts of effort necessary to develop a numerical model.

A majority of the documentation process included in the methodology is already in place within the CRWM Program and the Yucca Mountain Project (YMP). Procedures for developing models, performing experiments, and conducting peer reviews have been established as part of the quality assurance programs within these organizations. Modification of the procedures may require some slight modifications before they can be used for their use in the model validation methodology.

The following sections present general guidelines for developing a validation report for a performance-assessment model. Section 2.0 describes the construction of the model development section of the validation report and the types of information that should be provided. Section 3.0 addresses the use of data in the validation process and the specific aspects that should be considered in developing the validation report. Section 4.0 describes the general nature of the formal technical review (FTR) process, where the model development and supporting data (experimental and observational) are considered and the adequacy of the model for addressing the problem under consideration is judged. The three components are iterative and a certain amount of overlap is necessary, particularly between the model development and the laboratory and field investigation activities.

The appendix contains definitions of the terms used throughout the methodology sections. These terms have been used in many ways in the technical and regulatory literature. To avoid confusion and misunderstanding, the reader is encouraged to review these terms before proceeding.

Draft - Version 0.0

VALIDATION METHODOLOGY FOR PERFORMANCE-ASSESSMENT MODELS

2.0 DOCUMENTATION OF MODEL DEVELOPMENT

The objectives of the activities in this component are to document 1) the process used to develop a performance-assessment model and 2) the technical review committee's appraisal of the model's suitability (validity) for its intended use. While only the 2.0 activities are discussed in this section, the sequence of activities during model development and testing is likely to be more similar to that shown in Figure 1.

The guidelines for computer code documentation provided by Silling (1983) should be used to accomplish much of the first objective. However, in order to address the specific needs of model validation, the documentation should also identify the assumptions and hypotheses that constitute the conceptual model. Furthermore, the scientific basis or justification for incorporating each of the assumptions in the model should be provided, along with interpretations of laboratory and field investigations related to the modeling effort.

This record of model development provides the basis for one or more formal reviews by a technical committee (discussed in Section 4.0). The review committee assesses the appropriateness and suitability of the model for answering the question being addressed and the validity of the model and its various components. The record documents the review committee's appraisal(s) and the actions planned or taken to respond to any inadequacies identified by the committee.

The record of model development should also contain reference to the internal reviews of the performance assessment model that were performed within the CRWM program. Existing quality assurance procedures for model development, peer review, etc., form the basis for much of the validation methodology.

2.1 STATE PROBLEM BEING ADDRESSED AND THE REGULATORY BASIS

The Site Characterization Plan (SCP) (DOE, 1988) contains an issues hierarchy that states the questions the DOE feels must be resolved about the performance of the mined geologic disposal system (i.e., the waste package, the engineered repository, and the natural system at the site) to demonstrate compliance with the applicable federal regulations. These issues are broadly divided into performance issues and design issues. The performance issues generally address questions about compliance with regulatory requirements for the performance of the disposal system and comprise the majority of the problems being addressed in performance assessment.

The documentation of model development must include a description of the problem being addressed. This step recognizes that performance-assessment models are problem-specific and application-dependent. A model developed specifically to address whether a particular regulatory requirement is met is not likely to be adequate or appropriate for addressing another problem without significant modification.

2.2 DESCRIBE THE CONCEPTUAL MODEL OF THE SYSTEM

A performance-assessment model can be described by its four parts: model geometry; initial and boundary conditions; physicochemical processes; and input parameters. The complexity of most physical systems requires that the system description be described simply enough that only the most important features are considered explicitly. The "most important" features are those that are expected to control or influence the intended application of the model.

The system is simplified by idealizing these model components through appropriate simplifying assumptions and approximations. The most critical assumptions must be justified by supporting information, such as experimental results, field observations, etc. For those assumptions that cannot be technically defended because of insufficient data, laboratory or field investigations should be proposed and a schedule for carrying them out submitted.

The objective of this activity is to describe the conceptual model(s) and the underlying assumptions, provide the technical basis for the assumptions, and assess their relative importance in influencing the behavior (results) of the model. The activity comprises four subactivities, shown in Figure 2 and described in detail below.

2.2.1 Identify the Assumptions, Hypotheses and Limitations(a)

System idealizations (i.e., geometry, initial and boundary conditions, processes, and parameters) usually can be described by a set of hypotheses that constitute the conceptual model for the system and its interaction with the external environment. The hypotheses composing the conceptual model are not valid a priori and must be tested and evaluated continually as part of the overall model-validation process. That a hypothesis can be shown to be invalid for some or all applications of a model gives rise to the notion of alternative conceptual models. These models consist of alternative or additional hypotheses whose inclusion within the model may result in a more correct or accurate representation of the system. Consequently, the

(a) Hypotheses and assumptions are considered synonymous for the rest of this report.

evolution of a conceptual model for a system is an ongoing iterative process during model development, validation, and application.

In general, the hypotheses constituting a conceptual model are selected to be as simple as possible and to be compatible with the known or anticipated features of the system (that is, for the known properties, processes, and structure of the system). The hypotheses also must be consistent; that is, they must not lead to any contradictions within the model. Furthermore, the hypotheses must be complete to the extent that they account not only for all currently known major features of the system, but also for subsequent rapid, dynamic response or for slow time-evolution of the system produced by changing external environmental conditions (Russell, 1948, p. 311).

For the four parts of the conceptual model and for the mathematical model developed for that conceptual model, the hypotheses and assumptions that are relevant to each part must be provided.[a] These include the following information as relevant for the particular model:

- model geometry, including the dimensionality, coordinate system, and special features (e.g., a hydrologic model might include a number of hydrogeologic units intersected by faults, fracture zones, etc.)
- process(es) (e.g., groundwater flow, diffusion, dissolution) being modeled; for coupled processes, the type of coupling
- time-variability of each process
- boundary and initial conditions
- mathematical formulation of the problem (equations and terms used)
- mathematical solution techniques to be employed (e.g., numerical, analytical, algebraic)
- the model's input, output and internal model property and parameter data, including their spatial and temporal variability, distributions and dependencies, etc.
- dependency of model on the input from other models and the need for this model's output by other models.

(a) The conceptual model for the system can be made up of any number of conceptual models for the individual processes, model geometry, conditions, etc.

2.2.2 Provide Technical Basis for Each Assumption[a]

The technical basis must be provided for each assumption used to describe the conceptual model. This information is included to simplify the task of evaluating the sources of uncertainty in the model and to identify any additional information or data that is needed. The technical basis for an assumption may include the following:

- nature and importance of problem or issue to be resolved (e.g., assumption is inherent in the problem being addressed)
- level of understanding of process(es) to be modeled
- supporting data from laboratory or field investigations
- results of a sensitivity analysis
- use of conservative or worst-case conditions
- asymptotic or limiting solutions
- results of a statistical analysis of laboratory or field data
- case histories or referenced information
- professional judgment
- assumed conditions assigned in the problem definition
- availability and limitations of suitable mathematical (numerical and analytical) solution techniques
- limitations in the quality, quantity, spatial and temporal distribution, and accuracy of model input data

2.2.3 Assess the Significance of Each Assumption

The significance of each assumption must be assessed and quantified to the extent possible. For example, sensitivity analyses, bounding calculations, hypothesis testing using laboratory and field investigations, observations, etc., will be used to evaluate the significance of the assumptions in this activity. The objective is to prioritize the importance of the assumptions relative to the output of the model.

(a) This information can be provided in the record by providing references to pertinent source documents (see Section 2.5).

2.2.4 Identify Assumptions Needing Additional Supporting Information

For those assumptions where the supporting data is considered inconclusive or inadequate, additional information will be required. The decision will be based, in part, on how sensitive the model results are to the assumption.

The results from the activities discussed above provide the basis for many of the laboratory and field investigations. This information is used to either support or refute the hypotheses. This process is illustrated in Figure 1 by arrows showing the flow of activities from the model development activities discussed above to those contained in the "Laboratory and Field Investigations" component. Laboratory and field investigation which are discussed in Section 3.0, play a major role in the validation process. Activities include the "traditional" validation experiments where model results are compared with the output from the model. The record of model development includes the logic for performing laboratory and field investigations and the results of the activities. Although the investigations are separate from the model development process, they are obviously closely tied and can provide most of the information to validate the model. The model development process is an interative one that relies on input and support from laboratory and field investigations (see Figure 1).

2.3 TEST HYPOTHESES OR COMPARE MODEL OUTPUT WITH DATA FROM LABORATORY AND FIELD INVESTIGATIONS

Laboratory and field investigations may be used either to test a hypothesis or to provide data for comparison with the output from a model. These uses will be discussed separately.

As mentioned previously, the complexity of the performance-assessment models and their application (predicting the behavior of the repository systems for long time periods) requires that the assumptions contained in the model also be validated (or tested) to the extent practical. Such validation can be accomplished using existing information or by performing additional investigations (described in section 3.0). This activity presumes that subactivity 2.2.4 identified hypotheses which required additional supporting information and that labcratory or field investigations were performed.[a]

Investigations also provide data for comparison with the output of a model (predicted behavior). Model here refers both to the submodels (e.g. process model) and the assembled model (processes, boundary and initial conditions, geometry, and input parameters). The criteria used to evaluate the overall suitability of the model for simulating the processes involved must be included.

(a) Field investigations include observational data from outcrop studies, trenches, natural analogs, etc.

Note that activity 2.3 is a decision point (see Figure 3). Here, a judgment is made whether the results from the investigation support or refute a modeling assumption or, when comparing results against the predicted results from the model, whether the variance is considered acceptable. The record of this interative process notes what other configurations of the model that were considered and why they were incorporated or abandoned.

2.4 STATE RELATED PLANNED ACTIVITIES

The schedule for the site characterization and performance confirmation activities is such that some of the additional supporting information identified in subactivity 2.2.4 will not be available before a formal technical review is initiated. Furthermore, long-term performance monitoring programs will generate data to be considered in periodic assessments of model validation over the operational phase of the repository. The record of model development should contain a description and schedule for these activities so that the timing of the formal technical reviews can be optimized and reviewers will have an idea of what additional data will be available in the future.

2.5 COMPILE REFERENCES AND DATA SOURCES

The references should include all sources that influenced the development of the model. Sources may include case history studies, natural analog studies, professional journal articles describing the investigations performed, and internal and external reviews (including previous formal technical reviews). The objective is to provide the formal technical review committee (FTRC) with a comprehensive set of resources.

2.6 SUBMIT MODEL DEVELOPMENT RECORD TO FTRC

A model can be submitted for review by the FTRC at any time in its development and, therefore, may receive more than one review. Because they are more complex, system models are particularly likely to have several reviews.

2.7 IDENTIFY ACTIVITIES TO ADDRESS FTRC CONCERNS

The principal product of a formal technical review (FTR) will be a report that identifies needed improvements to the model, areas of unacceptable uncertainty, and perceived flaws in the laboratory and field investigations and their interpretation. Upon receipt of the FTRC report, the principal investigators involved in the development of the model should identify how the FTRC comments will be resolved and implement the necessary activities.

Activity 2.7 is the second decision point in the model validation methodology. Figure 4 illustrates the three possible outcomes. If the FTRC identifies what it considers to be major flaws in the model or in the design or interpretation of the laboratory and field investigations, significant changes in the conceptual model may be required (Section 2.2) and/or perform additional investigations (Section 3.0). If the FTRC concludes that only minor corrections are required, these changes and the level of effort required are likely to be relatively small. Finally, if the FTRC concludes that the model is valid for its intended use, the only additional work required will be a periodic review to verify that new information is consistent with the earlier data.

2.8 DESCRIBE HOW FTRC CONCERNS WERE ADDRESSED

The objective of this activity is to document the steps taken to resolve any deficiencies that may be identified in the FTRC report (Section 4.5). The FTRC report identifies needed improvements to the model and concerns over the design or interpretation of the laboratory and field investigations. This documentation will form the basis for the subsequent formal technical review and should be submitted as part of the model development record (Section 2.6).

3.0 RECORD OF EXPERIMENTAL AND FIELD OBSERVATION DATA

This component of the methodology addresses the details of the investigations that were performed to support model development and validation. These investigations include experiments and observations to reduce the number of competing hypotheses, and those commonly associated with model validation where the results from an investigation are compared with model predictions. The activities in this component include not only laboratory and field experiments but also observational information collected during field reconnaissance studies, natural analog studies, etc. A summary of this information should be provided in the record developed in Section 2.0.

The activities in this component parallel the first several activities in 2.0. The resulting record should include a statement of the problem being addressed, a description of the experimental design or investigation, a statement of the assumptions made and the limitations of the investigation, the technical basis for the assumptions, and an interpretation of the data collected. Each of these components is described below.

3.1 STATEMENT OF THE PROBLEM BEING ADDRESSED

The design of an experiment or field investigation is determined, to a large extent, by the nature of the question being addressed. In general, the design will reflect the rationale the experimenter has used to relate the facts that he hoped to learn from the investigation with those he was relatively certain. Therefore, the objective should be clearly stated for each of the experiments and investigations that was performed. The relationship between the investigation and the hypothesis, conceptual model or submodel (e.g., process model) being addressed should be identified.

3.2 DESCRIBE THE DESIGN OF THE INVESTIGATION

The following information should be documented for each experiment, field study, or natural analog investigation:

- geometry and scale
- initial and boundary conditions
- processes involved
- physical properties
- duration
- locations where data were collected
- equipment and instrumentation used
- procedures and methods used (e.g., sample collection, storage and preparation).

3.2.1 State Assumptions, Hypotheses and Limitations

The simplifying assumptions associated with the investigation must be explicitly stated as they relate to the processes and the conditions being modeled. The investigation should be explained in terms of how representative it is of the conceptual or mathematical model it supports and the differences and limitations present.

3.2.2 Provide Technical Basis for Each Assumption

The basis for each assumption must be explicitly stated (similar to Section 2.2.2). Of particular importance are applicable similitude laws. For example, a scaled-down experiment may be a true representation of one, but not all, of the processes involved. This is especially the case if the coupling of the processes is nonlinear. In that case, the results of the experiment may be applicable to one of the processes, but not the others. The similitude laws can be used to determine to which processes the experiment is applicable.

3.2.3 Assess the Significance of Each Assumption

This subactivity is similar to that in Section 2.2.3. The objective is to address and quantify, where possible, the consequence of each of the assumptions, including assumptions related to data analysis. For example, the displacement between two points can be used to estimate changes in the stress state. The algorithm used to calculate the stresses will depend on a number of assumptions such as the relative homogeneity of the material, whether it behaves as an elastic or plastic continuum, etc. The sensitivity of the results to the assumed conditions can be estimated by comparing the results with those from a bounding calculation.

3.2.4 Identify assumptions needing additional supporting information

Additional efforts may be required to obtain supporting information for those assumptions that have a large influence on the interpretation of the investigation and lacking sufficient evidence to technically support them. Continuing with the example in Section 3.2.3, additional laboratory experiments could be proposed to verify the assumptions of homogeneity and linear elasticity for the material involved.

3.3 PRESENT AND INTERPRET DATA

A record documenting the data that were collected and the methods for analyzing the results must be provided for each of the investigations used to support model development and testing. The record should also identify the various sources of uncertainty in the data and those that were explicitly considered in the interpretation of the results.

4.0 FORMAL TECHNICAL REVIEW

The third component of the model validation methodology involves formal technical review. The FTR's objectives are to evaluate the adequacy of the models and the supporting experimental data for the licensing process and to recommend improvements in the models and data if deficiencies are found. It is important to emphasize that the role of FTRs is to identify needed improvements, not correct them. The FTR provides the information that the modelers, experimenters and their managers can use to begin corrective measures, but the FTR committee members should not become involved in the actual process of correcting model or experimental inadequacies. This distinction is made to ensure independent review, that is, to avoid a situation of reviewing "our" material versus "their" material.

The FTR process generally consist of a sequence of reviews. At the outset, the level of acceptability of the individual hypotheses or components of the model will vary. Some model components will likely be deemed technically adequate while others may be identified as requiring significant work. In these early review stages, the main task of the FTRC is to identify gross defects in the model and to identify a set of issues to be resolved. A working group made up of the modelers, experimenters and others will address these items by performing additional experiments, making changes to various aspects of the model, and documenting these activities in the model development record.

After the technical issues of the committee are resolved, the model is certified as valid for its intended application. As new information becomes available during site characterization and performance confirmation, the model will be reviewed periodically to determine continued applicability and appropriateness and, if it is no longer applicable, to indicate what changes are necessary to make the model consistent with the new information.

The number of FTRs and the amount of time spent on them will depend on several factors:

- the regulatory requirement being addressed and the role of the performance-assessment model in demonstrating compliance
- the complexity of the model, experiments and required data
- the degree of validation achieved at each iteration of the validation process
- the competence and experience of the modelers, experimenters and reviewers
- the extent and nature of the communications among modelers, experimenters and reviewers
- the efficient and effective management and allocation of funds for the entire validation process.

The focus of an FTR is the technical content of the performance-assessment model, not its context in the program (e.g., schedule, manpower, and budget). While there is a technical component to cost and schedule associated with a performance-assessment model, it should not come under consideration as part of the review. However, the FTR is likely to provide important information bearing on the cost, schedule, manpower, and other management concerns related to model development and validation.

The criteria used to decide the validity of a performance assessment model will be determined by the FTRC. The criteria will consist of qualitative and quantitative measures to assist the committee in reaching a consensus and communicate to the modelers and experimenters the specifics of any weaknesses identified in model development or the supporting investigations.

It is important to recognize the actual behavior of complex natural systems will never be completely known because of the spatial and temporal variability involved and limitations in technology for characterizing them. Thus, the accuracy achievable via modeling can only reach a certain level. If the achievable accuracy of the models is inconsistent with the expectations placed on them by the regulations, the licensing process will be seriously hampered.

The validation methodology presented in this document recognizes this limitation and seeks to continuously improve the model through laboratory and field investigations and technical reviews over the operational lifetime of the repository. However, additional work is needed to develop a generic set of qualitative and quantitative criteria for deciding when a model is valid and to evaluate the likelihood the criteria, particularly the quantitative ones, can be met by the various performance assessment models under development. Such an effort should be initiated as soon as possible to verify that the expectations being placed on the models are achievable.

4.1 ORGANIZATION OF THE FTRC

Because the primary function of the FTR is to provide reliable information about the validity of a performance-assessment model for license application, the responsibility for selecting review group members belongs to those DOE management personnel who are responsible for licensing. The DOE can delegate this responsibility to an outside group (e.g., the Nuclear Waste Technical Review Board, National Academy of Science).

The number of reviewers on the FTRC should be sufficient to ensure that the needed expertise and experience are present on the committee. Additional participants may have to be included should a needed expertise be overlooked. The general guideline is to limit the review team to six or fewer participants per model.

The review committee should comprise only qualified individuals for whom there is no conflict of interest in serving on the committee. The review participants should be from outside the OCRWM organization to ensure independent evaluation of the model under review.

4.2 ASSIGN ELEMENTS TO FTRC MEMBERS

The performance-assessment model should be reviewed by its functional elements or divisions such as geometry, initial and boundary conditions, input data used, assumptions, etc. The greater the complexity of the model, the greater the number of possible interactions and the greater the potential for possible inconsistencies in assumptions and boundary conditions. The chairman of the FTRC is responsible for assigning the functional elements.

4.3 STATEMENT OF THE REVIEW CRITERIA

The FTRC should approve a set of criteria to decide the validity of a performance assessment model. The criteria may consist of both qualitative and quantitative measures of the model and its components. The contents of the criteria is the responsibility of the FTRC.

The criteria are likely to vary between models due to differences in the types of uncertainty involved and the ability to quantify them. The modeling uncertainty associated with engineered systems, such as the waste package, is less than that for models of natural systems where the uncertainty from the spatial variability of the physicochemical properties, structure, state conditions, and processes are more difficult and sometimes impossible to quantify. Therefore, the allowable variance between the predicted and observed behavior (in laboratory and field investigations) may be greater for models of natural systems.

The criteria should be documented in the FTR report (Section 4.5) together with a statement identifying criteria that were not met. The objective is to provide the modeler and experimenter a clear understanding of the remaining issues to be resolved.

4.4 CONDUCT THE FTR

The detailed objectives of the FTR are to

- identify needed improvements in a performance-assessment model and experiments
- identify those parts of a performance-assessment models and experiments in which improvement is either not desired or not needed
- confirm the adequacy and degree of the model validation

The review meetings should be open to scrutiny to ensure the quality of the review. FTRC members should come to the review meetings prepared, having reviewed the record provided by the DOE and the assignments given to them by the FTR leader. The modelers and experimenters should summarize the record and clarify questions for the FTRC.

4.5 ISSUE FTR REPORT

The FTRC should prepare a report evaluating the status of the performance-assessment model and experiments reviewed. Criticisms should be reported accurately and specifically. The report should identify needed improvements in the model and experiments and identify those aspects of the model and experiments for which improvements are neither desired nor needed for model validation.

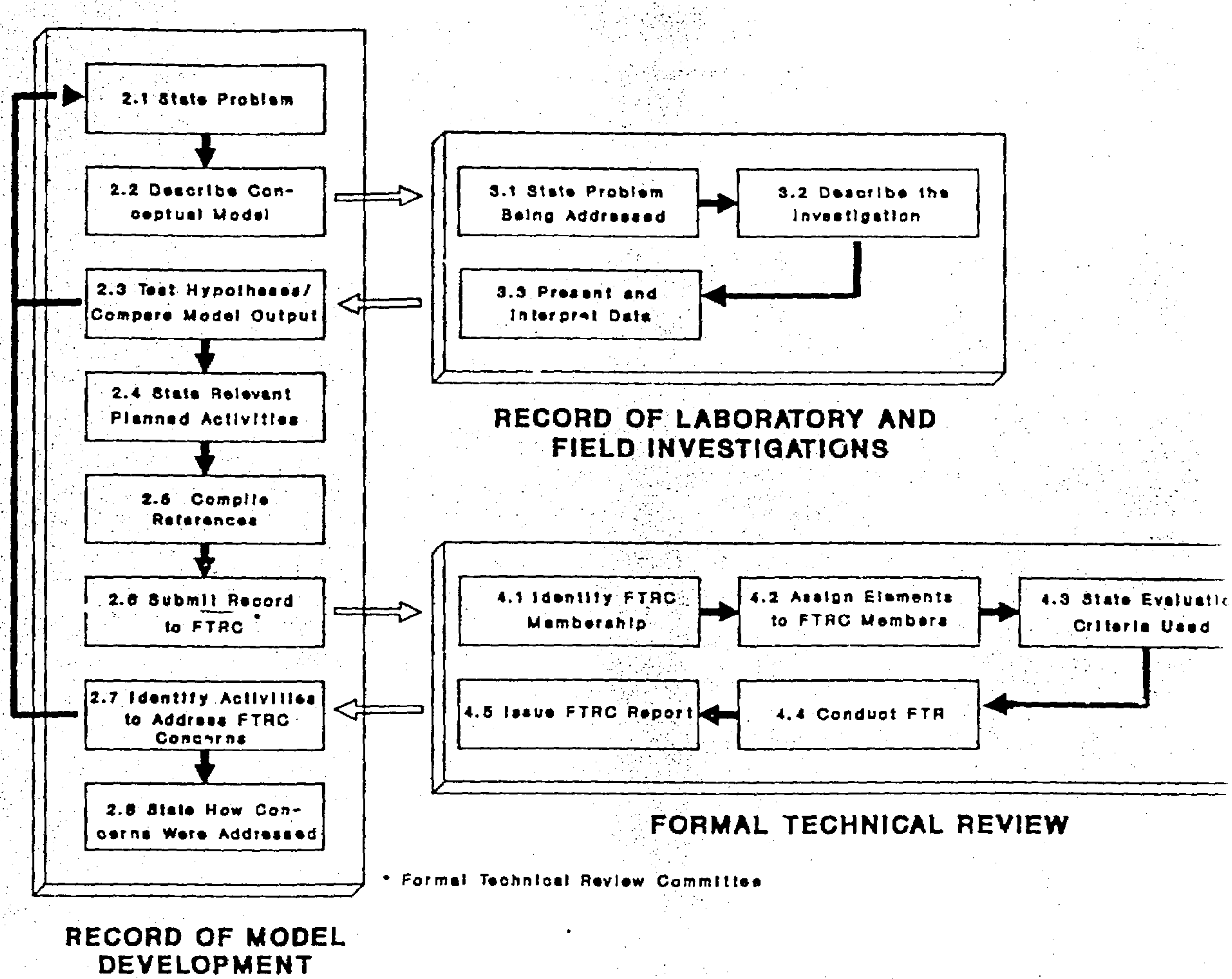

Figure 1. Schematic of the validation methodology for performance assessment models.

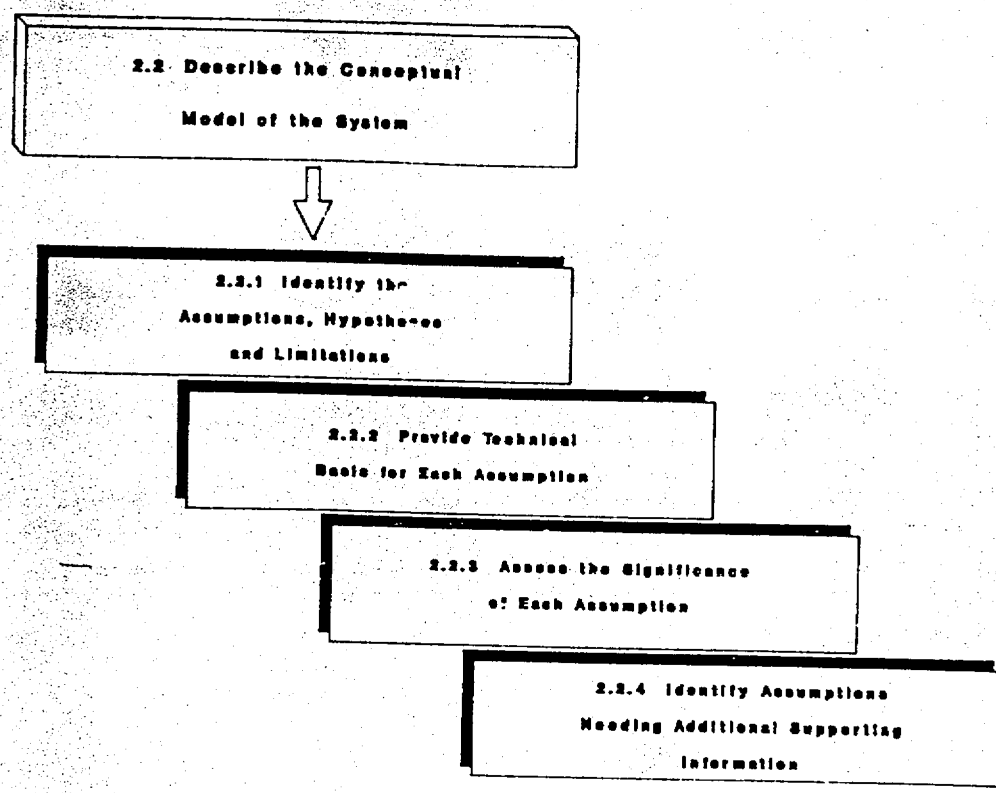

Figure 2. Subactivities to describe the conceptual model of a system.

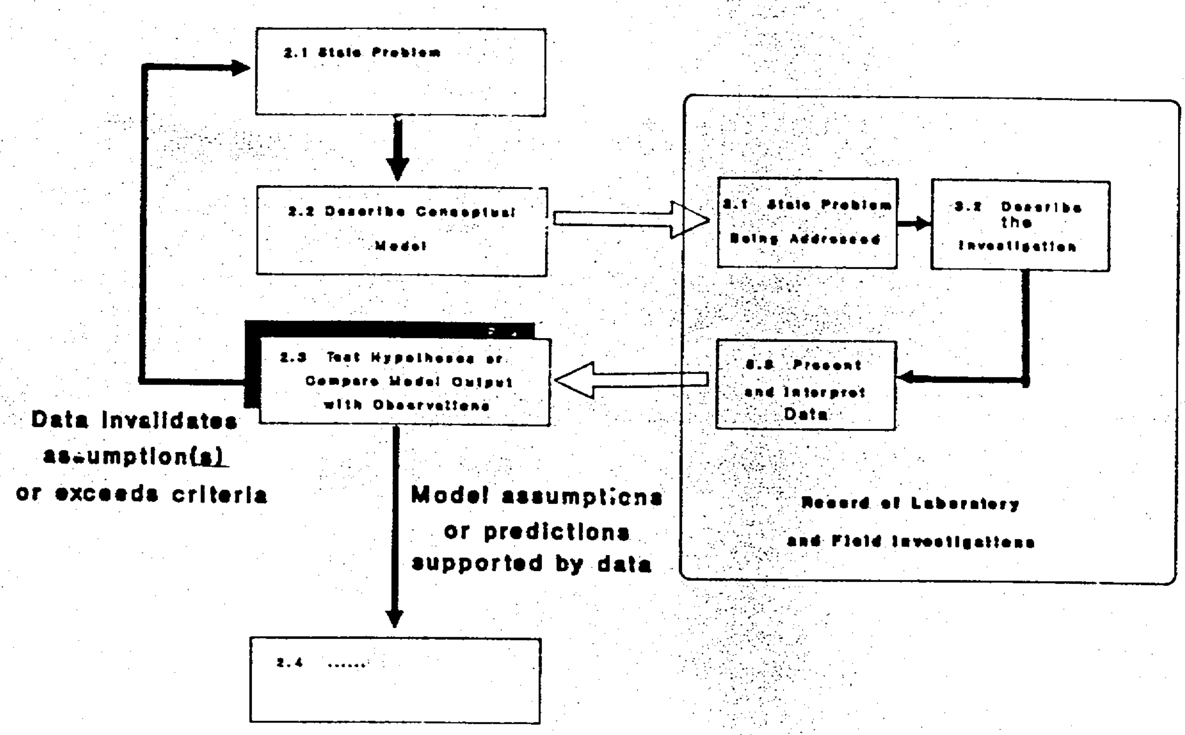

Figure 3. Activity 2.3 is a decision point. The schematic illustrates the iterative nature of model development process.

5.0 REFERENCES

Russell, B. 1948. Human Knowledge, Its Scope and Limits. Simon and Schuster, New York.

Silling, S. A., 1983. Final Technical Position on Documentation of Computer Codes for High-Level Waste Management. NUREG-0586, U.S. Nuclear Regulatory Commission, Washington, D.C.

U.S. Department of Energy (DOE). 1988. Site Characterization Plan, Yucca Mountain Site, Nevada Research and Development Area, Nevada, DOE/RW-0199, 9 volumes. Washington, D.C.

U.S. Nuclear Regulatory Commission (NRC). 1987. "Code of Federal Regulations, Title 10, Energy, Part 60, Disposal of High-Level Radioactive Wastes in Geologic Repositories," 10 CFR Part 60, pp. 627-658, U.S. Government Printing Office, Washington, D.C.

APPENDIX

Definition of Terms

These definitions were taken from the Yucca Mountain Site Characterization Plan (DOE, 1988), Section 8.3.5.20.

Performance Assessment

Performance assessment is the process of quantitatively evaluating system, subsystem, or component behavior relative to the containment or isolation of radioactive waste, to support the development of a high-level waste repository and to determine compliance with quantitative safety criteria.

Performance assessment refers to evaluations of risks and hazards to workers and the public in the preclosure phase of the repository (Section 8.3.5.1) and refers to evaluations of the behavior of the repository for the postclosure phase (Section 8.3.5.8). In particular, as articulated in Section 8.3.5.8, postclosure performance assessment addresses the resolution of Key Issue 1 in the issues hierarchy, which parallels the regulatory system-performance requirements. Thus, performance assessment is a type of systematic safety analysis that is used to (1) predict potential health and safety effects, (2) depict these effects in terms of magnitude and likelihood, (3) compare the results to acceptability standards, and (4) document the process and results in an appropriate and usable format.

Conceptual Model

A conceptual model is an abstraction of the relationships among the system and its component subsystems processes, geometric structures, and bounding environmental conditions. The conceptual model is a set of these relationships, selected from among a larger set of possible relationships and conditions, that is sufficient to describe the system for the intended application of the model to a preclosure safety or postclosure waste isolation assessment. Ideally, these relationships and their alternatives are expressed in terms of testable hypotheses.

Model

A model is a representation of a system that implements the conceptual model in terms of quantitatively linking key features or aspects of the conceptual model with important behaviors, such as containment and isolation. A quantitative model may range in complexity from simple, closed-form analytical solutions of one or more governing equations to numerical models that rely on sophisticated and complex computer codes and resources. Of necessity, mathematical models will only be applied to problems that are mathematically well posed, meaning that a solution does exist.

Code:

A code is a sequence of mathematical expressions and computer instructions written so that a computer can implement those instructions and solve the mathematical expressions as directed. A code, with appropriate data, implements the model, and running the code quantifies the predictions of the model.

Validation:

The concept of validation was defined by the International Atomic Energy Agency (IAEA, 1982) as follows:

> A conceptual model and the computer code derived from it are "validated" when it is confirmed that the conceptual model and the derived computer code provide a good representation of the actual processes occurring in the real system. Validation is thus carried out by comparison of calculations with field observations and experimental measurements.

This definition is adequate for many cases but is not strictly appropriate for the long-term and large-scale postclosure system performance predictions that cannot be compared with field measurements or replicated in a laboratory. The definition does, however, separate the validation problem into two aspects: (1) ascertaining when the model has achieved a good representation of the system and (2) comparing predictive results to appropriate observations and experimental results.

The definition being used in this document is that a model is valid when it is demonstrated that it is appropriate and adequate for the problem being addressed, that it was logically developed using the best available technology, that it can be supported by experimental and observational data, that the quality of the data is high, and that the limitations of the model are well understood.