MAY 11 1988

Dr. Nestor R. Ortiz, Acting Supervisor
Waste Management Systems
Division 6416
Sandia National Laboratories
Albuquerque, NM 87185

Dear Dr. Ortiz:

SUBJECT: COMMENTS ON THE SCOPING DOCUMENT "TREATMENT OF UNCERTAINTIES IN THE PERFORMANCE ASSESSMENT OF GEOLOGIC HIGH-LEVEL RADIOACTIVE WASTE REPOSITORIES" (FIN A-1165)

In July of 1987, R. M. Cranwell and E. J. Bonano submitted a scoping document entitled "Treatment of uncertainties in the performance assessment of geologic high-level radioactive waste repositories". This document was of use in preparing the FY88-89 Statement-of-Work (SOW) for FIN A-1165. In discussions between myself and R. M. Cranwell late last year, we agreed that this document would also be useful as a partial basis for completing subtask 2.2 of the current SOW, but that the draft formal report being requested would need to build substantially on the scoping document, as outlined in the SOW. In the February and March 1988 monthly reports for FIN A-1165, it was implied that lack of NRC comments on the scoping document was hindering completion of this subtask. I am now enclosing NRC staff comments on the scoping document, and look forward to efficient completion of the subtask 2.2 report. Because it is possible that sections of the scoping document will be used verbatim in the subtask 2.2 report, the enclosure includes both content- and style-related comments. I expect all comments to be formally addressed at the time the draft subtask 2.2 report is submitted to the NRC.

The action taken by this letter is considered to be within the scope of the current contract (FIN A-1165). No changes to cost or delivery of contracted products are authorized. Please notify me immediately if you believe this letter would result in a change to the cost or delivery of contracted products.

Sincerely,

/S/

Daniel A. Galson, Project Manager
Operations Branch
Division of High-Level Waste Management
Office of Nuclear Material Safety
and Safeguards

Enclosure: As stated

cc w/encl.:
P. Davis, Division 6416, SNL
E. Bonano, Division 6416, SNL
R. Cranwell, Division 6416, SNL

DISTRIBUTION: SEE NEXT PAGE

WH16
A-1165
WM-11

OFC :HLOB	:HLOB	:HLOB	:	:	:	:
DAG	RC					
NAME:DGalson	:DCodell	:SCoplan	:	:	:	:
DATE:05/06/88	:05/10/88	:05/11/88	:	:	:	:

[OFF]ICIAL RECORD COPY

DISTRIBUTION

Central File	R. Ballard	J. Pohle, HLTR
PDR	J. Bunting	P. Brooks, HLOB
LPDR	NMSS R/F	R. Codell, HLOB
CNWRA	HLOB R/F	D. Fehringer, HLSE
LSS	D. Galson	D. Chery, HLTR
R. E. Browning	S. Coplan	T. Verma, HLTR
M. J. Bell	P. Prestholt	J. Trapp, HLTR
B. J. Youngblood	J. Randall, RES	

NRC STAFF COMMENTS ON THE SCOPING DOCUMENT:

"TREATMENT OF UNCERTAINTIES IN THE PERFORMANCE ASSESSMENT OF GEOLOGIC HIGH-LEVEL RADIOACTIVE WASTE REPOSITORIES"

1. On page 3, first paragraph, it is stated that "Frequently preconceived notions about the behavior of the system resulting from past experiences with apparently similar systems can lead to serious errors in the conceptual model." Can examples be provided of when this has occurred?

2. On page 8, first full paragraph, the use of expert systems to estimate probabilities for conceptual models is discussed. It is unclear how such probabilities would be used, however. It would probably be inappropriate to average or combine such probabilities in any way. The word "probability" is too strong here. Most likely, calculations will need to be carried through to the end for all reasonably supportable conceptual models, and the model leading to the most conservative results used for regulatory purposes. (See also comment 8.)

3. On page 8, second paragraph, "validation" is discussed under the heading of mathematical model uncertainty. However, the process of validation also tests the conceptual model and the computer code. In effect, it can be considered as an overall test of "model" uncertainty.

4. On page 9, second paragraph, the use of natural analogues is discussed. However, an important use of natural analogues is not directly addressed: this is the ability of analogues to provide confidence that the relevant processes are being considered and, that once considered, are being appropriately modeled.

5. On page 10, second paragraph, it is stated that "The cost of computer time and the need to use large computer systems to run these codes may force the analyst to further truncate parameter sampling schemes." It is not clear that this will be the case. Even the cost of a Cray supercomputer dedicated to repository analyses would be trivial compared to the overall cost of repository development.

6. On page 22, first paragraph, the "aggregation" of expert opinions is discussed. It is unclear why this aggregation "must" be done, or how it should be done. Averaging would seem inappropriate as it leads to a loss of the ability to detect uncertainty. Any aggregation done must be done carefully and, even so, could only provide a rough estimate of uncertainty at best.

7. Terminology

The terms treatment, data and parameter should be defined explicitly and then used consistently in future reports. For example, the NRC staff has used the term "treating" as equivalent to "reducing". Apparently, Sandia staff uses the same term as equivalent to "expressing". The terms data and parameter are often used conjunctly in the report. The difference between the two should be clarified and distinctions made between sources of uncertainty as well as methods for expressing and reducing uncertainty with respect to each.

8. Section 2.2., Conceptual Model Uncertainty

Conceptual models are described as an area of uncertainty resulting from simplification of the real system so it can be presented with a tractable mathematical model and the fact that the "real" system is often poorly described (lack of information). The report notes further that techniques must be developed to reconcile results of multiple models. The conclusion reached is that the assignment of probabilities to each of the possible conceptual models is the main issue.

If a probability is assigned ultimately to each possible conceptual model, it is not clear how this probability (uncertainty) will be propagated through to the estimation of performance measures using the performance methodology developed under FIN A-1266 (RES). In the methodology, the full finite-difference model would be used to find dominant flow paths and then a fast ~~rapid~~ one-dimensional model would be used to carry out probabilistic parameter variation. The mechanics of this methodology do ~~does~~ not appear amenable to incorporating a probabilistic "distribution" of conceptual models (i.e., flow paths). Further, it is not clear whether assigning probabilities to conceptual models is a technical requirement and, therefore, an issue. Could not professional judgement be applied when comparing probabilistic output (of a performance measure) from two or more defensible conceptual models without assigning an arbitrary value of "probability"? The focus should be on the defensibility and reasonableness of the conceptual model(s) rather than the probability "value" or methodology by which it was estimated. This subject requires further thought and development. (See also comment 6.)

Section 3.3, Parameter and Data Uncertainty

I have several comments on this section as follows:

9. First, I note that one of the objectives of this scoping document is to provide potential techniques for quantifying and reducing uncertainty. This section stresses techniques for propagating parameter and data uncertainty to the estimation of performance measures. As indicated on page 19, these techniques require information about the values of the parameters themselves (either in the form of probability distribution functions, correlations or ranges). The report does not identify either procedures for quantifying the uncertainty in the value of parameters or procedures for reducing uncertainty in data/parameters. This area remains to be developed in future reports. It would appear logical that this should be a priority because our ability to implement the procedures discussed in this report depends largely on being able to provide defensible parameter distributions as input to the methodology.

10. Second, the first paragraph (page 11) puts forth the conclusion that using single values for parameters in the analysis is generally not acceptable. This conclusion is based on three points:

a) Because the complexity of the systems that need to be modeled and the temporal scales involved, the analyst is often confronted with the difficulty of deciding the values of parameters needed for the analysis;

b) Many of the parameters needed in the models will not be single-valued; and

c) There is likely to be greater uncertainty in obtaining single values for parameters than in defining a distribution of values.

While I agree that using single values for parameters is generally not acceptable because doing so fails to bound the uncertainty associated with those values, the three points supporting this conclusion need to be developed further in future reports. For example, analysts are always confronted with the difficulty of deciding the values of parameters needed for the analysis and as an argument the point is specious. The issue is not the problem of the analyst but the time and scale dependency of the parameter. Further, the point that parameters will not be single-valued needs to be developed further with respect to the time or scale dependency rather than from only the more limited problem of describing the variance of scale dependant parameters throughout the portion of the geologic setting of interest. Finally, the generalization that there is likely to be greater uncertainty in obtaining single values for parameters than defining a distribution of values appears questionable as stated. Perhaps this is because the intended meaning is not clear. There is support in the literature for assuming that the values of the parameters of interest are log normally distributed in nature. If the means of these distributions are derived from "single values" obtained during site testing, does it follow that to impose an assumed log normal distribution around those means will be less uncertain than the actual values themselves? It would seem that the generation of smooth distribution functions for the parameters does not necessarily quantify the uncertainty resulting from the time or scale dependency of parameters or reduce that uncertainty. A more in-depth discussion is necessary in future reports.

11. Third, this section presents several procedures for "propagating parameter and data uncertainty to the estimation of performance measures." These include statistical methods, stochastic methods, interpolation techniques and differential analysis techniques. In future reports it would be helpful to categorize these techniques by more specific objectives. These techniques do not all provide the same specific function. Techniques for developing the spatial variance of parameter values (or probability distribution functions) should be better separated from statistical/stochastic methods which require the former as input. A more precise "functional framework" would better show where each piece falls into place in the overall licensing methodology.

12. Fourth, the section is weak with respect to identifying techniques for developing probability distribution functions. Additionally, the report is skewed toward expressing the spatial variance of parameters (potentially scale-dependent) while overlooking how we can express and reduce the uncertainty which results from possible scale-dependency. Using geostatistics to evaluate spatial variance does not touch upon the scale-dependance issue.

13. Finally, the report should scope out potential uncertainties arising from use of "generic" data/parameters. NRC contractors have raised this issue previously with respect to Sandia's report on "Techniques for Determining the Probabilities of Events and Processes Affecting Geologic Repositories." In responding to such comments, Sandia noted the subject was beyond the scope of that report. However, Sandia did note that the NRC should be alert to the possibility that DOE may use generic data to condition models. This subject should begin to be developed with respect to uncertainty so as to allow future decisions to the degree such a practice would be acceptable in a license

application. Consideration of NRC positions on quality assurance should be included in scoping out this potential area of uncertainty.

14. Section 4 (Sensitivity Analysis)

The first sentence under this section (page 20) states: "One approach for reducing uncertainty in model parameters is the use of sensitivity analysis." Interpreted literally, this sentence is not true. While a sensitivity analysis is an appropriate approach for assessing which parameters cause the greatest impact on calculated performance measures and making decisions as to which parameters most need the uncertainty in their value "reduced", the technique does not deal directly with the fundamental uncertainty in their value. Imprecise language likely has resulted in the area of expressing and reducing uncertainty in data and parameters not to be thoroughly scoped out in this report. (See also comment 23.)

15. General: There is an underlying theme throughout this document that model validation and verification materially affects the ability of the modeler to do the performance assessments. I think that this is overstated, and that the most valuable contribution of models is to the understanding of the system being characterized. I have not seen convincing arguments that all of the benchmarking activities contribute materially to the assessments of performance. Identification of conceptual models and calibration of the computer code to the site are of utmost importance. A good example of this is NUREG/CR-3964 which studied conceptual models of the Hanford site and used the mathematical model to demonstrate various explanations for the observed phenomena.

16. Page 5 first paragraph - The meaning of the last sentence is not clear. Unless you are talking about time-dependent phenomena, I think that you would only have deterministic but distributed parameters in the field rather than random processes. Furthermore, I think that you overlooked a more important reason for scatter in that constitutive variable such as hydraulic conductivity cannot be measured, but only inferred from ~~field~~ data ~~variable~~ such as the response of a well to a pumping stress.

17. Page 6, third paragraph - Methods already exist for incorporating expert opinion and subjective data into performance assessments. See for example Merkhofer (1987).

18. Pages 6, 7 - Scenario probabilities - It is difficult to understand the definitions of axiomatic, "frequentist" and modeling approaches from your definitions. I only understood after going back to the original definitions in NUREG/CR-3964, where the explanation is much clearer.

19. Page 7 last paragraph - I think that expert system technology would be of little use in performance assessment modeling. Such techniques are of value when there are expert humans who already know the answers. This is not the state of affairs in performance assessments of high level waste repositories.

20. Page 8, uncertainty in mathematical model - See introductory remarks above. I think that model validation is oversold. The main application of models should be to develop an understanding of the site.

21. Page 10, third paragraph - Quality must be built into a computer code, not "quality assurance". Your explanation of quality assurance is actually "configuration management", which assures that you know what version you have, and that the version has been run on the standard problems. Quality assurance is more widely regarded as a techniques where the computer codes are developed using procedures that assure each sub-unit of the code work as stated. Poorly developed codes cannot be

quality assured, and no amount of verification or validation can assure that they will work flawlessly for some future set of inputs. This is not to say that configuration management is not important; it certainly is, but let's not confuse it with QA.

22. Page 18 - The descriptions of the differential analysis techniques are somewhat out of date. The local nature of the estimate is no longer necessary. The latest versions of the adjoint methods eliminate this weakness by using the results to fit a response surface much the same as the Monte Carlo methods, but possibly at much lower cost. Furthermore, compiler programs such as GRESS and ADJEN have been developed to add the necessary coding for the differential methods to a FORTRAN code automatically (Pin, 1987).

23. Page 20 Sensitivity Analyses - The differentiation between uncertainty analysis and sensitivity analysis is blurry. Your first sentence in fact is "One approach for reducing uncertainty in model parameters is the use of sensitivity analysis." Furthermore, the sensitivity of the performance measure may be determined but if you don't know the model input parameters, sensitivity analysis isn't going to help you reduce the uncertainties of those input parameters. (See also comment 14).

24. Page 24, Figure 1 - I don't think that this is an accurate flow chart on propagation of uncertainty in performance analysis. It doesn't make any sense. How does the scenario go into the "real system"? What is "system data"? Parameter is misspelled. I think that the flow chart must start with a loop which shows the interaction of the program for gathering field

data and reconciling it with the conceptual model.. Only when the conceptual model is nailed down should there be any testing of scenarios to predict performance. I think that this chart should be discarded and careful thought given to the process.

25. Section 3.3

In the section on methods for reduction of uncertainties, very little attention is given to methods for reducing measurement or observation errors. Section 3.3, which is ~~to talk~~ about parameters and data uncertainty, talks about a whole bunch of statistical techniques such as stochastic, kriging, etc. These techniques are meaningless if the basic input data is faulty. Garbage in equals garbage out. There should be a section which goes into methods for reducing measurement error.

26. On page 6: Modeling requires either axiomatic or frequentist input. Making a comparison between modeling and frequentist /axiomatic is a misleading comparison.

27. Sources of uncertainty should be identified at the element and/or subcomponent level of the Performance Assessment.

Treatment of uncertainty for critical sources at the element and/or subcomponent level should be discussed as:

- reduction of uncertainty
- Analysis of uncertainty
- Expression of uncertainty

	Page	Comment
28.	4	The section on "Computer Code Uncertainty" tries to do too much with too few words. This section mixes a discussion of extracting information from mathematical models by numerical methods with a discussion of computer programs that implement those methods. This section should be preceded with a section with a title like "Uncertainties Attributable to Procedures for Extracting Information from Mathematical Models." Topics covered should include exact solutions and numerical approximations to them, approximate solutions (such as the ones used in DNET), and numerical solutions. In connection with numerical solutions, the report can discuss truncation error (the difference between a differential equation and its numerical representation), stability of numerical algorithms, and convergence of the numerical solutions to exact solutions.
29.	4	Line 4: Replace "truncation" with "roundoff".
30.	4	Line 6: Replace the expression "imported numerical algorithms" with "canned computer programs" if that is what is meant.
31.	4	Line 9: Replace "the latter" with "user error".

	Page	Comment
32.	4	Section 2.3, line 3: Insert "any necessary" before "computer".
33.	6	Scenario Probabilities, lines 6 and 7: Replace "use analytical techniques and numerical models" with "have statistical and deterministic aspects".
34.	7	Section 3.2, lines 8 and 9: Replace "these data are quite massive" with "there are many such data".
35.	8	Uncertainty in Mathematical Model, line 2: Insert "mathematically" after "facility".
36.	8	Uncertainty in Mathematical Model, line 4: Insert "often" after "The equations".
37.	8	Uncertainty in Mathematical Model, line 10: Insert "mathematical" after "a".
38.	8	Uncertainty in Mathematical Model, lines 10 and 11: Delete "as embodied in a computer code".
39.	8	Uncertainty in Mathematical Model, paragraph 2, lines 3 and 4: Replace "accurate to the extent possible" with "realistic".
40.	10	Line 1: Replace "Computer Code" with "Solution Procedures and Their Implementations".
41.	10	Line 6: Insert "in solution procedures" after "uncertainties".
42.	10	Line 9: Insert "faulty solution procedures and" before "computational".
43.	10	Paragraph 3, line 1: Replace "is" with "provide".
44.	11	Last paragraph, lines 3 and 4: Replace "computer model" with "mathematical model".
45.	12	Next to last line: Replace "computer code" with "mathematical model".
46.	13	The bullet items on this page lack stylistic consistency. Some of the items are complete sentences and some are not. Some have subjects and some do not. One style should be used for all of the bullets.
47.	14	Line 1: Replace "A number of" with "Several".
48.	14	Stochastic Models, line 1: Replace "reduce" with "reducing".
49.	14	Stochastic Models, paragraph 2, first sentence: The definition of the inverse problem is not general enough. A good way to define an inverse problem is to compare it with a direct, or well posed, problem. In a direct problem, one is given a spatial domain and must predict dependent variables (e.g. head, concentration, or temperature) as a function of

Page Comment

space and time subject to given initial values of the dependent variables, given values of the dependent variables or their spatial derivatives on the boundary of the spatial domain, and given values of parameters (e.g. properties) that will control the spatial and temporal distribution of the dependent variables. In an inverse problem, one is given values of dependent variables over specified spatial and temporal domains and must determine one or more of the "givens" of the direct problem.

50. 15 Line 6: Replace "are" with "is".

51. 16 Line 1: Replace "nalysis" with "analysis".

52. 17 Line 1: Replace "these are generic" with "the".

53. 17 Line 2: Insert "have been described generally" after "models".

54. 17 Line 4: Replace "methods" with "stochastic models".

55. 17,18 The section on Interpolation Techniques would provide a good introduction to Stochastic Methods if it were moved to page 14 after the second paragraph.

56. 18 Differential Analysis Techniques, line 9: Replace "analytically" with "exactly".

57. 18 Differential Analysis Techniques, line 10: Replace "complex" with "complicated".

58. 19 Third bullet: Replace "computer" with "mathematical".

59. 19 Bullet items: See comment on page 13.

60. 19 Before the paragraph beginning with "From the discussion", insert the title "Concluding Remarks".

61. 20 Section 4, paragraph 3, lines 1 and 2: Replace "computer ... system" with "deterministic mathematical models".

62. 20 Paragraph 3: This paragraph repeats material from Section 3. It could be shortened by making appropriate references to Section 3.

63. 20 5th line from bottom: Replace "computer code" with "mathematical model".

64. 21 Line 4: Define "computer calculus".

65. 21 Section 5, paragraph 2, line 2: Replace "dependent" with "depending".

66. 23 Line 2: Replace "treatment" with "estimation".