SEP 11 1990

Dr. Evaristo J. Bonano, Supervisor
Waste Management Systems
Division 6416
Sandia National Laboratories
Albuquerque, NM 87185

Dear Dr. Bonano:

Enclosed are NRC staff comments on the draft report prepared under Subtask 3.2 of FIN A1165, Techniques for Estimating Probabilities of Events and Processes Affecting the Performance of Geologic Repository: Assessing Compliance With the EPA's Containment Requirements (40 CFR 191.13). The title of the report is "Techniques for Determining Probabilities of Events and Processes Affecting the Performance of Geologic Repositories: Volume 11 -- Suggested Approaches.

Comments are based on reviews by staff in the Division of High-Level Waste Management, the Probabilistic Risk Analysis Branch of the Office of Research, and the Office of General Counsel. They are directed toward needed corrections, revisions, and clarifications, but also include rather extensive statements of staff opinion to provide background for the comments.

The action taken by this is considered to be within the scope of the current contract (FIN A-1165). Please notify me immediately if you believe this letter would result in a change to cost or delivery of contracted products.

Sincerely,

Pauline Brooks
Technical Monitor, FIN A1165
Division of High-Level Waste Management, NMSS

Enclosure: As Stated
cc: P. Davis, SNL Div 6416
L. Price, SNL Div 6416
D. Tiktinsky, NMSS/PI

DISTRIBUTION:

Central Files	HLGP r/f	NMSS r/f
RBrowning, HLWM	BJYoungblood, HLWM	JBunting, HLEN
RBallard, HLGP	JLinehan, HLPD	SCoplan, HLGP
PBrooks, HLGP	PDR & LPDR	

OFC :	HLGP	:HLGP
NAME:	PBrooks/ga	:SCoplan
Date:	9/14/90	:9/10/90

A-1165
WM-11
NH1E

ENCLOSURE

NRC STAFF COMMENTS ON SNL DRAFT REPORT 3.2, TECHNIQUES FOR DETERMINING PROBABILITIES OF EVENTS AND PROCESSES AFFECTING PERFORMANCE OF GEOLOGIC REPOSITORIES: VOLUME II--SUGGESTED APPROACHES

GENERAL COMMENTS

1. A major concern expressed by NRC staff concerning this report is the emphasis on the use of expert judgment in determining probabilities without stating enough caveats about the need for "hard" data and valid models.

   The report states in many areas that it is not intended to serve as an example of a methodology for eliciting expert judgment. However, examples on the use of expert judgement dominate the report. In Chapter 5 there is little presented except a discussion of the URS report followed by an example of the use of expert judgment. In Chapter 6, the conclusion is that expert judgment must be used. These are examples of how the report appears to suggest that the primary method for obtaining probabilities will be "expert" opinion.

   Expert opinion has and will be used in NRC licensing procedures. However, in all cases that we are aware of, its acceptance in the licensing process was based on the apparent validity of the "hard" data which was the foundation for the expert opinion. While we agree that expert opinion will probably play a larger role in the licensing process for a high-level waste repository than has been used in past licensing actions, the need for "hard" data will not diminish. What is needed is an evaluation of the types of expected hard data and an evaluation of techniques which can be used by the experts to provide objective evidence that the probabilities are valid. Although the report gives some guidance in this area, far more is needed.

   The survey (Table 5-5) in the tectonics discussion seems to call only for the conclusions of the experts, with the factual basis and reasoning of those experts merely inferred by the analysts. This procedure will not be acceptable in a licensing proceeding. For a licensing board will "in all circumstances [have] the right, indeed the duty, to satisfy itself that the conclusions expressed by expert witnesses on significant safety or environmental questions have a solid foundation. To this end, Board members are free to examine the witnesses themselves respecting the basis for opinions which they express - including the methodology or assumptions underlying the analyses which led to those opinion." South Carolina Electric and Gas Company (Virgil C. Summer Nuclear Station, Unit 1), ALAB-663, 14 NRC 1140, 1163 (1981).

   Unbiased expert opinion cannot be obtained when the process used in the selection of experts does not adequately represent the major field being evaluated. The report has acknowledged that the selection of the experts was flawed. The decision to proceed without replacement of expertise in the field of geology biased the end results.

Unbiased expert opinion cannot be elicited if the experts have been given reports as reference which provide conclusions on areas in which the expert opinion is to be solicited. At best, what can be obtained is a peer review of the reference material given the assumptions presented in the reference material. The approach demonstrated in Chapter 5 was further flawed because inappropriate reference material was given to the "experts".

Unbiased expert opinion cannot be obtained if the personnel who are reviewing the input of the reviewers are unqualified or biased in their opinion. In such a case, the best that can be obtained is a reflection of the preconceived notions of the reviewer. The example might well be used as a basis for more extensive treatment of the pitfalls and possible difficulties in using expert judgment.

2. The report gives a lucid exposition of a Bayesian approach which has been extensively applied to probabilistic risk assessments of nuclear plants. The Bayesian approach presented in the subject report has the virtue of explicitly using all available information--from historical data, model results and expert judgment-- to estimate performance probabilities for geologic repositories. However, while it may have desirable theoretical properties, this complex Bayesian approach has the drawback of being more sophisticated than much of the information to which it is applied. As the report concedes, "a major objection that is usually raised against the methods discussed here is that they require too much information (b_i, σ_i, etc.), and that they are more sophisticated than the pool of expert estimates to which they are applied" (page 3-68) and "... the fact remains that assessing reasonable rules for b_i, σ_i, and ρ_{ij}, is very difficult" (page 3-71). As a consequence of this imbalance between the methodology and the information to which it would be applied, extensive sensitivity studies would be required to ascertain the extent to which the estimated performance probabilities are determined by the assumptions required by the Bayesian approach. To avoid having estimated performance probabilities with very large uncertainties, the analyst and/or decision-maker may be tempted to make assumptions which cannot be justified on scientific grounds.

The report goes into some detail on the sources of the difficulties in applying the Bayesian approach. However, by not presenting any real alternatives to the Bayesian approach, the report implies that there is no alternative and possibly makes it the more likely that the practitioner will succumb to the pitfalls of the Bayesian approach as presented in the report.

There are, in fact, several possible alternatives to the Bayesian approach which should be considered in cases where there is much scientific uncertainty. Like the Bayesian approach, these alternative approaches are based on all available information, but they also try to minimize the structure (i.e., assumptions and models) required to estimate performance probabilities. The alternative approach with perhaps the least structure is to simply present all available historical data and model results to a

panel of experts and, after extensive discussion and elicitation training, ask each expert to estimate an uncertainty distribution for each performance probability. The results for each expert can then be combined by a Monte Carlo procedure as in Bernreuter et al. (1989). The resultant uncertainty distribution is a substitute for the posterior in the Bayesian approach, and can be useful when there is much disagreeement within the expert panel. There are two other alternatives to the Bayesian approach when not enough information is available to justify using the Bayesian approach. One is Bayesian geostatistics, which is a formulation for updating prior soft information. While this approach is based on Bayesian updating, it is a non-parametric approach which does not require assuming an arbitrary or mathematically convenient likelihood function. A second alternative is the bounding approach, which shows how to propagate uncertainty when some input variables can only be characterized by bounds.

There are also complementary methods to consider that are based on the probability distribution approach. In addition to Bayesian geostatistics and the bounding approach, there are other modifications and alternatives to the basic probability distribution approach. All of these methodologies should be considered before deciding on an approach to estimating probabilities.

Alternatives to the Bayesian approach should be noted in Chapter 3.

SPECIFIC COMMENTS

1. Page 1-1, line 24. At this point, or later , on page 1-2 it is noted that the first volume reviewed the literature on techniques for predicting the occurrence of events and processes in five areas, but this volume treats only three of the five. It would be helpful to the reader if the reason for omitting two of the areas were stated in terms of the volume's purpose. (See also Editorial Suggestions on page 10 of this enclosure.)

2. Page 2-3. Screening of low probability events and scenarios should not "eliminate" things from consideration if it is done correctly. To assure adherence to mathematical requirements, it is necessary that those things "screened out" be moved into another category, probably one which is not subjected to a detailed analysis, or one which is handled as the base case. In this way it is possible to make sure that the probabilities add up to as close to one as possible.

3. Page 3-2. NOTE: While Poisson distribution is often assumed as correct, there is more and more evidence that this is not the correct distribution function for tectonic processes, especially when considering the time frame of interest for high-level waste analysis.

4. Page 3-51. The notes presented on this and the following pages are of interest as they provide some of the mathematical bases to demonstrate how the biases of the decision maker can skew the results.

5. Page 4-2. The fact that global climatic models are inaccurate predictors on the regional scale is one reason that regional models and the like are needed. See for example, the papers presented in FOCUS 89 and IHLRW Conference (90) on this subject.

6. Chapter 5. During previous discussion with Sandia personnel, the geology/geophysics staff commented on the "lack of realism" in the questionnaire and in the lack of validity in the resultant conclusions. One way of stating the basic facts relating to the methodology outlined for the tectonic setting is:

   1) three out of four geologists declined to participate. All three of those who declined to participate were in academia.

   2) only one respondent chose the geologic observations as a data base. This respondent's input to the evaluation was eventually rejected due to inconsistencies in his/her answers.

   3) a basic conclusion was that if there had been "a representative percentage of a large number of experts, the choice against the geologic observations of Whitney et al.[1986] as a basis for estimating the length of the seismic cycle on the Windy Wash Fault would seem reasonable."

These points demonstrate that any probability developed from an expert opinion approach must be examined very carefully in light of flaws that can be hidden in the evaluation. Specifically, the probability of an event that resulted from this analysis would have as a basis the facts that:

1) The majority of potential experts who might be described as academics and who might have expanded the distribution of answers declined to participate.

2) The one respondent who chose to use the basic data available and whose responses to questions 9-15 seemed to be the most consistent with available data, gave inconsistent responses to other questions and so was eliminated from the survey.

3) Apparently, if the answers to the questions were consistent with the URS (1987) report, then rejection of the basic data as a reference was deemed acceptable.

From the foregoing, one might conclude that this particular poll of experts resulted in an analysis that may be seriously flawed. Therefore, any methodology for developing probabilities should include a process that provides for an in-depth characterization (from a scientific as well as a statistical perspective) of the analysis to illustrate any potentially significant hidden factors that might be incorporated in a probability determination.

7. Page 5-1 to 5-2. This section could be deleted without significant adverse effect on the thesis.

8. Page 5-1 to 5-2 last paragraph. Page 5-1, line 26 states that "most of the significant faults in the vicinity of the proposed repository have been identified and mapped." However, on Page 5-2 the report notes that "the ability to relate earthquakes to specific faults and identification of the type of faulting involved in earthquake generation is in a very primitive stage of understanding" and that "regional seismic records of earthquakes are sparse." In view of the "primitive understanding" and "sparse records" we are concerned about the accuracy of the statement that "most of the significant faults... have been identified and mapped." We recommend some additional qualification of this statement.

9. Page 5-3 to 5-4. The statement that the ranges are basement-cored automatically provides the "accepted" tectonic model which is being used by the person reviewing the "expert" opinion, places constraints on the characteristics of the fault; places constraints on the "expected" earthquake; and places constraints on the probability of occurrence. Most potential tectonic models which are in the recent literature for the Yucca Mountain region use some type of "detachment" or rely on some type of wrenching, and many of the the models incorporate aspects of both. While there still is the concept of a "basement-cored range" in some of the

models, this is by far a minority opinion. By providing the experts the accepted tectonic model this study has biased the results with out-of-date information.

10. Page 5-4. The Windy Wash fault is north trending and is one of the youngest faults documented at Yucca Mountain. How is this north trending fault related to the northeast trending features which are described as the youngest features in the region?

11. Page 5-4. What assumed relationship is there between the Mine Mountain structural zone and the Rock Valley structural zone? In addition, does this mean that the magmatic bodies in the area of the site such as those which are responsible for Lathrop Wells cone are related to the Mine Mountain structural zone?

12. Page 5-5. Wasn't Mehrer talking about faults which didn't rupture the surface in areas of upward weakening crusts? How did the Garlock fault studies, and studies of faults in the Las Vegas Shear zone (which did rupture the surface) confirm Mehrer's model?

13. Page 5-6, line 5. More strictly speaking, the historic seismological record is short.

14. Page 5-6 and 5-7. Section 5.1.3 should be clarified. It appears to say that the faulting at Yucca Mountain. has nothing whatsoever to do with the faulting in California. Considering the location of the site in relation to the Death Valley faults, The Walker Lane, the Las Vegas Shear zone, etc, this may be a difficult view to justify. Deep geophysics in the area could be used to suggest that this zone is a transition between the "California" type tectonics and the true Basin and Range type tectonics.

15. Page 5-15. What is the basis for the statement that the Windy Wash fault is analogous to the Ghost Dance fault? For example, is there any evidence that the Windy Wash Fault zone displays the scissors type displacement that the Ghost Dance fault displays?

16. Pages 5-18 through 5-25. Section 5.4.2. While this section is caveated with statements that its purpose is to demonstrate a procedure and that it is not important that there be the best or right model, the data bring out many specific questions which relate to the present data base. For example:

    How does this procedure on modeling tie into the horizontal component of movement? The procedures used to date have only had information on the vertical component of movement and have basically ignored the horizontal component. Is this important or can we ignore the horizontal component?

How does this modeling take into account the non-periodicity of faulting and seismicity, i.e., clustering? What is the right distribution function? How can we decide which is the right function, and what effect is there if we use the wrong function?

How does this modeling tie into various models of the tectonic framework? The way this information is presented implies that there is no need to understand the tectonics. All that is needed is to play games with statistics. How do we handle two or more different tectonic models with varied probability results, each of which (at least on the surface) have equal validity? How, from a regulatory perspective, do we handle two different models which result in two different consequence models? Is there a mathematical procedure we can use which relates to tectonics, a mathematical procedure which relates to decision theory, or is there an alternative?

17. Page 5-30, line 4-6. When trying to get "expert" opinion, it is very important that the respondents not be biased by the information presented. In this case, the URS report contains curves. The result was an assessment of that curve, based on the information presented in that report. Thus, the results may have been invalidated prior to even asking the experts a question.

18. Page 5-34 and 5-35, Question 1 and 2. To those who have been working with nuclear facilities, maximum credible earthquake is normally an earthquake which has been determined using Appendix A criteria. The assumption that there is a specific relationship with the characteristic earthquake, such as suggested by reference to Anderson, is not valid unless the terms maximum credible earthquake and characteristic earthquake are specifically defined. They have different meanings for different people in different contexts.

    Question 3. It is obvious from the discussion on the top of page 5-36 that the authors of this section have already been biased into accepting the URS data as the " correct" answer.

    Question 5. Along with other problems, this question and the response ignore the data which suggest the existence of "clustering" of tectonic events and the suggestions that the assumption of a Poisson process is invalid for the time frame of interest.

    Question 8. This response further demonstrates that the reviewer has a preconceived notion of the seismology of Yucca Mountain and that any answer which is given by an expert who disagrees with the person reviewing the questionnaire will be ignored or downgraded.

19. Page 5-46. Experimental results indicate that people tend to overestimate their level of confidence by about a factor of 2. Accordingly, a claimed 90% range is likely to be closer to a true 50%-confidence interval than to an 80%-confidence interval (Meyer and Booker (1985). p. 49).

20. Page 6-3, lines 1-14. The same reference (NRC 1983) also states that "intrusion for such purposes would have to be reviewed in the licensing process if the particular circumstances are sufficiently credible to warrant consideration." In other words, how deliberate intrusion will be handled is not as cut and dried as this section would seem to imply.

21. Page 6-3, line 15-24. The NRC has not ruled on anything related to this subject. The reference is a draft which is based on an EPA standard which is under revision.

22. Page 6-7, lines 9-16. A society which has the technology to drill boreholes several thousands of feet deep and would be likely to have a technology which allows them to understand radioactivity. This example appears to contradict item 4 on page 6-7.

23. Pages 6-13 and 6-14. The last sentence on page 6-13 (going to 6-14) is one of the many areas in this report where it appears that "subjective" judgement is being suggested as the main method for obtaining probabilities.

24. Page 6-15, and page 6-27. The drilling rates given by EPA are non-mandatory guidance and the DOE is going to have to use a drilling rate which is acceptable to the NRC. This drilling rate may be much higher that then EPA numbers and, at present, is not constrained by anything EPA or NRC has said.

25. Page 6-18, line 8. It is not clear that having experts supply the weights for a composite drilling rate is any less speculative than Harbaugh's two approaches. Furthermore, averaging over drilling rates which differ by several orders of magnitude can yield misleading estimates of the probability of human intrusion. It would be preferable to present the decision-maker with the consequences of the various drilling rates, along with the experts' subjective weights, and then encourage the decision-maker to use decision-theoretic techniques to decide whether or not the results are acceptable.

26. Page 6-29, line 6. The Central Limit Theorem applies only if there are a sufficiently large number of terms and no small number of terms dominates the sum. These conditions, along with the likely ske wness of the distributions, implies that the composite rate would most likely not be normally distributed. However, this is not a drawback, as the distribution of the composite rate could be established by Monte Carlo sampling.

27. Chapter 6. The human intrusion chapter erroneously implies that merely because some analysis conforms with EPA's guidance for implementation (rather than its numerical standard), it would be "legally justifiable" (p. 6-15). The guidance by its very terms is not binding upon NRC, and this agency has consistently taken the position that because it deals with matters of implementation it cannot take precedence over the independent judgment of the Commission in applying the standard itself.

    Moreover, insofar as human intrusion is concerned, the report fails to give weight to the assumption contained in the NRC definition of "unanticipated processes and events." Those assumptions will play a major role in the analysis of both the probability and the consequences of human intrusion events.

28. Page 7-2, lines 9-12. This section reinforces the previously stated opinion that the procedure used was flawed as the report the influence that the URS report had on the final results.

29. Page 7-2, line 22-23. We agree that it is quite bothersome that the data from the one study referenced on the Windy Wash was thrown out.

REFERENCES

These are references for General Comment 1.

Bernreuter, D.L., J.B. Savy, R.W. Mensing, and J.C. Chen (1989). "Seismic Hazard Characterization of 69 Nuchlear Plant Sites East of the Rocky Mountains." NUREG/CR-5250, UCID-21517. Lawrence Livermore National Laboratory, Livermore, California.

Meyer, M.A. and J.A. Booker (1990). "Eliciting and Analyzing Expert Judgment." NUREG/CR-5424, LA-11667-MS. Los Alamos National Laboratory, Los Alamos, New Mexico.

EDITORIAL SUGGESTIONS

Some editorial suggestions are given below for your consideration.

1. Some clarification of the terms used to describe the topics covered in this report seems to be needed. In various parts of the report, different terms seem to be used to name the same topics and this could be misleading.

   The Table of Contents (pages ii to iv) leads the reader to three areas of application of the methodology: Section 4.0 "...Climate," Section 5.0 "...Tectonic Events," and Section 6.0 "Human Intrusion."

   The stated Purpose of the report (as given in Section 1.1, L 25, 26) indicates that the report "...looks at three of the five areas [covered in the first report] (resource exploration, climatology, tectonics and seismicity) and..."

   The three topics, as given in the Scope (p 1.2), are climatology (L 22), tectonics (L 24), and resource exploration (L 25).

2. The document does not have an executive summary. Some readers would be served well by a two to three page summary (up front). The executive summary should be more detailed than the existing Summary and Conclusions, which is presented as Section 7.c. The following facts along with those given from page 1-2 (line 21) to page 1-3 (line 2) might be appropriate input to use when the first or second paragraph of an Executive Summary for this report is being drafted:

   This volume, the second of a two volume set, addresses generic approaches for calculating probabilities for events associated with deep geologic disposal of HLW. [The name of the first volume, and something about it might go here.] The methodology for these calculations is demonstrated in this volume for the areas of climate change, tectonics, and human intrusion. These topics are part of the five main topics covered in the first volume: resource exploration, climatology, tectonics and seismicity, seismic hazard assessment, and volcanology.

   The purpose of this report is to "outline a methodology for calculating probabilities." [State why this this needed.] A mathematic basis for combining information from the main sources of information (historical data, models, and expert judgement) using Bayes Theorem has been developed; discussions of the application of this method are given for three areas. Three areas were chosen because the nature of the sources of data expected to be available for each of these areas is sufficiently different as to furnish variations for demonstrating the suitability of the methodology.

The Executive Summary would also serve well if it contained some of the details concerning the available information and the predicted behavior.

3. Page 2-2, Line 7. Change "to estimate" to "for estimation of."

   Reason: As worded, the reader must determine whether "to estimate" modifies "methods" or "to devise." "Methods for estimation of" seems less ambiguous.