REVIEW OF THE HIGH-LEVEL NUCLEAR WASTE REPOSITORY SITE ANALYSIS

FINAL REPORT

Prepared by:

ECO Northwest

Prepared for:

Washington State Department of Ecology

December 19, 1986

TABLE OF CONTENTS

Review of the High-Level Nuclear Waste Repository Site Analysis

1.0 Introduction

High-level radioactive wastes produced by nuclear power plants operating in the United States currently are stored using procedures considered to be temporary and adequate only for the short run. The Nuclear Waste Policy Act (NWPA), passed in 1982, outlines how the federal government proposes to dispose of this material. A key to this process is the identification of a single preferred site for a nuclear repository, designed (within capacity constraints) for permanent storage of all high-level wastes produced in the United States.

Draft environmental assessments of the nine leading sites were released in December, 1984, and used as the basis for narrowing the choice to five sites. These five sites were then evaluated using three alternative methods, and a subset of three sites--Yucca Mountain (Nevada), Deaf Smith (Texas) and Hanford (Washington)--were selected for further intensive study, a process called "characterization". However, the methods used by the Department of Energy in going from five sites to three were widely criticized, by private citizens as well as by organizations such as the National Academy of Science. As a result, in late 1985 a second study--employing the more rigorous evaluation methodology of multiattribute utility analysis--was commissioned.

Citing the results of the second study as well as additional analysis undertaken within the department, Secretary Herrington of

the DOE then recommended to the President, on May 28 1986, that three sites be advanced for the characterization that is required under the NWPA. Such characterization will involve numerous engineering investigations of a site, including the construction of an exploratory shaft, tests of the adjacent rock, and subsurface tunnelling. The three sites proposed by the Secretary, and subsequently approved by the President, again are at Yucca Mountain (Nevada), Deaf Smith County (Texas), and Hanford (Washington).

The present report comments on only the most recent portions of the repository-siting process, and focuses on two studies released by the Department of Energy (DOE) in May, 1986. The first is a multiattribute utility analysis of the five selected sites (DOE/RW-0074), titled *A Multiattribute Utility Analysis of Sites Nominated for Characterization for the First Radioactive-Waste Repository--A Decision-Aiding Methodology*. The second is the Recommendations report that was simultaneously released by the Secretary (DOE/S-0048), titled *Recommendation by the Secretary of Energy of Candidate Sites for Site Characterization for the First Radioactive-Waste Repository*. In general terms, the purpose of our report is to provide Washington State's Nuclear Waste Board with an independent review of the decision-aiding methodology employed by the Secretary of Energy to select candidate sites for characterization for the first high-level nuclear waste repository. More specifically, its purpose is, first, to briefly explain and critically review the assumptions and approach of the multiattribute utility analysis and, second,

to examine the assumptions and credibility of the Recommendations Report.

Our comments therefore are organized under these two general headings, pertaining to the multiattribute utility analysis itself and to the official Recommendations report. The reader should keep in mind that our purpose in this brief report is to present a straightforward and comprehensive review of a sophisticated evaluation methodology that has been applied in an exceptionally complex decision environment. Thus, we at times may be guilty of oversimplification; however, our intent throughout has been to faithfully represent both the evaluation methodology and its application to this specific facility-siting context.

2.0 Comments related to the multiattribute utility analysis

2.1 The choice of multiattribute utility analysis as the evaluation methodology

Multiattribute utility technology (hereafter MAUT) was chosen as the single methodology for evaluating the five sites nominated as suitable for characterization. In this section we ask three preliminary questions--what is MAUT?, why MAUT?, and how was this MAUT structured?--before proceeding in the next sections to ask how well this multiattribute utility analysis was carried out.

2.1.1 What is MAUT?

Multiattribute utility analysis is a decision-aiding technology that is based on a well-documented theoretical structure developed over the past forty years as part of the broader field known as decision analysis (e.g., von Neumann and Morgenstern, 1947; Pratt,

Raiffa and Schlaifer, 1964; Keeney and Raiffa, 1976). It is designed to organize and present evaluative information (Edwards and Newman, 1982), with the special distinction that it is able to combine both judgmental (subjective) and factual (objective) assessments as part of the same evaluative structure. For example, the critical objects of evaluation are determined explicitly at the beginning of a multiattribute analysis, with the idea that this effectively identifies the most important elements of the decision function. Then at the end of the analysis, a specific action (e.g., selection of the preferred site) is recommended following the analysis of inputs from, and sensitivity tests of the trade-offs between, all designated attributes.

A typical MAUT proceeds sequentially though a series of steps, as shown in Table 1. These steps involve:

1. Identify all key considerations for the decision:
   a) What are the objectives or goals?
   b) What options should be considered?
   c) What aspects or attributes need to be considered? These aspects both describe the options and show the degree to which each option attains the goals.
   d) Who are the stakeholders, that is, the people whose values are relevant to the decision?
2. Model the decision: Choose the form for combining the separate pieces of fact and value to measure the adequacy of each option.
3. Measure all needed inputs for the model:
   a) Assess the facts: this is a detailed description of each option in terms of each attribute

TABLE 1: STEPS TO FOLLOW IN CONDUCTING A MULTIATTRIBUTE UTILITY ANALYSIS

1. Identify key considerations:

   - objectives
   - options
   - attributes
   - stakeholders

2. Model the decision

3. Determine required inputs:

   - assess the facts
   - assess the uncertainties
   - assess the utilities and value trade-offs
   - verify independence assumptions

4. Derive total utility measure for each option

5. Conduct sensitivity analyses

6. Recommend a decision

b) Assess the uncertainties in critical inputs

c) Assess the values or utilities for each attribute and the value tradeoffs between attributes

d) Test whether the model chosen in Step 2 is appropriate by verifying certain independence assumptions.

4. Combine all the elements, arriving at a total utility measure for each option.
5. Do sensitivity analyses to see which inputs, if changed, would alter the relative rankings of the total utility measures across the options.
6. Use the insights gained from the first five steps to make a decision.

Multiattribute utility analysis therefore provides a method for quantifying complex preferences among alternatives that vary on multiple, and often conflicting, objectives. It can clarify individual or public values, as a means to gaining a better understanding of people's differing perspectives on controversial options, and it can help to resolve those conflicts by providing an explicit value structuring process that compares multidimensional options in commensurate terms.

2.1.2 Why MAUT?

There are two sets of reasons why multiattribute utility analysis works well in the context of siting a high-level nuclear waste repository. A first set of reasons relate to the logical and theoretical properties of a MAUT: we make five points here. First, its assumptions are explicit: this makes it a good candidate for evaluation in a controversial policy setting

because differing perspectives can be represented easily by recalculating the results using different inputs. Second, it provides a basis for comparing unlike attributes of a decision through the development of value trade-offs between attributes. Whereas a benefit-cost analysis might employ a multiple accounts structure, with trade-offs between the accounts strictly forbidden, MAUT makes the trade-offs between attributes explicit; any attribute can be expressed in terms of any other. Third, MAUT yields the most easily understood output--a number. The procedure itself cannot make decisions, any more so than any other decision-aiding evaluative structure can, but it does provide a method for unambiguously ranking competing options. Fourth, MAUT easily incorporates uncertainty regarding the possible outcomes resulting from the choice of an option. Fifth, sensitivity analyses within a MAUT identify those aspects of the decision that are important, that is, the parts that capture significant differences between options.

A second set of reasons why MAUT is appropriate has to do with the fact the MAUT is a familiar decision aid to use in facility-sitings and in decisions involving risk or uncertainty. Numerous examples of relevant applications of MAUT can be found in the published literature in the field, for example, by Edwards and von Winterfeldt (1986) or by Keeney and Raiffa (1976).

2.1.3 How was this MAUT structured?

In the course of any analysis, decisions need to be made by the analysts and client in order to operationalize each stage of the assessment procedure. In this MAUT analysis of alternative

high-level nuclear waste repository sites, several key elements of the decision process (shown in Table 2) merit attention.

First, the overall objective--to minimize the adverse impacts of a repository--was divided into two lower-level objectives: minimize adverse preclosure impacts, and minimize adverse postclosure impacts. More detailed objectives then were derived from the DOE's system guidelines, published in 1985: for postclosure impacts, the objective is to minimize adverse health and safety effects; for preclosure impacts, the objective is to minimize adverse environmental, socioeconomic, health and safety effects, and costs.

For the postclosure impacts, two performance measures were used: cumulative releases of radionuclides to the accessible environment during the first 10,000 years after repository closure, and cumulative radionuclide releases to the environment during the period 10,000 to 100,000 years after repository closure. For the preclosure analysis, fourteen performance measures (attributes) were developed; each is the measure of one objective, as shown in Table 3. Some of the measures are natural quantitative scales (e.g., dollars or lives lost); others are constructed scales (i.e., the labels on the scales are verbal descriptions of outcomes). In this analysis constructed scales were developed for four performance measures (aesthetic, biological, socioeconomic and archeological/historical/cultural impacts), with the associated scores assigned by a panel of experts whose determinations reflected complex "influence"

TABLE 2: KEY ELEMENTS OF THE DOE MULTIATTRIBUTE UTILITY ANALYSIS

1. Definition of the objectives:
   - distinction between postclosure and preclosure impacts
   - definition of performance measures
   - natural vs. constructed scales
2. Judgments of fact vs. judgments of value
3. Definition of the stakeholder groups
4. Importance given to postclosure over preclosure concerns:
   - identification of scenarios
   - assessment of scenario probabilities
5. Selection of the form of utility functions:
   - postclosure: single-attribute utility functions
   - preclosure: additive multiattribute utility functions
   - selection of postclosure/preclosure scaling factors
6. Sensitivity analyses of all important inputs
   - optimistic/pessimistic/base-case conditions
   - omission of certain objectives
   - alternative value trade-offs
   - assumption of identical attribute scores

TABLE 3: OBJECTIVES AND PERFORMANCE MEASURES USED IN PRECLOSURE ANALYSIS (from Table 4-1)

OBJECTIVE	PERFORMANCE MEASURE
	HEALTH-AND-SAFETY IMPACTS
1. Minimize worker health effects from radiation exposure at the repository	X_1: repository-worker radiological fatalities
2. Minimize public health effects from radiation exposure at the repository	X_2: public radiological fatalities from repository
3. Minimize worker health effects from nonradiological causes at the repository	X_3: repository-worker nonradiological fatalities
4. Minimize public health effects from nonradiological causes at the repository	X_4: public nonradiological fatalities from repository
5. Minimize worker health effects from radiation exposure in waste transportation	X_5: transportation-worker radiological fatalities
6. Minimize public health effects from radiation exposure in waste transportation	X_6: public radiological fatalities from transportation
7. Minimize worker health effects from nonradiological causes in waste transportation	X_7: transportation-worker nonradiological fatalities
8. Minimize public health effects from nonradiological causes in waste transportation	X_8: public nonradiological from transportation
	ENVIRONMENTAL IMPACTS
9. Minimize adverse aesthetic impacts	X_9: constructed scale
10. Minimize adverse archaeological,	X_{10}: constructed scale
11. Minimize adverse biological impacts	X_{11}: constructed scale
	SOCIOECONOMIC IMPACTS
12. Minimize adverse socioeconomic impacts	X_{12}: constructed scale
	ECONOMIC IMPACTS
13. Minimize repository costs	X_{13}: millions of dollars
14. Minimize waste-transportation costs	X_{14}: millions of dollars

diagrams indicating the inter-relationships between the most relevant factors.

Second, the analysis required two kinds of judgments: judgments of fact (for example, concerning the identification of the 14 preclosure performance measures) and judgments of value (for example, concerning the relative trade-offs between these measures). The two types of judgments are carefully differentiated in this analysis.

Third, the relevant stakeholder groups were identified. The DOE analysis is unusual here in that only a single stakeholder group, composed of four project managers employed by the DOE, was used as the source of all values information. This is in contrast to the more usual case where a MAUT analysis actively seeks to define the concept of "public" or "societal" values by defining and soliciting views from a number of different groups of people expected to hold reasonably coherent but differing views on the problem under consideration. Because the values and utilities needed as inputs to the MAUT model are derived from structured elicitation sessions with the stakeholders, the four DOE managers are implicitly assumed to be the source of all information about the nation's relevant values. In other words, they are viewed as having values representative of the nation.

Fourth, following the DOE guidelines, primary importance was placed on postclosure considerations over preclosure concerns. Scenarios that could potentially pose a risk to performance of the repository then were defined for the postclosure analysis of the five nominated sites. A panel of technical experts identified 12 disruptive conditions that could affect performance (i.e., adversely

affect human health and safety) and assigned probabilities to each; two additional scenarios also were included, a first giving the expected (baseline) conditions and a second showing a catch-all scenario that includes all "unexpected features".

Uncertainty was not included in the preclosure model, presumably because the relevant considerations (based on the 14 performance measures) were thought to be defined with more assurance.

Fifth, the form of the overall utility function was selected. For the postclosure case, simple expected utility models were used for both the 10,000 year and the 10,000 to 100,000 year radionuclide releases. For each, the technical experts judged the amount of release expected under each scenario (i.e., each future event description, such as an earthquake) for each site. The utility (i.e., the worth or value) of the releases was judged to be linear with the amount of release. The expected utility for each site was the sum of the utilities for each scenario, weighted by the probability, for that site. Finally, the two expected utilities were combined using relative weights provided by the four DOE project managers.

For the preclosure case, additive multiattribute utility models were developed. To do this, the technical experts assigned a score on each of the 14 attributes to each of the sites, based on their knowledge of the sites. The four DOE managers assessed the utilities for each scale and the relative weights for each scale. The overall preclosure utility for a

site was the sum of the 14 utilities for that site, each weighted by its assessed relative weight.

Finally, the analysts explored the relative weights to be assigned to the pre- and post-closure components, in order to add the two parts (each weighted accordingly) to arrive at a single overall utility for each site. In fact, no one pair of these weights was ever selected by the DOE project managers; it was found that over a reasonable range of possible values for these two weights, the sites were ranked the same and, thus, the analysts did not believe it was necessary to assess these weights precisely.

Sixth, sensitivity analyses of all important inputs were carried out. For example, site utilities were calculated under optimistic, pessimistic and base-case assumptions for preclosure impacts and for postclosure scores and disruption possibilities. Rankings under the sensitivity tests remained quite stable; for example, the Hanford site in all cases ranked fifth regardless of the relative weights that were assigned. Other sensitivity analyses systematically omitted certain of the objectives (e.g., costs) from the analysis, tested different value trade-offs between the two postclosure periods (by changing the weights) and between key attributes (e.g., by changing the ratios comparing the value of worker to public fatalities), and investigated how the results would vary when different sites were assumed to show identical impacts for one or more objectives. Over a wide range of sensitivity tests, the following rankings were fairly stable:

(a) for postclosure impacts: Davis Canyon, Richton Dome, Deaf Smith, Yucca Mountain, and Hanford;

(b) for preclosure impacts: Yucca Mountain, Richton Dome, Deaf Smith, Davis Canyon, and Hanford.

2.2 The choice of the performance criteria

Performance criteria include both postclosure and preclosure considerations. According to the DOE siting guidelines, preclosure criteria are segmented into three groups, in decreasing order of importance: radiological safety; environment (aesthetic, biological, and archaeological/historical/cultural impacts), socioeconomics and transportation; and cost (of siting, construction, operation and closure). Preclosure performance criteria provide direct measures of the associated objectives, to the extent that they directly assess changes in the key issues of concern (e.g., changes in statistical fatalities). Postclosure performance criteria, however, deal with so-called proxy measures--in this case, radionuclide emissions--in order to avoid estimation problems associated with the unusually long period of the analysis.

Scales used to score the performance criteria used in this analysis are of two distinct types. Most are based on "natural" scales, which implies that they enjoy common usage: costs, for example, are typically denoted in dollar terms. Several of the performance criteria, however, employ "constructed" scales that are specifically developed for the problem at hand because suitable natural scales do not exist. The utilities associated with these scales reflect the judgments of the four DOE assessors; the construction of the scales and the ratings of the sites on each scale reflect the judgments of the expert panels. In the MAUT analysis there are four constructed scales, covering

socioeconomic concerns and the three sources of "environmental" impacts. Because these four performance measures are the only ones on which Hanford receives other than a last-place rating under the base-case estimates (as shown in Table 4-6 of the MAUT report), we are encouraged to look more closely at the details of their construction.

The most interesting of the constructed (preclosure) scales is that concerning socioeconomic impacts. This scale emphasizes the effects of the in-migration of repository workers on the local economy, the public infrastructure (e.g., schools), local transportation and existing land uses. However, we believe that this view of socio-economic effects is narrow. In particular, a number of factors that could figure prominently as impacts of a nuclear waste repository are omitted. These include adverse impacts on tourism (e.g., campers in Washington state, or conventioners in Nevada) and adverse impacts on business development (e.g., due to fears associated with catastrophic risk potential). In a larger sense, the adverse socioeconomic impacts associated with possible stigma effects--whereby a fairly large geographic area around the repository site could be marked as undesirable--are entirely ignored in this analysis, despite their possible pre-eminence as a source of socioeconomic impacts (or as justification for claims of economic compensation). In addition, the creation of an additional 5000 jobs related to repository operation is viewed as disruptive to "existing business patterns," despite the fact that facility operation is

anticipated to continue over a 50-70 year period (which compares well to that of many conventional resource-based ventures).

With respect to the natural scales, the most obvious concern stems from the exclusion of non-fatal health effects. The analysis states that the omission is appropriate because of the expected correlation between morbidity and mortality effects. However, there are two reasons for concern. First, the number of worker fatalities varies between sites, with Hanford showing the most fatalities and Richton Dome the least (the base-case equivalent consequence impacts show a total of 120 deaths for Hanford and 64 for Richton Dome). If a correlation is assumed, those sites ranking worst in fatal effects also would rank worst in injuries. Furthermore, the Hanford site is arguably the most difficult of the five in which to construct a waste repository, and it has the highest transportation mileage and costs. We are not experts on technical issues concerning construction or transportation injuries. However, it seems likely that different types of sites would yield different death-to-injury ratios (although apparently the experts used here did not think so), which means that the omission of injuries would be inappropriate. Second, disregarding non-fatalities might have biased the value trade-off judgments made later (by the DOE assessors) because these judges were given profiles descriptively rich regarding, for example, archaeological, historical and cultural impacts but descriptively sparse in death information. Judges in such a situation may not have realized that "non-radiological worker deaths" really meant "non-radiological worker deaths and all the health and safety and injury claims that inevitably co-occur with x deaths". This

might lead to underestimating the relative value of the death-indexed performance measures.

The most interesting aspect of the postclosure scales is the relation of the single performance score (associated with health and safety effects) to the EPA guidelines. All five sites are shown to substantially exceed (i.e., are safer than required by) the EPA guidelines, and this appears to be used as the rationale for only partially considering (because there is "no basis for discrimination") several features of the accessible environment in the analysis despite the fact that substantial variation exists between sites--for example, with regard to the estimated travel time for groundwater, or differences in the pathways that releases could take (such as the quality and uses of the water body into which radionuclides could be released).

Nevertheless, the Hanford site is ranked approximately an order of magnitude worse than any of the other sites in terms of the base-case expected postclosure releases (Table 3-6). This is largely due to the greater uncertainty associated with long-term storage considerations at Hanford; specifically, Scenario 2 (Table 3-2), which covers "unexpected [adverse] features," was judged by the panel of experts to have a higher probability of occurrence at Hanford than at any of the other sites. In fact, the base-case probability assessed for this scenario is fully 2.4% (or 1 in 40), as shown in Table 3-3 of the MAUT report. Because the probability of this scenario (which, by definition, is an exhaustive "all other" category, as required for the probability assessments) is a direct reflection of scientific

uncertainty regarding what might go wrong, even a 1 in 40 chance appears uncomfortably high. Furthermore, as the weight placed on later releases is increased, the significance of releases from the Hanford site is increased (i.e., its expected utility decreases) because relatively more deaths are occurring over the 90,000 year late-postclosure period (see Figures 3-17 and 3-18 in the MAUT report).

Adding to this effect--and influencing the results in the same direction--is the fact that the base-case scorings given for the repository cost attribute are acknowledged to "underestimate the potential for higher costs" because total repository costs have both "increased significantly in recent years" and "more often than not" are underestimated in cost projections due to "delays, ...legal circumstances, and other unexpected conditions" (page G-60). This suggests that the estimates given for the cost of the Hanford site, which ranks considerably worse (i.e., is more expensive to construct) than any of the other four sites, are likely to be biased downwards and that utility measures reflecting the overall preclosure scores for the sites will show Hanford as a more attractive (or, equivalently, less unattractive) alternative than it actually is likely to be.

2.3 The assignment of weights

Weights are assigned to the sets of postclosure and preclosure objectives by the four selected DOE judges. These weights have a significant effect on the utility estimates of performance criteria, because they establish the trade-offs by which gains in

one attribute can be balanced directly against losses in another. As clearly explained in the analysis, the weights appear as scaling factors in the multiattribute utility function: a weight of one ($K_1 = 1$) placed on on-site worker fatalities due to radiation exposure, for example, means that the death has been assessed as equivalent to an additional cost of $1 million. For a constructed scale, such as that for socioeconomics, the weight of five ($K_{12} = 5$) means that the judges feel it is worth $5 million to reduce socioeconomic impacts by 1% (this is equivalent to saying it is worth $500 million to reduce socioeconomic impacts from the best to the worst level).

Questions regarding the appropriateness of eliciting values from a small panel of only four experts, all of whom are DOE employees, already have been raised. The Meek Report (page 20; (Subcommittee on General Oversight, Memorandum of October 21) adds the further insight that the combined weight given to the four socioeconomic and environmental attributes (in terms of dollar equivalents for best-to-worst changes in each scale) more than doubled in going from the August, 1985, MAUT feasibility study to the May, 1986, MAUT report. Because the DOE officials on the values panel probably had access to both the conclusions and the interpretation of the earlier study, it may be that their assessed weightings given in the later MAUT analysis unconsciously (and incorrectly) were influenced by the earlier findings; at a minimum, the shift in reported values is curious.

The next question is whether the preclosure and postclosure weights are sensible. In several cases, the answer appears to be that they are not, particularly with regard to some of the

preclosure weights. Two examples (as shown in Tables 4-7 and 4-9 of the MAUT report, and reproduced in Table 4 here) are worth noting. First, consider the trade-offs shown between the aesthetic scale (Table 5) and those depicting worker fatalities. The entire aesthetic scale (i.e., a movement from best to worst conditions) is valued at $100 million; thus, it may also equivalently be said that going from best to worst on the aesthetics scale is equal in value to losing 100 repository-worker lives. Going from one to two major aesthetic effects (level four to five) is equivalent to losing an additional 34 worker lives. This seems to us a very high price to pay for aesthetic impacts (such as "significant contrast with the visual setting") associated with a site deliberately chosen so as to be removed from major population centers. A lower weight for the aesthetics scale relative to the lives-lost scales would be necessary to reflect our values properly.

A second example concerns the trade-offs associated with the socioeconomics scale. Going from best to worst on this scale is equivalent to 500 worker lives. Going from a 0 to 2 rating on this scale (i.e., from essentially no adverse effects to "minor impacts" because "repository effects are somewhat incompatible with existing land uses") is judged to be worth fully 100 worker lives (or 25 public lives). Again, the point is that the weights associated with the socioeconomic scale appear to be unduly high, thereby decreasing the relative importance of all other attributes (including the health and safety impacts, considered by the DOE guidelines to be the most important) and (because it

TABLE 4: PRECLOSURE VALUE TRADEOFFS (from Table 4-7)

PERFORMANCE MEASURE	VALUE TRADEOFF
1. Repository-worker radiological fatalities	1.0
2. Public radiological fatalities from repository	4.0
3. Repository-worker nonradiological fatalities	1.0
4. Public nonradiological fatalities from repository	4.0
5. Transportation-worker radiological fatalities	1.0
6. Public radiological fatalities from transportation	4.0
7. Transportation-worker nonradiological fatalities	1.0
8. Public nonradiological fatalities from transportation	4.0
9. Aesthetic impacts (see Table 4-2)	1.0
10. Archaeological, etc. impacts (see Table 4-3)	0.2
11. Biological impacts (see Table 4-4)	0.3
12. Socioeconomic impacts (see Table 4-5)	5.0
13. Repository cost (millions of dollars)	1.0
14. Transportation cost (millions of dollars)	1.0

TABLE 5: PERFORMANCE MEASURES FOR ADVERSE AESTHETIC IMPACTS (from Table 4-2)

IMPACT LEVEL	AESTHETIC IMPACTS IN THE AFFECTED AREA
0	None
1	One minor effect
2	Two minor effects
3	Three minor effects
4	One major effect
5	Two major effects
6	Three major effects

Where a minor effect is defined as the following (e.g.):

- The locations of residences, ... major vistas ... or public highways are such that these points are on the project's line of sight but are within a visual setting that would not significantly contrast with the project;

- No key observation points or sensitive-receptor areas on the line of sight or within audible distance of the project attract many visitors.

Where a major effect is defined as the following (e.g.):

- The locations of residence,... major vistas ... or public highways are such that these points are on the project's line of sight and are within a visual setting that would significantly contrast with the project;

- Some key observation points or sensitive-receptor areas on the line of sight or within audible distance of the project attract many visitors.

scores relatively well on these scales) improving the overall score given to the Hanford site.

With this in mind, it appears that the sensitivity tests (shown on page 4-31 of the MAUT report) are done in the wrong direction. These analyses show changes in the preclosure scores of the five repository sites if socioeconomic impacts are doubled in importance (with K = 10 rather than 5). It makes more sense, and would lead to conclusions less favorable to the Hanford site, if sensitivity analyses were done showing socioeconomic impacts as less important; for example, using K = 2 or even 0.5.

Table 6 shows the result of one possible recalculation of the base-case site impacts, using revised weights for the socioeconomic and aesthetic scales but leaving all other non-cost performance scores the same (i.e., the transportation cost and repository cost performance measures are excluded). As shown in the table, we have changed the weight on the aesthetic scale from 1 to 0.5 (i.e., aesthetic considerations count for one-half as much in the overall analysis) and we have changed the weight on the socioeconomic scale from 5 to 1. These two changes--which seem to us to be very reasonable--significantly change the rankings of the five nominated repository sites: Hanford, which is ranked first in the non-cost base-case calculations presented in the MAUT report, is ranked fourth (i.e., next-to-worst) when these revised weights are employed.

TABLE 6: NON-COST BASE-CASE EQUIVALENT-CONSEQUENCE IMPACTS, SHOWING REVISED VALUE WEIGHTINGS

PERFORMANCE MEASURE	VALUE TRADE-OFF	RD	DS	DC	YM	HN
Worker and Public Fatalities		64.0	74.0	80.0	83.0	120.0
Aesthetic Impacts	1.0	33.0	33.0	100.0	33.0	3.0
Archaeological, etc. Impacts	0.2	1.2	2.4	11.2	4.6	1.2
Biological Impacts	0.3	4.5	3.6	8.7	3.0	3.6
Socioeconomic Impacts	5.0	100.0	80.0	100.0	30.0	15.0
Total Noncosts1		203.0	193.0	300.0	154.0	142.0
*Aesthetic Impacts	0.5	16.5	16.5	50.0	16.5	1.5
Archaeological, etc. Impacts	0.2	1.2	2.4	11.2	4.6	1.2
Biological Impacts	0.3	9.5	3.6	8.7	3.0	3.6
*Socioeconomic Impacts	1.0	20.0	16.0	20.0	6.0	3.0
Revised Total Noncosts2		107.0	113.0	170.0	114.0	129.0

Note: Sites with lower scores are preferred.

*Shows scales with weightings changed.

1Hanford shown to be best site.

2Hanford shown to be next-to-worst site.

2.4. The adequacy and implications of the sensitivity analysis

The purpose of sensitivity analysis is to expose the uncertainty associated with the evaluation of alternative actions. In the siting of nuclear waste repositories, there is uncertainty about (a) the performance of the sites on individual siting objectives, (b) the relative weighting of the different objectives, and (c) the forms of the utility functions. In generating the base-case ranking of the five sites, the report adopts a series of assumptions about site performance, attribute weights, and the utility functions. The report then uses sensitivity analysis to test the sensitivity of the site ranking to changes in these assumptions. The question addressed by sensitivity analysis is: Will the adoption of reasonable alternative assumptions change the ranking and/or the relative scoring of the five sites?

If the adoption of alternative assumptions causes significant changes in the scoring of sites, then the base-case computations do not accurately represent the relative merits of the sites. Because the base-case results reflect only one set of assumptions, they are correctly viewed as only one of a potentially large set of possible ranking results. Because there is uncertainty about which set of assumptions is correct, there is uncertainty about which set of ranking results would correspond most closely to real-world conditions were a site to be developed as the selected high-level waste repository. The sensitivity analysis translates uncertainty about the underlying

assumptions into uncertainty in the ranking results, thereby exposing the uncertainty associated with the evaluation process.

The report employs three techniques in its sensitivity analysis. First, in order to account for the uncertainty in the assignment of weights, total scores (the weighted sum of the attribute scores) are computed with different sets of weights. Changes in these weights could, in principle, change the ranking and/or the relative scores of the five potential sites. In fact, the changes in attribute weightings explored by the analysts did not affect either the ranking of sites or their relative scores.

The second sensitivity technique accounts for uncertainty in the scores of individual sites on the performance measures. There is uncertainty surrounding these scores for two reasons. First, there is limited knowledge about the features of the potential sites. Second, it is possible that disruptive events (acts of nature) may occur. To account for this type of uncertainty, DOE computes scores for three cases: in addition to the base-case rankings, which incorporate the "best estimates" of the features of the sites and the probabilities of disruptive events, there are rankings for a "pessimistic" case and an "optimistic" case. The pessimistic case computes site scores under the assumption that certain detrimental events (considered by analysts to be possible) will occur. Similarly, the optimistic case assumes that certain advantageous events will occur. Only the ranking of Yucca Mountain is affected by these changes; Hanford remains the lowest-ranked site.

The third technique accounts for uncertainty in the form of the utility function. There are two issues. First is the issue

of whether society is risk-averse, risk-neutral, or risk-prone. DOE handles this uncertainty about risk aversion by performing computations with different measures of risk aversion. Second (and closely related) is the issue of whether the marginal disutility of radionuclide releases increases or decreases as the volume of releases increases. The report handles this uncertainty about the disutility of releases by performing computations with different functional relationships. The sensitivity analysis suggests that the site rankings are not sensitive to reasonable changes in (a) the degree of risk aversion and (b) the marginal disutility of releases.

2.4.1 Preclosure Sensitivity Analysis

The conclusion from the preclosure sensitivity analysis is that the base-case ranking is a good representation of the relative merits of the five alternative sites. DOE tested the sensitivity of the ranking to changes in attribute weights, attribute scores, and the parameters of the utility function. For the changes made (those considered reasonable), there were no significant effects on either the ranking of sites (see Table 7) or on their relative scores.

DOE considered several alternatives to the base-case weighting scheme, none of which caused significant changes in the ranking results. The rankings were recomputed with changes in the weights assigned to several objectives, including fatalities (public fatalities, worker fatalities, radiological fatalities, and nonradiological fatalities), socioeconomic impacts, and costs. Under the alternative weighting schemes, the relative

TABLE 7: RANKINGS OF NOMINATED REPOSITORY SITES

==

1. For postclosure impacts:

   Davis Canyon
   Richton Dome
   Deaf Smith
   Yucca Mountain
   Hanford

2. For preclosure impacts:

   Yucca Mountain
   Richton Dome
   Deaf Smith
   Davis Canyon
   Hanford

3. Composite overall ranking:

   Yucca Mountain
   Richton Dome
   Deaf Smith
   Davis Canyon
   Hanford

4. Recommendation report ranking:

   Yucca Mountain
   Deaf Smith
   Hanford

==

scores of sites are similar to the relative scores obtained under the base-case assumptions. The reason for the small differences in relative scores is that costs (siting costs and transportation costs) dominate the non-cost criteria in the overall scoring of sites. Unless the weight assigned to costs is close to zero, changes in weighting schemes do not cause significant changes in the relative scores of the alternative sites.

As reported in the previous section, and summarized in Table 6, we have done a further sensitivity analysis based on the non-cost preclosure performance measures that radically alters the ranking of the nominated sites. Specifically, the original non-cost preclosure ranking was Hanford (best), Yucca Mountain, Deaf Smith, Richton Dome and Davis Canyon (worst). Our new sensitivity analysis yields the ranking Richton Dome (best), Deaf Smith, Yucca Mountain, Hanford and Davis Canyon (worst)--i.e., Hanford has moved from best to next-to-worst. We would normally advise that not too much attention be given to rankings based on only a subset of the attributes. However, the Recommendations report (discussed later in this report) places great emphasis on Hanford's superiority on just this non-cost component of the analysis, apparently without realizing the sensitivity of its conclusions to reasonable changes in the socioeconomic and aesthetics weightings.

The sensitivity analysis of attribute scores (the pessimist-optimist analysis) given in the MAUT report concludes that replacement of the base-case scores with either the optimistic scores or the pessimistic scores does not yield significant changes in the rankings. The scoring is insensitive to changes

in the degree of optimism because, for most of the siting guidelines, the five sites have similar ranges of scores for the three cases. In moving from the base-case to the pessimistic case, the attribute scores for all the sites decrease by similar percentages, meaning that the relative scores are not affected by changes in the degree of optimism.

Changes in the utility function do not generate significant changes in the preclosure ranking of sites. Changes in risk aversion do not change either the final ranking or the relative scores of the sites. The reason again is that, while most of the uncertainty is related to the performance of sites with respect to non-cost objectives, the cost attributes dominate the preclosure evaluation.

The general conclusion from the preclosure sensitivity analysis is that the base-case evaluation provides a robust representation of the average ranking (showing, in order, Yucca Mountain, Richton Dome, Deaf Smith, Davis Canyon, and Hanford) and likely performances at the five potential sites. The existing uncertainty (about attribute weights, scores, or the form of the utility function) does not really matter because the adoption of plausible alternatives to the base-case assumptions does not cause significant changes in the rankings or in the relative scores of the potential sites.

2.4.2 Postclosure Sensitivity Analysis

Given the postclosure sensitivity analysis, it appears that the base-case results do not fully represent the relative merits of the five potential sites. This is because changes in

some of the underlying assumptions cause significant changes in the relative scores of the five sites. Consequently, in order to get a complete picture of the likely performances of the sites, one must look beyond the base-case results. The sensitivity analysis suggests that there are circumstances under which Hanford is not only the lowest ranked site on postclosure guidelines, but that it is worse than all other sites by a wide margin.

The consideration of uncertainty in attribute scores causes significant changes in the relative scores of the five repository sites, although not in their ranks. Following the base-case postclosure analysis, the computed utility scores (shown in Table 3-6) are as follows: Davis Canyon, 99.99, Richton Dome, 99.99, Deaf Smith, 99.98, Yucca Mountain, 99.98, and Hanford, 99.76. However, even here the differences are not trivial when expressed in terms of expected (or "equivalent") releases per 10,000 years: Hanford's base-case expected releases (at 2.41 x 10-3) are fully an order of magnitude larger than the release-level expected at the next-worst sites (Yucca Mountain, at 2.35 x 10-4, and Deaf Smith, at 2.33 x 10-4). Given that this variation in releases occurs over a period of 10,000 years, it is not clear that the difference (e.g., in public fatalities or in public illnesses) is trivial despite the acceptance of all these levels under the EPA guidelines.

The adoption of the optimistic scenario causes trivial changes in scores, since most of the scores in the base-case are already near the maximum (100). In contrast, the relative scores

change dramatically under the pessimistic scenario: the scores of Davis Canyon, Richton Dome, and Deaf Smith are virtually unaffected, but Hanford's score drops to 63, and Yucca Mountain's score drops to 86. While Hanford's base-case score is over 99 percent of the score of Deaf Smith (the third-ranked), site, Hanford's score is only 64 percent as high as that of the third-ranked site under the pessimistic scenario.

Hanford's performance under the pessimistic scenario reflects the fact that the expected consequences of Hanford's pessimistic scenario are relatively large. The Hanford site has (a) relatively large probabilities of detrimental events (as noted in Section 2.3) and (b) relatively large consequences from such detrimental events. As a result, if one considers the full range of possible outcomes (optimistic, base case, and pessimistic) it is reasonable to conclude that, not only is Hanford ranked last, but it is worse than the other sites by a large margin. Because the base-case analysis ignores the possibility of the pessimistic outcome, reliance on this portion of the probability distribution incorrectly suggests that differences between the sites on postclosure objectives is small.

Changes in the relative weights do not generate significant changes in the ranking or scoring of the five sites. The relevant weights are those for early radiological releases (0-10,000 years) and later releases (10,000-100,000 years). Because the predicted early and later releases are highly correlated, changes in the weights have similar effects on each case. As a result, there are only small changes in relative scores, and no changes in rankings.

Changes in the utility functions also fail to generate significant changes in the ranking or scoring of the five repository sites. Under the base-case scenario, the volume of releases is relatively small. In addition, differences across sites in the volume of releases are relatively small. As a result, changes in the degree of risk aversion do not cause significant changes in the relative scores of the sites. Similarly, changes in the form of the utility function (changes in the relationship between the marginal disutility of releases and the volume of releases) do not cause significant changes in the ranking or the relative scores.

Our general conclusion of the postclosure sensitivity analysis is that reliance on the base-case ranking misrepresents the relative merits of the five potential sites. While the adoption of the pessimistic scenario does not change the postclosure ranking of the sites, it does change the relative scores of the sites: while the scores of the top three sites are virtually unchanged, Hanford's score decreases from 99.76 to about 63. The report minimizes these results and suggests that, based on the base-case postclosure analysis, there are trivial differences in the performance of sites on postclosure guidelines. This is simply wrong: there are circumstances (and these are circumstances considered possible, although unlikely, by the technical experts) under which Hanford's postclosure score is only about 64 percent of the score of the third-ranked site.

2.5 Independence

MAUT is a credible evaluation method only if the performance criteria are carefully chosen. In the MAUT models used here, a particular site's total score is the sum of the scores on the individual objectives. In order to prevent double-counting, the set of attributes must not be redundant, that is, the individual attributes must be independent of one another. If the attributes were in fact redundant, double-counting would occur and the ranking generated by MAUT would be impossible to interpret.

The performance measures used in the MAUT report were chosen to make independence among guidelines likely. On the postclosure side, the two attributes (early and later radionuclide releases) are logically non-overlapping. Similarly, the preclosure guidelines are designed in such a way that there are no obvious cases of redundancy (e.g., worker and public fatalities are examined separately; radiological and nonradiological fatalities are examined separately).

A quite different set of independence criteria is imposed by the additive models used in this MAUT. These criteria, tested by a number of formal techniques, refer to independence of values. For example, they investigate whether the value associated with decreasing the number of non-radiological worker deaths from 15 to 14 varies as, for example, the aesthetic impact of the repository changes. Negative answers to questions like this were used to establish independence. We could detect no implausibilities in these tests as performed.

3.0 Comments related to the Recommendations report

The DOE siting guidelines, issued in November 1984, represented the first step in the siting process outlined by the 1982 NWPA. The nomination of (at least) five sites for characterization was the second step. This section of our report comments on the third step of the siting process, which is the recommendation by the Secretary of Energy that characterization proceed at three of the nominated sites.

This recommendation is said to rely in part on the multiattribute analysis of the five sites, discussed in the previous section of this report. That analysis documents the overall utility ranking of the sites to be (1) Yucca Mountain, (2) Richton Dome, (3) Deaf Smith, (4) Davis Canyon, and (5) Hanford. As described by the DOE, the Secretary's recommendation also is based on a "portfolio analysis", which reflects the provisions in the siting guidelines for diversity of rock types and geohydrologic settings. The final recommendation of the Secretary, based primarily on these two factors, is that the following three sites be characterized: Yucca Mountain, Deaf Smith, and Hanford.

There is an obvious discrepancy in these two rankings, in that the sites recommended by the Secretary are ranked first, third, and fifth in the detailed MAUT analysis. In this section we examine the logic that we believe has led to this choice.

First, a comment on process. In contrast to multiattribute utility analysis, which has both a strong theoretical basis and a generally accepted protocol, portfolio analysis is not a technique per se but rather a generic term referring to

situations in which a decision maker must choose a number of options from some larger set rather than select a single option (Edwards and Newman, 1982). Thus, construction of an evaluative procedure that considers each cluster of options, as a group, is needed. Evaluations of single sites is relevant but not sufficient for an adequate portfolio choice.

In this case, the problem (as outlined in the NWPA) is to pick the best set of three sites from the group of five. The best site (Yucca Mountain) probably should be one of them, but the best three sites considered individually need not be the best three sites considered as a portfolio (e.g., due to the need for rock diversity, or considerations related to maximizing the amount of information gained from the characterization process). Because the three salt sites do not fit the NWPA requirements, there appear to be nine combinations of three sites that satisfy the portfolio requirements.

Two questions follow. First, could any of the concerns not considered until the portfolio-analysis stage--specifically, concerns related to rock diversity--have been handled directly as part of a multiattribute utility analysis? The answer, it seems clear, is yes. In fact, if diversity had been considered along with the set of other objectives, then trade-offs between diversity considerations and other objectives (e.g., non-worker fatalities, or cost) also could have been considered directly. Thus, diversity would have been weighted relative to other concerns rather than having been introduced as a paramount

consideration at the end of the MAUT analysis and, by implication, free to override the conclusions of that analysis.

In addition, diversity could have been more formally defined (as was done for the environmental and socio-economics scales) were it introduced as part of the MAUT. For example, if diversity is as important as the Recommendations report suggests, then perhaps other types of rock media (such as granitic formations) beyond those that remain in the final set of five sites should have been considered; the inclusion of diversity considerations as a concern at earlier stages of the site-selection process might have altered the final set of five nominated repository sites. The earlier introduction of diversity considerations also would have both improved consistency and saved costs: if diversity is defined as three different rock media, then it is clear that both Hanford and Yucca Mountain will be recommended from the list of five sites because the other three candidates are all salt sites. Thus (taking this narrow view), the MAUT analysis need have considered only the three salt sites and not the two non-salt sites.

Second, has the DOE (in the Recommendations report) carefully documented what was done to come up with the portfolio recommendations? The answer here is no, and in fact the contrast between the clarity of exposition evidenced in the MAUT analysis and the near-total lack of information given in the Recommendations volume is striking. This is particularly important because, as noted above, a properly conducted portfolio analysis would consider all possible groupings of three sites,

considered as if the three were a single option. This is a sufficiently complex process--requiring, for example, that one address the anticipated degree of independence of the information gained regarding site costs--that an entirely separate analysis needs to be conducted (i.e., in addition to the MAUT analysis that looks at the five sites individually). There is no evidence in the Recommendations report that the DOE actually conducted such an analysis.

The Recommendations report concludes that, because the repository and transportation cost performance measures "are the least important of all guideline subgroups" (page 5), it is reasonable to examine relative base-case impacts of the five sites in the absence of these considerations. This omission is not defensible for three reasons. First, the statement that transportation costs are in the least important guideline grouping is not correct; transportation is stated to be of second importance (see page G-58 of the MAUT analysis), just as are the environmental and socio-economic considerations. Second, these attributes of the alternative repository sites are integral to the analysis and cannot simply be excluded on the basis of hindsight. If transportation and siting costs are important, then the overall utility scores should reflect these concerns; if they are not, then the original MAUT analysis is incorrect and the information obtained from the values panel (e.g., concerning the relative weights of the 14 selected preclosure attributes) should be amended. Third, the emphasis in the Recommendations report on the 12 non-cost preclosure attributes to justify Hanford's

inclusion is unwarranted in view of the extreme sensitivity (noted above) of the relative rankings of sites to changes in the weightings of these attributes. We have shown that quite reasonable changes in two of these weights mean that Hanford ranks fourth, rather than first, on this subset of attributes.

4.0 Conclusions

We have presented our comments on two reports, published by the DOE, that are of critical importance to the siting of a high-level nuclear waste repository. The multiattribute utility analysis of the five nominated sites is extremely well done. Moreover, it carefully documents its assumptions and its conclusions, and straightforwardly tells the interested reader what has been done. Nevertheless, we believe that in several instances either the inputs to the analysis undertaken in the MAUT study or the interpretation of its results serve to unfairly prejudice the analysis in favor of the Hanford site. Our most important criticisms include the following points:

- the four DOE project managers who composed the values panel do not seem to us to represent sufficiently the diversity of relevant stakeholders
- the exclusion of non-fatal health effects may have biased the relative values expressed in the analysis
- the socio-economic impacts scale is, in our view, too limited in scope
- some of the value trade-offs are surprising, in particular, the value of deaths relative to the value of other non-cost attributes seems low to us and, in this

connection, some of the sensitivity analyses are incomplete

- reliance on the base-case postclosure rankings misrepresents the relative merits of the nominated sites because it ignores plausible event scenarios.

The Recommendations report, in contrast, fails to document either its assumptions or its conclusions. It purports to have conducted analyses of all relevant combinations of the possible sets of sites, taken three at a time, but does not inform the reader as to how this was done; furthermore, it makes several assumptions regarding the importance of the various attributes of the analysis that cannot be supported by the data provided in the multiattribute utility study. Whereas the MAUT report provides a sound basis on which to begin consideration of the Nuclear Waste Policy Act mandate, the Recommendations report is a travesty of nearly everything that decision-aiding methods stand for.

REFERENCES

Edwards, W. & Newman, J. R. (1982). Multiattribute Evaluation. Sage Publications. Beverly Hills.

Edwards, W. & von Winterfeldt, D. (1986) Public values in risk communication. Risk Analysis.

Keeney, R. L. & Raiffa, H. (1972). Decisions with Multiple Objectives: Preferences and Value Tradeoffs. John Wiley & Sons, New York.

Pratt, J. W., Raiffa, H. & Schlaifer, R. O. (1965). Introduction to Statistical Decision Theory. McGraw-Hill, New York.

Von Neumann, J. & Morgenstern, O. (1947). Theory of Games and Economic Behavior. Princeton University Press, Princeton, N.J.