3/24/80

DR. 460

LA-8052-MS

Informal Report

An Evaluation Procedure for Radioactive Waste Treatment Processes

University of California

LOS ALAMOS SCIENTIFIC LABORATORY
Post Office Box 1663 Los Alamos, New Mexico 87545

LA-8052-MS
Informal Report

UC-70
Issued: November 1979

An Evaluation Procedure for Radioactive Waste Treatment Processes

William J. Whitty

CONTENTS

AN EVALUATION PROCEDURE FOR RADIOACTIVE WASTE TREATMENT PROCESSES

by

William J. Whitty

ABSTRACT

An aspect of the Los Alamos Scientific Laboratory's nuclear waste management R&D programs has been to develop an evaluation procedure for radioactive waste treatment processes. This report describes our process evaluation method. Process worth is expressed as a numerical index called the Figure-of-Merit (FOM), which is computed using a hierarchical, linear, additive, scoring model with constant criteria weights and nonlinear value functions. A numerical example is used to demonstrate the procedure and to point out some of its strengths and weaknesses. Potential modifications and extensions are discussed, and an extensive reference list is included.

I. INTRODUCTION

The radioactively contaminated wastes generated by nuclear operations require special handling and storage or disposal to minimize the hazard potential in the environment. Department of Energy (DOE) waste management goals specify that the maximum volume reduction of solid wastes be made that is feasible in terms of existing technical, environmental, and economic considerations.

There is a wide variety of potential[1-6] waste treatment processes; a mixed blessing when one is faced with selecting a process to solve a particular waste handling problem. A number of treatment options may be feasible, but they may differ in ways that make a clear-cut selection difficult. Process selection is better understood and simplified when a formal procedure using well-defined criteria is applied.

Toward this end, systems evaluation methods were investigated as a part of the effort at the Los Alamos Scientific Laboratory (LASL) to evaluate alternative nuclear waste treatment processes. The evaluation procedure selected provides the means for structuring complex problems so that each important factor can be evaluated separately. The relative worths of each factor are then combined into a single number that represents the worth of the total system. Finally, the evaluator may find that by considering the problem a step at a time, as is required in following this procedure, he has a better understanding of the overall problem.

The evaluation procedure aids waste management personnel in the ranking and selection of treatment processes. It also provides a decision-making tool applicable at most levels of R&D management. The evaluation concepts presented here may have a wider range of application; when tailored to a specific need, the evaluation procedure can be applied to many situations involving the comparison and evaluation of complex alternatives.

II. METHOD OF EVALUATION

For an evaluation to be valid, every process in a particular evaluation must be one designed to treat waste of a known compositional range. The handling of the treated waste will be similarly prescribed.

After the problem has been identified, detailed evaluation criteria are defined which reflect the major areas of importance for this application. Next, the properties of the alternative processes are linked to the evaluation criteria by performance measures. These measures are expressed as numbers and are called levels of performance.

For example, if the main objective of a treatment process is volume reduction, then the effectiveness of the process in reducing the volume of waste would be a valid criterion. The performance measure selected could be the volume reduction factor, or the ratio of the original waste volume to the final volume. Thus a particular process under evaluation might have a volume reduction factor of 25 to 1, and this would be the level of performance.

For multiple-criteria decision problems, the performance measures frequently have dissimilar units and a problem arises in finding some means of relating the various levels of performance to a common unit of measure. In many cases, the actual worth of a particular level of performance is not the number itself, but a subjective value judgment made by the evaluator. Thus for different people the perceived value of a level of performance, called worth, may differ, and in addition the worth may vary depending on the specific application.

The relationships between levels of performance and performance worth are called value functions in decision problems where uncertainty is not taken into account. (They are called utility functions in decision problems under uncertainty.) The transformation of levels of performance by value functions to worth places all performance measures on a common unit of measure. Finally, the value functions are combined to produce a single scalar of overall process worth.

Process worth is called the Figure-of-Merit (FOM). The FOM is a function of the performance measures and is represented by

$$\text{FOM} = U(X_j \mid j = 1, \ldots, m), \tag{1}$$

where U is some function of the m performance measures, X_j. The numerical values of the performance measures, or levels of performance, can be displayed in vector form as

$$x = (x_1, x_2, \ldots, x_m). \tag{2}$$

The basic assumption is made that, for all practical purposes, the set of performance measures adequately satisfies the set of evaluation criteria and the vector x can be specified.

The procedure used to compute the FOM is to partition the m-component problem into m single-component problems, with each being easier to solve than the original. The solutions are then combined. The foundation of this procedure is a linear additive model of the form

$$\begin{aligned}\text{FOM} &= u(x_1, x_2, \ldots, x_m) \\ &= \sum_{j=1}^{m} w_j u_j(x_j),\end{aligned} \tag{3}$$

where $u_j(x_j)$ is the one-dimensional worth of a particular level of performance, x_j, and w_j is a positive scaling constant. The w_j's, called criteria weights, express the relative importance of the criteria to the overall FOM. The x_j's, $u_j(x_j)$'s, and w_j's are numbers. The choice of the range of the w_j's and $u_j(x_j)$'s is arbitrary, but it is convenient for them to be

$$0 \leq u_j(x_j) \leq 1, \tag{4}$$

$$0 \leq w_j < 1 \text{ for } j = 1, \ldots, m, \tag{5}$$

and

$$\sum_{j=1}^{m} w_j = 1. \tag{6}$$

The Constant-sum Method[7] was selected for calculating criteria weights. One hundred points are distributed between every pair of criteria and a set of weights is produced by appropriate arithmetic operations. The Churchman-Ackoff Method was used to check for consistency.[8] This method has the evaluator make comparisons between one criterion and several combined criteria. The comparisons are then used to adjust the weights, if necessary. This procedure is detailed in Appendix A.

The evaluation procedure is completed by determining the worth of specific levels of performance for each performance measure. These value functions are constructed by having each evaluator sketch a curve (on graph paper) that represents his adjudged performance worth for each of various levels of performance. (See Appendix B for details.) Then Eq. (3) is used to combine the weights and value functions to produce the FOM.

III. IMPLEMENTATION EXAMPLE

To test the method and to determine if LASL waste management personnel could effectively apply the concepts, a fictitious problem, based on the LASL waste management system, was defined. A group of waste management personnel developed a set of criteria and performance measures for the sample problem. Weights were assigned to the criteria and value functions determined for each performance measure. Finally, two processes were compared using the procedures and numerical values developed in the previous steps.

The sample problem consisted of selecting production-scale volume reduction processes for combustible, low-level radioactive waste generated at LASL. The waste considered for treatment arrives at an imaginary input boundary; at LASL, this is the process receiving point. Output is the primary waste product of the volume reduction process plus secondary waste generated by the process. Combustible secondary waste is returned to the process for reduction. Liquid components of the secondary waste stream are treated separately and the treatment residue is added to the volume of solid secondary waste to make up the total volume of treated waste. All other post-volume reduction processes, such as waste stabilization, are not included in this evaluation because the needed waste form criteria presently do not exist. Secondary components released to the environment are assumed to be held to within regulated limits.

The evaluation procedure is for production-scale processes and a base requirement is that a process must meet all applicable health and safety regulations before an evaluation is conducted. Process residue characteristics were not included in this evaluation exercise because of uncertainties in potential stabilization processes; and because only combustible, low-level, and transuranic wastes were considered, recovery of radionuclides was neglected. However, the scope of the evaluation can be expanded to include the form and composition of the process residue by the addition of appropriate criteria as this information becomes available.

For this application, the treatment process is defined to include the following subsystems.

1. On-site feed preparation (sorting, shredding, etc.).
2. Feed introduction (pneumatic conveyor, auger, ram feeder, etc.).
3. Waste reactor (oxidizer, etc.).
4. Product removal and packaging.
5. Effluent cleanup (liquid and off-gas cleaning trains used before release of effluents).

A. Criteria Selection and Performance Measures

The relationships among the criteria selected for this particular application are shown in Fig. 1. The definitions of the criteria and performance measures follow.

1. *Effectiveness* is the ability of the process to meet the volume reduction objective for the test waste matrix (see Appendix C). The performance measure for effectiveness is the Volume Reduction Factor, or the ratio of the original waste volume to the volume of treated waste.

2. *Flexibility* is the ability of the process to accept a wide range of waste compositions at different processing rates. The main criterion is composed of four subcriteria: Handling Normal Components, Solids Throughput, Handling Noncombustibles, and Liquid Throughput. Normal Components include polyvinyl chloride (PVC), paper and rags, polyethylene, rubber, and dense cellulosics. A composite formula relating percent of feed and feed rate is used as a performance measure (see Appendix D). The performance measure for Solids Throughput is the Rated Capacity (kg/h) of the unit. Noncombustible waste is included because a process should be able to cope with the inadvertent introduction of small amounts without serious problems; the performance measure is the Maximum Percentage of Noncombustibles permitted in the feed. Although this

evaluation is for solid waste treatment process, consideration has also been given to handling liquids. The performance measure for Liquid Throughput is the Maximum Injection Rate for organic liquids (see Appendix D).

3. *Availability* is the frequency with which the process operates satisfactorily. Inclusive in this consideration are the systems' maintainability and reliability. The direct measure is the percentage of uptime; however, it is expressed as the percentage of downtime as this is a more familiar measure. Downtime includes all time during which the equipment is not available for use.

4. *Operability* is the level of operational complexity in terms of equipment, controls, and manpower. The performance measure is the Minimum Number of Parameters that must be controlled for the process to operate and meet process and health and safety requirements. Controlled ancillary or extra variables, not necessary for process operation, are not included.

5. *Resource Use* involves energy, operational consumption of scarce strategic materials, water, and land requirements. Performance measures for these subcriteria are Categories of Use—extensive, very high, high, moderate, low, very low, none.

B. Criteria Weights

Ten LASL waste management personnel applied the Constant-sum Method to assign weights to the five main criteria. This was followed by the Churchman-Ackoff consistency check. Where the consistency check produced results that were incompatible with the weights, the weights were adjusted slightly until there was agreement. When both sets of results were in agreement, the evaluator was finished with that set of criteria. If his results did not agree, the weights and consistency check were reexamined until there was agreement.

These steps were applied to all sets of subcriteria. The weights for each criterion were then averaged to produce group weights. Group consensus and transitivity of group orderings were confirmed before the weights were averaged (see Appendix E). The final set of weights for the main criteria is presented in Table I. (All other weights are given in Appendix F.)

TABLE I

MAIN CRITERIA WEIGHTS

Criterion	Weight
Effectiveness	0.35
Flexibility	0.22
Availability	0.20
Operability	0.13
Resource Use	0.10
Total	1

C. Value Functions

Each evaluator was asked a series of questions designed to determine the shapes of the value functions. Then he was asked to sketch a curve that represented his value functions for each performance measure. Next, the individual value functions were averaged to produce group value functions for each performance measure. The evaluators were assembled and discussions were held about the coordinate pairs. Modifications were made when the majority felt they were necessary. The group value function for the Volume Reduction Factor is shown in Fig. 2. (All other value functions are given in Appendix G.) The value functions under Resource Use exhibited as much variation within themselves as among the other value functions under the same main criterion. Therefore, one value function was prepared by averaging the group value functions for Energy, Scarce Strategic Materials, Water, and Land.

D. Evaluating Two Hypothetical Alternatives

Incineration and mechanical compaction are two volume reduction alternatives. Two hypothetical systems are used here to exemplify the evaluation procedure.

A hypothetical example of each processes' performance level estimates is shown in Table II. Table III shows the worth of each separate performance level. Table IV gives a scale factor for each performance measure, which is determined by multiplying all weights down a path in the criteria tree (Fig. 1). For

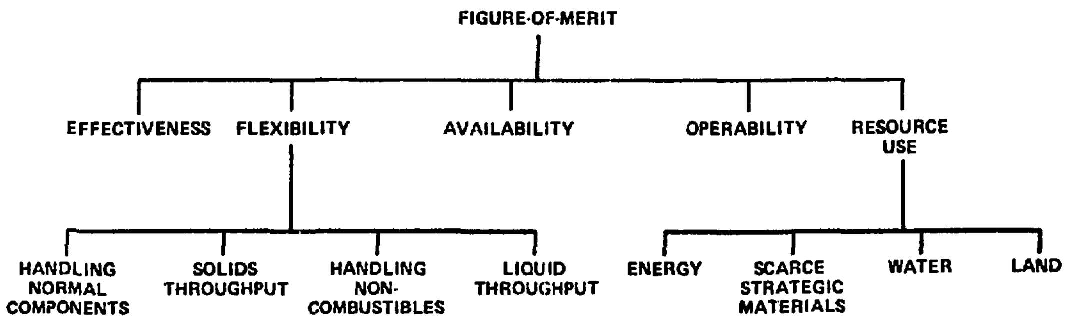

Fig.1.

Criteria structure.

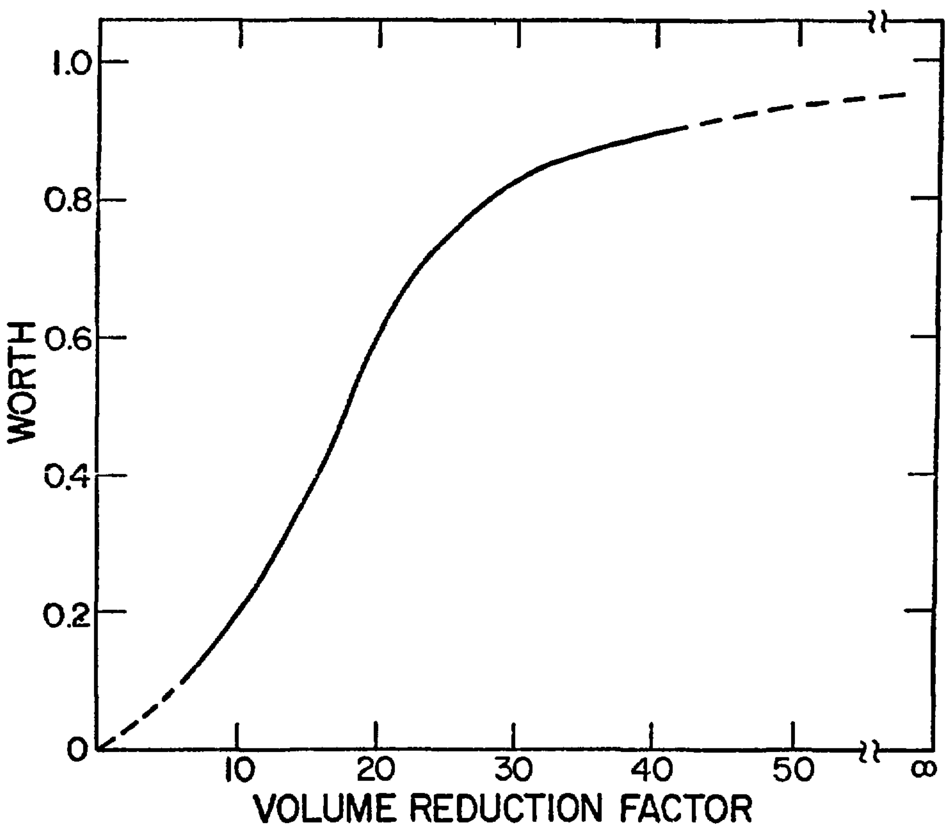

Fig. 2.

Group value function for the Volume Reduction Factor.

TABLE II

PERFORMANCE MEASURE ESTIMATES FOR TWO HYPOTHETICAL SYSTEMS

	Incinerator	Compactor
Volume reduction factor	25:1	5:1
Percent downtime	30	10
Minimum number of parameters	20	6
Normal components (R_iF_i) (see Appendix D)		
PVC	1	1
Paper & rags	2	1
Rubber	1	1
Polyethylene	1	1
Dense cellulosics	0.8	1
Rated capacity (kg/h)	50	200
Maximum percent of noncombustibles	5	10
Liquid injection rate	0	0
Energy use	moderate	very low
Use of scarce strategic materials	none	none
Water use	low	very low
Land use	very low	very low

TABLE III

WORTH OF PERFORMANCE FOR TWO HYPOTHETICAL SYSTEMS

	Incinerator	Compactor
Volume reduction factor	0.74	0.08
Percent downtime	0.80	0.95
Minimum number of parameters	0.84	0.99
Normal components (R_iF_i) (see Appendix D)		
PVC	0.60	0.60
Paper & rags	1	0.60
Rubber	0.60	0.60
Polyethylene	0.60	0.60
Dense cellulosics	0.50	0.60
Rated capacity (kg/h)	0.95	0.99
Maximum percent of noncombustibles	0.46	0.90
Liquid injection rate	0	0
Energy use	0.67	0.92
Use of scarce strategic materials	1	1
Water use	0.82	0.92
Land use	0.92	0.92

example, the path from Flexibility to Handling Normal Components has weights of 0.22 (Table I) and 0.46 (Table F-I), respectively, and the scale factor, or final weight, for Handling Normal Components is 0.10 (Table IV).

Table V shows the worth of performance (Table III) multipled by the corresponding scale factor (Table IV). The final line of Table V shows the FOMs for the hypothetical incinerator and compactor. Thus the significant difference between these two systems is the amount of volume reduction.

IV. DISCUSSION

LASL waste management personnel found this evaluation procedure to be a useful way to quantify subjective evaluations. They felt that the method for computing the FOM was reasonable and straightforward. Selecting and defining criteria and specifying performance measures required the most time. All members of the evaluation group were satisfied with the final definitions and found no difficulty in producing the weights and value functions. Most felt that a primary benefit of the procedure was that it helped them structure the complex evaluation in a systematic way.

Before selecting any procedure for evaluating system worth, it is necessary to specify the reason for the evaluation, the relationships among criteria, and the availability of related data.[9] For this application, operational considerations dictated that the evaluation procedure be relatively easy to understand and explain, easy to implement, economical, reasonably stable against small errors in estimates, and easy to modify for slightly different conditions or applications. The procedure also had to be independent of the number of alternatives.

A linear additive model of the type described in Section II satisfies these requirements. A more sophisticated model probably would not significantly increase the accuracy of the FOM,[10] and would be more difficult to apply. User acceptability was also a consideration in selecting the methods for determining weights and value functions.

The most important assumption related to Eq. (3) was that the performance measures be value independent, so that addition could be used for combining the worth of separate performance levels.[11-21] Value independence means that an evaluator's preferences for a specific level of one performance measure do not influence his preferences for levels of any other measure. However, value independence is difficult to maintain in actual situations. Yntema and Torgerson[10] found that whereas dependencies may not significantly affect the ordering of alternatives, they may distort the resulting FOMs. In some cases measures can be dependent, but by carefully combining some and/or eliminating others, a minimum distortion can be maintained. Miller[20] and Gustafson et al.[16] advocate eliminating performance measures where substantial interactions occur. Care in defining criteria and performance measures can eliminate many problems (see, for example, Fishburn[22]). Precautions taken when the value functions are being drawn can also do much to minimize the effects of interactions (see Appendix B). (See also Fishburn,[14,15,23] and Keeney.[18,19])

Additive models are not as sensitive to errors in component estimates as multiplicative ones; the former tend to smooth out errors whereas the latter tend to magnify them. This study used a hierarchical additive model where weights were multiplied at each level of the hierarchy (see Fig. 1 and Table IV). The resulting error in weights increases as the number of levels in the hierarchy increases.[16] For small weights, errors are not significant. (Extremely small weights may even restrict the contribution of the associated performance measures, and the argument has been made that they should be eliminated. For example, Tables IV and V suggest that Land Use might be excluded.)

The major areas for sensitivity analysis are related to highly weighted performance measures and the effects of errors in the construction of value functions. The FOM can be significantly affected when the performance levels occur at, or close to, large slopes or slope changes (see Appendix H).

V. EXTENSIONS AND REVISIONS

For any particular form of Eq. (3), all alternatives that have identical FOMs are considered equal. Final selection of a process can be made by considering the FOMs together with total process costs and intangible or political considerations. In the presence of costs and FOMs alone, the choice might be made by forming FOM-to-cost ratios and selecting the process with the largest ratio.[16,20,21,24,25] This is the familiar cost-effectiveness formulation

TABLE IV

SCALE FACTORS (WEIGHTS) FOR PERFORMANCE MEASURES

Volume reduction factor	0.35
Percent downtime	0.20
Minimum number of parameters	0.13
Normal components (see Appendix D)	0.10
Rated capacity (kg/h)	0.06
Maximum percent of noncombustibles	0.03
Liquid injection rate	0.03
Energy use	0.04
Use of scarce strategic materials	0.03
Water use	0.02
Land use	0.01
Total	1

TABLE V

WEIGHTED WORTH OF PERFORMANCE FOR TWO HYPOTHETICAL SYSTEMS

	Incinerator	Compactor
Volume reduction factor	0.26	0.03
Percent downtime	0.16	0.19
Minimum number of parameters	0.11	0.13
Normal components (see Appendix D)	0.07	0.06
Rated capacity (kg/h)	0.06	0.06
Maximum percent of noncombustibles	0.01	0.03
Energy use	0.03	0.04
Use of scarce strategic materials	0.03	0.03
Water use	0.02	0.02
Land use	0.01	0.01
Figure-of-Merit	0.76	0.60

where the FOM is the index of overall process effectiveness. (This is distinct from the criterion effectiveness.) This formulation assumes that the FOM is measured on a ratio scale and the relationship between the FOMs and costs is linear.

The sample analysis assumed that the performance levels were known and that the only errors were in the measurement of the levels, the estimates of the weights, and the estimates of the value functions. Disregarding all other factors, if costs were deterministic, then processes with equal cost-effectiveness ratios would be equivalent alternatives. At this stage, distortion of the FOMs due to dependencies can be serious because it can change the alternative selected. More on cost-effectiveness can be found in articles by Fishburn,[23] Dean and Roepcke,[24] and Seiler.[25]

At very early stages of an R&D selection effort, cost could be included in the set of evaluation criteria[16,22,27-30] to give a preliminary evaluation. As more reliable performance and cost estimates are obtained, cost should be removed from the criteria set and used as a resource constraint. As the R&D effort evolves, the evaluation model should become more sophisticated, commensurate with reliability of the data. Remember that the selection model will not only change from application to application, but will also evolve for a specific application as the R&D effort progresses.

The degree of refinement of the evaluation procedure will depend on the particular objective. For example, if the FOMs were average values, then the next step might be to retain the individual weights and value functions and use simulation to construct FOM distributions.[27] Averaging weights and value functions is valid only when the interests and backgrounds of the evaluators are similar. Even then, individual differences are masked.[27,31,32] Sometimes two distinct groups will be involved in the selection of an alternative. For example, a set of alternative projects is evaluated by R&D engineers, but the final choice for a full scale R&D effort is made by a management team. This situation can be handled in several ways. All individuals could take part in the exercise and form FOM distributions rather than averages. Another prospect would be for the management personnel to provide the criteria weights and the R&D staff to specify the value functions.[33] As with all other aspects of the evaluation method, the path followed depends on the application.

The next step to be taken to extend the evaluation method would be to include the uncertainty of the performance level estimates. These variations, as well as the variations in weights and value functions, could be used to compute expected values and standard deviations for the FOMs.[34] Alternatively, the individual estimates could be preserved and used by themselves or with probability figures to construct FOM distributions.

VI. CONCLUSIONS

The emphasis of this document is to outline the procedure developed and used for evaluating radioactive waste treatment processes at LASL. The use of the procedure, as well as the degree of refinement or extension; depends on the particular objective. Essentially, the concepts described here have a much wider sphere of application than the example presented. Keeney,[35] states, "One could speculate ... that open-mindedness and willingness to think hard about consequences are more important for correctly assessing preferences than any formal quantitative education."

The evaluation procedure, if not completely satisfactory, at least forces the evaluator to clarify and quantify factors important to the overall objective. One of the primary benefits of the procedure is that a systematic approach was taken to quantify the problem, which in turn allowed the evaluators to communicate the rationale used. This may be the first step in a series of discussions regarding the evaluation procedure and disagreements pertaining to it.

APPENDIX A

CALCULATING CRITERIA WEIGHTS

There are several methods for transforming experts' judgments into relative weights. Eckenrode[36] and Bartlett et al.[37] found no significant differences among the techniques they investigated. Since no method is clearly superior, the preferred method in any application depends on the intended use of the scale (ordinal, interval, or ratio level of measurement[38]), the time required to use it, the subjects' mental attitudes, their understanding of the overall problem, their perception of the instructions for weighting the criteria, and their understanding of the criteria definitions.

Five methods for producing an interval scale were presented to LASL waste management personnel to select one or two for further investigation. Only one method could be eliminated. Therefore, a small experiment was conducted to see if one method would emerge with a broad user appeal. Three methods and modifications are described here along with an experiment to select the most acceptable of these for use by LASL waste management personnel. The weights for the example implementation are also presented.

I. WEIGHTING METHODS

The Churchman-Ackoff Method[8] of successive comparisons and the modifications to this method are common techniques for determining weights. The criteria are ordered and the most important criterion is arbitrarily assigned the number one. All other criteria are compared to the most important criterion and assigned numbers relative to it. For example, there are four criteria, C_1, C_2, C_3, C_4, listed in order of importance. A hypothetical evaluator assigns tentative importance numbers, s, to the criteria as follows:

Criterion	Relative Importance(s)
C_1	1
C_2	0.7
C_3	0.4
C_4	0.2

Next, successive comparisons are made with a criterion and combined criteria to readjust the relative importance numbers. Our fictitious evaluator feels that C_1 is less important than all other criteria combined; therefore, no modification need be made ($s_1 < s_2 + s_3 + s_4$). Then he compares C_1 with the combination C_2 and C_3 and feels that C_1 is of equal importance to the pair, so s_1 is changed to 1.1. The first comparison is reexamined and is found to still hold ($1.1 < 1.3$). On the next comparison, C_2 is felt to be less important than C_3 and C_4, so s_2 is changed to 0.5. On examination of the previous comparisons, it can be seen that s_1 needs to be changed to 0.9. The new set of importance numbers are 0.9, 0.5, 0.4, 0.2. All importance numbers are normalized by their sum to produce the final weights 0.45, 0.25, 0.20, 0.10. Turban and Metersky[21] modified this procedure by asking judges to allocate 100 points among all criteria.

Klee[17] assigns a ratio of importance to pairs of overlapping criteria. The ratio assigned to a pair represents the importance of the first criterion in the pair to the importance of the second criterion (also see Dooley[12]). This procedure is followed until each criterion has been compared to the criterion immediately below it on the list. For example, C_1 is considered 1.8 times as important as C_2, C_2 is thought to be 1.3 times as important as C_3, and C_3, 2.0 times as important as C_4. This gives the weights 0.46, 0.25, 0.19, and 0.10. Litchfield et al.[27] modified this procedure by having subjects assign a number to the first criterion in a pair and then assign a number to the second criterion that represents its relative importance to the first criterion. The two numbers for each overlapping pair are used to form the ratios that Klee assigns directly. For example, C_1 is assigned 100 points and C_2 is assigned 55 points. Next, C_2 is assigned 100 points and the evaluator assigns 80 points to C_3. Finally, 100 points are assigned to C_3 and C_4 is thought to be worth 50 points in relation to it. The final weights are the same as before.

Weights are calculated by the Constant-sum Method[7] in which an evaluator distributes 100 points between every pair of criteria. For example, our fictitious evaluator assigns 60 points to C_1 and 40 to C_2. Then 70 to C_1 and 30 to C_3, 85 to C_1 and 15 to C_4. Likewise, 55 to C_2 and 45 to C_3, 70 to C_2 and 30 to C_4. Finally, he allocates 70 points to C_3 and 30 to C_4. The final weights are as before.

The Constant-sum Method requires m(m-1)/2 evaluations, whereas the other procedures mentioned above require one fewer evaluations (m-1) than criteria (m). This procedure produces more estimates than necessary since only m-1 are needed to construct the weights. However, all estimates are used to calculate composite scale values.

With the exception of the Constant-sum Method, each of the weighting techniques can be summarized as follows.

1. If a number is assigned to each criterion in a paired comparison evaluation, form a ratio, $r_{j+1,j}$, from the numbers, which represents the relative importance of C_{j+1} to C_j, for j = 1, ..., m-1.

2. If the ratios are assigned directly, $r_{j+1,j}$ is as defined in item 1 above.

3. Define a set of numbers $\{s_j | j = 1, ..., m\}$, then obtain their values from the recursion relationship

$$s_{j+1} = s_j r_{j+1,j}, \quad s_1 = 1, \; j = 1, \ldots, m-1. \qquad \text{(A-1)}$$

The Churchman-Ackoff Method sets s_1 equal to one and has the evaluator specify the other s_j's directly. Equation (A-2) is used to normalize the weights.

$$w_j = \frac{s_j}{\sum s_k}, \; j, k = 1, \ldots, m. \qquad \text{(A-2)}$$

The Constant-sum Method uses m(m-1)/2 ratios. If p_{jk} is the number of points awarded to C_j when compared to C_k, the ratio of importance of C_j to C_k is given by

$$r_{jk} = \frac{p_{jk}}{p_{kj}}, \; j, k = 1, \ldots, m, \quad j \neq k. \qquad \text{(A-3)}$$

By not requiring the criteria to be ordered, Eq. (A-1) shows that r_{jk} is an estimate of the ratio s_j/s_k. In addition to the direct estimate of r_{jk}, there are m-2 indirect estimates of s_j/s_k. The final step involves calculating some composite estimate of r_{jk}. Torgerson[7] suggests using the geometric mean of the estimates. The geometric mean possesses the property $r_{jk} = 1/r_{kj}$, which is consistent with Eq. (A-1) and Eq. (A-3).

Torgerson's least squares formulation for estimating the scale is given by

$$s_j = \text{antilog}\left(\frac{1}{m} \sum_{\substack{k=1 \\ k \neq j}}^{m} z_{jk}\right), \quad j = 1, \ldots, m \qquad \text{(A-4)}$$

where z_{jk} is the log of r_{jk}. The s_j's have a geometric mean of one and are converted to the normalized form by Eq. (A-2).

All of the above methods assume a constant weight over all performance levels. This should not be a problem if the performance measure ranges are small. In addition, these methods force transitive orderings. However, since the Constant-sum Method uses comparisons of all pairs, it can be used as a check on transitivity before computing the weights. The extra information on transitivity and the additional criteria estimates are at the expense of increased administration time.

II. SELECTION OF A WEIGHTING PROCEDURE

Turban and Metersky's constant-sum modification of the Churchman-Ackoff procedure was considered for this application, but it was, in practice, confusing to users and was dropped. Other methods considered were the ratio technique used by Klee, the Constant-sum Method, and the Churchman-Ackoff Method. Some inconsistencies appear in the literature regarding the administration time and simplicity of the Churchman-Ackoff Method,[28,38,39] especially regarding the consistency checking procedure. However, no matter which method is used, some type of consistency check should be made. Therefore, it was decided to investigate the Churchman-Ackoff Method temporarily ignoring the consistency checking aspect. After some discussion, several engineers favored Litchfield's[27] modification of Klee's method. Further, they suggested that one of the ratings in each pair be set at 100.

Because there was some difference of opinion on which method to use, it was decided to conduct a small experiment to see if one method would emerge with a broad user appeal. Initially, four process engineers ranked the five main criteria from randomized lists. Group consensus was obtained by defining the ranking for the group to be the ordering given by the sum of the ranks for each criterion, after Kendall.[40] There was general agreement that the ranking was realistic, and the Kendall Coefficient of Concordance was significant at the 0.01 level.[39,40] This means that the probability of stating that the coefficient is not due to chance, when in fact it is, is 0.01.

Because there was group agreement, the next step was to have each participant use each method once. The weighting sequences were randomized in a Latin Square design format[41] to separate possible differences in attitudes toward the different weighting methods as a function of the time of use and the order of use by subjects. The primary aim of the experiment was to determine the subjects' attitudes toward the weighting methods; a secondary aim was to see if the weights were similar from method to method.

As the experiment progressed over a 6-week period, the subjects tended to vary about the order of the criteria. While the experiment was being conducted, criteria definitions and performance measures were revised, although principally in wording rather than substance. It was clear that the subjects were changing their minds because of increased exposure to the overall project as a result of the group discussions. After completing the original ranking and using two methods, each subject appeared to have selected an ordering. The Constant-sum Method was used to verify that subjects were behaving transitively, and the Kendall Coefficient of Concordance was used to determine the degree of agreement among the subjects' rankings. There was good agreement, and all coefficients were significant at the 0.01 level except one that was significant at the 0.05 level. Average weights were similar among methods.

Comments were solicited from each subject after his use of each weighting method. All subjects liked the Constant-sum and Churchman-Ackoff Methods; the ratio (Klee) and rating (Litchfield) schemes met with equal favor and disfavor. The Constant-sum Method required more time (9 min) than the Churchman-Ackoff Method (3 min), but the difference would not be a hindrance in an operational situation. The short time required for all methods seemed to indicate that not enough attention was being paid to the task; however, this premise went unsubstantiated because subjects indicated, at each stage, that they were satisfied with the ranks and weights. Several weeks later the subjects completed the Churchman-Ackoff consistency checking procedure. All found the procedure easy to follow and were able to complete it within 15 min.

The Constant-sum Method seemed to produce higher weights for the most important criterion. If Gustafson et al.[18] are correct in assuming that people are conservative in assigning weights, then the most important criterion would receive less weight and the Constant-sum Method might have correctly estimated the true weights. Although the Churchman-Ackoff estimated weights (without the consistency checking procedure) were only first approximations to the true weights, it was not thought that the consistency check would significantly change the weights. (The consistency check applied at the end of the experiment was not applicable because of the change in orderings that occurred.) Three of the original subjects, and another individual who was not previously involved with the project, completed both methods and the consistency check in an attempt to resolve this discrepancy. In all four cases the Constant-sum Method produced closer estimates to the final weights. The final weights were calculated by applying the consistency checking procedure to both methods and were identical or nearly identical for all four individuals. The modifications to the original weights were as small as possible within the constraints set by the consistency check. When several sets of weights were possible, the one selected was the set where the sums of squares of the differences of the original and modified weights were minimal.

Although both methods seemed to produce the same weights when used with the consistency checking procedure, it was decided to use the Constant-sum Method because of its ability to produce multiple estimates, to check for transitivity, and produce a ratio scale.[7] The Churchman-Ackoff Method was applied later mainly to check for consistency. However, both sets of weights were used to resolve differences when they occurred.

III. DETERMINATION OF EXAMPLE WEIGHTS

Before the weights could be calculated, it was necessary to expose the individuals who had not yet been involved in the project to the Constant-sum and Churchman-Ackoff Methods so that they all would be at approximately the same place on a learning curve. After they were well acquainted with the techniques to be employed, they ranked the main criteria. Shortly thereafter, they used the Constant-sum Method to assign weights. At this stage, the ranking provided by the weights was compared to the original ranking. Each evaluator was checked for transitive ordering by examining the constant-sum pairs. Both rankings were in agreement and there were no intransitivities.

A few days later, the Churchman-Ackoff Method was used to calculate the weights. After another few days the consistency check was used. When the consistency checking procedure produced results that were not compatible with the weights, the weights were adjusted slightly until there was agreement. Each individual was presented with his modified weights and the outcome of the consistency check. If he agreed with these results, then he was finished with that set of criteria. If he did not agree, the weights and the consistency check were reexamined until there was agreement.

Each of these steps was carried out for all sets of criteria. At no time did any individual know the results produced by any other individual. Before the weights could be averaged to produce a set of group weights for each set of criteria, it was necessary to check for group consistency in rankings and for group intransitivity. There were no intransitivities and there was group agreement on the orderings of all four sets. Group consensus was checked by the Kendall Coefficient of Concordance. All coefficients were significant at the 0.05 level or better (see Appendix E). Remember, the significance level is a probability and not the value of the coefficient.

When the average weights were calculated, one set of criteria had a slightly different ordering than the group ordering by ranks. The group discussed this difference, and two individuals claimed that they misunderstood the criteria in question. They agreed with their original ranking but changed their weights. This produced an average set of weights which had the same rank order as the group ordering by ranks.

APPENDIX B

PERFORMANCE WORTH-VALUE FUNCTIONS

In deterministic situations, performance measures are related to a common unit of measure by value functions. The actual transformation is accomplished by questioning individuals concerned with the specific application. Miller[20] calls this transformation a scoring function and defines it as, "a mathematical rule which assigns a unique worth score, in points, to every possible value of some physical performance measure."

Value functions are unique up to a monotonic linear transformation. If $v_j(X_j)$ is a value function for the performance measure X_j, then so is $u_j(X_j)$, if for any positive constant, a

$$u_j(X_j) = av_j(X_j) + b_j. \quad \text{(B-1)}$$

This means that value functions can be represented on an interval scale. An interval scale has the properties that adjacent numbers are separated by a constant arbitrary unit of measurement and the ratio of any two intervals is independent of the unit of measurement and an arbitrary zero point. A preliminary condition is that objects measured on the scale can be ordered. Because there is no natural origin, the origin may be chosen in any convenient manner. If the origin is set to zero, the least preferred or worst level of X_j is set to zero, and the best level is set to one, then

$$0 \leq u_j(X_j) \leq 1. \quad \text{(B-2)}$$

The next step requires that, given the end points or range of X_j, enough coordinates be estimated so that a graph can be constructed. If there is a natural zero a ratio scale may be formed, i.e., all b_j's equal zero. This can be justified if worth is considered to be the fractional attainment of an objective.

Fishburn[42] presents some methods for estimating the functions, as well as methods for estimating weights. One of these has the subject estimate the worth of the level of performance midway between the best and worst levels. This point now separates two intervals and midpoints are estimated for these in a similar manner. This continues until there are enough points to plot a graph. Huber[43] and Dee et al.[44] outline approaches for constructing value functions, and detailed step-by-step instructions are presented by Miller.[20] Miller suggests that a rough sketch of the function be drawn, based on questions posed to evaluators, and then estimates of performance levels be solicited for 10% intervals on the worth scale. Litchfield et al.[27] have subjects estimate the worth of four levels of performance.

Trial value functions were constructed for three performance measures following the general outline of Miller's procedure. The best level of performance was assigned a worth of one and the worst level was assigned zero. If a performance measure has no natural or logical upper or lower bound, then a practical operations bound must be specified. Miller suggests using 50% or more in excess of the anticipated maximum performance level as an operational upper bound. Huber suggests using the practical bound, or a number just beyond it, as the natural bound. Two of the three performance measures in this trial possessed no logical bound. When the bounds were set at about 50% more than the anticipated maximum, the size of the graph became very large along the abscissa and the form of the curves and inflection points became distorted. This bothered four individuals who were trying to estimate interior points. Miller's suggestion of estimating performance for every 10% change in worth, estimating midpoints as Fishburn suggests, and Litchfield's fixed levels were also unworkable in this context. What seemed to be happening was that subjects had a rough idea of what the shape of a curve should look like, picked some reference point that had some individual intuitive appeal, and then used this point and the end points to complete the graph. When the operational upper bound was placed at 50% or so greater than the anticipated maximum, they had trouble estimating points in between because of the scale extension. When points were estimated, they were so far from the bound selected it was almost impossible to sketch a curve that had meaning beyond the last point estimated. Therefore, the upper bounds were set much closer to the practical upper bound, as Huber recommends. The subjects generally completed the portions of the curves near the boundaries first, then they picked interior points with no discernible pattern.

The results of this exercise at constructing value functions were very similar for all four individuals. Each subject was satisfied that his curve was realistic and that those of the others were only slightly different from his own. Consistency was evaluated by asking questions similar to, "How much better is level i than level j"? Only slight modifications of the graphs were necessary.

When a subject is completing a graph it is important that he consider only the performance measure at hand. The most desirable level for the other measures is commonly used.[27] Miller recommends that the value functions be constructed as if there were no other performance measures. If two or more performance measures are varied at the same time, the independence of the value functions will be suspect. Therefore, subjects were instructed to consider all other performance measures to be at their best possible levels. After each graph was completed, and before checking consistency, the subject was asked if he had considered the level of any other measure while he was estimating the value function. All subjects said that they had not.

The test for scoring performance resulted in the following prescription for constructing value functions in the local situation.

1. Present each performance measure to the evaluator individually, without him knowing the results of any other evaluator.

2. Advise him to consider only the performance measure under investigation and assume that all other performance measures are at their best attainable levels.

3. Follow the first stage of Miller's procedure to determine the general shape of the curve.

4. If the evaluators disagree on the boundaries or shape of the curve call them together and resolve the problem.

5. Plot the level of performance on the horizontal axis and the worth of performance on the vertical axis. Mark the upper and lower bounds and their appropriate worths. If there is no natural upper bound, indicate some practical upper bound and assign a number just beyond it as a practical operational natural bound. Subdivide the vertical axis in tenths and the horizontal axis in some convenient equally spaced unit.

6. Have the user fill in the interior of the graph by whatever means that allows him to feel comfortable.

7. When the graph is completed, get the user's assurance that no other performance measures were considered. If they were, the graph should be redrawn.

8. Present the subject with ratios of worth from the graph and ask him if these ratios reflect his judgment. For example, if Percent Downtime is the performance measure, if 30% downtime is worth 0.6 and 60% is worth 0.2, then 30% should be 3 times as satisfactory as 60%. Alternatively, select several levels of performance and ask the evaluator to specify the ratios. (This procedure is valid for a ratio scale. If the evaluator is willing to assume only an interval scale, then differences should be used, i.e., 0.4 related to 30%.)

9. If the results of step 8 are not consistent with the graph, modify the graph until the evaluator and investigator are satisfied that a valid value function has been constructed.

10. Prepare a composite value function by averaging the points for all evaluators.

11. Present the composite value function to the assembled evaluators. If the group does not agree that the value function is realistic, make modifications until it is acceptable.

APPENDIX C

WASTE COMPOSITION

Wastes vary widely among facilities but may be categorized on the basis of origin, i.e. laboratory, line or process, and construction. Table C-I indicates the range of waste compositions expected in these categories, based on a 1973 LASL survey.[46]

Laboratory wastes include wastes from support areas in the immediate vicinity of, as well as, specific-laboratory refuse. Process wastes typically result from glovebox production operations; construction wastes originate with the disassembly of decommissioned facilities. The bulk of routine, low alpha-activity waste is expected to originate in the laboratory areas.

TABLE C-I

TYPICAL WASTE COMPOSITION

Material	Lab (wt%)	Process (wt%)	Constr. (wt%)
Paper & rags	10-60	10-40	5-20
Plastics	5-60	30-40	5-30
Rubber	5-30	10-45	5-20
Lumber	-	-	5-30
Dirt & concrete	-	-	5-60
Metallics & glass	5-50	30-40	10-50

APPENDIX D

PERFORMANCE MEASURES FOR NORMAL COMPONENTS AND LIQUID THROUGHPUT

I. NORMAL COMPONENTS

A design *basis feed is associated with each* process. Components normally include polyvinyl chloride (PVC), paper and rags, rubber, polyethylene, and dense cellulosics (lumber, computer paper, et(` The performance measure for each component .s the effective feed rate for the component as a fraction of the rated capacity (R_i) multiplied by the maximum fraction of the component (F_i) permitted in the feed. A linear relationship is assumed between performance level ($R_i F_i$) and worth from 0 to 1.5 on the $R_i F_i$ scale (see Fig. D-1). To compute the combined worth for normal components the worth of each component is multiplied by a weight, then all are added.

For example, a process may be designed to handle 100% PVC at the rated capacity, 100% paper and rags at 200% of the rated capacity, 100% rubber at the rated capacity, 100% polyethylene at the rated capacity, and 100% dense cellulosics at 80% of the rated capacity. Details are shown below, where u_i is the worth of $R_i F_i$ from Fig. D-1 for component i, and w_i is its weighting factor.

Component	R_i	F_i	$R_i F_i$	u_i	w_i
PVC	1	1	1	0.6	0.29
Paper & rags	2	1	2	1	0.23
Rubber	1	1	1	0.6	0.21
Polyethylene	1	1	1	0.6	0.17
Dense cellulosics	0.8	1	0.8	0.5	0.10

Then the total worth for normal components is 0.68

$$\left(\sum w_i u_i\right).$$

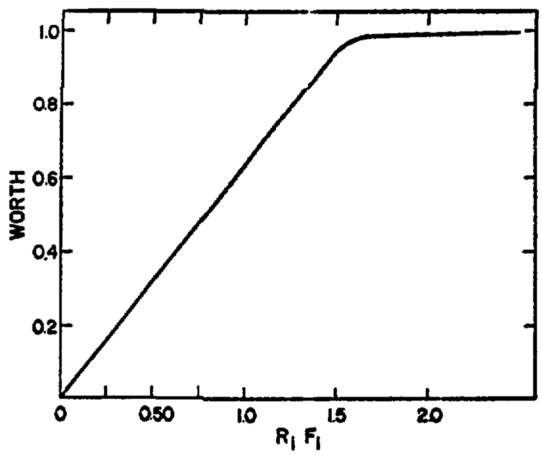

Fig. D-1.

Group Value Function for all performance measures under the subcriterion Handling Normal Components.

II. LIQUID THROUGHPUT

The calculations for determining a high volume of liquid, which would be worth a score of approximately 0.75, are based on a controlled-air incinerator. The limiting factor for combustion of solid *waste is the off-gas cleanup system. Of the major* components, polyethylene will produce the maximum volume of off-gas. If it is assumed that 45 kg/h of polyethylene can be oxidized, then, with other assumptions not covered here, the off-gas system, when liquid alone is treated, can handle a calculable amount of liquid. Calculating the volume of n-hexane that generates the predetermined off-gas volume results in approximately 1.5 ℓ/min of n-hexane.

APPENDIX E

GROUP CONSENSUS

Majority rule is a common method of reaching a consensus. However, it is possible for individuals to behave transitively, but under majority rule the group ranking may be intransitive. For example, if a person prefers a to b, b to c, and a to c, the comparisons are transitive. If he prefers a to b, b to c, and c to a then the comparisons are intransitive, and a valid ranking is not possible. Fortunately, in this study there were no cases of intransitivity either for individuals or for the group. After a set of criteria weights was determined for all individuals, the weights were replaced by ranks and the degree of group consensus was determined by the Kendall Coefficient of Concordance. The ranking convention followed was to replace the largest weight by one and the smallest weight by m(the number of criteria), etc.

The coefficient of concordance is a measure of the overall agreement among k judges in ranking m items (criteria) and is given by

$$W = \frac{S}{S_{max}}, \qquad \text{(E-1)}$$

where S is the sum of the squares of the deviations of the actual rank totals from the average rank sum and S_{max} is the maximum possible value of S.

The value of S is computed by

$$S = \sum_{j=1}^{m} \left(t_j - \frac{k(m+1)}{2}\right)^2, \qquad \text{(E-2)}$$

where t_j is the sum of the ranks for criterion j and the average rank sum is $k(m+1)/2$. The maximum value of S is attained when judges are in perfect agreement. When this occurs the sum of the ranks for any criterion j is

$$t_j = kj, \qquad \text{(E-3)}$$

Substituting Eq. (E-3) in Eq. (E-2) gives the maximum value of S

$$S_{max} = \frac{k^2(m^3 - m)}{12}. \qquad \text{(E-4)}$$

For any other rankings S will be less than S_{max}. If *there is complete disagreement, all* t_j*'s will be equal* to the average rank total and, therefore, S and W equal zero. With perfect agreement W equals one.

If some group ranks are tied, a correction can be made for S_{max}. In addition, a correction for continuity can be made if considered necessary. The significance of W can be tested by using tables of critical values of S or the Chi-square distribution, depending on the number of items. For further information on corrections for ties and continuity, and on tests of significance, see Kendall[40] or Siegel.[38]

A significant value for W suggests that there is agreement among the judges. Then the best estimate of the correct group ranking is given by the order of the sums of ranks, t_j. That is, the most important criterion will have the smallest t_j and the least important one will have the largest t_j. If the sums of ranks are equal for any criteria, the tie is broken by assigning the higher rank (lower number) to the criterion with the smaller variation. This can be determined by summing the squares of the ranks and assigning the higher rank to the criterion with the smaller sum of squares. There were no tied group ranks in this study. The coefficient of concordance (W) for all sets of criteria and the average Spearman Rank Correlation Coefficient ($r_{s_{av}}$) which is a linear function of W, are as follows.

	W	$r_{s_{av}}$
Main criteria	0.85*	0.83
Flexibility subcriteria	0.72*	0.69
Normal components subcriteria	0.47*	0.41
Resource use subcriteria	0.35	0.28

All Ws were significant at the 0.05 level. The starred (*) Ws were significantly different from chance at the 0.01 level or higher. The average rank correlation coefficient is given because more readers probably are more familiar with correlation coefficients than the coefficient of concordance.

APPENDIX F

GROUP WEIGHTS

TABLE F-1

FLEXIBILITY SUBCRITERIA WEIGHTS

Criterion	Weight
Normal components	0.46
Solids throughput	0.25
Handling noncombustibles	0.15
Liquid throughput	0.14
Total	1

TABLE F-II

NORMAL COMPONENTS SUBCRITERIA WEIGHTS

Criterion	Weight
PVC	0.29
Paper & rags	0.23
Rubber	0.21
Polyethylene	0.17
Dense combustibles	0.10
Total	1

TABLE F-III

RESOURCE USE SUBCRITERIA WEIGHTS

Criterion	Weight
Energy	0.39
Scarce strategic materials	0.25
Water	0.21
Land	0.15
Total	1

APPENDIX G

GROUP VALUE FUNCTIONS

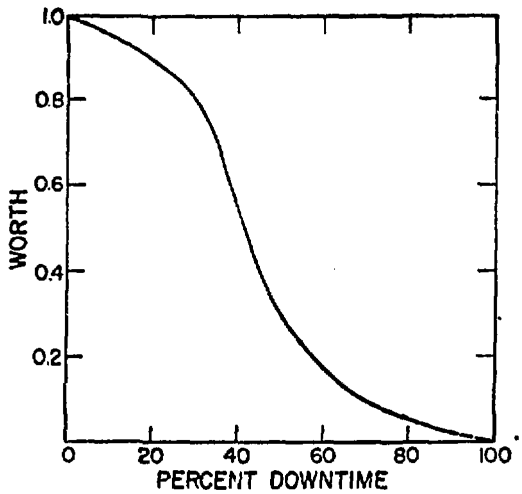

Fig. G-1.
Group value function for the Percent Downtime.

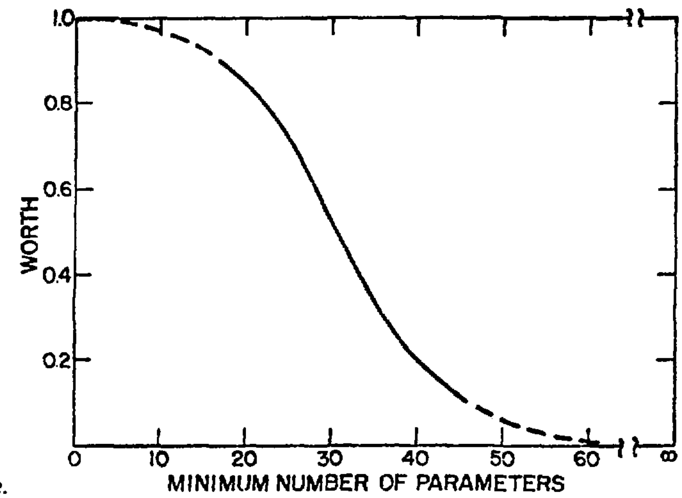

Fig. G-2.
Group value function for the Minimum Number of Parameters that must be controlled.

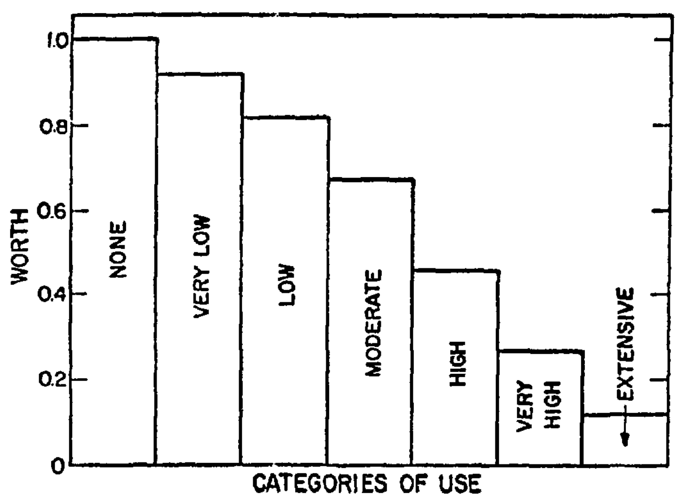

Fig. G-3.
Group value function for four Performance Measures under the criterion Resource Use.

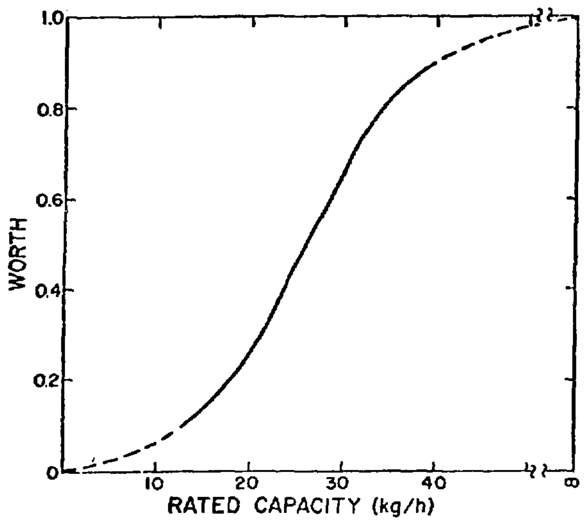

Fig. G-4.
Group value function for Rated Capacity.

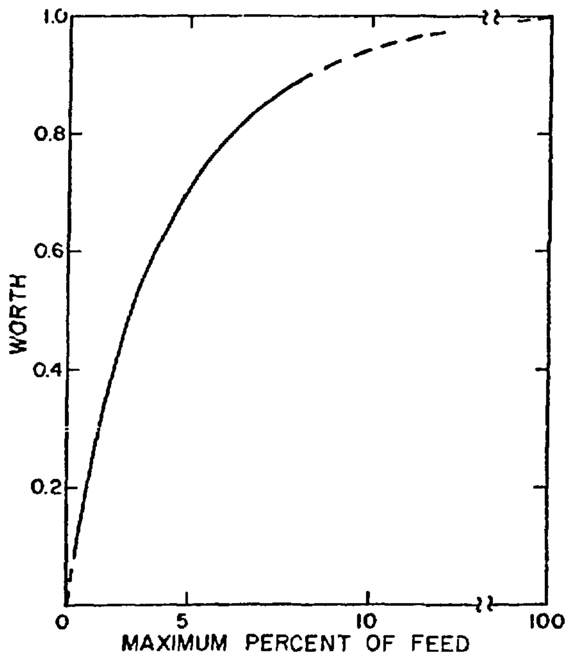

Fig. G-5.
Group value function for Maximum Percent of Noncombustibles permitted in the feed.

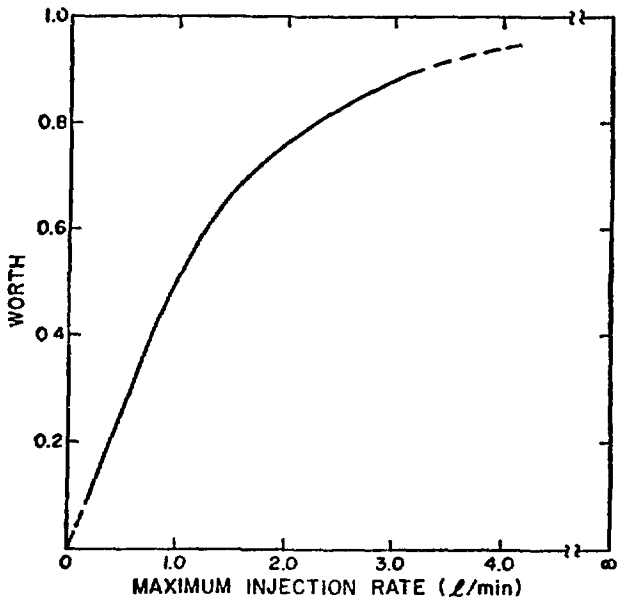

Fig. G-6.
Group value function for Maximum Injection Rate for Organic Liquids.

APPENDIX H

SENSITIVITY ANALYSIS

Sensitivity analysis is essentially operations performed to evaluate variations in the objective function resulting from variations in system parameters. Assume an objective function is

$$U = f(p_i \mid i = 1, \ldots, k), \qquad \text{(H-1)}$$

where p_i is one of the k parameters. Then *one* sensitivity function is the total differential

$$dU = \sum_{i=1}^{k} \frac{\partial f}{\partial p} dp_i . \qquad \text{(H-2)}$$

Alternatively, individual terms or the partial differentials could be used.

For example, if U is the FOM additive formula and k equals the number of performance measures, then if the weights are assumed to be known exactly, or with negligible error,

$$dFOM_X = \sum_{i=1}^{k} w_i \frac{\partial [u_i(X_i)]}{\partial X_i} dX_i . \qquad \text{(H-3)}$$

However, if the performance level estimates are known exactly, or with negligible error, and the p_i's are weights, Eq. (H-2) yields

$$dFOM_{w_i} = \sum_{i=1}^{k} u_i(x_i) dw_i . \qquad \text{(H-4)}$$

As before, X_i and functions of X_i are variables, but x_i and functions where x_i replaces X_i are numbers. The subscript on dFOM indicates the variable parameter.

Equation (H-2) can be conveniently represented in vector terminology

$$dFOM = \mathbf{grad\ f} \cdot dR , \qquad \text{(H-5)}$$

where **grad f** is the gradient and **dR** is the vector of differentials. Before any numerical sensitivity analysis is performed, the magnitude of changes in the FOM for small changes in the performance levels and weights can be examined by inspection of the gradient. Equations (H-3) and (H-4) show that the major concern is in the estimate of the performance levels and the construction of the value functions. The FOM can be significantly affected when the performance levels occur at, or close to, large slopes or slope changes in the value functions. The concern increases as the weight increases. Tables II and IV in the text show that the two performance measures representing 55% of the weight have levels at slope changes (see Fig. 2 and Fig. G-1).

Tilanus[46] used elasticities, which are common economic sensitivity functions. Elasticity is the ratio of the relative change in system response to the relative change in a system parameter.[47,48] The elasticity of the FOM with respect to parameter p is given by

$$E(FOM, p) = \frac{\partial FOM}{\partial p} \frac{p}{FOM} . \qquad \text{(H-6)}$$

Engineers familiar with feedback control theory will recognize Eq. (H-6) as the sensitivity of the transfer function with respect to a specific system parameter.[49,50]

If the system parameter is the weight w_i, then

$$E(FOM, w_i) = \frac{w_i u_i(x_i)}{FOM} , \qquad \text{(H-7)}$$

and if the parameter is the performance measure X_i, then

$$E(FOM, X_i) = w_i \frac{\partial [u_i(X_i)]}{\partial X_i} \frac{X_i}{FOM} . \qquad \text{(H-8)}$$

Equations (H-7) and (H-8), like Eqs. (H-3) and (H-4), show that the value functions will cause the greatest problem. If the performance levels are subject to error, these sensitivity functions will indicate whether or not the FOM might be affected. Equations (H-3) and (H-4) or Eqs. (H-7) and (H-8) can be used to find the trouble spots before numbers are used.

Numerical solutions could be obtained by replacing the differentials with increments and the variables by numbers of interest. Partial derivatives could be estimated or calculated if curve fitting techniques were used for the value functions. Special attention would be required for discrete performance levels and worth scales if numbers were to be substituted in Eq. (H-2) or (H-6). The approach to sensitivity analysis, as well as the form of Eq. (1), the criteria, performance measures, weights, and value functions, depends on your objective in using the evaluation procedure and on the specific application under consideration.

ACKNOWLEDGMENTS

The author is indebted to many LASL waste management personnel for their extensive and diligent effort in defining evaluation criteria, and determining weights and value functions. Special thanks go to A. S. Neuls, G. Maestas, A. Zerwekh, R. A. Koenig, Jr.,and L. C. Borduin for taking part in additional experiments and providing guidance and criticism, and to V. J. McCabe for her patience and expert help in the technical editing and preparation of this document.

The author is solely responsible for the technical content related to the evaluation procedure and any errors in the report and would be glad to be informed of them.

REFERENCES

1. B. L. Perkins, "Incineration Facilities for Treatment of Radioactive Wastes: A Review," Los Alamos Scientific Laboratory report LA-6252 (July 1976).

2. D. L. Ziegler, L. L. Richey, A. J. Johnson, A. L. Teter, P. K. Feng, D. J. Peterson, and L. J. Miele, "Status Report on the Fluidized Bed Incinerator System for U.S. Energy Research and Development Administration Waste," Rockwell International, Atomics International Division, Rocky Flats Plant report RFP-2540 (March 1977).

3. L. C. Borduin, W. E. Draper, R. A. Koenig, A. S. Neuls, and C. L. Warner, "Controlled-Air Incineration of Transuranic-Contaminated Solid Waste," Proc. ERDA 14th Air Cleaning Conf., Sun Valley, Idaho, 1976, (1977), Vol. 1, pp.36-49.

4. R. E. Lerch and J. R. Devine, "Waste Volume Reduction by Acid Digestion," Hanford Engineering Development Laboratory report HEDL-SA-963 (June 1975).

5. D. E. McKenzie, L. F. Grantham, and R. B. Paulson, "Volume Reduction of Radioactive β-γ and TRU Waste by Molten Salt Combustion," in *Management of Low-Level Radioactive Waste*, Proc. Symp. Management of Low-Level Radioactive Waste, Atlanta, Georgia, May 23-27, 1977 (Office of Interdisciplinary Programs, Georgia Institute of Technology, in press).

6. S. S. Freeman, L. M. Klinger, D. F. Luthy, and J. E. Todd, "The Mound Laboratory Cyclone Incinerator for Radioactive Waste," in *Management of Low-Level Radioactive Waste*, Proc. Symp. Management of Low-Level Radioactive Waste, Atlanta, Georgia, May 23-27, 1977 (Office of Interdisciplinary Programs, Georgia Institute of Technology, in press).

7. W. S. Torgerson, *Theory and Methods of Scaling* (John Wiley and Sons, Inc., New York, 1958), pp. 104-112.

8. C. W. Churchman and R. L. Ackoff, "An Approximate Measure of Value," Oper. Res. **2**, 172-187 (1954).

9. H. Terry, "Comparative Evaluation of Performance Using Multiple Criteria," Mgt. Sci.**9**, 431-442 (1963).

10. D. B. Yntema and W. S. Torgerson, "Man-Computer Cooperation in Decisions Requiring Common Sense," IRE Trans. on Human Factors in Electronics **HFE-2**, 20-26 (1961).

11. E. Adams and R. Fagot, "A Model of Riskless Choice," Behavioral Sci. **4**, 1-10 (1959).

12. J. E. Dooley, "Decisions on Social and Technological Tasks Incorporating Expression of Preference and Environmental Insult," Mgt. Sci. **20**, 912-920 (1974).

13. J. S. Dyer, W. Farrell, and P. Bradley, "Utility Functions for Test Performance," Mgt. Sci. **20**, Part I, 507-519 (1973).

14. P. C. Fishburn, "Independence in Utility Theory with Whole Product Sets," Oper. Res. **13**, 28-45 (1965).

15. P. C. Fishburn, "Utility Theory," Mgt. Sci. **14**, 335-378 (1968).

16. D. H. Gustafson, G. K. Pai, and G. C. Kramer, "A Weighted Aggregate Approach to R&D Project Selection," AIIE Trans. **3**, 22-31 (1971).

17. A. J. Klee, "The Role of Decision Models in the Evaluation of Competing Environmental Health Alternatives," Mgt. Sci. **18**, B-52-67 (1971).

18. R. L. Keeney, "Utility Functions for Multiatributed Consequences," Mgt. Sci. **18**, Part I, 276-287 (1972).

19. R. L. Keeney, "A Decision Analysis with Multiple Objectives: The Mexico City Airport," Bell J. of Econ. and Mgt. Sci. **4**, 101-117 (1973).

20. J. R. Miller III, *Professional Decision- Making, A Procedure for Evaluating Complex Alternatives* (Praeger Publishers, New York, 1970).

21. E. Turban and M. L. Metersky, "Utility Theory Applied to Multivariable System Effectiveness Evaluation," Mgt. Sci. **17**, B-817-828 (1971).

22. P. C. Fishburn, *Decision and Value Theory* (John Wiley and Sons Inc., New York, 1964), pp. 328-342.

23. P. C. Fishburn, "Additive Utilities with Finite Sets: Applications in the Management Sciences," Nav. Res. Log. Q. **14**, 1-13 (1967).

24. B. V. Dean and L. A. Roepcke, "Cost Effectiveness in R&D Resource Allocation," IEEE Trans. on Eng. Mgt. **EM-16**, 222-242 (1969).

25. J. N. Warfield, *An Assault on Complexity* (Battelle Memorial Institute, Columbus, Ohio, 1973), pp. 17-1 - 17-2.

26. K. Seiler III, *Introduction to Systems Cost-Effectiveness* (Wiley-Interscience, New York, 1969).

27. J. W. Litchfield, J. V. Hansen, and L. C. Beck, "A Research and Development Decision Model Incorporating Utility Theory and Measurement of Social Values," IEEE Trans. on Systems, Man, and Cybernetics **SMC-6**, 400-410 (1976).

28. P. G. Goodwin, "A Method for Evaluation of Subsystem Alternate Designs," IEEE Trans. on Eng. Mgt. **EM-19**, 12-21 (1972).

29. C. M. Mottley and R. D. Newton, "The Selection of Projects for Industrial Research," Oper. Res. **7**, 740-751 (1959).

30. J. Fasal, "Forced Decisions for Value," Product Eng. **36**, 84-86 (1965).

31. H. J. Einhorn, "The Use of Nonlinear, Noncompensatory Models in Decision Making, "Psychol. Bull. **73**, 221-230 (1970).

32. G. P. Huber, V. K. Sahney, and D. L. Ford, "A Study of Subjective Evaluation Models," Behavioral Sci. **14**, 483-489 (1969).

33. J. B. Burnham, M. H. Karr, G. L. Wilfert, W. S. Maynard, S. M. Nealey, G. Jones, I. Jones, B. Gray, N. Parker, J. Coe, N. Geitner, "A Technique for Environmental Decision Making Using Quantified Social and Aesthetic Values," Battelle, Pacific Northwest Laboratories report BNWL-1787 (February 1974).

34. D. R. J. White, D. L. Scott, and R. N. Schulz, "POED-A Method of Evaluating System Performance," IEEE Trans. on Eng. Mgt. **EM-10**, 177-182 (1963).

35. R. L. Keeney, "An Illustrated Procedure for Assessing Multi-attributed Utility Functions," Sloan Mgt. Rev. **14**, 37-50 (1972).

36. R. T. Eckenrode, "Weighting Multiple Criteria," Mgt. Sci. **12**, 180-192 (1965).

37. C. J. Bartlett, E. Heermann, and S. Rettig, "A Comparison of Six Different Scaling Techniques," J. Soc. Psychol. **51**, 343-348 (1960).

38. S. Siegel, *Nonparametric Statistics for the Behavioral Sciences* (McGraw-Hill Book Company, Inc., New York, 1956), pp. 23-29, 229-239.

39. D. H. Stimson, "Utility Measurement in Public Health Decision Making," Mgt. Sci. **16**, B-17-30 (1969).

40. M. G. Kendall, *Rank Correlation Methods* (4th Ed., Charles Griffin and Co. Ltd., London, 1975), pp. 94-102.

41. W. G. Cochran and G. M. Cox, *Experimental Designs* (2nd Ed., John Wiley and Sons, Inc., New York, 1957), pp. 117-119.

42. P. C. Fishburn, "Methods of Estimating Additive Utilities," Mgt. Sci. **13**, 435-453 (1967).

43. G. P Huber, "Multi-Attribute Utility Models: A Review of Field and Field-Like Studies," Mgt. Sci. **20**, 1393-1402 (1974).

44. N. Dee, J. K. Baker, N. L. Drobny, K. M. Duke, and D. C. Fahringer, "Environmental Evaluation System for Water Resource Planning," Battelle Columbus Laboratories report PB-208 822 (January 1972).

45. E. B. Fowler, J. L. Warren, K. A. Pashman, and J.W. Healy, "Transuranic Waste Research and Development Program," Los Alamos Scientific Laboratory report LA-5281-MS (May 1973).

46. C. B. Tilanus, "Measuring Operating Efficiency," Oper. Res. Q. **26**, 63-69 (1975). Great Britain.

47. J. R. McGuigan and R. C. Moyer, *Managerial Economics: Private and Public Sector Decision Analysis* (The Dryden Press, Hinsdale, Illinois, 1975). pp 156-166.

48. J. M. Henderson and R. E. Quandt, *Microeconomic Theory: A Mathematical Approach* (2nd Ed., McGraw-Hill Book Co., New York, 1971), pp. 27, 57.

49. A. D. Hall, *A Methodology for Systems Engineering* (D. Van Nostrand Co., Inc., New York, 1962), pp. 184-189.

50. J. J. DiStefano III, A. R. Stubberud, and I. J. Williams, *Theory and Problems of Feedback and Control Systems* (Schaum's Outline Series, McGraw-Hill Book Co., New York, 1967), pp. 167-178.