LEGIBILITY NOTICE

A major purpose of the Technical Information Center is to provide the broadest dissemination possible of information contained in DOE's Research and Development Reports to business, industry, the academic community, and federal, state and local governments.

Although a small portion of this report is not reproducible, it is being made available to expedite the availability of information on the research discussed herein.

ORNL-6428
Dist. Category: UC-70 (Waste)

Engineering Physics and Mathematics Division

DETERMINISTIC UNCERTAINTY ANALYSIS

Brian A. Worley

ORNL--6428

DE88 006139

Date Published - December 1987

Prepared for the
National High-Level Waste Management Program
U.S. Department of Energy

Prepared by the
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831
operated by
MARTIN MARIETTA ENERGY SYSTEMS, INC.
for the
U.S. DEPARTMENT OF ENERGY
under contract DE-AC05-84OR21400

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Uncertainties of computer results are of primary interest in applications such as high-level waste repository performance assessment in which experimental validation is not possible or practical. Because of the complex computational structure of large computer models, and because of the large number of input and data parameters associated with such models, to date almost all uncertainty analysis of computer results has been performed using a statistical approach. This paper presents a deterministic uncertainty analysis (DUA) method for calculating uncertainties that has the potential to significantly reduce the number of computer runs compared to conventional statistical analysis. The method is based upon the availability of derivative and sensitivity data such as that calculated using the well known direct or adjoint sensitivity analysis techniques. Formation of response surfaces using derivative data and the propagation of input probability distributions are discussed relative to their role in the DUA method. A sample problem that models the flow of water through a borehole is used as a basis to compare the cumulative distribution function of the flow rate as calculated by the standard statistical methods and the DUA method. Propogation of uncertainties by the DUA method is compared for ten cases in which the number of reference model runs was varied from one to ten. The DUA method gives a more accurate representation of the true cumulative distribution of the flow rate based upon as few as two model executions compared to fifty model executions using a statistical approach.

I. INTRODUCTION

The Office of Nuclear Waste Isolation (ONWI) is performing sensitivity and uncertainty studies as part of its performance assessment of a high-level nuclear waste repository in salt.[1,2] The role of the sensitivity analysis is to provide a means to limit the scope of the more complicated problem of quantifying uncertainties. Uncertainty analyses will be performed to support design reliability studies, to produce a cost-benefit analysis in conjunction with cost estimates, to insure compliance with regulatory criteria, and to help identify important research and development needs.

For quantification of uncertainties in computer-generated results, the problem can be expressed more precisely as the propagation of input uncertainties through models by the laws of probability to obtain output uncertainties. Uncertainties of computer results are of primary interest in applications such as repository performance assessment in which experimental validation is not possible or practical. Because of the complicated nature of the computational structure of large computer models, and because of the large number of input and data parameters associated with such models, to date almost all uncertainty analysis of computer results has been performed using a statistical approach. The purpose of this work is to present an alternate deterministic approach that retains the characteristics of analytically computing result uncertainties based upon parameter probability distributions.

II. BACKGROUND

The analytical propagation of input uncertainties through a calculational model is unfeasible, if not impossible, for all but the most simple models. The difficulty lies in mapping probability density functions from a multidimensional space of input parameters to the one-dimensional output distribution function. To circumvent this problem, the most common approach is to randomly sample the input distributions and then calculate the model output of interest, constructing a probability distribution of the output by rerunning the model for each sample set of input parameters. The input probability distributions and any parameter correlations are handled, in a statistical sense, in the sampling procedure.[3,4] The information available from probability propagation is lost, but hopefully the sampling procedure will lead to an output distribution that is representative of that which would result from the actual propagation of input probability distributions. As the number of sampling sets increases, the difference between the calculated and "true" output distribution diminishes. The problems occur in practice when the number of runs of the computer model needed to assure a large enough statistical sample becomes too expensive.

Another approach is to discretize the input probabilities into histograms and evaluate the model output of interest for all possibilities of parameter combinations to form a probability tree.[5] All parameter correlations are incorporated into the tree probability structure. This method does not rely on random sampling and probabilities are easily propagated in probability trees by simple multiplication. The histogram probability distributions are not actually propagated, but rather mean or

endpoint parameter values are used. This method is quite feasible for models with a small number of parameters or even for a large number of input parameters if the model is simple (inexpensive). Again the problem arises when the computer model has numerous input parameters and/or is expensive to run.

A third approach is the response surface method in which the computer model is replaced with a simple analytical expression.[6] The expression is constructed by fitting the computed values of the model output to the corresponding input parameters, or more generally, to chosen functions of the input parameters. The uncertainty in the computed value of the expression is then determined in the usual statistical sense by sampling of the input distributions. The advantage of replacing the model with the response surface is the drastically reduced computational time to compute the expression result compared to running the computer model. The disadvantage is the introduction of error in the calculated output by replacement of the model with a simple expression.

This paper introduces a method for calculating the uncertainty in computer model results that is analytic (deterministic) in principal and that is firmly based on the model equations. The method combines the characteristics of the response surface method and probability trees. Statistical sampling is not required and probabilities are propagated analytically within discretized numerical meshes that encompass the parameter space. This approach is referred to as the Deterministic Uncertainty Analysis (DUA) method.

III. DETERMINISTIC UNCERTAINTY ANALYSIS METHOD

The approach underlying the deterministic calculation of uncertainties in the DUA method relies upon (1) a replacement of the computer model with an analytical function relating the responses of interest to the parameters of interest and (2) discretizing the parameter space and calculating the expected value of the response within each discrete parameter space "mesh." The parameters of interest are chosen to be those that are "uncertain," meaning that they have known or assumed probability distributions. The parameters of interest may often include the entire set of data used by the computer model.

This deterministic approach differs from the response surface methods in two ways. First, the analytical function is constructed based upon the response value of interest as well as the partial derivatives of the response with respect to each of the parameters. The classical response surface method constructs the surface (analytical expression) based only upon the response value at each parameter space point. Thus the degrees of freedom with which to fit the response to the parameter values is much greater in the DUA method than in the response surface methods. There is of course no reason to distinguish the DUA method from response surface methods if the response surfaces are constructed using derivative data; but in the classical response surface methods these derivatives are assumed to be unavailable, most likely since the response surface methods grew out of experimental design fields in which only set points of the control variables (parameters in our terminology) and the experimental measured values (responses) are known. As the response surface methods became to be used for replacement of large, complex computer models, again the derivative information was not used because of the difficulty of calculating partial

derivatives chained through complex computational paths. However, the development of efficient methods for calculating derivatives and sensitivities for large-scale computer model results has progressed steadly based upon a firm theoretical foundation.[7,8,9,10,11,12] Moreover, now new techniques for calculating derivatives within existing computer models based upon computer calculus are available.[13,14,15] For these reasons, the calculation of derivative information for the purpose of improving the formation of response surfaces is both practical and cost effective. This availability of derivative information is a key component in the DUA method.

The second feature that distinguishes the DUA method from response surface methods, and a feature that it has in common with probability tree methods, is that the entire parameter space is spanned. In the response surface methods, the construction of response surfaces has been primarily used to dramatically increase the number of sampling points in a statistical determination of response probability distributions since the evaluation of the analytical response surface expression is much less expensive to obtain than the corresponding computer model result. However, only by spanning the entire parameter space can probabilities be propagated, either through a computer model or through an analytical expression. Spanning the entire parameter space is practical only if the discretization of parameter probability distributions is performed over a reasonably large mesh. In probability tree methods, for example, the probability distributions are typically replaced with the high and low values of the distribution. The DUA method extends the probability tree methods into a more rigorous propagation of probabilities in two ways: (1) Since an analytical expression relates the response to the parameters, the expected value of the response

over each discretized mesh can be calculated analytically and thus gives a more meaningful value than just a single sampling point within the mesh. (2) Because the computer model is replaced with an analytical expression, a finer mesh size can be constructed over the parameter space and a more accurate representation of the parameter probability distributions of most interest can be obtained.

Another point to be made in favor of propagating probabilities through an expression that only approximates the original computer model, but one that makes the DUA method possible for computer models with a large number of parameters, is the integral nature of the probability distribution of the response. The probability distribution of the response of interest is an integral quantity and errors introduced by replacement of the computer model by an analytical expression are most often washed out when the parameter space is completely spanned.

IV. FORMATION OF RESPONSE SURFACE USING DERIVATIVE INFORMATION

The DUA method replaces the computer model with an analytical expression by relating the response of interest as calculated by the computer model to the parameter values by techniques that incorporate knowledge of the partial derivatives of the response with respect to the parameters of interest. The simplest scheme is linear extrapolation from reference space points to each mesh of the discretized parameter space. Within each mesh the response surface is linear with respect to the parameters, and the calculation of the expected value of the response within the mesh, given parameter probability functions, is straightforward. Various extrapolation schemes are possible; several are discussed and compared in the sample problem discussion.

A more general approach for construction of a response surface is a least-squares fitting technique. In Appendix A, linear least-squares fitting is reviewed, and possible schemes for incorporating derivative information into the standard fitting technique are presented. Basically, to construct a response surface to a given order of expansion, the use of derivative information reduces the number of computer runs required to uniquely determine the expansion coefficients by a factor of approximately 1/K, where K is the number of parameters. One can either construct a global response surface or define local response surfaces over subregions of the parameter space. A local fit of the response values and derivatives using a low-order function may be more desirable than a global fit using a higher order function to fit a large portion of the data because a higher order fit involving many response points may result in a very radically behaved function in the parameter space not near the fitted points. For this reason most of our research to date has focused upon either local

fitting or linear extrapolation from reference parameter space points. By careful selection of these parameter space points for which model results will be obtained, the number of computer runs can be held to a small fraction ($<< 1/K$) of the number required for the conventional construction of a response surface.

Finally, sensitivity analysis plays an important role in the formation of the response surface by eliminating those parameters that have a negligible effect on the result of interest based on their sensitivities and uncertainty ranges. Also, the derivative information from the reference model runs can be used to identify the occurrence of parameters that occur exclusively in a given combination. Such identification reduces the parameter space by replacement of the individual parameters with the particular combination. For example, if the derivative of the response with respect to each of two parameters is the same at each reference space point sampled, the two parameters most likely appear in the model as a sum, and a single parameter representing the sum of the two can be used in the formation of the response surface in place of the two individual parameters. The sample problem exemplifies these uses of sensitivity and derivative data in the formation of the response surface.

V. PROPAGATION OF PROBABILITIES

The propagation of parameter probability distributions from the multidimensional parameter space to the singly-dimensioned result space is determined by the governing system of equations and the input variable probability density functions (pdf's). In theory, this propagation can be performed analytically by convolution of the integral of the parameter space into a discrete number of integrals of the singly-dimensioned response space, in which each integral is over a monotonically changing function representing the result. However, because the identification of the convolution integrals, in particular the limits of the integrals, is virtually impossible for all but the simplest problems, and because the model equations are nonlinear and complexly intertwined in general, the propagation of probability distributions through computer model cannot be treated analytically in the strictest sense.

The propagation of parameter probability distributions in the DUA approach is performed by discretizing the K-dimensional parameter space (K = number of parameters) into L meshes, each mesh denoted by m_ℓ. The probability of mesh m_ℓ occurring within the entire parameter space, $p(m_\ell)$, is calculated as well as the expected value of the response function within the mesh, $E(r_\ell)$, where r_ℓ represents the response function within m_ℓ. The probability $p(m_\ell)$ is assigned to $E(r_\ell)$ to obtain the probability of $E(r_\ell)$ within the discrete space of expected values. The pairs of $p(m_\ell)$ and $E(r_\ell)$ are reordered such that $E(r_1) < E(r_2) < \ldots E(r_L)$ and as such constitute the probability distribution function of the response r over the parameter space. The cumulative distribution function of r, F(r), is the running sum of the reordered $p(m_\ell)$ paired with the corresponding value of $E(r_\ell)$. In

the limit as $L \Rightarrow \infty$, $F(r)$ approaches the true cumulative distribution function of r as calculated using the response function.

Let the functional form of the response within m_ℓ be given by

$$r_\ell \approx g_\ell(x) \tag{1}$$

where $g_\ell(x)$ is the response surface function within m_ℓ resulting either from a fitting procedure or from a linear expansion from one or more reference space points. The vector x is the K-dimensional parameter column vector given by $x = \{x_1, x_2, \ldots, x_K\}^T$, where the superscript T denotes the transpose. Given the joint probability function of x as $P(x) = P(x_1, x_2, \ldots x_K)$, the probability that $x \subset m_\ell$ is given by

$$p(m_\ell) = p(x \subset m_\ell) = \int_{m_\ell} P(x)\, dx \quad , \tag{2}$$

and the expected value of the response r within m_ℓ, $E(r_\ell)$, is

$$E(r_\ell) = \int_{m_\ell} g_\ell(x)\, P(x)\, dx \,/ p(m_\ell) \quad . \tag{3}$$

The values of $p(m_\ell)$ and $E(r_\ell)$ as calculated by Eqs. (2) and (3) are used to construct the probability density function and cumulative distributions function of the response r.

VI. SAMPLE PROBLEM

Reference 16 describes a sample problem that exemplifies the use of uncertainty analysis in high-level waste applications. The sample problem consists of three coupled equations with eight input parameters and three dependent variables. The analysis focuses on one of the three dependent variables as the response of interest, and statistical techniques are used to calculate the cumulative distribution of the flow rate given probability distributions for the eight input parameters.

The governing equations describe the downward flow of water through a borehole that is drilled from the ground surface through two aquifers. For a fully penetrating well and no ground-water gradient, the steady-state flow through the upper aquifer into a borehole is given by

$$Q = \frac{2\pi(H_u - H_{wu})T_u}{\ell n(r/r_w)} \quad , \tag{4}$$

where

Q = flow, m^3/yr

T_u = transmissivity of upper aquifer, m^2/yr

H_u = potentiometric head of upper aquifer, m

H_{wu} = steady-state potentiometric head in borehole at upper aquifer, m

r = radius of influence, m

r_w = radius of borehole, m.

Similarly, the steady-state flow from the borehole to the lower aquifer is given by

$$Q = \frac{-2\pi(H_\ell - H_{w\ell})T_\ell}{\ell n(r/r_w)} \quad , \tag{5}$$

where

T_ℓ = transmissivity of lower aquifer, m^2/yr

H_ℓ = potentiometric head of lower aquifer, m

$H_{w\ell}$ = steady-state potentiometric head in borehole at lower aquifer, m

The flow of water through the borehole is assumed to be laminar and isothermal and is given by

$$Q = \frac{\pi r_w^2 (H_{wu} - H_{w\ell}) K_w}{L} \quad , \tag{6}$$

where

K_w = hydraulic conductivity of borehole, m/yr.

L = length of borehole, m.

In Eqs. (4-6), Q, H_{wu}, and $H_{w\ell}$ are dependent variables; the flow rate of water, Q, is the response of interest. The uncertainty problem is to calculate the cumulative distribution function of Q, F(Q), given the probability density functions of the eight input parameters r_w, r, T_u, T_ℓ, H_u, H_ℓ, K_w, and L. The probability density functions of these eight parameters are shown in Table B.1 of Appendix B, which is taken from Ref. 16.

The standard statistical approach for calculating F(Q) is to define a design matrix based upon the pdf's of the parameters. Several sampling procedures are available for determining a suitable design matrix. For this problem, Ref. 16 investigates the formation of design matrices based upon the Latin Hypercube Sampling (LHS) procedure using 10 and 50 design points. The 10-point and 50-point LHS design matrices are given in Tables B.2 and B.3 of Appendix B, along with the calculated values of Q, H_{wu}, and $H_{w\ell}$.

The choice of the sets of input parameters in a design matrix hopefully accounts for the parameter pdf's and any parameter correlations such that each calculated value of the response is of equal probability. Thus, the probability of a calculated response is 1/N, where N is the number of input sets in the design matrix and formation of F(Q) is performed by ranking the values of Q from lowest to highest and apportioning a probability of 1/N to each value. Figure B.1 of Appendix B shows the plots of F(Q) resulting from the use of the 10-point and 50-point LHS design matrices. Clearly, the 10-point LHS design matrix does not result in a good approximation of F(Q) when compared to the 50-point set. It is important to keep in mind that a design matrix based on N input sets requires that the computer model be run N times to determine F(Q). For this sample problem, the "computer model" consists of Eqs. (4-6) and these were solved 10 times for the 10-point design matrix and 50 times for the 50-point design matrix.

VII. APPLICATION OF DUA TO THE SAMPLE PROBLEM

The DUA method was applied to this sample problem and the results compared to the published statistical results of Ref. 16. The choice of N reference points from which the response surface is formed in the application of DUA to this problem was chosen to be a subset of the 10-point LHS design matrix. For each reference point i, i=1,...,N, defined by the 8-dimensional parameter vector $x_i = (r_w^i, r^i, T_u^i, T_\ell^i, H_u^i, H_\ell^i, K_w^i, L^i)^T$, the derivative vector $q_i = [(\partial Q/\partial r_w)_i, (\partial Q/\partial r)_i, (\partial Q/\partial T_u)_i, (\partial Q/\partial T_\ell)_i, (\partial Q/\partial H_u)_i, (\partial Q/\partial H_\ell)_i, (\partial Q/\partial K_w)_i, (\partial Q/\partial L)_i]^T$, and the response $Q(x_i)$ were calculated. First-order sensitivities of Q with respect to each parameter, defined by $(\partial Q/\partial x_k)_i\ (x_{k,i}/Q(x_i))$, were also computed. Parameters with sensitivities and uncertainty ranges such that their influence on Q was negligible were dropped from the parameter space for the purpose of forming the response surface. (The values of the derivatives and sensitivities for the 10-point LHS design matrix are listed in Table B.4.) As a result, the parameters T_u, r, and T_ℓ were not used in the formation of the response surface. In addition, the derivatives of H_u and H_ℓ are equal but opposite in sign indicating that the two parameters occur in the model in the combination of (H_u-H_ℓ). In fact, in this case we can verify this relationship between Q and H_u and H_ℓ by direct solution of Eqs. (4-6) for Q; the solution is given by $Q = 2\pi\ T_u(H_u - H_\ell)/\{[\ell n(r/r_w)][1 + 2LT_u/(\ell n(r/r_w)r_w^2K_w) + T_u/T_\ell]\}$. Therefore these two parameters were replaced by the single parameter (H_u-H_ℓ). The response surface was then constructed based on only the four significant parameters (H_u-H_ℓ), r_w, K_w, and L. This reduction in the number of significant parameters and combination of parameters illustrates the role of sensitivity analysis in the DUA method as applied to this sample problem.

In this application, the response surface was formed by linear extrapolations of Q from the reference points. Three extrapolation schemes were first tested. The accuracy of each of the tested extrapolation schemes was evaluated by comparison of the extrapolated value of Q to the actual calculated value at the space points making up the 50-point LHS design matrix.

The three extrapolation schemes, denoted ES1, ES2, and ES3, are described below. To account for the fact that the dimensionality of the parameters varies, the metric $d_{i\ell}$ used to define the "distance" between a reference space point x_i and the space point x_ℓ to which the response is to be extrapolated was chosen as

$$d_{i\ell} = \frac{\sum_{k=1}^{K} \left(\frac{X_{k,\ell} - X_{k,i}}{X_{k,i}} \right)^2 S_{k,i}^2}{\sum_{k=1}^{K} S_{k,i}^2} , \qquad (7a)$$

where $X_{k,\ell}$ and $X_{k,i}$ are the elements of x_ℓ and x_i, and where $S_{k,i}$ is the sensitivity of the response Q with respect to parameter k at point i, defined as $S_{k,i} = (\partial r/\partial x_k)_i \, (x_{k,i}/Q_i)$. Note that $x_{k,i}$ cannot equal zero if Eq. (7a) is used. However, if $x_{k,i} = 0$, $d_{i\ell}$ can be alternatively expressed, using the definition of $S_{k,i}$, as

$$d_{i,\ell} = \frac{\sum_{k=1}^{K} (x_{k,\ell} - x_{k,i})^2 \, (\partial Q/\partial x_k)_i^2}{Q_i^2 \sum_{k=1}^{K} (S_{k,i})^2} . \qquad (7b)$$

Equation (7b) can be used if the $x_{k,i}$ can take on zero values and Q_i^2 does not equal zero; in most applications the responses of interest are

non-zero. Equations (7a) and (7b) remove the dimensionality of each parameter by normalization of the parameter value change to its value at the reference point. Also, $d_{i\ell}$ as defined above makes use of sensitivities to weight the more important parameters.

The calculated value of the flow rate by extrapolation from point x_i to point x_ℓ, $Q(x_\ell)$, for each of the three extrapolation procedures is:

$$\text{ES1}: \quad Q(x_\ell) = Q(x_i) + (x_\ell - x_i)^T q_i \quad , \qquad i \ni d_{i\ell} = \min\,(d_{n\ell}) \ , \ n = 1,N \tag{8a}$$

$$\text{ES2}: \quad Q(x_\ell) = 1/2(Q(x_i) + Q(x_j)) + 1/2\{(x_\ell - x_i)^T q_i + (x_\ell - x_j)^T q_j\} \ ; \qquad i \ni d_{i\ell} = \min(d_{n1}); \ j \ni d_{j\ell} = \min(d_{n\ell}), \ n \neq i \ . \tag{8b}$$

$$\text{ES3}: \quad Q(x_\ell) = \frac{d_{i\ell}}{d_{i\ell} + d_{j\ell}} \, (Q(x_j) + (x_\ell - x_j)^T q_j) + \frac{d_{j\ell}}{d_{i\ell} + d_{j\ell}} \, (Q(x_i) + (x_\ell - x_i)^T q_i) \ ; \qquad i \text{ and } j \text{ as in ES2} \ . \tag{8c}$$

Recall that the vector q_i is the derivative vector which has as its elements the partial derivatives of the response with respect to the parameters.

Scheme ES1 simply extrapolates from the nearest point as defined by the metric $d_{i\ell}$ in Eq. (7b). ES2 averages the extrapolated value of the nearest two points, and ES3 weights the extrapolated value of Q of the two nearest points by their respective closeness to the point of interest.

The vectors x_i and x_ℓ in Eqs. (8a,8b,8c) are of dimension K where K (used in Eqs. (7a) and (7b)) is equal to the total number of problem parameters if the sensitivity and parameter range data is not used to reduce the parameter space. However, as discussed previously, in most applications, many parameters may have a negligible effect upon the response, and an examination of the sensitivity data and parameter ranges can be used to reduce the parameter space before the response surface is formed. In this sample problem, for example, the parameter space was reduced from eight to four parameters, and Eq. (7b) and Eqs. (8a,8b,8c) were evaluated for the reduced parameter set.

These three schemes were compared for ten different cases in which the number of reference points N was varied from one to ten. In general, the number of reference points (model runs) in the DUA method would initially be one or two. The CDF of the response of interest would be calculated and additional reference points added one by one until the CDF did not change appreciably. The choice of reference points could be determined dynamically based upon knowledge of the sensitivities and derivatives already calculated, or statistically using techniques such as LHS (but where the number of reference points is very small compared to the number required for statistical determination of the CDF). For this sample problem, however, the reference points used in the DUA approach for the ten cases were a near-optimal selected subset of the 10-point LHS design set. Algorithms for choosing optimal reference points are a problem for further investigation, particularly dynamic schemes based upon prior knowledge. The extrapolated points x_ℓ were chosen to be the parameter space points of the 50-point LHS design matrix. The results are summarized in Table 1.

Table 1. PERCENT DIFFERENCE BETWEEN CALCULATED AND EXTRAPOLATED VALUES FOR EXTRAPOLATION FROM N POINTS TO THE 50-POINT LHS SPACE POINTS

VALUES SHOWN ARE FOR EXTRAPOLATION SCHEME ES1*

Number of Points From Which the Extrapolations Are Performed

Space Point	N=10	N=9	N=8	N=7	N=6	N=5	N=4	N=3	N=2	N=1
1	0.3				-1.1	-4.5	-0.9			
2	-0.2									
3	-0.6			-0.4						
4	2.1									-48.2
5	-0.5									
6	-1.4									-11.0
7	-0.9									
8	-0.1				0.5	0				
9	0.9									
10	-5.6			-3.0				1.8	0.6	
11	-2.6							-3.8		
12	-8.0							-14.5	-2.0	-57.8
13	-0.2					0.4				
14	-1.6		0							-16.4
15	-6.0	0.9								-4.1
16	-0.1						8.4			
17	-0.4			-1.0						
18	3.6									-52.9
19	1.2		-2.4				-9.9	-0.8	-59.7	
20	-0.5									
21	0.8									-31.9
22	-0.2					-2.0	2.0			
23	-17.9					-3.7	-10.5			
24	0.2				-6.1					
25	-0.8		9.7					0.7		-2.5
26	-0.8					1.7				
27	-1.4						-10.0			
28	0.6					-1.7	-2.0		9.0	
29	-1.6									
30	-1.1									
31	0.1							-0.6		-17.7

Table 1. (Continued).

Space Point	N-10	N-9	N-8	N-7	N-6	N-5	N-4	N-3	N-2	N-1
32	-0.6									
33	0.5								-34.2	
34	-0.2									
35	0.1	1.2								
36	-1.8			-7.3				-0.1		
37	1.8						-2.4			
38	1.7			-0.7	-1.8					
39	0.4				-2.4					
40	-19.1							-44.1	5.5	-92.6
41	-1.0		0.7					-0.6		-22.8
42	-2.0			-2.4					2.0	
43	-0.2				-5.9					
44	0.4					2.7				
45	0.1									
46	-1.9				-6.3	0.8				
47	0				-0.3					
48	0.3							-14.8		
49	-1.5									
50	-11.1					-0.7				
R (ES1)*	177.9	168.0	191.6	211.0	361.4	359.6	373.4	486.4	379.8	2868.6
R (ES2)	300.0	345.2	544.4	558.6	618.7	2132.9	2389.0	2614.0	3009.0	2156.0
R (ES3)	281.8	288.8	1098.0	439.2	645.5	2508.4	3028.0	2614.0	3009.0	4126.0

*ES1 refers to extrapolation from nearest point; ES2 refers to extrapolation from nearest two points with each point weighted equally; ES3 refers to extrapolation from nearest two points with the points weighted by their respective proximity to the extrapolated point.

R is a measure of the fit of the extrapolated points to the 50 LHS points;

$$R = \sum_{\ell=1}^{50} (Q(x_\ell) - Q_c(x_\ell))^2 \quad .$$

Blank entries indicate that the value is unchanged from the value at the left.

(Table 1 is based upon using a metric $d_{i\ell}$ that did not weight with sensitivities but the values listed in the table are not appreciably different than if Eq. (7a) had been used.) The columns represent the results of extrapolation using all 10 of the input parameter sets of the 10-point LHS design matrix (far left column) down to extrapolation from a single input parameter set of the 10-point LHS design matrix (far right column). The numbers at the extreme far left represent the 50 input parameter sets of the 50-point LHS design matrix listed in Table B.3. The entries in Table 1 are the percent differences between the extrapolated value of $Q(x_\ell)$ from Eq. (8a) and the actual value of Q, $Q_c(x_\ell)$, at the 50 points as calculated by the model. Blank entries indicate no change from the value of the percent difference from that of the column to the immediate left. A blank value thus indicates that the same reference point was chosen from which to extrapolate as in the previous case. Below each column is listed the variable R which represents the sum of the squares of the differences between $Q(x_\ell)$ and the actual value of Q for each of the three extrapolation schemes. For this problem, scheme ES1 gives the most accurate prediction of the flow rate. Using information from the second closest point does not improve the extrapolated value of Q; and weighting based upon distance to the desired space point is better than straight averaging in only a few cases. Using 7 to 10 reference points has approximately the same accuracy, then the accuracy drops somewhat when using 2 to 6 reference points. However, examination of R for the best extrapolation scheme, ES1, reveals that the decrease is not monotonic as the number of reference points increases, indicating only that the selection of the reference parameter space point to eliminate from case to case was not optimal. (The actual values of Q were assumed to be unknown.)

Extrapolation scheme ES1 was chosen for the formation of the response surface over which the parameter probabilities were to be propagated. The entire parameter space of significant parameters as identified in the sensitivity analysis (r_w, H_u-H_ℓ, K_w, and L) was divided into L discrete, nonoverlapping meshes, m_ℓ, ℓ= .,,,L. The expected value of Q within each mesh, $E(Q_\ell)$, was determined by replacing Q_ℓ for $g_\ell(x)$ in Eq. (3). Here, $Q_\ell = Q(x)$, $x \subset m_\ell$, where within m_ℓ, Q(x) was calculated using Eq. (8a). The mesh probability $p(m_\ell)$ was calculated from Eq.(2) using the parameter probability distributions from Table B.1 (note that $p(H_u\text{-}H_\ell)$ had to be calculated separately using the individual distributions given in Table B.1). For calculating the values of $d_{i\ell}$ needed for the extrapolation scheme, the value of x_ℓ chosen for each m_ℓ was E(x), $x \subset m_\ell$. The probability assigned to each $E(Q_\ell)$, ℓ=1,,,L, was the corresponding value of $p(m_\ell)$. As discussed earlier, in the DUA method the number of meshes, L, is chosen such that the entire reduced parameter space is covered by nonoverlapping discrete meshes and therefore the entire probability space is complete in that

$$\sum_{\ell=1}^{L} p(m_\ell) = \sum_{\ell=1}^{L} p(E(Q_\ell)) = 1.$$

As a benchmark against which a comparison of the DUA method and the statistical results from Ref. 16 could be compared, the sample problem model was executed 2304 times in order to determine the "true" CDF of Q for this problem. A comparison of this benchmark 2304-pt CDF to the statistical 50-point CDF from Ref. 16 is shown in Fig. 1. The CDF based upon the 50-point LHS design matrix is a fairly accurate represention of the true CDF of Q. DUA method results were obtained for three cases. The first case propagates probabilities over a 144-mesh parameter space

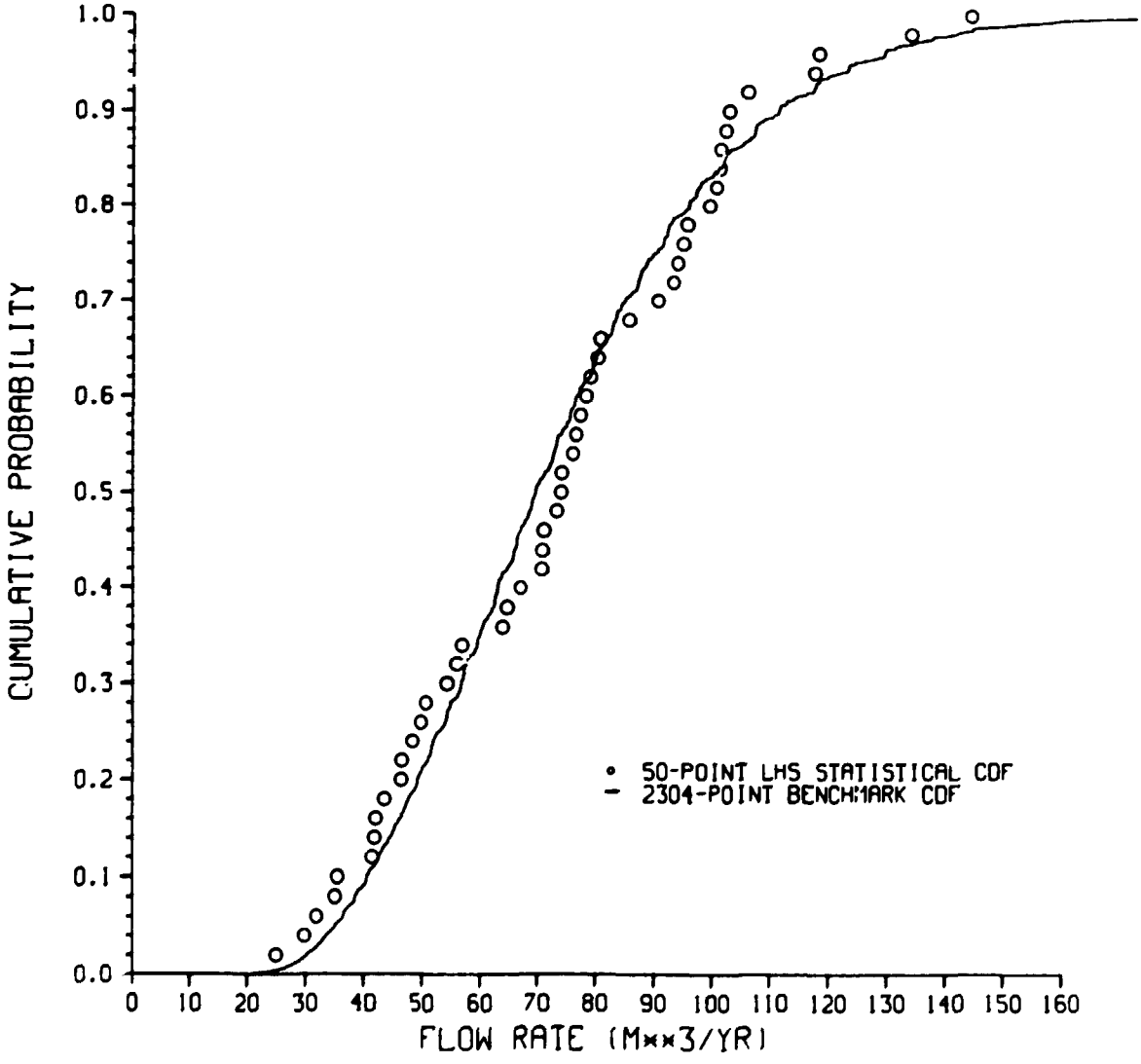

Fig. 1. Comparison of the 50-Point LHS CDF to the 2304-Point Benchmark CDF.

by extrapolation from two reference model runs, and the second case by extrapolation from 10 reference model runs. Figure 2 compares the CDF's for these two DUA cases to the benchmark 2304-pt CDF. The DUA method for both cases gives a somewhat more accurate representation of the true CDF than the use of a 50-point LHS design matrix. The curves in Fig. 2 also indicate the integral nature of a CDF: although individual values of Q may be inaccurately predicted using a response surface (see Table 1), the CDF of Q is accurately represented. In addition, the accuracy of the CDF was only slightly improved for this problem by increasing the number of computer runs from two to ten.

The effect of the number of discretized meshes over which probabilities are propagated was evaluated in the third DUA case by increasing the number of meshes from 144 to 2304. The resulting CDF based upon the two-point extrapolated response surface is compared to the benchmark CDF in Fig. 3, in which only every fortieth point of the DUA-generated curve is plotted. For this problem, the increase in the number of meshes from 144 to 2304 leads to a somewhat more accurate representation of the true CDF. The apparent discontinuous nature of the deterministic curves in Fig. 2 arises as a result of duplications in the response for different values of H_u and H_ℓ. These singularities tend to disappear with increasing number of discretized meshes. Thus, the number of meshes over which probabilities are to be propagated should be as high as possible within constraints of the cost effectiveness of constructing the response surface and propagating the probabilities.

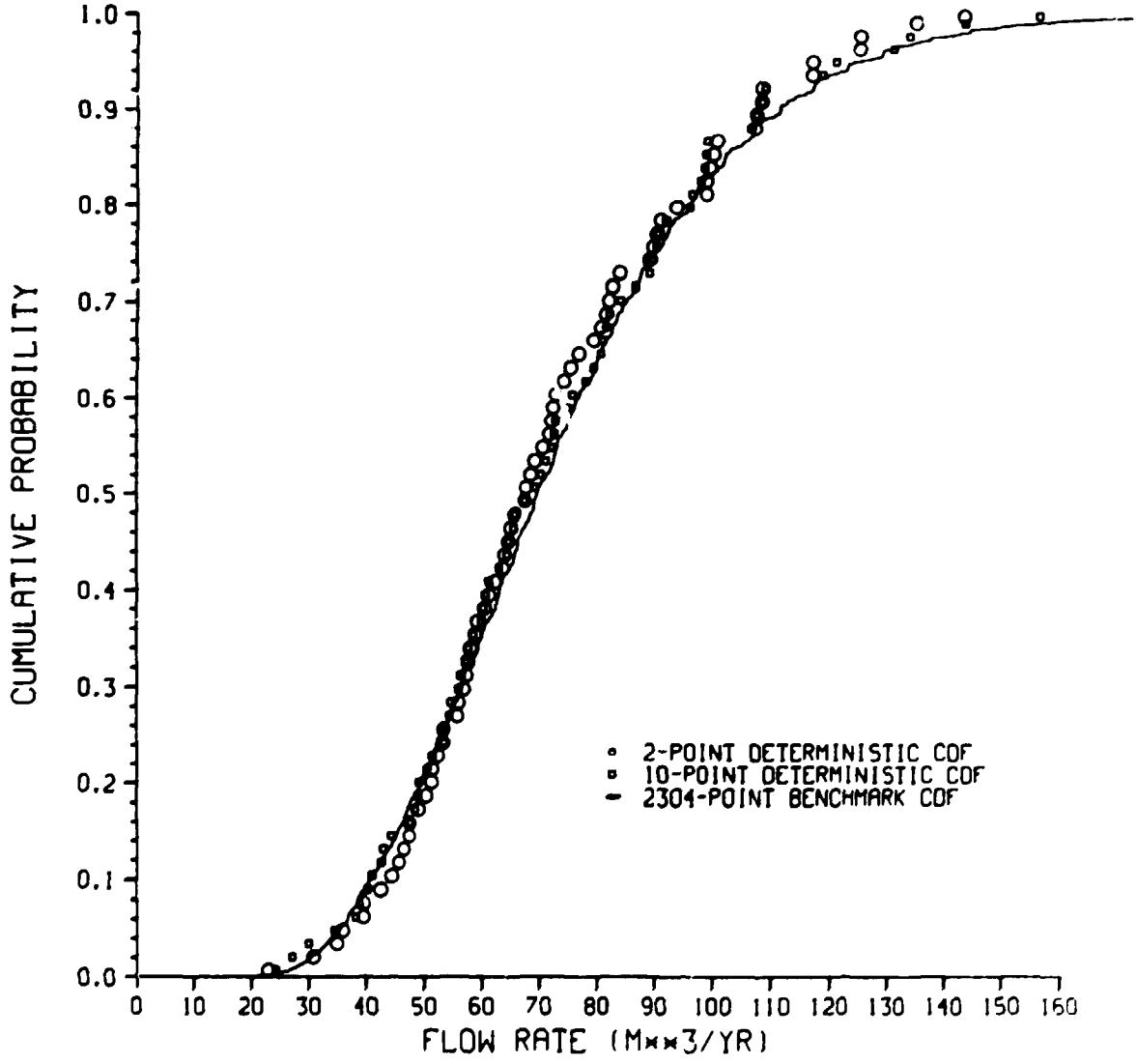

Fig. 2. Comparison of 2-Point and 10-Point Deterministic CDF by Extrapolation to 144 Discrete Meshes with the 2304-Point Benchmark CDF.

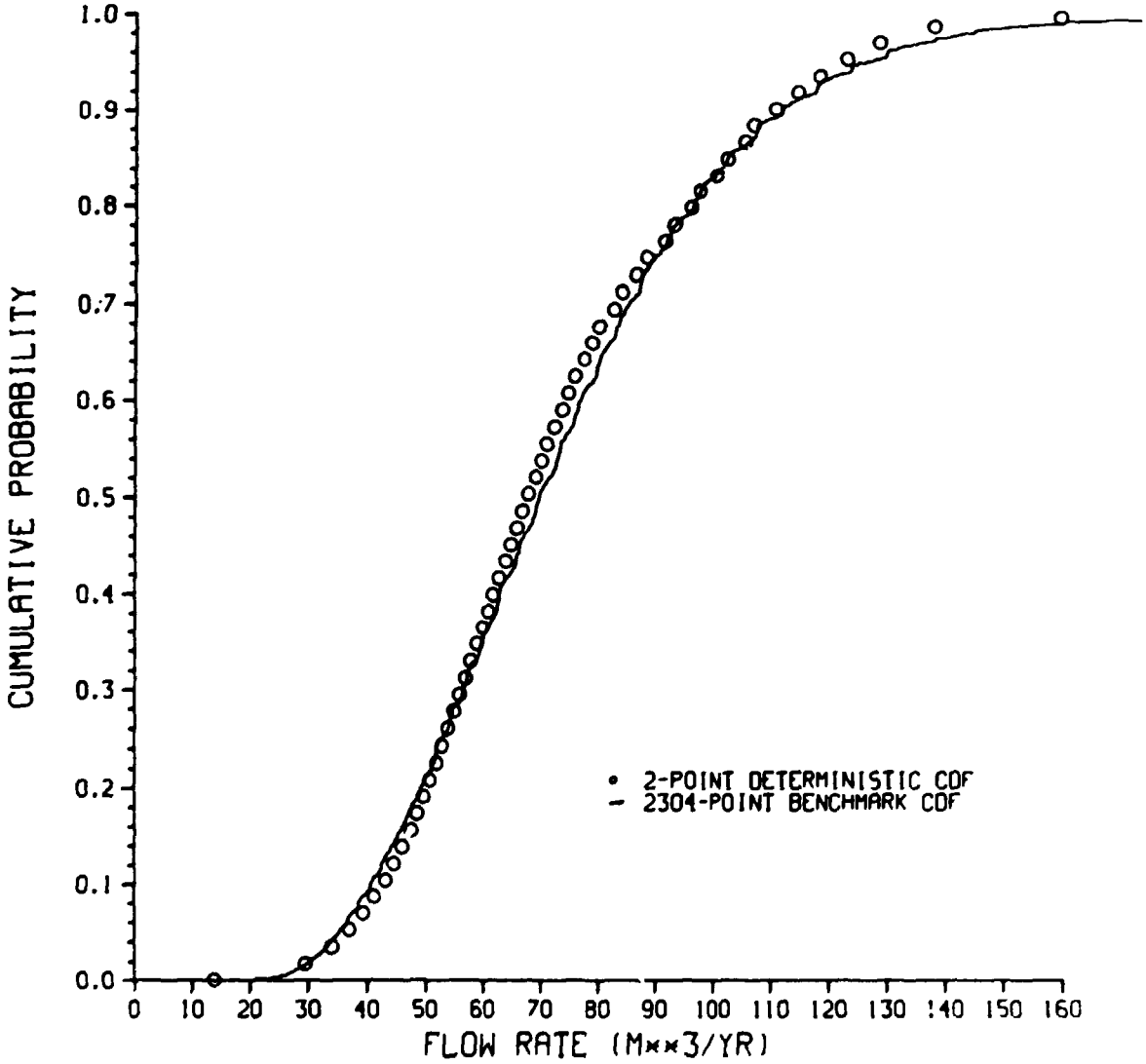

Fig. 3. Comparison of the Deterministic CDF Based on an Extrapolation from Two Points to 2304 Discrete Meshes Compared to the 2304-Point Benchmark CDF. (For clarity, every fortieth point is plotted for the extrapolated values.)

VIII. CONCLUSIONS

The availability of derivative information gives a much more complete basis for creating an accurate response surface than does the use of response values alone. The response surface could be sampled in place of the original model, or the propagation of probabilities can be performed in a rigorous fashion over the response surface. The latter procedure results in a strictly deterministic method of probability propagation. The sample problem results show that simple linear extrapolation from two space points produces a CDF of the response of interest that more closely matches a benchmark 2304-point CDF than does the CDF based upon a 50-point LHS design matrix. The reduction in model runs by a factor of 25 and the increased accuracy in calculating the CDF of the response of interest are strong evidence that a substantial savings in computational cost is possible using derivative information. This reduction is offset by the additional cost of calculating derivatives, but the deterministic calculation of model derivatives has been shown in the published literature to be both feasible and cost efficient for large-scale computer models. The availability of automated precompilers for adding derivative-taking capability to existing models makes the DUA approach even more practical. The strong analytical foundations of propagating probabilities deterministically is another desirable feature of the DUA approach.

ACKNOWLEDGEMENTS

The author wishes to express his appreciation to E. M. Oblow for the many stimulating conversations that lead to this paper. R. E. Maerker performed the calculations on which the plotted results are based, as well as provided valuable insight into the extrapolation schemes. The author also wishes to thank F. G. Pin, C. R. Weisbin, and F. C. Maienschein for their continued support of this work. The patience and expert document production by K. S. Harber are greatly appreciated.

REFERENCES

1. Thomas, R. E., "Uncertainty Analysis," ONWI-380, Office of Nuclear Waste Isolation, Battelle Memorial Institute, Columbus, Ohio (1982).

2. Harper, W. V., "Sensitivity/Uncertainty Analysis Techniques for Nonstochastic Computer Codes," ONWI-444, Office of Nuclear Waste Isolation, Battelle Memorial Institute, Columbus, Ohio (1983).

3. Hicks, C. R., *Fundamental Concepts in the Design of Experiments*, Holt, Rinehart, and Winston, New York, NY (1973).

4. Iman, R. L., and W. J. Conover, "A Distribution Free Approach to Inducing Rank Correlation Among Input Variables," *Communications in Statistics* B11(3), 311-334 (1982).

5. Sutcliffe, W. G., "Uncertainty Analysis: An Illustration from Nuclear Waste Package Development," *Nucl. and Chem. Waste Manag.* 5, 131-140 (1984).

6. Myers, R. H., *Response Surface Methodology*, Allyn and Bacon, Inc., Boston, Massachusetts (1971).

7. Tomovic, R. and M. Vukobratovic, *General Sensitivity Theory*, Elsevier North-Holland, Inc., New York (1972).

8. Greenspan, E., "Developments in Perturbation Theory," *Advances in Nuclear Science and Technology*, Vol. 9, Academic Press, New York (1976).

9. Oblow, E. M., *Nucl. Sci. Eng.* 68, 322 (1978).

10. Frank, P. M., *Introduction to System Sensitivity Theory*, Academic Press, New York (1978).

11. Cacuci, D. G., *J. Math. Phys.* (NY), 22, 2794 (1981).

12. Cacuci, D. G., *J. Math. Phys.* (NY), 22, 2803 (1981).

13. Oblow, E. M., "An Automated Procedure for Sensitivity Analysis Using Computer Calculus," ORNL/TM-8776, Oak Ridge National Laboratory, Oak Ridge, Tennessee (1983).

14. Oblow, E. M., "GRESS, Gradient-Enhanced Software System - Version D User's Guide," ORNL/TM-9658, Oak Ridge National Laboratory, Oak Ridge, Tennessee (1985).

15. Pin, F. G., E. M. Oblow, J. E. Horwedel, and J. L. Lucius, "ADGEN: An Automated Adjoint Code Generator for Large-Scale Sensitivity Analysis," *Trans. Am. Nucl. Soc.* 55, 311 (1987).

16. Harper, W. V. and S. K. Gupta, "Sensitivity/Uncertainty Analysis of a Borehole Scenario Comparing Latin Hypercube Sampling and Deterministic Sensitivity Approaches," BMI/ONWI-516, Office of Nuclear Waste Isolation, Battelle Memorial Institute, Columbus, Ohio (1983).

APPENDIX A - FITTING PROCEDURES USING DERIVATIVE INFORMATION

The DUA method replaces the computer model with an analytical expression by relating the response of interest as calculated by the computer model to the parameter values by fitting techniques that incorporate knowledge of the partial derivatives of the response with respect to the parameters of interest. Before discussing fitting techniques that use derivative information, a review of fitting strategies using only response values will be helpful.

A. Linear Least Squares Fitting of Response Values.

Let the computer model response of interest r be represented by a function f of the k parameters such that $r=f(x_1,x_2,\ldots x_k)$. Defining the parameter space vector at point i as $\mathbf{x}_i = (x_1,x_2,\ldots x_k)$, the corresponding observed (calculated) response r_i is given by

$$r_i = f(\mathbf{x}_i) . \tag{A.1}$$

Note that r_i refers to a chosen response of interest, r, at the parameter space point i and not to another response. For most computer models, the functional form of f is generally too complex to express in closed form. A response function $g(\mathbf{x}_i)$ is sought which closely approximates f over some domain of the parameter space. Thus the observed response r_i is to be related to $\mathbf{x}_i$ by the relation

$$r_i \approx g(\mathbf{x}_i) . \tag{A.2}$$

The basic procedure of linear least-squares fitting begins by choosing $g(x_i)$ to be a linear combination of simple functions of the parameters. Denoting these simple expansion functions as $g_j(x)$, then

$$g(x_i) = \sum_{j=1}^{J} c_j g_j(x) \tag{A.3}$$

where the selection of the g_j and J determines the order and completeness of the expansion and the c_j are the expansion coefficients to be determined. Eq. (A.3) becomes

$$r_i = \sum_{j=1}^{J} c_j g_j(x_i) \quad . \tag{A.4}$$

Defining

$$r = (r_1, r_2, \ldots, r_N)^T$$

$$G(x) = \begin{bmatrix} g_1(x_1) & \cdots & g_J(x_1) \\ \vdots & & \vdots \\ g_1(x_N) & & g_J(x_N) \end{bmatrix}$$

and $$c = (c_1, c_2, \ldots, c_J)^T \quad ,$$

then Eq. (A.4) can be written in matrix form as

$$r = G\, c \quad . \tag{A.5}$$

Note that r is an N-dimensional vector of the values of a single response for N parameter space points, not a vector of different responses. The matrix G is an NxJ matrix whose elements are all known, being simple functions of preselected parameter space points. The vector c is a J-dimensional vector of constants to be determined. If J=N, and if G is

non-singular, the vector c can be solved for directly and Eq. (A.3) will yield a function g(x) that exactly reproduces r_i for the corresponding x_i. As an example, for K=2, N=6, and J=6, a fit of six points in a two-dimensional parameter space can be performed by expanding the response in a perfect quadratic form such that

$$r = c_1 + c_2x_1 + c_3x_2 + c_4x_1^2 + c_5x_2^2 + c_6x_1x_2 \quad . \tag{A.6}$$

In general, the relative values of N, J, and K will determine the order and completeness of the expansion defined by Eq. (A.3). Often, high-order expansions do not prove to be practical because g(x) exhibits unrealistic behavior "between" and "outside of" the selected parameter space points x_i. Therefore, for many applications a low-order expansion of g(x) is chosen and N will be greater than J. For this case Eq. (A.5) is overdetermined and a least-squares fitting technique must be used. In fact, "fitting techniques" are often associated with solving for c assuming Eq. (A.5) is overdetermined. Least-squares fitting determines the vector c that gives the best fit of the members of r to the space points x_i, i=1,N, in a least-squares sense. Defining the sum of the squares of the differences between r and Gc as L,

$$L = (r - Gc)^T (r - Gc) \tag{A.7a}$$

or

$$L = r^Tr - 2c^TG^Tr + c^TG^TGc \quad . \tag{A.7b}$$

The vector c that results in L being a minimum is given by

$$\frac{\partial L}{\partial c} = -2G^Tr + 2\ G^TGc \quad . \tag{A.8}$$

Setting $\frac{\partial L}{\partial c} = 0$,

$$c = (G^T G)^{-1} G^T r \quad . \tag{A.9}$$

Equation (A.9) defines the basic fitting solution of the least-squares technique when fitting response values.

B. Fitting Schemes Using Derivative Information.

One approach to fitting values and derivatives simultaneously is to differentiate Eq. (A.5) with respect to a parameter of interest,

$$\frac{\partial r}{\partial x_k} = G_k c \tag{A.10}$$

where x_k is a parameter of interest and $G_k = G_k(x) = \frac{\partial G}{\partial x_k}$, G being the same matrix as the one in Eq.(A.5). Defining $p_k = \frac{\partial r}{\partial x_k}$,

$$p = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_k \end{bmatrix}$$

and $G' = G'(x) = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_k \end{bmatrix}$,

then

$$p = G'c \quad . \tag{A.11}$$

Note that p, the vector of observed (calculated) partial derivatives of the response with respect to all parameters of interest at all of the parameter space points, is of length NK and G' is an NKxJ matrix. If the derivatives are expanded to first order in the parameters, J=K+1. Thus, for most cases

NK>J and Eq. (A.11) is overdetermined, and a least squares fitting procedure could be used. However, the objective for this particular procedure is to solve Eqs. (A.5) and (A.11) simultaneously. The G matrix of response expansion function values and the G' matrix of the evaluated derivatives of these functions have the same number of columns (the derivative of a constant element of the G matrix is a constant equal to zero in the G' matrix). Thus Eqs. (A.5) and (A.11) can be augmented to produce the single relationship

$$\frac{r}{p} = \frac{G}{G'} c \quad . \tag{A.12}$$

Eq. (A.12) is the system of equations that describe the fitting of both response values and derivatives. There are N responses and NK derivatives to be fitted. The vector $\frac{r}{p}$ is N(K+1) x 1 and the matrix $\frac{G}{G'}$ is N(K+1) x J. The solution of the J length vector c based upon Eq. (A.12) is discussed for two cases:

1. The matrix $\frac{G}{G'}$ is square.

For this case, J=N(K+1) and the solution for c is unique if $\left|\frac{G}{G'}\right|$ is nonsingular. In such a case, all N response values and NK derivatives are reproduced for the response surface defined by Eq. (A.5) where c is the solution to Eq. (A.12). However, the $\left|\frac{G}{G'}\right|$ matrix has been found to be singular for some simple situations. For example, consider the quadratic expansion in two-parameter space (K=2) given by Eq. (A.6). Fitting the responses and derivatives of two space points (N=2), then N(K+1)=6=J, and c can be solved for exactly if $\left|\frac{G}{G'}\right|$ is nonsingular. For this case, $\left|\frac{r}{p}\right|$ and $\left|\frac{G}{G'}\right|$ are given by

$$\frac{\mathbf{r}}{\mathbf{p}} = \begin{bmatrix} r_1 \\ r_2 \\ \frac{\partial r}{\partial x_1}\,1 \\ \frac{\partial r}{\partial x_1}\,2 \\ \frac{\partial r}{\partial x_2}\,1 \\ \frac{\partial r}{\partial x_2}\,2 \end{bmatrix} \tag{A.13a}$$

and

$$\frac{\mathbf{G}}{\mathbf{G}'} = \begin{bmatrix} 1 & x_1(1) & x_2(1) & x_1^2(1) & x_2^2(1) & x_1(1)x_2(1) \\ 1 & x_1(2) & x_2(2) & x_1^2(2) & x_2^2(2) & x_1(2)x_2(2) \\ 0 & 1 & 0 & 2x_1(1) & 0 & x_2(1) \\ 0 & 1 & 0 & 2x_1(2) & 0 & x_2(2) \\ 0 & 0 & 1 & 0 & 2x_2(1) & x_1(1) \\ 0 & 0 & 1 & 0 & 2x_2(2) & x_1(2) \end{bmatrix} \tag{A.13b}$$

The determinant of $\frac{\mathbf{G}}{\mathbf{G}'}$ can be shown to be algebraically equal to zero with a matrix rank of five for this case. In particular, for this case, $\mathbf{r} = h\left(\frac{\partial r}{\partial x_1}, \frac{\partial r}{\partial x_2}\right)$. Physically interpreted, for this case all the derivatives are reproduced by the fit, and integration of the derivatives introduces only one constant back into the form of the response expansion. Therefore only one response point can be fit. Or both responses could be fit and differentiating the expressions only introduces enough degrees of freedom to fit 3 of the 4 known derivatives. Interestingly, had a functional form other than x_1x_2 been chosen for $g_6(\mathbf{x})$, the determinant would not have been zero.

2. Eq. (A.12) is overdetermined.

If Eq. (A.12) is overdetermined, some fitting procedure must be used that will lead to the best fit of the responses while incorporating the

derivative information. But both responses and derivatives should not be fitted using Eq. (A.12) as is because the response values have units different than the derivatives. In fact, the derivatives most likely have units different from each other and even the derivatives defined by Eq. (A.11) must be normalized if the system of equations is overdetermined and a fitting of the derivatives is to be performed. The fitting of the sensitivities would eliminate the normalization problem since the sensitivities are unitless, but the response function itself would not be addressed. The main objective is still to identify a response surface that would closely reproduce the observed vector of responses r. One possible approach is to choose the functions g_i that make up the matrix G and solve Eq. (A.9) for c, calculate the derivatives based upon differentiation of Eq. (A.10), and calculate the sum of the squares of differences in the calculated and observed values of the sensitivities. The process could be repeated for various choices of the functions g_i until a set of g_i are identified that closely match the responses and the sensitivities. Yet another method would be to treat some of the derivative equations as constraints and solve Eq. (A.9) with these constraint conditions. Numerous other possibilities exist for treating Eq. (A.12) for the case where the equations are overdetermined, but no superior method has been identified and future research in this area is needed.

C. Local Fitting and Expansion Based on Derivative Information.

A local fit of the response values and derivatives using a low-order function may be more desirable than a global fit using a higher order function to fit a large portion of the data for two reasons: (1) the inversion of the matrix $G = G/G'$ in Eq. (A.9) is expensive for large values of NK,

and (2) a higher order fit involving many response points may result in a very radically behaved function in the parameter space not near the fitted points. But another basic difference exists between conventional fitting to response values and fitting using derivative information that makes local fitting very attractive. Use of derivative information permits a fit to a fewer number of points for a fixed order of expansion as compared to conventional fitting of just the data. For example, the quadratic expansion in two parameters defined by Eq. (A.6) required six data points to uniquely determine the expansion coefficients based on fitting six response values. However, with the use of derivative information, only two points are required to uniquely determine five of the six expansion coefficients (see discussion in previous section). In the limit of local fitting by linear extrapolation from reference space points, the fitting procedure is a series of linear expansions.

APPENDIX B - SAMPLE PROBLEM DATA FROM REFERENCE

Table B.1.* Input Parameter Probability Distributions.

Input Parameter	Range	Distribution
r_w	0.05 to 0.15 m	Normal (μ = 0.10, σ = 0.0161812)†
r	100 to 50,000 m	Lognormal (μ' = 7.71. σ' = 1.0056)**
T_u	63,070 to 115,600 m^2/yr	Uniform
H_u	990 to 1,110 m	Uniform
T_l	63.1 to 116 m^2/yr	Uniform
H_l	700 to 820 m	Uniform
L	1,120 to 1,680 m	Uniform
K_w	9,855 to 12,045 m/yr	Uniform

*From Ref. 16.
†μ, σ are the mean and standard deviation, respectively, of r_w.
**μ', σ' are the mean and standard deviation, respectively, of the ln (r) (which is normally distributed).

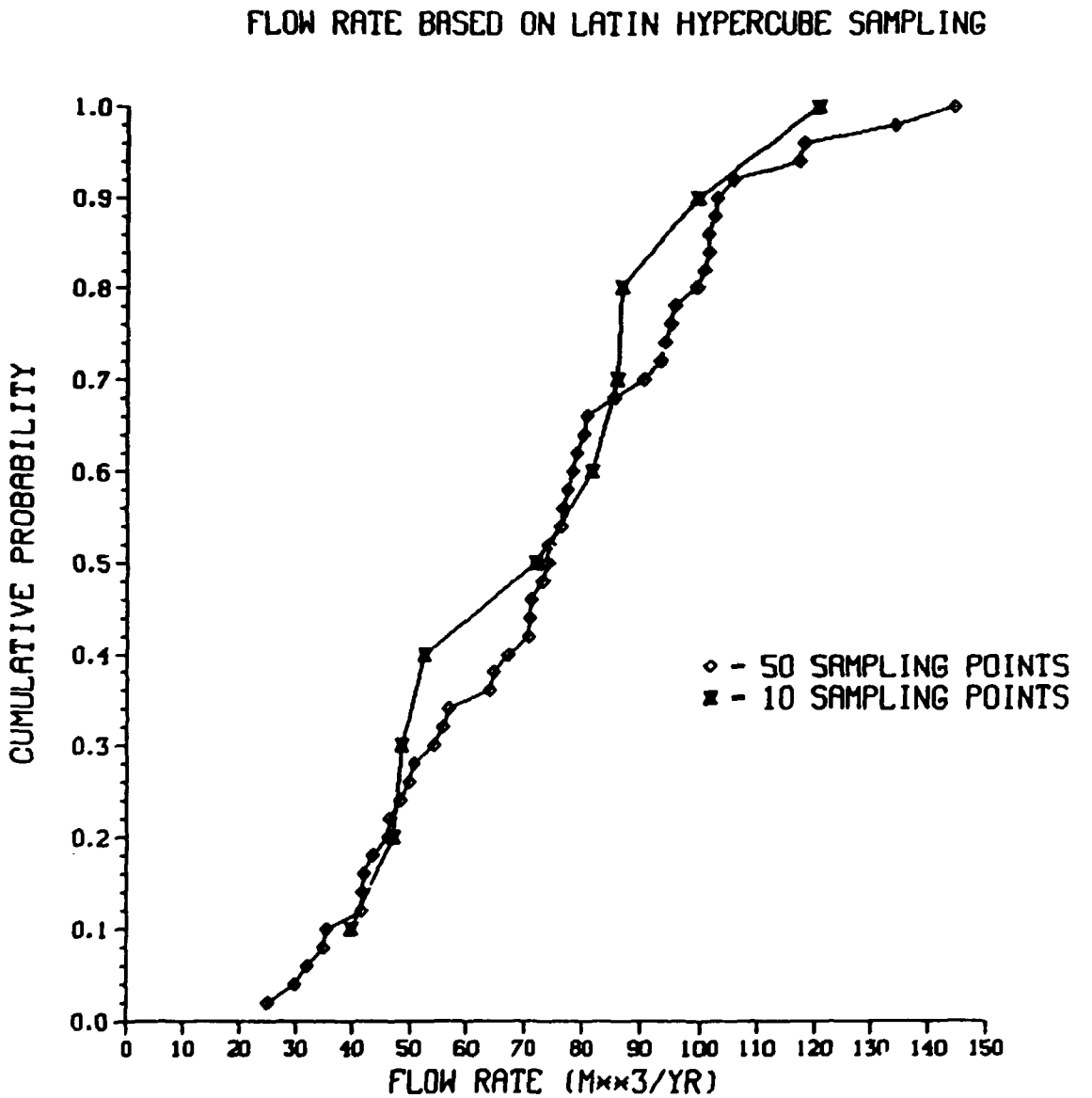

Fig. B.1. Cumulative Distribution Functions for 10-Point and 50-Point LHS Design Matrices. From Data Published in Ref. 16.

Table B.2.* LHS Boreflow Results, N = 10.

Run No.	r_w	r	T_u	H_u	T_l	H_l	L	K_w	H_{wl}	H_{wu}	Q
1	0.8609E-01	2948.0	0.8337E+05	1044.0	107.0	783.0	1250.0	0.1001E+05	783.754	1044.0	48.5247
2	0.1050	5194.0	0.8840E+05	1093.0	67.7	788.0	1181.0	0.1126E+05	790.538	1093.0	99.8815
3	0.1180	1358.0	0.9471E+05	993.0	98.6	758.0	1466.0	0.1174E+05	759.236	992.999	81.8886
4	0.9050E-01	240.0	0.1091E+06	1037.0	81.2	811.0	1534.0	0.1052E+05	811.614	1037.0	39.7706
5	0.9287E-01	1861.0	0.9219E+05	1101.1	103.0	711.0	1575.0	0.1075E+05	712.101	1101.0	71.9224
6	0.1250	1165.0	0.1015E+06	1055.0	86.5	734.0	1325.0	0.1023E+05	736.033	1055.0	120.885
7	0.9778E-01	902.0	0.6798E+05	1072.0	76.9	715.0	1380.0	0.1116E+05	716.631	1072.0	86.3202
8	0.1100	2616.0	0.6944E+05	1085.0	113.0	799.0	1429.0	0.1147E+05	800.233	1085.0	86.8866
9	0.1000	14690.0	0.7427E+05	1009.0	70.8	768.0	1672.0	0.1043E+05	769.257	1009.0	46.9832
10	0.7457E-01	8017.0	0.1153E+06	1025.0	90.6	738.0	1141.0	0.1202E+05	739.071	1025.0	52.6204

*From Ref. 16.

Table B.3.* LHS Boreflow Results, N = 50.

Run No.	r_w	r	T_u	H_u	T_l	H_l	L	K_w	H_{wl}	H_{wu}	Q
1	0.1200	307.0	0.7133E+05	1070.0	87.4	704.0	1504.0	0.1080E+05	705.691	1070.0	118.347
2	0.1210	7787.0	0.8976E+05	1088.0	88.5	812.0	1672.0	0.1071E+05	813.610	1089.0	80.8426
3	0.1040	2684.0	0.7764E+05	1092.0	91.4	782.0	1245.0	0.1120E+05	783.667	1092.0	94.2502
4	0.7928E-01	455.0	0.6463E+05	1046.0	63.1	801.0	1556.0	0.1147E+05	801.776	1046.0	35.5482
5	0.1170	480.0	0.9364E+05	1034.0	64.2	730.0	1430.0	0.1129E+05	732.114	1034.0	102.499
6	0.1010	1508.0	0.1029E+06	1006.0	114.0	737.0	1590.0	0.1051E+05	737.762	1006.0	56.8220
7	0.1130	16040.0	0.1038E+06	1004.0	103.0	714.0	1352.0	0.1179E+05	715.848	1004.0	100.801
8	0.1120	1121.0	0.8194E+05	1036.0	105.0	702.0	1282.0	0.1008E+05	703.439	1036.0	103.045
9	0.1000	24370.0	0.1060E+06	1023.0	84.1	711.0	1480.0	0.1177E+05	712.819	1023.0	77.4957
10	0.8917E-01	3967.0	0.1044E+06	1063.0	82.1	750.0	1154.0	0.9943E+04	751.392	1063.0	67.0667
11	0.9504E-01	1798.0	0.7483E+05	997.0	95.3	753.0	1138.0	0.1067E+05	754.063	996.999	64.6362
12	0.7425E-01	4686.0	0.1081E+06	1017.0	104.0	778.0	1311.0	0.1014E+05	778.540	1017.0	31.9442
13	0.9985E-01	1985.0	0.7907E+05	1044.0	71.2	725.0	1211.0	0.1106E+05	727.006	1044.0	90.6787
14	0.9564E-01	750.0	0.9121E+05	1025.0	113.0	793.0	1544.0	0.1156E+05	793.629	1025.0	49.7792
15	0.1150	6951.0	0.9751E+05	1008.0	69.2	799.0	1523.0	0.9870E+04	800.415	1008.0	55.8928
16	0.9209E-01	11710.0	0.1019E+06	1109.0	80.9	723.0	1525.0	0.1198E+05	724.859	1109.0	80.3990
17	0.1100	274.0	0.1120E+06	1039.0	112.0	745.0	1257.0	0.1075E+05	746.058	1039.0	95.2331
18	0.7821E-01	3134.0	0.8075E+05	1009.0	73.7	734.0	1651.0	0.1100E+05	734.803	1009.0	35.1060
19	0.7562E-01	2346.0	0.1071E+06	993.0	81.3	810.0	1303.0	0.1185E+05	810.603	993.0	29.7998
20	0.1080	2134.0	0.1134E+06	1072.0	103.0	819.0	1392.0	0.1116E+05	820.131	1072.0	73.9938
21	0.1030	2869.0	0.7306E+05	1014.0	110.0	813.0	1623.0	0.1017E+05	813.620	1014.0	41.8485
22	0.1040	10170.0	0.8365E+05	1079.0	100.0	727.0	1453.0	0.1171E+05	728.754	1079.0	95.9132
23	0.8143E-01	1359.0	0.6610E+05	1053.0	86.9	747.0	1500.0	0.1141E+05	747.861	1053.0	48.3514
24	0.1300	6148.0	0.9541E+05	1058.0	95.7	781.0	1179.0	0.1086E+05	783.404	1058.0	134.290
25	0.9691E-01	840.0	0.9264E+05	992.0	75.2	816.0	1226.0	0.1204E+05	816.973	991.999	50.7137

*From Ref. 16; continued on next page.

Table B.3.* (Continued).

Run No.	r_w	r	T_u	H_u	T_l	H_l	L	K_w	H_{wl}	H_{wu}	Q
26	0.9251E-01	664.0	0.1009E+06	1099.0	90.2	785.0	1379.0	0.1166E+05	786.114	1099.0	71.1289
27	0.9083E-01	2403.0	0.8688E+05	1104.0	107.0	795.0	1602.0	0.1089E+05	795.822	1104.0	54.2967
28	0.8480E-01	2023.0	0.9936E+05	1088.0	106.0	720.0	1235.0	0.1134E+05	721.152	1088.0	76.0979
29	0.1110	985.0	0.9662E+05	1052.0	77.4	754.0	1636.0	0.1093E+05	755.434	1052.0	76.6925
30	0.1180	1720.0	0.1149E+06	1012.0	76.3	716.0	1406.0	0.1161E+05	718.123	1012.0	106.151
31	0.8566E-01	4939.0	0.9829E+05	995.0	69.8	742.0	1298.0	0.1036E+05	743.158	994.999	46.3359
32	0.1050	1293.0	0.8778E+05	1101.0	108.0	797.0	1478.0	0.1032E+05	798.017	1101.0	73.2740
33	0.8969E-01	3494.0	0.1139E+06	1042.0	101.0	790.0	1601.0	0.1061E+05	790.701	1042.0	42.0874
34	0.1090	1894.0	0.6707E+05	1061.0	84.4	789.0	1437.0	0.1126E+05	790.457	1061.0	79.1258
35	0.1250	4345.0	0.6927E+05	1021.0	78.6	762.0	1669.0	0.1135E+05	763.818	1021.0	85.8513
36	0.9419E-01	4163.0	0.9238E+05	1107.0	66.7	804.0	1123.0	0.1049E+05	806.000	1107.0	78.3645
37	0.9908E-01	779.0	0.1102E+06	1084.0	65.3	758.0	1569.0	0.1000E+05	759.395	1084.0	63.8048
38	0.1150	5981.0	0.7600E+05	1077.0	92.9	807.0	1151.0	0.1029E+05	808.853	1077.0	99.5990
39	0.1270	1220.0	0.7380E+05	1049.0	91.6	769.0	1207.0	0.1007E+05	770.873	1049.0	117.576
40	0.6327E-01	3603.0	0.7221E+05	1032.0	73.0	740.0	1459.0	0.9940E+04	740.596	1032.0	24.9673
41	0.8654E-01	1035.0	0.8923E+05	1000.0	99.0	774.0	1336.0	0.1097E+05	774.657	999.999	43.5338
42	0.8840E-01	1412.0	0.8486E+05	1083.0	111.0	775.0	1171.0	0.1152E+05	776.029	1083.0	74.1388
43	0.1380	2896.0	0.8553E+05	1028.0	79.1	731.0	1267.0	0.1041E+05	733.895	1028.0	144.571
44	0.9766E-01	3368.0	0.7792E+05	1075.0	97.0	709.0	1632.0	0.1057E+05	710.214	1075.0	70.7905
45	0.1070	910.0	0.8256E+05	1067.0	71.9	765.0	1377.0	0.1192E+05	766.872	1067.0	93.4468
46	0.9668E-01	5456.0	0.6340E+05	1066.0	58.4	718.0	1191.0	0.1190E+05	719.797	1066.0	101.575
47	0.1070	1569.0	0.1083E+06	1096.0	67.7	771.0	1328.0	0.1161E+05	773.289	1096.0	101.476
48	0.8296E-01	8875.0	0.6732E+05	1056.0	85.6	757.0	1420.0	0.1024E+05	758.000	1056.0	46.4636
49	0.1020	2523.0	0.7028E+05	1031.0	115.0	763.0	1365.0	0.1108E+05	763.992	1031.0	70.8401
50	0.7149E-01	605.0	0.1109E+06	1093.0	93.6	706.0	1562.0	0.1044E+05	706.638	1093.0	41.4624

*From Ref. 16.

Table B.4.* Derivatives and Sensitivities of the Response Q with Respect to Each Parameter for the 10-Point LHS Design Matrix.

Parameter	Value	$\frac{\partial R}{\partial \alpha}$	$\frac{\partial R}{\partial \alpha}\frac{\alpha}{R}$
1	R=Q = 48.52464	H_{w1} = 783.7536	H_{wu} = 1043.999
T_u	83370.00	0.2156935E-08	0.3705821E-05
H_u	1044.00	0.1859182	4.000000
r	2948.00	-0.4557747E-05	-0.2758951E-03
r_w	0.08609	1124.082	1.994291
T_l	107.00	0.1309451E-02	0.2887424E-02
H_l	783.00	-0.1859182	-3.000000
K_w	10010.00	0.4833502E-02	0.9971089
L	1250.00	-0.3870748E-01	-0.9971089
2	R=Q = 99.88144	H_{w1} = 790.5381	H_{wu} = 1092.998
T_u	88400.00	0.7200674E-08	0.6372952E-05
H_u	1093.00	0.3274801	3.583607
r	5194.00	-0.1481603E-04	-0.7704580E-03
r_w	0.105	1886.844	1.983538
T_l	67.70	0.1227723E-01	0.8321550E-02
H_l	788.00	-0.3274801	-2.583607
K_w	11260.00	0.8796593E-02	0.9916721
L	1181.00	-0.8386930E-01	-0.9916721
3	R=Q = 81.88855	H_{w1} = 759.2360	H_{wu} = 992.9987
T_u	94710.00	0.4734320E-08	0.5475582E-05
H_u	993.00	0.3484619	4.225532
r	1358.00	-0.3395267E-04	-0.5630546E-03
r_w	0.118	1380.910	1.989868
T_l	98.60	0.4368129E-02	0.5259557E-02
H_l	758.00	-0.3484619	-3.225532
K_w	11740.00	0.6938450E-02	0.9947350
L	1466.00	-0.5556440E-01	-0.9947350
4	R=Q = 39.77062	H_{w1} = 811.6145	H_{wu} = 1037.000
T_u	109100.00	0.7377022E-09	0.2023687E-05
H_u	1037.00	0.1759762	4.588496
r	240.00	-0.5719953E-04	-0.3451766E-03
r_w	0.0905	877.3659	1.996489
T_l	81.20	0.1331737E-02	0.2719018E-02
H_l	811.00	-0.1759762	-3.588496
K_w	10520.00	0.3770191E-02	0.9972790
L	1534.00	-0.2585555E-01	-0.9972790

*From Ref. 16.

Table B.4. (Continued).

Parameter	Value	$\frac{\partial R}{\partial \alpha}$	$\frac{\partial R}{\partial \alpha}\frac{\alpha}{R}$
5	R=Q = 71.92241	H_{wl} = 712.1008	H_{wu} = 1100.999
T_u	92190.00	0.2460308E-08	0.3153618E-05
H_u	1101.00	0.1844164	2.823077
r	1861.00	-0.1102517E-04	-0.2852775E-03
r_w	0.09287	1544.613	1.994485
T_l	103.00	0.1970982E-02	0.2822641E-02
H_l	711.00	-0.1844164	-1.823077
K_w	10750.00	0.6671550E-02	0.9971741
L	1575.00	-0.4553598E-01	-0.9971741
6	R=Q = 120.8852	H_{wl} = 736.0329	H_{wu} = 1054.998
T_u	101500.00	0.6427935E-08	0.5397149E-05
H_u	1055.00	0.3765893	3.286604
r	1165.00	-0.7195967E-04	-0.6034929E-03
r_w	0.125	1922.578	1.988021
T_l	86.50	0.8850572E-02	0.6333071E-02
H_l	734.00	-0.3765893	-2.286604
K_w	10230.00	0.1174183E-01	0.9936614
L	1325.00	-0.9065580E-01	-0.9936614
7	R=Q = 86.32016	H_{wl} = 716.6310	H_{wu} = 1071.998
T_u	67980.00	0.6562483E-08	0.5168174E-05
H_u	1072.00	0.2417931	3.002801
r	902.00	-0.4794403E-04	-0.5009897E-03
r_w	0.09778	1757.957	1.991343
T_l	76.90	0.5128352E-02	0.4568693E-02
H_l	715.00	-0.2417931	-2.002801
K_w	11160.00	0.7699403E-02	0.9954261
L	1380.00	-0.6226474E-01	-0.9954261
8	R=Q = 86.88659	H_{wl} = 800.2331	H_{wu} = 1084.998
T_u	69440.00	0.8779252E-08	0.7016403E-05
H_u	1085.00	0.3037992	3.793706
r	2616.00	-0.1423474E-04	-0.4285827E-03
r_w	0.11	1573.013	1.991463
T_l	113.00	0.3315279E-02	0.4311673E-02
H_l	799.00	-0.3037993	-2.793706
K_w	11470.00	0.7542402E-02	0.9956813
L	1429.00	-0.6053978E-01	-0.9956813

Table B.4. (Continued).

Parameter	Value	$\frac{\partial R}{\partial \alpha}$	$\frac{\partial R}{\partial \alpha}\frac{\alpha}{R}$
9	R=Q = 46.98318	H_{wl} = 769.2566	H_{wu} = 1008.999
T_u	74270.00	0.31442505E-08	0.4970363E-05
H_u	1009.00	0.1949509	4.186722
r	14690.00	-0.1402964E-05	-0.4386581E-03
r_w	0.10	934.7742	1.989593
T_l	70.80	0.3460010E-02	0.5213967E-02
H_l	768.00	-0.1949509	-3.186722
K_w	10430.00	0.4481110E-02	0.9947811
L	1672.00	-0.2795333E-02	-0.9947811
10	R=Q = 52.62038	H_{wl} = 739.0709	H_{wu} = 1024.999
T_u	115300.00	0.1338124E-08	0.2932052E-05
H_u	1025.00	0.1833462	3.571428
r	8017.00	-0.2115667E-05	-0.3223333E-03
r_w	0.07457	1406.067	1.992582
T_l	90.60	0.2167197E-02	0.3731408E-02
H_l	738.00	-0.1833462	-2.571428
K_w	12020.00	0.4361387E-02	0.9962656
L	1141.00	-0.4594555E-01	-0.9962656

ORNL-6428
Dist. Category: UC-70 (Waste)

INTERNAL DISTRIBUTION

1. R. G. Alsmiller, Jr.
2. S. M. Babcock
3. D. L. Barnett
4. M. Beckerman
5. B. L. Burks
6. A. G. Croff
7. J. R. Einstein
8. C. W. Glover
9. E. C. Halbert
10. J. E. Horwedel
11. W. R. Hamel
12. D. M. Jollay
13. J. P. Jones
14-18. R. E. Maerker
19. F. C. Maienschein
20. R. C. Mann
21. J. R. Merriman
22. L. J. Mezga
23-27. E. M. Oblow
28. L. E. Parker
29-33. F. G. Pin
34. D. B. Reister
35. T. H. Row
36. S. H. Stow
37. L. D. Trowbridge
38-42. C. R. Weisbin
43-47. B. A. Worley
48. R. Q. Wright
49. A. Zucker
50. J. J. Dorning (Consultant)
51. G. H. Golub (Consultant)
52. R. M. Haralick (Consultant)
53. D. Steiner (Consultant)
54. RSIC
55-56. Central Research Library
57. Division Advisory Committee
58. Laboratory Records, RC
59. ORNL Y-12 Technical Sec.
60. ORNL Patent Office
61-65. EP&MD Reports Office

EXTERNAL DISTRIBUTION

66. Office of Assistant Manager for Energy Research and Development, DOE-ORO, Oak Ridge, TN 37830

67. Sandy Perkins, Research and Waste Management Division, U.S. DOE-ORO, Oak Ridge, TN 37830

68. G. Appel, Salt Repository Project Office, 110 N. 25 Mile Ave., Hereford, TX 79045

Rockwell Hanford Co., P.O. Box 800, Richland, WA 99352

69. L. C. Brown
70. P. M. Clifton
71. C. DeFigh-Price
72. C. J. Geier
73. C. McLane
74. B. Sager
75. J. Sonnicksen
76. R. Wild
77. D. Wood

Battelle Memorial Institute, Office of Nuclear Waste Isolation, P.O. Box 2360, Hereford, TX 79045

78. W. A. Carbiener
79. V. Guvanasen
80. G. Gahimer
81. W. V. Harper
82. J. F. Kircher
83. G. McNulty
84. R. E. Thomas
85. ONWI Library

Pacific Northwest Laboratory, P.O. Box 999, Richland, WA 99352

86. M. Apted
87. C. Cole
88. P. Doctor
89. R. Erikson
90. M. Foley
91. A. Liebetrau
92. M. Piepho
93. S. Sneider

Sandia National Laboratories, P.O. Box 5800, Albuquerque, NM 87185

94. R. Cranwell
95. J. C. Helton
96. R. L. Iman

U.S. DOE-BWIPO, Richland, WA 99352

97. D. H. Dahlem
98. P. E. Lamont
99. M. K. Thompson

100. Donald Alexander, U.S. DOE, Civilian Radioactive Waste Management Program, 1000 Independence Ave., SW, Washington, DC 20585

101. R. E. Allan, Ontario Hydro, 700 University Ave., H18D16, Toronto, Ontario, M5G1X6, CANADA

102. C. Koplik, The Analytic Sciences Corporation, 1 Jacob Way, Reading, MA 10867

103. W. J. Kurzeka, Manager, Commercial Waste Tech. Unit, Rockwell Hanford Operations, P.O. Box 800, Richland, WA 99352

104. D. Langmuir, Department of Chemistry and Geochemistry, Colorado School of Mines, Golden, CO 80401

105. W. L. Lee, R. F. Weston, Inc., 2301 Research Blvd., Rockville, MD 20850

106. J. A. Lieberman, OTHA, Inc., P.O. Box 651, Glen Echo, MD 20812

107. J. O. Neff, U.S. Department of Energy, Richland Operations Office, 505 King Avenue, Columbus, OH 43201

108. T. H. Pigford, Department of Nuclear Engineering, University of California, Berkeley, CA 94720

109. J. Rhoderick, U.S. DOE/OCRWM, 1000 Independence Ave., SW, Washington, DC 20585

110. S. Simonson, Massachusetts Institute of Technology, Room 24-304, 77 Massachusetts Ave., Cambridge, MA 02139

111. J. Sykes, Dept. of Civil Engineering, University of Waterloo, Waterloo, Ontario, N2L3G1, CANADA

112. D. Veneziano, Massachusetts Institute of Technology, Room 1-382, 77 Massachusetts Ave., Cambridge, MA 02139

113-363. For distribution as shown in TID-4500, Distribution category UC-70 - Nuclear Waste Management