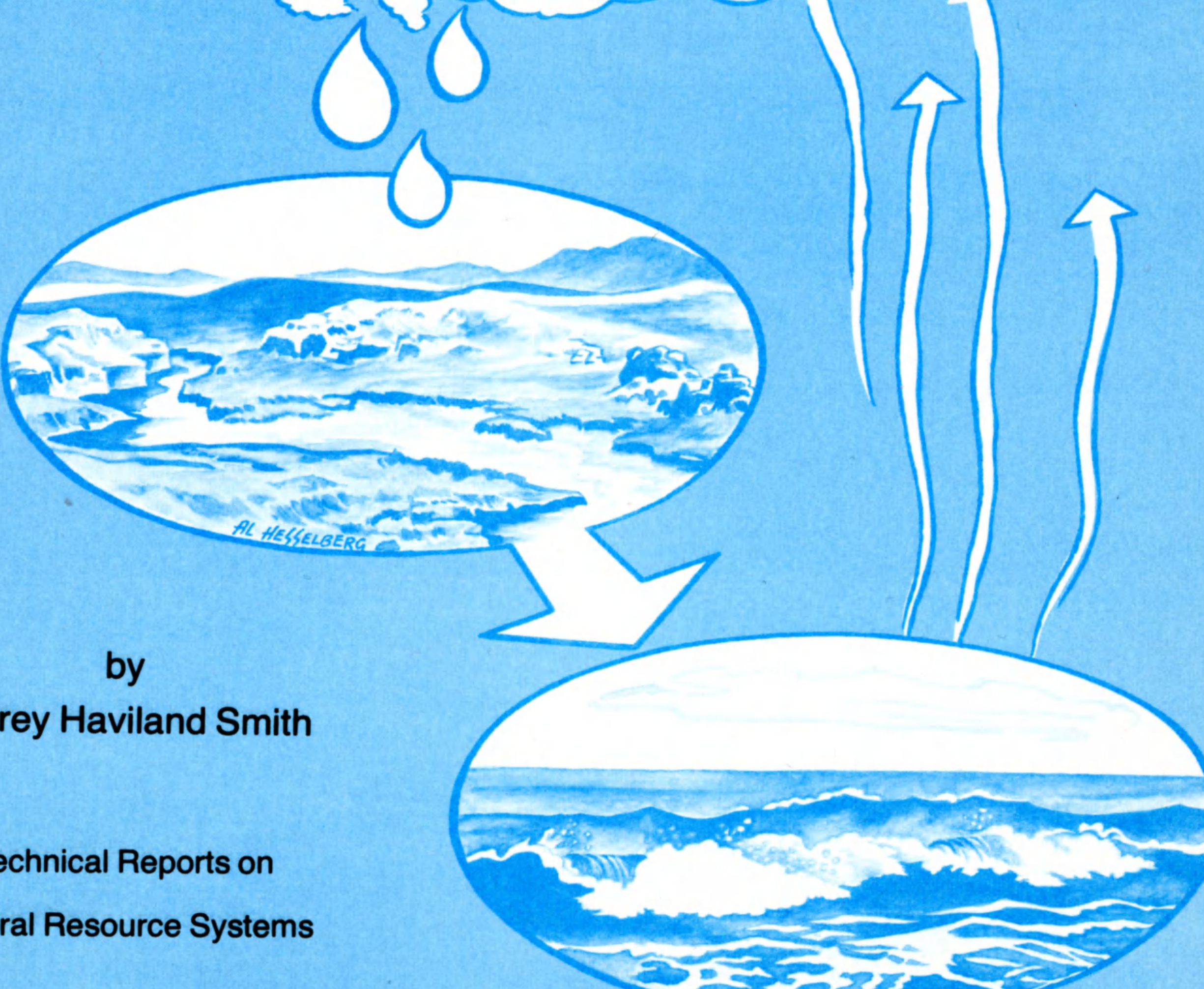
Technical Report No. 24

July, 1975

A Stochastic Sediment Yield Model for Bayesian Decision Analysis Applied to Multipurpose Reservoir Design



by

Jeffrey Haviland Smith

Technical Reports on

Natural Resource Systems

The University of Arizona
Tucson, Arizona 85721

A STOCHASTIC SEDIMENT YIELD MODEL FOR BAYESIAN DECISION ANALYSIS APPLIED TO MULTIPURPOSE RESERVOIR DESIGN

by

Jeffrey Haviland Smith

Natural Resource Systems Technical Report
No. 24

July 1975

Collaborative Effort Between
Department of Systems & Industrial Engineering
Department of Hydrology & Water Resources
School of Renewable Natural Resources

University of Arizona
Tucson, Arizona 85721

PREFACE

This report constitutes the Master of Science thesis of the same title completed by the author in January 1975, and accepted by the Department of Systems and Industrial Engineering. It is the result of two joint research projects, one on "Decision Analysis of Watershed Management Alternatives," supported in part by the United States Department of the Interior, Office of Water Resources Research, as authorized under the Water Resources Research Act of 1964, and one on "Practical Uses of Bayesian Decision Theory in Hydrology and Engineering," under a grant from the National Science Foundation (No. GK-35791).

Lucien Duckstein and Donald R. Davis
Jointly with the Departments of
Systems & Industrial Engineering
and Hydrology & Water Resources

Martin M. Fogel
School of Renewable Natural
Resources

ACKNOWLEDGMENTS

The author is grateful to Prof. V. K. Gupta for his willingness and continuous interest in the classification of certain concepts related to the stochastic processes involved. I would also like to thank Profs. Donald R. Davis and Lucien Duckstein for their constant, persistent interest and help in the theoretical aspects and developments of this thesis. Their encouragement was always appreciated. I would like to thank Prof. Martin M. Fogel for his insight between the models presented and the natural phenomena they were to represent. Also, I would like to thank him for the information regarding the statement that the rainfall mechanisms of the Atterbury and Charleston watersheds could be assumed identical. I would like to express my appreciation to the graduate students and faculty of the Department of Systems and Industrial Engineering who made my graduate experience the success that it was. I am grateful for the financial support provided by the National Science Foundation (Grant GK-35791), "Practical Uses of Bayesian Decision Theory in Hydrology and Engineering."

I would like to thank my wife, Ginny, whose understanding, perserverance, and cooperation made this work possible.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

ABSTRACT

This thesis presents a methodology for obtaining the optimal design capacity for sediment yield in multipurpose reservoir design. A stochastic model is presented for the prediction of sediment yield in a semi-arid watershed based on rainfall data and watershed characteristics. Uncertainty stems from each of the random variables used in the model, namely, rainfall amount, storm duration, runoff, peak flow rate, and number of events per season.

Using the stochastic sediment yield model for N-seasons, a Bayesian decision analysis is carried out for a dam site in southern Arizona. Extensive numerical analyses and simplifying assumptions are made to facilitate finding the optimal solution. The model has applications in the planning of reservoirs and dams where the effective lifetime of the facility may be evaluated in terms of storage capacity and of the effects of land management on the watershed. Experimental data from the Atterbury watershed are used to calibrate the model and to evaluate uncertainties associated with our knowledge of the parameters of the joint distribution of rainfall and storm duration used in calculating the sediment yield amount.

CHAPTER 1

THE EVENT-BASED SEDIMENT YIELD MODEL

This thesis will be concerned with determining the optimal size of a multipurpose reservoir based on the sedimentation characteristics of the watershed upstream from this reservoir. Thus, this problem is approached in three stages. More precisely, given the rainfall, storm duraticn, and economic characteristics of the watershed in question, the decision to be made is how large should the reservoir be designed to allow for the accumulation of sediment over the projected lifetime of the project. The first stage involves the development of an event-based stochastic model of the sediment yield process. Second, the extension of this model from a single event to a seasonal one and then to an N-season model to obtain the total sedimentation over an N-year period of lifetime. Finally, due to our uncertain knowledge about the parameters α and β of the bivariate distribution of rainfall and storm duration, a Bayesian analysis is carried out to assess the uncertainties associated with these variables and the effect on the decision variable. This goal is achieved through the use of the N-season model, together with the concepts of Bayesian decision theory to find the optimal design capacity for sediment yield.

In the planning and development of multipurpose reservoirs, it is necessary to make an allocation of storage space for sediment accumulation. This accumulation affects the design of the dam in that it influences the active storage capacity requirements, outlet sill elevations, recreational facilities, and backwater conditions (Borland and Miller, 1958). It is extremely important that sedimentation be accurately estimated as it is a very expensive process to remove the sediment once it has settled into the facility.

This work will rely on the theoretical and empirical works of others, particularly on sediment results by Wischmeier (1959, 1960), Wischmeier and Smith (1960, 1965), and Wischmeier, Smith, and Uhland (1958); on Bayesian decision theory by Davis (1971), Davis, Kisiel, and Duckstein (1972), Raiffa and Schlaifer (1961), De Groot (1970), and Feller (1966); and on rainfall by Duckstein, Fogel, and Kisiel (1972).

Background of the Problem

A sediment event is defined as the product of erosion from any runoff-producing storm over a watershed. In other words, a sediment event is of sufficient intensity to the point that runoff occurs. The independent variables are rainfall amount from the storm ($\underset{\sim}{x}_1$, in inches), the duration of the storm over time ($\underset{\sim}{x}_2$, in hours), and the time of concentration (a_2, in hours). The time of concentration is defined as the average length of time it takes for the precipitation to reach the outlet of the watershed. In addition to these variables, the model must take into account the individual characteristics of the watershed under study.

In the past, a variety of methods have been used to obtain estimates of sediment yield. Among others (Task Committee on Sedimentation, 1973), there are the area-increment method and the empirical area-reduction method (Borland and Miller, 1958). The universal soil-loss equation proposed by Wischmeier et al. (1958) is currently being used by the U. S. Soil Conservation Service. All of the work mentioned thus far has been on a deterministic level. Woolhiser and Todorovic (1974) have developed a stochastic model of the sediment yield process where the total seasonal (yearly) sediment yield is treated as the sum of a random number of random events (variables), assumed to be mutually independent and identically distributed. Their approach will be used in part here. A literature review tends to show that there is no widely accepted method of computing sediment yield in dam or reservoir design and that there are no methods that treat the sediment yield estimation problem in a probabilistic fashion based on the actual independent variables involved and the individual watershed characteristics.

This exposition takes the deterministic model of Wischmeier et al. (1958) modified for individual events and, using a joint distribution of rainfall ($\underset{\sim}{x}_1$) and storm duration ($\underset{\sim}{x}_2$), obtains the probability density function (pdf) of sediment yield of a per event basis. This model was chosen because it accounts for runoff and peak flow rate in terms of rainfall amount and storm duration. It also accounts for the time of concentration, but rather than considering it a random variable, a constant average value is used in this study.

The next section presents the actual derivation of the sediment yield model. Also presented is an illustrative example.

Sediment Model

The universal soil-loss equation mentioned earlier has provided reasonable estimates of sedimentation in the past and will be the choice here for the randomization. This equation as modified by Williams and Hahn (1973) gives the sediment yield on a per event basis. It is a function of watershed characteristics and practices and runoff volume and peak flow rate. As mentioned before, an event is defined here as the sediment products of a runoff-producing type storm. Since the runoff volume and peak flow rate are direct functions of rainfall ($\underset{\sim}{x}_1$) and storm duration ($\underset{\sim}{x}_2$), these two random variables are of major concern. The modified universal soil-loss equation is:

$$Z = 95(Qq_p)^{.56}KCP(LS), \text{ tons} \tag{1}$$

where Z = sediment yield in tons,
Q = runoff volume in acre-feet,
q_p = peak flow rate in cfs,
K = soil erodibility factor,
C = cropping-management factor,

P = erosion control practive factor, and
LS = slope length and gradient factor.

Values for K, C, P, and LS may be computed, using the algorithms outlined by Williams and Berndt (1972). These algorithms make use of available data on soil type, topographic maps, and on-site estimates. Due to the form of the rainfall and duration data available, a conversion constant (to be defined later) is introduced to convert Z from tons to acre-feet, and Q in inches to acre-feet. The values of Q and q_p are computed from the Soil Conservation Service formulas:

$$Q = \frac{\underset{\sim}{x}_1^{\,2}}{(\underset{\sim}{x}_1 + s)}, \text{ acre-inches} \tag{2}$$

where Q = runoff volume in acre-inches,
$\underset{\sim}{x}_1$ = effective rainfall in inches (rainfall less a constant initial abstraction), and
S = watershed infiltration constant;

and

$$q_p = \frac{484AQ}{(a_1\underset{\sim}{x}_2 + a_2)}, \text{ cfs} \tag{3}$$

where A = drainage area of watershed in square miles,
a_1 = constant (.50),
$\underset{\sim}{x}_2$ = storm duration in hours, and
a_2 = time of concentration of storm in hours (assumed constant for a given watershed).

By substituting Equations 2 and 3 into Equation 1, and defining the conversion constant a_o as $a_o = 484A^2(640)/12$, where area is in square miles, we get

$$Z = \left(\frac{a_o\underset{\sim}{x}_1^{\,4}}{(\underset{\sim}{x}_1 + S)^2(a_1\underset{\sim}{x}_2 + a_2)}\right)^{.56} W, \text{ acre-feet} \tag{4}$$

where W = 95KCP(LS)(2000)/(mean sediment density x 4.356×10^4). To obtain the distribution function of sediment yield, we need the joint distribution of rainfall and storm duration $f(\underset{\sim}{x}_1,\underset{\sim}{x}_2)$. Crovelli (1971) has proposed a bivariate gamma pdf. This distribution possesses many of the properties consistent with the empirical properties of certain storms. The data from the Atterbury Experimental Watershed were used to conduct a Kolmogorov-Smirnov goodness-of-fit test on the marginal distributions of rainfall and storm duration. The distributions could not be rejected at the 10% level of significance, so the bivariate gamma pdf was assumed to be an acceptable fit of the random properties of storms over the

Atterbury watershed. The parameters of this distribution were estimated by the method of maximum likelihood of the marginal distributions. The bivariate gamma pdf of rainfall amount and storm duration is

$$f(x_1,x_2) = \begin{cases} \alpha\beta \exp - \beta x_1(1 - \exp - \alpha x_2) & \text{if } 0 \leq x_2 \leq \frac{\beta}{\alpha} x_1 \\ \alpha\beta \exp - \alpha x_2(1 - \text{esp} - \beta x_1) & \text{if } 0 \leq \frac{\beta}{\alpha} x_1 \leq x_2 \end{cases} \tag{5}$$

where $\alpha,\beta > 0$, $0 < x_1 < \infty$, and $0 < x_2 < \infty$.

Given n pairs of rainfall and duration values (R_i,D_i) with n_1 pairs such that $\alpha D_i < \beta R_i$ and n_2 pairs such that $\alpha D_i > \beta R_i$ and $n = n_1 + n_2$, the likelihood values may be computed from

$$\begin{aligned} \text{LIKELIH}(\alpha,\beta) = {} & [\alpha^n \exp - \alpha \sum_{j=1}^{n_2} D_j \prod_{i=1}^{n_1} (1 - \exp - \alpha D_i)] \\ & [\beta^n \exp - \beta \sum_{i=1}^{n_1} R_i \prod_{j=1}^{n_2} (1 - \exp - \beta R_j)] \end{aligned} \tag{6}$$

To obtain $F_e(Z)$, it is necessary to integrate the joint pdf in terms of the conditional and marginal distributions.

$$\begin{aligned} F_e(\underset{\sim}{z} < Z) &= \int_o^\infty P(\underset{\sim}{z} < Z \; x_1 = x) f(x_1) dx_1 \\ &= \int_o^\infty P([\frac{a_o x_1^4}{(x_1 + S)^2 (a_1 x_2 + a_2)}]^{.56} W < Z | x_1 = x) f(x_1) dx_1 \\ &= \int_o^\infty P(x_2 > \frac{c x_1^4}{(x_1 + S)^2} - d | x_1 = x) f(x_1) dx_1 \end{aligned}$$

where $c = a_o/a_1(W/Z)^{1/.56}$ and $d = a_2/a_1$. We also let

$$\psi(x_1, Z) = \frac{cx_1^{4}}{(x_1 + S)^2} - d$$

and we have

$$F_e(\underset{\sim}{z} < Z) = \int_{o}^{\infty} \int_{\psi(x_1,Z)}^{\infty} f(x_1,x_2)dx_2dx_1 \qquad (7)$$

Since $\psi(x_1,Z)$ is defined over both regions of Equation 5, Equation 7 must be broken into four separate regions. Figure 1 illustrates the joint pdf $f(x_1,x_2)$ and the sediment function $\psi(x_1,Z)$. Note that the curve is defined on both sides of $x_2 = \frac{\beta}{\alpha} x_1$, so the region of integration from $\psi(x_1,Z)$ to infinity must be partitioned and the boundary points x_1* and x_2** are obtained by solving the equations $\psi(x_1,Z) = \frac{\beta}{\alpha} x_1$ and $\psi(x_1,Z) = 0$, respectively. The resulting expression is then

$$\begin{aligned} F_e(\underset{\sim}{z} < Z) = & \int_{o}^{x_1^{**}} \int_{o}^{\frac{\beta}{\alpha} x_1} \alpha\beta\exp - \beta x_1(1 - \exp - \alpha x_2)dx_2dx_1 \\ & + \int_{x_1^{**}}^{x_1^{*}} \int_{\psi(x_1,Z)}^{\frac{\beta}{\alpha} x_1} \alpha\beta\exp - \beta x_1(1 - \exp - \alpha x_2)dx_2dx_1 \\ & + \int_{o}^{x_1^{*}} \int_{\frac{\beta}{\alpha} x_1}^{\infty} \alpha\beta\exp - \alpha x_2(1 - \exp - \beta x_1)dx_2dx_1 \\ & + \int_{x_1^{*}}^{\infty} \int_{\psi(x_1,Z)}^{\infty} \alpha\beta\exp - \alpha x_2(1 - \exp - \beta x_1)dx_2dx_1 \end{aligned}$$

It was not possible to obtain closed form expressions for all of the above terms; however, the simplest form obtained was

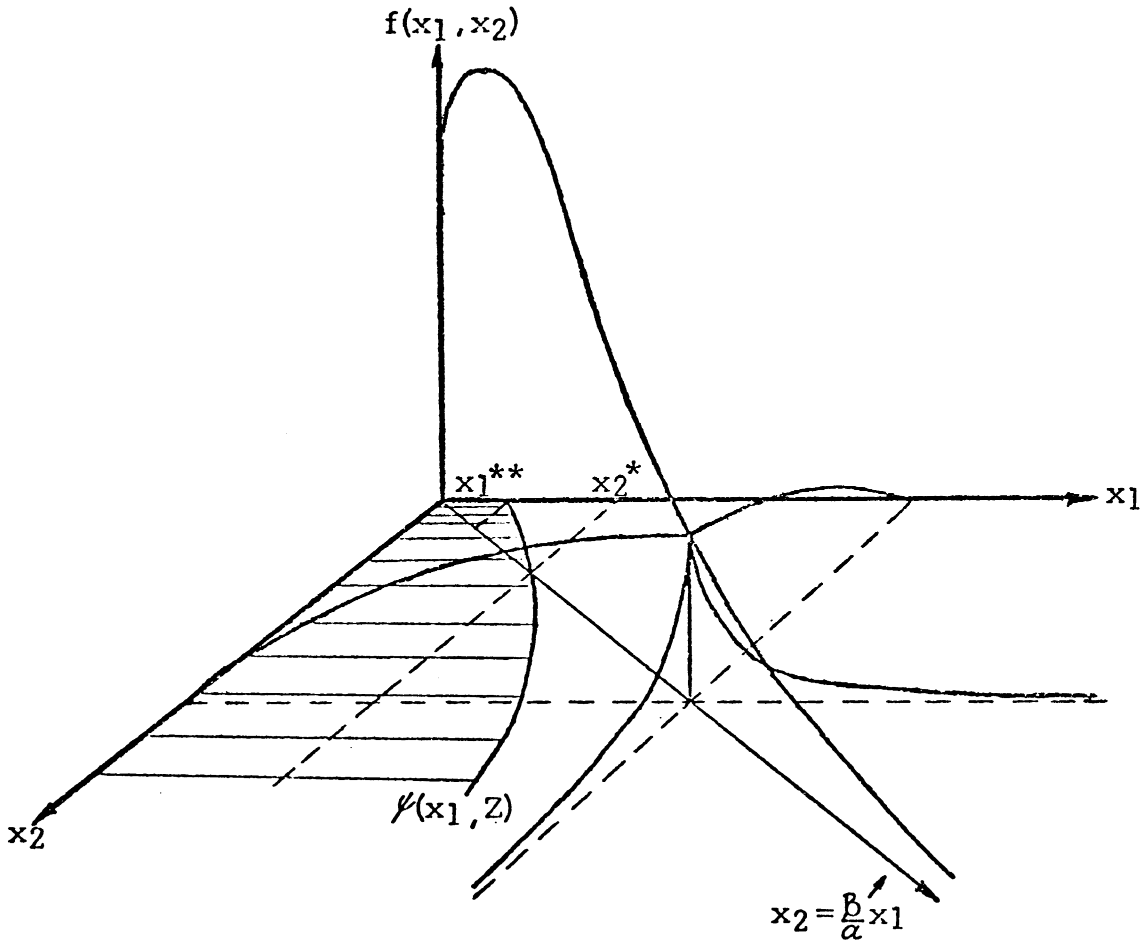


Figure 1. Bivariate gamma distribution of rainfall amount and storm duration

x_1 = rainfall amount; x_2 = storm duration. Note the sediment yield function, $\psi(x_1, Z)$.

$$F_e(\underset{\sim}{z} < Z) = 1 - \exp - \beta x_1^*(\beta x_1^* + 1) + (1 + \alpha d)(\exp - \beta x_1^{**}$$

$$- \exp - \beta x_1^*) + \alpha c \exp - \beta x_1^{**} [\frac{2}{\beta} (\beta(x_1^{**} + S)$$

$$+ 1)(2S - \frac{1}{\beta}) - (x_1^{**} + S)^2 - S^2(6 + \frac{\beta S^2}{x_1^{**} + S})]$$

$$+ \alpha c \exp - \beta x_1^* [\frac{2}{\beta} (\beta(x_1^* + S) + 1)(\frac{1}{\beta} - 2S)$$

$$+ (x_1^* + S)^2 + S^2 (6 + \frac{\beta S^2}{x_1^* + S})] + \alpha\beta c S^3 (\beta S$$

$$+ 4) \exp \beta S [\ln \frac{x_1^* + S}{x_2^{**} + S} + \sum_{k=1}^{\infty} \frac{(-\beta)^k}{k \cdot k!} [x_1^* + S)^k$$

$$- (x_1^{**} + S)^k]] + \beta[\int_{x_1^*}^{\infty} \exp - \alpha\psi(x_1,Z)dx_1$$

$$- \int_{x_1^{**}}^{\infty} \exp (-\alpha\psi(x_1,Z) - \beta x_1)dx_1] \qquad (8)$$

The procedure to compute $F_e(Z)$ is to first calculate the constant

$$c = \frac{a_o}{a_1} (W/Z)^{1/.56}$$

then to calculate the roots x_1^* and x_2^{**}, to evaluate the constant terms in Equation 8, the series term, and, finally, the integral terms. Convergence of the series term to 20 decimal places was obtained within 60 to 70 terms. The two integrals were evaluated using Simpson's rule in which the upper bound on the absolute error between successive iterations was arbitrarily preset at 10^{-8}. The two integrands are exponentially asymptotic and converge so rapidly that the contribution of their sum was less than 10^{-3}.

Data from the Atterbury watershed were used to compute Equation 8. Mean rainfall and storm duration values were available for individual rainfall events over a period of 14 years. These values were used to compute the marginal maximum likelihood estimates of α and β. Figure 2 illustrates the distribution function obtained from Equation 8 for these particular parameter values. Computations of 56 values consumed approximately 10 to 15 seconds of computer time, depending on the convergence rate of the integration routine. The pdf of sediment yield for a single event was calculated from the values of $F_e(Z)$, using the definition of the derivative (Figure 3). Thus, it is possible to compute the probability of Z units of sediment given only n pieces of data which are used

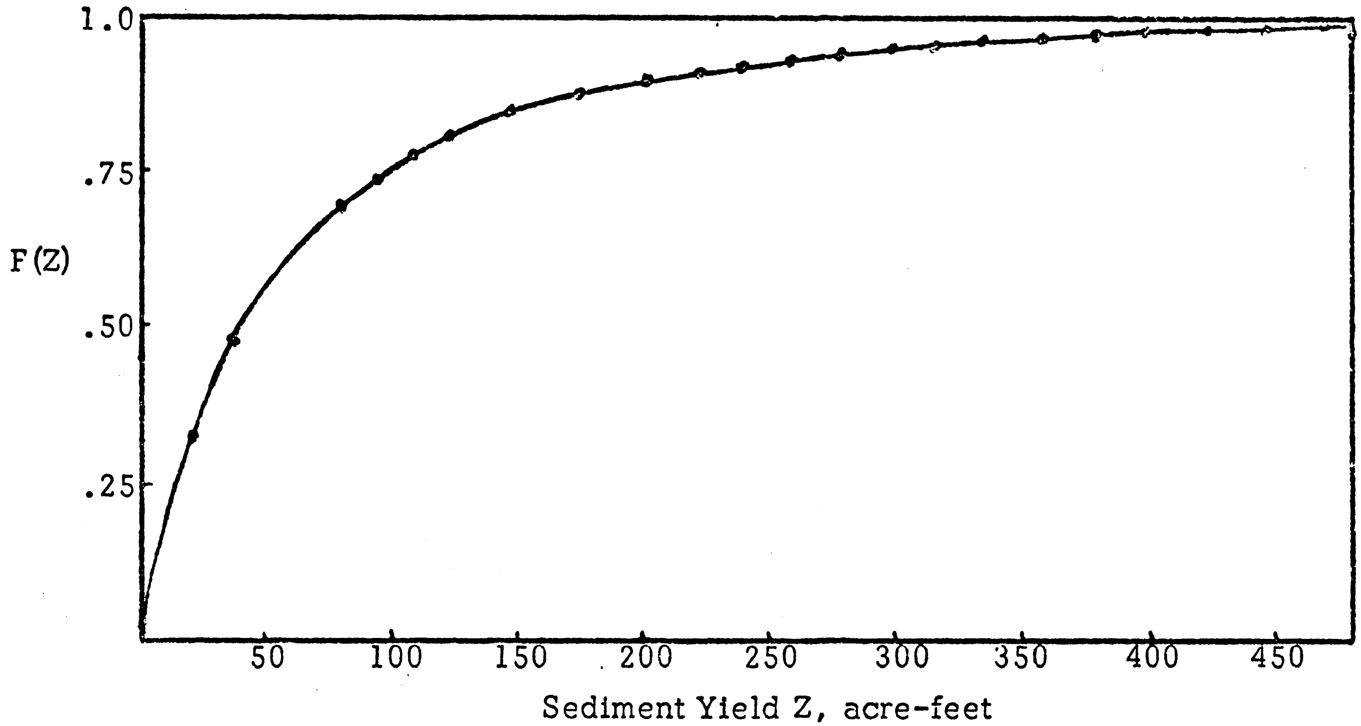


Figure 2. Cumulative distribution function of single event, using the maximum likelihood estimators of marginal distributions of rainfall and storm duration

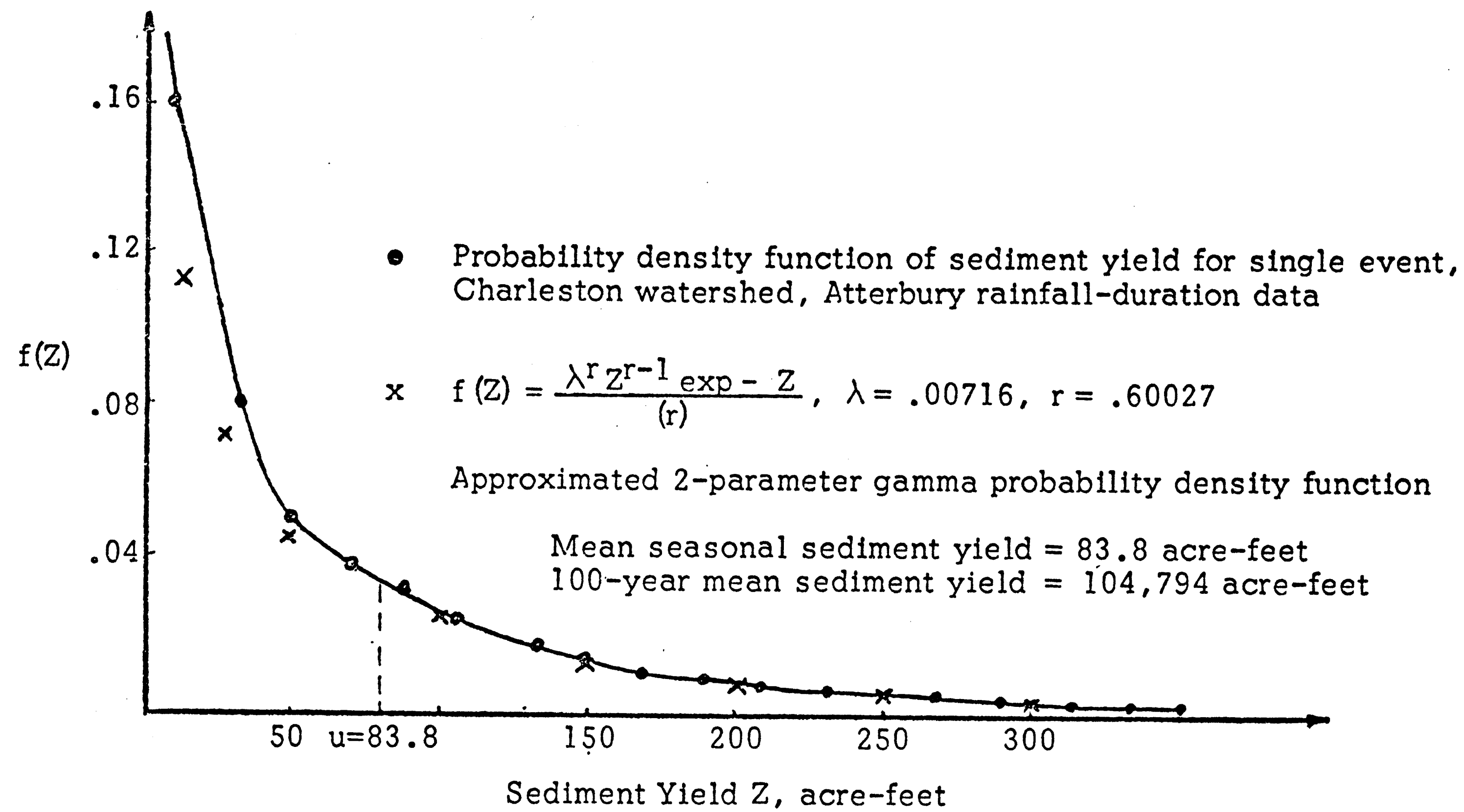


Figure 3. Probability density function of sediment yield for a single event

to estimate the parameters α and β. The mean and variance for this pdf may be calculated by using the Riemann-Stieljes definition of integration. The first two moments of $F_e(Z)$ are defined by

$$E(Z) = \int_o^\infty Z \, dF_e(Z) \cong \sum_{i=1}^{n} Z \, [(F_e(Z_{i+1}) - F_e(Z_i)]$$

and

$$E[(Z - u)^2] = \int_o^\infty Z^2 \, dF_e(Z) \cong \sum_{i=1}^{n} Z^2 \, [F_e(Z_{i+1}) - F_e(Z_i)] - [E(Z)]^2$$

After computing the mean and variance of $F_e(Z)$, the parameters of a two-parameter gamma distribution are fit to the mean and variance. The two-parameter gamma pdf is

$$F_e(Z) = \frac{\lambda_e^{r_e} Z^{r_e - 1} \exp - \lambda_e Z}{\Gamma(r_e)}, \quad Z > 0$$

with

$$\left. \begin{aligned} \mu_e &= \frac{r_e}{\lambda_e} \\ \sigma_e^2 &= \frac{r_e}{\lambda_e^2} \end{aligned} \right\} \qquad \begin{aligned} r_e &= \frac{\mu_e^2}{\sigma_e} \\ \lambda_e &= \frac{\mu_e}{\sigma_e} \end{aligned}$$

Figure 3 also shows the empirical pdf and the approximation by the gamma-2 pdf, with $\lambda_e = .00716$, $r_e = .60027$, and $\mu_e = 83.800$ acre-feet.

The next step will be to extend the event pdf to a seasonal distribution and finally to a N-season model. This is necessary in order that we may use Bayesian decision theory, since the decision being made is dependent on the number of seasons to be considered in the lifetime of the project.

CHAPTER 2

EXTENSION OF THE EVENT-BASED MODEL TO A SEASONAL SEDIMENT MODEL

Here, the distribution computed in the last chapter is first extended to a seasonal pdf, then to a pdf over N seasons, where N is the lifetime of the facility under consideration.

Since the individual sediment events are hypothesized to be gamma-2, the seasonal distribution of sediment yield may be defined as the sum of n mutually independent and identically distributed gamma-2 events. This pdf is represented as

$$f_s(Z) = \frac{\lambda_e^{nr_e} Z^{nr_e - 1} \exp(-\lambda_e Z)}{\Gamma(nr_e)} \tag{9}$$

with $\mu_s = \frac{nr_e}{\lambda_e}$ and $\sigma_s^2 = \frac{nr_e}{\lambda_e^2}$. The critical assumption is that the number of events per season is always the same; i.e., n is a fixed quantity from season to season. In order to be more realistic, n should be taken as a random variable. Todorovic and Yevjevich (1969) and Duckstein et al. (1972) have shown that seasonal rainfall events for many areas, including the Atterbury watershed, are distributed in a Poisson fashion. Using this result, the sum of n gamma-2 events defined earlier may be convoluted with the Poisson probability of n events occurring (Feller, 1966). Thus, the seasonal distribution becomes

$$f_s(Z) = e^{-\tilde{n}} \sum_{k=1}^{\infty} \frac{(\tilde{n})^k}{k!} \frac{\lambda_e^{kr_e} Z^{kr_e - 1} \exp(-\lambda_e Z)}{\Gamma(kr_e)} \tag{10}$$

where $\tilde{n}$ is the average number of storms per season (= 12.5 for Charleston watershed). The index starts at zero because the gamma function in the denominator is undefined at k = 0. It is heuristically assumed that the probability of zero events is extremely small.

The seasonal mean and variance, μ_s and σ_s^2, may be calculated directly without knowing the parameters of the actual seasonal distribution of $f_s(Z)$. Benjamin and Cornell (1970) outline the formula for the mean and variance of the pdf of a sum of a random number of random events mutually independent and identically distributed as:

$$\mu_s = E(n)E(Z)$$

$$\sigma_s^2 = E(n)\ \text{var}(Z) + \text{var}(n)\ [E(Z)]^2$$

The same method may be used to obtain the mean and variance of the distribution of N seasons. The first two moments of this pdf are:

$$\mu_L = N \cdot \mu_s$$
$$\sigma_L^2 = N \cdot \sigma_s^2 \qquad (11)$$

The value obtained for μ_L is the design estimate on which the decision is made; i.e., the volume of sediment that should be allowed for in designing the reservoir.

To illustrate the above, data from Atterbury Experimental Watershed is used. Throughout the example, the maximum likelihood estimates of the marginal distributions are used.

Example: Atterbury Data Applied to Charleston Watershed.

$\alpha = 1.84$	$\beta = 2.58$
$K = 0.60$	$A = 1220 \text{ mi}^2$
$C = 0.80$	$a_o = 3.83576 \times 10^{10}$
$P = 0.10$	$a_1 = 0.50$
$LS = 0.50$	$a_2 = 6.989$
$S = 2.50$	

Figure 2 illustrates the distribution function for the event-based case. The solid line on Figure 3 shows the pdf obtained from Figure 2, using the definition of derivative. The crosses indicate the 2-parameter gamma approximation. Note the extremely good fit in the tail. Both curves tend to infinity as Z goes to zero, although the differences become slightly more pronounced. Several curves were calculated to substantiate the 2-parameter gamma approximation. The seasonal mean and variance are

$$\mu_s = E(n)E(Z) = 1{,}050$$
$$\sigma_s^2 = E(n)\ \mathrm{var}(Z) + E(Z)^2\ \mathrm{var}(n) = 234{,}000$$

and the lifetime mean and variance for N = 100 years are

$$\mu_L = 100\ \mu_s = 105{,}000$$
$$\sigma_L^2 = 100\ \sigma_s^2 = 23{,}400{,}000$$

Note here that the coefficient of variation ($100\ \sigma/\mu$) is around 4 percent. This fact will be of major importance in the decision theory chapters.

To summarize, a model for sediment yield has been developed, which accounts for the probabilistic nature of the process through the random variables-rainfall amount $\underset{\sim}{x}_1$ and storm duration $\underset{\sim}{x}_2$. The model has been extended to provide a means of estimating the amount of sedimentation first on a seasonal basis, then over a period of N seasons.

The study will now proceed to the decision theory aspects of the sediment design problem. The next chapter will present the Bayesian decision methodology in general form before its application to the sedimentation problem in Chapter 4.

CHAPTER 3

MAKING THE OPTIMAL DECISION

Davis et al. (1972) have adapted methodology developed by Raiffa and Schlaifer (1961) and Howard (1966) for finding optimal alternatives to hydrologic design problems through the use of Bayesian decision theory. The first step in the application of this theory is the identification of the goal; that is, at the same time the decision to be made and its alternatives, a, must be defined. Next, a goal or loss function $L(a,\theta)$ must be constructed in which the state or uncertain variables, θ, must be selected. The state variables in this study are the uncertain parameters α and β of the joint distribution of rainfall and storm duration, so that the symbol θ is a vector of the state variables α and β. To make the decision, it is necessary (1) to calculate the expected value of the goal function for each alternative, and (2) to choose an alternative to minimize the expected value of the goal function. It then remains to evaluate the decision and determine the expected opportunity loss due to the uncertain parameters in the problem. The calculation of the expected value of the goal-loss function for each alternative is defined as the risk of that alternative. Given the state and decision variables, their respective pdf's, and the loss function, the Bayes solution is obtained by choosing the alternative a^* that minimizes the risk:

$$BR(a^*) = \min_a \int L(a,\theta)f(\theta)d\theta \tag{12}$$

Next, the decision must be evaluated. If the true values, θ_t, of the state variables were known, the alternative chosen, a_t, would be the one that minimized the loss function for θ_t:

$$L(a_t,\theta_t) = \min_a L(a,\theta_t)$$

The decision has been made to use the alternative a^*, which may be a nonoptimal choice. In making this selection, an opportunity loss (OL) is suffered:

$$OL(a^*,\theta_t) = L(a^*,\theta_t) - L(a_t,\theta_t)$$

The value of θ_t is not known, but the pdf, $f(\theta)$, is known, so an expected opportunity loss (EOL) may be calculated:

$$EOL(a) = \int[L(a^*,\theta) - \min_a L(a,\theta)]\, f(\theta)d\theta$$

The expected opportunity loss represents the expected value of perfect information and may be used to judge the effect of uncertainty as embodied in $f(\theta)$ on the performance of the project. The pdf $f(\theta)$ is known as the posterior distribution. It is calculated from Bayes rule, using the likelihood function of the state variable LIKELIH(θ) and the prior distribution of the state variables, i.e., PRIOR(θ). The actual expression which yields the posterior distribution will be given later.

The next step is to define the loss function, prior, and posterior distributions of the state variables in the context of the sediment yield problem.

The loss function defined by Jacobi (1971) is used for assessing the expected value of additional information. This particular type of loss function, known as the linear terminal opportunity loss function, is:

$$L(a,\overline{Z}) = \begin{cases} K_o(a - \overline{Z}) & \text{if } a > \overline{Z} \text{ (overdesign)} \\ K_u(\overline{Z} - a) & \text{if } a < \overline{Z} \text{ (underdesign)} \end{cases} \qquad (13)$$

The function proposed by Jacobi was modified by a scalar amount to account for the economic losses associated with decreasing water storage over time. This constant is embedded in constants K_u and K_o, which are defined later.

The posterior pdf of θ is defined by Bayes rule as:

$$\text{POST}(\theta|\text{data}) = \frac{\text{PRIOR}(\theta)\ \text{LIKELIH}(\theta)|\text{data})}{\int[\text{numerator}]d\theta} \qquad (14)$$

Having defined the critical elements of the problem, it is now necessary to conduct the search for the optimal solution. This step has been greatly simplified by the linearity property of the loss function. According to Raiffa and Schlaifer (1961, pp. 195-197), by taking partial expectations over the distribution on the decision variable and the loss function, the expression which must be solved represents a cumulative distribution over the decision variable. The equation to be solved is: find a* such that

$$F(a^*) = \int_\theta \int_o^{a^*} f(Z|\theta)\text{POST}(\theta|\text{data})dZd\theta = \frac{K_u}{K_u + K_o} \qquad (15)$$

where K_u and K_o are defined in Equation 13. As can be seen in Figure 13, the loss function depends on the alternative chosen, a, and the actual state of nature $\overline{Z}$. The variable $\overline{Z}$ is the decision variable in the sediment yield problem and represents the mean number of acre-feet of sediment of the pdf of sedimentation for an N-year period. This quantity is also equivalent to the u_L term in Equation 11. While the pdf $f(Z|\theta)$can be computed through an extension of Equation 9, the use of an indicator function greatly reduces the computation time required to generate the pdf $f(Z|\theta)$. Returning to Equation 15, it can be seen that POST(θ|data) and $f(Z|\theta)$ are known and K_u and K_o are defined; integration of Equation 15 is carried out for each a* until the value $K_u/(K_u + K_o)$ is achieved within some prespecified error bound ε. In other words, the problem becomes: find a* such that

$$a* = F^{-1}\left[\frac{K_u}{K_u + K_o}\right]$$

or

$$F(a*) - \frac{K_u}{K_u + K_o} \leq \varepsilon$$

or

$$\int_{\theta} \int_{o}^{a*} f(Z|\theta) POST(\theta|data) dZd\theta - \frac{K_u}{K_u + K_o} \leq \varepsilon \qquad (16)$$

It now remains to apply the above equation to the sediment yield problem. The next chapter will consist mainly of the above concepts adapted to the problem with basic notational changes as required.

CHAPTER 4

ADAPTATION OF BAYESIAN METHODOLOGY TO THE SEDIMENTATION PROBLEM

The purpose of this chapter is to apply the concepts presented in the last section to the sediment yield problem. After defining the loss function and the distributions in a Bayesian framework, a discussion of the various computational problems and simplification is undertaken and an actual case study is presented.

First, the constants K_u and K_o are defined in order to illustrate completely the elements of the loss function. The (a) term is the alternative chosen and the $\bar{Z}$ term is the state of nature or actual amount of sedimentation over the lifetime of the project. These quantities appear in the definition of the general loss function in Equation 13. Naturally, the decision maker would like to choose $a = \bar{Z}$, but this is where uncertainty arises. The constant terms in the loss function are:

$$K_o = P \cdot K_1 (1 + i)^n - K_{ws}$$

where P = proportionality factor between total sediment load and suspended load (= 1.1),
K_1 = unit cost of construction (= \$150/acre-ft),
i = interest rate of borrowed dollars (4.78%),
n = years between commencement of loan and start of reservoir operation (= 10 yr), and
K_{ws} = water storage value (= \$53/acre-ft);

and

$$K_u = P \cdot K_2 \left[\frac{1}{(1 + r)^M} + \frac{1}{(1 + r)^{2M}} + \cdots + \frac{1}{(1 + r)^{N-M}} + \frac{1}{(1 + r)^N} \right] + K_{ws} \tag{17}$$

where K_2 = unit cost for removal of sediment (\$1700/acre-ft),
r = interest rate for discounting (= 4.78%), and
M = constant time interval between removals (25 yr).

The definition of the posterior distribution is difficult in that the property of "natural conjugates" is not present here; that is, the posterior and prior distributions are not of the same family. It is thus necessary to integrate the denominator of Equation 14 everytime the pdf is updated. Also present here is the problem that the form of the pdf of the prior distribution is not known. As a consequence, it is assumed that no prior knowledge of the data exists. In effect, we are assuming that the prior pdf is a uniform distribution. There is a conflict here since there is no guarantee that the variables range over a finite interval. A numerical simplification to be illustrated later was used to reduce the interval to a finite range where a uniform distribution is defined. Basically, this was achieved by restricting the regions of integration in Equations 14 and 15 to an area of most dense likelihoods. In other words, instead of integrating over the entire space, integration is performed only over the most likely area in

$$\text{POST}(\alpha,\beta|\text{data}) = \frac{1\cdot\text{LIKELIH}(\alpha,\beta|\text{data})}{\int_\beta\int_\alpha 1\cdot\text{LIKELIH}(\alpha,\beta|\text{data})d\alpha d\beta}$$

Replacing the posterior distribution in Equation 15 by its more explicit representation, the problem becomes: find a* such that

$$\frac{\int_\beta\int_\alpha\int_o^{a^*} f(Z|\alpha,\beta)\text{LIKELIH}(\alpha,\beta|\text{data})dZd\alpha d\beta}{\int_\beta\int_\alpha \text{LIKELIH}(\alpha,\beta|\text{data})d\alpha d\beta} - \frac{K_u}{K_u + K_o} \leq \varepsilon \tag{18}$$

For notational purposes defined:

$$C = \int_\beta\int_\alpha \text{LIKELIH}(\alpha,\beta)d\alpha d\beta \tag{19}$$

At this point, the determination of the form of the pdf $f(Z|\alpha,\beta)$ has not been pursued. However, it was found that it was not necessary to compute the pdf, since an indicator function simplification was introduced based on a result obtained in the section on development of the sediment yield model which eliminated the necessity of having to calculate the explicit probabilities. This simplification will be explored later. Since $f(Z|\alpha,\beta)$ depends on α and β, each time the values α and β are change in the numerical integration of Equation 18, a new pdf of Z must be calculated. This difficulty may be overcome quite readily as shown in the next section where the simplifications mentioned here are illustrated in greater detail.

CHAPTER 5

NUMERICAL CONSIDERATIONS ASSOCIATED WITH THE COMPUTATION OF EQUATION 18

Since the values α and β may take on a possible range from zero to infinity, Equation 18 becomes: find a* such that

$$F(a^*) = \frac{1}{C} \int_0^\infty \int_0^\infty \int_0^{a^*} f(Z|\alpha,\beta) \text{LIKELIH}(\alpha,\beta)|\text{data}) dZ d\alpha d\beta - \frac{K_u}{K_u - K_o} \leq \varepsilon$$

There are two major methods of numerical simplification:

1. Restrict α and β to a finite area to uphold the uniform prior assumption and to reduce the regions of integration.

2. Use linear regression to obtain the relationship between $\overline{Z}$ and α and β, since $\overline{Z}$ had to be recalculated every time α or β changed in the numerical integration routine. Then, an indicator function is used to avoid the computation of the pdf $f(Z|\alpha,\beta)$ every time α or β changes.

In the first simplification, the infinite regions of integration were restricted to the most likely area in terms of the likelihood function. This technique has been used successfully by Yakowitz, Duckstein and Kisiel (1974). This area was obtained by repeatedly evaluating the likelihood function (Equation 6) over a coarse grid and defining the area as shown in Figure 4. The rectangular boundaries were arbitrarily fit as shown, due to the simple forms associated with the straight lines. From Figure 4, the parameters α and β range over the following regions:

$$.4324\beta + .2932 \leq \alpha \leq .4324\beta + 1.4932$$

$$1.75 \leq \beta \leq 3.60 \qquad (20)$$

So now the regions of integration are much more compatible with the uniform distribution assumption.

The Indicator Function

The discussion of the second set of numerical simplifications is somewhat more involved. Since $f(Z|\alpha,\beta)$ must be calculated everytime α and β change in the numerical integration, an indicator function is introduced to avoid this calculation.

The basis for using this special function is the fact that the coefficient of variation found earlier is small, which indicates a highly peaked distribution. In fact, to use the indicator function, it is assumed that there is a spike about $\overline{Z}$ somewhere in or out of the interval

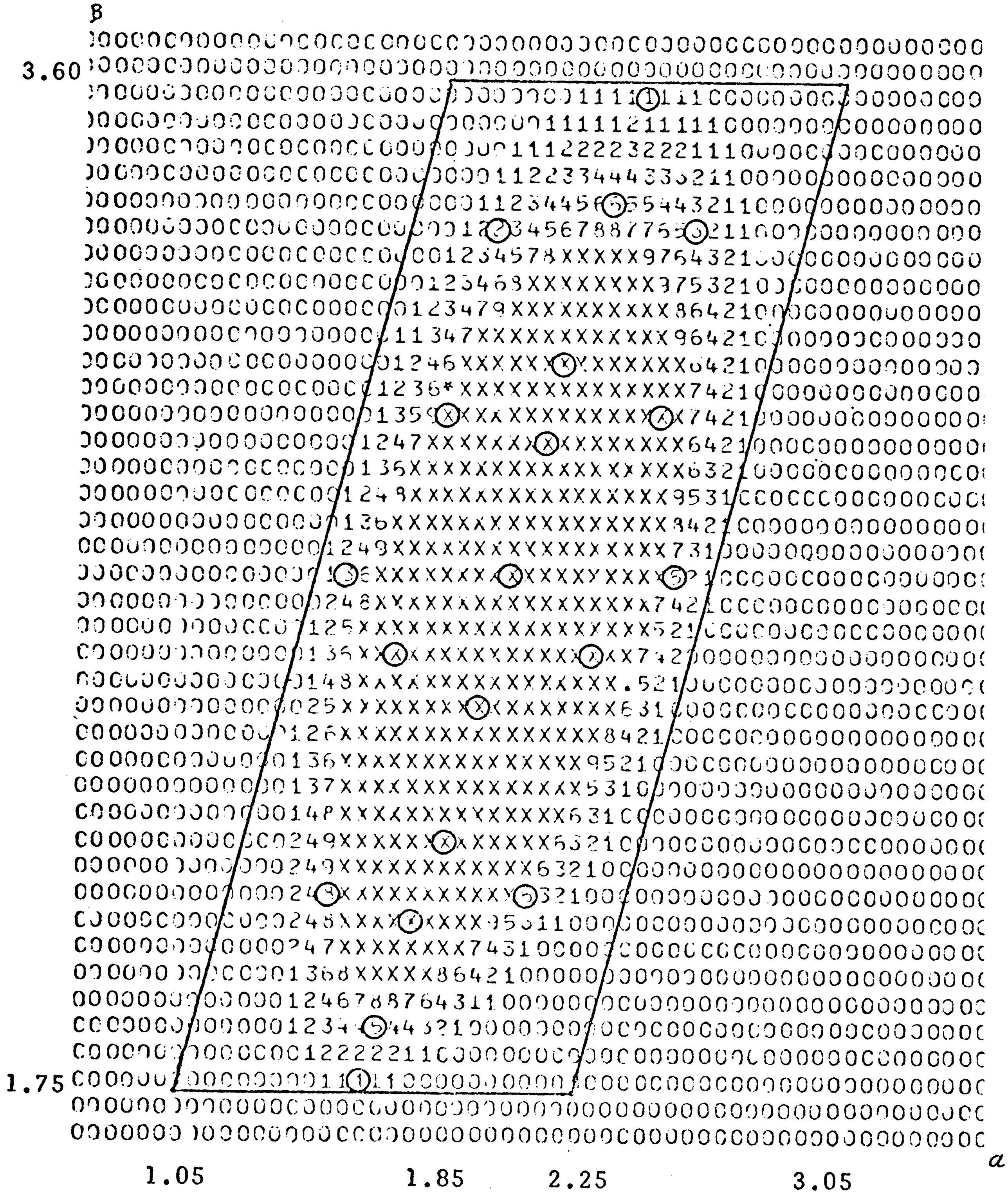


Figure 4. Likelihood map of α and β

Grid size = 0.05.

[0,a*]. If the pdf $f(Z|\alpha,\beta)$ lies within the region [0,a*], the assumption that it is a spike indicates that since it is a pdf, the value of the integral must be one (Figure 5). If the value of $\overline{Z}$ is outside the interval [0,a*], then it is assumed the entire pdf lies outside the interval and the integration from zero to z* yields a zero (Figure 6). Thus, in the numerical integration of Equation 18, each time α or β changes value, the inner integral over Z is either a zero or a one rather than a specific probability. Since this pdf does depend on α and β, however, a relationship between α, β, and $\overline{Z}$ is needed so that a $\overline{Z}$ value may be calculated rapidly from the α and β values to be used in the indicator function. The mathematical form of the indicator function is:

$$I(\overline{Z}_{[0,a^*]}|\alpha,\beta) = \begin{cases} 0 \text{ if } \overline{Z} > a^* \\ 1 \text{ if } \overline{Z} \leq a^* \end{cases}$$

Once the relationship between α, β, and $\overline{Z}$ is found, Equation 18 is reduced from a triple integration to a double integration, which represents an appreciable savings in computing time.

The Regression Relationship

The next obvious question is how can the values of $\overline{Z}$ be calculated quickly from the changing α and β values. As the numerical integration procedure used here was Simpson's rule, the β value is held constant while the integration takes place over the variable α and then Z; then β is incremented and the process is repeated until all the limits are reached. Since it takes 10 seconds of computer time to calculate a single $\overline{Z}$ value from an (α,β) pair, if 400 integration points are used by the Simpson routine (and this is conservative), the evaluation of Equation 18 takes at least 4,000 seconds or 66 minutes of computer time for each a* value used in the search; then many iterations may be necessary. Since such a computation is economically infeasible, the following simplification is introduced. A set of 20 points is extracted from the likelihood map (circles on Figure 4); this set is felt to be a representative sample of the region. In this manner, the actual calculation of the $\overline{Z}$ values associated with the 20 pairs consumes a total of 200 seconds. From this point, a setpwise linear regression program was used to obtain the coefficients of the least squares polynomial through the region. The form of the polynomial is:

$$\overline{Z} = \rho_u(\alpha,\beta) = a_o \cdot \alpha + a_1 \cdot \beta + a_2 \cdot \alpha^2 + a_3 \cdot \beta^2 + a_4 \alpha\beta + a_5 \qquad (21)$$

Thus, in the program to evaluate Equation 18 as α and β change, the value of $\overline{Z}$ is calculated using Equation 21 and an indicator function subroutine makes the test on the position of $\overline{Z}$ relative to [0,a*] and returns either a zero or a one. It is necessary to check the smoothness of Equation 21 prior to using it to insure that the polynomial will accurately represent the relationship. The constant C was evaluated in 9 seconds, using the Simpson's rule program mentioned above. This program allows the user to

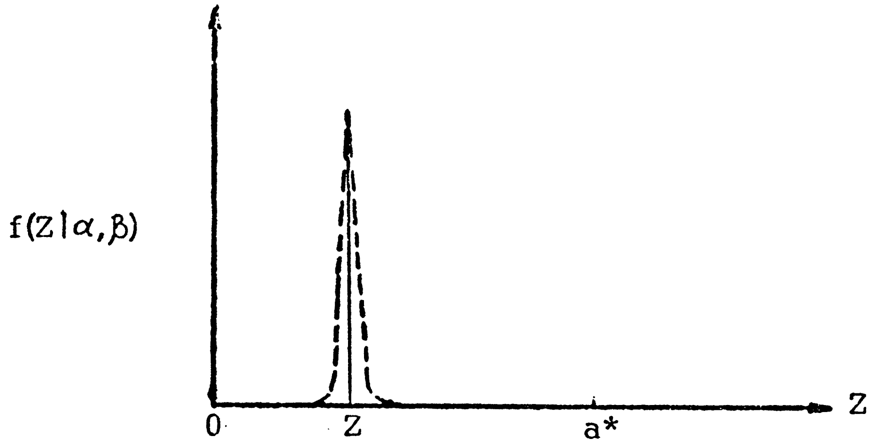


Figure 5. Indicator function $I(\overline{Z}_{[0,a^*]} | \alpha,\beta) = 1$ for all $\overline{Z} \leq a^*$

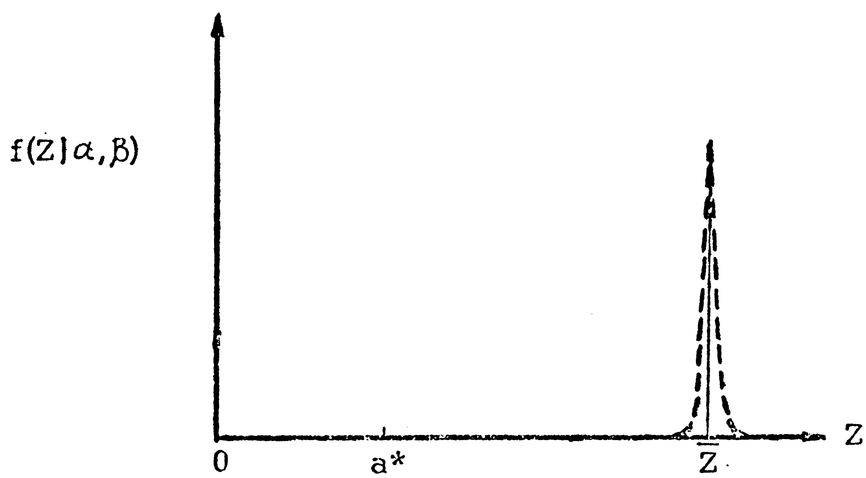


Figure 6. Indicator function $I(\overline{Z}_{[0,a^*]} | \alpha,\beta) = 0$ for all $\overline{Z} \geq a^*$

prescribe the absolute error between successive iterations of the integration. The calculation of C utilizes an error bound of 10^{-3}.

The above simplifications reduced Equation 18 to the following expression: find a* such that

$$F(a^*) = \frac{1}{C} \int_{\beta_1}^{\beta_2} \int_{\alpha_1,(\beta)}^{\alpha_2,(\beta)} I(\overline{Z}_{[0,a^*]} | \alpha,\beta)[\text{Equation 6}]d\alpha d\beta$$

$$- \frac{K_u}{K_u + K_o} \leq \varepsilon \tag{22}$$

In the actual program which performed the search for a*, the following strategy was utilized

$$\text{If } \left| F(a_i^*) - \frac{K_u}{K_u + K_o} \right| < \varepsilon, \text{ stop;} \tag{23}$$

$$\text{If } \left\{ F(a_i^*) - \frac{K_u}{K_u + K_o} \right\} \equiv d \geq \varepsilon, \text{ set } a_{i+1}^* = \begin{cases} a_i^*(1+d) \text{ if } d < 0 \\ a_i^*(1-d) \text{ if } d > 0 \end{cases}$$

To summarize, the steps of a general purpose algorithm are defined to provide guidelines for the use of this methodology for multipurpose reservoir design with respect to sedimentation. The case study following this outline is an illustration of both the methodology and the theory. The algorithm consists of nine specific steps:

1. Define the watershed constants, K, C, P, LS, drainage area A, infiltration constant S, a_1, a_2, and the mean sediment density. Calculate W = 95KCP(LS)2000/(mean sediment density x 4.356×10^4) and $a_o = 484A^2(640)/12$.

2. Compute the likelihood map over α and β to determine the integration regions.

3. Extract a representative sample of points from the region.

4. Calculate the N-year lifetime means for each of these points.

5. Fit a polynomial $\rho_u(\alpha,\beta)$ to these points.

6. Check $\rho_u(\alpha,\beta)$ to insure smoothness over the integration regions.

7. Define the linear terminal loss function and the values of its associated parameters K_u and K_o.

8. Set a* and calculate

$$F(a^*) = \frac{1}{C} \int_{\beta_1}^{\beta_2} \int_{\alpha_1(\beta)}^{\alpha_2(\beta)} I(\overline{Z}_{[0,a^*]} | \alpha,\beta) \mathrm{LIKELIH}(\alpha,\beta | \mathrm{data}) d\alpha d\beta$$

9. Search for a* under the strategy outlined in Equation 23 until $F(a^*) - K_u/(K_u + K_o) < \varepsilon$, then stop.

Certain steps would require more computation if a different watershed were being examined; however, this extra work would be of a minimal nature.

CHAPTER 6

CASE STUDY -- CHARLESTON DAM SITE

The Charleston Dam site is approximately 24 miles south of Benson, Arizona, on the San Pedro River. This area was recognized as a dam site as early as 1909 (Schwalen, 1961). However, since the instigation of the first Central Arizona Project Report in 1941, the dam has been awaiting construction in conjunction with the Tucson Aqueduct to provide additional water resources to the city of Tucson, Arizona, and to help reduce the rapid depletion of the groundwater supply now taking place. The original dam design allows for 238,000 acre-feet of water below the spillway crest. Of this, 116,000 acre-feet are alloted for flood protection.

As a basis for comparison, the results of the study made by Schwalen (1961) will be used to evaluate the expected amount of sedimentation over the lifetime of the project. Schwalen estimated the useful life of the project at 200 years based on an annual estimate of 630 acre-feet, using a suspended sediment density of 70 pounds per cubic foot. The method consisted of taking actual measurements of the sediment at the dam site and obtaining sediment yield in tons per day for each day in the 3-month summer rainy season. Almost all of the sediment deposited in the proposed reservoir would occur during this time. The drainage into the facility comes from the Charleston watershed with an area of 1,220 square miles.

Since rainfall and storm duration data are not available for the Charleston watershed on a per event basis, data from the Atterbury Experimental Watershed are used. This is justified on the grounds that rainfall and duration characteristics of the convective storms occurring over both watersheds were the same (M. M. Fogel, School for Renewable Natural Resources, University of Arizona, personal communication, 1974).

This case study will follow the algorithm outlined in the last chapter. At each step, the computational difficulties and programs will be discussed as they arise.

Step 1. Definition of Watershed Constants

For the Charleston Dam site, the values of K, C, P, and LS were estimated by a hydrologist rather than using the algorithms proposed by Williams and Hahn (1973). The accuracy of the estimates was considered to be of minimal importance since the purpose of the study is the definition of a method.

Soil-erodibility factor	K = 0.60
Cropping-management factor	C = 0.80
Erosion control practice factor	P = 0.10
Slope length and gradient factor	LS = 0.50
Infiltration constant	S = 2.50

Peak flow equation constant	$a_1 = 0.50$
Time of concentration constant	$a_2 = 6.989$
Conversion constant	$a_o = 3.83576 \times 10^{10}$
Combined constant	$W = 1.49547 \times 10^{-3}$

Step 2. Computation of the Likelihood Map

Using Equation 6 and the Atterbury data, the likelihood map of the parameters α and β was calculated. As can be seen on Figure 4, the boundaries of the most dense section of the likelihood map are well-defined. This figure also shows the parallelogram boundary section around the integration region. This distance between each likelihood value is exactly 0.05, and the corners of the parallelogram define lines shown on the graph. These lines determine the inner limits of the integration over α. Likewise, the top and bottom of the parallelogram define the limits over β. The limits of integration are equivalent to those in Equation 20.

Steps 3 to 6. Sample Points from Likelihood Region

The circled points in Figure 4 represent the data values used in computing the polynomial for the indicator function. For each α and β value circled, the lifetime mean sedimentation is calculated. These values are shown in Table 1. After the points and means are used to compute the polynomial $\rho_u(\alpha,\beta)$, the polynomial is evaluated to insure that the function is smooth. Table 1 also shows the calculated polynomial values. The calculated polynomial is:

$$\overline{Z} = \rho_u(\alpha,\beta) = .224\alpha - 27.403\beta + .197\alpha^2 + 3.796\beta^2 - .298\alpha\beta + 56.358$$

Step 7. Definition of Loss Function

The loss function as defined in Equation 13 for the values given in Equation 17 becomes:

$$L(a,\overline{Z}) = \begin{cases} 210.19\ (a - \overline{Z}) & \text{if } a > \overline{Z} \text{ (overdesign)} \\ 888.89\ (\overline{Z} - a) & \text{if } a < \overline{Z} \text{ (underdesign)} \end{cases}$$

i.e., $K_u = 888.89$ and $K_o = 210.19$ and $K_u/(K_u + K_o) = 0.8087$.

Table 1. Data points for regression 100-year mean sedimentation as a function of α and β

Polynomial fit also given.

α	β	$\overline{Z}_L$	$\rho_u(\alpha,\beta)$
1.60	1.75	20.593×10^4	20.009
1.65	1.85	18.773	18.600
1.50	2.10	14.983	15.335
1.75	2.05	15.769	16.005
1.55	2.70	9.548	9.542
1.70	2.55	10.648	10.752
1.84	2.58	10.479	10.500
1.85	2.20	13,968	14.249
1.85	3.00	7.882	7.665
1.95	2.45	11.548	11.701
2.0	3.35	6.380	6.306
2.05	2.70	9.661	9.606
2.10	2.10	15.255	15.519
2.15	2.95	8.226	7,975
2.20	3.10	7.529	7.217
2.30	2.55	10.828	10.903
2.35	3.40	6.231	6.209
2.45	3.60	5.389	5.907
2.50	3.0	8.029	7.786
2.55	2.70	9.802	9.769
2.60	3.35	6.444	6.384

Step 8. Calculation of F(a*)

The value of C is the quantity defined in Equation 19. Using Simpson's rule, this value was found to be: $C \cong 5.2266 \times 10^{-17}$. The error bound ε, was taken to be 10^{-3}. The initial estimate of a* was taken to be 5.0.

Step 9. Finding the Solution

The optimum point at which the search stopped was 0.8087. The computation of the optimum alternative required 41 seconds of computer time and the results were printed out at each iteration. Five iterations of a* were required before the optimum was reached. Using the method of Schwalen (1961), his estimate would yield 63,000 acre-feet over a 100-year period. The optimum calculated by the Bayesian method is about twice that value at 127,570 acre-feet. The difference is significant in that the Bayesian estimate reflects the economic characteristics of the process.

As a check, the expected opportunity loss (EOL) is calculated to assure that the theory is correct. The expected opportunity loss is calculated from the following expression:

$$EOL(a) = \frac{1}{C} \int_{\beta_1}^{\beta_2} \int_{\alpha_1(\beta)}^{\alpha_2(\beta)} L(a,\overline{Z}) LIKELIH(\alpha,\beta \mid data)\, d\alpha\, d\beta \qquad (24)$$

The integration is carried out over all values of the optimum a* for different alternatives, a. Figure 7 illustrates the curve generated by repeatedly solving Equation 24. The optimum point found by this method was 127,000 acre-feet, which compares accurately with the value found by the search method. This additional evidence reinforces the conclusion that the estimate according to Schwalen's (1961) method may be too low. The mean of the 100-year pdf was calculated using the maximum likelihood estimates of the parameters α and β. This value was found to be 104,000 acre-feet. The mean was also calculated using numerical integration. This expression was:

$$\overline{Z}_{Bayes} = \frac{1}{C} \int_{\beta_1}^{\beta_2} \int_{\alpha_1(\beta)}^{\alpha_2(\beta)} (\overline{Z} \mid \alpha,\beta) LIKELIH(\alpha,\beta \mid data)\, d\alpha\, d\beta$$

The value of this expression was 106,000 acre-feet. This indicates that the estimate obtained which ignores the uncertainty in the process is conservative.

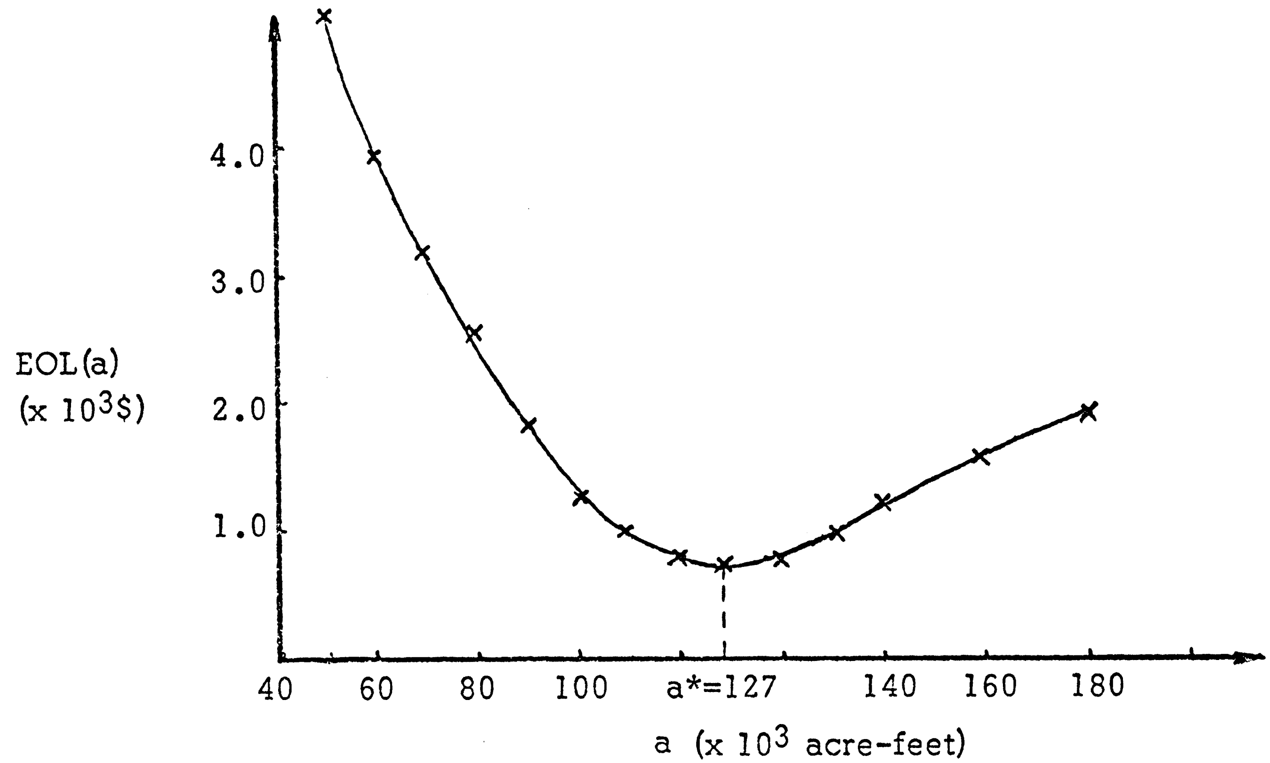


Figure 7. Expected opportunity loss for alternative a

CHAPTER 7

DISCUSSION AND CONCLUSIONS

The proposed method and model have several shortcomings. First, the Bayesian approach utilizes a linear loss function when, in fact, the losses may bear no resemblance to linearity. Due to some of the simplifications, the optimal solution may be inaccurate. The use of the indicator function may not have been valid. It can be shown that the pdf of the 100-year distribution of sediment approaches the normal distribution in the limit (Gupta, 1973); it might have been better to substitute a normal distribution instead of the indicator function. To do so would have compounded the numerical problems. The assumption that the prior distribution on the state variables is uniform may not have been entirely correct. One of the more difficult problems with the model itself is the lack of data. Only 33 events were available from which to obtain the estimates of α and β. As the number of data points was small, the maximum likelihood estimates of the parameters for the joint distribution of rainfall and duration, α and β, could not be obtained due to the nonconvergence of the maximum likelihood-estimating equations. The fact that the actual data from the Charleston watershed was not used may also have caused some variation. The elements of the universal soil-loss equation were designed for small watersheds, certainly smaller than the Charleston watershed. The random variable of time of concentration, a_2, probably should not have been constant. The estimation of the watershed constants in the equation is also subject to a great deal of variation.

The application of this type of analysis to a process in which individual events distributed in a gamma fashion and where an accumulation of some quantity is involved should be sought. There may be bacteriological population growth problems in which there are accumulations distributed in a gamma fashion where the event would be a reproductive cycle. It would be a natural extension of this work to analyze the sensitivity of the optimal solution to physical parameters, such as the watershed constants, or economic ones, such as interest rates and time between loan and reservoir operation, etc. Other Bayesian quantities, such as the value of perfect information or the expected net gain or sampling, could be computed. As an alternative approach to generating the event-based distribution from the conditional and marginal distributions, the transformation defined by Equation 4 of random variables may actually be performed to see if the pdf is indeed gamma. There are numerical difficulties associated with this approach, but it would provide a valid check of the gamma-2 approximation assumption in that the same results should be expected.

The approach taken in this thesis is based on empirical and heuristic considerations. It was felt that the model developed here represents nature more accurately in making an optimal choice than many other tools of the reservoir designer. Not only does it represent nature, but it also takes into account the economic considerations involved, which, for example, the method of Schwalen (1961) does not.

In conclusion, the following points have been demonstrated:

1. For the Atterbury Experimental Watershed, a bivariate gamma distribution could not be rejected for the joint pdf of rainfall depth $\underset{\sim}{x}_1$ and storm duration $\underset{\sim}{x}_2$ of the storms observed over that area.

2. The pdf of sediment yield for an individual storm is approximately distributed as a 2-parameter gamma pdf.

3. Since storms were found to arrive in a Poisson manner, the mean and variance of sediment yield for N seasons can be found without actually determining the pdf itself.

4. Restricting the regions of integration to the most likely values proved to be a great help in reducing computer times. Any loss of accuracy was felt to be minimal in terms of computer time savings.

5. The assumption that the distribution of Z is highly peaked may be justified due to the low coefficient of variation of the pdf. Furthermore, the variance decreases as more data are obtained with time.

6. The use of a linear regression model to relate the lifetime mean and the rainfall-duration parameters, α and β, was found to be acceptable because of the low residual sums of squares. Also, the error between the estimate with the largest deviation and the true value was approximately 2 percent.

7. The optimal design sediment capacity for the Charleston Dam is approximately 127,570 acre-feet in a 100-year period; that is, 127,570 acre-feet should be allotted to sedimentation over that time in order that flood protection would not be diminished during the project lifetime. An estimate of 127,000 acre-feet was obtained as the optimum design, using the expected opportunity loss method. This estimate served to verify the theory presented here. Thus, the two estimates indicate that the sedimentation allocation should be greater than previously thought. The calculation of the means with uncertainty and without uncertainty also supports this hypothesis.

NOTATION

A	Drainage area of watershed (square miles).
a	Alternative action or decision (may also be used to represent set of all possible decisions).
a^*	Optional decision.
a_o	Conversion constant $(484A^2(640)/12)$.
a_1	Constant (= 0.40).
a_2	Mean time of concentration of storm in hours.
C	Cropping-management factor.
K	Soil erodibility factor.
K_u, K_o	Coefficients of linear loss function.
LS	Slope length and gradient factor.
$\tilde{n}$	Mean number of events per season.
P	Erosion control practice factor.
Q	Runoff volume in inches.
q_p	Peak flow rate in cfs.
S	Watershed infiltration constant.
$\underset{\sim}{x}_1$	Random variable of rainfall amount in inches (R).
$\underset{\sim}{x}_2$	Random variable of storm duration in hours (D).
Z	Sediment yield amount in acre-feet.
$\bar{Z}_{Bayes}$	Bayes estimate of mean sediment yield for the lifetime distribution.
α	Parameter of $f(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$.
β	Parameter of $f(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$.
ε	Error bound on optimizing search accuracy.
λ_e, r_e^2	Parameters of 2-parameter gamma pdf $f_e(\underset{\sim}{Z})$.
μ_e, σ_e^2	Mean and variance of single event distribution $f_e(\underset{\sim}{Z})$.

μ_s, σ_s^2	Mean and variance of seasonal distribution $f_s(\underset{\sim}{Z})$.
$\mu_L(=\overline{Z}), \sigma_L^2$	Mean and variance of lifetime distribution $f(Z\|\alpha,\beta)$.
$\rho_u(\alpha,\beta)$	Least squares regression line for Z.
$\theta = \langle\alpha,\beta\rangle$	State variable vector.
$f(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$	Joint bivariate gamma pdf of rainfall amount and storm duration.
$f_e(\underset{\sim}{Z})$	Probability density function of sediment yield Z for a single event.
$F_e(\underset{\sim}{Z})$	Cumulative distribution function of sediment yield Z for a single event.
$f_s(\underset{\sim}{Z})$	Probability density function of sediment yield Z for a single season.
$f(Z\|\alpha,\beta)$	Lifetime probability density function of sediment yield Z.
$F(Z<a^*), F(a^*)$	Cumulative distribution of decision variable Z.
$LIKELIH(\alpha,\beta)$	Likelihood function of parameters α and β, parameters of the joint pdf of rainfall and duration.
$L(a,\theta)$	Economic loss or goal function for alternative a, depending on state variable θ.
EOL(a)	Expected opportunity loss for alternative a.
$PRIOR(\theta)$	Prior pdf of state variable θ.
$POST(\theta\|data)$, $f(\theta)$, $POST(\alpha,\beta\|data)$	Posterior distribution of state variables.
$I(Z_{[0,a^*]}\|\alpha,\beta)$	Indicator function for $f(Z\|\alpha,\beta)$ about $\overline{Z}$.

REFERENCES

Benjamin, J. R., and C. A. Cornell. 1970. Probability, Statistics, and Decision for Civil Engineers. McGraw-Hill, New York.

Borland, W. M., and C. R. Miller. 1958. Distribution of sediment in large reservoirs, Journal of the Hydraulics Division, ASCE, Vol. 1, No. HY2, April, p. 1587.

Crovelli, R. A. 1971. Stochastic models for precipitation. Ph.D. dissertation, Colorado State University, Ft. Collins.

Davis, D. R. 1971. Decision making under uncertainty in hydrology. Natural Resource Systems Technical Report No. 2, Department of Hydrology and Water Resources, University of Arizona, Tucson.

Davis, D. R., C. C. Kisiel, and L. Duckstein. 1972. Bayesian decision theory applied to design in hydrology. Water Resources Research, Vol. 8, No. 1, February, pp. 33-41.

De Groot, M. A. 1970. Optimal Statistical Decision. McGraw-Hill, New York.

Duckstein, L., M. Fogel, and C. C. Kisiel. 1972. A stochastic model of runoff-producing rainfall for summer type storms. Water Resources Research, Vol. 8, No. 2, April, pp. 410-421.

Feller, W. 1966. An Introduction to Probability Theory and Its Applications, Vol. 1. John Wiley, New York.

Gupta, V. K. 1973. A stochastic approach to space-time modeling of rainfall. Natural Resources Systems Technical Report No. 18, Department of Hydrology and Water Resources, University of Arizona, Tucson.

Howard, R. 1966. Decision analysis: applied decision theory. Proceedings, Fourth International Conference on Operations Research. John Wiley, New York, pp. 55-71.

Jacobi, S. 1971. Economic worth of sediment load data in a statistical decision framework. Ph.D. dissertation, Colorado State University, Ft. Collins.

Raiffa, H., and R. Schlaifer. 1961. Applied Statistical Decision Theory. Harvard University Press, Cambridge, Massachusetts.

Schwalen, H. C. 1961. Suspended sediment and chemical analysis of the San Pedro River at Charleston, Arizona. Report No. 202, Arizona Agricultural Experiment Station.

Task Committee for Preparation of Manual on Sedimentation of the Committee on Sedimentation of the Hydraulics Division. 1973. Journal of the Hydraulics Division, ASCE, Vol. 99, No. HY4, April.

Todorovic, P., and V. Yevjevich. 1969. Stochastic process of precipitation. Colorado State University Hydrology Paper No. 35.

Williams, J. R., and H. D. Berndt. 1972. Sediment yield computed with universal equation. Journal of the Hydraulics Division, ASCE, Vol. 98, No. HY12, December, pp. 2087-2098.

Williams, J. R., and R. W. Hahn. 1973. Hymo: problem-oriented computer language for hydrologic modeling, user's manual. Agricultural Research Service, Technical Publication ARS-S-9.

Wischmeier, W. H. 1959. A rainfall erosion index for a universal soil-loss equation. Proceedings, American Soil Science Society, Vol. 23, No. 3, pp. 246-249.

Wischmeier, W. H. 1960. Cropping-management factor evaluation for a universal soil-loss equation. Proceedings, American Soil Science Society, Vol. 24, July-August, pp. 418-424.

Wischmeier, W. H., and D. D. Smith. 1960. A universal soil-loss equation to guide conservation farm planning. Proceedings, 7th International Congress of Soil Science, Madison, Wisconsin.

Wischmeier, W. H., and D. D. Smith. 1965. Predicting rainfall-erosion losses from cropland east of the Rocky Mountains. Agriculture Handbook 282, Agricultural Research Service.

Wischmeier, W. H., D. D. Smith, and R. E. Uhland. 1958. Evaluation of factors in the soil-loss equation. Agri. Engr., Vol. 39, pp. 458-462.

Woolhiser, D. A., and P. Todorovic. 1974. A stochastic model of sediment yield for ephemeral streams. Proceedings, USDA-IASPS Symposium on Statistical Hydrology, Tucson, Arizona. USDA Misc. Pub. 1275, pp. 232-246.

Yakowitz, S., L. Duckstein, and C. C. Kisiel. 1974. Decision analysis of a gamma hydrologic variate. Water Resources Research, Vol. 10, No. 4, August, pp. 695-704.