

## RESERVOIR DESIGN UNDER RANDOM SEDIMENT YIELD

by

L. Duckstein,<sup>1</sup> F. Szidarovszky,<sup>2</sup> S. Yakowitz<sup>3</sup>

### INTRODUCTION

The purpose of this paper is to provide a Bayesian methodology for designing a reservoir subject to random sedimentation and for evaluating the worth of data used in that decision. The approach, which is illustrated by the case of the Charleston Watershed in Southern Arizona, is essentially a continuation of the research described in Szidarovszky *et al.* (1976), where no decision analysis was performed. Furthermore, the elements of the present model are the same as those in Smith *et al.* (1974), but the methodology is different, as it will be explained later.

The design of a storage reservoir subject to sedimentation requires that sediment yield be forecasted for the useful lifetime of the structure (Jacobi, 1971). An underdesign of the reservoir leads to a shortening of the lifetime (McHenry, 1974) or to additional removal costs, while an overdesign constitutes an opportunity loss.

Sediment yield is most commonly estimated by a linear regression of mean annual yield on watershed and climatic characteristics (Flaxman, 1972; Nordin and Sabol, 1973; Jansen and Painter, 1974; Guymon, 1974; McPherson, 1975). The reservoir is designed on the basis of these estimated annual quantities; however, as pointed out in Weber *et al.* (1976), serious questions may arise about the validity of such regression results when the wrong transformation is used or when the assumptions of the model are violated. Furthermore, sediment yield is, generally speaking, a random variable since it is the result of random precipitation phenomena. This is especially true in ephemeral flow regions, where sediment accumulated over one or more seasons may be modelled as the sum  $S$  of a random number of random sediment yield events (Woolhiser and Todorovic, 1974). Then, for a Bayesian analysis accounting for the uncertainty on  $S$ , the distribution function (DF) of  $S$ ,  $F_S(s)$ , must be estimated. Rare are the cases when available time series data of sediment yield events are long enough (and stationary!) to estimate  $F_S(s)$ ; on the other hand, since precipitation records are usually adequate, it is possible to use models transforming rainfall into sediment yield to estimate the DF of  $S$ .

Sediment produced by an individual precipitation event may be estimated by the use of a physical model (Bennett, 1974; Renard and Laursen, 1975) or an empirical model such as the Universal Soil Loss Equation (Wischmeier and Smith, 1960; Onstad and Foster, 1976). The latter equation is in fact used here, as it was done in Smith *et al.* (1974), with the following fundamental difference: whereas Smith *et al.* calculated the DF of  $S$  by a transformation of random variables, a simulation approach is used in the present investigation. The next section of the paper describes the methodology, from input analysis to decision-making, which is then applied to the Charleston Watershed in the Application section.

### METHODOLOGY

The methodology section is divided into two parts:

1. The modeling of rainfall and of sediment yield, which leads to the generation of synthetic time series of sedimentation events.
2. The decision-making phase, which includes choosing a reservoir design and evaluating the decision in terms of cost of uncertainty.

Modeling rainfall and sediment yield events. (a) Rainfall events. In a semi-arid region, sedimentation occurs mostly in summer; then it is only necessary to use summer season precipitation models and data to estimate yearly sediment yield. The stochastic rainfall model assumes a Poisson process of precipitation events of relatively short duration (Duckstein *et al.*, 1972); events are defined by a pair of random variables  $R = (X_1, X_2)$ , where  $X_1$  is the effective rainfall depth and  $X_2$  the duration. The joint probability density function (pdf) of  $R$  is assumed to have the form proposed by Crovelli (1971):

<sup>1</sup>Professor on joint appointment, Depts. of Systems & Industrial Engineering and Hydrology & Water Resources, University of Arizona, Tucson, AZ 85721.

<sup>2</sup>Professor, Dept. of Numerical Methods, Eötvös University, Muzeum körút, Budapest VIII, Hungary.

<sup>3</sup>Associate Professor, Dept. of Systems & Industrial Engineering, University of Arizona, Tucson, AZ 85721.

$$f(x_1, x_2) = \begin{cases} \alpha\beta e^{-\beta x_1} (1 - e^{-\alpha x_2}), & \text{if } 0 \leq x_2 \leq \frac{\beta}{\alpha} x_1 \\ \alpha\beta e^{-\alpha x_2} (1 - e^{-\beta x_1}), & \text{if } 0 \leq \frac{\beta}{\alpha} x_1 \leq x_2, \end{cases} \quad (1)$$

where  $\alpha$ ,  $\beta$ ,  $x_1$ , and  $x_2$  are all positive. The values of the parameters  $\alpha$  and  $\beta$  are estimated by use of the maximum likelihood method for the marginal distributions. As shown in Crovelli (1971), the marginal density functions are:

$$f_1(x_1) = \beta^2 x_1 e^{-\beta x_1} \quad (2)$$

$$f_2(x_2) = \alpha^2 x_2 e^{-\alpha x_2} \quad (3)$$

A separate maximum likelihood estimation of  $\alpha$  and  $\beta$ , offers no particular difficulty and is given in Szidarovszky *et al.* (1976). Marginal distributions are used so that the estimated values  $\hat{\alpha}$  and  $\hat{\beta}$  may be evaluated by the Kolmogorov-Smirnov goodness-of-fit test, which has only been developed for scalar random variables. A simulation method is used to compute random bivariate values of  $\underline{R}$  as in Szidarovszky *et al.* (1976).

Next, since a Poisson process is an acceptable hypothesis for summer rainfall (Gupta, 1973), the time interval  $t$  between two rainfall events is assumed to be exponentially distributed random variable with the distribution function (Fogel and Duckstein, 1969; Smith, 1975).

$$F(t) = 1 - e^{-\lambda t}, \quad (4)$$

where  $\frac{1}{\lambda}$  is the expected number of storm events per year. Random values of the time interval can be computed from uniformly distributed random numbers using a transformation method. Thus, by generating triplets  $(x_1, x_2, t)$ , synthetic time series of rainfall events are obtained.

(b) Sediment yield per event. A sediment event is defined here as the sediment yield from any runoff-producing rainfall event. The Universal Soil-Loss Equation in its modified version (Wischmeier and Smith, 1965), is taken as a basis for estimating sediment yield per simulated rainfall event:

$$Z_c = 95(Qq_p)^{0.56} K C P L_s \quad (\text{tons})$$

where

$Z_c$  = Sediment yield per event 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 practice factor

$L_s$  = Slope length and gradient factor

Algorithms developed by Williams and Hann (1973) are used for the calculation of the values of  $K$ ,  $C$ ,  $P$ , and  $L_s$ . The values of  $Q$  and  $q_p$  are computed from the Soil Conservation Service formulas (Kent, 1968).

$$Q = \frac{x_1^2}{x_1 + S}, \quad (\text{inches}) \quad (5)$$

where

$Q$  = Runoff volume in inches

$x_1$  = Effective rainfall in inches (rainfall less a constant initial abstraction)

$S$  = Watershed infiltration constant

and

$$Q_p = \frac{484AQ}{a_1\lambda_2^2 + a_2}, \quad (\text{cfs}) \quad (6)$$

where

A = Drainage area of watershed in square miles

$a_1$  = Constant (.50)

$\lambda_2$  = Storm duration in hours

$a_2$  = Time of concentration of storm in hours (assumed constant for a given watershed)

Combining equations (4), (5), and (6) yields

$$Z = \frac{a_0\lambda_1^4}{(\lambda_1 + S)^2(a_1\lambda_2 + a_2)} 0.56 W, \quad (\text{acre-feet}) \quad (7)$$

where

$$a_0 = 484A^2 (5280 \text{ ft/mi})^2 / (4.356 \times 10^4 \text{ ft}^2/\text{acre}) / (12 \text{ in/ft}) \quad (8)$$

$$W = 95KPL_5(2000 \text{ lb/ton}) / (\text{mean sediment density} = 70 \text{ lb/ft}^3) \div (4.356 \times 10^4 \text{ ft}^2/\text{acre}) \quad (9)$$

In equation (7)  $a_0$ ,  $a_1$ ,  $a_2$ , S, W are assumed constant and  $Z$  is a random variable which can be considered as a transformed variable of the two dimensional variate  $(\lambda_1, \lambda_2)$ . All the elements necessary to generate a time series of sediment events have now been defined and the following algorithm can be applied to simulate the accumulated sediment yield during N years.

- (1) Let T: = 0, ZZ: = 0.
- (2) Simulate a realization t of the exponential by distributed random variable  $\lambda$  with parameter  $\lambda$  and let T = T+t.
- (3) If T > N years, then go to (5). If T ≤ N years, then go to (4).
- (4) Simulate a bivariate realization of  $(\lambda_1, \lambda_2)$ . Using equation (7), calculate the corresponding value of Z, let ZZ: = ZZ+Z, and go to (2).
- (5) The value of ZZ is a sample element of the accumulated sediment yield.

When enough sample elements of ZZ have been simulated, an empirical DF usable in decision analysis is obtained.

Decision-making. (a) Objective function under natural uncertainty. The loss function derived by Smith (1975) on the basis of the work of Jacobi (1971) is being used here; namely,

$$L(a, Z) = \begin{cases} 210.19(a-Z) & \text{if } a > Z \text{ (overdesign)} \\ 888.89(Z-a) & \text{if } a < Z \text{ (underdesign)} \end{cases} \quad (10)$$

where a is the decision or action in acre-feet. This is an opportunity loss type of objective function since there may exist an optimum action for which  $E(L(a, Z)) = 0$ .

The objective function G(a) is ordinarily the expected value of  $L(a, Z)$  under natural uncertainty. In the present study, G(a) was computed by generating 1000 samples of total sediment yield  $Z_i$  as described above and calculating the arithmetic average

$$\begin{aligned} G(a) &= E_Z[L(a, Z)] \\ &= \frac{1}{1000} \sum_{i=1}^{1000} L(a, Z_i) \end{aligned} \quad (11)$$

These calculations were repeated for given discrete values of a. The optimum value  $a^*$  of the decision a and  $G(a^*)$  of the objective function G(a) were obtained by parabolic interpolation between the three lowest points of G(a).

(b) Goal function under natural and parameter uncertainty stems from estimating the parameters  $\alpha$ ,  $\beta$  of pdf (1) from a finite (and small) sample. This sample or parameter uncertainty leads to considering

that  $\alpha$  and  $\beta$  are random variables  $\alpha, \beta$  distributed according to a prior pdf  $f_{\alpha, \beta}(\alpha, \beta)$  (Davis et al., 1972). The objective function  $G(a)$  is now called a goal function and is written as

$$g(a, \alpha, \beta) = \int_0^{\infty} L(a, Z) f(Z | \alpha, \beta) dZ \quad (12)$$

and the Bayes Risk (BR) is

$$BR(a) = \int_0^{\infty} \int_0^{\infty} g(a, \alpha, \beta) f_{\alpha, \beta}(\alpha, \beta) d\alpha d\beta \quad (13)$$

The Bayes decision  $a^*$  minimizes BR

$$BR(a^*) = \min_a BR(a) \quad (14)$$

It is theoretically appropriate and computationally convenient to choose a conjugate prior pdf  $f_{\alpha, \beta}(\alpha, \beta)$  (Raiffa and Schlaifer, 1961). In the present study, the estimations of  $\alpha$  and  $\beta$  were based on a univariate gamma distribution; thus, the results derived in Yakowitz et al. (1974) concerning conjugate densities for pdf (2) and (3) can be used, namely, that the conjugates are, respectively:

$$\frac{\beta^{2n} y_1^{2n+1} e^{-y_1}}{\Gamma(2n+1)} \quad \text{and} \quad \frac{\alpha^{2n} y_2^{2n+1} e^{-y_2}}{\Gamma(2n+1)}, \quad (15)$$

where  $y_1 = \sum_{k=1}^n x_1(k)$  and  $y_2 = \sum_{k=1}^n x_2(k)$ ;  $x_1(k), x_2(k), k = 1, \dots, n$  denote simultaneous sample elements of the variables  $(X_1, X_2)$ .

With the above elements, the simulation method leading to a Bayes decision consists of the following steps:

- (1) Let  $T = 0, ZZ = 0$ .
- (2) Simulate random value of  $\lambda$  and using this  $\lambda$  as a random value of  $t$ , and let  $T = T+t$ .
- (3) If  $T > N$  years, then go to (6). If  $T \leq N$  years then go to (4).
- (4) Simulate random values of  $\alpha, \beta$  using the pdf's (15) and the method of generating two-parameter gamma distributions shown (Szidarovszky et al., 1976). Go to (5).
- (5) Using the generated values of  $\alpha, \beta$  simulate the random simultaneous value of  $(X_1, X_2)$ , as in step (4) of the section on Modeling Rainfall and Sediment Yield Events, and calculate the corresponding value of  $Z$ , and let  $ZZ = Z$ . Go to (2).
- (6) The value of  $ZZ$  is a sample element of the accumulated sediment yield. Go to 1 until  $f_Z(Z | \alpha, \beta)$  is defined, then go to (7).
- (7) Calculate  $g(a | \alpha, \beta)$  as shown in Eq. 11 for  $G(a)$ . Calculate the Bayes Risk (Eq. 13). Go to (8).
- (8) Find the Bayes decision  $a^*$  and minimum Bayes Risk  $BR(a^*)$ .

Note that the uncertainty in the interarrival time between events (exponential parameters) which describes the counting distribution of precipitation events has been taken into account in Step 2 of the aforementioned procedure. The methodology is now set to study parameter uncertainty in the following three cases:

- Case I, uncertainty in the bivariate parameters of the pdf of  $(X_1, X_2)$  only.
- Case II, uncertainty in the exponential parameter of interarrival time  $t$ .
- Case III, uncertainty in both sets of parameters.

(c) Sensitivity of the decision to available data. The simulations were run under the successive assumptions of record sample size of  $N = 5, 10, 30, 60$ , and  $90$ . A run was also made without uncertainty which is equivalent to assuming an infinitely long sample record.

A complete evaluation of the decision can theoretically be made in terms of expected opportunity loss, which represents the worth of perfect information, and (expected) opportunity loss, which gives an economic measure of the economic worth of another sample point (Davis et al., 1972; Musy and Duckstein, 1976). For this purpose, one would have to improve the efficiency of the simulation and computations,

test the computing time becomes prohibitive.

### APPLICATION

For the case study of the Charleston Watershed in Southern Arizona, the data are:

$$\begin{aligned} K &= 0.60; & C &= 0.80; & P &= 0.10; \\ L_S &= 0.50; & S &= 2.50; & A &= 1220; \\ a_1 &= 0.50; & a_2 &= 6.989. \end{aligned}$$

Using equations (8) and (9) yields

$$a_0 = 3.83576 \times 10^{10} \text{ and } W = 1.49547 \times 10^{-3}.$$

The parameters  $\alpha$  and  $\beta$  are estimated from 91 sample elements obtained during 90-day summer periods (1 July-30 September) ( $m = 91$ )

$$\hat{\alpha} = 1.84 \text{ and } \hat{\beta} = 2.58.$$

Applying a Kolmogorov-Smirnov goodness-of-fit test, the marginal distributions could not be rejected at the 10% level of significance. The parameter  $\lambda$  of the exponential interval time between storm events has an estimate of  $\hat{\lambda} = 0.08$ . The lifetime of a project is taken as 100 years.

Since the sample elements  $X_1(k)$ ,  $X_2(k)$ ,  $k = 1, \dots, 91$ , are given with two decimal places, an upper bound of 0.001 has been chosen for the accuracy of the simulation algorithm.

Using the methodology described previously, 1000 1000-sample elements of accumulated sediment yield are simulated. The mean and standard deviation of these sample elements is calculated assuming that  $N = 5, 10, 30, 60, 90, \infty$  sample points are available in the three cases of sample uncertainty described in the previous section. The results are shown in Table 1; the standard deviation of the 100-year sediment production increases as more uncertainty is introduced into the model, and decreases as the record becomes longer, rapidly at first, leveling off afterwards. These results are also illustrated in Figure 1, where the variance has been plotted to increase the resolution. The mean shows the same trend as the standard deviation but varies only a little; this is due to the choice of the loss function (Equation 10) (piecewise linear with a cusp).

TABLE 1  
MEAN AND STANDARD DEVIATION OF 100-YEAR SEDIMENT PRODUCTION,  
IN THOUSANDS OF ACRE-FEET

N	CASE I Uncertainty in the Bivariate Parameters		CASE II Uncertainty in the Exponential Parameter		CASE III Total Uncertainty	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
5	110.7	92.30	102.3	68.95	108.6	10.97
10	106.9	71.43	105.3	60.90	109.4	86.64
30	105.7	55.63	104.5	51.79	105.0	58.63
60	107.1	52.63	102.4	51.22	103.9	52.13
90	105.0	51.45	103.4	50.77	104.1	51.35
$\infty$	103.3	48.46	(no uncertainty)			

Next, Table 2 shows the Bayes decision and minimum Bayes Risk under the same assumptions on record length and type of uncertainty cases as in Table 1. Figure 2 illustrates how the entries in the last row of Table 2 are calculated ( $a^* = 143.6 \times 10^3$  acre-feet,  $BR(a^*) = \$16.01 \times 10^6$ ). The Bayes decision and Bayes Risk are quite sensitive to N which reflects the effect of the variance. Figures 3 and 4, which are graphical representations of the results in Table 2, were drafted for the purpose of dramatizing the effect of information available in the decision of showing that simulation results are subject to small jumps (Figure 3 especially). Note that the results are commensurate with those of Smith *et al.* (1974). The computing times were as follows:

N	5	10	30	60	90
Computer Time (sec.)	69.0	74.8	102.3	144.8	180.5

TABLE 2  
BAYES DECISION AND OPTIMUM BAYES RISK

N	CASE I Uncertainty in the Bivariate Parameters		CASE II Uncertainty in the Exponential Parameter		CASE III Total Uncertainty	
	Bayes Decision $10^3$ acre-feet	Bayes Risk $10^6$ \$	Bayes Decision $10^3$ acre-feet	Bayes Risk $10^6$ \$	Bayes Decision $10^3$ acre-feet	Bayes Risk $10^6$ \$
5	163.1	31.32	153.3	23.40	165.2	35.93
10	153.2	24.18	153.2	20.35	159.2	27.90
30	146.4	18.92	148.5	16.87	152.4	19.68
60	148.0	17.87	146.1	16.71	146.0	17.27
90	143.4	17.03	145.8	16.53	146.0	17.20
$\infty$	143.6	16.01	(no uncertainty)			

The main points reached in this investigation are:

- (1) Synthetic time series of bivariate rainfall events can be combined with an empirical sediment yield equation (such as the Universal Soil Loss Equation) to design a dam subject to this natural uncertainty in sedimentation.
- (2) The methodology is capable of accounting for any combination of sample uncertainty (rainfall, counting pdf, or both) in the design, leading to a Bayes decision.
- (3) The mean 100-year sediment yield varied little with simulated record length N while both Bayes decision and Bayes risk decrease strongly when N is increased; this shows the advantage of a Bayes approach over the mere use of mean estimates, namely, that of translating the variance of the record into economics.

#### ACKNOWLEDGEMENTS

The work upon which this paper is based was supported in part by funds provided by the United States Department of the Interior, Office of Water Resources Research, as authorized under the Water Resources Act of 1964, Project No. B-043AZ, entitled, "Practical Use of Decision Theory to Assess Uncertainties about Actions Affecting the Environment," as well as National Science Foundation Grant No. GF-38183, entitled, "Cooperative Research on Decision-making under Uncertainty in Hydrologic and Other Resource Systems."

#### REFERENCES CITED

- Bennett, J. P., "Concepts of Mathematical Modeling of Sediment Yield, Water Resources Research, 10(3), 1974, pp. 485-492.
- Crovelli, R. A., "Stochastic Models for Precipitation," Ph.D. Dissertation, Colorado State University, Fort Collins, 1971.
- Davis, D. R., C. C. Kisiel and L. Duckstein, "Bayesian Decision Theory Applied to Design in Hydrology," Water Resources Research, 8(1), 1972, pp. 33-41.
- Duckstein, L., M. Fogel and C. Kisiel, "A Stochastic Model of Runoff Producing Rainfall for Summer-type Storms," Water Resources Research, 8(2), 1972, pp. 410-421.
- Flaxman, E. M., "Predicting Sediment Yield in Western United States," Journal of the Hydraulics Division, ASCE, 98(HY12), 1974, pp. 41-51.
- Gupta, V. K., "A Stochastic Approach to Space-time Modeling of Rainfall," Technical Report No. 18, Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, 1973.
- Guymon, G. L., "Regional Sediment Yield Analysis of Alaska Streams," Journal of the Hydraulics Division, ASCE, 100(HY1), 1974, pp. 41-51.
- Jacobi, S., "Economic Worth of Sediment Load Data in a Statistical Decision Framework," Ph.D. Dissertation, Colorado State University, Fort Collins, 1971.
- Jansen, J. M. L. and R. B. Painter, "Predicting Sediment Yield from Climate and Topography," Journal of Hydrology, 21(4), 1974, pp. 371-380.
- McHenry, J. R., "Reservoir Sedimentation," Water Resources Bulletin, AWRA, 10(2), 1974, pp. 329-337.
- McPherson, H. J., "Sediment Yields from Intermediate-sized Stream Basins in Southern Alberta," Journal of Hydrology, 25(314), 1975, pp. 243-257.
- Musy, A. and L. Duckstein, "A Bayesian Approach to Tile Drain Design," to appear, ASCE Journal of Irrigation and Drainage, 1976.
- Nordin, C. F. and G. V. Sabol, "Estimating Average Sediment Yields from Annual Streamflow and Sediment Records," Proceedings of the IAMR, International Symposium of River Mechanics, Bangkok, Thailand, Vol. 1, January, 1973, pp. 93-104.
- Onstad, C. A. and G. R. Foster, "Erosion Modeling on a Watershed," Transactions of the American Society of Agricultural Engineers, 18(2), 1975, pp. 288-292.
- Raiffa, H. and R. Schlaifer, Applied Statistical Decision Theory, Harvard University Press, Cambridge, Mass., 1961.
- Renard, K. G. and E. M. Laursen, "Dynamic Behavior Model of Ephemeral Stream," Journal of the Hydraulic Division, ASCE, 101(HY5), 1975, pp. 511-528.
- Smith, J. H., "A Stochastic Sediment Yield Model for Bayesian Decision Analysis Applied to Multipurpose Reservoir Design," Master's Thesis, Dept. of Systems & Industrial Engineering, University of Arizona, Tucson, Arizona, 1975.
- Smith, J., M. Fogel and L. Duckstein, "Uncertainty in Sediment Yield from a Semi-arid Watershed," Proceedings, 18th Annual Meeting of the Arizona Academy of Sciences, Flagstaff, Arizona, April, 1974, pp. 258-268.
- Szidarovszky, F., S. Yakowitz and R. Krzysztofowicz, "A Bayes Approach for Simulating Sediment Yield," to appear, Journal of Hydrological Sciences, 1976.
- Weber, J., M. Fogel and L. Duckstein, "The Use of Multiple Regression Models in Predicting Sediment Yield," to appear, Water Resources Bulletin, 1976.
- Wischmeier, W. H. and D. D. Smith, "A Universal Soil-Loss Equation to Guide Conservation Farm Planning," Proceedings, 7th International Congress of Soil Science, Madison, Wisconsin, 1960.
- Wischmeier, W. H. and D. D. Smith, "Predicting Rainfall Erosion Losses from Cropland East of the Rocky Mountains," Agriculture Handbook, 282(ARS), May, 1965.
- Woolhiser, D. A. and P. Todorovic, "A Stochastic Model of Sediment Yield for Ephemeral Streams," Proceedings, USDA-IASPS, Symposium on Statistical Hydrology, Tucson, Arizona, 1971; USDA Misc. Publication No. 1275, 1974, pp. 295-308.

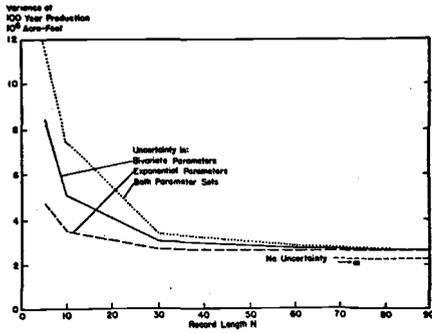


Figure 1. Variance of total sediment yield as a function of sample record length.

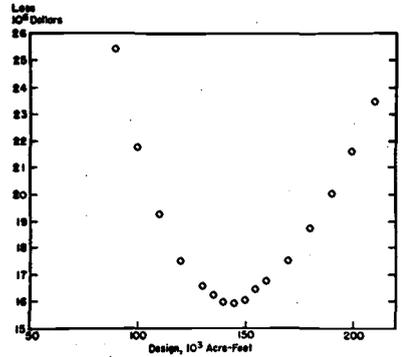


Figure 2. Objective function under natural uncertainty.

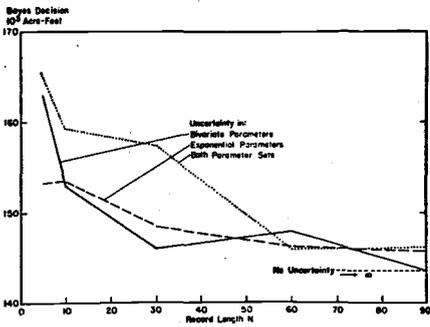


Figure 3. Bayes decision  $a^*$  versus sample record length.

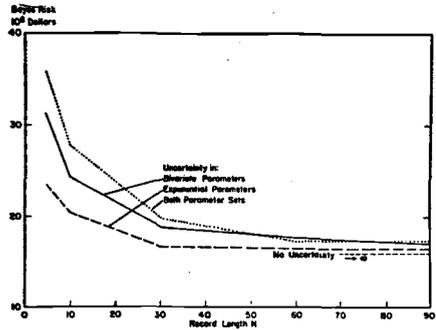


Figure 4. Bayes Risk versus sample record length.