# MAXIMIZATION OF NET BENEFIT FROM A STREAMGAGE

by

Marshall E. Moss

---

A Thesis Submitted to the Faculty of the

DEPARTMENT OF CIVIL ENGINEERING

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

1969

## STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Marshall E. Moss

## APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

E. M. Laursen
E. M. LAURSEN
Professor of Civil Engineering

Dec 6, 1968
Date

# ACKNOWLEDGMENTS

The author is sincerely grateful for the advice and critical review of this thesis by several of his colleagues within the U.S. Geological Survey, in particular David R. Dawdy and Dr. N.C. Matalas. The guidance of the thesis director, Dr. E. M. Laursen, and the review by the other committee members, Professors Carmody and Lewis, is greatly appreciated.

Special commendations are extended to typists, Joan Barnes, Glenn Golden, Ines Gore and Joan Wilcox.

## TABLE OF CONTENTS

TABLE OF CONTENTS--*Continued*

# LIST OF ILLUSTRATIONS

# LIST OF TABLES

Table | Page

# ABSTRACT

A scheme that will aid managers of streamgaging stations in making decisions of how long to operate a particular station and how frequent to make discharge measurements during the period of operation, is presented. This scheme is based on the definition of the present value of data and an economic marginal analysis. An example of the scheme using data from the gaging station at Arroyo Seco near Soledad, California in a conservation-reservoir design yielded a range in optimum length of gaging station record from 9 to 56 years. The nine year value was the most economically efficient. Because of the stability of the stage-discharge relation, the optimum frequency of discharge measurement at this station was indeterminant.

The sensitivity of the optimum length of record to its parameters was tested. The lower limit was found to be highly sensitive to a parameter that reflected both the streamflow characteristics and the anticipated level of development of the stream. The parameters, interest rate and cost of obtaining the data, were found to be

somewhat less sensitive. In all practical cases the upper limit of the record was determined by the planned date of initiation of the conservation project.

# CHAPTER 1

# INTRODUCTION

In classifying natural resources, surface water streams can be designated flow resources. To make the best use of a flow resource, an estimate of its flow characteristics must be determined. Gaging a surface water stream through a period of time is the predominant method of estimating its characteristics. Because this period is not of infinite length, the record of flow can only be considered a time sample. The accuracy of the estimate is a function of both the accuracy of the individual measurements in the record and the length of time through which the sample is obtained.

The prime decision facing the operator of a gaging station network is: how does he expend his money and time so that he can obtain the most economically efficient estimate of the flow characteristics of the streams that he is gaging; that is, how long should he operate the stations and how accurate should the operation be?

Langbein (13)[1] investigated a scheme based on correlation and regression analyses between a project design

---

[1] Numerals in parentheses refer to corresponding items in the List of References.

station and a long term base station. His scheme yielded an estimate of the optimum length of time to operate the project design station. Two prerequisites to the use of his method, however, are: the availability of a base station of sufficient record length hydrologically proximate to the design station; and the arbitrary determination of the desired accuracy of the determination of the mean annual runoff at the project design station. In a later study Fiering (7) presented a model that selects an optimal set of base stations from a larger set of stations that are being operated within a specified area. The decision criterion used by Fiering is the maximization of the information content of the gaging network. The maximization process is constrained only by budgetary limitations. This method yields no information about the length of time that a new station should be operated. Neither of these studies relate to the problem of accuracy requirements for the individual determinations within the record.

Application of operations research to the definition of decision models in the water resources field has been successfully accomplished by Allison (1). In his study Allison defined the optimum types and quantities of data that were required to construct a model of a ground-water basin that would aid ground-water managers in making the best use of their resource. In a similar manner this paper presents a scheme that will aid the managers of gaging stations in making the proper decisions with respect to the optimum length of record and the optimum frequency of discharge measurements. The optimum length of record and its associated optimum frequency of measurement will be designated the optimum operating procedure. In determining this procedure the scheme analyzes only the economic efficiency of an operation, and no budgetary or administrative constraints are considered. The criterion for determining the highest value of economic efficiency is the maximization of the present worth of the net benefits derived from operating a station.

The following sections present the scheme in detail. Because this study is economically oriented, the key to its applicability is the accurate determination of monetary values for the costs and benefits accrued from the operation of a gage. The procedures used to define these monetary values are not exclusive. The scheme presented here is compatible with any procedures that define the costs and/or the benefits with sufficient accuracy.

# CHAPTER 2

# THE COST OF DATA

The costs of obtaining streamflow data include costs of reconnaissance, construction of the gage, construction of an artificial control, field measurements, computation and publication of the record, and overhead. For the purpose of this study all costs except those of field measurements and for the construction of a control are considered fixed for a particular gaging station. The total cost of obtaining a streamflow record of N years length can be stated algebraically,

$$C_t = C_i + dC_c + N\ (mC_m + C_r) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

in which $C_t$ = the total cost; $C_i$ = the initial costs of reconnaissance and construction of the gage; $C_c$ = the cost of constructing an artificial control; d = a decision variable whose value is either one or zero; m = the number of discharge measurements made per year; $C_m$ = the cost of one discharge measurement; $C_r$ = the annual costs of overhead and record computation and publication. $C_r$ is assumed to be independent of the frequency of discharge measurements. This assumption is not completely valid because the cost of computing the discharge record can be influenced to a minor extent by the number of discharge measurements made during the period of record being computed. The factors $C_i$, $C_c$, $C_m$, and $C_r$ are assumed to be time invariant; therefore the marginal cost of obtaining an additional year of record is,

$$C_a = \partial C_t / \partial N = mC_m + C_r \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

The accuracy of computation of the discharge record is a function of two factors that enter into the cost analysis. Accuracy is directly related to the frequency of the discharge measurements. A method of graphically relating the standard error of computation of discharge, which is the inverse of accuracy, to the time period between discharge measurements has been presented by Burkham and Dawdy (3). The period between measurements is the reciprocal of the frequency of measurement. In a manner less amenable to precise description, the accuracy also depends upon the stability of the control of the stage vs. discharge relation.

## Cost vs. Accuracy

The error analysis procedure of Burkham and Dawdy was developed for alluvial streams. It can be used, however, on any stream where accuracy is a function of frequency of discharge measurement. The only necessary modification of this procedure is that values of the standard error of computation must be determined in dimensionless units of mean discharge. The procedure requires an existing record of discharge measurements and stream stage at the gage site; therefore this scheme cannot be used at the site until an initial period of record is available for analysis.

Implementation of the analysis requires that a fixed number of the discharge measurements be randomly selected for use in computing the discharge record. This set of measurements is referred to as the analysis group. The remaining measurements are utilized as a control group to evaluate the accuracy of the computed record. If the residuals between the computed and measured discharges in the control group are homoscedastic - their variance remains constant with respect to the measured discharge - the standard deviation of the residuals is an estimate of the standard error of computation for an instantaneous determination of discharge. This estimate is biased, however, because in addition to the computational error the standard deviation of the residuals contains a component of the error of measurement. The probable value of this measurement error was determined by Carter and Anderson (4). An unbiased estimate of the standard error of computation of an instantaneous discharge is

$$S_c = (S_R^2 - S_m^2)^{1/2} \text{ .........................} \quad (3)$$

in which $S_R$ = the standard deviation of the residuals, and $S_m$ = the standard error of measurement.

If the residuals are not homoscedastic, they are grouped by ranges of measured discharge over which homoscedasticity can be assumed to occur. The standard error of computation is then estimated from the pooled standard deviations within the groups. This estimate must also be corrected for the bias of the measurement error by use of equation 3, in which the pooled variance is substituted for $S_R^2$. The pooled variance is obtained by an equation that is a slight modification of the equation given by Dixon and Massey (6) for pooled estimates of variances. This estimate is

$$V_R = \sum_{i=1}^{k} P_i S_i^2 / P_s \text{ .........................} \quad (4)$$

in which i = an index of the discharge group; k = the number of nonzero groups; $P_i$ = the part of time the discharge is in range i; $S_i$ = the standard deviation of the residuals in group i; $P_s$ = the summation of the $P_i$'s, That part of the time during which no flow occurs is not considered in equation 4. It is not considered because it is assumed that zero discharge is determined with sufficient accuracy that the errors of computation that occur in this discharge group are negligible.

Standard errors of computation are computed for several sets of discharges in the analysis group. The size of each set is selected so that it represents a different frequency of measurement. The number of sets is made large enough so that a graphical relation between $S_c$ and the period between measurements can be established.

An equation for the relation between the standard error of computation of the mean discharge for a period of time was given by Burkham and Dawdy. For the mean annual discharge,

$$S_A = (S_c^2/m)^{1/2} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

in which $S_A$ = the standard error of computation of the annual mean discharge. Because m is equal to 365 divided by the period between measurements, expressed in days, equation 5 can be written thus

$$S_A = (p\ S_c^2/365)^{1/2} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (5a)$$

in which p = the average period of time between discharge measurements. From the form of equation 5a it is apparent that the relation of $S_A$ vs. p can be extracted from the graph of $S_c$ vs. p.

The preceding error analysis is based on the assumption that an artificial control is not built at the gaging station. Under this assumption the variable d in equation 1 would be equal to zero. The factors $C_i$, $C_m$, and $C_r$ are evaluated through a simple cost analysis of former gages that have been operated under field conditions similar to those of the gage being investigated. The annual cost of obtaining a record with a computational error of a specified amount is determined from the combination of the graphical relation of $S_A$ to p and equation 2. With a length of record and a requisite computational error specified the total cost of obtaining a record can be derived from a similar combination using equation 1 in place of equation 2. These relations of annual cost and total cost are displayed graphically.

## The Effects of an Artificial Control

If an artificial control is considered for the station, the foregoing analysis is usually inapplicable. The stability of the discharge vs. stage relation with an artificial control is such that the accuracy of the computation usually becomes relatively fixed, and often is no longer a function of the frequency of measurement. The accuracy of computation for the artificially controlled condition is determined by past experience with a similar control, by a hydraulic modeling procedure, or by analysis of the record obtained after the control is built. In this case the decision on the frequency of measurement becomes a subjective one made by the manager of the station, and an optimal frequency cannot be determined. The cost equations 1 and 2 are still applicable, however the decision variable d assumes the value of 1.0 when an artificial control is analyzed, and the variable m is based on the decision of the manager. A cost estimate of the construction of the control is obtained to evaluate $C_c$.

Even with an artificial control, the accuracy of the record is occasionally a function of the frequency of discharge measurement. This phenomenon is caused by scour and fill conditions alternately taking place in the approach to the control. In this case the error analysis would be valid, and the costs would be determined from equations 1 and 2 with the decision variable d equal to 1.0.

# CHAPTER 3

# THE WORTH OF DATA

The worth of a gaging station record can be equated with the value of benefits foregone if the record were not available; therefore the worth of data is determined by its ultimate uses. Because all of these uses cannot be foreseen, a precise determination of the total monetary value cannot be made. Methods are currently being developed, however, that ascertain that portion of the total value that can be attributed to a specific anticipated use. The scheme of this paper is illustrated by using the worth valuation procedure of Dawdy et. al. (5); however, any other procedure or combination of procedures that describe the worth of a particular record can be substituted into this portion of the scheme.

The valuation procedure of Dawdy et. al. was contrived to estimate the worth of a record that is used primarily for the design of a water conservation reservoir. A long sequence - 500 years - of data is simulated from an estimate of the statistics of the flow at the station. This synthetic record is used to determine the relation of gross benefits from the project to the size of the dam. The relation of cost of constructing the dam and its associated facilities to the size of the dam is determined by making cost estimates for several sizes of reservoir. The optimum size of reservoir is determined from the 500-year record and the cost and benefit curves.

The 500-year synthetic record is partitioned into a set of fifty 10-year records. In other words, the first 10 years of the 500-year record are designated a record; the second 10 years, another record; etc. The 500-year record is then repartitioned into sets of twenty 25-year records, then 50-year records, and five 100-year records. A determination of the best reservoir size is made for the individual records in each set. For each size determination values of cost and benefit are read from the relations determined for the 500-year record. The differences between the benefits and costs for the individual sizes of reservoir are averaged within each set. The averages are the expected net benefits derived from the project if its design is based on a gaging station record of a length equal to the length of the records in the corresponding sets.

As the record length is increased the value of the expected net benefit increases. This increased benefit from a record of one length to a record of a greater length is the benefit foregone by not extending the record to the greater length. The value of the benefit foregone is the total worth of the data that would have been collected in the extension of the record. By using this procedure it is possible to determine the value of a record after specific time periods.

For the purpose of interpolating values of worth for lengths of record that were not analyzed in the above procedure, it is assumed that the worth of the record is a function of its information content about the pertinent statistics of the population of which it is a sample. Information content is defined by Fisher (9) as the inverse of the variance of a series of determinations. According to this definition, the information content of a record is the reciprocal of the variance of the sampling distribution of the population statistic being considered.

In referring to the selection of the proper size of a storage reservoir, Thomas and Burden (15) state "... the variance is by far the most important parameter even for relatively short record lengths." This statement indicates that the worth of a record used to design a storage reservoir is a function of the reciprocal of the variance of the sampling distribution of the variance of the record. Kendall (12) has defined the variance of the sampling distribution of the variance as

$$S_V^2 = 2V^2/N \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

in which V = the true variance of the streamflow population. Equation 6 is applicable only when there are no errors of measurement in the individual elements of the samples. Because the individual discharge determinations are not exact, the estimates of $S_V^2$ must be increased by substituting $V + S_A^2$ for V.

Equation 6 becomes

$$S_V^2 = 2(V + S_A^2)^2/N \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (6a)$$

and the equation for information content becomes

$$I = N/2(V + S_A^2)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (7)$$

in which I = the information content about the variance. If the worth of a record is a function of its information content, it can be stated - in the case of reservoir design -

$$W = f\ [N/(V + S_A^2)^2] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (8)$$

in which W = the worth of the record and f = an undefined function. The parameter V is estimated either from an existing record or a regional regression analysis, and the value of $S_A$ is obtained from the error analysis that was performed in the cost valuation. Values of W and N from the procedure of Dawdy et. al. are used in conjunction with the appropriate values of V and $S_A$ to empirically define the function f. Families of curves of W vs. N and $S_A$ are constructed to graphically illustrate this function.

The marginal or annual worth of a particular year of data is

$$M_j = W_j - W_{j-1} \qquad \text{.................................(9)}$$

in which $M_j$ = the marginal worth of the data collected in the year j and $W_j$ = the worth of a record of length equal to j years. When the values of N and $S_A$ are fixed, equations 8 and 9 yield values that are complementary with the cost figures obtained from equations 1 and 2.

# CHAPTER 4

# THE PRESENT WORTH OF NET BENEFIT

Up to this point the search for the optimum operating procedure has not accounted for the distribution of costs and benefits in time. This section deals with this factor because of the apparent diminution of the magnitude of future costs and benefits that is caused by man's preference for receiving his benefits now and because of his uncertainty of the future.

The net benefit derived from the operation of a streamgage is the excess worth of the data after the costs of operation are satisfied and can be stated mathematically

$$B_N = W - C_t \quad \text{.................................} (10)$$

in which $B_N$ = the net benefits derived from a streamgaging station. In practically all cases a net benefit must be anticipated for a project to be economically feasible.

Maximization of net benefit is the criterion of economic optimality if the stream of benefits and costs are physically restricted to constant annual rates. This criterion of optimality applies only to this specific type of series of costs and benefits, however.

As a general rule, the desirability of a time series of economic rewards and penalties can be measured by the present worth of the series. Grant (10) defined the present worth of a series of future incomes as "......the present investment which the future incomes would just repay with interest." A similar definition of the present worth of a series of future expenditures is: the present investment which would just liquidate the future expenditures by disbursement of the principal and interest. Grant gave an equation for the present worth of either a benefit or cost incurred in a future year. This equation is

$$P_W = P/(1 + R)^N \qquad (11)$$

in which $P_W$ = the present value of a future cost or benefit; P = the monetary value of the cost or benefit; R = the annual interest rate expressed as a fraction; and N = the distance in the future that P is incurred, in years. The term $(1 + R)^N$ is defined as the discount factor.

Economic optimality is achieved for any project by maximizing the present worth of the series of costs and benefits attributed to the project. The present worth of the series is equal to the present worth of the net benefit.

Present worth determinations are very sensitive to two factors, N and R. In this scheme N is one of the decision variables, and its value is determined by the analysis. The interest rate (R), on the other hand, is a constant in the analysis. The magnitude of R is dependent upon the interest rate at which the backers of the project can borrow money. Because most major water resources projects in the United States are financed by the federal government, the value of R used in this scheme will be the same as that used by government agencies in determining the present worth of their projects. Halmos (11) reported that this discount rate was to be raised to 4.625% in the year 1968; therefore, $(1.04625)^N$ is used as the discount factor in this study.

# CHAPTER 5

# THE OPTIMUM OPERATING PROCEDURE

The optimum operating procedure is defined as that frequency of measurement of discharge and that length of record that yields the maximum value of present worth of net benefit derived from the operation of a streamgage. The ease with which the determination and maximization of present worth of data is accomplished depends upon the type of use that is to be made of the data. For the purpose of a present worth analysis gaging stations can be categorized into three classes: water-management stations, project-design stations, and long-term stations. These three classes are discussed individually in the remainder of this section.

## Water-Management Stations

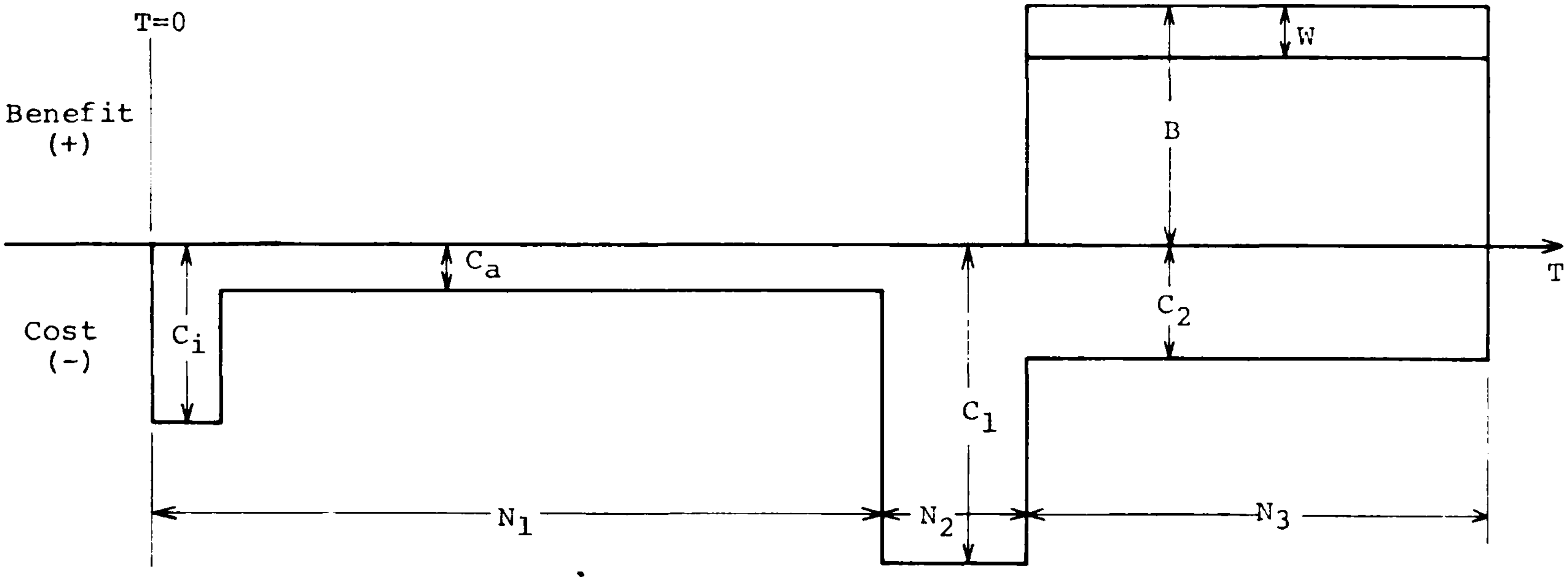
Water-management stations are used primarily in the operation and accounting procedures of an active water resources project. Benefits derived from this type of station are obtained in the same time period in which the costs are incurred, and the magnitudes of these costs and benefits tend to be constant for each year of operation. Because the stream of benefits and costs has an essentially constant annual rate, maximization of net benefits from this type of station is sufficient to insure that present worth is also a maximum.

Economic theory states that the net benefit is a maximum at the point where marginal cost equals marginal worth. In economic marginal analyses, the word marginal refers to the effects of the subsequent addition of a single unit of input to a process. In the previous references to marginal quantities, the unit of input has been an additional year of data. Marginal-time analysis is not useful, however, in the analysis of water-management stations, because the annual costs and benefits are constant. In this case the marginal input quantity is an additional discharge measurement within each year of record. The worth added to the data by this measurement is the marginal worth, and the marginal cost is the cost of the discharge measurement. The number of measurements per year at which the marginal worth equals $c_m$ defines the point at which maximum net benefit is obtained and, therefore, is the optimum frequency of measurement. The optimum length of record is set equal to the life of the project, because benefits and costs continue at an essentially constant rate as long as the project and the gage are in operation.

The use of this scheme for water-management stations is at present not possible because of the lack of definition of the worth of data for this type of station. Undoubtedly methods will be contrived that will define this quantity. As these methods become available they can be inserted into the scheme, and operation procedures for water-management stations can be optimized.

## Project-Design Stations

Project-design stations are operated primarily to obtain data upon which a water resources project is to be designed. Unlike the water-management station, the benefits derived from this class of station are not received at the same time that the costs of operation are encountered. The stream of benefits and costs of a conservation project, including the operation of a gage and the construction of a dam, are shown in figure 1. It can be seen that all benefits derived from the record are accrued subsequent to the operation of the gaging station. Because of this time lag in the reception of the benefits, an increase in the length of the streamgaging record increases the exponent of the discount factors by a magnitude equal to the increased length of the record. This exponential increase is applicable to the discount factor for each year in which benefits are accrued, and its effect on the present worth



B = Average annual benefit of the project
$C_1$ = Average annual cost of construction of the project
$C_2$ = Average annual operating cost
$C_i$ = Cost of constructing gage
$C_a$ = Annual cost of obtaining record

$N_1$ = Length of record
$N_2$ = Length of construction period
$N_3$ = Project economic life
T = Time
W = Average annual worth of the data

Figure 1.--Benefits and Costs of a Contemplated Reservoir Project.

of the stream of benefits becomes very major after only a few years of operation. Figure 1 can be reduced to an equivalent value of present worth by the summation of repetitive applications of equation 11 to each annual value of cost and benefit.

The decision to increase by one year the length of time that the gaging station is operated has multiple effects on the present worth of the project. The first consequence is an increase in the cost of a magnitude equal to the annual cost of operating the gage discounted from the end of the contemplated record back to the present. The second consequence is an increase in the average net benefits obtained from the project. Annual net benefits, though actually increased in magnitude, are received a year later, and this added year causes the discount factor of each annual net benefit to multiply by $1 + R$. The increase in the discount factor is the third effect of extending the record length. Of these three consequences only the second has a positive effect on the present worth; therefore, if an increase in record length is to be a favorable action, the increase in average net benefits must not only be greater than the cost of increasing the record length, but it must also offset the increased discount factors.

Because the worth of the data obtained at the gaging station is an integral part of the project, the present worth of the data cannot be determined without evaluating the present worth of the entire project. The worth of the data is the residual net benefit above that obtained from a project designed with no data; therefore, maximization of the present worth of the project automatically assures that the present worth of the data is at its greatest value. Maximization of the present worth of the project is accomplished by a series of computations of present worth for a given accuracy of computation. The series begins with one year of record, and each subsequent determination is made with an additional year of data. The series is continued to the point where the present worth begins to decrease. A series of present worth values is obtained for each of several values of standard error of computation, $S_A$ , and the maximum of each series is plotted against its associated standard error. The peak of this graphical relation yields the optimum value of the standard error of computation.

With the value of the optimal $S_A$ and the relation of m to $S_A$ that was defined in the section on the cost of data, the optimum frequency of measurement is determined. The optimum length of record is the value of N at the maximum of the present worth series of the optimum standard error of computation.

The preceding analysis is for the case in which the project is to be built as soon as sufficient data are available for design. If construction of the project is planned for some fixed date which is further in the future than the optimum length of record, a decision must be made on whether or not to gage the stream for the entire period from the present to the initiation of the project. This period is designated the waiting period. The decision criterion is still the maximization of the present worth of the net benefits. Present worth will not be as great for similar projects in this case as it was in the previous one because the benefits are not obtained throughout the optimal time of accrual.

Because the annual cost of operating the gaging station is constant, the year whose present value of cost is the greatest is the earliest one encountered; therefore, if it is not profitable to gage during the whole waiting period, the period of non-gaging precedes the period of gaging. In other words, the station should not be established until the latest possible time to obtain the optimal benefits from the record. A marginal analysis is performed with one year of data as the marginal input. The marginal input, however, is added just prior to the planned initiation of the record instead of after it. If gaging through the entire waiting period is a desirable action, the present worth of the data obtained in the first year must be greater than its cost. This decision statement can be expressed mathematically

$$P_{w,j} - P_{w,j-1} > mC_m + C_r \quad \text{.......................} \quad (12)$$

in which $P_{w,j}$ = the present worth of the benefits from a record j years in length and j = the length of the waiting period in years.

A graph with families of curves of $(P_{w,j} - P_{w,j-1})$ vs. N and $(mC_m + C_r)$ vs. N for several values of standard error of computation is used to define the optimum operating procedure. If the waiting period is smaller than the value of N at the intersection of a cost and present worth curve of equal $S_A$, equation 12 is satisfied, and the gage is operated through the entire waiting period. If j is greater than N at the intersection, the gage is operated only for the N years immediately prior to the initiation of the project.

The optimum frequency of measurement is determined by evaluating the present worths of records of several values of $S_A$ in order to define graphically the relation of present worth to $S_A$. The value of $S_A$ at the peak of this relation is used to determine the optimum frequency of measurement from the relation of m to $S_A$.

## Long-Term Stations

Stations of this class have three primary uses: (1) a basis for the extension of short-term records through correlation analysis, (2) the determination of long-range trends in the hydrology of an area, and (3) the definition of regional relations of streamflow characteristics. Many stations that are classified as long-term may also be used to design projects or used as water-management stations in measuring inflow to projects.

The intent of the long-term station is to operate continuously without a fixed termination date; therefore, no optimum length of record can be specified for this type of station. Because of the multiplicity of uses of its data, determination of the optimum frequency of measurement for a station in this category will be very difficult. This difficulty can be primarily attributed to the definition of the worth of data relation. The worth relation will be a function of the potential of the water resources in the area that the station represents for ultimate development and of the density of long-term stations in that area. At present this relation has not been defined, and the scheme cannot be applied to a station in this category.

# CHAPTER 6

# APPLICATION OF THE SCHEME

The scheme as presented requires certain input that must be secured before the analysis can be initiated. This input consists of estimates of the streamflow parameters and field measurements of discharge and stage. The availability of this information is dependent upon whether or not a station has existed in the past at the site in question; therefore, discussion of the application of the scheme is separated into two cases.

## Case I - Analysis of a Proposed Gage

In this case where no field data exist, the analysis must be delayed until field measurements of discharge and stage are obtained because methods are not available currently that estimate errors caused by shifting stage-vs-discharge relations. Two options are available to obtain this data: either the planned station can be constructed and full-scale operation begun, or a temporary staff gage can be installed to obtain stage measurements for the analysis. If the gage is to be used for water management or the design of a project that is to be constructed immediately after design data are obtained, the first option is the obvious choice. However, if the gage is to be a project-design station for a project that is planned for a future date, the present value of the cost of constructing the gage may indicate the installation of a temporary staff gage.

With either of the above options the measurements of discharge and stage should be obtained in as brief a time span as is practicable. In most instances, one year of concentrated effort should yield enough discharge measurements that include a sufficiently large range of discharges to make the error analysis feasible.

Estimates of streamflow characteristics for a nonexistent station are obtained from regional regression models such as that used for flood discharges by Benson (2). In the situation where regional models do not exist a regionalization study is a preliminary step to the scheme. These regional estimates are used in a generating model such as that of the Army Corps of Engineers (17), Fiering (8), or Young and Pisano (16) to obtain a 500-year synthetic discharge record that is used in the worth of data analysis.

These input data are inserted into the scheme, and the optimum operating procedure is determined. The measurement frequency is changed from the concentrated effort needed to obtain data for the analysis to that of the optimum operating procedure. If, after operating the gage for a few years at the indicated optimum frequency of discharge measurement, information indicates that the original data or estimates were in error, a new analysis is performed using the new information, and a revised optimum operating procedure is established.

### Case II - Analysis of an Existing Gage

In this case measured data are usually available for both the error analysis and the generation of the synthetic record. If the existing data are not sufficient, however, they are augmented by increased discharge measurements and/or regional analyses. Consideration of the use of regional estimates of streamflow characteristics at a station with an existing record creates a need for criteria for the decision of whether or not the regional estimates are better than the estimates obtained from the record. These criteria are set forth by Matalas and Gilroy (14).

The input data are processed through the scheme, and the optimum operating procedure is defined. Because of the availability of more information at the time of the analysis, only one determination of the optimum procedure should be required for this case.

# CHAPTER 7

## EXAMPLE - ARROYO SECO NEAR SOLEDAD, CALIFORNIA

An analysis of the existing gaging station on Arroyo Seco near Soledad, California is presented in this section as an illustration of the scheme. This station was chosen as an example primarily because worth of data values and discharge and stage measurements were readily available. The worth of data values were derived by Dawdy et. al. (5) with the assumption that the data was to be used only for the design of a conservation reservoir that was to yield a discharge equal to 50% of the average discharge of the stream. Because the intent of the analysis is only illustrative, the scheme is executed both for the case in which the project is planned for a future date and for the case in which the project is to be initiated immediately after data are available.

Arroyo Seco, a tributary of the Salinas River, is an intermittent stream that drains an area of 244 square miles on the east side of the coast range in central California. The gaging station is located at latitude 36°16'50" north and longitude 121°19'20", 1.5 miles downstream from the mouth of Vaqueros Creek, and 10 miles south of Soledad. U.S. Geological Survey records (18) state that the average discharge based on 65 years of record is 159 cubic feet per second (cfs), and the median of yearly mean discharges is 122 cfs. The maximum flood discharge observed during the period of record, 1901-66, was 28,300 cfs. Records for this station are considered to be of good accuracy by the Geological Survey, and there has been no regulation or large diversion from the stream above the station during the period of record. Figure 2 is a daily flow-duration curve for this station.

## Error Analysis

In order to obtain a sufficiently large sample of discharge measurements, the error analysis was performed on the 10-year period of record beginning October 1, 1950 and ending September 30, 1960. During this period 141 discharge measurements were made, and the frequency was 14.1 measurements per year. The range in discharge for these observations was from 0.97 cfs to 4,550 cfs. The average number of verticals per discharge measurement was 25, and the 0.2-.8 depth rule and the 0.6 depth rule were used to determine mean velocity with approximately equal frequency. According to figure 1 of Carter and Anderson (4), the standard error of measurement was 3%. This value was converted to its decimal equivalent and squared to obtain the value of $S_m^2$.

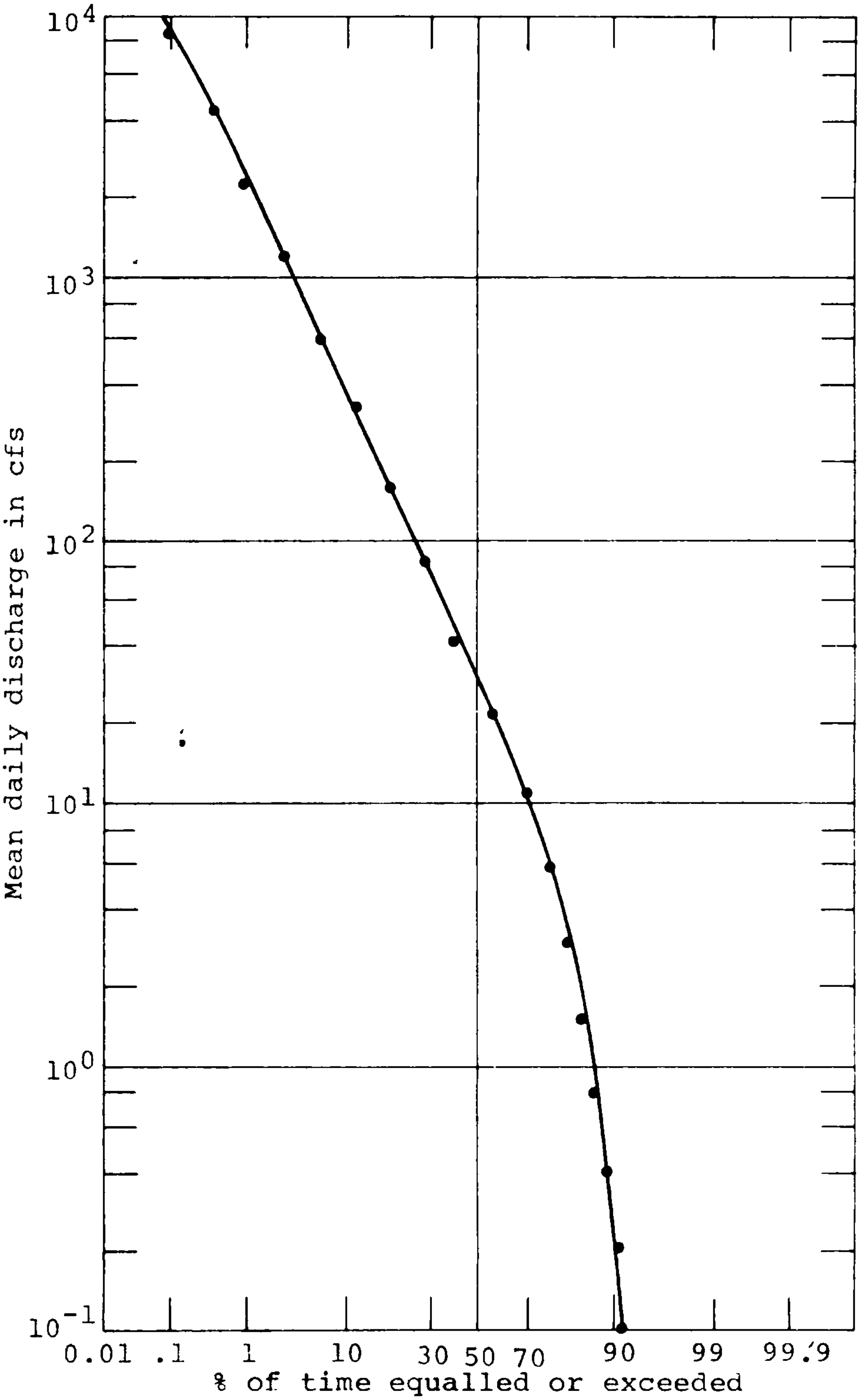


Figure 2.--Daily Flow-duration Curve for Arroyo Seco near Soledad, California.

Three sets of observations were chosen as the analysis groups for three individual determinations of $S_c$. The three sets represented frequencies of 3.5, 7.0, and 10.5 measurements per year. In order that the values of $S_c$ in the latter determinations were not affected by information gained in the first determination, the first set was made a subset of the second set, and the second set was made a subset of the third. Because preliminary analysis indicated that homoscedasticity was not present over the range of discharges experienced, each of the control groups was divided into four sub-groups on the basis of the magnitude of the measured discharge. Table 1 shows the limits of these sub-groups and the variances of the residuals within each sub-group. The residuals were expressed as ratios to the estimated mean annual discharge. The variances were pooled within each set using equation 4, and values of $S_c$ and $S_A$ were computed for each set. Table 2 and figures 3 and 4 illustrate the relations of each of these variables to the average period between discharge measurements.

The data point at a value of p equal to zero in figure 3 was obtained by equating the standard error of computation with the standard error of measurement. This equality is valid because the equation, $p = 0$, indicates continuous measurement of discharge, and the accuracy of

Table 1.--Sub-group Data.

| | Sub-group | Limits (cfs) | $P_i$ | $S_i^2 \times 10^3$ | $P_i\ S_i^2 \times 10^3$ |
|---|---|---|---|---|---|
| Set 1 | 1 | 0-15 | 0.30 | 8.63 | 2.59 |
| | 2 | 15-75 | .29 | 1.05 | .30 |
| | 3 | 75-200 | .16 | 1.82 | .29 |
| | 4 | >200 | .17<br>$\Sigma$=0.92 | 106. | 18.0<br>$\Sigma$=21.2 |
| Set 2 | 1 | 0-15 | .30 | 1.34 | .40 |
| | 2 | 15-75 | .29 | .72 | .21 |
| | 3 | 75-200 | .16 | 2.18 | .35 |
| | 4 | >200 | .17<br>$\Sigma$=0.92 | 96.8 | 16.4<br>$\Sigma$=17.4 |
| Set 3 | 1 | 0-15 | .30 | 1.05 | .32 |
| | 2 | 15-75 | .29 | .94 | .27 |
| | 3 | 75-200 | .16 | .56 | .09 |
| | 4 | >200 | .17<br>$\Sigma$=0.92 | 87.9 | 14.9<br>$\Sigma$=15.6 |

Table 2.--Standard Errors of Computation.

| Set | m (meas./yr.) | P (days) | $P_i\ S_i^2 \times 10^3$ | $V_R \times 10^3$ | $S_m^2 \times 10^3$ | $S_C$ (%) | $S_A$ (%) |
|---|---|---|---|---|---|---|---|
| 1 | 3.5 | 104 | 21.2 | 23.0 | 0.9 | 14.9 | 8.0 |
| 2 | 7.0 | 52 | 17.4 | 18.9 | .9 | 13.4 | 5.1 |
| 3 | 10.5 | 35 | 15.6 | 17.0 | .9 | 12.7 | 3.9 |

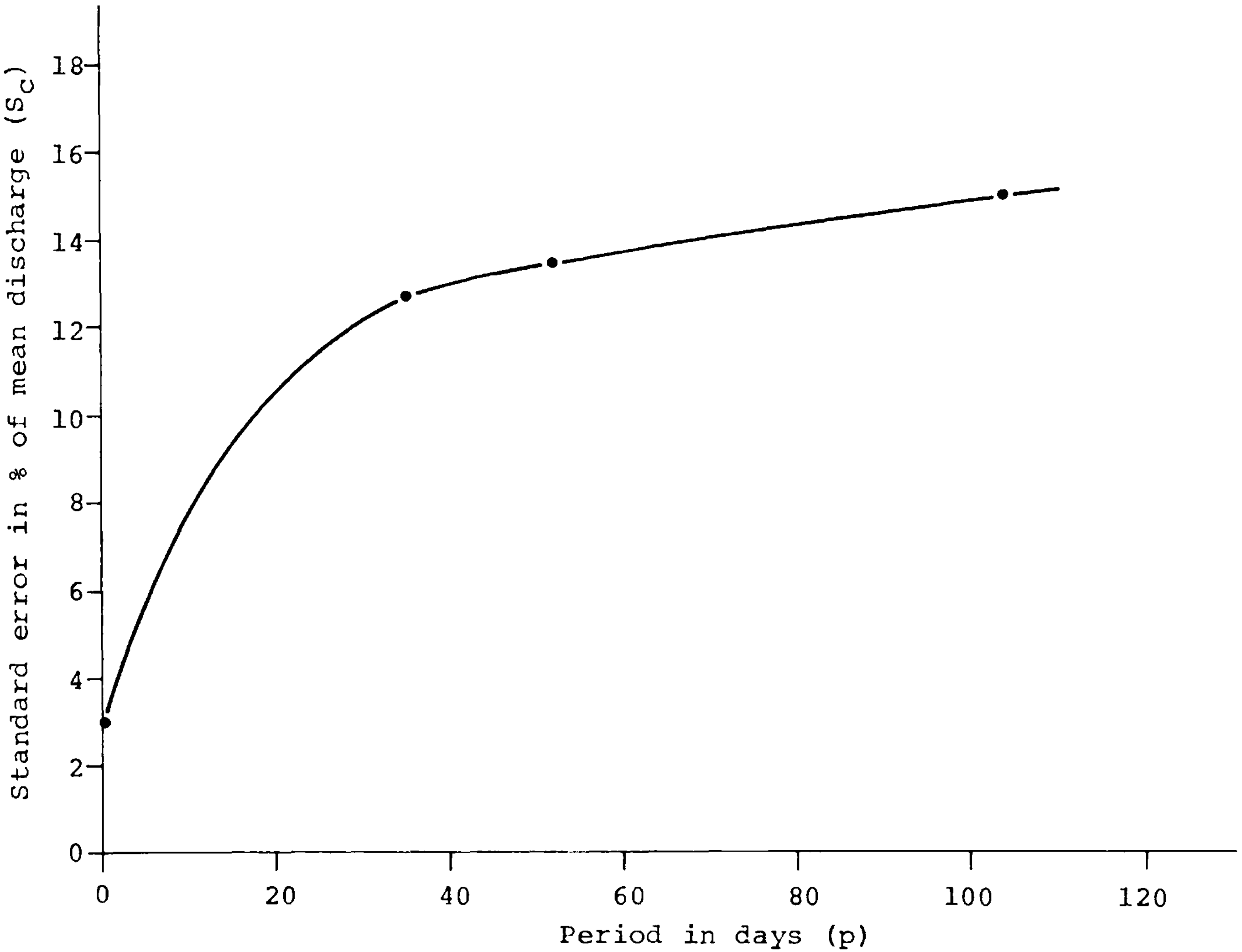


Figure 3.--The Relation of Error in Computing Instantaneous Discharge to the Average Period between Discharge Measurements.

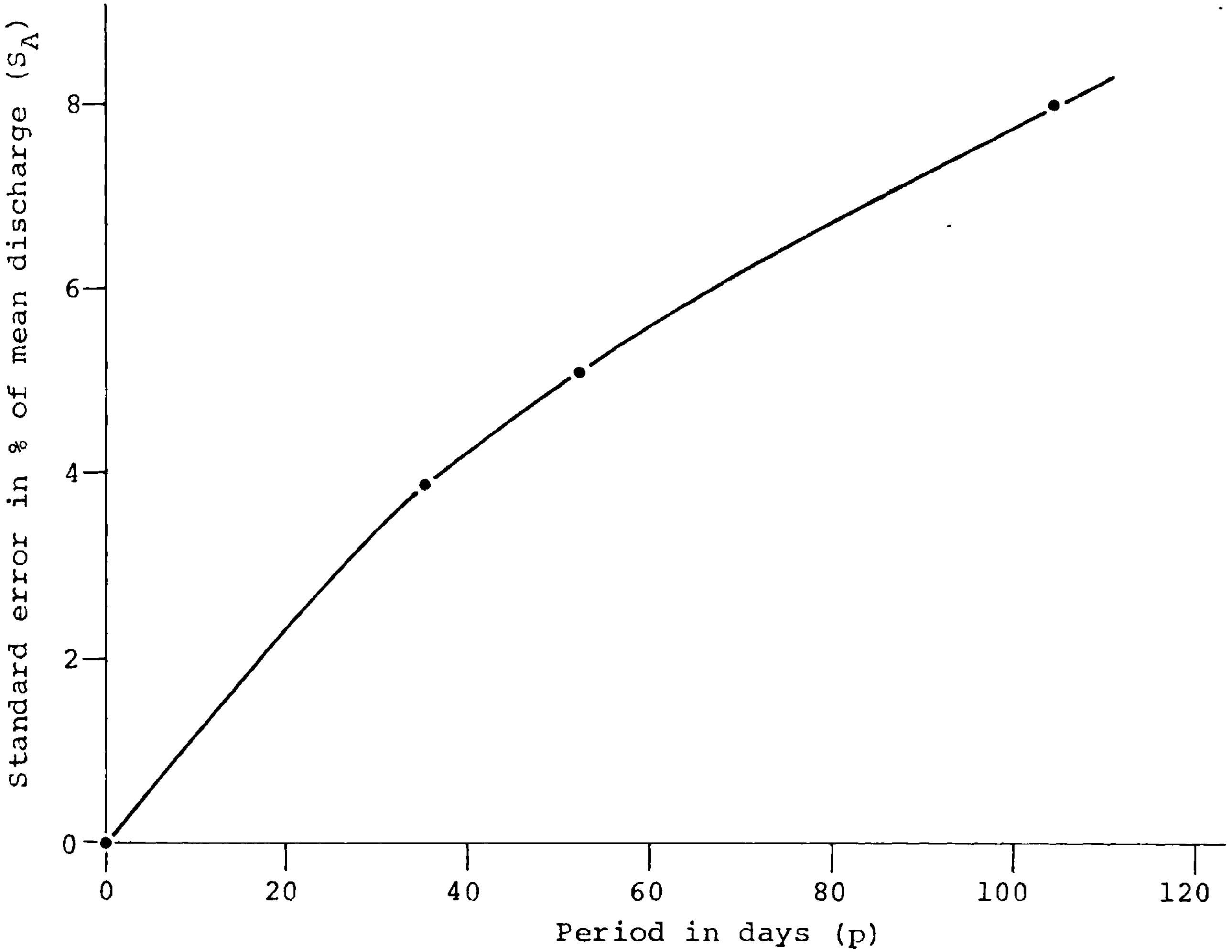


Figure 4.--The Relation of Error in Computing Annual Mean Discharge to the Average Period between Discharge Measurements.

determining the magnitude of discharge at any instant would be equal to the accuracy of the measurement made at that instant. The data point at the origin of figure 4 was obtained by solving equation 5a for a value of p equal to zero.

## Annual Cost of Data

The relation of the annual cost of data to the standard error of computation of the mean annual discharge was graphically derived by substituting m values extracted for a given level of accuracy from figure 4 into equation 2. The parameters $C_m$ and $C_r$ were estimated to be $50 and $1000, respectively. This relation is illustrated in figure 5.

## Worth of Data

Values of average net benefit that would be derived from conservation reservoirs constructed on Arroyo Seco whose designs are based on varying lengths of streamflow record were taken from Dawdy et. al. (5), and are presented in table 3. These data are values of present worth of the reservoir project at the point in time of its initiation. The net benefits were related to the worth of the streamflow records by subtracting from each value of net benefit a constant that was equal to the net benefit that could be obtained from the project if its design were

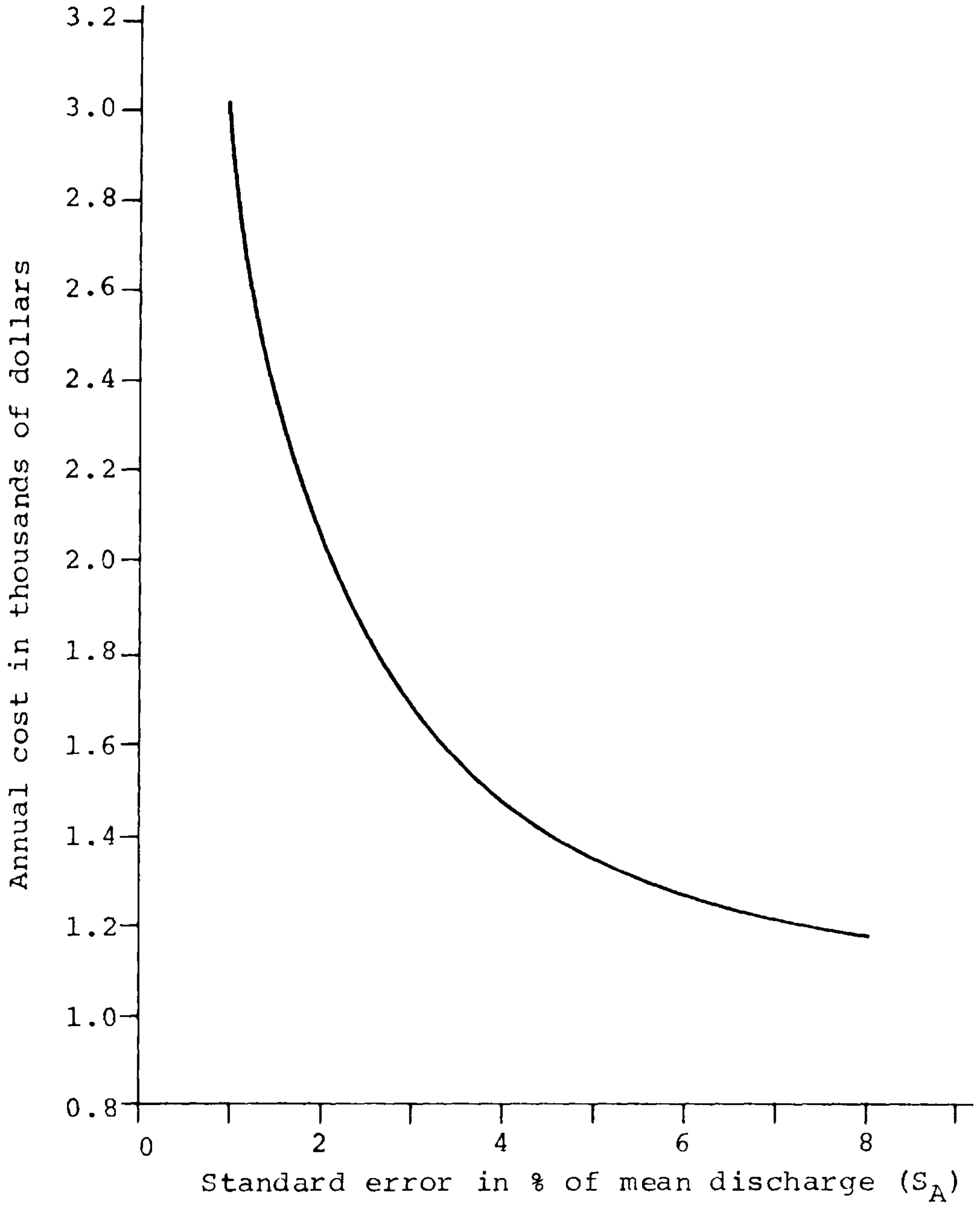


Figure 5.--Annual Cost of Obtaining Data vs. $S_A$.

Table 3.--Average Net Benefit from a Reservoir Project.

| Length of Record (N) | 10 | 25 | 50 | 100 | 500 |
|---|---|---|---|---|---|
| Net Benefit[a] | 15.1 | 17.9 | 18.6 | 19.5 | 20.5 |

[a] Net benefit is expressed in millions of dollars as the present worth of the project at its beginning.

based on no record at the gaging site. The magnitude of this constant was of no consequence in determining the optimum operating procedure, and it was assumed equal to zero.

The data in table 3 were derived from a constant value of $V + S_A^2$. This constant was estimated by equating it with the variance of the annual mean discharges of the period of record expressed as dimensionless ratios of the mean discharge. The magnitude of this constant was found to be 0.570. By the use of equation 7 the information content of a streamflow record concerning the variance of annual discharges at Arroyo Seco near Soledad was determined to be N/0.65 for the accuracy of computation that was used during the period of record.

Because N is the only parameter of information content that varies under the conditions analysed by Dawdy et. al., the function f of equation 8 was determined by empirically fitting a curve to the data of table 3. The data approximated a hyperbolic relation, and the best fit was found to be

$$P_{wi} = (20 \times 10^6) - (50 \times 10^6/N) \text{ ............ } (8a)$$

in which $P_{wi}$ = the present worth in dollars of the data at the initiation of the project. This relation is illustrated graphically in figure 6. Substitution of equation 8a into equation 8 yielded the identity

$$c2\ (V + S_A^2)^2 = 50 \times 10^6 \text{ ..................... } (13)$$

in which c = the coefficient of the reciprocal of information content that is implied by f. The coefficient c was evaluated by simplification of equation 13, and was found to equal 77 x $10^6$. The relation of present worth to information content became

$$P_{wi} = (20 \times 10^6) - 154 \times 10^6\ (V + S_A^2)^2/N \text{ ... } (8b).$$

In equation 8b, $S_A$ and N are the only independent variables. The parameter V is a constant for a particular station. For Arroyo Seco it was assumed to equal the difference between the estimate of $V + S_A^2$ and the value of $S_A^2$ determined for the existing record for the period 1950-60. During this period the frequency of measurement was 14.1 measurements per year, which is the equivalent of 26 days between measurements. $S_A$ was found to equal 3% according to figure 4, and V was set equal to 0.569.

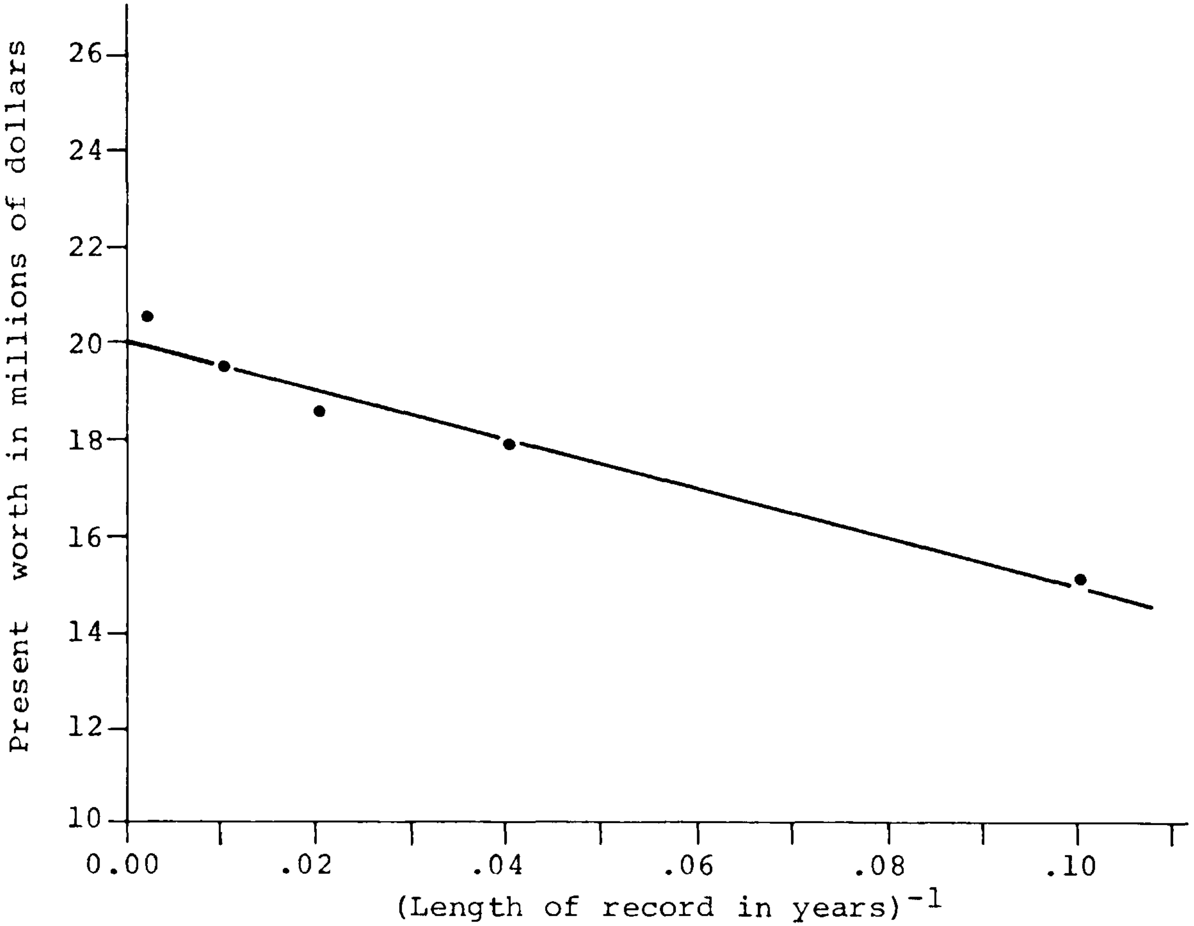


Figure 6.--Present Worth of Data at the Initiation of a 50% Yield Project.

The range of definition of $S_A$ was found in the section on error analysis to be from 0 to 8%. Within this range the term $(V + S_A^2)^2$ of equation 8b varies by less than 2% of its median value. Because this variation was considered insignificant when compared with the accuracy of determining V, the term was considered to be a constant. Thus, equation 8a was accepted as defining the present value of the record at the beginning of the project.

The present value of the marginal worth of an added year of record was found by taking the derivative of equation 8a with respect to N, and it was expressed

$$d\,P_{wi}/dN = 50 \times 10^6/N^2 \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\;(14).$$

This equation was converted to one that defines the present value at the beginning of the record by dividing equation 14 by the discount factor $(1.04625)^N$. The result of this conversion was

$$P_w = 50 \times 10^6/N^2\,(1.04625)^N \;\ldots\ldots\ldots\ldots\ldots\ldots\;(15).$$

Equation 15 is illustrated graphically in figure 7.

## Optimum Operating Procedure

Because of the insensitivity of the relation of worth of data to the accuracy of computation for Arroyo Seco near Soledad, the optimal frequency of measurement was not definable. This lack of sensitivity combined

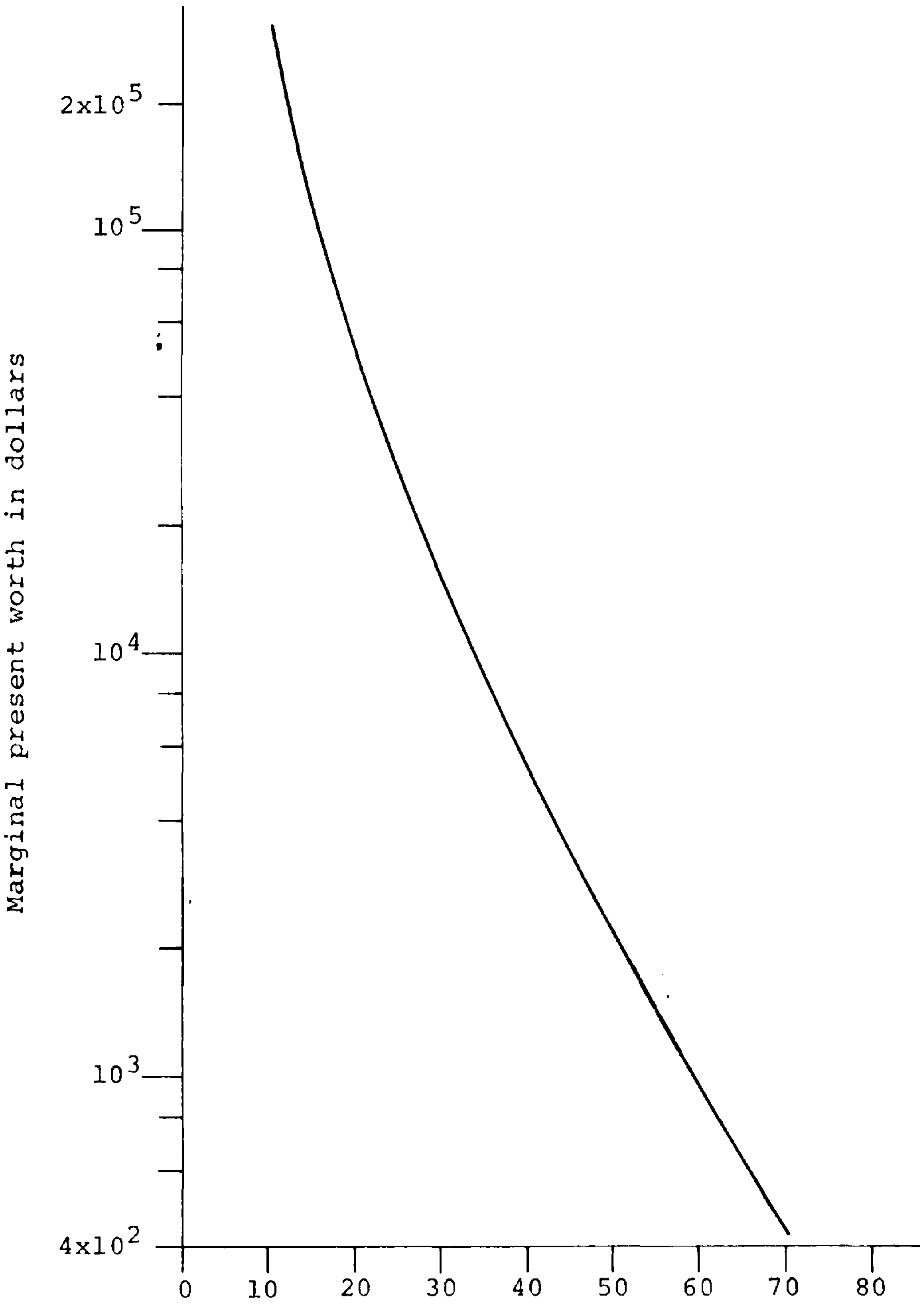


Figure 7.--Marginal Present Worth of Data at the Beginning of the Record.

with the fact that cost increases with accuracy did suggest, however, that a low frequency of measurement is preferable. In order to complete the illustration, a standard error of computation of annual mean discharge of 7% was arbitrarily chosen as the optimum accuracy. This standard error corresponded to a frequency of 4.3 measurements per year.

The optimum length of record for the design of a project that is planned for a specific future date was determined by equating the annual cost and the present value of the marginal worth of an additional year of data. This equation was stated mathematically by combining equations 2 and 15 thus

$$m\ C_m + C_r = 50 \times 10^6/N^2\ (1.04625)^N \ \ldots\ldots\ldots \ (16)$$

Equation 16 was solved graphically. A plot of marginal present value vs. length of record similar to figure 7 was constructed for the right side of the equation. The remainder of the equation was superimposed on the graph as a horizontal line whose ordinate value was determined by substituting real values for m, $C_m$, and $C_r$ in the left side of the equation. This graph is shown as figure 8. The value of N, 56.3 years, at the intersection of the two curves defined the maximum length of time that the gage would be operated. If the waiting period as defined in

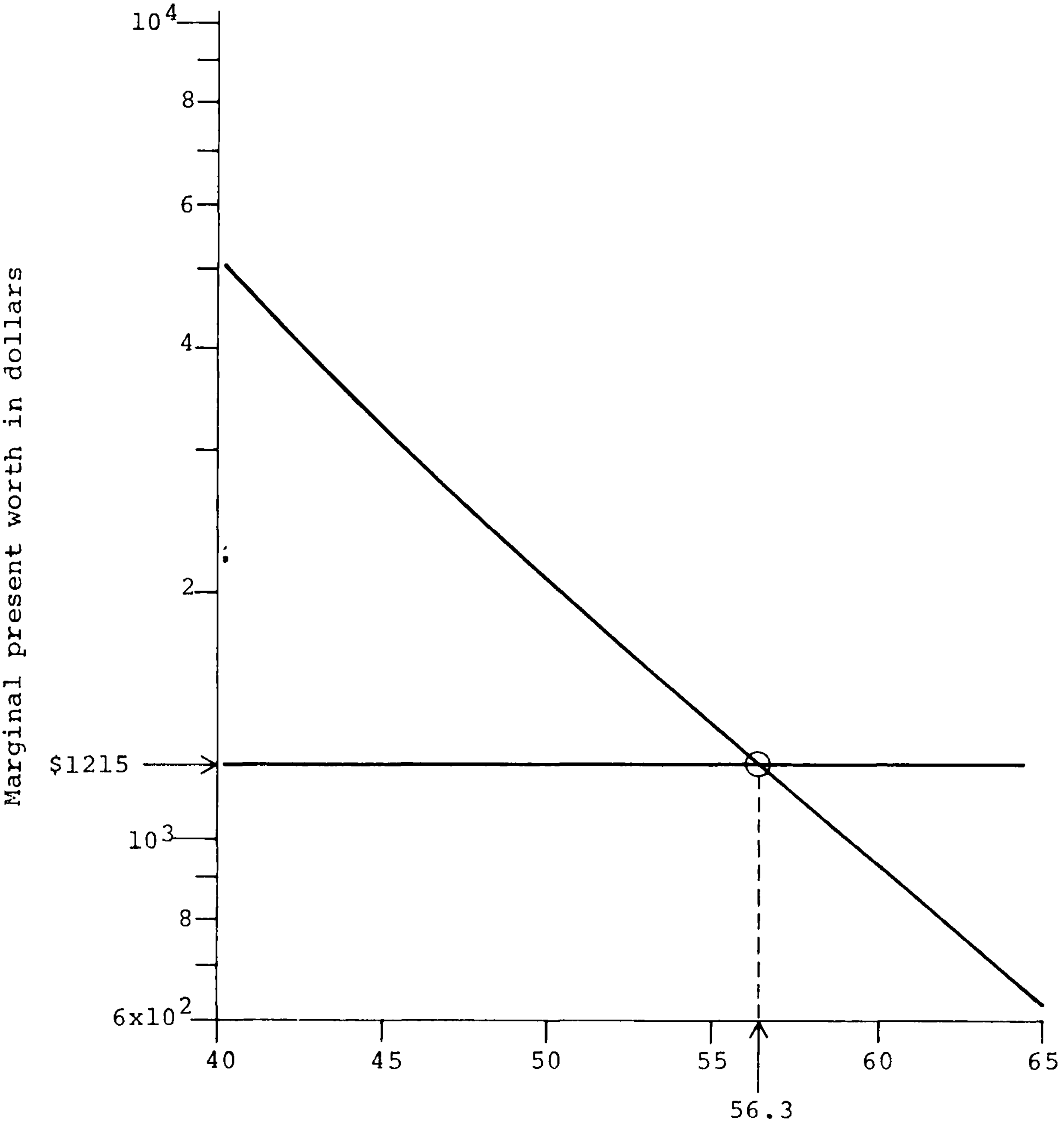


Figure 8.--Solution of Equation 16.

Chapter 5 is longer than 56 years, the gage is operated only during the last 56 years of the waiting period. If, however, the waiting period is shorter than 56 years, the gage will be operated as a project-design station only during the waiting period.

The optimum length of record, for the case in which the project is to be built immediately after design data are available, was found to be much shorter than 56 years. Under this condition, the optimum length of record was defined as the value of N at the maximum present worth of the project. A series of determinations of present worth was computed using equations 2, 8a, and 11 for values of N from 1 to 60 and for a value of $S_A$ equal to 7%. The results of these computations are shown in figure 9. The peak of this series occurs at N equal to nine years. The associated value of present worth, 9.6 million dollars, was the maximum value of present worth that could be obtained by any plan of development of Arroyo Seco at a yield of 50% of mean discharge.

The optimum value of nine years represented an extrapolation of one year below the lowest data point used to derive the worth function. Because of the relative flatness of the relation in the vicinity of the peak, shown in figure 9, this short extrapolation was not considered to be risky.

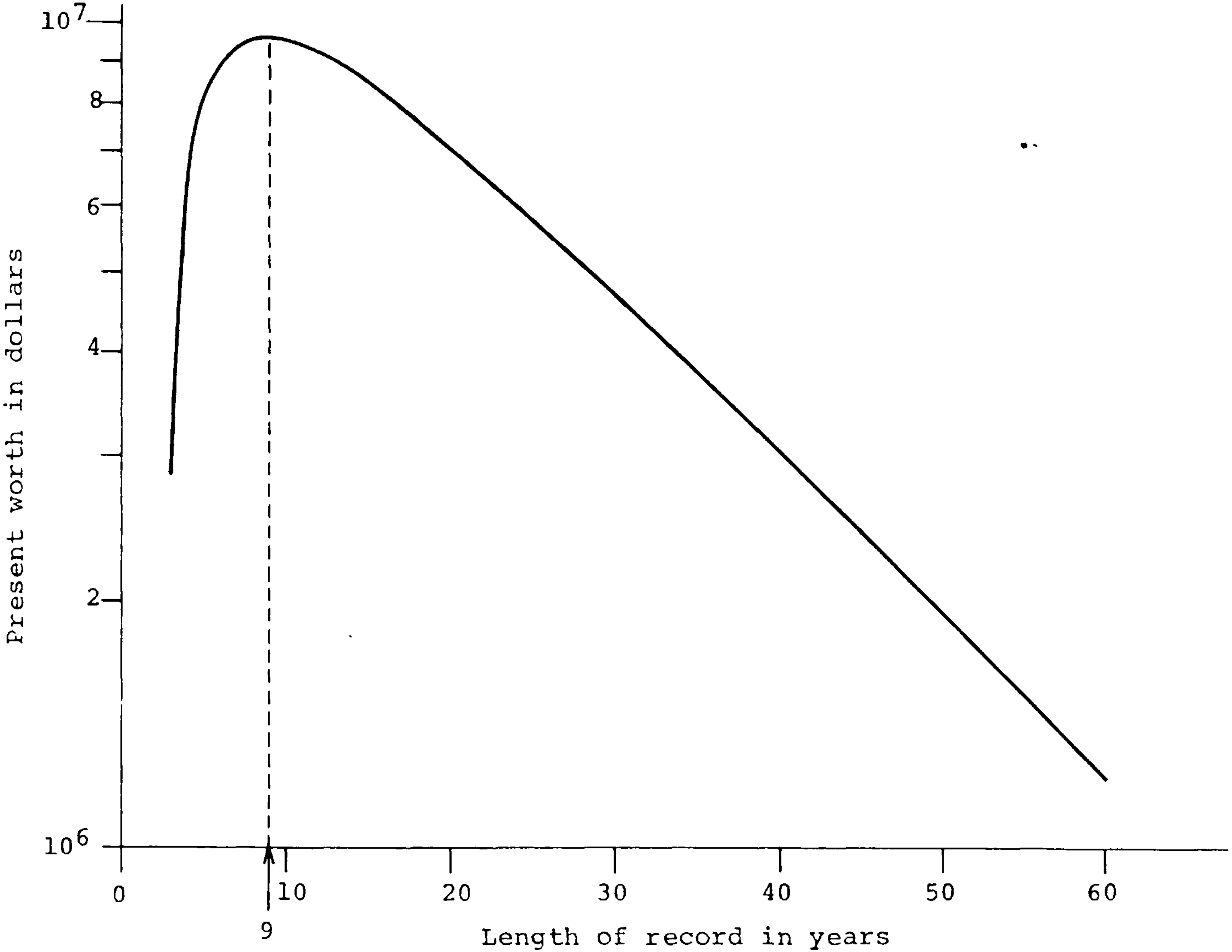


Figure 9.--Present Worth of Data at the Beginning of the Record.

The preceding analysis indicated that, if the gaging station, Arroyo Seco near Soledad, California, was established solely to provide data to design a reservoir that would deliver a constant flow at approximately 80 cfs, the best plan of development would be to gage the stream for nine years using a low frequency of discharge measurement, and to initiate the project immediately after the nine-year period of gaging. If financial constraints require that the construction of the project be delayed beyond the nine-year period, gaging of the stream would be economically beneficial up to a length of record equal to 56 years. This delay of the initiation of the project beyond the nine-year optimum reduces the present worth of the project because of the delay in the reception of the benefits.

# CHAPTER 8

# SENSITIVITY OF THE PARAMETERS

The optimum lengths of record of 9 and 56 years that were obtained in the preceding chapter are applicable only for the specified conditions of cost, discount rate, use of the data, and level of development of the stream. Changing any one of these parameters will possibly negate the results of the foregoing analysis and make reanalysis necessary. This chapter demonstrates the changes that can be expected in the values of the optimum length of record when these parameters are varied.

## General Equations of Worth, Cost, and Present Value

The effect of varying either the use of the data or the level of development of the stream or of varying both would be a change in the function f that relates the worth of the data to its information content. Variation of the use of the data cannot be evaluated quantitatively at present because studies of uses other than reservoir design have not been completed. It is expected that the form of the relation that was empirically determined in Chapter 7 will not necessarily remain constant for other uses of the data. This hyperbolic form should be sufficient, however, to describe the worth function for reservoir design at other levels of development of Arroyo Seco and for other streams. The coefficients of the relation should vary with the stream and the level of development. This latter hypothesis can be partially tested by analyzing data from Dawdy et. al. (5) for Arroyo Seco at a development of 70% of mean flow. Figure 10 is a plot of the 70% yield data. The straight line on figure 10 describes the equation,

$$W = 37 \times 10^6 - 200 \times 10^6/N \text{ ............... } (8c).$$

Equation 8c is the same hyperbolic form as equation 8a, and it is, therefore, assumed that this form is descriptive of the worth relation if the data are used in the design of a conservation reservoir. The general form of equation 8 is written

$$W = B_m - K/N = B_m \ (1-K'/N) \text{ ............... } (8d)$$

in which $B_m$ = the expected net benefit from a project designed using very long streamflow records; K = a coefficient that is related to the variability of flow and the level of development of the stream; and K' = the ratio of K to $B_m$.

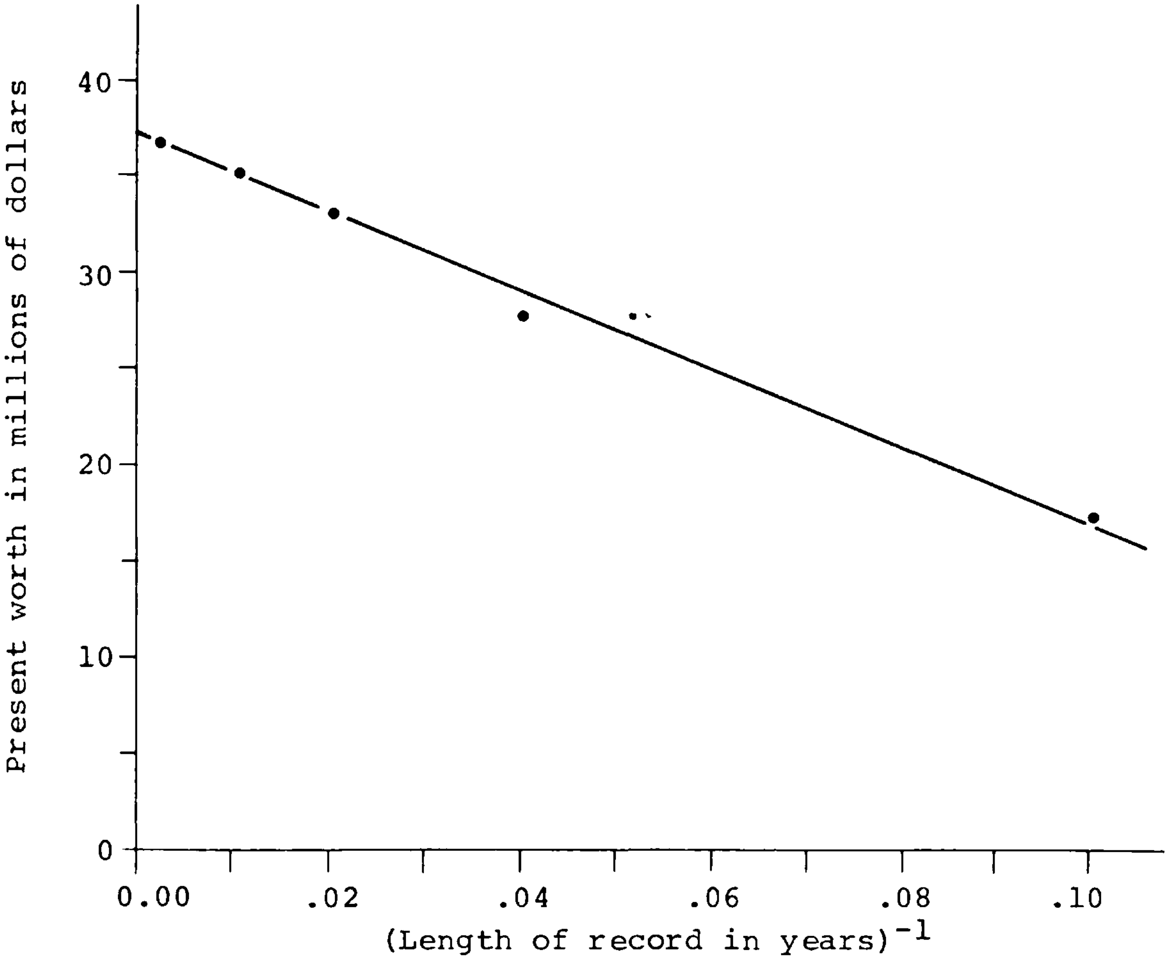


Figure 10.--Present Worth of Data at the Initiation of a 70% Yield Project.

The general equation for the present value of the worth of the data is obtained by substitution of equation 8d into equatin 11. This substitution yields

$$P_{WB} = B_m \ (1-K'/N)/(1+R)^N \ \ldots\ldots\ldots\ldots\ldots\ldots \ (11a)$$

in which $P_{WB}$ = the present value of the worth of the data.

Because the annual cost of obtaining data is incurred approximately uniformly throughout the year, the general form of the cost equation is an integral form of equation 1:

$$C_t = C_F + C_a \int_0^N dN \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ (1a)$$

in which $C_F$ = the sum of the fixed costs; and $C_a$ is defined by equation 2. The general equation of the present value of cost is

$$P_{WC} = C_F + C_a \int_0^N dN/(1+R)^N \ \ldots\ldots\ldots\ldots\ldots \ (11b),$$

and the general equation for the present value of net benefit of the data is

$$P_W = P_{WB} - P_{WC} \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ (11c).$$

## Variable Waiting Period

For the case in which the waiting period, as defined in Chapter 5, is a variable and is equal to the optimum length of record at the gaging station, the optimum length of record can be obtained by determining the maximum present value of the record. This maximum is found by expanding equation 11c, taking its partial derivative with respect to N, and equating this derivative with zero. This procedure results in the following equation

$$B_m \, [K' N^{-2} (1+R)^{-N} + K' N^{-1} (1+R)^{-N} \ln (1+R) - (1+R)^{-N} \ln (1+R)] - C_a (1+R)^{-N} = 0 \ldots\ldots\ldots (17).$$

Letting L equal the ratio of $C_a$ to $B_m$ and simplifying the result, equation 17 is rewritten

$$K' + K' N \ln (1+R) - N^2 \ln (1+R) - N^2 L = 0 \ldots\ldots (17a).$$

The value of the optimum length of record for any combination of values of the parameters, L, K', and R, can be determined by substituting the set of values into equation 17a, and solving for N.

The comparability between the use of equation 17a and the method presented in Chapter 5 for the determination of the optimum operating procedure was checked by substituting the values of $L = 0.00006$, $K' = 2.5$, and $R = 4.625\%$ that were obtained from the data in Chapter 7 into equation 17a. The resulting value of 8.8 years compares very favorably with the 9 year figure obtained in Chapter 7.

Equation 17a was used to determine the sensitivity of the value of N to the individual parameters. This sensitivity analysis was accomplished by determining the values of N that corresponded to a range of values for a particular parameter while holding the other parameters at fixed values.

## Effect of Cost of Data

Because of the small magnitude of the cost of data relative to the net benefit derived from its use, it was assumed that large variations of this parameter will have little effect on the optimum length of record. This assumption was checked by determining the value of L that would reduce N from 9 years to 8.5 years in the example of Chapter 7. A value of 8.5 years, along with values of $K' = 2.5$ and $R = 4.625\%$, was substituted into equation 17a. The resulting value of L was 0.0027. This value indicated that the ratio of maximum expected benefit from the data to the annual cost of obtaining the data would have to drop to approximately 400 for $C_a$ to become significant in the determination of the optimum length of record. At the current cost of obtaining streamflow data this ratio is probably representative of a very small irrigation operation. Because of the unlikelihood of a project of this size depending to a great extent on streamflow data, the cost of obtaining data was not considered pertinent in the remainder of this section.

## Effect of $K'$

$K'$ is a function of both the variability of flow at a prospective site and the projected level of development at that site. Values of $K'$ at Arroyo Seco near Soledad, California were found to be 2.5 for 50% yield development and 5.4 for 70% yield. For an interest rate of 4.625%, the optimal lengths of record for the two yields were found to be 8.8 years and 14.0 years, respectively.

At present it is not possible to define the function that relates $K'$ to flow variability and level of development. Because $K'$ appears to be a sensitive parameter in determining the optimum operating procedure, however, a graph of optimum length of record vs. $K'$ for a range of values of $K'$ that might be expected is given in figure 11. Figure 11 is based on an interest rate of 4.625%.

## Effect of Interest Rate

Because of the large effect that interest rate generally has in economics, it could be expected to be a sensitive parameter in the determination of the optimum length of record. Inspection of equation 17a, however,

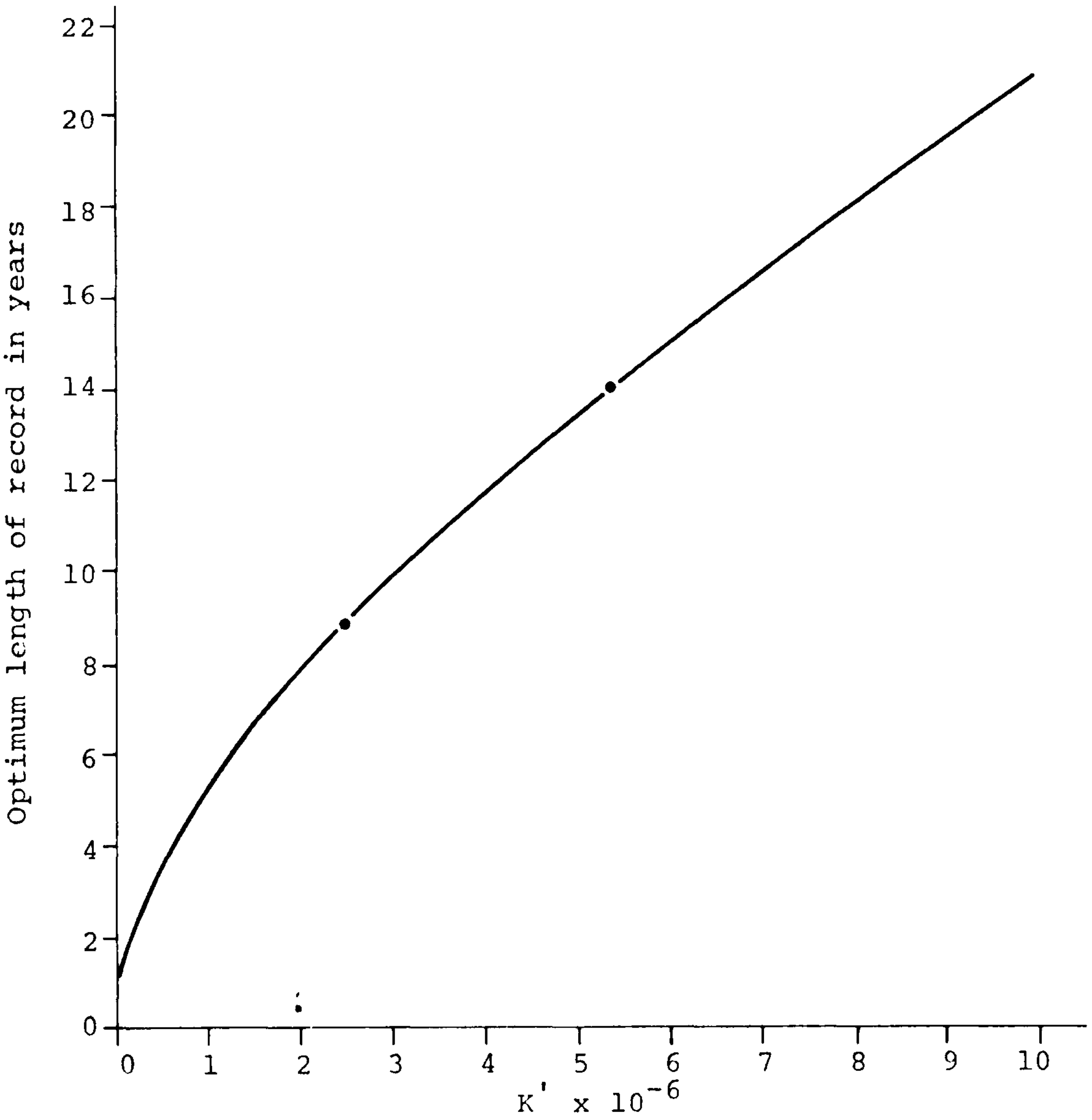


Figure 11.--Optimum Length of Record vs. K'.

indicates that R enters the relation only within a logarithmic operation. Its sensitivity is, therefore, somewhat diminished.

Interest rates used in discounting projects to their present worth have in the past ranged as low as 3%. The current trend, however, is for a gradual increase of R with time. The range of R's that might be encountered is from 3% to 7%. The sensitivity of N to R for this range is illustrated in figure 12. This illustration is based on the 50% level of development on Arroyo Seco.

The variation of N from 7.5 years to 10.5 years is much less than the variation caused by fluctuation of K'. It can be observed in figure 12 that the sensitivity of N to R is decreasing in the direction of increasing R, which is the direction of the current trend in interest rates.

## Hierarchy of Sensitivity

The hierarchy of sensitivity of the optimum length of record to its parameters can be extracted from figure 11 and 12 and from the subsection on the effects of costs. It is apparent from the ranges of N of figure 11 and 12 that the physical parameter K', which represents the flow variability and level of development, is much more sensitive than are the economic parameters of cost and interest

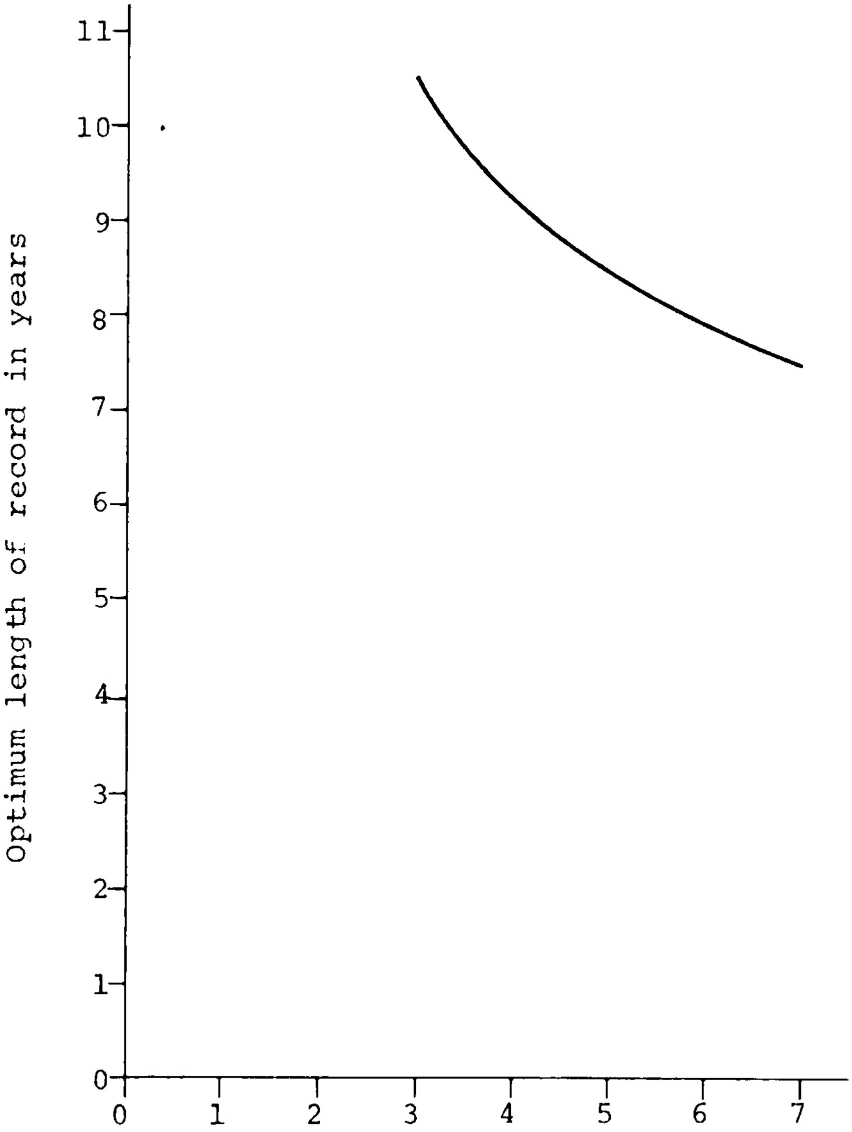


Figure 12.--Optimum Length of Record vs. Interest Rate

rate. Of the two economic parameters, interest rate is more effective in varying N than is the highly insensitive variable, cost of obtaining data.

## Fixed Waiting Period

If the length of the waiting period is fixed by considerations other than the obtaining of design data, the optimum length of streamflow record is determined from the general form of equation 16, which is

$$C_a = K/N^2\ (1+R)^N \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (16a).$$

Comparison of equation 16a with equation 17a shows that the optimum lengths of record for both the fixed and the variable waiting period are controlled by the same parameters, though in different forms. Equation 16a is used to test sensitivities of the parameters in a manner similar to that used in the preceding section.

### Effect of Cost of Data

The effect of changes in the optimum length of record for a range of annual costs from \$600 to \$6000 was determined graphically. This range of costs appears to bracket all values that might be encountered. The sloping line of figure 8 illustrates this effect for values of K equal to \$50 x $10^6$ and of R equal to 4.625%. The range in values of N for these parameters was found to be approximately from 65 years to 35 years. The optimum length of record for the fixed-waiting-period case is defined as the smaller of two quantities: N obtained from equation 16a or the length of the waiting period. Because it would be very unusual for the waiting period of a firmly committed project to be as long as even 35 years, the effect of annual cost of obtaining data on the optimum length of record would in most cases be nil.

## Effect of K

As with the case of $K'$, the relation of K to the physical properties of the streamflow and the contemplated project cannot at present be defined. Only two values of K are currently available: $\$50 \times 10^6$ for 50% development of Arroyo Seco, and $\$200 \times 10^6$ for 70% development of the same stream. A measure of the sensitivity of N to values of K was obtained, however, for a range of K values that appeared to be within the realm of expectancy. This range was from $\$10 \times 10^6$ to $\$400 \times 10^6$. Figure 13 illustrates the sensitivity of K for values of $C_a = \$1215$ and $R = 4.625\%$.

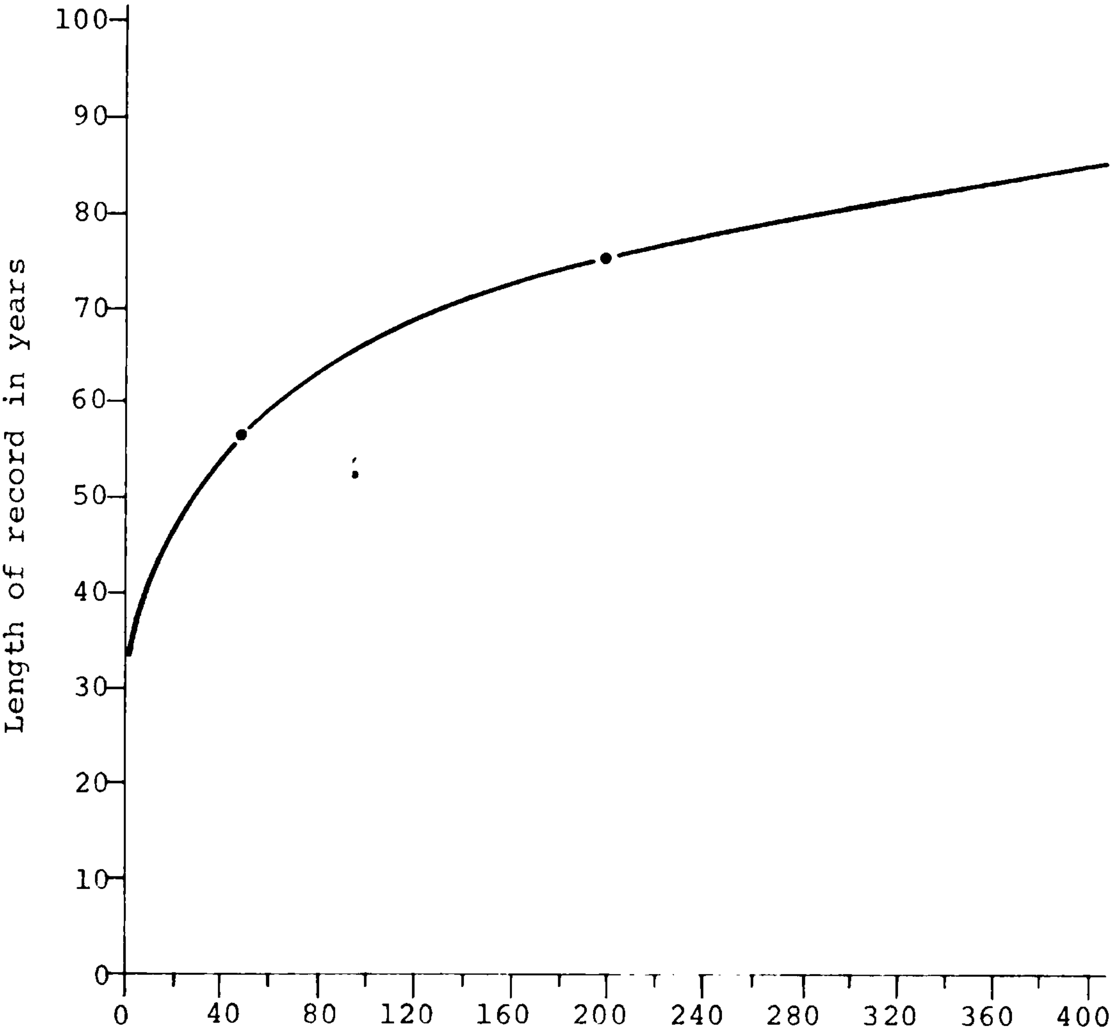


Figure 13.--Length of Record vs. K.

Inspection of figure 13 shows that all values of N for this range of K are greater than 35 years. By the same reasoning as was used in the preceding subsection, it can be stated that the value of K is very seldom critical in the determination of the optimum length of record.

### Effect of R

Interest rates from 3% to 7% were substituted into equation 16a along with values of $K = \$50 \times 10^6$ and $C_a = \$1215$ in order to test the sensitivity of the optimum length of record to the value of R. Figure 14 illustrates the values of N that were obtained from these substitutions. Inspection of figure 14 shows that all values of N are greater than could be expected for the length of the waiting period. Thus interest rate is also insignificant in the determination of the optimum length of record for a project with a fixed waiting period.

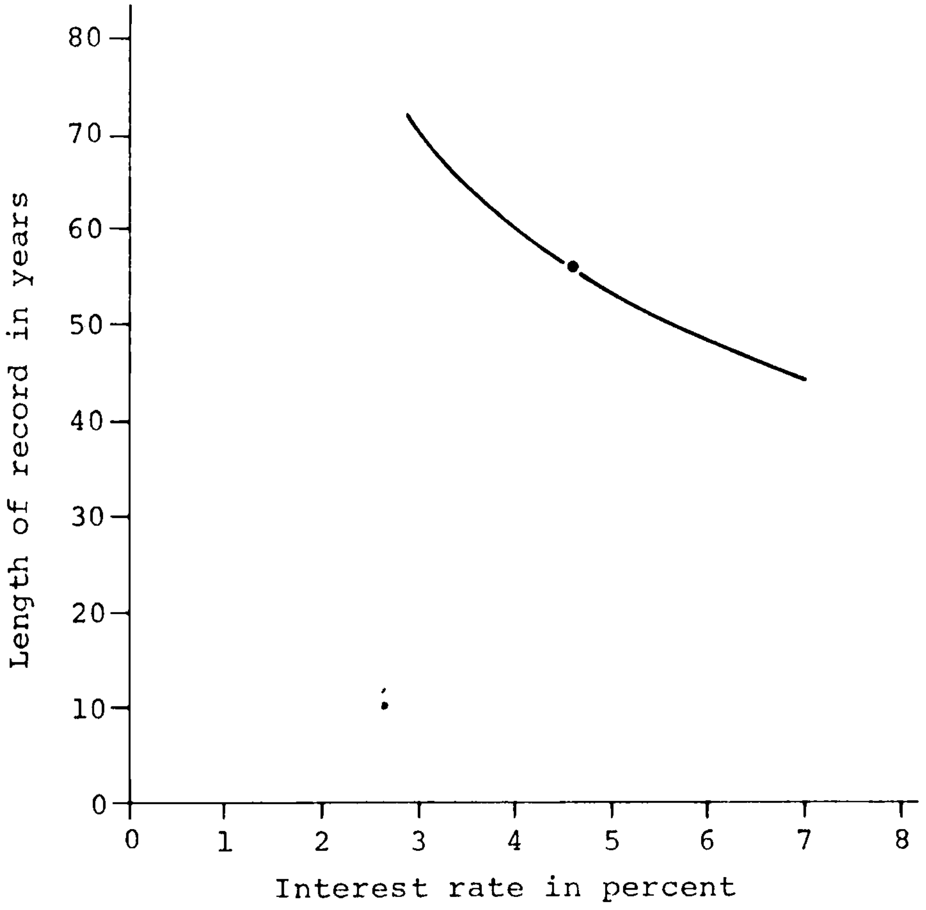


Figure 14.--Length of Record vs. Interest Rate.

# CHAPTER 9

## CONCLUSIONS AND SUGGESTIONS FOR FURTHER STUDY

Existing discharge and stage measurements and a worth of data study were used to illustrate and test the applicability of the scheme for determining the optimum operating procedure at the gaging station on Arroyo Seco near Soledad, California. Definition of the optimum operating procedure was accomplished only in part in the example. The optimum length of record was satisfactorily defined for two different economic conditions. The optimum frequency of discharge measurement was indeterminate; however, a qualitative indication of the analysis showed that the frequency currently being used is apparently too high.

The indeterminacy of the optimum frequency was ascribed to the small magnitude of the standard error of computation in comparison with the natural variance of the streamflow. This standard error was small because of the construction at the gage site of a small, artificial, rock control, which made the stage-discharge rating more stable than might have occurred under natural channel conditions.

This indeterminacy does not negate the use of the scheme, but it does demonstrate one limitation of its use.

The analysis of the optimum length of record yielded a range in values from 9 to 56 years. The 9 year value is the most economically efficient if the period of gaging is followed immediately by the initiation of the project. If the project is to be delayed for economic or political reasons beyond the nine year optimum, it is beneficial to continue gaging the stream so long as the delay is not greater than 56 years. The limiting values of 9 and 56 years, although based on an empirical equation for the worth of the data, appear to be reasonable and acceptable figures.

A sensitivity analysis was performed on the variables that effect the optimum operating procedure. The optimum length of record for the fixed-waiting-period case was concluded to be equal to the length of the waiting period because of the large values of N that were obtained for all combinations of the parameters tested. Furthermore, it was considered insensitive to hydrologic or economic considerations. For the variable-waiting-period case, however, N was found to be highly sensitive to the parameter K'. It was considered to be less sensitive to interest rate and annual cost of operating the gaging station than to K'.

The high sensitivity of N to K' requires that K' be defined accurately in order that an efficient estimate of the optimum length of record can be made. Dependency of K' on both the flow characteristics and the level of development of the stream presents difficulties in its definition. Evaluation of the flow characteristics requires some foresight into the size and location of the future project, while the size and location are to some degree dependent upon the unknown flow characteristics. Thus, dependence between the independent variables complicates the estimation of K'. A further complication develops from man's often limited ability to predict his needs for water even a short distance in the future.

The minimum values of the optimum length of record were of the order of magnitude of ten years. Even values that are this small are often too large to satisfy the political and social expediencies that might be attached to a water resources project.

An apparent solution to the problems of defining K' and meeting expediencies is the establishment of an efficient network of long-term streamgaging stations. Through regression analyses based on the network stations and the basin parameters, streamflow characteristics can be estimated at any point within the region covered by the network. This ability to transfer information on the streamflow characteristics makes it possible to define the most efficient sizes and locations of projects in the region and also makes possible the definition of K' through studies like that of Dawdy et. al (5). The definition of K' and the less sensitive parameters leads to an estimate of the optimum length of record. The political and social expediency of a shorter waiting period can also be met by synthesizing a record at a project sight that is equivalent to the optimum length of measured record. This synthetic record can be extracted from the records of the long-term network and a regression model based on a short record at the project site and the concurrent records at the network stations. A check of the equivalence of the synthetic record with the optimum length of record can be made by the procedure demonstrated by Langbein (13).

The proposal of a continuous, long-term gaging-station network is not by any means a new one. Its necessity is, however, reinforced by the conclusions drawn from this paper.

Several points that warrant further study were encountered in this paper. Among these are the following:

1. The scheme should be applied to a stream with a less stable channel to test the method of determining optimum frequency of measurement. A variation of at least 10% caused by the standard error of computation in the

reciprocal of information content would be desirable in the new study. This variation would require a definable standard error of computation whose square is at least 5% greater than the natural variance of the streamflow.

2. Tests of sensitivity of the error analysis procedure to the number of discharge and stage measurements used should be conducted.

3. Methods for determining the worth of data used for purposes other than the design of conservation dams should be developed in order that qualitative analyses can be made of the operating procedures used to obtain these data.

# APPENDIX I

## NOTATION

The following symbols are used in this paper:

$B$ = average annual benefit of the project;

$B_m$ = expected net benefit from a project designed using a very long streamflow record;

$B_N$ = net benefits derived from a streamgage;

$c$ = a coefficient;

$C_a$ = marginal cost of obtaining an additional year of record;

$C_c$ = cost of constructing an artificial control;

$C_F$ = sum of the fixed costs;

$C_i$ = initial costs of reconnaissance and construction of the gage;

$C_m$ = cost of one discharge measurement;

$C_r$ = annual costs of overhead and record computation and publication;

$C_t$ = total cost of operating a streamgage;

$C_1$ = average annual cost of construction of a project;

$C_2$ = average annual operating cost of a project;

$d$ = a decision variable;

$f$ = an undefined function;

$i$ = an index of a discharge group;

$I$ = information content;

$j$ = an index of length of record;

$k$ = number of nonzero groups of residuals;

$K$ = coefficient related to streamflow variability and level of development;

$K'$ = ratio of K to $B_m$;

$L$ = ratio of Ca to $B_m$;

$m$ = frequency of discharge measurement;

$M_J$ = marginal worth of data collected in the year j;

$N$ = length of record;

$p$ = average period of time between discharge measurements;

$P$ = monetary value of a cost or benefit;

$P_i$ = part of the time discharge is in the range i;

$P_s$ = summation of the $P_i$'s;

$P_w$ = present worth of a future cost or benefit;

$P_{WB}$ = present value of the worth of data;

$P_{WC}$ = present value of the cost of obtaining data;

$P_{wi}$ = present worth of data at the initiation of a project;

$R$ = interest rate;

$S_A$ = standard error of computation of annual mean discharge;

$S_c$ = standard error of computation of an instantaneous discharge determination;

$S_i$ = standard deviation of the residuals in a group;

$S_m$ = standard error of measurement;

$S_R$ = standard deviation of the residuals;

$S_v$ = standard deviation of the sampling distribution of the variance;

$T$ = time;

$V$ = true variance of the streamflow population;

$V_R$ = pooled estimate of the variance of the residuals;

$W$ = worth of a record.

# LIST OF REFERENCES

1. Allison, S. V., "Cost, Precision, and Value Relationship of Data Collection and Design Activities in Water Development Planning", Contribution No. 120, Water Resources Center, University of California, Berkeley, Calif., May, 1967.

2. Benson, M. A., "Factors Affecting the Occurrence of Floods in the Southwest", Water-Supply Paper 1580-D U.S. Geological Survey, 1964, pp. 46-67.

3. Burkham, D. E., and Dawdy, D. R., "Alluvial Streams - Error Analysis of Streamflow Data", Open-file Report, U.S. Geological Survey, 1968.

4. Carter, R. W., and Anderson, I. E., "Accuracy of Current Meter Measurements", Journal of the Hydraulics Division, ASCE, Vol. 89, No. HY 4, July, 1963, pp. 105-115.

5. Dawdy, D. R., Kubik, H. E., Beard, L. R., and Close, E. R., "Worth of Streamflow Data for Project Design - A Case Study", presented at the 49th Annual Meeting of Amer. Geophys. Union, Washington, D.C. April, 1968.

6. Dixon, W. J., and Massey, F. J., "The Variance: Estimation and Tests of Hypotheses", Introduction to Statistical Analysis, 2nd Edition, McGraw-Hill Book Co., Inc., New York, 1957, p. 109.

7. Fiering, M. B., "An Optimization Scheme for Gaging", Journal of Water Resources Research, Vol. 1, No. 4, 4th quarter 1965, pp. 463-470.

8. Fiering, M. B., "Markovian Flow Models", Streamflow Synthesis, Harvard University Press, Cambridge, Mass., 1967, pp. 28-65.

9. Fisher, R. W., "The Generalization of Null Hypotheses. Fiducial Probability", *The Design of Experiments*, 8th Edition, Hafner Publishing Co., New York, 1960, pp. 184-185.

10. Grant, E. L., *Principles of Engineering Economy*, The Ronald Press Company, New York, 1938, p. 46-93.

11. Halmos, E. E., Jr., "Byline Washington", *Civil Engineering - ASCE*, Vol. 38, No. 9, Sept. 1968, p. 128.

12. Kendall, M. G., *The Advanced Theory of Statistics*, 5th Edition, Vol. 1, Hafner Publishing Co., New York, 1952, p. 224.

13. Langbein, W. B., "How Long Should Gaging Stations be Operated?, *Proceedings 22nd. Annual Meeting of the Western Snow Conference*, Salt Lake City, Utah, April 1954.

14. Matalas, N. C., and Gilroy, E. J., "Some Comments on Regionalization in Hydrologic Studies", to be published in the *Journal of Water Resources Research*, Vol. 4, No. 6, Dec. 1968.

15. Thomas, H. A., Jr., and Burden, R. P., *Operations Research in Water Quality Management*, Harvard Water Resources Group, Cambridge, Mass., Feb., 1963, pp. 1-14 through 1-17.

16. Young, G. K., and Pisano, W. C., "Operational Hydrology Using Residuals", *Journal of the Hydraulics Division*, ASCE, Vol. 94, No. Hy 4, July, 1968, pp. 909-923.

17. "Monthly Streamflow Synthesis", *Hydrologic Engineering Center Computer Program 24-J2-L243*, U. S. Army Corps of Engineers, Sacramento, Calif., Sept., 1966.

18. "Water Resources Data for California 1965", Part 1, Vol. 1, U. S. Geological Survey, 1966, p. 300.