

USE OF STORAGE WATER IN A HYDROELECTRIC SYSTEM

by

John Dutton Conant Little

S.B., Massachusetts Institute  
of Technology (1948)

SUBMITTED IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY  
February, 1954

Signature of Author .....  
Department of Physics, October 11, 1954

Certified by .....  
Thesis Supervisor

Accepted by .....  
Chairman, Departmental Committee  
on Graduate Students



ABSTRACT

USE OF STORAGE WATER IN A HYDROELECTRIC SYSTEM

by

John Dutton Conant Little

Submitted to the Department of Physics on October 11, 1954 in partial fulfillment of the requirements for the degree of doctor of philosophy.

The big water reservoirs of a hydroelectric system collect water during high river flows for use during low flows. The problem considered is how to use such stored water in the face of uncertain future flow. A method for handling the problem has been formulated and used to obtain numerical results for a simple case on a high speed digital computer.

Best water use is taken to be that which minimizes the expected cost of operating the system. The required expected cost functions are set up and the problem shown to be somewhat like a business inventory problem. A solution was calculated on a digital computer for a model of a one dam hydro system patterned after the Grand Coulee Dam on the Columbia River. The solution was used to operate the model with the historical record of flows for the Columbia River. The results are compared with those obtained from a conventional operation of the system.

The water use determined by the minimum expected cost solution was found to give a slightly lower average cost than the conventional operation.

Thesis Supervisor: Professor P. M. Morse  
Title: Professor of Physics

ACKNOWLEDGMENT

The author thanks Professor Philip M. Morse for his help and supervision.

He thanks Professor Boehne for introducing him to the problem and members of the Bonneville Power Administration for much help in understanding hydro systems. He thanks Peter Schweitzer for many discussions.

This work has been supported in part by the Army Office of Ordnance Research, the Office of Naval Research, and the Bonneville Power Administration. Digital Computer Laboratory time was made possible by the Office of Naval Research.

TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT.....	2
ACKNOWLEDGMENT.....	3
TABLE OF CONTENTS.....	4
I. INTRODUCTION.....	7
II. DESCRIPTION OF THE PROBLEM.....	11
Hydro Systems.....	11
The Short Range Problem.....	15
The Long Range Problem.....	18
III. THE MODEL.....	24
IV. THEORY.....	34
Current Methods.....	34
Other Methods.....	36
Expected Value Method.....	38
Setting Up the Equations.....	40
Hydro Case.....	44
V. NUMERICAL METHODS.....	48
Probabilities.....	48
Optimization Calculation.....	56
Programming.....	65

	<u>Page</u>
VI. RESULTS OF THE CALCULATIONS.....	68
Kinds of Operation.....	68
Water Use Tables.....	71
Results of Operation with the Historical Flow.....	74
Curves of System Behavior.....	75
Average Costs.....	78
Other Results.....	83
Accuracy of the Calculation.....	87
Stability.....	89
VII. POSSIBLE GENERALIZATIONS OF THE MODEL.....	96
VIII. SUMMARY.....	101
APPENDIX.....	106
LIST OF FIGURES AND TABLES.....	116
BIBLIOGRAPHY.....	117
BIOGRAPHICAL NOTE.....	118

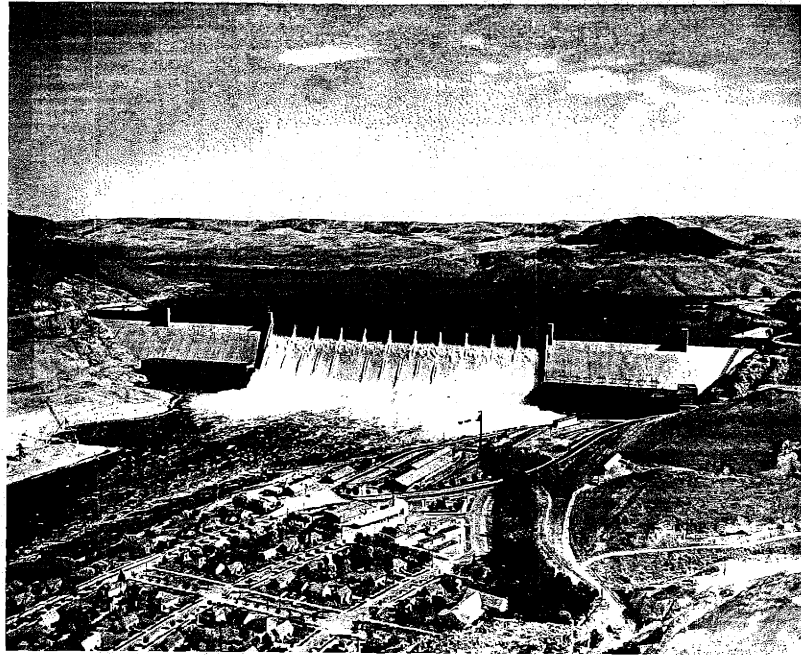


Fig. 1. Grand Coulee Dam, Columbia River, Washington.  
(Photograph courtesy of the Department of the Interior)

## I. INTRODUCTION

The big water reservoirs of a hydroelectric system collect water during high river flows for use during low flows. The problem considered in this paper is how to use such stored water in the face of uncertain future flow. A method for handling the problem has been formulated and used to obtain numerical results for a simple case on a high speed digital computer.

The use of stored water is taken to be an allocation problem not a dispatching problem; that is, the object is to find out how much storage water to allot for use during the next week or month rather than to fix the instantaneous generation of each plant throughout the day. This latter problem, short range power dispatching, is not a simple one and has been studied by various writers (1,2,3). The long range problem considered here suppresses the fine details of daily operation but introduces effects ignored in short range analyses, namely, the uncertainty of future stream flows and the variation of head at big reservoirs over a season.

After a year is over it is usually possible to look back and discover a way of using storage that would have been more effective. Unless weather prediction and hydrology improve to the point where future river flow can be exactly known, such effectiveness is unattainable. Until then we cannot speak of maximizing the return

from a system but only of maximizing the average return.

The method of treating the problem, therefore, is to maximize the expected (average) return, or, alternatively, minimize the expected cost of operating. However, the expected return for a year depends not on just one decision about using storage but on a whole set of decisions throughout the year. Furthermore, these decisions must depend on the particular conditions prevailing at the time they are made, at least on the quantity of water in storage and the prevailing river flow. Thus we do not seek one number for an answer, but, as it will turn out, a set of functions.

In performing the maximization, none of the techniques of elementary calculus, Lagrange multipliers, calculus of variations, or linear programming seem to be easily applicable. There is, however, a way of setting up the problem so that, in moderate complexity, it may be handled by machine computation. The point of view is not novel, having been used at least by Dvoretzky, Kiefer, and Wolfowitz (4) in discussions of inventory problems and by Bellman (5) in studies of dynamic programming. Indeed, the hydro problem treated here can be considered as a special case of some of the discussions of these writers.

The method has been used for numerical calculations on a one dam hydro system, some of whose characteristics are drawn from the Grand Coulee plant on the Columbia River. Results are presented



to show the response of the system to the historical flows of the Columbia River in the period 1914-52, when operation is determined so as to optimize the expected return as judged by a given criterion. These are compared with the results of operation determined by a simple example of current methods. Extremes of criteria are tried to show quantitatively the conflict between the requirements of maximum energy and minimum risk of power shortage.

A simple hydro system was deliberately chosen for the numerical work in order better to understand what was going on physically. The result is, however, that the particular numbers do not necessarily apply to any real situation in the Northwest. This difficulty is not inherent in the theory; its applicability is limited primarily by the degree of complexity which can be handled computationally.

The computer used to calculate the numbers in this paper is the Whirlwind I digital computer at M.I.T. As compared to other current computers, it has the property of being very fast for manipulating low accuracy ( $4\frac{1}{2}$  decimal digits) fixed point numbers. By somewhat more elaborate programming than might ordinarily be used, this property was utilized to make the computation rather fast. Tables for determining optimum system operation were calculated in times ranging from 5 minutes to  $\frac{1}{2}$  hour, depending on the accuracy sought.

Throughout the discussion, illustrative examples are drawn from the hydroelectric system of the Pacific Northwest. The background for these resulted from two valuable trips the writer made to Portland, Oregon where he talked at length with personnel of the Bonneville Power Administration (BPA) and briefly with personnel of the Northwest Power Pool Coordinating Group. The author hopes that he properly understood all he was told and in any case assumes responsibility for the statements in this paper.

## II. DESCRIPTION OF THE PROBLEM

This section discusses water use in hydroelectric systems with specific reference to the Pacific Northwest. The purpose is to sketch enough of a picture of a hydro system to find the storage problem, put it in relation to the rest of hydro operation, and lead toward a useful model for making calculations.

### Hydro Systems

A hydroelectric system is a device for providing electric energy to customers. It consists, roughly, of one or more generating plants run by water power, none or more other sources of power, wires to transport the energy, the customers, and some person or persons who operate the system.

The load of the system is the amount of power its customers are taking at a given instant, or in certain contexts, the amount of energy they require in a given interval of time. In regions heavily dependent on hydro, the amount of energy available is uncertain because future river flow is unknown. The load is then divided into firm and interruptible parts. Customers contracting for interruptible power can be required to stop using it on short notice. Firm power customers can expect to receive power unless water conditions become worse than those used in calculating

the amount of firm energy to be contracted.

Hydro plants can be divided into two classes: run-of-the-river plants and storage plants. Run-of-the-river plants cannot store any appreciable amounts of water but must use the river flow as it comes along. They may have what is known as pondage; that is, they have enough of a pond behind them so that during the peak load hours of the day they can draw down a little and during the slack night hours refill. Storage plants have reservoirs big enough to collect seasonal high flows for subsequent use. On the Columbia River, Bonneville is a typical run-of-the-river plant, Grand Coulee a large storage plant. During a low winter flow Bonneville could easily draft its maximum permitted draw-down of 4 ft in a day. Grand Coulee when filled could discharge at full gate for a month without inflow.

Most big hydro systems have steam plants for generating additional energy. Indeed a system with more hydro than steam is a rarity. The Pacific Northwest, where the growth of industry has not yet outstripped the water power potential, is one of these, supplying 90% of its load with hydro in an average year. Besides steam plants, there is often auxiliary power from interchange with other systems, and, in times of severe shortage, small amounts can be obtained from diverse sources. For instance,

in the 1952-53 Northwest power shortage a Navy Power Train tied 10 MW into the system.

Since customers want to rely on their source of electricity, an important property of a hydro system is how much energy it can produce in the dry season of a dry year. Clearly, however, if the region load gears itself to using no more energy than this, water will go to waste in wet and even average years.

The input of the system, water, fluctuates widely, whereas the output, energy to the load, is desired to be smooth. Storage reservoirs help by filtering out some of the seasonal fluctuations and in certain cases, part of the fluctuations having periods more than a year.

Two other ways of making use of high water are by interruptible loads and auxiliary power. Interruptible loads are a means of fitting the load to the vagaries of the water, but only a certain class of customers can use it. Auxiliary power, for instance steam generation, may cost more to operate than hydro, but in high water years need not be run. Its existence makes possible the contracting of more firm energy than otherwise. More of the water power is utilized but because steam plants must be built and occasionally run, the average cost is higher.

The Pacific Northwest is not a single system. There are more

than 30 sizable hydro plants on various rivers divided among local public utilities, private utilities, and the federal system. The biggest plants are on the Columbia River and are federal. The various systems are independently managed, but they are interconnected electrically and associated together in the Northwest Power Pool. The Pool is a voluntary organization which coordinates the necessarily interdependent generations of its members. For instance, detailed studies are made of how to meet the load in a critically dry year.

The money collected from the sale of energy in a big hydro system comes to many million dollars. Therefore, improvements in operation on the order of fractions of a percent can be financially significant.

The next paragraphs discuss water use under the headings of the short range and long range problems. The short range problem is defined as that of daily dispatching of power and is much concerned with peaking and with load variations. By the long range problem is meant the allocation of stored water for use in some period on the order of a week to a month. Here the uncertainty of river flow in coming weeks and the variation of head with drawdown become important.

### The Short Range Problem

As customers throw switches the hydro system must respond by increasing or decreasing the power generated. When there are many plants in the system, there are generally good and bad ways of dividing the generation among them. The problem of optimum dispatching has been worked on by various writers (1,2,3) and will only be outlined here in its main features.

First of all, peaking is expensive. During certain hours of the day the load is on the order of 40% greater than the average. Facilities must be available to meet it and inefficiencies in operation will be tolerated in order to deliver the energy when it is wanted. For example, the Bonneville plant peaks by drawing down its pond during the day and letting it fill up at night. The result is a head averaging perhaps one foot lower than the maximum of 64 feet. Thus more than 1% of the energy is sacrificed in order to peak. It might therefore seem desirable to take more peaking at Grand Coulee where the large reservoir prevents much daily fluctuation of head. But Grand Coulee is further from the heavy load centers, and transmission losses go up about as the square of the power. Besides, Grand Coulee has an operating limitation on its rate of change of discharge because of river bank erosion by its tailrace.

Such operating limitations hem in the short range problem. There are many others: A plant has only so many generators. Transmission lines can carry only so much power. The Bonneville pool can only be dropped 4 feet before ferryboats upstream may go aground. Where flood control is important, it always has priority over electric generation. Too low discharge may make riverside wells go dry or impede navigation. At the Kerr plant in Montana it is said that generation is not allowed to drop below a certain value because low discharge exposes islands in the river downstream, whereupon fishermen go out on them and may be drowned when the plant peaks.

Another complication of the short range problem is the routing of the water down a stream with several plants. The cut back of generation on Sunday at Grand Coulee puts a hole in the Columbia River that reaches Bonneville in the middle of the week when it is least desired. The Rock Island plant which follows Grand Coulee on the river, has not Grand Coulee's generating capacity and, if Grand Coulee is wide open, Rock Island is forced to spill water.

Uncertainty of river flow is not a difficulty in the short range problem. The water a plant uses tomorrow is in the river today somewhere upstream. Certain flashy streams are exceptions



but big rivers rise and fall slowly and if there is appreciable storage in plants along a river, the plants themselves dictate the flow.

The load on the other hand is subject to considerable short term variation. Cloud cover causes people to turn on their lights; cold weather, their heaters. Whether or not the small and medium user of power turns on his equipment is up to him and a statistical variation of load results. The aggregate is fairly regular, depends on the time of day, and is predictable within a few per cent from day to day. To allow for fluctuations a certain amount of spinning reserve is maintained.

The difficulty of finding the best short range operation probably depends strongly on the system. Best operation may be obvious. For instance, the marginal cost of steam may increase so rapidly that taking the peak with steam is prohibitively expensive. Then it may happen that a given hydro plant is the best for taking the peak and so, as much as fits the operating restrictions is taken there, the rest in other plants in order of efficiency. Bottlenecks in transmission may make it impossible to get more than a certain amount of power from a desirable source and so the maximum is taken. Thus, the best operation may be some combination of extremes determined by operating restrictions and fairly

easily discernible to some one familiar with the system. However, if a system is complicated and at the same time has flexibility of operation, the determination of best dispatching may be difficult.

#### The Long Range Problem

The long range problem is to decide how much stored water to use in the next week or month, taking into account the effect this may have on operation several weeks or months away. It is a problem of balancing present against future benefits.

There may not always be a problem. If the hydro generating plant has the capacity to use the highest flows the river produces and maximum energy is desired, it will be obtained by keeping the reservoir full at all times. In some cases flood control regulations may almost completely dictate reservoir operation. Also, if there is no flexibility in the load seen by a plant, in other words, if there is no interruptible power to switch on or off, or no steam or other hydro plants to take over the load, or if there is an unsellable power surplus, then the plant just generates the load it sees and the reservoir may be forgotten.

An example of a reservoir where good water use requires a

balance between present and future needs is Grand Coulee on the Columbia River. There the high spring flow always fills the reservoir. In winter the natural flow falls off sharply and so the stored water is drafted. If this is started too early in the fall, the head is soon materially reduced and the natural flow gives appreciably decreased power. If, in addition, the winter flows are low, storage may run out and a power shortage result. On the other hand, if the high water of spring arrives and there is water remaining in storage, it is too late to put that water to use. Best results require a balance between the advantages of early and late water use in the face of uncertain future flow.

Such uncertainty is an essential feature of the long range problem. During the 1952-53 season in the Northwest, for instance, the river flow in the fall was so low that in November part of the firm load was curtailed (6). There was water in storage which could have been used to carry the load but, if that water had been used and the flow had continued to decrease, much greater curtailments would have been forced later in the season. As it turned out, January brought large amounts of warm rain. Using hindsight the curtailment could have been avoided. But there was not then and there is not now any way to be sure of the size of future flow. In deciding the magnitude and timing

of the curtailment it was necessary to balance the desire to put off curtailment as long as possible in hopes of rain against the seriousness of greater curtailment later on in case flows decreased. Each of these possibilities had to be weighed, at least qualitatively, by the probabilities of the corresponding occurrences of stream flows. In many less dramatic problems of water use, the same balancing problems arise. The research reported here sought ways for getting the flow probabilities into decisions in a quantitative way.

The future load as well as flow is uncertain. The weather has some effect and perhaps more substantial, economic fluctuations affect the load. However, estimations of load six months in advance are usually correct to within 10% whereas the flow six months in advance is not known within a factor of two or more. Consequently, in this work the load has been assumed known, while the flow has been characterized by probability distributions.

With respect to uncertainty in river flow, Grand Coulee is fortunate in having a seemingly guaranteed annual refill. So lusty is the Columbia in spring and early summer that the minimum spring runoff on record would easily refill the usable storage of Grand Coulee's 150 mile long reservoir even with the plant generating at full gate. Hungry Horse Dam on the South Fork of the Flathead River, on the other hand, which has somewhat more

than half as much storage as Grand Coulee, could not be filled in one season if completely drawn down. Therefore, decisions about using this water have to take account of the needs of a year or more later.

The long range problem for multiple reservoirs is likely to involve water routing. The water from Hungry Horse eventually reaches the Columbia and is used at Grand Coulee and various downstream plants. Before it gets there, however, it passes through two large natural lakes. These lakes have dams at their outlets which can regulate the water level a few feet. Because of the large area of the lakes, these few feet represent a considerable and desirable addition to the storage of the system. On the other hand natural lakes have the property that it is hard to get the water out of them when they are low. This is not surprising for it helps explain their existence. Rivers, too, when low, will not carry water as fast as when high. This means that water released from Hungry Horse when the lakes and rivers are low may not get down where it will do the most good for an appreciable number of days. Such considerations as these can make determination of good operation of interconnected reservoirs complicated. One rule that often applies is that upstream reservoirs should be drafted before downstream so that head will be kept up downstream where the natural flow is usually larger.

Among the principal operating limitations that affect the long range problem are restrictions on maximum generation and minimum discharge. Flood control often dictates reservoir levels at certain times of year. Miscellaneous restrictions arise: Kootenai Lake just over the border in Canada is regulated by international agreement to protect some dike formed land at the head of the lake which is afflicted with excessive seepage at high lake levels. At another lake, the water level is being held within certain limits in an experiment to find out whether fluctuations are killing fish eggs.

Although allocating stored water is primarily an energy problem, peaking cannot be ignored. It is possible to have the energy available in the river or in storage but not be able to supply it for lack of peaking. The average energy can be available but during heavy load hours the generating capacity may not be available to meet the peak load. A more subtle situation is one where both peaking and energy are available but the load cannot be met. At a run-of-the-river plant, for instance, it may be possible to reach the peak but not hold it long enough before the allowed pondage runs out, even though the next night's refill would replenish it. Peaking requirements, then, can limit the use of available water.

To summarize: the long range problem is defined as that of allocating stored water for use over the next week or month. The advantages of immediate use of the energy must be balanced against those of future use, taking into account the uncertainty of future river flow. Generation and reservoir regulation are restricted by various operating limitations. Multiple reservoir systems increase the number of variables and introduce water routing to the problem. Peaking must be taken into account in so far as it restricts the use of available energy.

### III. THE MODEL

Mathematical model is a term currently applied to a quantitative description of an actual situation. The word model emphasizes that the description usually does not take into account every ramification of the real problem but that certain details thought unimportant or too difficult are omitted. The term also has the healthy effect of forcing a discrimination between whether a method is a good way to solve a model and whether a model and its solution are applicable to the real problem.

A simple model of a hydro system was picked for making a study of the long range problem. If the results are considered to warrant it, many additional factors can be introduced. The basic model is essentially an idealization and simplification of Grand Coulee.

A single reservoir is considered. It has a volume of usable storage about equal to that of Grand Coulee, 2,600 Kcfs days. The unit of volume used, Kcfs day, is the volume of water represented by 1000 cubic feet per second flowing for a day.

The reservoir is idealized to have vertical sides and so the volume of water in storage is proportional to its depth. The total usable draft in the model is 80 feet as is Grand Coulee's.



There is an additional head at Grand Coulee whose value varies a few percent with discharge because the tailrace elevation changes. The model takes the additional head as a constant 264 ft. To show what a massive block of concrete Grand Coulee is, we note that it is more than twice as high as Niagara Falls.

The maximum discharge through the turbines is taken as 70 Kcfs and constant. The real maximum at Grand Coulee varies with head and ranges from about 75 to 90 Kcfs. However, in the long range problem we wish to talk about the average discharge over perhaps two weeks or a month. As mentioned earlier, if a plant has to supply peaking, the maximum average energy it can deliver is lower than that implied by its full gate discharge. Further, peaking is more a problem of good times than bad times because when water is low, there is usually unused hydro capacity standing by which can take the daily peak. In the interests of simplicity, then, the maximum discharge was taken as a constant, somewhat less than the real maximum at low head and considerably lower than the maximum at high head.

Minimum discharge is taken as 20 Kcfs. This is only half the figure usually quoted for Grand Coulee, but was picked for reasons discussed below.

The energy generated in a time interval is taken as the product of the volume discharged, the average of the heads at the beginning and end of the interval, and a constant. The constant is approximated from tables for Grand Coulee and neglects variations of turbine efficiency with head and discharge.

The load is taken constant and equal to the generation at maximum discharge with full head. The load in most regions is not constant throughout the year. In the Northwest the load goes up 10 or 15% in winter. Choosing the load as the maximum hydro generation means the system is dominated by hydro but that water will not go to waste until the flow exceeds the maximum discharge.

A variable must be selected which determines storage use and thereby system operation. It may be called the decision variable. In this work, it is denoted by  $s$  and is the volume of storage water planned for use in the next time interval. A negative volume implies filling the reservoir. The volume does not necessarily have to be used because its magnitude is fixed before the flow actually comes. When the flow does come it may be discovered that use of the volume  $s$  would violate operating restrictions. For instance, if the flow is high, it may not be possible to use all the water planned because discharge would then be greater than permitted.

A different quantity could have been selected as decision variable. A tempting one is planned generation at a plant. The planned generation would determine the stored water used. Whether the planned generation was actually generated would depend on the actual flow that came and the operating restrictions. The planned generation was not used as the decision variable because storage use is quite convenient, and perhaps a little more conservative in that the amount of reserve being committed is directly under control.

A source of supplemental energy is assumed. It is taken to include the dropping of the interruptible load. In other words, as the water supply becomes tight and interruptible customers are cut off, we choose to think of the load as unchanged but a block of supplemental generation equal to the dropped load as added. The effect is the same. Beyond the interruptible load, the supplemental energy may be thought of as including steam generation, exchange with other areas, miscellaneous sources, and, if necessary, curtailment.

The performance of the system is to be judged by the cost of the supplemental energy. This cost may or may not be dollars that someone is paying, but is designed as an index to tell whether the results for one way of operating are better than

for another. The cost as a function of supplemental energy might behave as follows: it could have a constant slope in the range where the interruptible load is being cut off. This cost would represent the revenue lost by the dropping of the interruptible load. Thereafter the slope would increase slowly as steam energy was added to the generation, the most economical plant first; next, second and so on. After the steam and miscellaneous sources are exhausted, the cost of supplemental energy presumably rises very rapidly. This would represent the loss to the region of having a power shortage. It might not represent much financial loss to the management of the hydro system but there would be strong reasons for avoiding drastic shortages and it is assumed that these reasons could be put in some approximate quantitative form. It would be desirable to avoid all shortages but nature decides this so that it is only possible to operate with a given risk of shortage. The only way to decide what risk can be tolerated is to have a quantitative estimate of the consequences of shortage.

A power series was chosen to represent the cost of supplemental energy. The constant term, representing fixed costs, can have no effect on operation and is omitted. For simplicity, only first and second power terms are considered. The cost function as specified requires no rigid division between dropping the

interruptible load, adding steam generation, and curtailment. It turned out in making calculations that considerably more supplemental generation was sometimes required than would be available if the resources of the Northwest were scaled down into a system the size of the model. The implication is that the load selected is somewhat large for comparison with real operation of Grand Coulee.

The historical record of the flow of the Columbia River at Grand Coulee is used for deducing probability distributions of flow and for simulating system operation. Records from the U.S. Geological Survey, Bureau of Reclamation, and BPA for the daily flow in the years 1914-53 were made into rough weekly averages. These flows have drawbacks. In recent years an increasing amount of upstream storage has been developed and put into use. Winter releases of water from storage materially increased the flow at Grand Coulee at least during 1952-53 and perhaps for some earlier years. The task of finding out what upstream storage releases were made and separating out the flow which corresponds to those of earlier years was not undertaken. The result is that the interesting year 1952-53 is rather ordinary in the set of flows used.

The release of the upstream storage in winter adds enough water

to the flow at Grand Coulee to make the current 40 Kcfs minimum on discharge a relatively innocuous restriction. This is not true for the flows used in this study. In operating with the low flow years of record, the reservoir is forced merely to dole out the minimum discharge until the water is gone. This may be legitimate but it tells little about how a particular scheme of water use reacts under the important situation of low flow. Therefore, to produce more interesting and realistic results, the minimum discharge was reduced to 20 Kcfs.

We wish to characterize the river flows by probability densities. Future river flow is not completely predictable nor is it completely random. A river will keep flowing for a long time even if there is no rain. It is fed from water in the ground. Rain recharges the ground water and also contributes to river flow by direct overland runoff. Direct runoff and ground water depletion are distinct enough and measurable enough properties of a basin to be very useful in river flow prediction. A basin as big as that of the Columbia River, however, must be broken up into sub-basins for good results. Snowmelt has different characteristics from rainfall. In particular it is sensitive to temperature.

River flow probability densities in consecutive time intervals

are clearly not independent. The shorter the intervals, the more correlated the flows. As a first approximation the flows are considered to have simple Markov probability densities. That is, the probability density of flow in one interval is determined by the flow of the preceding interval. Prediction is involved in the sense that knowledge of the last interval's flow greatly cuts down the range of sensible possibilities for this interval's flow. Prediction is not involved in the sense that the exact flow is guessed. The probability density tells the range of possibilities that must be considered.

The following notation is introduced to describe the model:

- $i$  = index of time intervals.
- $t$  = length of a time interval.
- $d_i$  = volume of discharge through turbines in the  $i$ th interval.
- $D_{max}$  = maximum volume of discharge in an interval.
- $D_{min}$  = minimum volume of discharge in an interval.
- $V$  = volume of usable storage of the reservoir.
- $H$  = head corresponding to  $V$ .
- $H_0$  = additional head of plant.
- $h_i$  = head due to volume in usable storage at beginning of  $i$ th interval.
- $A$  = area of reservoir.

$v_i = Ah_i$  = volume in storage at beginning of  $i$ th interval.

$e_i$  = hydro energy generated in  $i$ th interval.

$L$  = energy load in an interval.

$g_i$  = supplemental energy generated in  $i$ th interval.

$C_i$  = cost of supplemental energy in  $i$ th interval.

The model is summarized:

$$V = 2600 \text{ Kcfs days.}$$

$$H = 80 \text{ ft.}$$

$$H_0 = 264 \text{ ft.}$$

$$D_{\max} = (70 \text{ Kcfs}) \times t.$$

$$D_{\min} = (20 \text{ Kcfs}) \times t.$$

$$e_i = p d_i \left( H_0 + \frac{h_i + h_{i+1}}{2} \right)$$

where  $p = 0.074$  megawatts/Kcfs ft.

$$L = p D_{\max} (H_0 + H).$$

$$g_i = L - e_i.$$

$$C_i = a_1 g_i + a_2 g_i^2.$$

In speaking about system operation, we often discuss the volume of water used and the amount of energy generated. However, the volume and the energy as numbers are not meaningful until the length of time involved is specified. Since the time is not the same in all discussions, it is often useful to speak of



the average flow rate in Kcfs and the average power in megawatts (MW) when comparative numbers are given.

#### IV. THEORY

In this section various approaches to the problem of water use in hydro systems are discussed. Thereafter the method used for the calculations of this paper is developed and applied to the model of the previous section.

##### Current Methods

Reservoirs in the Northwest and elsewhere are often operated according to rule curves. A rule curve is a graph which specifies the amount of water in a reservoir as a function of time throughout the year. They may be drawn up in such a way that, if the lowest flows in the historical record recur, the reservoirs will be operated so that the firm load can be met. In fact it is on the basis of such an analysis that firm load contracts are made. For the Columbia River the lowest winter flow on record is for 1936-37.

If the lowest flow on record does not recur, the rule curve based on it may not provide best use of the reservoir water. In the Northwest, the Power Pool makes median year rule curves as well as critical year rule curves. Neither of the flows used in making these studies will occur exactly as assumed, but the curves can be used as an operating guide. In addition

to taking these curves into account, BPA makes studies throughout the drawdown season based on the possibility that current flows will taper off toward critical flows according to the known hydrologic behavior of the river basin. Rule curves are then made which spread any energy surplus or deficiency uniformly over the rest of the drawdown season.

The determination of rule curves in the above manner is concrete and computationally not too difficult. Their use has been shown by experience to lead to good operation. However, since the flow is never exactly the same as that used in making the rule curves, most final decisions contain substantial elements of managerial judgment. Often final decisions may properly subordinate that which earlier was considered best water use to some pressing current problem.

The big advantage of a rule curve based on the lowest flow in the historical record is that the firm load is protected in case of recurrence of that flow. Small wonder, for that is how the firm load is defined. This did not prevent curtailment of firm power in the Northwest in 1952-53, however, when the winter run-off was considerably greater than the worst of record. This seeming paradox was the result of autumn flows below the historical worst, alleviated by heavy rain in January. No

criticism of the 1952 curtailment is intended. The only point is that the use of any particular fixed flow for determining rule curves and firm power is necessarily arbitrary and tends to give a misleading concreteness to a system operation which is really intimately involved with flow probabilities.

Basing the firm load and rule curves on the 1936-37 low flow does not imply that that firm load can always be met. What it should imply is that based on the best available analysis of stream flows and the objectives of the hydro system, operation based on 1936-37 rule curves will be serve these objectives in the long run.

Viewed this way, however, finding best operation is a rather complicated expected value problem involving the probabilities of stream flow. It is felt that methods for handling such problems have not been widely known in the hydro industry. With the help of machine computation it should be possible to bring river flow probabilities into decisions of storage water use in a quantitative way.

#### Other Methods

Cypser (7) has proposed and done some computing with a calculus of variations scheme for determining rule curves of interdependent

reservoirs when the river flow is known in advance. Johannessen (8), in a study of transmission losses, has used this method with a differential analyzer and has determined some rule curves for a specific case. However, these are fixed flow studies and so in an important respect are like those currently made and discussed above.

Lane (9) in a 1944 TVA report discusses reservoir operation taking into account flow probabilities. He does not give many details as to the theory behind the work and how he performs calculations. It appears, however, that certain computational decisions are aided by rules of thumb considered consistent with good operation. The paper has qualitative discussions of some important hydrologic and storage use problems involved in hydro operation.

The Monte Carlo method may be applied to the problem. Using the historical flows or synthetic flows made up using random numbers, the system may be operated on paper according to different rules of water use and the best selected. For instance the rule curve which minimizes the average cost of system operation could be found. This method is conceptually simple and good for finding quick rough answers. However, it is believed to require much more computation than the

method used below, if the same accuracy of answers is required.

Simon and Holt (10) have studied the inventory problem from the point of view of servomechanism theory. Work of this sort might be useful in the water use problem. There are enough non-linearities which seem essential to the problem so that analytic work would probably be difficult. Another possibility is the simulation of a hydro system on an analog computer. It is likely that rather complex systems could be set up. Different ways of operating them could be tried experimentally. However, if changes in energy of the order of fractions of a per cent are considered important, accuracy might be a problem.

#### Expected Value Method

It is assumed that there is some objective way of judging the performance of a system. The index of performance is called the cost of operating the system and operating decisions about using stored water are made so as to minimize the expected cost of future operation.

For purposes of the problem, time is broken up into intervals of some arbitrary length, two weeks, for example. At the beginning of each a decision is to be made about storage use in that interval. The decision is to be made knowing the current

reservoir head and the preceding two weeks flow. It is to be made in such a way that the expected cost of the rest of the year's operation is a minimum.

The mathematical treatment begins by setting up an expression for expected value of the cost. Manipulation of the expression shows that the only information lacking for finding optimum operation in one interval is an expected cost function for the succeeding interval. By assuming or knowing this function for the final interval, one may work backward to any interval. At each step functions are found which tell the proper use of stored water under the conditions that may arise in that interval.

Thus the theory reduces to setting up a considerable number of expected value functions and finding their minima. In this it is similar to many other stochastic problems. With proper labeling of variables the storage use problem may be considered an inventory problem and a special case of the rather general work of Dvoretzky, Kiefer and Wolfowitz (4). It differs from the usual business inventory problem in that the input, not the output, is the random variable. Furthermore reservoirs, unlike most warehouses, have the property that the more nearly they are filled, the more valuable is a unit in them because

the head is higher.

The hydro problem also fits into the general framework of optimization in multistage processes discussed by Bellman (5) in his papers on dynamic programming theory. Equation (4.16) below is a functional equation somewhat like those he discusses. Bellman points out that working with functional equations often has computational advantages over classical techniques for optimal allocation problems.

#### Setting Up The Equations

The expected cost expressions for determining best storage water use in a one reservoir hydroelectric system are set up below. The extension to multiple reservoirs increases the number of variables and the computational difficulty. However, the ideas involved do not change and so to simplify the expressions, the one reservoir case is treated. The first part is kept general enough to include the one commodity inventory problem.

- Let:
- $i = 0, 1, 2, \dots, N$  ; an index of time intervals.
  - $v_i$  = amount of commodity in storage at the beginning of the  $i$ th interval. The commodity is water in the hydro problem.
  - $x_i$  = random variable. This is the volume of river flow in the  $i$ th interval for the hydro problem. It would be amount of sales in a business inventory problem.



$s_i$  = decision variable. In the hydro problem as set up here, this will be the volume of stored water planned for use in the  $i$ th interval. In a business inventory case, this could be the amount of the commodity ordered.

$f(x_j, \dots, x_i | x_{i-1}, \dots, x_k)$  = joint probability density of  $x_j, x_{j-1}, \dots, x_i$  given the values of  $x_{i-1}, \dots, x_k$ .

$C_i$  = cost of operating the system in the  $i$ th interval as a function of any pertinent variables of  $i$ th and preceding intervals, in particular,  $x_i, v_i, s_i$ .

$K_i = \sum_{j=i}^N c_j$  = total cost from  $i$ th interval to end of time period considered.

$E_i$  = expected value of  $K_i$ . It can be a function of various variables whose values are known at the beginning of the  $i$ th interval.

First we make an observation about conditional probability densities. The definition (11) of the conditional probability of an event A given the occurrence of an event H, denoted  $P_r \{A | H\}$ , is stated in terms of the joint probability of the events, denoted  $P_r \{AH\}$ , and the probability of the event H, denoted  $P_r \{H\}$ :

$$P_r \{AH\} = P_r \{A | H\} P_r \{H\} \quad (4.1)$$

From this it follows that:

$$f(x_N, \dots, x_{i+1}, x_i | x_{i-1}, \dots, x_0) dx_N \dots dx_{i+1} dx_i = \quad (4.2)$$

$$f(x_N, \dots, x_{i+1} | x_i, x_{i-1}, \dots, x_0) dx_N \dots dx_{i+1} f(x_i | x_{i-1}, \dots, x_0) dx_i$$

Now by the definition of an expected value:

$$E_i = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} K_i f(x_N, \dots, x_{i+1}, x_i | x_{i-1}, \dots, x_0) dx_N \dots dx_i \quad (4.3)$$

where it is assumed that there is a way to pick  $s_j$  in each interval. Expanding:

$$E_i = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} [C_i + K_{i+1}] f(x_N, \dots, x_{i+1} | x_i, x_{i-1}, \dots, x_0) dx_N \dots dx_{i+1} f(x_i | x_{i-1}, \dots, x_0) dx_i \quad (4.4)$$

Since  $C_i$  is a function only of variables in the  $i^{\text{th}}$  interval or earlier,

$$E_i = \int_{-\infty}^{\infty} [C_i + E_{i+1}] f(x_i | x_{i-1}, \dots, x_0) dx_i \quad (4.5)$$

This is a recurrence equation for  $E_i$ . If we know  $E_N$ , we can work backward and find any  $E_i$ .

It has been assumed that there is some way to pick the decision variable  $s_i$ . This quantity is controllable and, with the random variable  $x_i$  and the conditions in the system at the start of the interval, determines the outcome of the interval, namely the cost and the conditions which start the next interval. The more variables that help determine  $s_i$ , presumably the lower the cost,

but on the other hand, the greater the computational complexity in finding  $s_i$ . Almost certainly  $s_i$  should depend on the amount of the commodity in stock, very likely on preceding values of the random variable, because they are harbingers of the future, and perhaps on other variables past or present. The future values of pertinent variables, being unknown when the value of  $s_i$  is decided upon, cannot help. If, for simplicity, we restrict ourselves to making decisions on the basis of the current amount in stock and preceding values of the random variable, we can write

$$s_i = s_i (v_i; x_{i-1}, x_{i-2}, \dots, x_0) \quad (4.6)$$

The outcome of system operation during an interval is the cost for the interval and the amount of stock for starting the next interval. It is assumed that these are determined by starting stock, decision variable, random variable, and by constant system parameters. Thus, the functions

$$\begin{aligned} v_{i+1} &= v_{i+1} (v_i, x_i, s_i) \\ C_i &= C_i (v_i, x_i, s_i) \end{aligned} \quad (4.7)$$

are known.

Conceivably past values of some of these variables could enter but this will not be considered here.

Rewriting equation (4.5) to exhibit functional dependence,

$$E_i(v_i; x_{i-1}, x_{i-2}, \dots, x_0; s_i) = \int_{-\infty}^{\infty} \{c_i(v_i, x_i, s_i) + E_{i+1}(v_{i+1}; x_i, x_{i-1}, \dots, x_0)\} f(x_i | x_{i-1}, \dots, x_0) dx_i \quad (4.8)$$

Assuming that we know  $E_{i+1}$ , we pick

$$s_i = s_i(v_i; x_{i-1}, x_{i-2}, \dots, x_0)$$

such that  $E_i$ , the expected cost for the rest of the time period, is minimum. This in turn determines  $E_i(v_i; x_{i-1}, x_{i-2}, \dots, x_0)$ , which is what is needed to find  $s_{i-1}$  and  $E_{i-1}$ . Knowing  $E_N$  we can thus work backward to find any  $s_i$ . By taking  $N$  far enough into the future,  $E_N$  can usually be considered constant or zero.

#### Hydro Case

The expressions above are now applied to the simple hydro model presented earlier.

River flow is the random variable  $x_1$ . Since it is assumed to have a simple Markov probability density,

$$f(x_i | x_{i-1}, \dots, x_0) = f(x_i | x_{i-1}) \quad (4.9)$$

These functions are to be determined by analysis of the historical flows.

The quantity  $s_i$  is the amount of storage water allocated for

possible use in the  $i^{\text{th}}$  interval at the beginning of that interval.

Its value is limited by the size of the reservoir:

$$v_i - V \leq s_i \leq v_i \quad (4.10)$$

The discharge through the turbines will be

$$d_i = x_i + s_i \quad (4.11)$$

unless this would violate some operating restriction. Because inequalities are involved, the complete functions  $d_i$  and  $v_{i+1}$  are best specified by a table:

Table 1. The functions  $d_i = d_i(v_i, x_i, s_i)$  and  $v_{i+1} = v_{i+1}(v_i, x_i, s_i)$ .

$x_i$	$d_i$	$v_{i+1}$
$V - v_i + D_{\max} \leq x_i \leq \infty$	$D_{\max}$	$V$
$D_{\max} - s_i \leq x_i \leq V - v_i + D_{\max}$	$D_{\max}$	$V + x_i - D_{\max}$
$D_{\min} - s_i \leq x_i \leq D_{\max} - s_i$	$x_i + s_i$	$v_i - s_i$
$D_{\min} - v_i \leq x_i \leq D_{\min} - s_i$	$D_{\min}$	$v_i + x_i - D_{\min}$
$-\infty \leq x_i \leq D_{\min} - v_i$	$v_i$	$0$

Knowing  $d_i$ ,  $v_i$ ,  $v_{i+1}$ , we can calculate the energy generated in the interval

$$e_i = p d_i \left[ H_0 + \frac{v_i + v_{i+1}}{2A} \right] \quad (4.12)$$

The supplemental energy needed to make up the load  $L$  is

$$g_i = L - e_i \quad (4.13)$$

The cost of the supplemental energy is

$$C_i = a_1 g_i + a_2 g_i^2 \quad (4.14)$$

Now, having specified the functions

$$\begin{aligned} v_{i+1} &= v_{i+1}(v_i, x_i, s_i) \\ C_i &= C_i(v_i, x_i, s_i) \end{aligned} \quad (4.15)$$

we must, for each  $i$ , find the function

$$s_i = s_i(v_i, x_{i-1})$$

such that

$$\begin{aligned} E_i(v_i, x_{i-1}; s_i) = & \\ & \int_{-\infty}^{\infty} \left\{ C_i(v_i, x_i, s_i) + E_{i+1}(v_{i+1}, x_i) \right\} f(x_i | x_{i-1}) dx_i \end{aligned} \quad (4.16)$$

is a minimum for each pair  $(v_i, x_{i-1})$ , and therefrom determine

$$E_i(v_i, x_{i-1})$$

for use in the preceding time interval. Since in our model the summer flow always fills the reservoir and completely meets the load, we may set  $E_N = 0$  where N, the final interval, is taken in summer.

## V. NUMERICAL METHODS

Calculations to find best use of stored water were performed on an automatic digital computer. The numerical techniques, therefore, stress generality. It is undesirable to have much manual intervention between the introduction of the data and program to the machine and the output of final results. Any change in the way numbers are processed which depends on how the numbers "look" must be expressed quantitatively and programmed into the calculation.

The computation consists of three parts: the calculation of river flow probabilities; the calculation of tables of best storage water use; and the calculation of actual system behavior with the historical flows. Some of the techniques and problems of the first two parts will be discussed. The last calculation is straightforward. At the end a short discussion of the programming is given.

### Probabilities

Standard techniques were used in calculating the conditional probabilities. Beyond recognizing that the distributions were skew and assuming that the flows in successive intervals were simple Markov, no detailed knowledge of the hydrology of the



river basin was introduced.

In practice the probabilities were not calculated as density functions but as conditional probability tables. For each interval,  $i$ , there was a table whose elements were of the following form:

$${}^{m_n} p_i = \text{Prob} \left\{ m_x - \frac{\Delta x}{2} \leq x_i \leq m_x + \frac{\Delta x}{2} \mid x_{i-1} = n_x \right\} \quad (5.1)$$

In words,  ${}^{m_n} p_i$  is the probability that, in the  $i$ th interval, the value of the flow,  $x_i$ , will lie in the range  $\Delta x$  about some value  $m_x$ , given the fact that the flow,  $x_{i-1}$ , in the preceding interval was  $n_x$ . For discussion purposes, we shall assume that the time interval is a month.

It is necessary to find the correlation between the flows in successive months. However, if we consider a frequency curve made from the forty years of data for, say, January, we find the distribution of flows is very skew. Furthermore the mean flow for January is different from that for December. It is therefore necessary to transform the data to get reliable correlations. The flows are put into unit normal form by an empirical transformation which we can represent by

$$u_i = u_i(x_i) \quad (5.2)$$

If  $i$  is January, any January flow,  $x_1$ , has a unit normal counterpart  $u_1$ . If there are  $N$  years of record, and the cumulative frequency curve for the January flows is denoted  $F_1(x_1)$ , then

$u_i$  is defined by:

$$\int_{-\infty}^{u_i} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \frac{F_i(x_i)}{N} \quad (5.3)$$

The expression at the left is the cumulative distribution of the normal probability function and is tabulated (12). In practice the raw frequency curve is smoothed somewhat in forming  $F_i(x_i)$ .

The correlation between months  $i$  and  $i-1$  is now:

$$h_{i,i-1} = \frac{\sum u_i u_{i-1}}{\sum u_{i-1}^2} \quad (5.4)$$

where the summation is over the years of record. The flows are assumed to be simple Markov in a manner defined by

$$u_i = h_{i,i-1} u_{i-1} + \epsilon_i \quad (5.5)$$

where  $\epsilon_i$  is a random variable assumed to be uncorrelated with  $u_{i-1}$ .

For each year of record we may calculate values of  $\epsilon_i$  from the above expression and thereby obtain sets of  $\epsilon_i$ . These will have some cumulative frequency distributions  $G_i(\epsilon_i)$ . Physically, we have taken  $u_{i-1}$ , last month's flow in unit normal form, predicted this month's flow to be  $h_{i,i-1} u_{i-1}$ , and found, by doing this for all years of record, the cumulative frequency distribution of the errors,  $\epsilon_i$ , to be  $G_i(\epsilon_i)$ . The distribution is skew because, knowing last month's flow, this month has a fairly rigid low flow limit but a wide range of high flow possibilities.

Having the functions  $u_i(x_i)$  and  $G_i(\epsilon_i)$  we may calculate the

conditional probabilities  ${}^{mn}p_i$ . The possible range of flows is divided up into intervals centered about a series of increasing flows,  ${}^0x, {}^1x, {}^2x, \dots, {}^jx, \dots$ . Because small flows are more important than large flows, the spacing is made to increase as  $j$  increases.

We let

$${}^jx = (\alpha)^j {}^0x, \quad \alpha > 1 \quad (5.6)$$

Now

$${}^{mn}P_i = \frac{1}{N} \left\{ G_i({}^{mn}\epsilon_i^+) - G_i({}^{mn}\epsilon_i^-) \right\} \quad (5.7)$$

where

$${}^{mn}\epsilon_i^+ = u_i({}^{mn}x + \frac{\Delta x}{2}) - h_{i-1} \dots u_{i-1}({}^{mn}x) \quad (5.8)$$

$${}^{mn}\epsilon_i^- = u_i({}^{mn}x - \frac{\Delta x}{2}) - h_{i-1} \dots u_{i-1}({}^{mn}x)$$

$$\Delta x = 2 \left( \frac{\alpha-1}{\alpha+1} \right) {}^{mn}x$$

The probability was taken zero for having any flows greater than or less than the extremes of the historical flows for that month. This was done for simplicity. There is still a non zero probability of obtaining a year drier than any on record. The reason is that the driest year of record does not contain the lowest flows on record for every month. The same situation holds for large flows.

After these probability tables had been calculated and used, two changes seemed desirable and were made. The first change concerned the mean flow. It was discovered that the mean flow for a month as calculated from the probabilities differed at times up to 9%

from the mean of the historical flows. Also it was found that the mean of next month's flow as a function of this month's flow did not vary smoothly. Neither fact is surprising, considering the small number of years of record and the empirical transformations involved in the calculation of the probabilities. However, since the expected energy in winter is closely proportional to the mean flow, some functions in later calculations became not smooth. In order to be sure that the reason lay in the probabilities and not in the later calculations, the flow means were smoothed. The mean of next month's flow as a function of this month's flow was fit to a cubic equation by least squares. The mean flow for a month was adjusted to be the same as that of the historical record. The changes in mean were made by shifting the probability distributions along the x axis.

Examples of the results of probability computations are shown in Fig. 2 and Table 2. Fig. 2a shows the mean flow by two week intervals from the historical record. Fig. 2b shows the correlations for unit normal flows between successive intervals through the year for two week and four week intervals. Table 2 shows a set of conditional probabilities. For this case,  $\sigma_x = 17.0$  Kcfs and  $\alpha = 1.181$ .

The other change in the probabilities was one of smoothing the probability densities. One interpretation of a column of a probability table is the histogram shown in solid line in Fig. 3. However, the dashed line was taken as the probability density in the calculations.

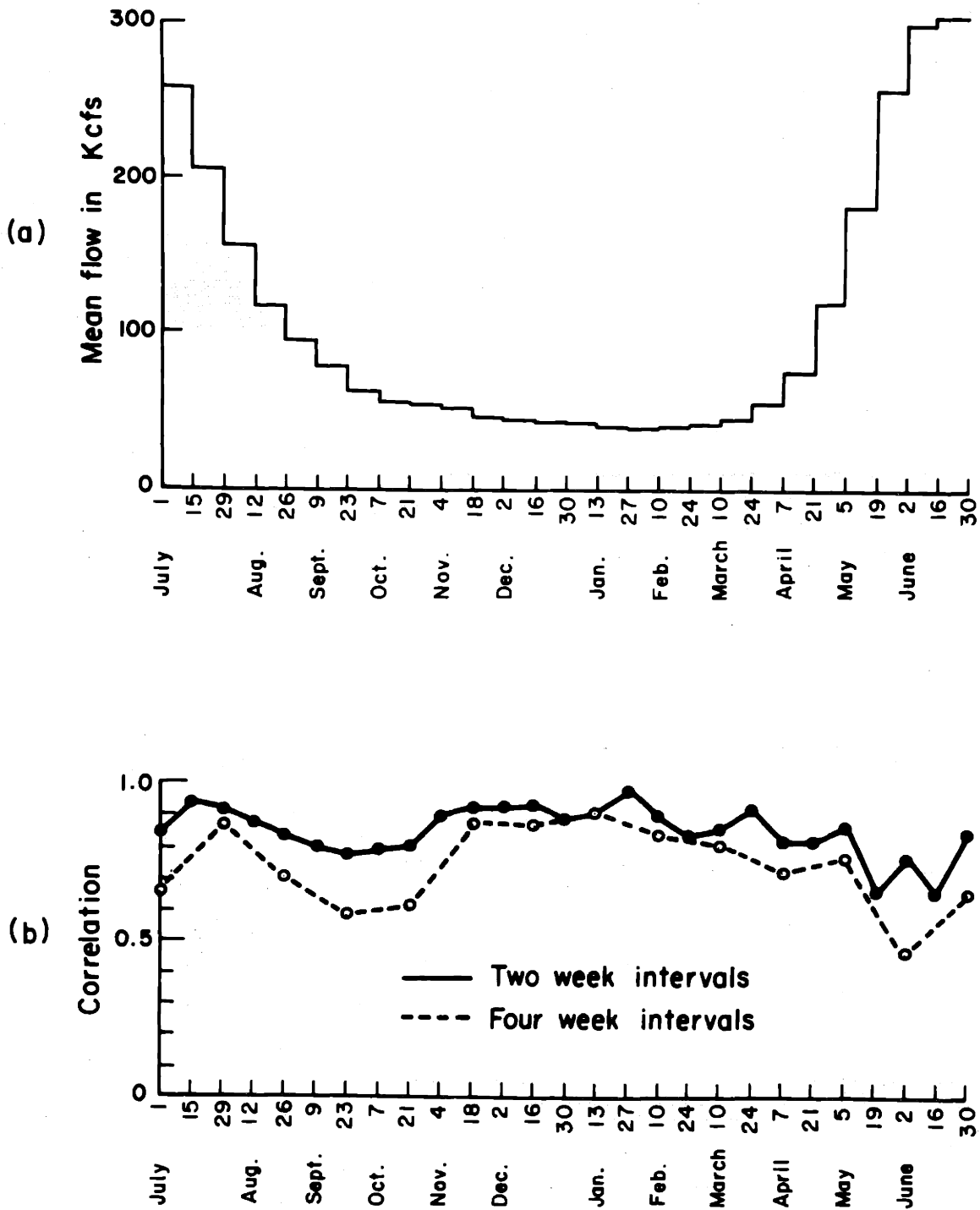


Fig. 2. Characteristics of the Columbia River at Grand Coulee.  
a) Mean flow, 1914-1953, by two week intervals.  
b) Correlations between unit normal flows in successive intervals.

$x_i$ in Kcfs	106	90	76	64	54	64	76	90	106
	.024	.040	.038	.179	.101	.179	.181		
	.026	.011	.038	.168	.077	.168	.169		
	.096	.179	.160	.273	.203	.273	.275		
.051	.433	.290	.331	.359	.583	.359	.354		
.077	.367	.479	.312	.022	.036	.022	.021		
.244	.055		.122						
.628	.46								
39		54	64	90	76	90	106		
33									

Table 2. A conditional probability table.

The table gives the probabilities  $nm p_i$  for  $i =$  the interval from October 21 to November 4. The following illustration shows how the table is read: If the river flow,  $x_{i-1}$ , has a value of 64 Kcfs during the interval October 7 to October 21, then the probability is .122 that the river flow,  $x_i$ , will have a value between 42 Kcfs and 50 Kcfs during the interval October 21 to November 4.

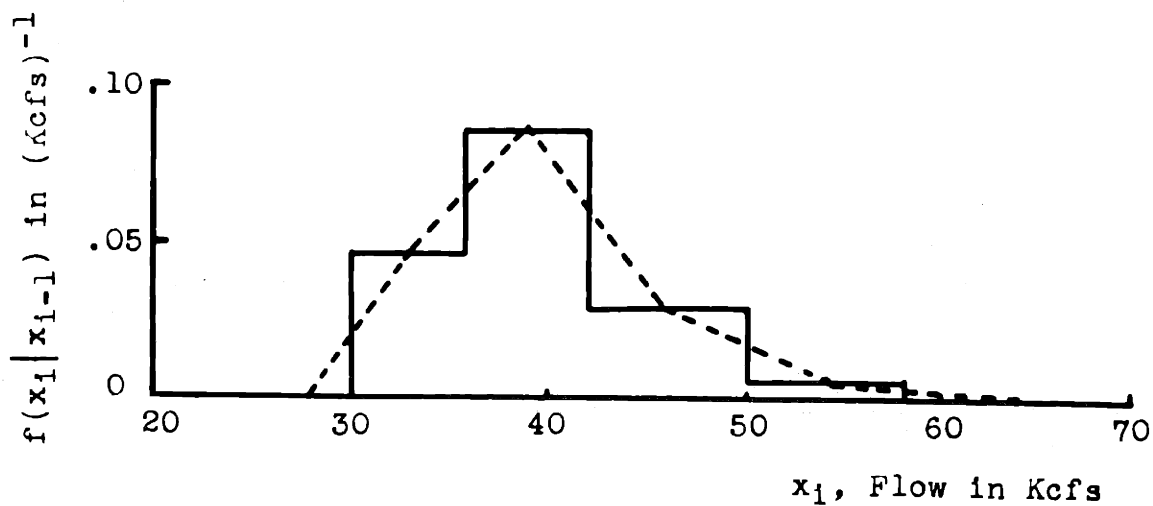


Fig. 3. An example of a conditional probability density,  $f(x_i | x_{i-1})$ . The solid curve is a rectangular probability density representing a column in the probability tables. The dashed curve is the continuous density used. For the case shown,  $x_{i-1} = 39$  Kcfs, and  $i =$  the interval from Dec. 30 to Jan. 13.

The ordinates at the end points of its straight line segments are given by:

$${}^m f = \frac{2 \alpha}{\alpha^2 - 1} \frac{{}^m P}{{}^m x} \quad (5.9)$$

In this expression the  $n$  and  $i$  subscripts have been suppressed. The mean and area under the curve are unchanged. The new probability density is continuous, which property is desirable.

In retrospect, it might have been sensible to have forced the probability densities into a smooth analytic form. This would make it somewhat easier to program the optimization calculation. Any lack of smoothness in functions calculated there could then be tracked down in that program.

The idiosyncracies of the probabilities calculated by the methods above are presumably caused by the historical data. One might expect that such idiosyncracies would therefore be an advantage when evaluating system operation with the historical flows. However, such smoothing as was done seemed, if anything, to improve the results of later calculations.

#### Optimization Calculation

Using equations (4.16) of section IV, the calculation proceeds as follows: Assume that the expected cost function  $E_{i+1}(V_{i+1}, X_i)$  is stored in the machine as a two dimensional table. For a given



pair  $(v_i, x_{i-1})$  the expected cost function

$$E_i(v_i, x_{i-1}; s_i) = \int_{-\infty}^{\infty} \{ C_i(v_i, x_i, s_i) + E_{i+1}(v_{i+1}, x_i) \} f(x_i | x_{i-1}) dx_i \quad (5.10)$$

is calculated for a number of  $s_i$  and the smallest  $E_i$  chosen. The corresponding value of  $s_i$  is stored in a table,  $s_i(v_i, x_{i-1})$ . This table represents the output of the calculation for the  $i$ th interval, since it specifies the use of stored water which will minimize the expected cost under any combination of conditions of  $v_i$  and  $x_{i-1}$  that may arise in the  $i$ th interval. The value of  $E_i$  found at the same time is stored in a table of  $E_i(v_i, x_{i-1})$  for use in the calculations for the preceding interval. When values of tabulated functions are required which are not directly in the tables, linear interpolation is used.

Thus, the basic task is the simple procedure of calculating considerable numbers of functions and picking out their minima. This would be a formidable job by hand but, on the other hand, is the sort of thing machines are built for. The remainder of this section will deal with some problems that arise in computing the functions.

In practice the integral of equation (5.10) is replaced by a sum using the trapezoidal rule. To simplify discussion, we shall assume we are working with a particular  $(v_i, x_{i-1})$  pair and so will

not indicate functional dependence on them. Let us also drop the subscripts denoting the interval and replace the quantity in the braces by the single letter I. Then equation (5.10) is

$$E(s) = \int_{-\infty}^{\infty} I(x, s) f(x) dx \quad (5.11)$$

and the trapezoidal rule approximation is

$$\tilde{E}(s) = \sum_{m=1}^M \frac{1}{2} [I(x^{m-1}, s) f^{m-1} + I(x^m, s) f^m] [x^m - x^{m-1}] \quad (5.12)$$

where there are  $(M+1)$  discrete values of  $x$  considered in the finite sum. When it is necessary to distinguish between a function and an approximation to it, we shall write a  $(\sim)$  over the latter. Since we can always make the range of  $x$  such that  $f(x)$  goes to zero at the ends, the above can be rewritten

$$\tilde{E}(s) = \sum_{m=1}^{M-1} \frac{1}{2} I(x^m, s) f^m [x^{m+1} - x^{m-1}] \quad (5.13)$$

which, by equation (5.9) can be written

$$\tilde{E}(s) = \sum_{m=1}^{M-1} I(x^m, s) f^m p \quad (5.14)$$

We seek the  $s$  such that  $E(s)$  is minimum. This will come when

$$\frac{dE}{ds} = 0 \tag{5.15}$$

unless, as often happens,  $E$  has its smallest value when  $s$  is at an extreme end of its limited range. Such a case is sometimes called corner minimum.

If  $E(s)$  is calculated for a number of  $s$  and the smallest one is picked, it would seem that we need not worry about the slope. Provided that  $E(s)$  is a fairly slowly varying function, a plausible way to search for the minimum is to sample  $E(s)$  with a coarse grid in  $s$  and then search more carefully about the best value. However, calculating  $E$  for values of  $s$  close together is tantamount to approximating the slope there by finite differences. Consequently, we can investigate how closely this technique finds the minimum of  $E(s)$  by seeing what sort of slope the calculated  $E(s)$  has.

It is next observed that

$$\frac{dI(x,s)}{ds} = 0 \tag{5.16}$$

unless  $x_b \leq x \leq x_t$  where

$$\begin{aligned} x_b &= D_{min} - s \\ x_t &= D_{max} - s \end{aligned} \tag{5.17}$$

This follows from the model. Physically it happens because if you try to use more storage water than a certain amount, no change in outcome occurs. This is because the discharge is limited to  $D_{max}$ . A similar situation occurs in trying to use too little water.

Now we discover that  $\frac{dE}{ds}$  resulting from equation (5.11)

$$\frac{dE}{ds} = \int_{x_b}^{x_t} \frac{dI(x,s)}{ds} f(x) dx \quad (5.18)$$

is not well approximated by the  $\frac{d\tilde{E}}{ds}$  implied by equation (5.14):

$$\frac{d\tilde{E}}{ds} = \sum_{\substack{m \\ x > x_b \\ m \\ x < x_t}} \frac{dI(x,s)}{ds} m p \quad (5.19)$$

The reason is that, as  $s$  changes, so do  $x_b$  and  $x_t$ . Expression (5.18) changes smoothly. Expression (5.19) undergoes a discontinuous change of magnitude  $(\frac{dI}{ds})_{mp}$  whenever  $x_b$  or  $x_t$  crosses an  $m_x$ . For the mesh of  $x$  used in the calculation here, the term added or dropped may be one of only four or five in the sum. Thus, delicate detection of minima is not possible.

The discontinuity can be reduced by decreasing the mesh size for  $x$  since this reduces the size of the individual  $m_p$  and thereby the discontinuity. However, it also increases the time required for the calculation.

A different method was used. The derivative was approximated from a difference of the function. However, the two values of the function involved were calculated in such a way that the difference was a better approximation than found above.

Let

$$\begin{aligned}
 \tilde{E}(s) = & \sum_{m \neq p, q} \gamma_m \left[ I(x_{m-1}, s)^{m-1} f + I(x_m, s)^m f \right] \left[ x_m - x_{m-1} \right] \\
 & + \frac{1}{2} \left[ I(x_b, s - \frac{\Delta s}{2}) f(x_b) + I(x_b, s)^{p-1} f \right] \left[ x_b - x_{p-1} \right] \\
 & + \frac{1}{2} \left[ I(x_b, s + \frac{\Delta s}{2}) f(x_b) + I(x_p, s)^p f \right] \left[ x_p - x_b \right] \\
 & + \frac{1}{2} \left[ I(x_t, s - \frac{\Delta s}{2}) f(x_t) + I(x_t, s)^{q-1} f \right] \left[ x_t - x_{q-1} \right] \\
 & + \frac{1}{2} \left[ I(x_t, s + \frac{\Delta s}{2}) f(x_t) + I(x_t, s)^q f \right] \left[ x_t - x_t \right]
 \end{aligned} \tag{5.20}$$

This is again a trapezoidal rule approximation to equation (5.11). However, the number of intervals has been increased by taking extra ordinates at  $x_b$  and  $x_t$ . Also, at these points the exact  $s$  has been replaced by  $(s - \frac{\Delta s}{2})$  on the low side and  $(s + \frac{\Delta s}{2})$  on the high side.

Now let

$$\left( \frac{d\tilde{E}}{ds} \right) = \frac{1}{\Delta s} \left[ \tilde{E}(s + \frac{\Delta s}{2}) - \tilde{E}(s - \frac{\Delta s}{2}) \right] \tag{5.21}$$

where the same  $x_b$  and  $x_t$  are used for both  $E$ 's. Then let

$$\left(\frac{\tilde{d}E}{ds}\right) = \frac{I(s + \frac{\Delta s}{2}) - I(s - \frac{\Delta s}{2})}{\Delta s}$$

or in certain cases: 
$$= \frac{I(s) - I(s - \Delta s)}{\Delta s} \tag{5.22}$$

or 
$$= \frac{I(s + \Delta s) - I(s)}{\Delta s}$$

Now using equations (5.16), (5.20), (5.21), and (5.22),

$$\begin{aligned} \left(\frac{\tilde{d}E}{ds}\right) &= \frac{1}{2} \left[ \frac{\tilde{d}I}{ds}(x_b, s) f(x_b) + \frac{\tilde{d}I}{ds}(x_t, s) f(x_t) \right] \underline{[x_t - x_b]} \\ &+ \sum_{\substack{m < q-1 \\ m > p}} \frac{1}{2} \left[ \frac{\tilde{d}I}{ds}(x_m, s) f(x_m) + \frac{\tilde{d}I}{ds}(x_{m+1}, s) f(x_{m+1}) \right] \underline{[x_{m+1} - x_m]} \\ &+ \frac{1}{2} \left[ \frac{\tilde{d}I}{ds}(x_{q-1}, s) f(x_{q-1}) + \frac{\tilde{d}I}{ds}(x_t, s) f(x_t) \right] \underline{[x_t - x_{q-1}]} \end{aligned} \tag{5.23}$$

which is a trapezoidal rule approximation to the equation (5.18). It

is similar to equation (5.19) except for the underlined terms. These

terms make  $\left(\frac{\tilde{d}E}{ds}\right)$  vary smoothly as  $s$  causes  $x_b$  or  $x_t$  to pass an  $m_x$ .

The slope  $\left(\frac{\tilde{d}E}{ds}\right)$  may be used to find minima. The best  $s$  is that for which a slightly smaller value gives a negative slope, a slightly larger value, a positive slope. In case of more than one minimum, including corner minima, the value of  $E$  will tell which is smaller.

The notation above became somewhat involved but the essential point is that considerable care had to be taken to avoid calculating an incorrect discontinuity into the slope of the function  $E(s)$ .

As a final comment in connection with numerical methods, the formal derivative of equation (5.10) could have been taken.

$$\frac{dE_i}{ds_i}(v_i, x_{i-1}; s_i) = \int_{-\infty}^{\infty} \left\{ \frac{dC_i}{ds_i} + \frac{dE_{i+1}}{dv_{i+1}}(v_{i+1}, x_i) \frac{dv_{i+1}}{ds_i} \right\} f(x_i | x_{i-1}) dx_i \quad (5.24)$$

By using this expression we can evaluate the derivative directly instead of by differences. The functions  $\frac{dC_i}{ds_i}$ ,  $\frac{dv_{i+1}}{ds_i}$  can be expressed analytically in our model. Thus assuming  $\frac{dE_{i+1}}{dv_{i+1}}$  is known, we can find  $s_i$  and then, since

$$\frac{dE_i}{dv_i}(v_i, x_{i-1}) = \int_{-\infty}^{\infty} \left\{ \frac{dC_i}{dv_i} + \frac{dE_{i+1}}{dv_{i+1}}(v_{i+1}, x_i) \frac{dv_{i+1}}{dv_i} \right\} f(x_i | x_{i-1}) dx_i \quad (5.25)$$

we can find  $\frac{d E_i}{d v_i}$  and so forth. However, there is an added complication in that

$$\frac{d C_i}{d v_i} = \frac{\partial C_i}{\partial v_i} + \frac{\partial C_i}{\partial s_i} \frac{d s_i}{d v_i} \tag{5.26}$$

$$\frac{d v_{i+1}}{d v_i} = \frac{\partial v_{i+1}}{\partial v_i} + \frac{\partial v_{i+1}}{\partial s_i} \frac{d s_i}{d v_i}$$

although having found  $s_i$ ,  $\frac{d s_i}{d v_i}$  can be found. By assuming a function for  $\frac{d E_N}{d v_N}$  we can work backward as before but this time conducting the calculation in terms of the derivatives instead of the expected cost functions themselves. After all it is the functions  $s_i(v_i, x_{i-1})$  which are the desired output of the calculation. One troublesome situation is the possibility of multiple minima. This could be handled by integrating  $\frac{d E_i}{d s_i}$  over  $s_i$  to find out which was lowest.

The reason for mentioning this method of calculation is that it avoids the differencing found in the methods above and therefore seems to put less stringent requirements on the accuracy required of the machine. Perhaps too, the above expressions would give more



accurate values of  $\frac{dE}{ds}$  than the differencing of  $E(s)$  for the same mesh of dependent variables.

The reasons for not using the method are the multiple minima possibility, the increased complication, and a curiosity to compare the expected cost with the average cost of operating with the historical flows.

#### Programming

The current state of the art of computer programming is such that a program is rather specific to the computer being used. The calculations were made on the Whirlwind I computer at M.I.T. The machine is, by present standards, very fast for handling  $4\frac{1}{2}$  decimal digit fixed point numbers, but it is considerably easier to program when its slower, more accurate floating point interpretative subroutines are used. It was decided at the outset to undertake the more complicated programming and risk some difficulty with accuracy in order to avoid worry over the time required for the calculation. As it turned out, the times involved were on the order of five minutes for calculating the probabilities, five to thirty minutes depending on the mesh for calculating the operating tables, and one to five minutes to operate the system with the historical flows, depending on the amount of computer output desired. These times were in a range which made it fairly easy to debug and experiment with the program.

It was early discovered that although  $4\frac{1}{2}$  decimal digits might be sufficient for accuracy, they did not cover a sufficient range of magnitude to accommodate the numbers encountered. Rather than revert to floating point computation, a device was used which might occasionally be useful elsewhere. It was found that in the most repetitive and therefore most time consuming parts of the program, it was only necessary to pick up numbers from the tables. The putting away occurred much less frequently. By storing not a number but its square root, the range of numbers handled could be increased to 9 decimal digits with negligible loss of accuracy. Reconstructing the number required but squaring the value in the tables. Taking the square root required more time but occurred relatively infrequently.

During the calculation, tables of functions were computed, stored, and then used in further calculations. The time for the whole calculation depends strongly on the size of the tables. It was therefore worth some effort to eliminate unnecessary values.

An example of this is the way the tables depend upon the range of flows. The range of the historical flow of the Columbia River at Grand Coulee is from about 17 to 570 Kcfs but for any given time of year it is much less than this. From the programming point of view it would be easiest to compute and store tables covering the full range, even if the values were useless. However, considerable

time and storage can be saved by making the start and finish of the tables flexible. While this is not too difficult to do, it is the kind of complication which is not undertaken unless, as in this case, it is justified.

The combined length of the final programs used was about 4000 single address instructions, exclusive of tables and subroutines provided by the Digital Computer Laboratory.

## VI. RESULTS OF THE CALCULATIONS

Calculations were made to try out the theory of section IV on the hydro model of section III. Solutions of the problem of storage water use were found for two different cost functions. They take the form of sets of tables for determining reservoir operation. In addition, for comparative purposes, a rule curve was computed by standard methods. However, the tables and rule curve are rather sterile when it comes to understanding how the system would behave under real flows. Consequently, the 39 year record of historical flows was made the input for the system and operation was computed for each of the three cases. The results will be discussed.

### Kinds of Operation

Calculations were made for three cases.

#### Case 1. Minimum expected cost, linear cost function.

The basic model assumes a cost function of the form:

$$C_i = a_1 q_i + a_2 q_i^2 \quad (6.1)$$

where  $g_i$  is the supplemental energy required in the  $i$ th interval. Case 1 corresponds to

$$a_1 = 0$$

(6.2)

The particular value of  $a_1$  makes no difference in determining best operation provided that it is greater than zero.

This is because decisions are made on the basis of relative cost. In the calculation it is only necessary to know whether one decision is less expensive than another, not exactly how much in dollars.

When cost is linear, minimum cost is the same as maximum hydro energy. This follows because cost is then proportional to the total supplemental energy demanded, irrespective of the size of the individual chunks and the way they are spread out in time. Case 1 is therefore not too realistic, and certainly does not apply in the Northwest. However, since it represents maximum expected energy, it is an important reference point.

Case 2. Minimum expected cost, pure quadratic cost function.

This case corresponds to

$$a_1 = 0 \tag{6.3}$$

in equation (6.1). Here too the particular value of  $a_2$  does not affect operation provided it is greater than zero. A quadratic cost function means that the cost of a unit of supplemental energy increases with the amount required. Such an increase usually occurs in practice, either because less efficient steam plants must be used or in extreme situations because load is curtailed. Whereas a real hydro system might have a much more complicated cost function, the pure quadratic assumed for case 2 is in keeping with the simplicity of our model.

Case 3. Rule curve operation.

A rule curve was fashioned for the model from the lowest flow on record, 1936-37. Water will be used from storage down to but not below the rule curve level specified for the particular time of year. The curve is constructed so that for the 1936-37 flow the supplemental energy required will be constant throughout the drawdown season. An exception is made at the beginning and end of the season. At these times rational water use demands somewhat lower supplemental generation. The rule curve is taken as an example of a simple reservoir operation which has been shown by experience to be good.

### Water Use Tables

The output of the optimization calculation is a set of tables of the functions  $s_i(V_i, x_{i-1})$ . They determine the volume of storage water which is to be used in each time interval. Once calculated they need not be recalculated unless the model is changed.

In Fig. 4 a table for a given  $i$  is plotted as a family of curves. Also plotted is the corresponding data for a rule curve. The curves determine operation as follows: If the flow in the preceding interval was 46 Kcfs and if the reservoir at the start of the interval contains a volume of water 0.80  $V$ , then  $s_i$ , the volume of storage water planned for use is 0.077  $V$ . This represents an average flow of about 14.3 Kcfs over the two week interval. As the interval unfolds, the water is added to the natural river flow to provide the discharge for power. If the sum of  $s_i$  and the natural flow implies the violation of an operating limitation, the proposed storage use must be modified. For instance, if the sum exceeds  $D_{max}$ , only  $D_{max}$  can be discharged and if the sum is less than  $D_{min}$ ,  $D_{min}$  must be discharged anyway, provided that it is available.

The same procedure applies for determining operation for the rule curve, Case 3. Here, however, the flow in the preceding

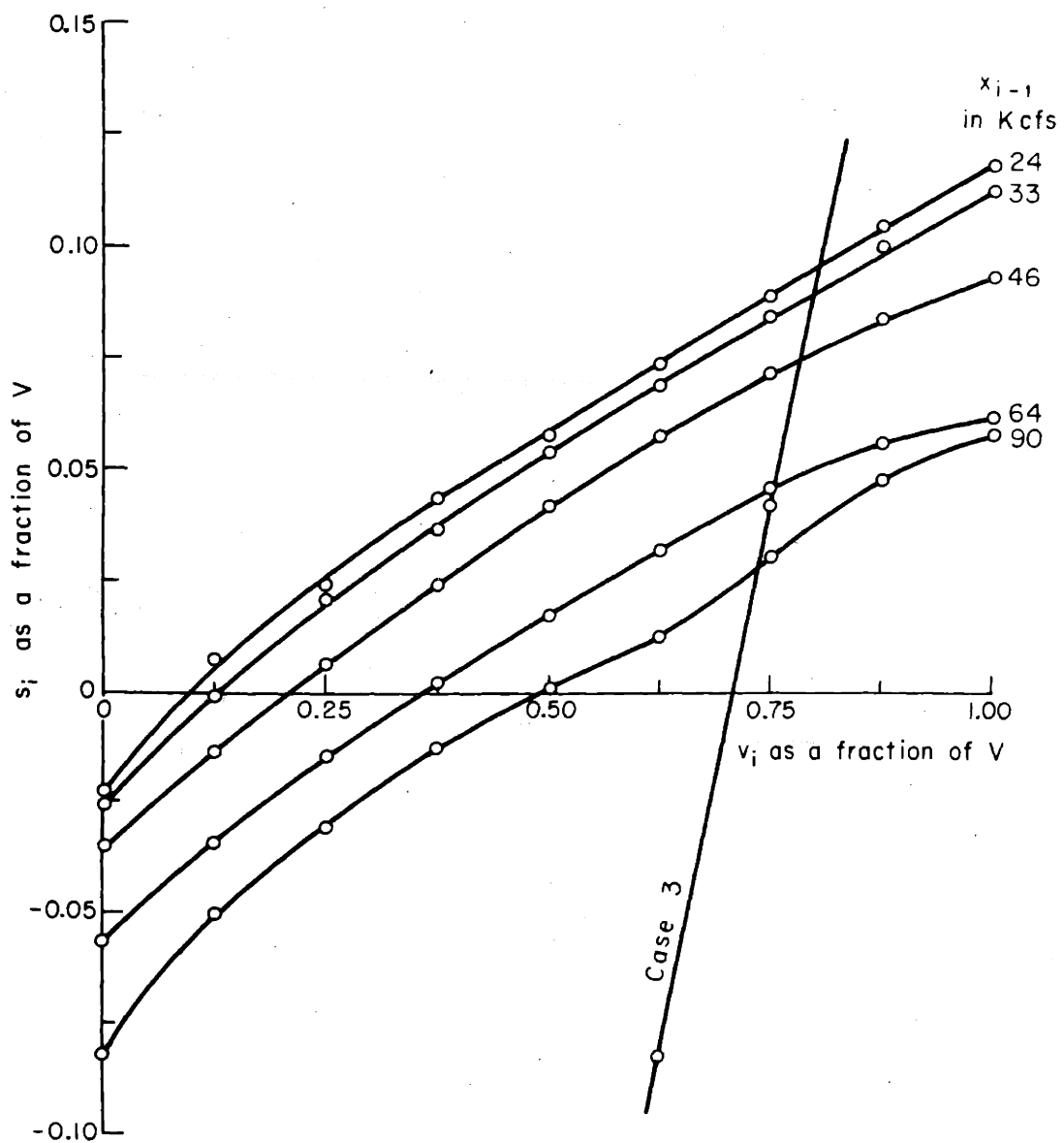


Fig. 4. An example of curves for determining volume of storage water use. The family of curves represents the function  $s_1(v_i, x_{i-1})$  for Case 2, 1 is the interval from Dec. 10 to Dec. 30. The corresponding information for the rule curve, Case 3, is also plotted.



interval is immaterial. For  $v_i = 0.80$  V,  $s_i = 0.094$  V.

This corresponds to a discharge of 17.5 Kcfs to be added to the natural flow during the interval.

The line shown for rule curve operation is not the rule curve itself but the point for the  $i$ th interval translated into a form comparable to the curves of  $s$ . The rule curve line is deceptive in that rule curve operation does not ordinarily permit the reservoir level to fall below some given value which, for the beginning of this interval, happens to be  $v_i = 0.80$  V. Thus the line shown for rule curve operation is used almost exclusively to the right of 0.80 on the graph.

A few further comments may be made about the curves of Fig. 4. First it is seen that for  $v_i = 0.80$  V, the rule curve value, the Case 2 curves do not recommend using as much water as the rule curve in this interval unless the flow,  $x_{i-1}$ , is rather low, about 28 Kcfs. Second, the volume of water allotted from storage in this interval decreases as  $x_{i-1}$  increases. This is a move to preserve water and head and does not happen in all intervals. In particular, situations arise toward the end of the drawdown season where good flow means that the danger of power shortage is past and water should be pulled out of storage rapidly. Finally, some of the unevenness of the points in the

Case 2 curves may be valid, but some is probably due to the limited accuracy of the calculation.

#### Results of Operation with the Historical Flows

The principal features of the results are:

1. The differences between the various methods of operation, as judged by the criteria used, were fairly small. Variations range from tenths of a percent to a few percent. This was to some extent expected. The amount of water which actually comes down the river is the first order effect. The manipulation of the water, provided it is sensible, will have a smaller effect. Its economic importance stems from the amount of money involved in a big operation.
2. The differences although small do exist. The expected value method for determining reservoir operation did what it was supposed to do. That is, the results of operation, as judged by a given criterion, were best for operation determined by an expected value calculation using that criterion.
3. Some effects become apparent only when averaged over a number of years. A method of operation which, compared to another, decreases cost in the long run may increase cost for some of the individual years. This possibility is basic to the concept of expected value and arises from the probabilistic

nature of the flows. The operation which in retrospect would have been best is attainable in practice only when future flow is perfectly known.

4. Operation was often at an extreme determined by operating limitations. That is, for much of the year, the reservoir was full, or empty, or discharge was maximum or minimum.

#### Curves of System Behavior

In Fig. 6 and continued in the Appendix are curves relating to the system behavior under the input of the historical flows. The curves are composed of dots on an oscilloscope plotted and photographed by the Whirlwind I computer. Reading across the page are sets for the same flow year for cases 1,2, and 3 respectively. The curves show the historical flow, the volume of storage water planned for use in the interval, the supplemental energy, and the reservoir volume, all as functions of time through the year. Further details are explained in Fig. 5.

The pictures are too gross to show up fine differences in operation but are excellent for seeing what the system is doing macroscopically. Characteristic are: the decreasing flow in winter followed by spring flood, the drawdown of the reservoir in winter and refill in spring, and the hump of supplemental energy required in winter.

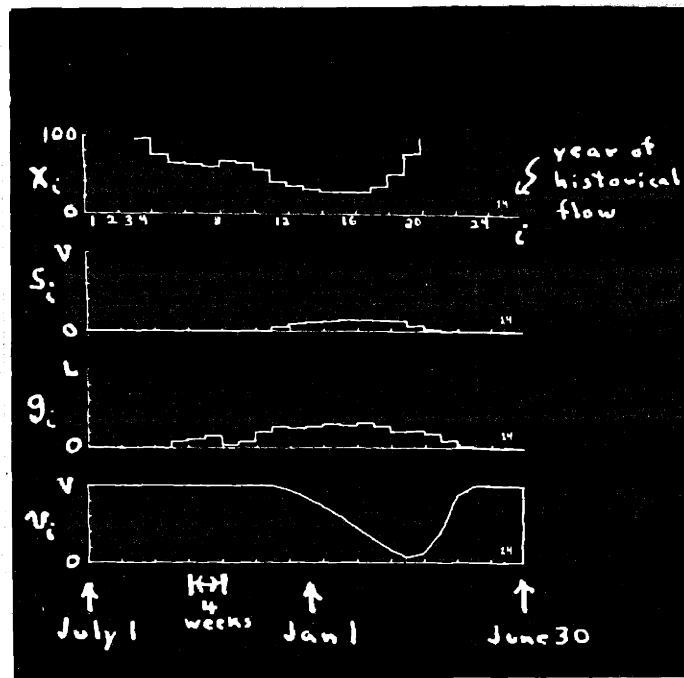


Fig. 5. Explanation for oscilloscope pictures.  
 $x_i$  = natural river flow in Kcfs.  
 $s_i$  = planned volume of storage water use.  
 $g_i$  = supplemental generation.  
 $v_i$  = volume of water in storage.  
 $i$  = index of two week intervals.  
 $V$  = volume of full reservoir, 2600 Kcfs days.  
 $L$  = full load, 1780 megawatts.  
Flows greater than 100 Kcfs are off scale.

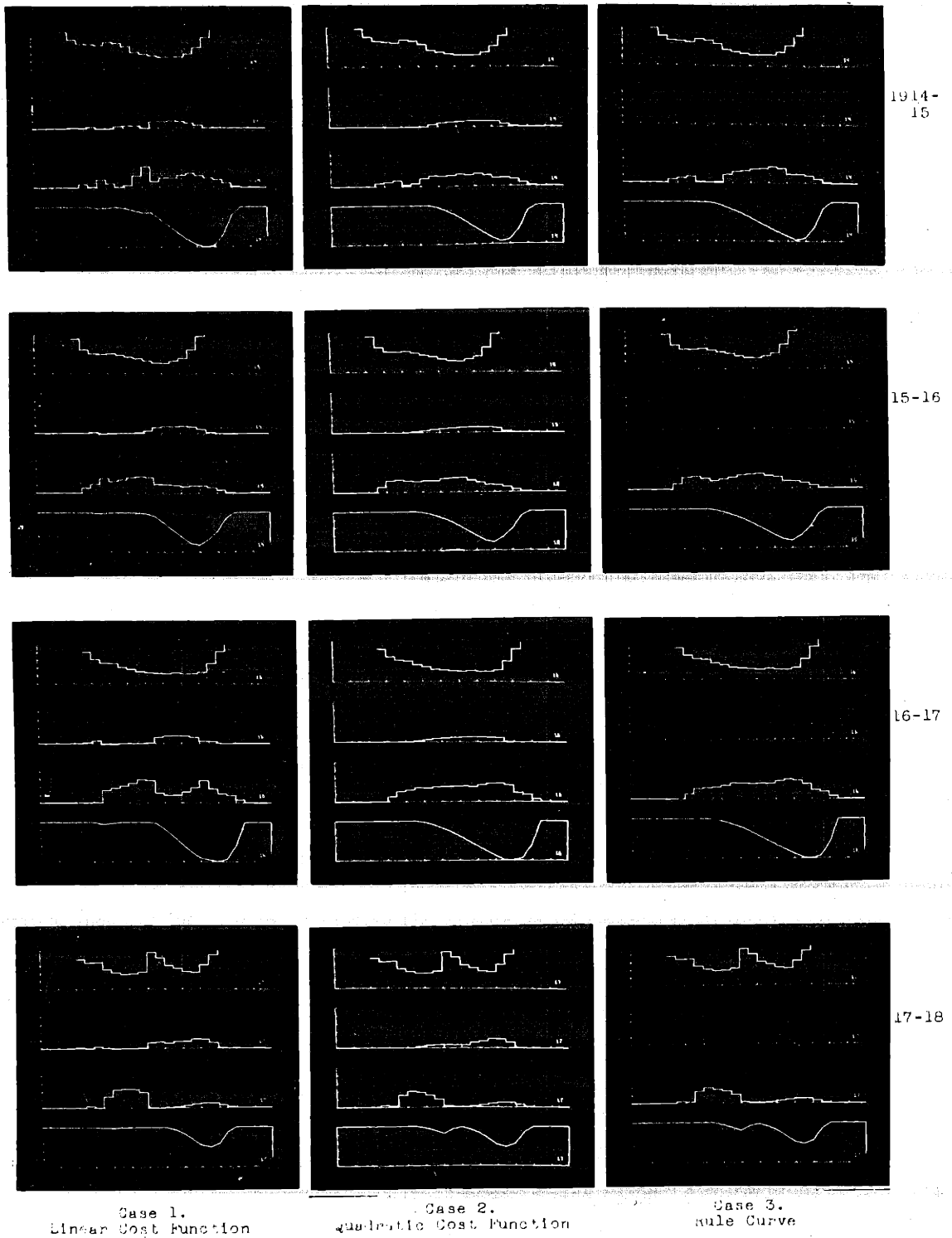


Fig. 6. Oscilloscope pictures of system behavior with historical flows (continued in Appendix).

In comparing the three methods of determining use of the stored water, the following distinguishing features appear. In Case 1, linear cost and maximum expected energy, the supplemental generation varies considerably from interval to interval. There is no penalty for this and so it is not surprising. Cases 2 and 3 on the other hand show more even generation. In case 2 this is caused by the quadratic cost of supplemental energy. In Case 3 the rule curve doles out the water fairly evenly through the winter season.

The rule curve, of course, does a fine job of smoothing out 36-37. However, in other years the supplemental generation is much more uneven.

Case 1 tries to get the most hydro energy out by doing two things: holding the head late into the fall, and trying to empty the reservoir early in the spring. The result is often large supplemental generation in the fall and late winter.

#### Average Costs

The three cases of operation with the historical flows are compared on the basis of two criteria. The first is the 39 year average supplemental power  $\bar{g}$ . This number is obtained from the supplemental energy total of the 39 years of operation. If the

cost function is linear, the average cost may be found by multiplying  $\bar{g}$  by some constant. If the cost function is quadratic,  $\bar{g}$  is not directly proportional to cost but still is proportional to the average supplemental energy required. The other criterion is  $\overline{g^2}$ , the average of the supplemental power squared. If the cost function is pure quadratic, the average cost may be found by multiplying  $\overline{g^2}$  by some constant. If the cost is not quadratic,  $\overline{g^2}$  is not proportional to cost but may be regarded as a measure of the variability of the supplemental generation. In fact, in mathematical terminology  $\overline{g^2}$  is proportional to the second moment about the origin of the set of all biweekly supplemental energies. Similarly  $\bar{g}$  is proportional to the first moment.

The quantities  $\bar{g}$  and  $\overline{g^2}$  are proportional to average cost if the cost functions are, respectively, linear and quadratic. Another interesting interpretation of these quantities can be made. It is found that as  $\bar{g}$  is minimized  $\overline{g^2}$  increases and vice versa. As  $\bar{g}$  decreases, more hydro energy is obtained from the system. On the other hand, as  $\overline{g^2}$  correspondingly increases, it implies an increase in the number of occasions when the supplemental generation is large. If we call periods of large supplemental generation power shortages, the conflict between minimum  $\bar{g}$  and  $\overline{g^2}$  may be regarded as a conflict between maximum hydro energy and minimum risk of shortage. In order to have a basis for operation some

weighing of the two factors must be decided upon. This, of course, is just what a cost function does. As stated earlier, a much greater variety of cost functions is possible than the pure linear and pure quadratic cases considered here.

Table 3 shows what happens to  $\bar{g}$  and  $\bar{g}^2$  in the three cases of operation calculated. Consider first Case 1, linear cost. When operation is such as to minimize expected cost, the 39 year average power required from supplemental generation,  $\bar{g}$ , is a fraction, 0.1520, of the total load of 1780 MW. This is 1.4% less than required for the operation of Case 2, and 1.4% less than required by the rule curve operation, Case 3. If cost is linear, these changes are changes in cost. Thus when Case 1 is judged by the criterion used in determining its water use, its operation is superior to the others.

We have talked of cost but not in terms of dollars. As stated above, the exact value of  $a_1$  does not make any difference to operation. However, to find what order of magnitude of money is involved in this model, we might take the cost of supplemental energy as 3 mils/kilowatt-hour. This would mean an average annual supplemental energy bill of about \$7,000,000, 1% of which is \$70,000. As discussed earlier, the model does not represent any actual situation in the Northwest and the same must be said for any dollar cost presented.



	Case 1	Case 2	Case 3
$\bar{g}$	.1520	.1542	.1542
$\overline{g^2}$	.05522	.04910	.04956

(a)

	Case 1	Case 2	Case 3
Change of $\bar{g}$ from Case 2 value	-1.4%	0	0
Change of $\overline{g^2}$ from Case 2 value	12.4%	0	0.9%

(b)

Table 3. The 39 year averages of supplemental power and supplemental power squared ( $\bar{g}$  and  $\overline{g^2}$ ) for the three cases of operation.

- a) Averages  $\bar{g}$  and  $\overline{g^2}$  expressed as fractions of full load values (1780 MW and 3.17 x10<sup>6</sup> MW<sup>2</sup> respectively).
- b) The same data expressed as percentage change from Case 2 results.

A discussion similar to the above applies to Case 2, quadratic cost function. Using this operation,  $\bar{g}^2$ , the average value of the power squared over the 39 year historical record is a fraction, 0.04910, of the value it would have if all power came from supplemental generation. On the other hand for the operation defined by Case 1,  $\bar{g}^2$  would be 12.4% greater, and, for the rule curve, 0.9% greater. If the cost function is quadratic, these changes are differences in cost. Again we might take the dollar value of a 1% change to be about \$70,000. Although the cost is not linear, this number is consistent with the idea that, whatever the shape of the cost function, there is only a certain range of average, long term costs that will keep a hydro system in business and that 1% of this is somewhere in the indicated range.

Case 3 is the rule curve. The criterion on which it is based is that there should be constant supplemental generation in the drawdown period of 1936-37. This criterion is certainly fulfilled. However, the rule curve results are the same or worse than Case 2 by both linear and quadratic cost criteria. The differences, however, are small.

The averages,  $\bar{g}$  and  $\bar{g}^2$  over the 39 year record have been used for judging system performance. Two related numbers, the expected values of  $g$  and  $g^2$ , can be obtained from the optimization

calculations involved in Cases 1 and 2. If the probabilities used in the calculation were perfect, the calculation itself very accurate, and the number of years in the historical record many more than 39, then the expected values from the calculation should agree closely with the averages from the historical record. These conditions are not well satisfied. The result is that the corresponding numbers differ by a few percent. They are shown in Table 4. It is noted that the changes in the expected values for the two cases are similar to the changes found in the average values of Table 3.

#### Other Results

Figure 7a shows frequency histograms made up from the sets of the 39 values of  $g$ , the average supplemental generation for the year. The three cases are seen to be quite similar. The reason is that the niche into which each year falls in this plot is determined mostly by the first order effect, the amount of winter runoff for the year. Wet years have low average supplemental power and fall near the origin; dry years appear out at the right. The effect of different reservoir operation is to shift the mean a few percent and does not show up much on this plot. Figure 7b shows similar histograms for  $g^2$ , the average square of the supplemental generation. The abscissas of Fig. 7a may be read as cost if cost is linear; the abscissas

	Case 1	Case 2
E(g)	.1488	.1517
E(g <sup>2</sup> )	.05618	.04827

(a)

	Case 1	Case 2
Change of E(g) from Case 2 value	-1.9%	0
Change of E(g <sup>2</sup> ) from Case 2 value	16.4%	0

(b)

Table 4. The expected values of supplemental power and supplemental power squared as calculated from the probabilities.

a) Expected values E(g) and E(g<sup>2</sup>) as fractions of full load values (1780 MW and 3.17x10<sup>6</sup> MW<sup>2</sup> respectively).

b) The same data expressed as percentage change from Case 2 results.

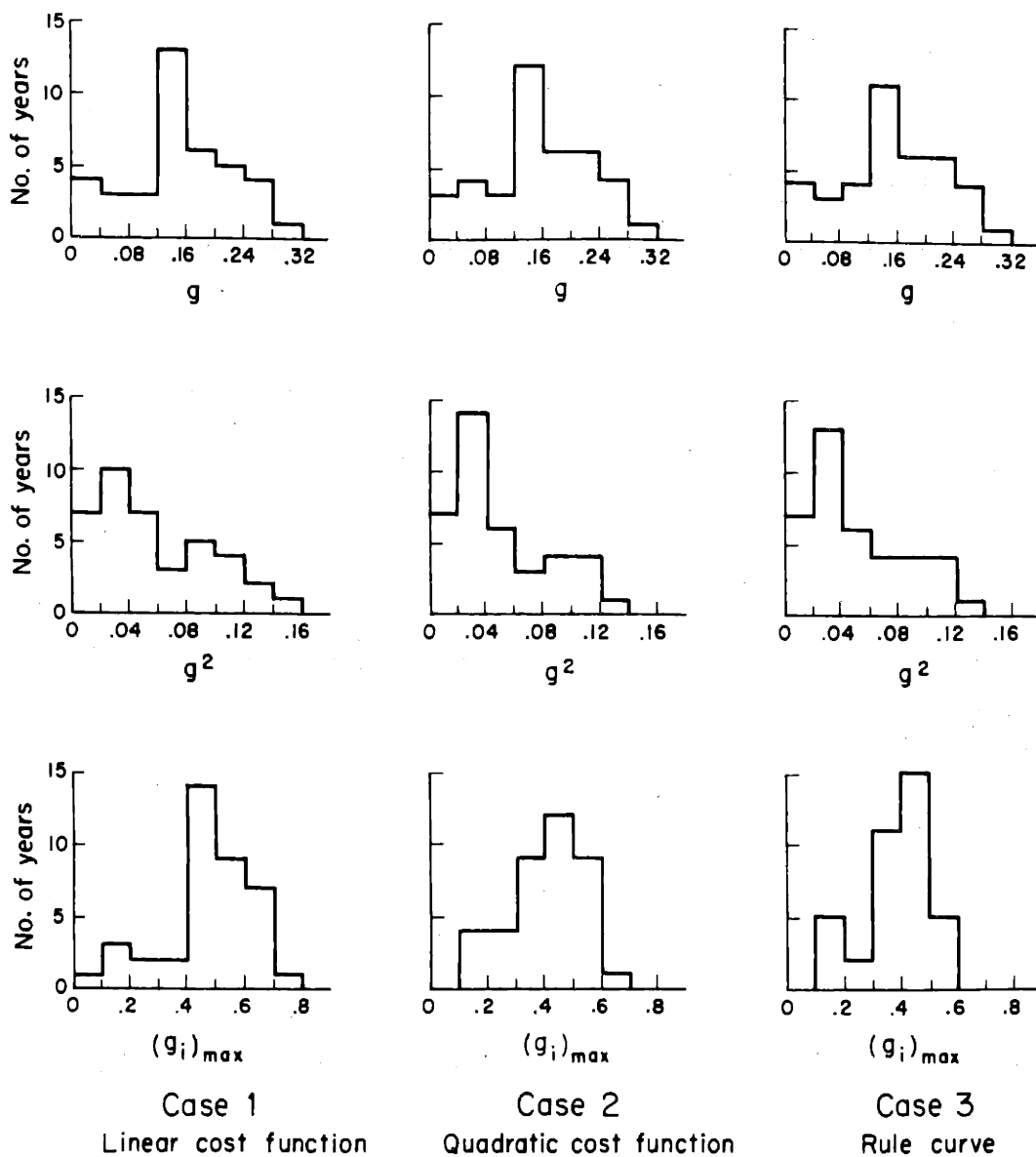


Fig. 7. Frequency histograms of  $g$ ,  $g^2$ , and  $(g_i)_{max}$  from the 39 years of flow for the three cases of operation. The ordinate is the number of years for which the particular quantity had a value between adjacent marks on the abscissa. Distributions are given for a)  $g$ , the average supplemental power during each year, b)  $g^2$ , average square of supplemental power, and c)  $(g_i)_{max}$ , largest  $g_i$  during each year. The quantities are given as fractions of their full load values.

of Fig. 7b as cost if cost is quadratic.

These plots show the variability in energy and cost due to randomness in river flow. The wide range of possibilities which nature can produce rather dwarfs the effects of different reservoir operations.

Figure 7c shows frequency histograms for  $(g_1)_{\max}$ , the largest value of supplemental generation during a year. Here the rule curve operation gives somewhat lower values than the quadratic cost operation. Perhaps then the rule curve in this model best corresponds to a cost function which rises more steeply than a quadratic. However, since the rule curve is not made up on the basis of a specific cost function, and since no calculations have been made with functions containing higher powers than second, we cannot conclude much.

An interesting characteristic of the operation of the model is the fraction of time that operation is not at an extreme limited by restrictions. The operating limitations are maximum discharge, minimum discharge, full reservoir, and empty reservoir. The reason for interest in the fraction is that it is some indication of how much of the time water use is difficult to determine. Of course, it is not necessarily easy to decide that operation should be at a limiting restriction, but, for

instance, in spring and summer the only rational thing to do is to maintain maximum discharge. The fraction of time during a year that operation was not at an extreme ranged from 0 to 0.42 and averaged 0.19, 0.28, and 0.31 for Cases 1, 2, and 3 respectively. The differences between cases are due in part to successively earlier times for starting drawdown.

#### Accuracy of the Calculation

Once the numbers which represent the model and the historical flows are put into the computer they may be invested with almost any desired degree of accuracy. This is advantageous for computing differences. Differences can be significant even though they come from numbers artificially increased in accuracy. For instance, the product of the head and the volume of flow is proportional to energy. A real flow may not be known to better than 5%, a real head to 1%. However, if the head can somehow be increased by a foot, the energy will be increased. By consistently attributing extra significant figures to head and flow, the energy before and after can be calculated and the difference made a valid measure of the gain.

In operating the system with the historical flows it is only necessary to carry that accuracy beyond which any differences detected would be of doubtful significance.

The accuracy of the optimization calculation is a somewhat different problem. The question is whether the calculation can be improved. There is some reason to believe that the calculations presented here can be. In the first place all the knowledge of the flows which is available to the optimization calculation must be embodied in the probabilities. These were calculated in a way which stressed generality over hydrological niceties. In particular the correlation between low flows is greater than between high flows, a fact not well taken into account. Furthermore the probabilities, once calculated, were then subjected to a smoothing which, although considered beneficial, might better have been planned into the original calculation.

Another accuracy problem in the optimization concerns the approximation of differentials by differences. We write

$$\frac{dE}{ds} \approx \frac{\Delta E}{\Delta s} = \frac{E(s + \frac{\Delta s}{2}) - E(s - \frac{\Delta s}{2})}{\Delta s} \quad (6.4)$$

We seek  $\frac{dE}{ds} = 0$ , except when best operation is at either the largest or smallest  $s$ , in which case the accuracy problem is not so acute. Since  $\Delta s$  is a constant of the calculation, comparisons are made on the basis of  $\Delta E$ . By decreasing  $\Delta s$  we should obtain increasingly good approximations to the



derivative at  $s$ . However, the differences  $\Delta E$  decrease at the same time and, for a range about the minimum, may become smaller than the smallest numbers the computer is detecting. Within this range,  $\Delta E$  equals zero, and the best  $s$  is indeterminate. In the computations reported above, the compromise was to take  $\Delta s$  to be a volume of water about  $3\frac{1}{2}\%$  of  $D_{max}$ .

### Stability

Related to the problem of finding the minima accurately is the question of what happens when operation is not at the minimum. This can easily occur. For one thing, the calculations may not have been perfect so that the recommended operation does not bring about true minimum cost for the model. Then again, even a good model will differ from a real situation so that its best operation may not be exactly right for the real case. Finally, in real operation disturbances inevitably arise which cause the best laid plans to be modified.

Therefore we want to know how flat the minimum of the expected cost curve is as a function of changes in operation.

Unfortunately, for each of the many values of  $s_i(v_i, x_{i-1})$  in the tables, there corresponds a different curve whose minimum has been found. These different minima vary in character.

Some are not even true minima but corner minima. Thus no single curve is representative of the lot whereas it would be desirable to characterize the whole optimization. The following rough way of doing so was used. The cost of system operation was calculated from the historical flows but, as values of  $s$  were selected from the tables for deciding the storage use in each time interval, a small number was added to  $s$ . This bias made a consistent change in favor of greater discharge, or, if negative, less discharge. The bias is a forced deviation from recommended water use. However, to force an added amount of drawdown early in the fall when the reservoir should be held full would be contrary to common sense. Consequently, when the tables called for an extreme value of  $s$  implying a corner minimum, the bias was not applied.

By calculating the 39 year average cost for a number of values of bias, the solid curve of Fig. 8a was constructed for Case 2. It shows that a small positive bias seems to improve operation slightly. Thus, at least with respect to the historical flows, the operating tables can be improved somewhat. As larger bias, positive or negative, was applied, the average cost increased as is shown quantitatively by the curve.

The same bias treatment was given the rule curve operation.

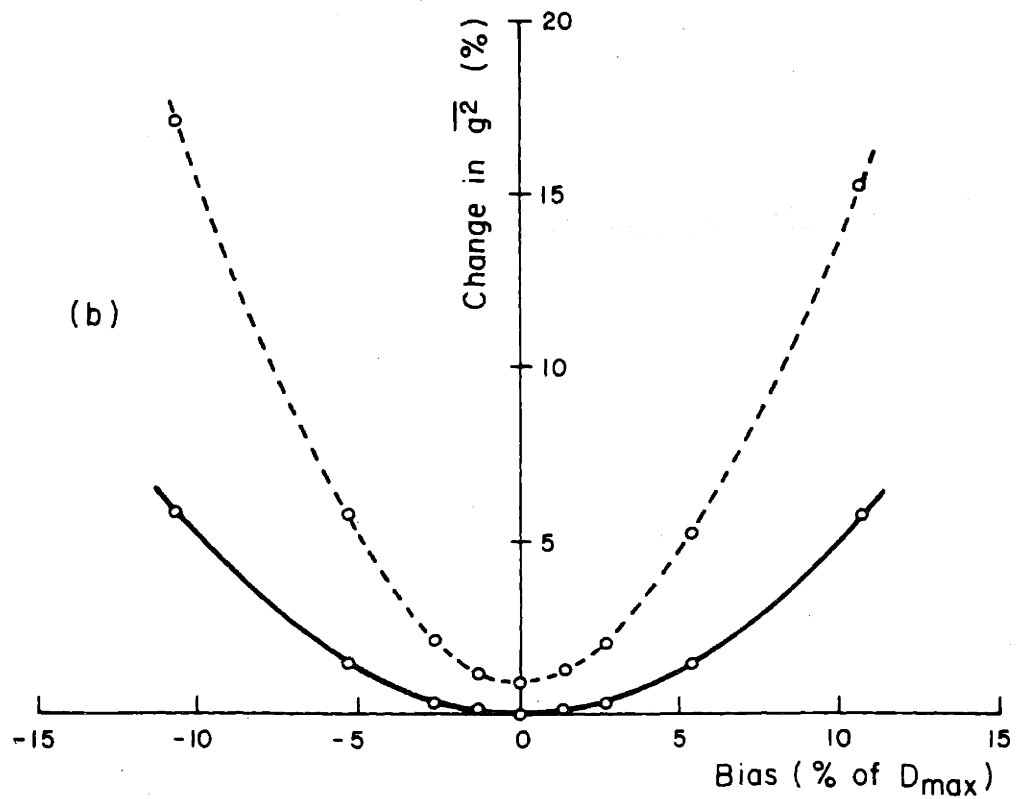
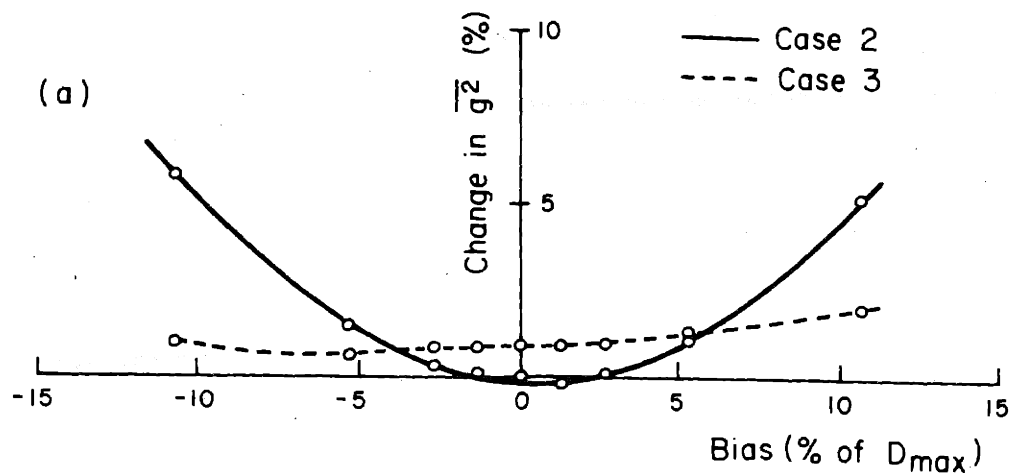


Fig. 8. The effect on  $\overline{g^2}$  of departure from recommended water use in Cases 2 and 3. For a quadratic cost function,  $\overline{g^2}$  is proportional to average cost. a) Bias consistently added to recommended storage use. b) Bias alternately added and subtracted from recommended storage use.

The result is the dashed curve of Fig. 8a. The cost is decreased for some values of negative bias. It is seen that, although the Case 2 cost is lower than rule curve cost near zero bias, the latter is less sensitive to the bias and does not rise so rapidly with bias.

The reason for the difference may be found by considering how the two cases handle the bias. For the rule curve the first application of the bias, if positive, pulls the reservoir below the rule curve level. In the next interval, the rule curve calls for that storage use which will bring the reservoir back onto the curve by the end of the interval. This volume is again changed by the bias and the interval again ends with the reservoir below the rule curve by a volume equal to the bias. Thus water use is very similar to that of rule curve lowered by an amount equal to the bias. Except for the first and last interval of the drawdown season, the volume of water planned for use is the same with and without bias. The result, shown in Fig. 8a is that the average cost is not changed much.

The bias is treated differently by the minimum expected cost operation of Case 2. Once again a positive bias pushes the reservoir below a level which might be considered appropriate for the time of year. However, instead of trying to correct

all the trouble by the end of the next interval, the tables call for a water use that implies spreading the correction over the rest of the year. This is done under the assumption that in future intervals water use will be such as to minimize expected cost, i.e., according to the later tables. Therefore, when the bias continues to be applied in subsequent intervals, the water use continues to be changed fairly strongly by the bias. The graph shows the resulting increase in cost.

Suppose, however, that bias is not consistently applied in the same direction. For instance, in rule curve operation this might result from a decision to draft the reservoir a little below the rule curve in one interval and then to build up above it in the next. While such alternation would probably not continue long in practice, it is a concrete departure from recommended operation with which to make calculations. The results are shown in Fig. 8b for a bias which, starting with the first time interval, is alternately added to and subtracted from the volume of storage water planned for use. Corner minima values are excepted. Case 2 operation behaves almost as before. On the other hand the rigid operation called for by the rule curve fares badly. After a positive bias causes excessive drawdown, the rule curve automatically calls for cutback of water use in the next interval and, since this

is then coupled with a negative bias, a certain amount of oscillation is introduced and with it high cost.

In practice reservoirs are drawn below rule curves only rarely so that the above might seem too hypothetical. In fact, there is some tendency in controlling reservoirs to keep a safety cushion of water above the rule curve. However, if such operation of the reservoir implies the existence of an unwritten rule curve a little higher and a little more flexible than the written one, then the above discussion and curves are pertinent.

Another point concerns the curves for Case 2. In the course of experimenting with the calculation a number of such curves were made. As the calculations were improved and the minimum decreased, the bottom of the curve was pushed down more than the sides. The improved calculation then gave a less flat curve but since, near the minimum, it lay wholly below the old curve, deviations from recommended operation were still a little lower in cost than formerly. Thus it is plausible to assume that, if the calculations can be further improved, the flatness of the minimum may decrease, but that this will not cause higher costs than before for corresponding deviations from recommended operation.

The curves of Fig. 8 may give an exaggerated impression of the seriousness of deviations from recommended water use because the curves represent the results of applying the bias at every permissible opportunity. If deviations occur less frequently, the increase in average cost will be less. Furthermore the range of bias plotted is large. A deviation of 10% of  $D_{\max}$  is a drastic change when it occurs in a dry month of winter. Thus departures from recommended operation are most likely to come in the flatter portions of the curves and should have the cost significance shown only if consistently maintained.

## VII. POSSIBLE GENERALIZATIONS OF THE MODEL

Some of the simplifications of a real system which were made in the model may easily be removed. For instance, the complications of non vertical reservoir walls, variation of tailrace with discharge, change of maximum discharge with head, and variation of turbine efficiency with head can be calculated by the machine without difficulty by putting them in tabular or functional form. However, in forming such functions, it must be remembered that it is not the instantaneous variation of tailrace with discharge that comes into the long range problem but an average value which takes into account daily fluctuations. Similarly maximum discharge refers to maximum average discharge for energy over the time interval and may be limited by peaking capacity. It can vary with the time of year. The load can also vary during the year.

The complications which increase the difficulty of the problem are those which introduce new variables into the decision. In the model used here the decision depended upon three variables: the time of year, the volume in storage, and the flow in the preceding interval. A complication that would seem to require another variable is the fact that interruptible customers do not like to be cut off and on in rapid alternation. If they are



coming on, they wish to know that they can stay on long enough to defray start-up costs. By making the amount of interruptible load a variable, this could probably be taken into account.

The difficulty which arises when more variables are introduced is that the time required for the computation becomes large. The addition of a new variable adds a new dimension to the tables being calculated and means just that many more points to be computed.

The load may be introduced as a random variable if deemed necessary. This would add an integration to the expected value expression and therefore adds about the same complication as a new variable.

Multiple reservoir problems introduce new variables and also interdependent decision functions. The theory as presented is easy enough to extend to multiple reservoirs. However, the calculations implied by a several reservoir system would be formidable if done in the same way as those of this paper. These methods stressed simplicity and generality. The complete range of permissible values for  $s$  was examined at least coarsely even though many of the points could have been predicted in advance not to be minima. Such brute force methods were in keeping with the desire to reduce programming and eliminate manual intervention

as long as the machine time did not become too large. However, for a bigger job, the machine time could probably be materially reduced by always starting the search at a value of  $s$  thought to imply sensible operation. A multiple reservoir problem must be thought out carefully in setting up the model to be sure each reservoir has enough flexibility of operation to warrant optimization. Nevertheless, it is likely that imagination will be required for finding practicable ways to handle multiple reservoir problems.

An interesting use may be made of the generality of the expected value formulation of the problem. In the model of this paper the criterion of performance was a power series cost function. Other criteria are possible and in some ways desirable. For instance, it is a difficult task to evaluate the loss to a region due to a power shortage. Rather than express such a loss as a cost it might be desirable to determine operation in the following manner: minimize the expected cost, provided that the probability of requiring a supplemental generation greater than  $G$  is less than  $P$ . If the probability is greater than  $P$ , minimize the probability.

The calculation to do this can be set up from equation (4.8). Expected value functions have been worked out there for cost, and the same procedure can be applied to probability. Let us

denote the expected cost by  $E_1$  as before and the probability of supplemental generation greater than  $G$  by  $E_1'$ . Then the calculation is to pick

$$S_i = S_i(v_i, x_{i-1}) \tag{7.1}$$

such that

$$E_i(v_i, x_{i-1}; s_i) = \int_{-\infty}^{\infty} \{ c_i(v_i, x_i, s_i) + E_{i+1}(v_{i+1}, x_i) \} f(x_i | x_{i-1}) dx_i \tag{7.2}$$

is minimum, provided that

$$E_i'(v_i, x_{i-1}; s_i) = \int_{-\infty}^{x_i \rightarrow g_i = G} (1) f(x_i | x_{i-1}) dx_i + \int_{x_i \rightarrow g_i = G}^{\infty} E_{i+1}'(v_{i+1}, x_i) f(x_i | x_{i-1}) dx_i \leq P \tag{7.3}$$

If not, minimize  $E_1'$ .

Another interesting possibility involves flood control. Equations similar to those above could be worked out for minimizing the

cost of hydro operation as long as the probability of flood was less than  $P$  and, at other times, minimizing the probability of flood.

### VIII. SUMMARY

The problem has been: how should the stored water in a hydroelectric system be used when future river flow is uncertain. Best water use was taken to be that which minimized the expected cost of operating the system. Expressions for the expected cost functions were set up. The problem was then solved for a simple hydro system by finding the minima of the expected cost functions on a digital computer. The solution consists of tables of storage water use which were stored in the computer and used to operate the system under the input of the 39 years of historical flow. The results were compared to those coming from a simple rule curve operation.

From the point of view of the general problem of the optimization of expected return in multistage processes, the following features have been interesting. First, sensible numerical results required that the probability distributions for successive time intervals could not be assumed to be the same nor could they be assumed independent. Second, the solution contains some "true" minima and some "corner" minima. The latter arise from operating restrictions in the form of inequalities and lead to an operation held back at some point by a restriction. Some optimization problems involve minima of only one type and the

techniques used to deal with them often make use of this fact. In the problem here the minima are found essentially by calculating the functions and picking out the minima so that both kinds can be treated about the same. Another feature of the problem has been the ability of the digital computer to handle a considerable amount of the complexity of the real problem.

From the point of view of determining water use in hydro systems, the method presented here differs materially from those in current use. The expected value method brings in the uncertainty of future river flow by probability distributions and judges system performance by a cost function. The rule curve method, as presented here, takes into account uncertainty of future flow by preparing for the worst year on record and judges that preparation by the evenness of the supplemental generation that would result in that year. The fact that this results in good performance in other years is perhaps fortuitous but, if it did not, some other method of determining operation would probably have been developed. However, the exact procedure for determining the rule can be controversial. For instance, it may be argued that the supplemental generation in the low flow year should not be constant but should increase somewhat throughout the drawdown season, because river flow is almost always better than the recorded worst. Or it may be

argued that little bumps in the worst recorded flow should be smoothed out. Such arguments are really concerned with flow probabilities and expected costs. The expected value method seeks to avoid such arguments by starting out with the flow probabilities and a cost function and using them to find the best operation.

The tables of water use calculated here differ from a rule curve by the addition of another variable to the decision. The volume of storage use is picked on the basis of the recent flow as well as on the volume currently in storage. This was not mandatory but seemed desirable and feasible. It presumably makes better decisions possible.

Comparisons have been made of the results of operating the model with the 39 years of historical flow for three different cases. The most interesting cases are minimum expected cost operation where the cost function is quadratic and rule curve operation. The behavior of the system is found to be very similar in the two cases. If the cost function is quadratic as assumed, the expected value method gives an average cost 0.9% lower than the rule curve.

A good argument can be made that the differences between the two methods are not large enough in consideration of the large

annual flow fluctuations to warrant use of the more complicated method. It would be an excellent result of the calculation if it could be concluded that for this model a simple rule curve operation came as close to best operation as is necessary.

After all in actual system operation qualitative judgments are bound to enter at the last minute to meet changing circumstances so that water use would probably not be exactly as planned.

These arguments are not entirely convincing. Although we have relied upon the rule curve to show that a hitherto untried method produces sensible operation, the techniques conventionally employed for finding the rule curve do not tell where to look for improvement. The expected cost method, on the other hand, requires a quantitative index, such as cost, for judging its performance. For the model and cost function used here, the expected value method does better than the simple rule curve by an amount which is of some interest and seems capable of improvement. Finally, calculations using the historical flows indicate that expected value operation is stable to departures from recommended water use.

The results found are specific to the model used and it is certain that there are many hydro systems without the flexibility of operation needed to permit much optimization. However,

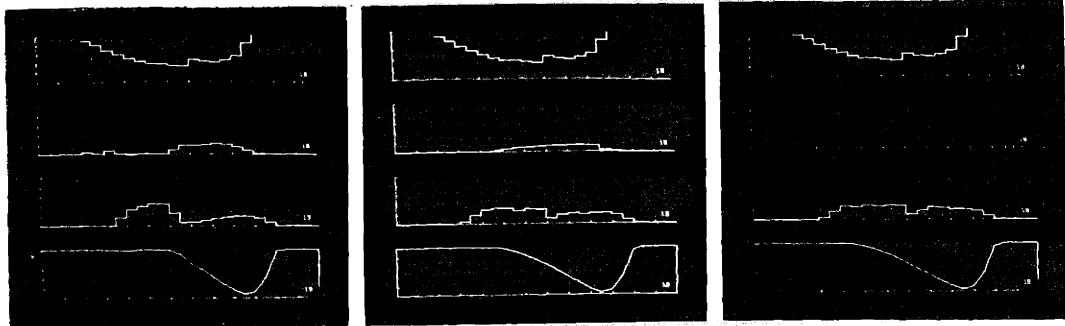


it seems likely that there are systems where expected value techniques can provide practical information on the use of stored water in the face of uncertain future flow.

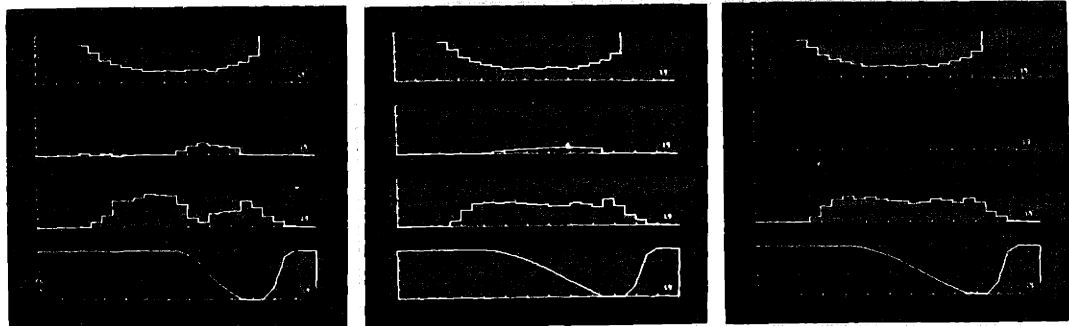
APPENDIX

Fig. 6. Oscilloscope pictures of system behavior with historical flows, 1918-19 through 1952-53.

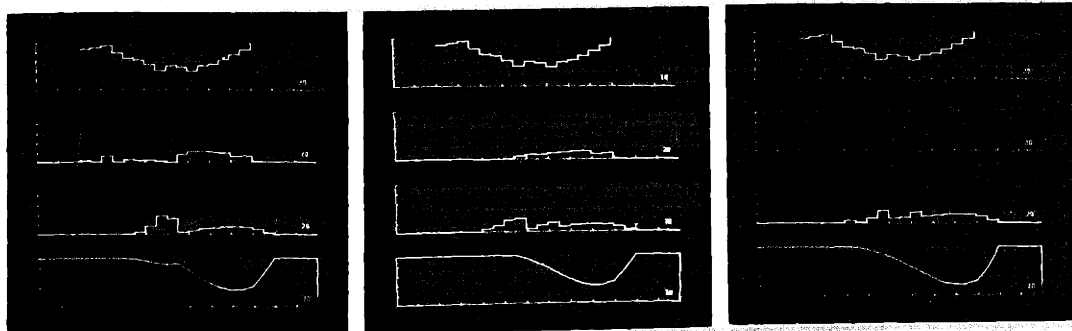
(Continued from p. 77)



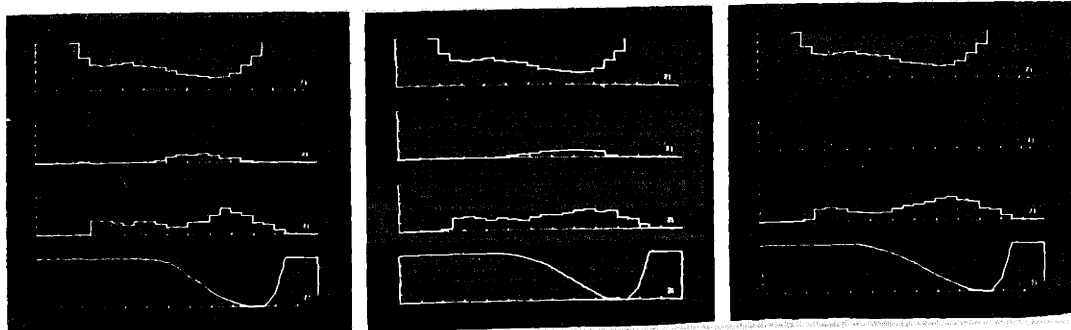
1918-  
19



19-20



20-21

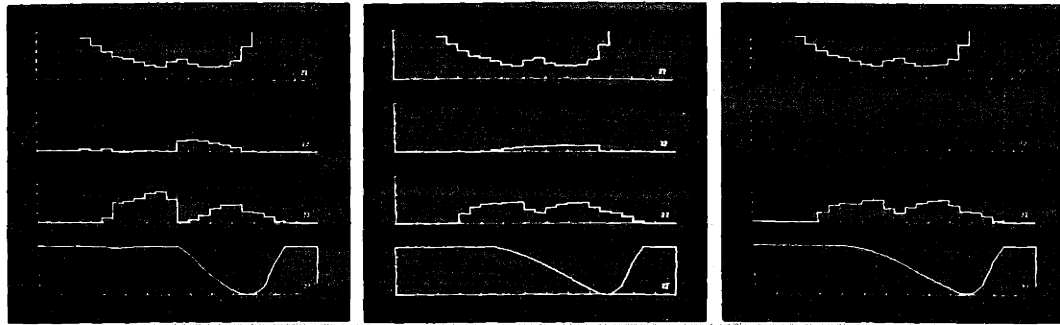


21-22

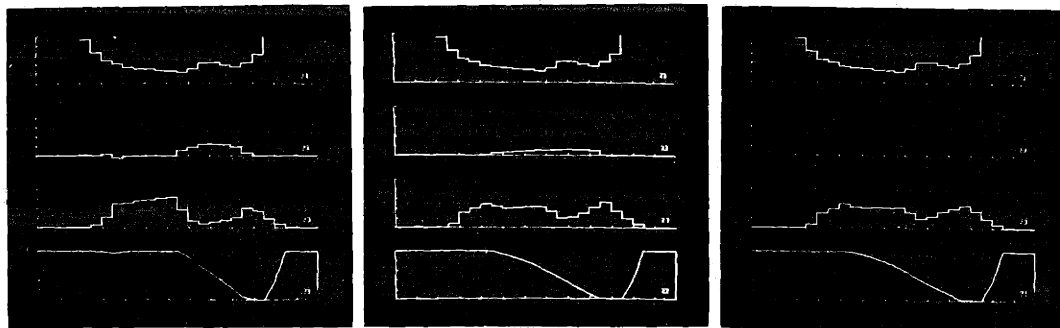
Case 1.  
Linear Cost Function

Case 2.  
Quadratic Cost Function

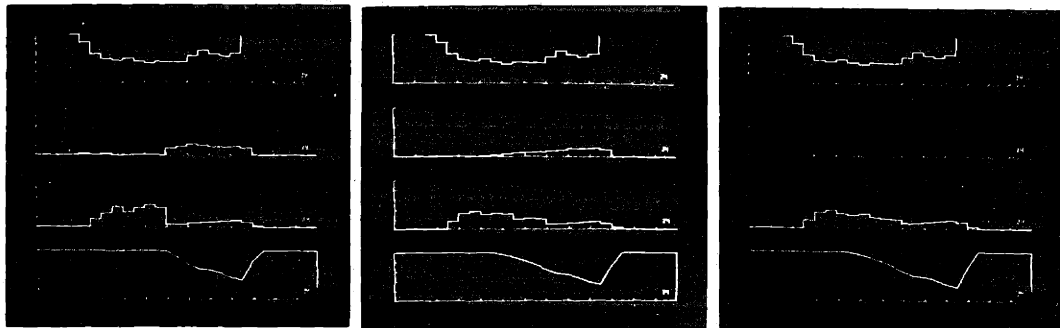
Case 3.  
Sigmoid Curve



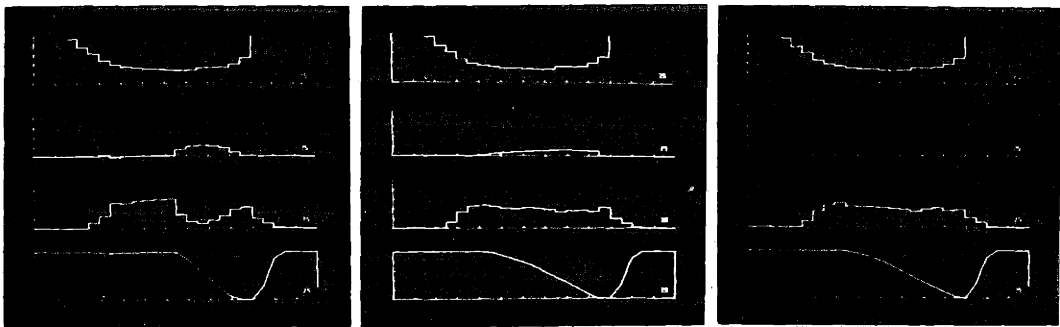
1922-23



23-24



24-25

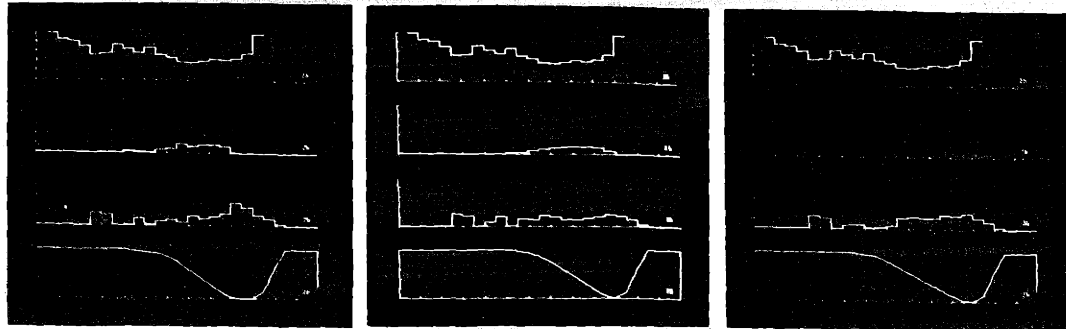


25-26

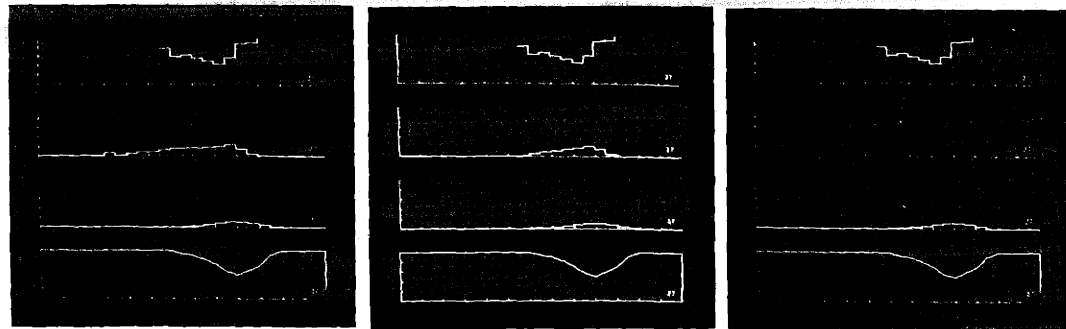
Case 1.  
Linear Cost Function

Case 2.  
quadratic Cost Function

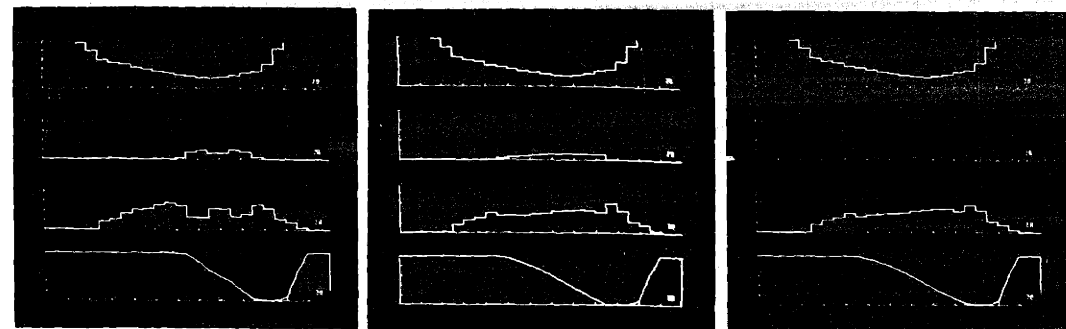
Case 3.  
rule Curve



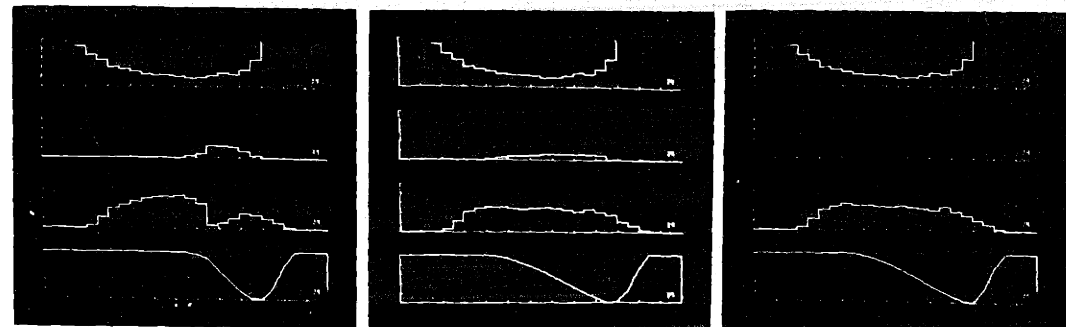
1986-  
27



27-29



28-29

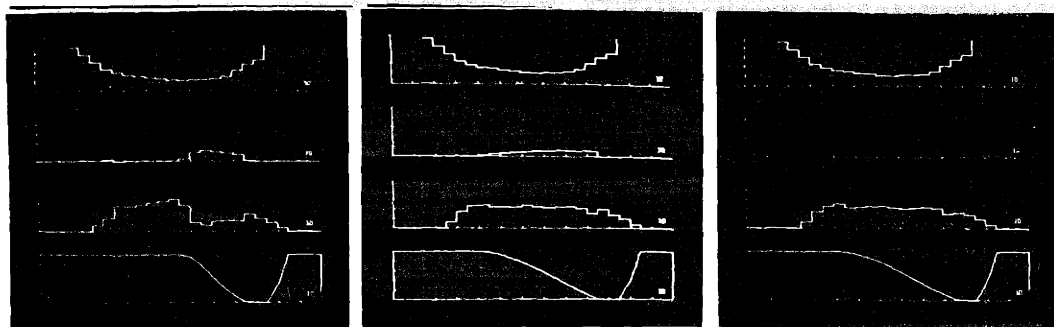


29-30

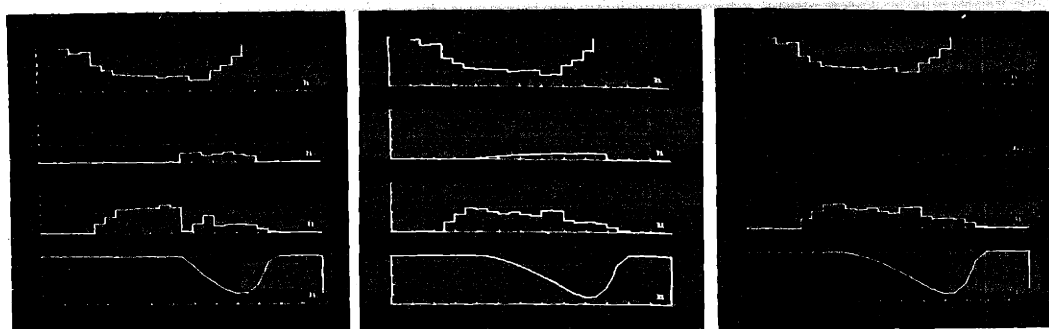
Case 1.  
Linear Cost Function

Case 2.  
quadratic Cost Function

Case 3.  
rule Curve

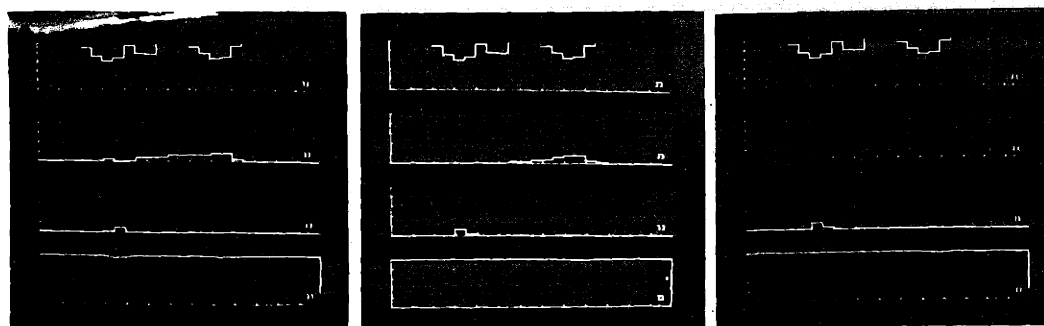


1930-31



31-32

32-33

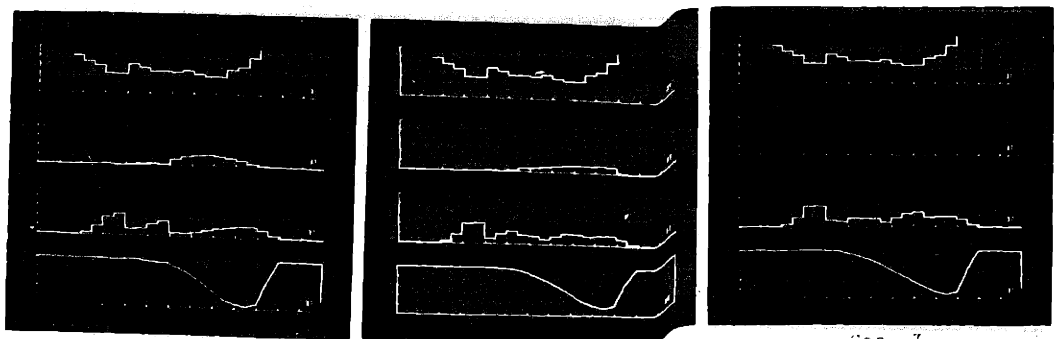
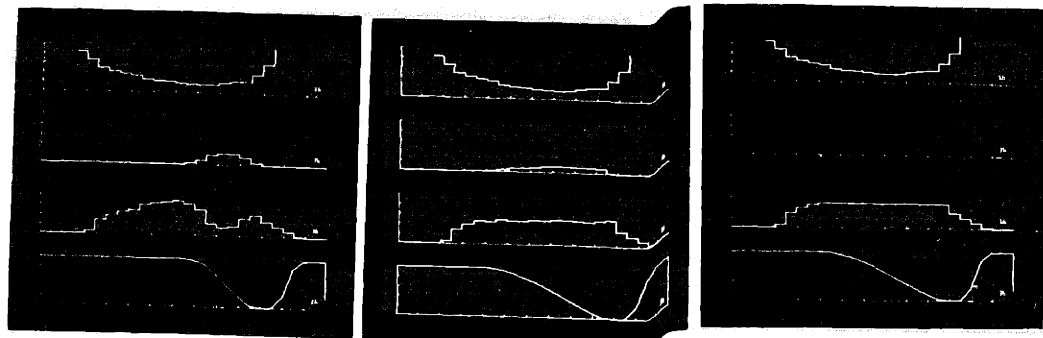
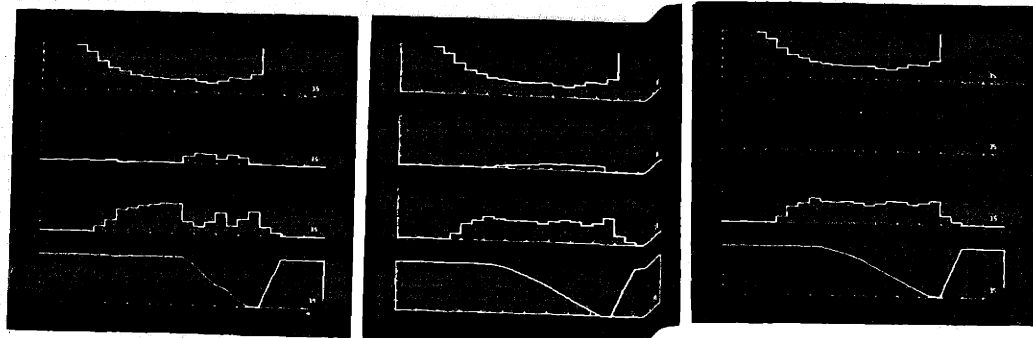
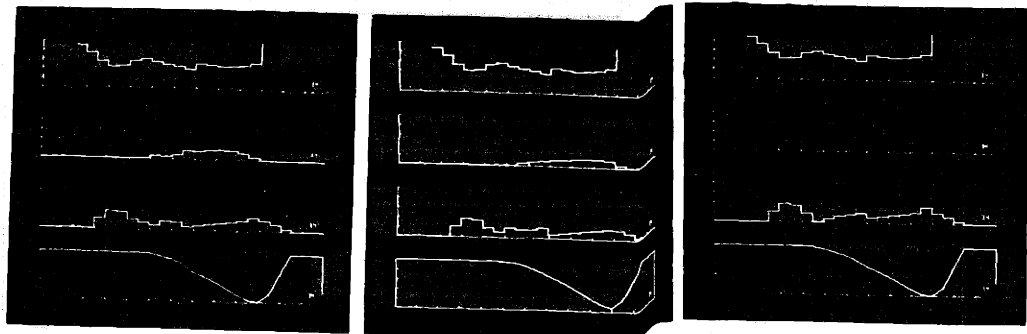


33-34

Case 1.  
Linear Cost Function

Case 2.  
Quadratic Cost Function

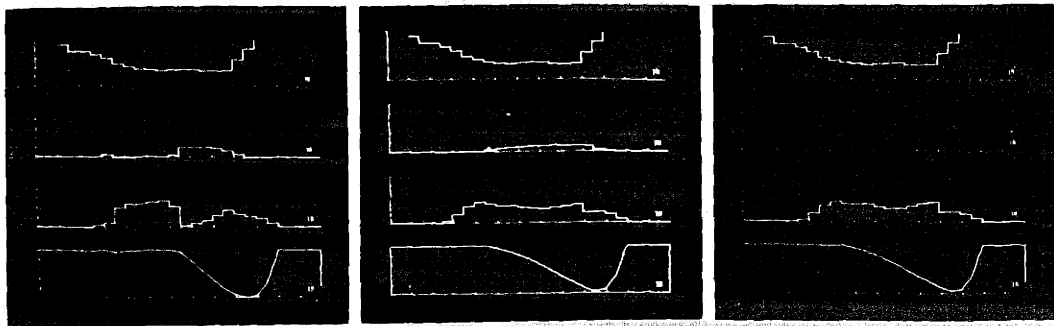
Case 3.  
Rule Curve



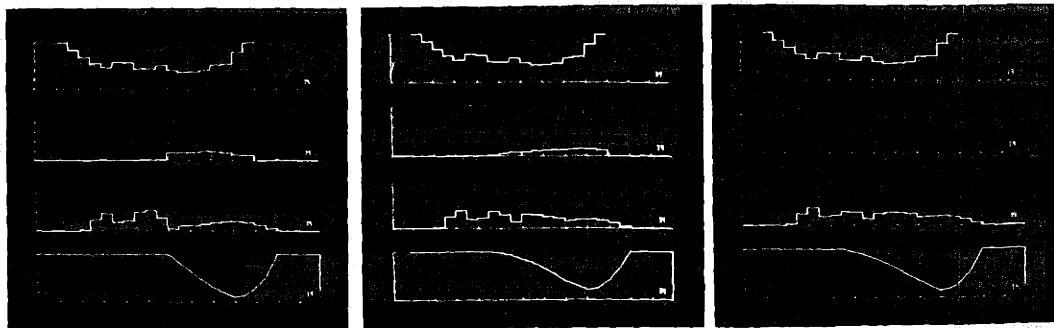
Case 1.  
Linear Cost Function

Case 2.  
Quadratic Cost Function

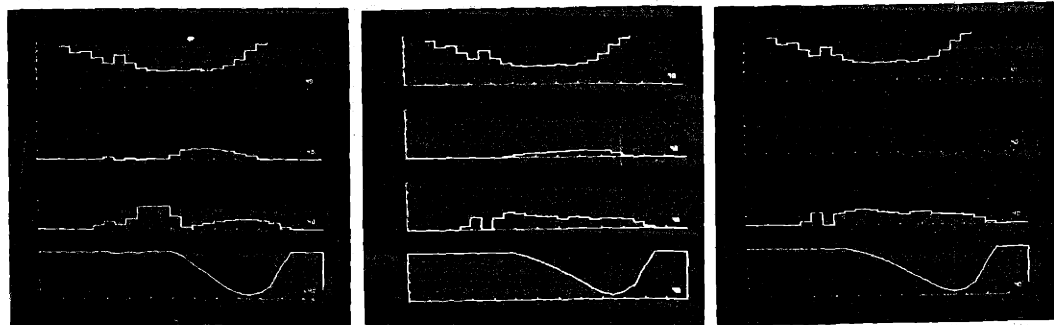
Case 3.  
Rule Curve



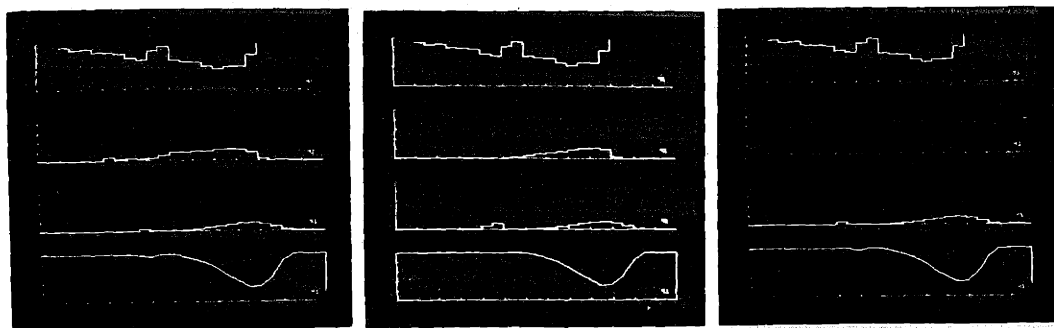
1938-39



39-40



40-41



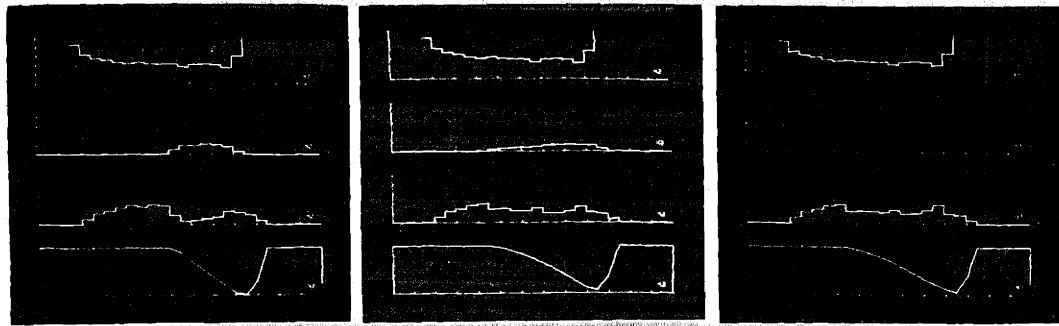
41-42

Case 1.  
Linear Cost Function

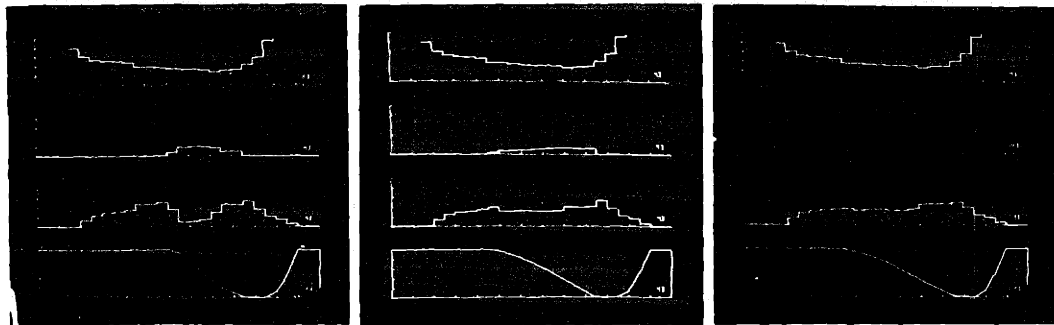
Case 2.  
Quadratic Cost Function

Case 3.  
hule Curve

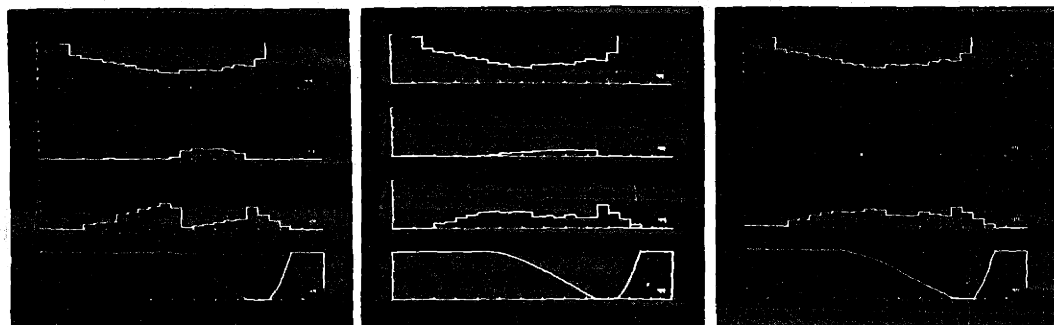




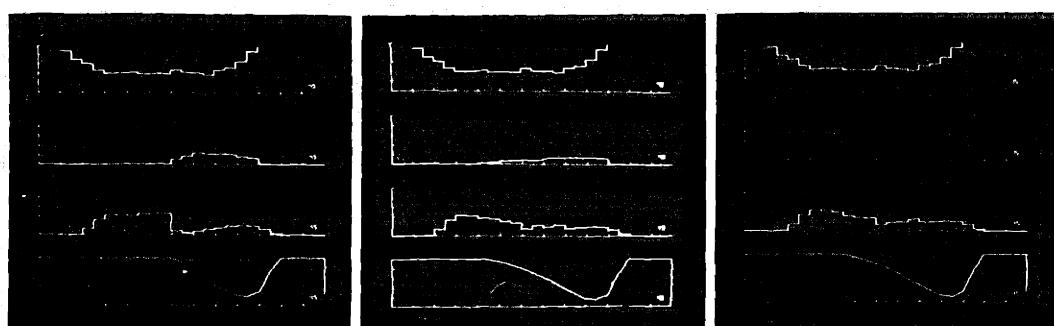
1942-43



43-44



44-45

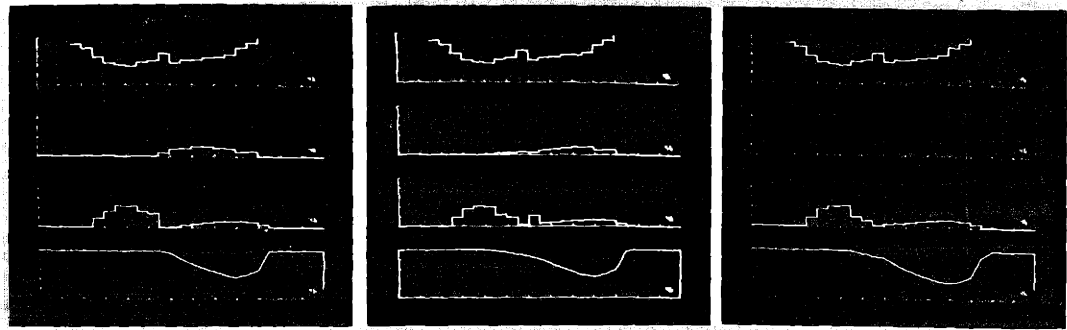


45-46

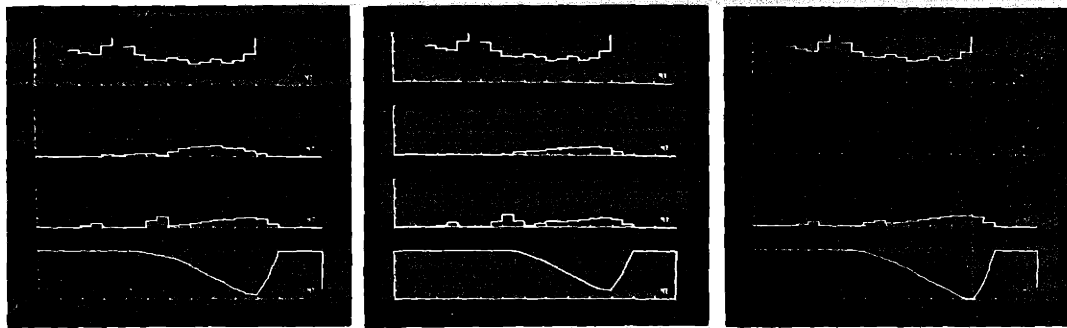
Case 1.  
Linear Cost Function

Case 2.  
Quadratic Cost Function

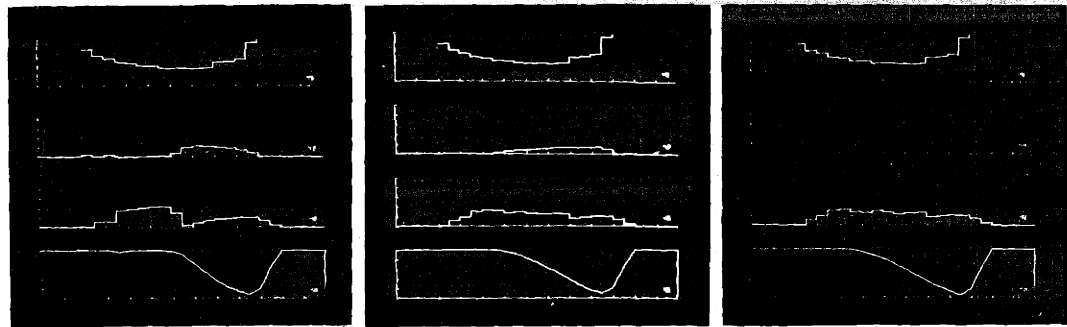
Case 3.  
Mule Curve



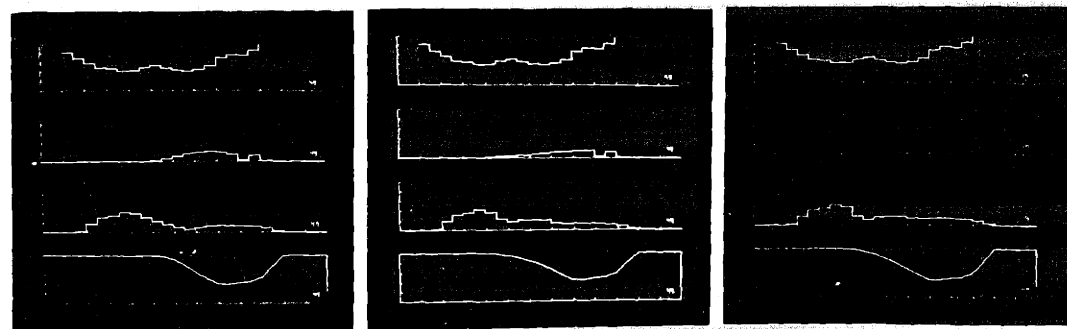
1946-47



47-48



48-49

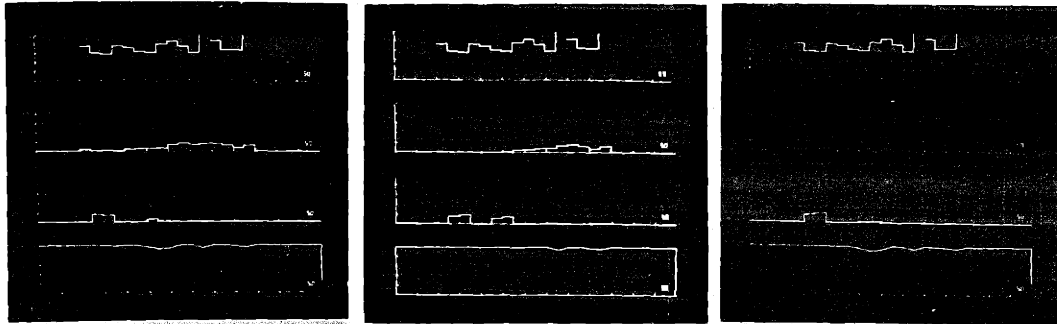


49-50

Case 1.  
Linear Cost Function

Case 2.  
Quadratic Cost Function

Case 3.  
Rule Curve



1950-51

51-52

52-53

Case 1.  
Linear Cost Function

Case 2.  
quadratic Cost Function

Case 3.  
rule Curve

LIST OF FIGURES AND TABLES

	<u>Page</u>
Fig. 1. Grand Coulee Dam, Columbia River, Washington.....	6
Fig. 2. Characteristics of the Columbia River at Grand Coulee: mean flow and cor- relations.....	53
Fig. 3. An example of a conditional probability density.....	55
Fig. 4. An example of curves for determining volume of stored water use.....	72
Fig. 5. Explanation of oscilloscope pictures.....	76
Fig. 6. Oscilloscope pictures of system behavior with historical flows:	
1914-15 through 1917-18.....	77
1918-19 through 1952-53.....	106
Fig. 7. Frequency histograms of $g$ , $g^2$ , and $(g_1)_{\max}$ from the 39 years of flow record for the three cases of operation.....	85
Fig. 8. The effect on $\overline{g^2}$ of departure from recommended water use.....	91
Table 1. The functions $d_i = d_i(v_i, x_i, s_i)$ and $v_{i+1} = v_{i+1}(v_i, x_i, s_i)$ .....	45
Table 2. A conditional probability table.....	54
Table 3. The 39 year averages of supplemental power and supplemental power squared for the three cases of operation.....	81
Table 4. The expected values of supplemental power and supplemental power squared as calculated from the probabilities.....	84

BIBLIOGRAPHY

- (1) W. G. Chandler, D. L. Dandeno, A. F. Glinn, and L. K. Kirchmayer, Trans. AIEE, Pwr. Apparatus Syst. No. 8, 1057-65 (Oct. 1953).
- (2) J. J. M. Carey, M. S. Thesis, Elec. Eng. Dept., M.I.T. (1953).
- (3) S. Stage, Control of Economic Loading of a Big Hydro and Thermal Power System, Swedish Water Power Assn. publication 400 (1948:3), translated by J. I. Holbeck, Bonneville Power Administration, Portland, Oregon, 1951.
- (4) A. Dvoretzky, J. Kiefer, and J. Wolfowitz, Econometrica 20, April and July (1952).
- (5) R. Bellman, J. Opns. Res. Soc. Am. 2, 275-288 (1954).
- (6) J. P. Jolliffe, Trans. AIEE, Pwr. Apparatus Syst. No. 9, 1243-8 (Dec. 1953).
- (7) R. J. Cypser, Sc. D. Thesis, Elec. Eng. Dept., M.I.T. (1953).
- (8) P. R. Johannessen, M.S. Thesis, Elec. Eng. Dept., M.I.T. (1953).
- (9) E. W. Lane, The Control of the TVA Storage Reservoirs for the Most Economical Power Production, Tennessee Valley Authority, Knoxville, Tennessee, 1944.
- (10) H. A. Simon and C. C. Holt, J. Opns. Res. Soc. Am. 2, 289-301 (1954).
- (11) W. Feller, An Introduction to Probability Theory and its Applications, John Wiley and Sons, Inc., New York, 1951, p. 78.
- (12) R. S. Burington, Handbook of Mathematical Tables and Formulas, Handbook Publishers, Inc., Ohio, 1945.

BIOGRAPHICAL NOTE

John D. C. Little was born Feb. 1, 1928 in Boston, grew up in Andover, Mass., and graduated from Phillips Academy in Andover in 1945. He entered MIT and in September 1948 graduated with an SB in physics. He then spent ten months hitchhiking through the country, a year working for General Electric Company and returned to MIT for graduate work in physics in 1950. He was a half time research assistant in physics with Project Lincoln, then with the Machine Computation Group, and finally with the Operations Research Group. He became a member of Sigma Xi in 1954.