University of Washington
Department of Civil and Environmental Engineering

WATER RESOURCES PLANNING AND MANAGEMENT: OPTIMIZATION MODELING OF RESERVOIR SYSTEMS

Mark Hedrich Houck
Bruce R. Cleland

Water Resources Series
Technical Report No. 54
October 1977

Seattle, Washington
98195

Department of Civil Engineering
University of Washington
Seattle, Washington 98195

WATER RESOURCES PLANNING AND MANAGEMENT: OPTIMIZATION MODELING OF RESERVOIR SYSTEMS

Mark Hedrich Houck
Bruce R. Cleland

Water Resources Series
Technical Report No. 54

October, 1977

Charles W. Harris Hydraulics Laboratory
Department of Civil Engineering
University of Washington
Seattle, Washington 98195

WATER RESOURCES MANAGEMENT AND PLANNING:
OPTIMIZATION MODELING OF MULTIPLE RESERVOIR SYSTEMS

by

Mark Hedrich Houck and Bruce R. Cleland

Technical Report No. 54

October 1977

Project Completion Report: Safe Yield and Policy Optimization Models

Principal Investigator: Mark Hedrich Houck, Research Assistant Professor of Civil Engineering, University of Washington

OWRT Project Number: A-086-WASH

OWRT Agreement Number: 14-34-0001-7102

Project Period: October 1, 1976 to September 30, 1977

TABLE OF CONTENTS

ABSTRACT

The class of optimization models which can be used as tools by the water resource planner or manager is examined. Five models are described and discussed in terms of their mathematical structure, computational burden or cost of solution, method of incorporating the stochastic nature of streamflows, ability to identify general optimal operating policies, and ability to identify optimal reservoir system designs. The five models are called: Multi-Year Deterministic Linear Program, Multi-Year Deterministic Dynamic Program, Explicitly Stochastic Linear Program, Explicitly Stochastic Dynamic Program, and Chance Constrained Stochastic Linear Program. The Explicitly Stochastic Dynamic Program and The Chance Constrained Stochastic Linear Program are superior to the other models in terms of general usefulness. The Chance Constrained Stochastic Linear Program is the best of the models in most planning situations.

Keywords: River Basin Planning*, Water Resources Management*, Systems Engineering, Mathematical Programming*, Optimization*, Stochastic Hydrology, Multiple Objective Programming, Decision Making, Yakima River.

ACKNOWLEDGEMENTS

The first chapter and appendix are primarily the work of the second author; the remaining chapters are based on research performed by the first author. Both authors are indebted to Professor Stephen J. Burges for his interest and questioning during the course of this project. His influence helped us refine and modify our understanding of the problems under investigation.

The authors are grateful for the assistance of many persons in preparing this report and carrying out the research upon which it is based. We extend our thanks to each of them and attribute to them any sage insights and passages found in this report; for ourselves, we reserve all credit for errors, contradictions and confusing passages.

The work upon which this report is based was supported in part by funds provided by the United States Department of the Interior, Office of Water Research and Technology as authorized under the Water Research Act of 1964 through the State of Washington Water Research Center.

1. INTRODUCTION AND THREE OPTIMIZATION MODELS

1.1 Outline of Chapters

Since 1962, when Maass, et al. first reported on water resources systems analysis, many different mathematical models have been developed. Optimization models which can be used to assist in managing and designing multi-purpose reservoir systems comprise one important subset of the mathematical models.

This report includes descriptions and comparisons of these optimization models. Three of the models are briefly discussed in the remaining sections of this chapter. The usefulness of each of these is limited by either immense computational costs of solution or an inability to include explicitly the stochastic nature of streamflows. Two other models, which are superior to those described in chapter one, are discussed in chapters two and three. An explicitly stochastic, dynamic program is the subject of chapter two and a chance-constrained, explicitly stochastic, linear program is the subject of chapter three. A summary of the previous chapters and conclusions are included in chapter four. Finally, a brief description of the problems associated with data handling, data analysis and model formulation is included in a special appendix.

1.2 Three Optimization Models

The three optimization models described in sections 1.3, 1.4 and 1.5 are called the Multi-Year Deterministic Linear Program, the Multi-Year Deterministic Dynamic Program, and the Explicitly Stochastic Linear program. Each of these models was formulated, constructed and solved (the Explicitly Stochastic Linear Program was not solved due to its prohibitively large solution cost) for a two reservoir system located in the Upper Yakima River Basin. The objective used for each model was: maximize the water supply potential of the reservoir system. The water supply potential is the ability of the reservoir system to provide consistently, over a period of time, a volume of water for downstream use with a specified reliability. Because each of the models has a different structure, the method of incorporating this objective is different for each one.

Even though a single purpose for the reservoir system is examined, the limitations of the models become apparent. The Multi-Year Deterministic Linear Program does not explicitly incorporate the stochastic properties of streamflows, does not directly generate general operating policies, and does become very expensive or computationally impossible to solve when large systems are considered. The Multi-Year Deterministic Dynamic Program exhibits these same characteristics and cannot be used to identify optimal designs. The Explicitly Stochastic Linear Program does explicitly incorporate the streamflow's stochastic properties and does provide general operating rules. However, unless major, possibly distorting, assumptions are made, this model's computational burden or cost of solution becomes prohibitively large when any more than two reservoirs are considered.

1.3 Multi-Year Deterministic Linear Program

The Multi-Year Deterministic Linear Program (M-YDLP) was the first optimizatic model examined during the course of this research project. The foundation of this model was presented by Maass, et al. (1962) and its formulation is perhaps the simplest of the optimization models. Watermeyer and Thomas (1962) described the continuity equation - a simple statement that the change in reservoir storage equals the difference between inflow to and outflow from the reservoir. If the change in reservoir stoı age is measured over some time period, the continuity equation would take the form:

$$S_{t+1} - S_t = I_t - R_t \qquad \forall t \qquad (1.1)$$

where S_t = reservoir storage at the beginning of time period t

S_{t+1} = reservoir storage at the beginning of time period t+1

I_t = inflow to the reservoir during time period t

R_t = release from the reservoir during time period t

Rearrangement of the continuity equation so that all variables are on the left hanı side yields the following:

$$S_{t+1} - S_t + R_t = I_t \qquad \forall t \qquad (1.2)$$

Watermeyer and Thomas (1962) recognized that by choosing some representative sequence of inflows (I_t for $t = 1,\ldots,T$) the continuity equation could become a portion of a linear program. By using a sequence of flows, several years long, the variation and stochasticity of the streamflows are implicitly considered in the M-YDLP model. Additional restrictions on the possible values for storage are needed to assure a realistic solution to the problem. Each storage level, S_t (for $t = 1,\ldots,T$), is required not to exceed the dam capacity, V, and to remain non-negative. A similar non-negativity restriction is necessary for each release, R_t (for $t = 1,\ldots T$). The constraints stated thus far simply describe the functioning of a simple reservoir for a specific sequence of flows.

The development of an objective function is the remaining requirement. The definition of water supply potential that proved to be most suitable for the M-YDLP model was the minimum flow below a system of dams which is met in all time periods of interest or:

$$R_t \geq MINR \qquad \forall t \tag{1.3}$$

where MINR = release from the reservoir which can be met in all time periods. The objective of the M-YDLP is to find the greatest value of this minimum release: hence, the objective function is Maximize MINR.

The area selected for testing the models discussed in this chapter was located in the Upper Yakima River and consisted of two reservoir sites, Lake Keechelus (site 1) and Lake Kachess (site 2). When multiple structures are considered, it is necessary to include the continuity equations for each potential structural site in the model. When each structure is located so that the outflow from one does not subsequently enter another, it is only necessary to include constraint 1.2 for each time period and each site. Using the Upper Yakima River as an example, it is possible to demonstrate the range of forms needed to consider multiple sites.

By inserting an initial subscript s, which designates the appropriate site number, the continuity equations for the area under study are:

$$S_{1t+1} - S_{1t} + R_{1t} = I_{1t} \quad \forall t \tag{1.4}$$

$$S_{2t+1} - S_{2t} + R_{2t} = I_{2t} \quad \forall t \tag{1.5}$$

The values of the inflows (I_{st}) or the sequence of flows which are used in the model are of critical importance. The sequence should fully represent the stochasticity of the actual streamflows but because only a limited number of time periods can be included in the model, this is impossible. If extreme events are included in the sequence, they may lead to reservoir operations which are based on flow variability which is greater than actually occurs. The avoidance of extreme events in the sequence, however, may tend to produce the opposite result. Possibly the best way to overcome these difficulties is to model as many years as are economically or computationally feasible.

In addition to the continuity equations other restrictions on the variables are needed. All of the variables must be restricted to non-negative values and the storage at any time must not be allowed to exceed the dam capacity, V_s.

$$S_{1t} \leq V_1 \quad \forall t \tag{1.6}$$

$$S_{2t} \leq V_2 \quad \forall t \tag{1.7}$$

In order to assure that at the end of the time period modeled (end of period T), the water resource is not in a better or worse condition than at the beginning of time period one, another constraint can be added to the linear program.

$$S_{11} - S_{1T} = 0 \tag{1.8}$$

$$S_{21} - S_{2T} = 0 \tag{1.9}$$

This assures that the storage at each site at the end of the last time period equals the storage at the site at the beginning of period one. By setting t+1 equal to 1 when t = T in constraints 1.4 and 1.5, the desired result is obtained. Finally, the combined releases of all sites must be equal to or greater than the minimum release of the system of reservoirs in all time periods.

$$(R_{1t} + R_{2t}) - MINR \geq 0 \quad \forall t \tag{1.10}$$

The method of obtaining operating policies is described by Young (1967). The solution of the linear program designates the optimal releases from a system of reservoirs for some length of time. Assuming that the stochasticity of the streamflow process is implicit in the sequence of streamflows used in the model, it is possible that a significant relationship can be found between the optimal releases and the streamflows. In fact, a relationship may be found between the optimal releases, past storage volumes, the season of the year, and streamflows. Young (1967) recommends that the relationship be found through a regression analysis where release is the dependent variable and previous streamflows, past storages and season of the year are included in the set of independent variables. This approach was instrumental in extending the power of the M-YDLP model even though Young did not base his work on a linear program. He introduced the regression method of obtaining operating rules from the solution of a multi-year deterministic dynamic programming model.

The formulation of the M-YDLP may be summarized as follows:

$$\text{Maximize} \quad \text{MINR} \tag{1.11}$$

$$\text{subject to:} \quad S_{st+1} - S_{st} + R_{st} = I_{st} \qquad \forall s, \forall t \tag{1.12}$$

$$S_{s1} - S_{sT} = 0 \qquad \forall s \tag{1.13}$$

$$S_{st} \leq V_s \qquad \forall s, \forall t \tag{1.14}$$

$$\sum_{s=1}^{m} R_{st} - \text{MINR} \geq 0 \qquad \forall t \tag{1.15}$$

$$\text{all variables} \geq 0$$

The size of the M-YDLP model is sensitive to the number of sites and time periods which are modeled. The number of constraints in the M-YDLP model (as described above) is:

$$T * (2m + 1) + m = (\ell * Y)(2 * m + 1) + m$$

where m = number of reservoir sites in the model

ℓ = number of seasons per year

Y = number of years modeled

T = number of time periods used for model analysis $(\ell * Y)$

After experimenting with the various linear programming packages available at the University of Washington Computer Center, it was found that MPOS (Multi-Purpose Optimization System) offered the best software to solve the M-YDLP optimization model. However, the maximum number of constraint equations which could be input to the program was 250: this required the full amount of central memory availabe on the CDC 6400 computer. (Other linear programming packages available at the University of Washington were only able to accept a maximum of 100 constraint equations.) Utilizing MPOS, it was found that equation 1.14 did not count in the constraint set, but fell into a category called boundary equations for this particular software package. Thus, under MPOS, the relationship which expresses the number of constraints in the M-YDLP model becomes:

$$T * (m + 1) + m = (\ell * Y)\ (m + 1) + m$$

Because 250 constraint equations were limiting for a 2-site M-YDLP analysis, mathematical substitutions were used wherever possible to reduce the number of constraint equations. For example, equation 1.10 can also be expressed as:

$$R_{1t} + R_{2t} = MINR + IR_t \quad \forall t \tag{1.16}$$

where IR_t = the amount by which the total reservoir release exceeds the safe yield (MINR) for time period t.

This constraint equation can then be eliminated by solving for R_{1t} and substituting the result into equation 1.4 yielding:

$$S_{1t+1} - S_{1t} + MINR + IR_t - R_{2t} = I_{1t} \quad \forall t \tag{1.17}$$

An attempt was also made to solve equation 1.17 for R_{2t} and to substitute the result into equation 1.5 giving a further reduction in the number of constraint equations. However, performing this step was found to cause a violation in continuity at one of the reservoir sites because individual inflows at each site were also combined. When the reservoir storage levels which were output from the M-YDLP were substituted back into the continuity equations for each site to determine the recommended reservoir release, negative values were found to result. Thus, by performing simple mathematical substitutions to reduce the number of constraint equations, the non-negativity required for the release variables was also eliminated. Therefore, the final reduced form of the M-YDLP for a 2-site reservoir system is:

$$\text{Maximize MINR} \tag{1.18}$$

$$\text{subject to:} \quad S_{1t+1} - S_{1t} + \text{MINR} + IR_t - R_{2t} = I_{1t} \qquad \forall t \tag{1.19}$$

$$S_{2t+1} - S_{2t} + R_{2t} = I_{2t} \qquad \forall t \tag{1.20}$$

$$S_{11} - S_{1T} = 0 \tag{1.21}$$

$$S_{21} - S_{2T} = 0 \tag{1.22}$$

$$S_{1t} \leq V_1 \qquad \forall t \tag{1.23}$$

$$S_{2t} \leq V_2 \qquad \forall t \tag{1.24}$$

$$\text{all variables} \geq 0$$

And the relationship which expresses the number of constraints in the M-YDLP model is:

$$m * (T + 1) = m * [(\ell * Y) + 1]$$

The next problem addressed was the determination of the length of time intervals to be used in the M-YDLP analysis. The reservoir sites studied in the Upper Yakima River are used primarily for irrigation over a seven month period from April to October. It was sought to achieve an operating policy that is consistent with feasible management of the present system. Crucial factors in attaining consistency are the length of the time interval and the degree of knowledge of future hydrologic events. A 50-year historical record of monthly flows provided by the Bureau of Reclamation was examined to determine the length of the time interval which would yield the most meaningful results from the optimization models. It was found, as a result of both the irrigation usage of the system and the seasonal variability of system inflows, that monthly time intervals during the irrigation season offered the best hope for meaningful results. The remaining five months were lumped into one time interval in order to reduce the number of constraint equations required for the model and because no irrigation demand exists on the system during this period. Hence, for a 2-site system and 8 seasons per year, the maximum number of years which could be studied in any particular M-YDLP analysis using MPOS on the CDC 6400 computer (250 constraint equations maximum) was found to be fifteen.

As stated previously, the sequence of inflows used in the model should fully represent the stochasticity of the actual streamflows. The computational limitations of the M-YDLP model did not allow the luxury of analyzing a sequence of flows longer than 15 years in any one run without making other sacrifices such as reducing the number of seasons per year. By using 15 year sequences, the existence of extreme events in a particular trace could tend to lead to operating policies based on flow variability which is greater than actually occurs. For this particular study, the problem could have been overcome for a 2-site reservoir system with a better linear programming package on a larger computer (e.g., the IBM 370 Mathematical Programming System can handle in excess of 1000 constraint equations). However, even with the best available computer technology, the computational burden on the M-YDLP becomes almost insurmountable as the number of reservoir sites under study is increased, as the length of the period of analysis becomes greater, and as more reservoir functions are examined.

1.4 Multi-Year Deterministic Dynamic Program

The Multi-Year Deterministic Dynamic Programming (M-YDDP) model is useful strictly as a policy model: it can be used to identify operating policies but not system designs. It is assumed that the stochasticity of natural streamflows is implicit in the sequence of flows used in the model. The output of the model is a set of optimal decisions (releases) and a set of optimal storage levels associated with the streamflow sequence input to the model. Operating rules can be obtained by performing a regression analysis with the releases as the dependent variable and previous storage levels, streamflows, and the season of the year as independent variables.

Young (1967) introduced the regression method of obtaining operating rules and demonstrated its power. He also used the dynamic programming model whuch is here called the M-YDDP but which he referred to as the forward-looking deterministic algorithm. A Monte Carlo technique was employed by Young (1967) to obtain sufficient data for the regression analysis. A method of generating sequences of streamflows which exhibited the same statistical properties as the historical streamflows (operational or synthetic hydrology) was used to provide sets of flows each of which was considered to be equally likely. The dynamic program was solved for each set of flows and the results from all of the solutions were used in the regression analysis.

Buras (1966), Hall and Dracup (1970), and Becker and Yeh (1974) provided discussions on the use of dynamic programming models in the development of water resources. In particular, they examined the dynamic program with deterministic inputs for single and multiple dam sites. Multiple uses of the dam system were also considered and possible methods for quantifying economic benefits and losses associated with the dam system operations were discussed.

The bases of the dynamic program are the recursion relation and the return function. The stages of the M-YDDP model are time periods (the same 8 seasons outlined in the discussion of the M-YDLP model for the Upper Yakima System) and the states are storage volumes. The storage capacity at each dam in the system is broken into as many discrete values as desired. The larger the number of discrete storage volumes which are considered, the greater will be the accuracy of the model. The state of the dam system is described by the discrete storage levels at all sites. If S_{st} represents the potential discrete storage volumes at site s at the beginning of time period t then $(S_{1t}, S_{2t}, \ldots, S_{st})$ represents the state of the system at the beginning of period t.

Movement from one system state at the beginning of period t to another system state at the beginning of time period t+1 is controlled by the continuity equation used in the M-YDLP model. Continuity is a statement that the change in storage in a reservoir over a time period equals the difference between inflow and outflow during the time period. If R_{st} equals the release from site s during period t and I_{st} equals the inflow to site s during period t, then the continuity equation is:

$$R_{st} = S_{st} + I_{st} - S_{st+1} \tag{1.25}$$

Therefore, if inflow and beginning and ending storages for any time period are known, the release can be determined.

The development of the cost function used in the recursion relation centered on whether or not the target release, MINR, is met. The cost function, C_{ij}, equals the cost of moving from the discrete storage increment S_i in time period t to storage increment S_j in time period t+1.

$$C_{ij} = \begin{cases} 0 \text{ if } R_{ij} = S_i + I_t - S_j \geq \text{MINR} \\ 1 \text{ if } R_{ij} = S_i + I_t - S_j < \text{MINR} \end{cases} \tag{1.26}$$

More simple stated, the cost function is the number of times the target release, MINR, is missed. The function $f_t(S_i)$ equals the number of times the target release has been missed during t time periods of operation when the storage at the beginning of season t is S_i. Thus, the recursion relation for the M-YDDP becomes:

$$f_{t+1}\ (S_j) = \min_i\ \{ C_{ij} + f_t\ (S_i)\} \tag{1.27}$$

The boundary condition for the dynamic program is $f_o(S_i)$ equals zero for all storage volumes S_i. Solution of the recursion relation results in the minimization of the number of times that the target release is not met.

Unlike the M-YDLP, it was found that the definition of water supply potential used in the M-YDDP model could be expanded to be the minimum flow below a system of dams which is met α percentage of the time; thus, introducing a system reliability factor. For any value of MINR, find the ratio of the minimum number of misses in T seasons and T. This ratio is an estimate of the probability of missing the target.

Utilizing the M-YDDP to find the optimal operating policy assumed prior knowledge of the target release, MINR. To find an appropriate value for MINR, therefore, required the use of a search procedure in addition to solving the recursion relation. To use the M-YDDP to find the appropriate target release, a binary search technique was employed. By performing five iterations of the search technique for each 50 year sequence of inflows which was examined, 2^5 or 32 increments or values of MINR could be tested. The initial monthly target release, MINR, which was analyzed was determined by dividing the combined mean annual inflows to both Lake Keechelus and Lake Kachess by the number of months the system was used for irrigation. To determine the range over which the search would be performed, the coefficient of variation which is a measure of dispersion was utilized. Let

$$IT = \text{initial target examined} = \frac{\text{combined mean annual flow}}{\text{no. of months irrigated}}$$

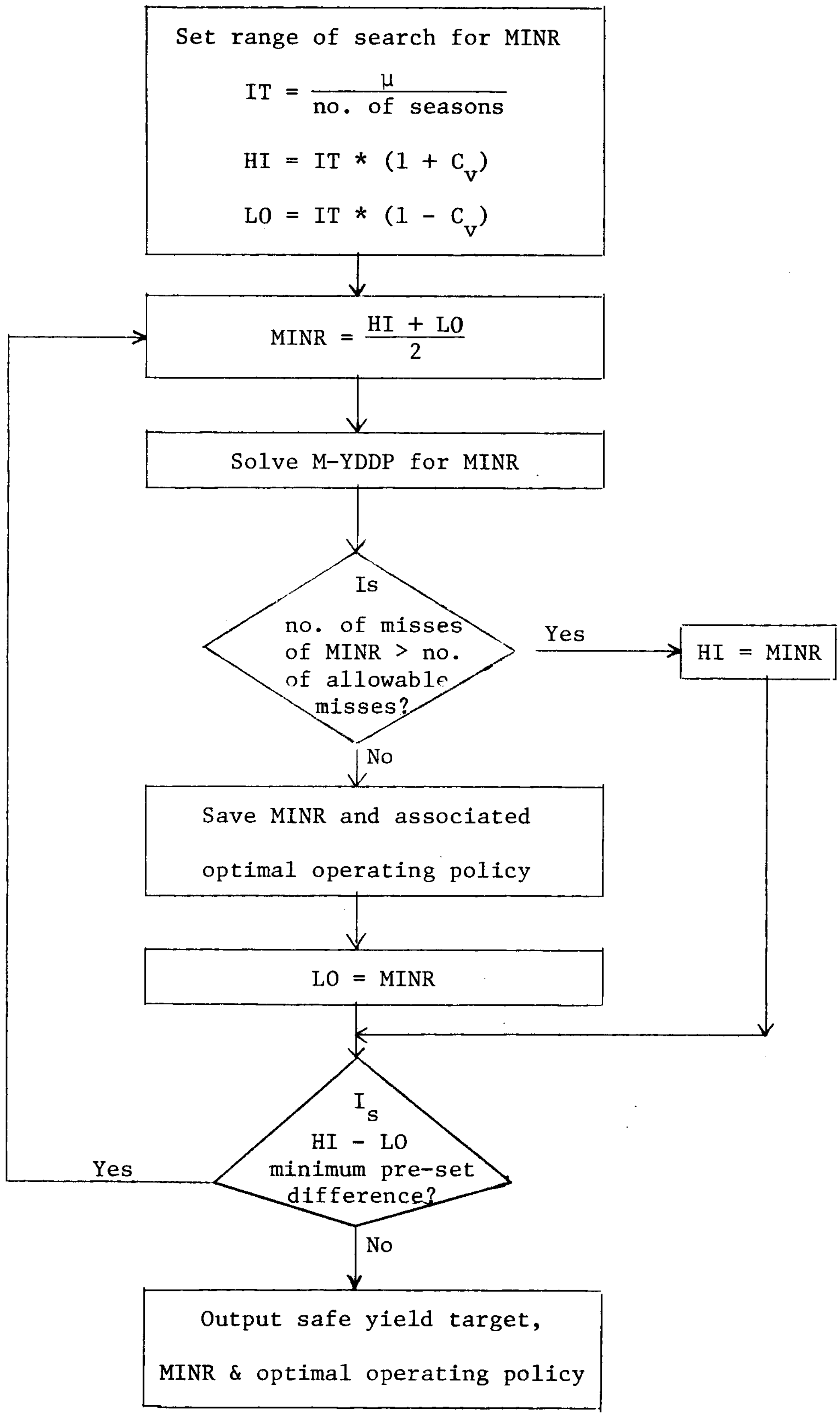
$$C_V = \text{coefficient of variation} = \frac{\text{standard deviation}}{\text{mean annual flow}}$$

The range examined was IT $\pm$ (C_V*IT). For all 50-year traces analyzed, the optimum target for the M-YDDP model was found to fall within this range.

The flow chart on page 12 outlines the method by which the binary search was utilized with the M-YDDP model to determine the water supply potential of the 2-site reservoir system.

The computational burden associated with the M-YDDP model is dependent on the number of dam sites, the number of discrete storage volume increments that are considered for each reservoir, and the number of time periods. The number of times that the recursion relation must be solved is approximately equal to the number of stages multiplied by the square of the number of states. If ℓ is the number of seasons per year, Y is the number of years, m is the number of dam sites and d is the number of discrete storage levels at each dam. Then the recursion relation must be solved $d^{2m} * \ell * Y$ times. As an example of how sensitive the M-YDDP is to the number of sites with respect to required computer time for model solution, for one-site (8 seasons per year, 50-years, 5 storage increments), the recursion relation must be solved 10,000 times. All else held constant (8 seasons per year, 50 years, 5 storage increments), the 2-site recursion relation must be solved 250,000 times. Obviously, for a water resource system larger than the two sites, use of the M-YDDP can lead to exorbitant costs.

Another aspect of the solution costs of the M-YDDP model is evident from the recursion relation. In order to sequentially solve the recursion relation, immense amounts of computer storage may be required. One method which can be used


Set range of search for MINR
IT = μ / no. of seasons
HI = IT * (1 + C_v)
LO = IT * (1 - C_v)
MINR = (HI + LO) / 2
Solve M-YDDP for MINR
Is no. of misses of MINR > no. of allowable misses?
Yes
HI = MINR
No
Save MINR and associated optimal operating policy
LO = MINR
Is HI - LO minimum pre-set difference?
Yes
No
Output safe yield target, MINR & optimal operating policy

to retain the optimal solution is to store the values of $f_t(S)$ for all storage volumes, for all time periods and for all sites. This means storing $(d)^m * \ell * Y$ values and does not include the storage requirements of the model code or the cost function. As an example of storage requirements with respect to the number of sites, the 1-site example outlined above requires 2000 words of central memory. The same 2-site case requires 20,000 words or ten times the amount of computer storage. So the M-YDDP model increases in storage and computational costs as the number of sites, the number of time periods and the number of discrete storage volume increments increase.

In contast to the M-YDLP, the M-YDDP allows more flexibility in the definition of water supply potential as well as providing the capability of examining longer sequences of inflows which can lead to more meaningful results. However, dividing each reservoir into discrete storage increments does lead to some loss in accuracy. As with the M-YDLP, the largest drawbacks of the M-YDDP are the non-explicit incorporation of the streamflows' stochastic properties, the indirect method of obtaining general operating rules and the large computational burden and storage requirements. It was found that mathematical substitutions and modulus arithmetic could be utilized to reduce these needs, but the subsequent gains were more than offset by the non-linear increase of the computer requirements for the multi-site cases.

1.5 Explicitly Stochastic Linear Program

The Explicitly Stochastic Linear Programming (ESLP) model is a subset of a nonlinear program that can be used as a design and/or a policy model and utilizes a discrete Markov process as the explicit stochastic description of the streamflow. The nonlinear program is structured in a special way so that if a partial design is specified, the remaining variables and equations form a linear program which can be used as a policy model. If a policy is specified, the remaining variables and equations in the nonlinear program form a linear program which can be used as a design model.

The design and policy models can be used separately or in tandem. When the models are used together, an iterative procedure is employed to find an approximation of the optimal solution of the nonlinear program. The procedure is to alternate between solving each of the linear programs, using the solution of one as input to the following one. Because the value of the objective for each model is nondecreasing during this alternation process and the objectives are bounded, the process terminates (no change in the objective function values) in a finite number of iterations. The termination point may be the global optimum of the nonlinear program in which case the optimal design and policy are found or the termination point may be a local (non-global) optimum. There is no method to determine whether the termination point is the global optimum or a local optimum. And if a local optimum is obtained, there is no method of determining how far from the global optimum it is

The historical development as well as extensions and improvements of the ESLP model are fully discussed by Loucks (1969), Gablinger and Loucks (1970), Loucks and Falkson (1970), Jacoby and Loucks (1972), and Houck and Cohon (1977). Loucks (1969) also pointed out two major difficulties of this type of model. As the complexity and scale of the system of potential sites increase, the data available become totally insufficient to estimate the parameters needed, and the size of the model grows so large that the model is unsolvable by present computer methods.

A discrete Markov process is used in the ESLP to describe the streamflow process. This requires that the continuum of streamflows at each site be broken into intervals and a characteristic flow for each interval be chosen to represent all flows in the interval. This discretization process is performed for all seasons of the year as well as for each site. For the Upper Yakima River System, the cumulative distribution functions (CDF) defined by historical streamflow data were broken into five intervals and the average flow within each interval of the CDF was used to represent the characteristic flow of that interval.

The ranges associated with each characteristic flow are important in defining the lag-one Markov process that will be used to describe explicitly the stochasticity of streamflows. It is necessary to determine the conditional probabilities of transitions from each of the characteristic flows in one season to each characteristic flow in the next season. The transition probability is a measure of the likelihood that a flow will follow a given flow in the previous season. It can be approximated by computing the relative frequency with which this occurs in the historical data or by using the historical data to calibrate a particular probability distribution function. For the Upper Yakima study, sequences of streamflows which exhibited the same statistical properties as the 50-year historical streamflows were generated from the data provided by the Bureau of Reclamation. These generated flows as well as the historical data were used to compute the relative frequencies.

P_{ij}^{t} will represent the probability of a transition to the j^{th} characteristic flow in season t+1, $I_{j,t+1}$, from the i^{th} characteristic flow in season t, I_{it}. (P_{ij}^{t} = Prob $[I_{j,t+1} \mid I_{it}]$). If two dam sites are under consideration and one is denoted by primed subscripts, then $P_{ii'jj'}^{t}$ represents the probability of a transition to the characteristic flows $I_{j,t+1}$ and $I_{j',t+1}$ at both sites in season t+1 from the characteristic flows, I_{it} and $I_{i't}$ in season t.

Storage volume is also intervalized and characteristic storage volumes are chosen to represent each interval or range of volumes. S_{kt} and $S_{k't}$ will represent the k^{th} and k'^{th} characteristic storage volumes at the beginning of season t at the two dam sites. Similarly, $\ell(\ell')$ and m(m') will represent the characteristic storage intervals at the beginning of seasons t+1 and t+2 respectively. Releases or drafts are broken into intervals based on beginning and ending storages of the period and flow during the period. Thus, there will be a release or draft for each combination of possible beginning storage, possible ending storage, and possible flow for each season t. The reservoir releases or drafts will be represented by $R_{ki\ell t}$ which is the release in season t resulting from a beginning storage S_{kt}, an ending storage $S_{\ell t+1}$, and a streamflow, I_{it}.

The formulation of the nonlinear program consists of continuity constraints,

constraints which relate transition and marginal probabilities and which limit the values of the probabilities, definitional constraints for the water supply potential or minimum target release, constraints which define the allowable values for storages, and an objective function. The development of the model is presented in this order.

The continuity equation is a mass balance on water at each dam site.

$$S_{\ell,t+1} = S_{kt} + I_{it} - R_{ki\ell t} \qquad \forall\ k,i,\ell,t \tag{1.28}$$

$$S_{\ell',t+1} = S_{k't} + I_{i't} - R_{k'i'\ell't} \qquad \forall\ k',i',\ell',t \tag{1.29}$$

This equation must hold for each initial and final storage volume, each streamflow, each season, and each site.

The drafts, $R_{ki\ell t}$ and $R_{k'i'\ell't}$, are the important variables as far as the water supply potential of the planning problem is concerned. The releases determine the extent to which the targets are met. In order to compute the minimum release, MINR, the probability of each draft must be found. The transition probabilities, $P^t_{ii'jj'}$ are needed to define the values of $P_{kk'ii'\ell\ell't}$, the probability that the release $(R_{ki\ell t} + R_{k'i'\ell't})$ occurs. Because streamflows from season to season are related through the transition probabilities, and releases are functions of streamflows, the release from season to season and the values of $P_{kk'ii'\ell\ell't}$ and $P_{\ell\ell'jj'mm't+1}$ (the probability that a total draft, $R_{\ell jmt+1} + R_{\ell'j'm't+1}$ is made) are related.

$$\sum_{k}\sum_{k'}\sum_{i}\sum_{i'} P_{kk'ii'\ell\ell't}\,(P^t_{ii'jj'}) = \sum_{m}\sum_{m'} P_{\ell\ell'jj'mm't+1}$$

$$\forall \ell,\ell',j,j',t \tag{1.30}$$

The right hand side of equation 1.30 equals the probability that inflows I_{jt+1} and $I_{j't+1}$ occur during season t+1 and the storages at the beginning of season t+1 are $S_{\ell t+1}$ and $S_{\ell't+1}$. By summing over all storage levels k and k' on the left hand side of equation 1.30, it becomes:

$$\sum_{i}\sum_{i'} \text{Prob}[S_{\ell t+1} \text{ and } S_{\ell' t+1} \text{ and } I_{it} \text{ and } I_{i't}]*$$

$$\text{Prob}[I_{jt+1} \text{ and } I_{j't+1} \mid I_{it} \text{ and } I_{i't}]$$

By summing over all values of i and i' the left hand side becomes the probability of beginning season t+1 with storage volumes of $S_{\ell t+1}$ and $S_{\ell' t+1}$, and having inflows of I_{jt+1} and $I_{j't+1}$. Thus, equation 1.30 assures that the discrete, lag-one Markov process is incorporated in the ESLP model.

Further constraints are necessary to restrict the set of joint probabilities. It is necessary to assure that the probability of some combination of releases occurring equals one in each season and that all probabilities are non-negative.

$$\sum_{k}\sum_{k'}\sum_{i}\sum_{i'}\sum_{\ell}\sum_{\ell'} P_{kk'ii'\ell\ell't} = 1.0 \qquad \forall t \tag{1.31}$$

$$P_{kk'ii'\ell\ell't} \geq 0 \qquad \forall kk'ii'\ell\ell't \tag{1.32}$$

The water supply potential can be measured in several ways: one of these is to assure that the minimum release or target release, MINR, is always met. In other models, this restriction has been formulated by simply writing a constraint, for each possible release, that specifies MINR is less than or equal to that release. It becomes necessary in the ESLP model to include also the probability of occurrence of each possible release in these constraints. If the probability of a particular release equals zero, it is not necessary for that release to exceed or equal MINR.

$$P_{kk'ii'\ell\ell't} * (R_{ki\ell t} + R_{k'i'\ell't} - \text{MINR}) \geq 0$$

$$\forall k,k',i,i',\ell,\ell',t \tag{1.33}$$

The final constraints which are necessary define the restrictions on storage volumes. No storage may exceed the dam capacity and no storage may take on negative values. This is accomplished by restricting the storage levels to be non-decreasing with increases in indices k or k'. The largest storage is restricted to values not greater than the dam capacity and all storage volumes are restricted to values not less than zero.

$$S_{kt} \geq 0 \qquad \forall\ k,t \tag{1.34}$$

$$S_{k't} \geq 0 \qquad \forall\ k',t \tag{1.35}$$

$$S_{kt} - S_{k+1,t} \leq 0 \qquad k = 1,2,\ldots,K-1 \quad \forall t \tag{1.36}$$

$$S_{k't} - S_{k'+1,t} \leq 0 \qquad k' = 1,2,\ldots,K'-1 \ \forall t \tag{1.37}$$

$$S_{Kt} - V_1 \leq 0 \quad \forall t \tag{1.38}$$

$$S_{K't} - V_2 \leq 0 \quad \forall t \tag{1.39}$$

The objective function is: Maximize MINR.

The nonlinear program consists of constraints 1.28 through 1.39, an objective of Maximize MINR, and non-negativity restrictions on all variables. The linear program which is the design model is formed by specifying values for $P_{kk'ii'\ell\ell't}$ such that constraints 1.30, 1.31 and 1.32 are satisfied. The remaining constraints, variables, and the objective function define a linear program.

The policy model, a linear program, is formed by specifying values of $R_{ki\ell+}$, $R_{k'i'\ell't}$ and MINR as obtained by solution of the design model. The remaining constraints, variables and objective of the nonlinear program define a linear program. Unfortunately, the objective value will not increase because MINR is assigned a specific value. However, it is possible to obtain different values for the probabilities by using an objective such as:

$$\text{Maximize} \sum_{t}\sum_{k}\sum_{k'}\sum_{i}\sum_{i'}\sum_{\ell}\sum_{\ell'} (k+k'+i+i'-\ell-\ell')P_{kk'ii'\ell\ell't} \ .$$

The result will be to maximize the sum of weighted probabilities, where the weights favor large releases and penalize small releases. This particular objective is only used as an example to demonstrate a method of obtaining a possibly superior solution. With a different set of values for the probabilities, the next design model may result in a higher value for MINR. (For a much more detailed discussion of how the design and policy models can be formed and sequentially solved, see Houck and Cohon (1977)).

The operating rules are conditional probabilities defined by the joint probabilities specified in the solution of either linear program. If the storages at the beginning of season t are known and the streamflows for season t are known or predicted, then the decisions to be made are how much water to release or what storage volume to strive for at the beginning of the next season at each dam. These decisions are equivalent because storage and release are related through the continuity constraint. The probability of being in storage intervals ℓ and ℓ' at the beginning of season t+1 given initial storage intervals k and k' and flow in intervals i and i' for season t is a statement of the operating rule. This probability $P[\ell\ell' | kk'ii't]$ is related to the release probabilities found in the program by:

$$P[\ell\ell' | kk'ii't] = \frac{P_{kk'ii'\ell\ell't}}{\sum\sum_{\ell\ell'} P_{kk'ii'\ell\ell't}} \tag{1.40}$$

The cost of solution or the computational burden of the policy and design models is large and can be estimated by the numbers of constraints in the linear programs. The design model for the 2-site, Upper Yakima River System consists of 127,080 constraints. Many of these constraints can be eliminated because many of the probabilities $P_{kk'ii'\ell\ell't}$ will equal zero. If a particular set of them equals zero then the events described in constraints 1.28, 1.29 and 1.33 may never occur: hence, it is not necessary to include those constraints in the linear program. It would be possible to reduce the number of constraints to 7000 or fewer. For a three dam system the number of constraints is much higher. Assuming that as many of the probabilities as possible equal zero, the number of constraints exceeds 100,000. The 3-site design model is unsolvable with current computer technology and the 2-site model is at the limits of solvability. So even for very simple planning problems the ESLP model has excessive computational requirements.

2. EXPLICITLY STOCHASTIC DYNAMIC PROGRAMMING MODEL

2.1 Introduction

In the course of this research project, the Explicitly Stochastic Dynamic Programming Model which is discussed in this chapter was found to be superior in several ways to the optimization models described in chapter 1. Therefore, more testing was performed on this model than the previous ones. Moreover, an example case study utilizing the dynamic program as well as multi-objective display techniques was completed. The objective of the study was to determine the impacts of increased hydroelectric energy production on overall reservoir system operation.

The objectives of this chapter are to describe the dynamic program structure, and to show how multi-objective analysis can be used in a water resource planning problem. A short review of the use of optimization modeling in the management of reservoir systems will be presented along with a description of the Explicitly Stochastic Dynamic Program. This will include a discussion of different uses of a system and the management conflicts involved. A description of the example reservoir system's configuration, streamflow characteristics and economic benefit and loss functions will follow. Finally, a discussion of the impacts of increased energy demands on the example reservoir system's operation will be presented.

The example reservoir system comprises two reservoirs situated on the Lehigh River System in Pennsylvania. The reservoirs are multiple purpose in that they are used for flood control, water supply and power production. This reservoir system was chosen as the example because it is managed to meet multiple and conflicting demands, this system design was selected by a screening model (a chance constrained linear programming model patterned on ReVelle, et al. (1969)) as the optimal design for a reservoir system in this river basin, and substantial economic data on costs and benefits of all uses of the system are readily available from Hufschmidt and Fiering (1966).

2.2 Related Work

Maass, et al. (1962) first reported on water resources systems analysis.

They discussed the multiple uses of reservoir systems which include flood control, water supply for domestic, industrial and agricultural purposes, recreation, and hydroelectric energy production, and the conflicts between the uses. For example, if the purpose of a reservoir system is to provide a dependable supply of water to a downstream site, then the optimal operating policy would be to keep the reservoirs as filled as possible while releasing sufficient water to meet demand. If the purpose of a reservoir system is to control downstream flooding, then the optimal operating policy would be to keep the reservoirs as empty as possible so that the maximum storage will be available for catching any flood that occurs.

Unlike the single uses of flood control and water supply which have obvious optimal reservoir operating policies (at least for single reservoirs) hydroelectric power production does not. Hydroelectric power production is a function of the head acting on the turbines (depth of water above the turbines) and the rate of flow through the turbines (Hufschmidt and Fiering (1966) and Doland (1954)). Because the streamflows entering a reservoir vary over time it is necessary to vary the head and release from the reservoir in order to meet a relatively constant power demand. It is not obvious how this should be done to obtain a maximum power output from the reservoir.

It is even less evident how a reservoir or system of reservoirs should be operated if multiple uses of the system are demanded. Maass, et al. (1962) described the first use of an optimization model, a linear program, to find the optimal operating rules for a multipurpose reservoir. An economic benefit function was derived for each use of the reservoir and an objective function of maximization of net economic benefits was employed. The streamflow description was similar to that used in the M-YDLP model.

Since this first modeling effort, many optimization models for determining both design and operating rules for multipurpose, multiple reservoir systems have been proposed and discussed in the literature: three of these have already been discussed in chapter one.

The set of optimization models can be divided into two subsets: linear programs and dynamic programs. The linear programming models are restricted because nonlinearities, especially the nonlinear relation of hydroelectric power production to reservoir head and release, can be extremely expensive to approximate within the linear programming format. Although the dynamic programs can easily accomodate these nonlinearities they are constrained because certain management decisions must be made before the model is solved (e.g., targeted demand levels must be specified). Hence, there exists a tradeoff between the linear and dynamic programming models.

Another categorization of the set of optimization models is defined by the method used to incorporate the stochastic properties of the streamflows. Several models use artificial or synthetic sequences of streamflow based on the historical streamflow record (e.g., M-YDLP and M-YDDP). The model is solved for a particular sequence of streamflows so that the operating rules which are determined are dependent on the streamflow sequence: no general operating rules are determined.

The remaining models incorporate the stochasticity of the streamflows by directly using their probability distribution functions. Butcher (1971) and Loucks (1969) explicitly use a discrete, lag-one, Markov process to describe streamflows' stochasticity. ReVelle, et al., (1969) make direct use of the cumulative distribution functions of seasonal streamflows in the "Linear Decision Rule" model. Because none of these models is based on a particular streamflow sequence, each of them yields general operating rules for reservoir systems.

Hydroelectric power production cannot be examined without also considering other forms of electric energy production. Approximately fifteen percent of this nation's electric energy is currently supplied by hydropower; fossil fuel and nuclear power plants supply most of the remaining eighty-five percent (Armstrong, 1976).

Energy Alternatives: A Comparative Analysis, prepared by The Science and Public Policy Program of the University of Oklahoma in 1975, contains a detailed description of the potential energy sources and methods of electric energy production which will be used in the foreseeable future. One important characteristic of fossil fuel and nuclear power plants is that they operate most efficiently at a constant output level. These plants cannot adjust their output levels over short periods without significantly reducing efficiency. Hydroelectric power production can, however, vary substantially without a significant efficiency loss: each turbine can be turned on or off easily. Because electric energy demands exhibit daily variations; hydroelectric, fossil fuel and nuclear power plants can be operated systematically to great advantage. The thermal plants operate at constant output levels at maximum efficiencies and the hydroelectric power plants operate to supply all additional electric energy demands.

Caselton and Russell (1976) have explored the impact that a reservoir used for hydroelectric energy production and a thermal power plant can have on each other. They limited their investigation to a single thermal power plant and one,single-use reservoir in an attempt to discover how the two facilities could be used effectively together. No attempt was made to analyze the impacts which reservoir operations would have on downstream flooding, recreation, water supply or any other reservoir uses.

2.3 Optimization Model

The Explicitly Stochastic Dynamic Programming Model was originally presented by Butcher (1971). The dynamic program's stages consist of seasonal time periods (3 per year) and the states consist of a combination of storage volume at the beginning of the period and streamflow during the period. Each state is defined by S_{skt} (k^{th} storage interval at the beginning of season t at site s) and I_{sit} (i^{th} streamflow interval during season t at site s).

The streamflow and storage volumes used are discrete values which are chosen to approximate the continuous ranges of flows and storages. The range of streamflows passing each site has been divided into intervals and the mean of each interval has been used to represent the streamflows for that interval (Figure 2.1). Similarly a particular set of discrete storage volumes is used to represent all possible storage volumes for each reservoir.

A discrete, lag-one, Markov process is used to describe streamflows to the system. This description is incorporated into the model with the use of a set of conditional joint probabilities, $P[I_{sjt+1}\ \forall s]$, which represent the probability of a particular set of streamflows at all sites in season t+1 conditioned on a set of streamflows at all sites in season t. Several methods of estimating these transition probabilities can be used. Assuming that a particular probability distribution governs the transitions, the available historical data can be used to estimate the parameters of the distribution. They can also be approximated by the relative frequencies of events recorded in the data. It should be noted that while the relative frequency approximated for a single transition from a limited data base may not be accurate, the set of relative frequencies may adequately represent the transition process because it is only used to find the expected value of a series of benefits.

Movement from one stage to the next as defined by the recursion relation (see below) is backward in time. Each transition is tested for feasibility using the continuity equation and all combinations of streamflow and storage producing negative releases are rejected. The label which is placed on each state, $f_n(I_{sit}, S_{skt}\ \forall s)$, is the value of expected benefits obtainable from optimal operation of the system over n time periods; where n equals the number of dynamic program stages that have been considered when the label is assigned.

The return function, BENEFIT $(S_{skt}, S_{s\ell t+1}, I_{sit}\ \forall s)$ can be any function that defines the value of expected benefits which accrue for a particular set of releases $D_{ski\ell t}$ (release at site s for a transition with beginning storage, S_{skt}, ending storage $S_{s\ell t+1}$ and inflow I_{sit} in season t). The return function used in this model combines the long and short term benefits (and disbenefits) gained from a release as follows: long term benefits are defined in terms of the expected

Season	Interval	Bear Creek	Beltzville
1	1	624.4	182.8
	2	887.3	233.0
	3	1129.0	259.3
	4	1344.7	335.2
	5	1938.4	553.4
2	1	1455.6	373.9
	2	1760.5	469.7
	3	2035.4	540.6
	4	2300.2	615.4
	5	2643.8	786.9
3	1	287.8	86.5
	2	435.1	135.9
	3	626.8	204.6
	4	868.0	271.9
	5	1614.4	529.3

(All flows are in units of 100 acre-feet, Season 1: Oct.-Jan., Season 2: Feb.-May, Season 3: June-Sept)

Figure 2.1 Discrete Streamflows

annual value of meeting certain target levels of water supply availability, and power production. Short term benefits (or disbenefits) are defined in terms of penalties or rewards for falling short or exceeding the target levels. In those cases where target levels are missed by large margins, a disaster may be declared. Benefits from flood control are defined each season as a function of the system freeboard. A detailed description of the respective benefit functions is included below.

The operating policy designated by the model is a set of rules specifying the storage levels at the beginning of the next season for each combination of storage levels at the beginning of the present season and streamflows during the season.

The recursion relation is used to find the optimal set of rules. An initial set of rules is formed based on the immediate return obtained from those operations. Subsequent solutions of the recursion relation find operating policies which increase the value of expected economic benefits.

The recursion relation is:

$$f_n(I_{sit}, S_{skt}\ \forall s) = \underset{S_{s\ell t+1}\ \forall s}{\text{maximum}} \{ \text{BENEFIT}(I_{sit}, S_{skt}, S_{s\ell t+1}\ \forall s) +$$

$$\underset{\substack{\uparrow \\ \text{all possible streamflow} \\ \text{combinations in season} \\ t+1,\ (I_{sjt+1}\ \forall s)}}{\Sigma} P[I_{sjt+1}\ \forall s \mid I_{sit}\ \forall s] * f_{n-1}(I_{sjt+1}, S_{s\ell t+1}\ \forall s) \} \qquad (2.1)$$

The only feasible values of ($S_{s\ell t+1}\ \forall s$) are those such that $D_{skilt} = S_{skt} + I_{sit} - S_{\ell t+1} \geq 0$ for all sites. BENEFIT, the probabilities and $f_n(\cdot)$ have already been defined. The subscript t is related to n by the following:

$$t = (\bar{t}*10^6 - n)(\text{modulo } \bar{t}) + 1 \tag{2.2}$$

$$t+1 = (\bar{t}*10^6 - t - n + 1)(\text{modulo } \bar{t}) + 1 \tag{2.3}$$

where $\bar{t}$ equals the number of seasons per year.

The solution proceeds by solving the recursion relation for n equal to 1, 2, 3.... The procedure is continued until the same operating rules are defined for each season for two succeeding years. At such a point the program has converged on the optimal solution.

2.4 Example Reservoir System

The example reservoir system consists of two dams in a parallel configuration in the Lehigh River Basin (Figure 2.2). Bear Creek reservoir has a capacity of 200,000 acre-feet. In addition, the Beltzville Pipe has been constructed to convey water from the Bear Creek reservoir to the Beltzville reservoir and to generate power at the Beltzville Pipe power plant. The Beltzville Pipe power plant and Beltzville power plant have capacities of 110 and 20 megawatts respectively. (The actual design specifications of this system differ from those described here.) For this example, three uses of the reservoir system are considered: flood control, water supply, and hydroelectric power production. All of the economic data used in this research were adapted from Hufschmidt and Fiering (1966). The economic benefits associated with flood control in each season are a function of the freeboard available at the two dams. Three flood damage sites exist in the Lehigh River basin and the economic benefit functions for each site are shown in Figures 2.3, 2.4, and 2.5.

The economic benefits associated with a change in the supply of water are

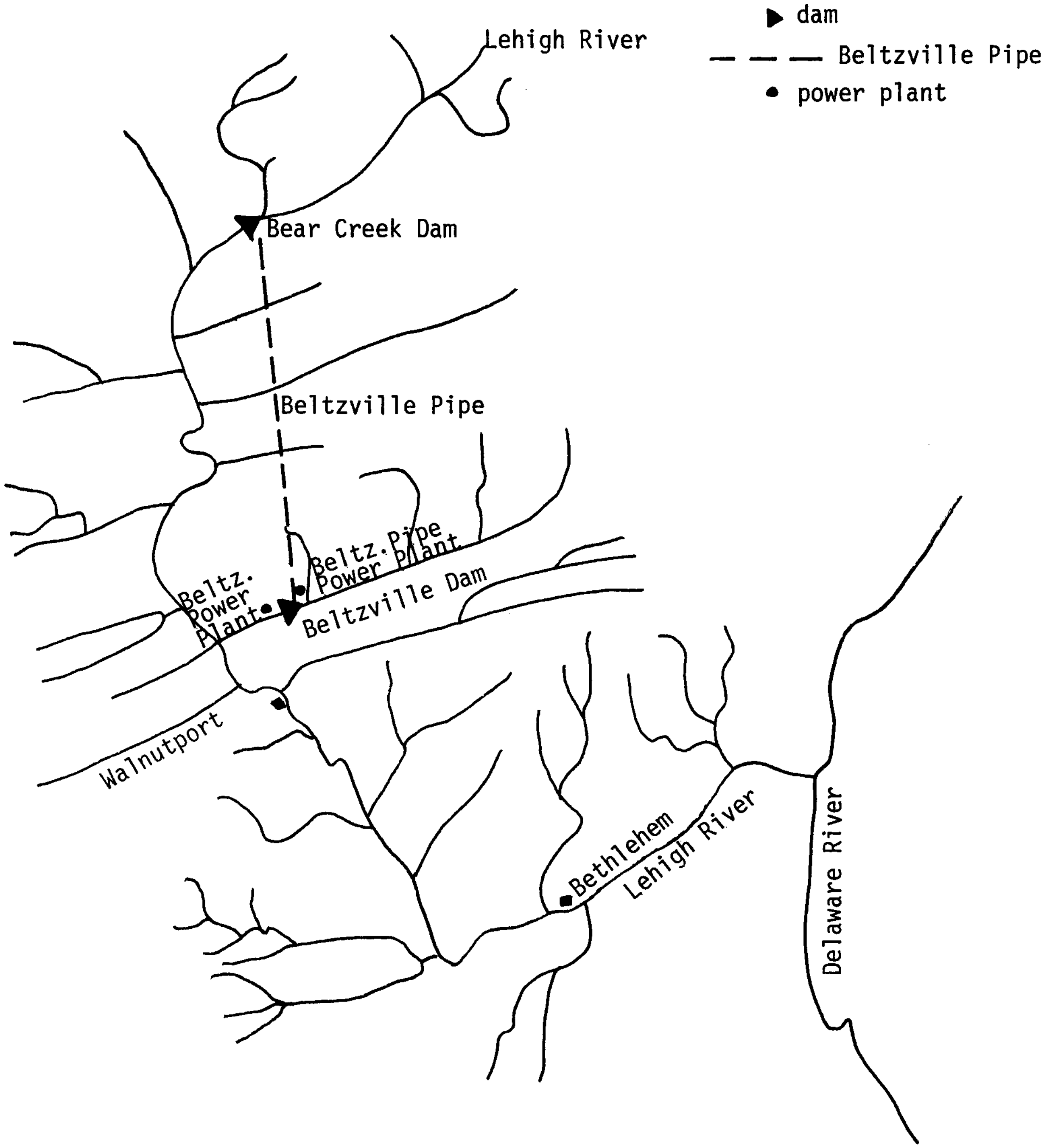


Figure 2.2. Lehigh River Basin

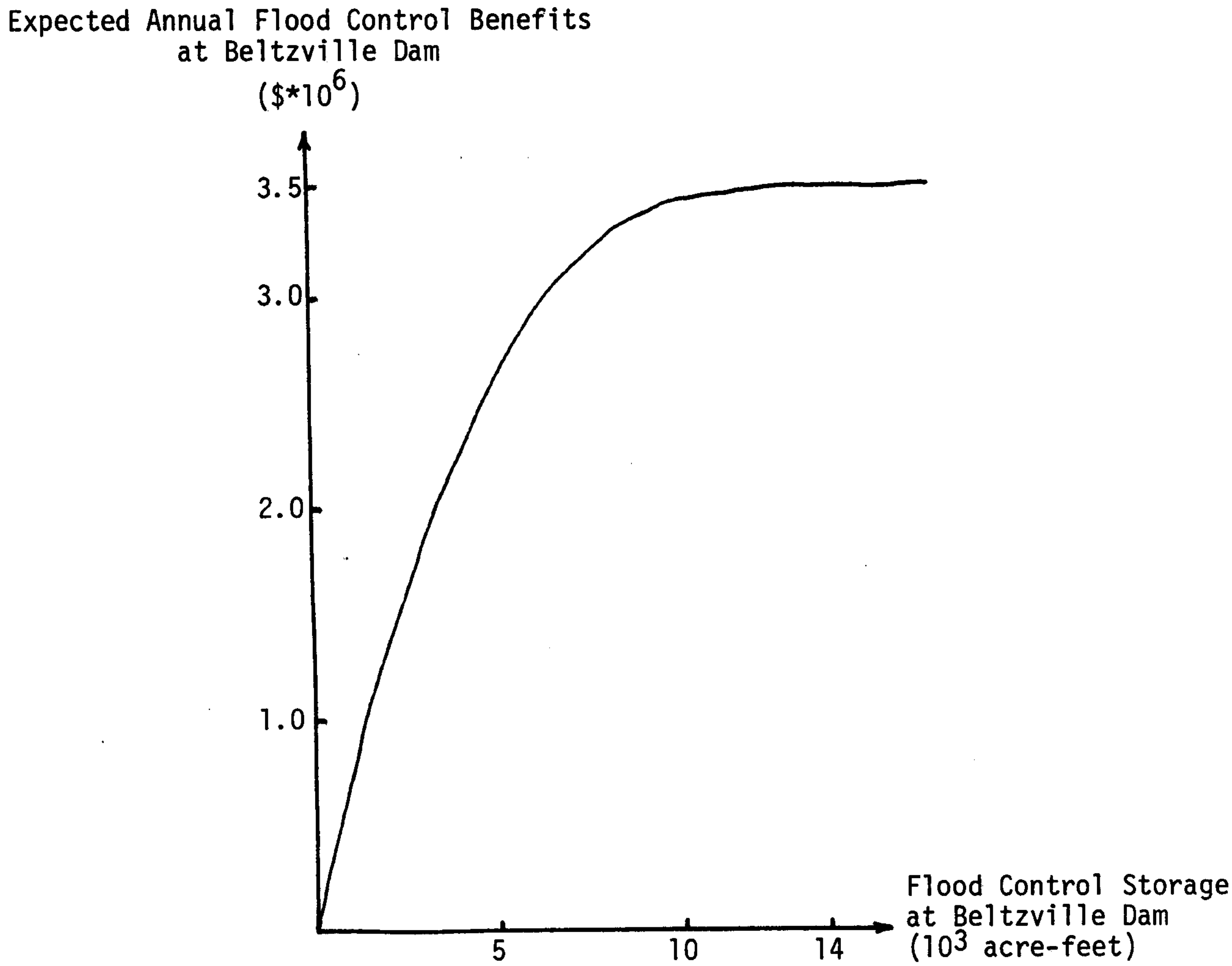


Figure 2.3. Flood Control Benefits Below Beltzville Dam

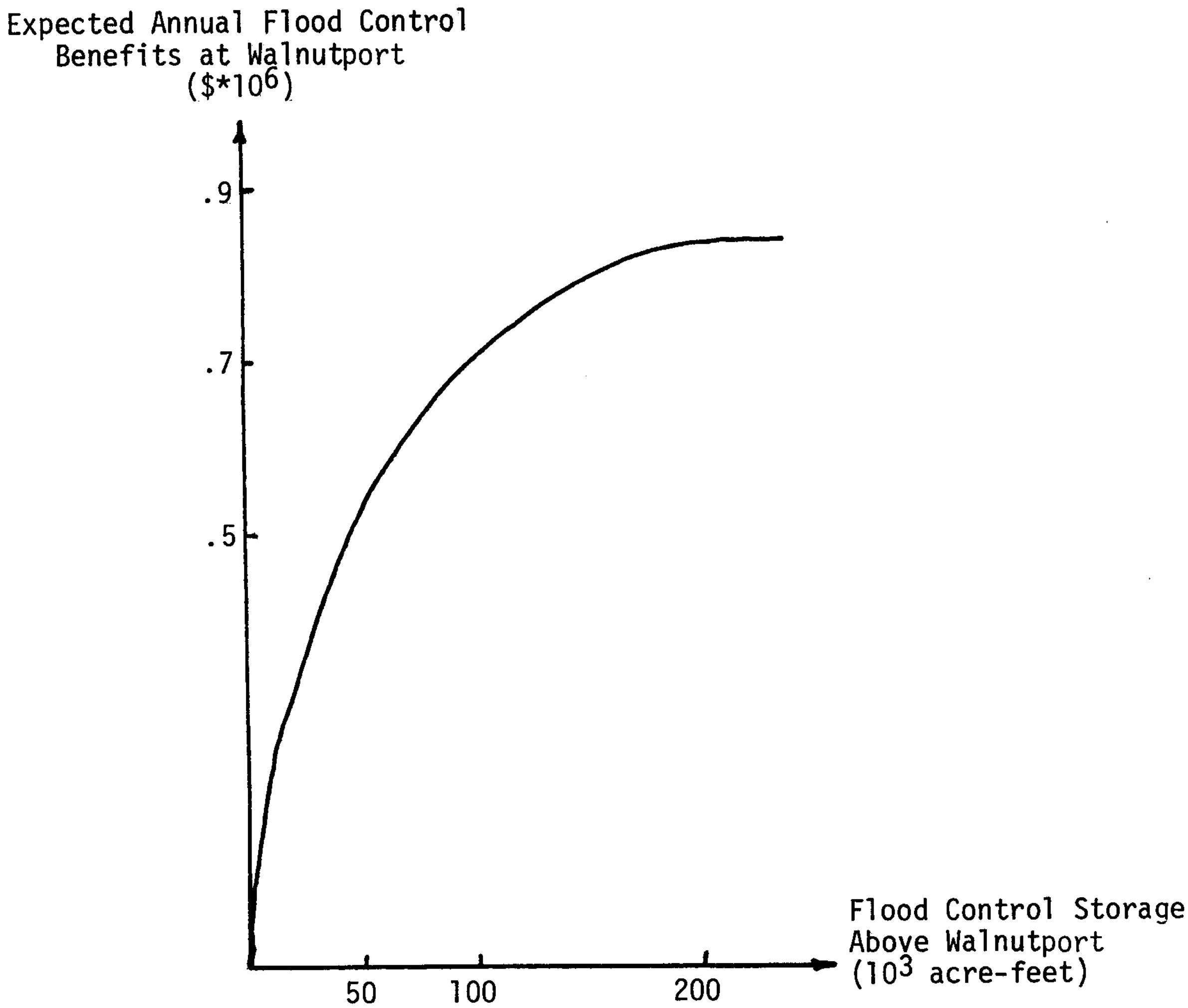


Figure 2.4. Flood Control Benefits at Walnutport

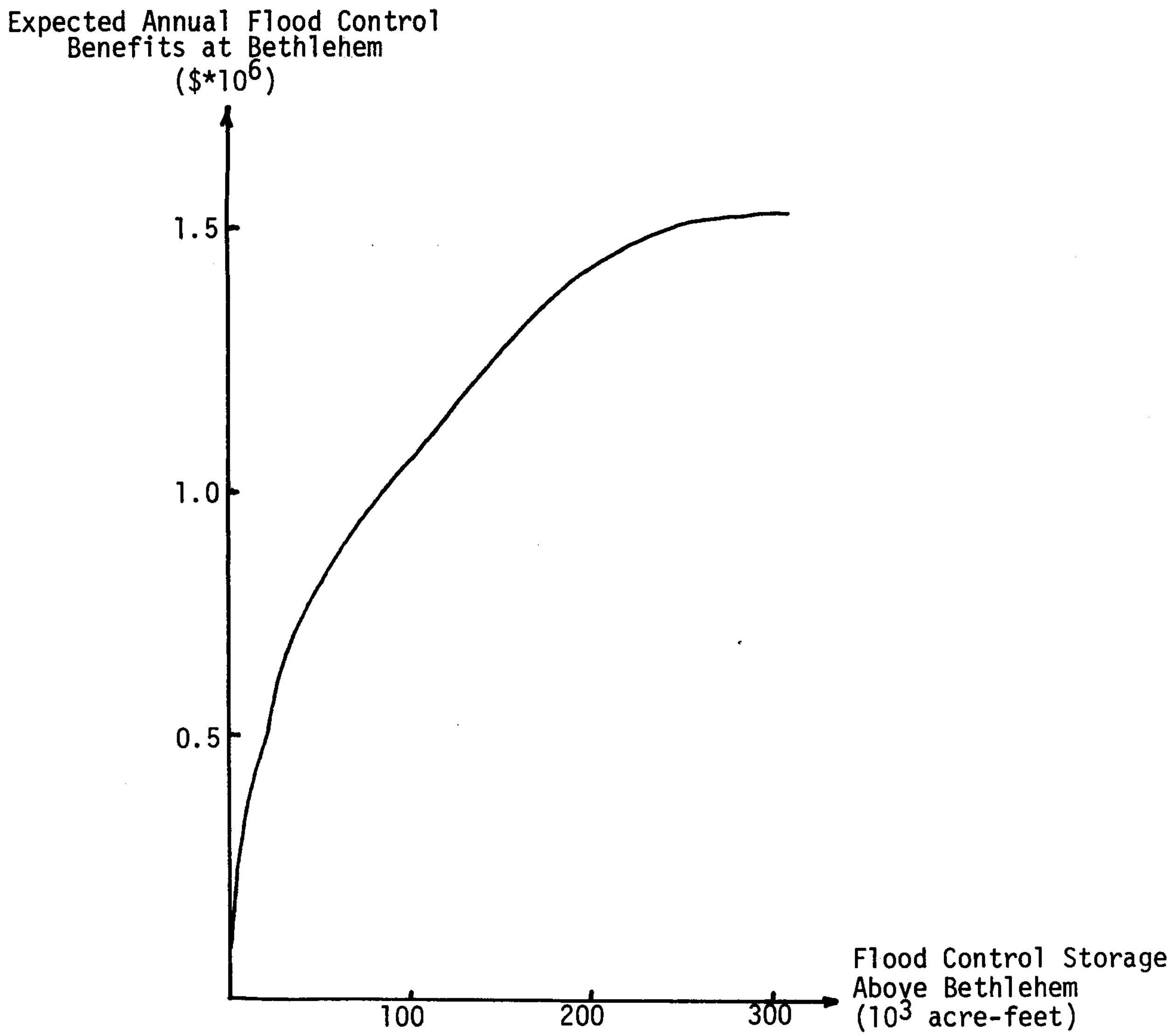


Figure 2.5. Flood Control Benefits at Bethlehem

stated as a function of a target and deviations from the target. The target represents an amount of water which can vary seasonally but which is anticipated and expected by water demanders. Long term planning of water use is based on the target and long term benefits are defined as a function of it (Figure 2.6). During actual operations it is possible that quantities of water either greater than or less than the target will be supplied to the water users. Short term benefits and losses are functions of the deviations of the amount of water supplied from the target. Figure 2.7 shows the short term loss per unit of deficit, as a function of the water supply deficit expressed as a percentage of the target.

Hydroelectric power production benefits are also defined in terms of seasonal targets and long term benefits, and deviations from the target and short term losses and benefits. Figure 2.8 shows the long term economic benefits as a function of an annual energy target. Hufschmidt and Fiering (1966) stated that short term losses amount to fifteen dollars per megawatt of deficit for any deficit up to twenty-five percent of the target. Deficits greater than twenty-five percent were considered disasters.

In order to investigate the effects which increased demands for hydroelectric energy production have on the management of the example reservoir system, the optimization model was solved for 33 different combinations of water supply target and energy production target. In the next section the results of these solutions are presented and discussed.

2.5 Results

The Explicitly Stochastic Dynamic Program was solved for thirty-three different combinations of water supply and energy production targets. The

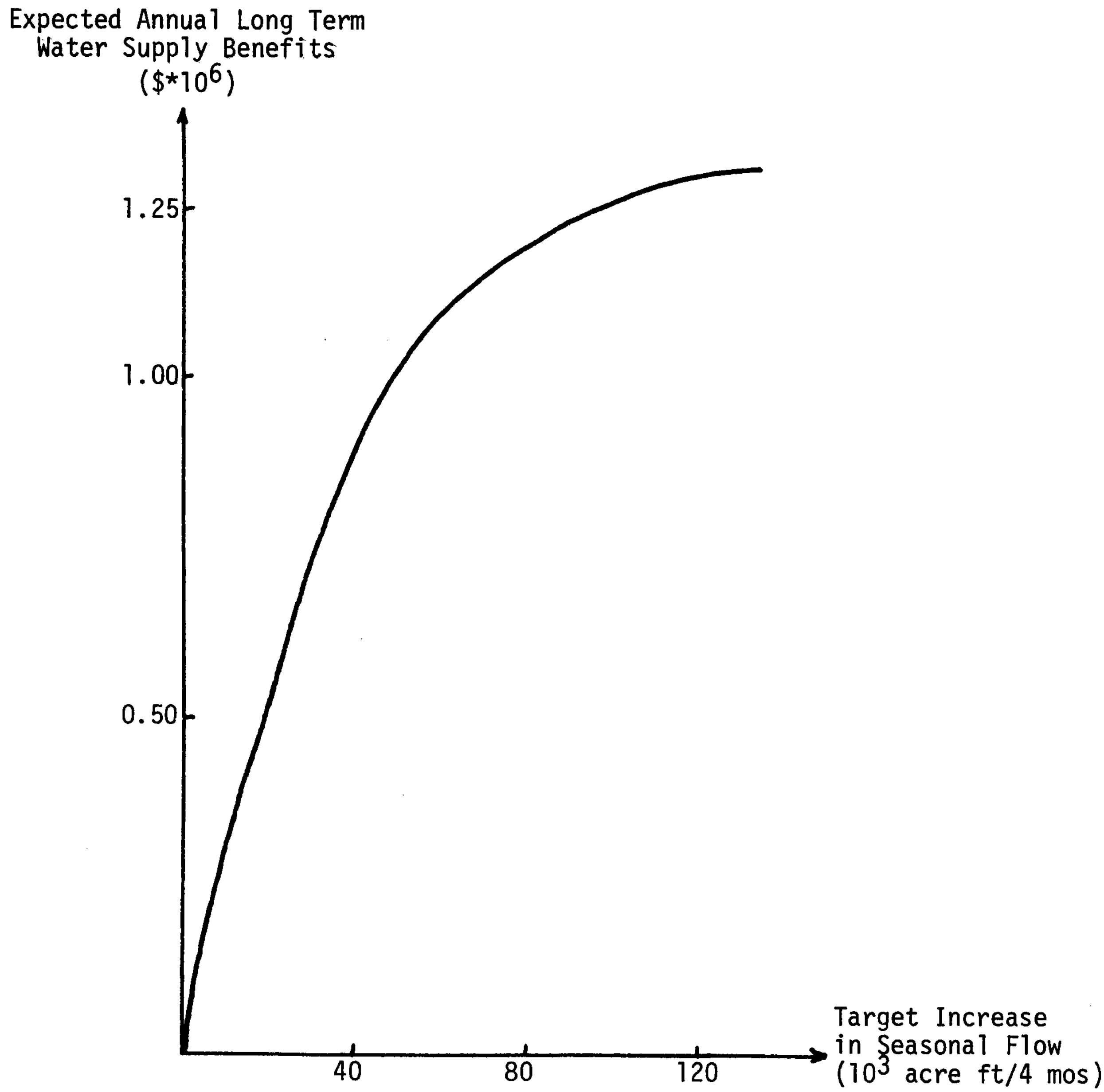


Figure 2.6. Long Term Water Supply Benefits

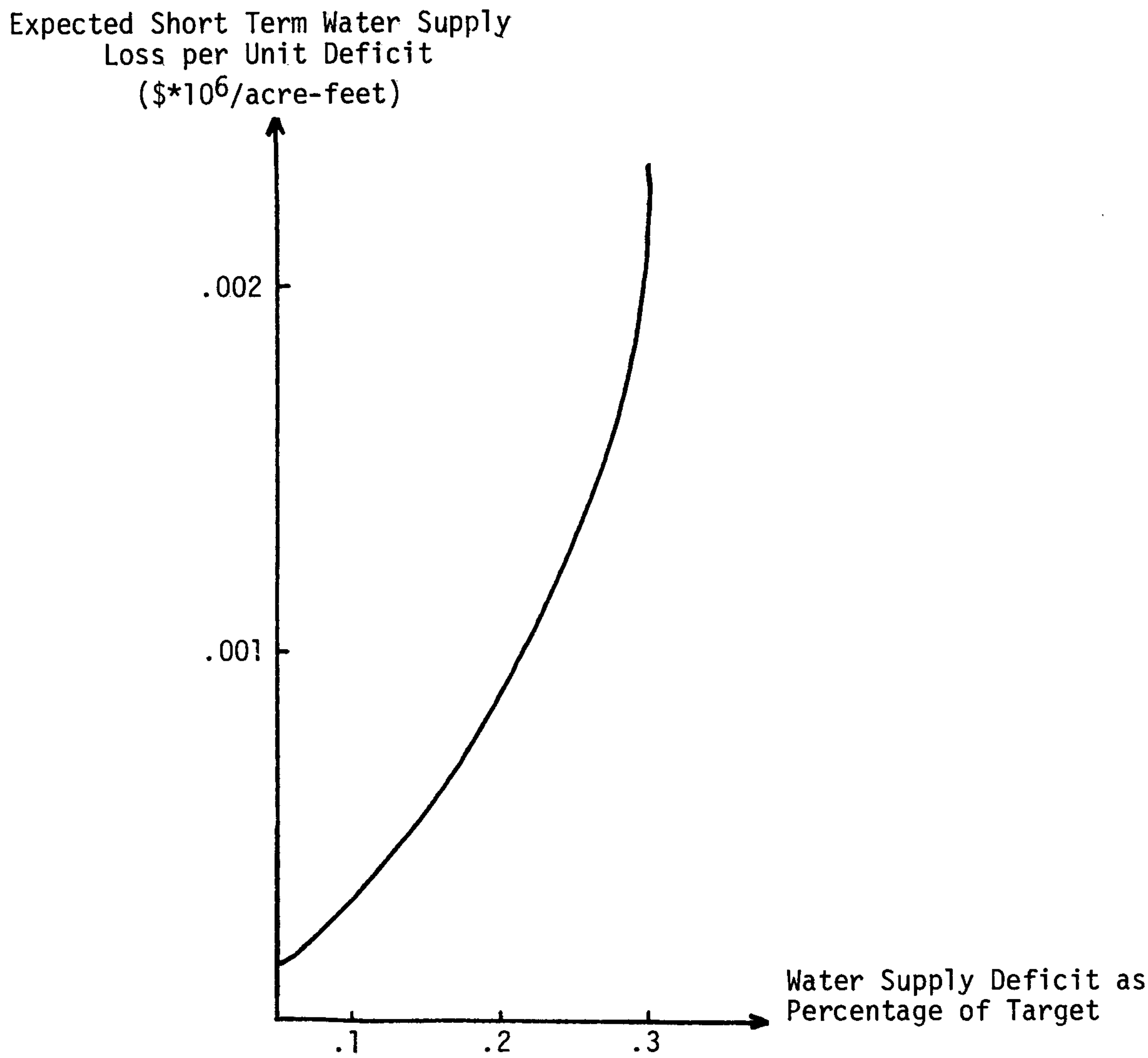


Figure 2.7. Water Supply Short Term Loss

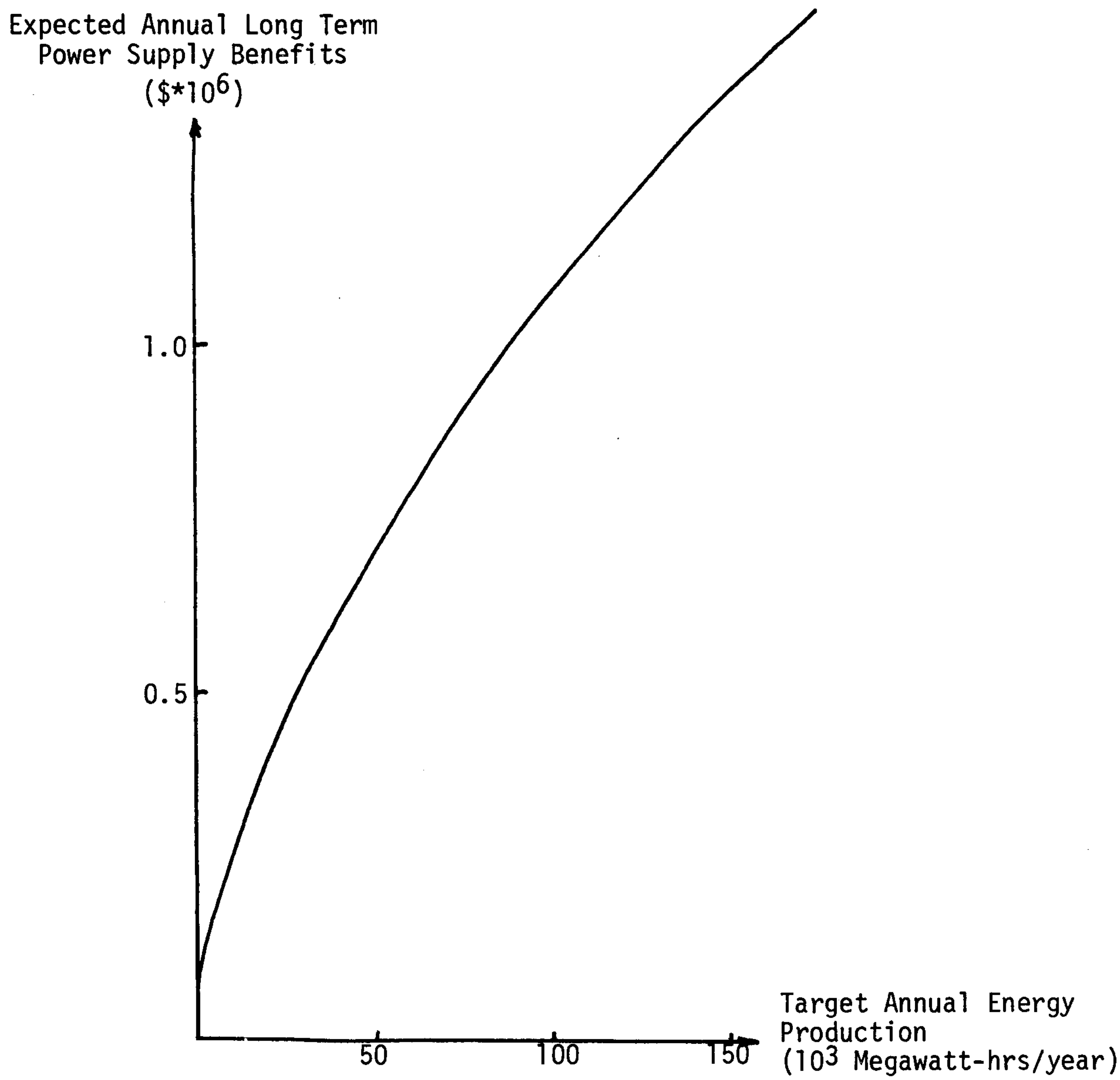


Figure 2.8. Long Term Energy Production Benefits

solutions can be displayed in numerous ways, each of which conveys different information to the observer and provides answers to different questions. Six displays have been generated to demonstrate the technique of multiobjective analysis.

Display 1. How do total net economic benefits vary as a function of the energy production target and the demand pattern of the system? Obviously, as the energy target changes there will be a change in the maximum total benefits obtainable from operation of the reservoir system. The energy target however does not indicate whether the power is generated at a relatively constant level (base load) or at varying levels (peaking).

The load factor which equals the ratio of average to peak power production can be used to indicate the variability of the power production. A load factor of .2 indicates that the power plant is used for meeting peak demands: the peak power production is five times the average. A load factor of .8 indicates that the power plant is being used to meet base loads: the peak power production is only twenty-five percent higher than average production.

The capacity of a power plant must be great enough to allow the peak production. Hence, for a fixed capacity power plant there are many possible combinations of energy target and load factor.

Figure 2.9 shows the relationship between the energy production target and the expected value of annual benefits obtainable from the reservoir system. The water supply target is fixed at a level of $60*10^3$ acre-feet/4 months. As the load factor increases and the energy target decreases, the maximum expected annual benefits vary.

Display 2. How do water supply benefits vary as a function of the energy target? Because water supply and power production may be conflicting uses of

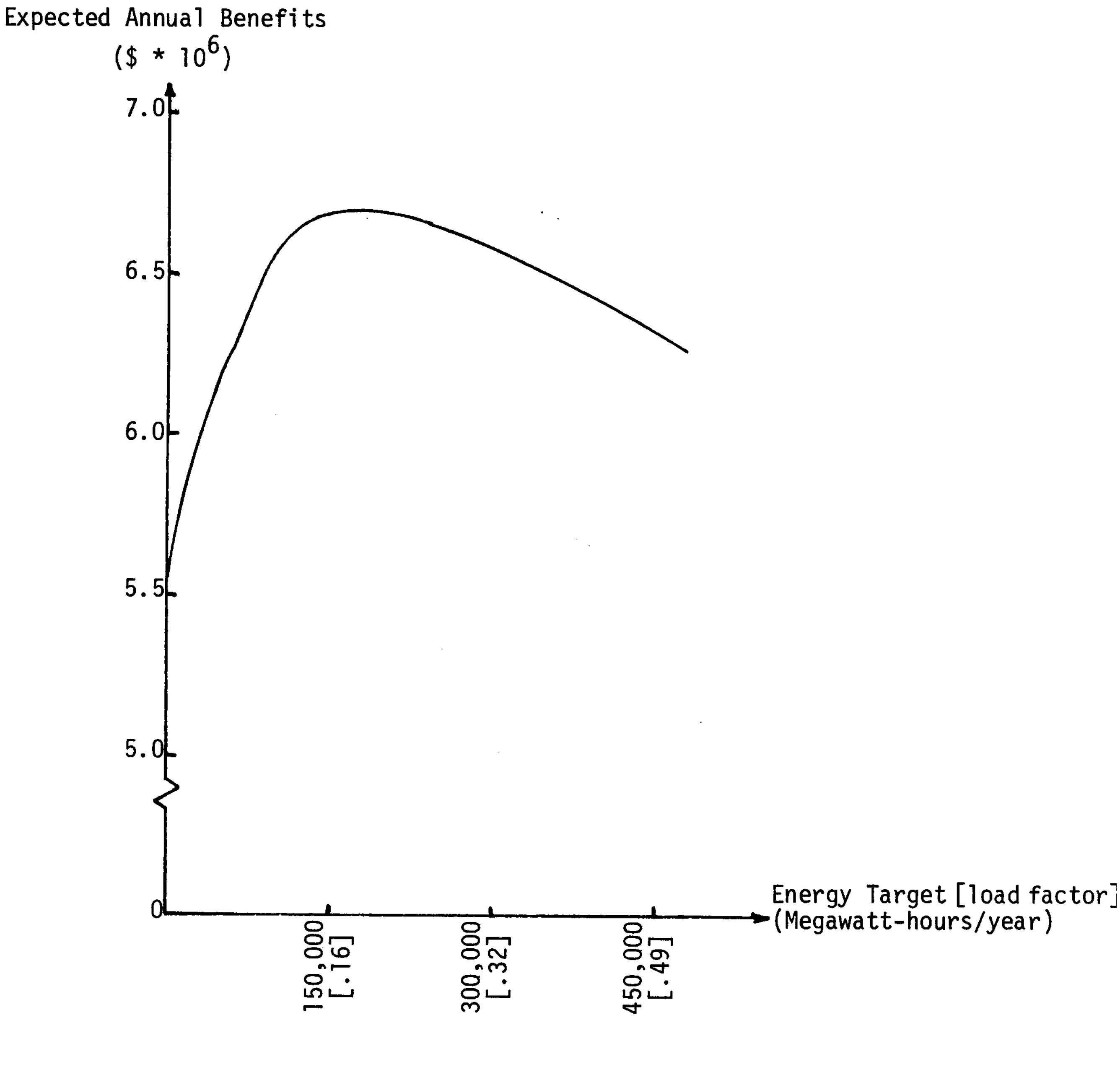


Figure 2.9. Expected Annual Benefits v. Energy Target
Water Supply Target = 60 * 10^3 acre-feet/4 months

the reservoir system, it is expected that as the energy target increases a decrease in expected annual water supply benefits would occur. Figure 2.10 shows the relationship between the energy target, water supply target and water supply benefits.

Display 3. How do flood control benefits vary as a function of the energy target? Figure 2.11 is similar to Figure 2.10 except that the vertical axis represents expected annual flood control benefits. As the energy target increases the benefits from flood control decrease because energy production and flood control conflict in terms of reservoir system management.

Display 4. How do benefits from power production vary as a function of the energy target? For a fixed water supply target it is expected that a single energy target will provide the maximum power benefits. Figure 2.12 shows this relationship holds for several water supply targets.

Display 5. How do total benefits vary as a function of the energy target and the water supply target? Figure 2.13 shows a family of contours or isopleths representing equal total expected benefits over a wide range of water supply and energy targets.

2.6 Conclusions

It is evident from these displays that flood control is the dominant use of this system. Both power production and water supply are much less beneficial economically than flood control. Not until targets for energy production and water supply are increased to levels significantly above the anticipated levels do benefits from these uses approach flood control benefits.

Even though total economic benefits are dominated by flood control benefits, the system displays expected characteristics. This is most evident in Figure 2.13. As either the water supply target or energy production target

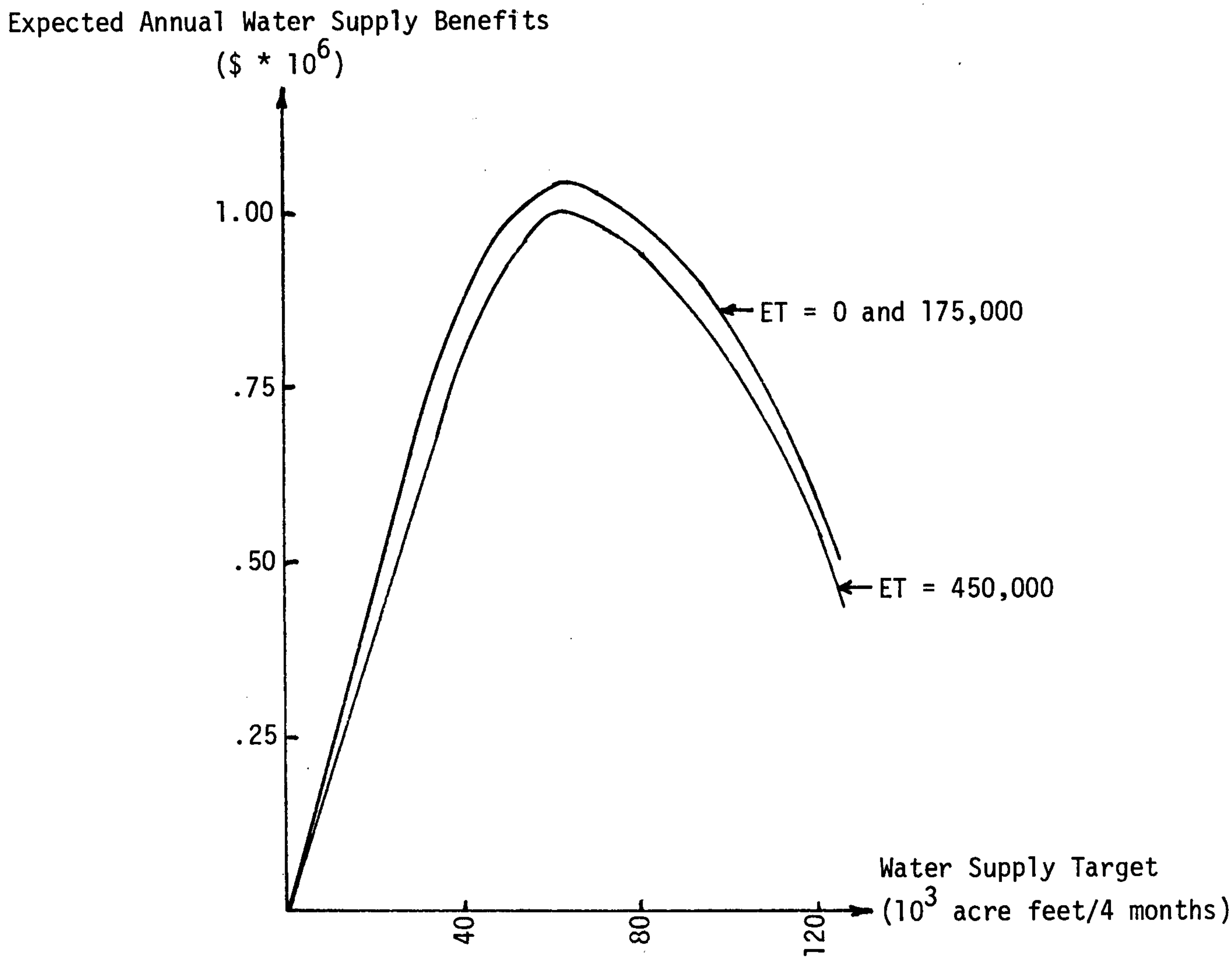


Figure 2.10. Expected Annual Water Supply Benefits v. Water Supply Target
ET = Energy Target (Megawatt-hours/year)

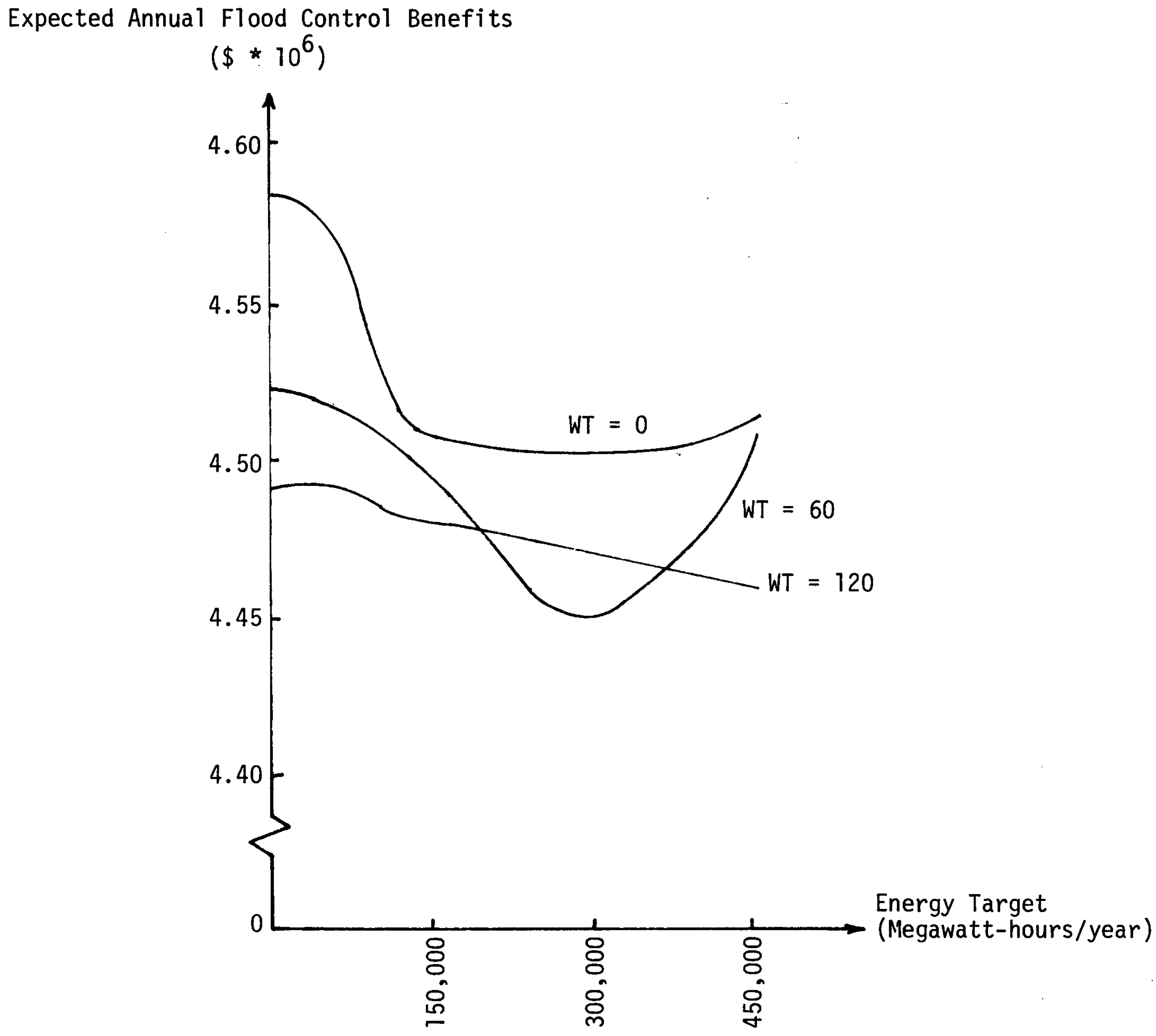


Figure 2.11. Expected Annual Flood Control Benefits v. Energy Target
WT = Water Supply Target (10^3 acre feet/4 months)

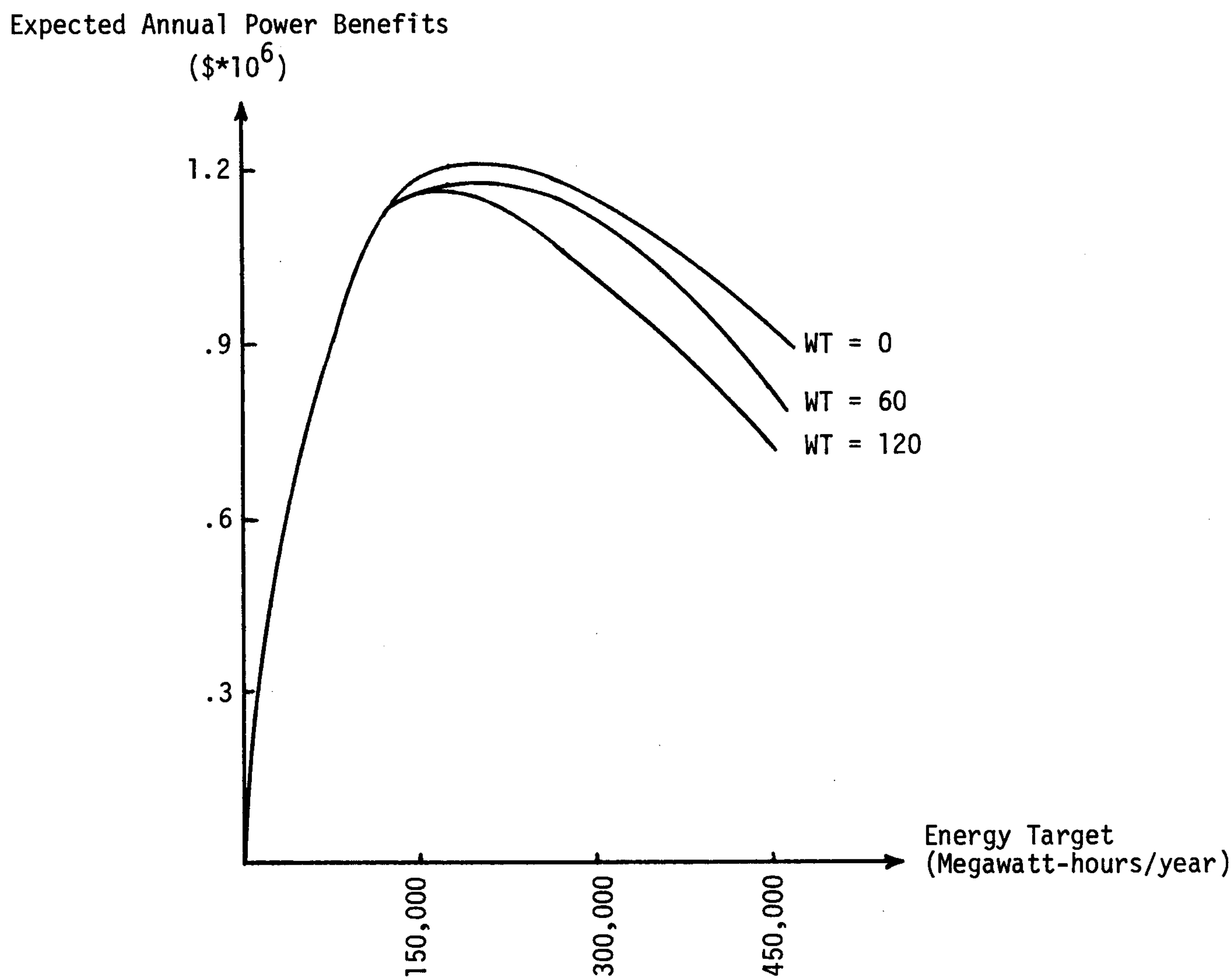


Figure 2.12. Expected Annual Power Benefits v. Energy Target
WT = Water Supply Target (10^3 acre-feet/4 months)

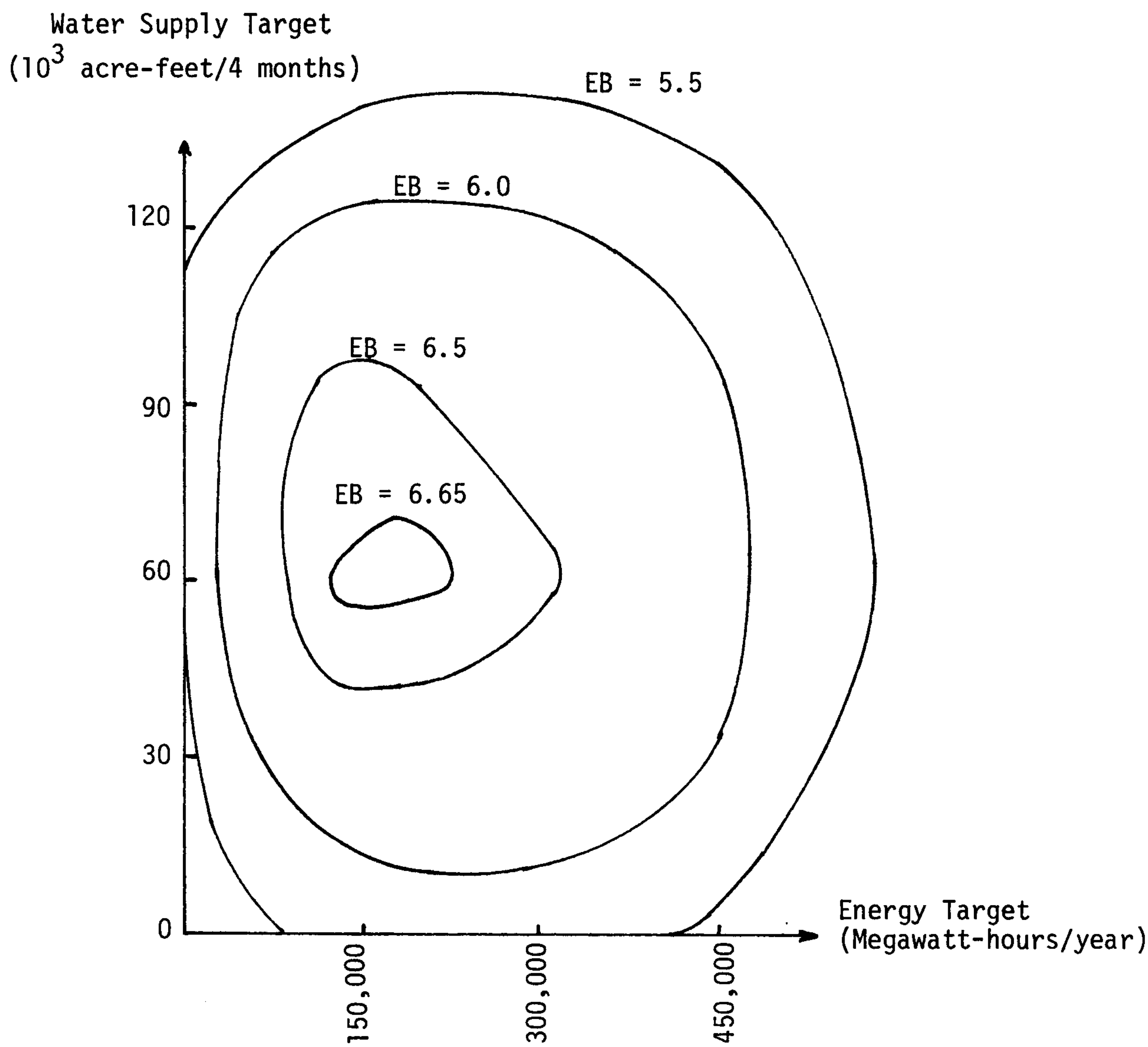


Figure 2.13. Expected Annual Benefit Isopleths
EB = Expected Annual Benefits ($\$ * 10^6$)

increases and the other is fixed, total benefits obtainable from the system increase and then decrease.

Future investigation must be undertaken to determine how management of the system is affected by changes in the economic benefit functions. Presumably as the demand for hydroelectric power production increases there will be a concomitant increase in both long term benefits and short term benefits and losses. Hence, anticipated increases in energy demand will have a greater impact on optimal reservoir management than indicated here.

3. CHANCE CONSTRAINED STOCHASTIC LINEAR PROGRAMMING MODEL

3.1 Introduction

An examination of optimization models described in the preceding chapters and used for the design and/or management of multiple purpose, multiple reservoir systems shows that those models which more accurately represent the system-- principally by explicitly incorporating the streamflows' stochasticities and by allowing general operating rules to be determined -- become computationally unsolvable when a moderately complex reservoir system is considered. In fact, those models which have explicitly stochastic streamflow descriptors and do not overly restrict the operating policy prior to solution cannot be used to consider systems with more than two or three reservoirs without major limiting assumptions being made. The other models which are less accurate are also less expensive to solve although their costs rise as the complexity of the system increases. Hence, there exists a tradeoff between computational burden and accuracy of the model (see Houck and Cohon (1976)). Therefore, it is possible to construct a model for a large scale system but the solution of the model may not provide any useful information because the model does not accurately represent the system.

One of the optimization models which was introduced by ReVelle, Joeres, and Kirby (1969) and discussed by many authors since then (for example: Eastman and ReVelle (1973), Gundelach and ReVelle (1975), Houck (1975), Loucks and Dorfman (1975), Joeres, et al., (1971)) can assist in design and/or management problems for multiple reservoir systems. This model, the Linear Decision Rule (LDR) model, does not explicitly incorporate the stochasticity of streamflows and it does limit the operating policy prior to solution of the model. Although these restrictions are significant, the LDR allows the formation of a very small linear programming model for large complex reservoir systems (Houck, Cohon and ReVelle (1977)). If methods can be found to incorporate the stochastic properties of the streamflows, relax the restrictions on operating rules, and retain the small computational burden of the model; the result may be an accurate, economically and computationally feasible, and generally applicable model.

Loucks and Dorfman (1975) demonstrated that an LDR model as proposed by ReVelle, et al., (1969) is conservative. That is, the operating rules designated by the LDR model when tested in a simulation model result in a management policy that exceeds the limitations specified in the LDR model. Loucks and Dorfman (1975) propose the use of a different LDR to make the model less conservative. Gundelach and ReVelle (1975) also propose an extended LDR which results in a less conservative model. Neither of these efforts remedies the cause of the conservative nature of the LDR model, however. They do not attempt to incorporate conditional streamflow cumulative distribution functions (CDF) in the model, but continue to use unconditioned CDFs in the chance constraints.

The next section of this paper deals with a simple method of remedying the cause of the conservativeness of the LDR model; incorporating explicitly the streamflows' stochasticities; and relaxing the restrictions on the operating rules. All of these are accomplished by introducing multiple linear decision rules, each conditioned on any desired seasons' streamflows, in the linear program. The forms of either of the extended LDRs proposed by Gundelach and ReVelle (1975) or Loucks and Dorfman (1975) can be used in this method. Moreover, the resulting linear program retains a relatively small computational burden even for large reservoir systems.

3.2 Multiple Linear Decision Rules Model

The LDR to be used in this model is: $X_t^{ij} = S_t^j - b_t^i$.

X_t^{ij} = release in season t conditioned on streamflows in interval i in season t-1 and interval j in season t-2

S_t^j = storage volume at the beginning of season t conditioned on streamflow in interval j in season t-2

b_t^i = decision parameter in season t conditioned on streamflow in interval j in season t-1

R_t^i = streamflow in season t conditioned on streamflow in interval i in season t-1

The streamflow intervals in each season represent a portion of the continuum of possible streamflows and can be defined in any desired manner.

The continuity equation -- a mass balance on water in storage in the reservoir -- is: $S^i_{t+1} = S^j_t + R^i_t - X^{ij}_t$. Substituting the LDR in the continuity equation yields: $S^i_{t+1} = R^i_t + b^i_t$ or $S^j_t = R^j_{t-1} + b^j_{t-1}$. Thus both storage and release are defined in terms of the decision parameters and previous streamflows. This also means that storage and release are positively correlated and that in each season they are actually conditioned on storage levels, releases and streamflows in previous seasons.

The chance constrained linear program can be formulated with the usual physical restrictions. Any performance requirements or economic measures as described by Loucks and Dorfman (1975), Houck, et al., (1977), or as described in chapter two can also be included in the linear program. The simplest reservoir model which will be used as an example here includes restrictions on minimum and maximum permissible storage levels and a minimum permissible release level. The objective will be to minimize the capacity of the reservoir required to meet these restrictions.

The restriction on minimum permissible release will be used to demonstrate the procedure required to formulate the constraints of the linear program. The first step is to write a set of chance constraints that defines the probability that each possible release in a season exceeds the minimum allowable release.

$$P\left[\begin{array}{l}\text{release in season } t \text{ conditioned on} \\ \text{streamflows in interval } i \text{ in season } t\text{-}1 \text{ and} \\ \text{interval } j \text{ in season } t\text{-}2\end{array} \geq X_{min}\right] = \alpha^{ij}_t \tag{3.1}$$

$$P\left[X^{ij}_t \geq X_{min}\right] = \alpha^{ij}_t \tag{3.2}$$

$$P\left[R^j_{t-1} + b^j_{t-1} - b^i_t \geq X_{min}\right] = \alpha^{ij}_t \tag{3.3}$$

$$P\left[R^j_{t-1} \leq X_{min} - b^j_{t-1} + b^i_t\right] = 1 - \alpha^{ij}_t \tag{3.4}$$

If R^{j}_{t-1} is a random variable, the left hand side of the equation is the cumulative distribution function of streamflows in season t-1 conditioned on streamflows in the previous season occuring in interval j. This CDF can be represented by $F_{jt-1}(\cdot)$ and equation 3.4 can be restated as:

$$F_{jt-1}(X_{min} - b^{j}_{t-1} + b^{i}_{t}) = 1 - \alpha^{ij}_{t} \tag{3.5}$$

If $F^{-1}_{jt-1}(\cdot)$ is the inverse conditional CDF then 3.4 can again be restated:

$$X_{min} - b^{j}_{t-1} + b^{i}_{t} - F^{-1}_{jt-1}\ (1 - \alpha^{ij}_{t}) = 0 \quad \forall i,j,t \tag{3.6}$$

Finally, if the negative inverse CDF is convex, it can be piecewise linearly approximated and the regular SIMPLEX algorithm can be used to determine a global optimal solution.

A normal restriction on release is that 95% of all releases in each season must exceed X_{min}. Because X^{ij}_{t} is the release conditioned on previous streamflows in season t-1 and t-2 being in intervals i and j respectively, it is possible to write a restriction on α^{ij}_{t}. The probability of streamflows in seasons t-1 and t-2 being in intervals i and j is represented by P^{ij}_{t} and can be estimated from historical streamflow data. The actual reliability with which release in season t exceeds X_{min} is the expected value of the α^{ij}_{t} s (left hand side of equation 3.7). And this reliability must exceed α which may equal .9, .95, .99 or some other value.

$$\sum_{i} \sum_{j} P^{ij}_{t}\ \alpha^{ij}_{t} \geq \alpha \quad \forall t \tag{3.7}$$

Equation 3.7 places restrictions on the possible values of α^{ij}_{t} and therefore on the range over which $-F^{-1}_{jt-1}\ (1 - \alpha^{ij}_{t})$ must be convex. The smallest possible value for any single α^{ij}_{t} occurs when all other α^{ij}_{t} take on values of one. Then the minimum possible value of α^{ij}_{t} equals $[(\alpha - 1 + P^{ij}_{t})/P^{ij}_{t}]$. If $\alpha = .95$ and all $P^{ij}_{t} \geq .1$, the smallest possible value for any α_t^{ij} is .5. Because the minimum release would not normally be met 100% of the time by all but one of the LDRs, the actual values of α_t^{ij} would exceed, perhaps greatly exceed, the minimum value. Over the range of 0.5 to 1.0 for α_t^{ij} or 0.0 to 0.5 for $(1 - \alpha_t^{ij})$, the negative inverse conditional CDF of streamflows will usually be convex. Hence, constraint 3.5 can be accomodated easily within a linear program.

The same procedure can be followed to assure a minimum storage restriction.

$$P\begin{bmatrix}\text{storage at the beginning of season t} \\ \text{conditioned on streamflows in interval} \geq S_{min} \\ \text{j in season t-2}\end{bmatrix} = \beta_t^j \tag{3.8}$$

$$P\left[S_t^j \geq S_{min}\right] = \beta_t^j \tag{3.9}$$

$$P\left[R_{t-1}^j + b_{t-1}^j \geq S_{min}\right] = \beta_t^j \tag{3.10}$$

$$F_{jt-1}(S_{min} - b_{t-1}^j) = 1 - \beta_t^j \tag{3.11}$$

$$S_{min} - b_{t-1}^j - F_{jt-1}^{-1}(1 - \beta_t^j) = 0 \quad \forall j, t \tag{3.12}$$

Again, if the negative inverse conditional CDF is convex it can be piecewise linearly approximated and the regular SIMPLEX algorithm used. The probability of storage S_t^j occuring is the probability that in season t-2 the streamflow occured in interval j, P_t^j. The reliability with which storage exceeds the minimum storage is β and equation 3.13 specifies that this requirement is met.

$$\sum_j P_t^j \beta_t^j \geq \beta \quad \forall t \tag{3.13}$$

So the minimum value for one β_t^j is $((\beta - 1 + P_t^j)/P_t^j)$ or .5 if the minimum value for any P_t^j is .1 and β equals .95. Over the range 0.0 to 0.5 for $(1 - \beta_t^j)$, the negative inverse conditional CDF is usually convex.

The last restriction is on maximum storage levels.

$$P \left[\begin{array}{ll} \text{storage at the beginning of season } t & \text{reservoir} \\ \text{conditioned on streamflow in interval } j \leq & \text{storage} \\ \text{in season } t\text{-}2 & \text{capacity} \end{array} \right] = \gamma_t^j \tag{3.14}$$

$$P \left[S_t^j \leq CAP \right] = \gamma_t^j \tag{3.15}$$

$$P \left[R_{t-1}^j + b_{t-1}^j \leq CAP \right] = \gamma_t^j \tag{3.16}$$

$$P \left[R_{t-1}^j \leq CAP - b_{t-1}^j \right] = \gamma_t^j \tag{3.17}$$

$$CAP - b_{t-1}^j - F_{jt-1}^{-1}(\gamma_t^j) = 0 \quad \forall j,t \tag{3.18}$$

The reliability with which the storage must not exceed the dam capacity is γ and the requirement is defined in equation 3.19.

$$\sum_j P_t^j \gamma_t^j \geq \gamma \quad \forall t \tag{3.19}$$

The minimum value for any γ_t^j is the same as the minimum possible value of β_t^j. The range over which the negative conditional CDF must be examined is .5 to 1.0 under the assumption stated above. Over this range the negative conditional CDF is usually concave. However, this is the shape for which piecewise linear approximation of the curve and the regular SIMPLEX algorithm will produce a global optimum.

A linear program can be formed with an objective of MINIMIZE CAP; constraints 3.6, 3.7, 3.12, 3.13, 3.18, and 3.19; and nonnegativity restrictions on all variables except the LDR decision parameters. This linear program can be solved using the regular SIMPLEX algorithm to obtain a global optimum. The restrictions placed on operating rules prior to solution are greatly reduced from previous LDR models. Although each potential rule has the form of the original linear decision rule, there are multiple rules for each season. Thus the restrictiveness due to the form of the rule is reduced because as operating conditions change the operating rule can change.

The multiple linear decision rule model has been developed in one form; the releases have been conditioned on the previous two seasons' streamflows and storages have been conditioned on the streamflow of two months previous. There are many other ways to formulate this type of model however. Both storage and release can be conditioned on the previous two, three, four or more season's streamflows; on only the previous season's streamflow; on the next season's streamflow and some set of previous seasons' streamflows; or on some other combination of future and/or past seasons' streamflows. Because in each of these multiple LDR models the storages and releases are functions of the same variables, the releases and storages are actually functions of previous releases, storages, and streamflows; and perhaps anticipated future releases, storages and streamflows. The choice of which events to condition the releases and storages on is dependent on how accurately future streamflow predictions can be made, how accurately the model is required to represent the system, and how large a computational burden can be supported.

The stochastic properties of the streamflows have been incorporated in the linear program in the form of conditional CDFs as defined by the multiple linear decision rules. The number of seasons of previous streamflow which are used to condition the current seasons streamflow, as well as the accuracy of measurements of previous seasons' streamflows used for conditioning, are controlled by the analyst.

The example model's computational burden is relatively small. The number of constraints in this single site example is 386 if monthly seasons are modeled and streamflows are broken into five intervals in each season. The model can be reduced significantly if the operating rule in season t is only conditioned on streamflows in the previous season. And, in either case, the number of constraints does not increase exponentially with the number of facilities considered.

3.3 A Multiple Linear Decision Rules Model Example

The multiple LDR model developed in the previous section was constructed and solved for a portion of the Yakima River System in eastern Washington. A single dam site, the present location of Lake Kachess, was examined. The streamflow data used to generate the conditional cumulative distribution functions (figures 3.1, 3.2, 3.3) were provided by the Yakima Project Office of the Bureau of Reclamation. Each of the conditional CDFs was approximated in the model by six linear pieces; three pieces for the 0 to 50 percentile portion and three pieces for the 50 - 100 percentile portion.

Due to limited computing capabilities only three seasons were modeled and the number of streamflow intervals (i or j) on which release and storage were conditioned was two. The streamflow volumes that divided the two intervals in each season are shown in Table 3.1. The values assigned to the reliability levels for assuring the minimum release (α), assuring the minimum storage volume (β), and assuring the maximum storage volume (γ) were all .90. The minimum storage level (S_{min}) was assigned a value of $24.67 * 10^6$ m^3 and the minimum seasonal release (X_{min}) was assigned a value of $30.84 * 10^6$ m^3.

The linear program comprised 35 constraints and 87 variables, 72 of which had finite upper bounds. Solution time was approximately 5 C.P.U. seconds using the Multi-Purpose Optimization System resident on a CDC 6400 computer. The total cost per solution was less than one dollar.

The minimal dam capacity required to meet the specified restrictions is $1.24 * 10^6$ m^3. Tables 3.2, 3.3, and 3.4 contain the values of the multiple LDR parameters b_t^i, and the release and storage reliabilities α_t^{ij}, β_t^j, γ_t^i designated in the solution of the linear program. From examination of Table 3.3 it is evident that the minimum release is met with a reliability exceeding .90 in all three seasons. Storage levels exceed the minimum storage level with reliabilities equal to .90, .90, and .965 in seasons

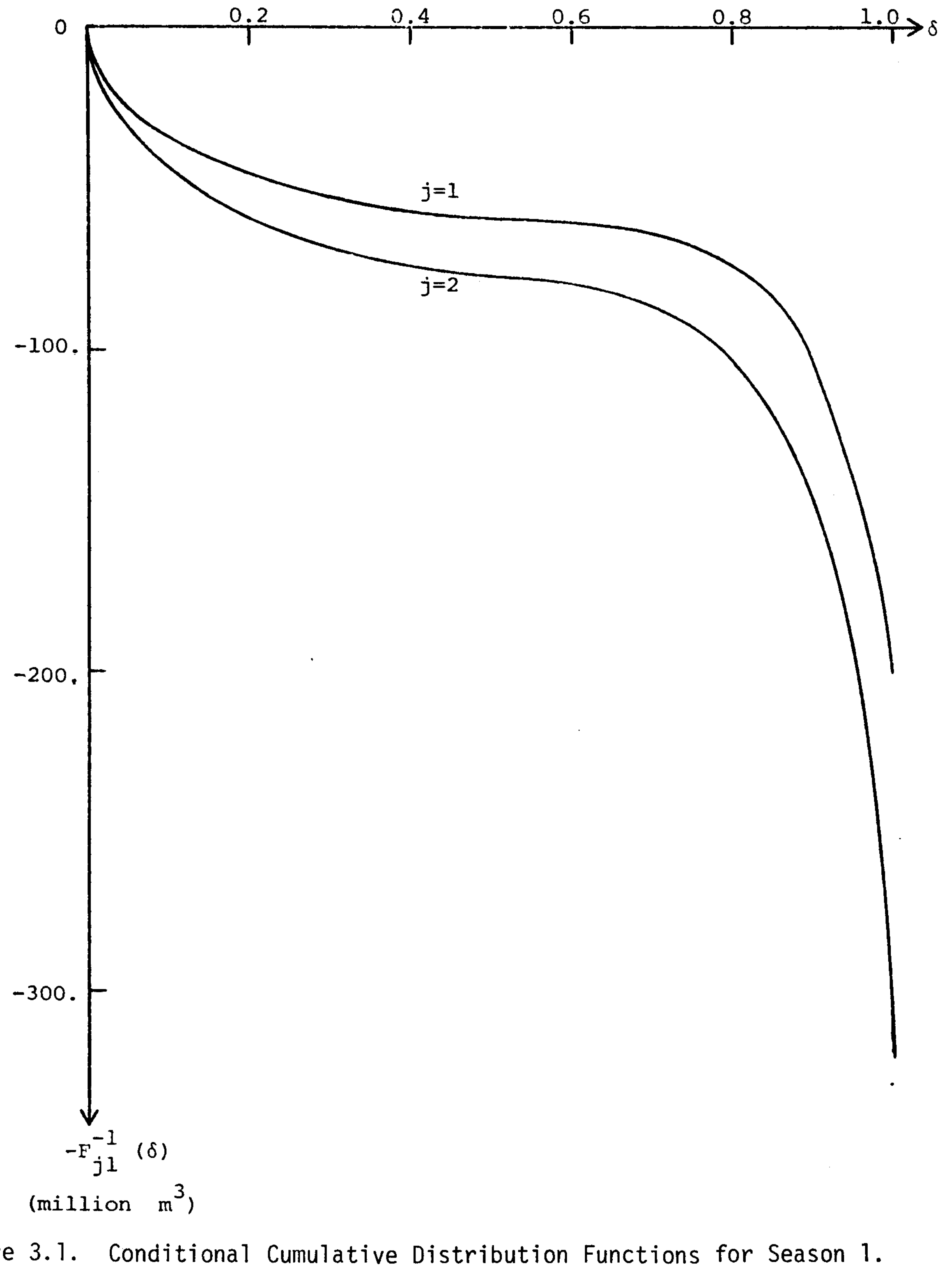


Figure 3.1. Conditional Cumulative Distribution Functions for Season 1.

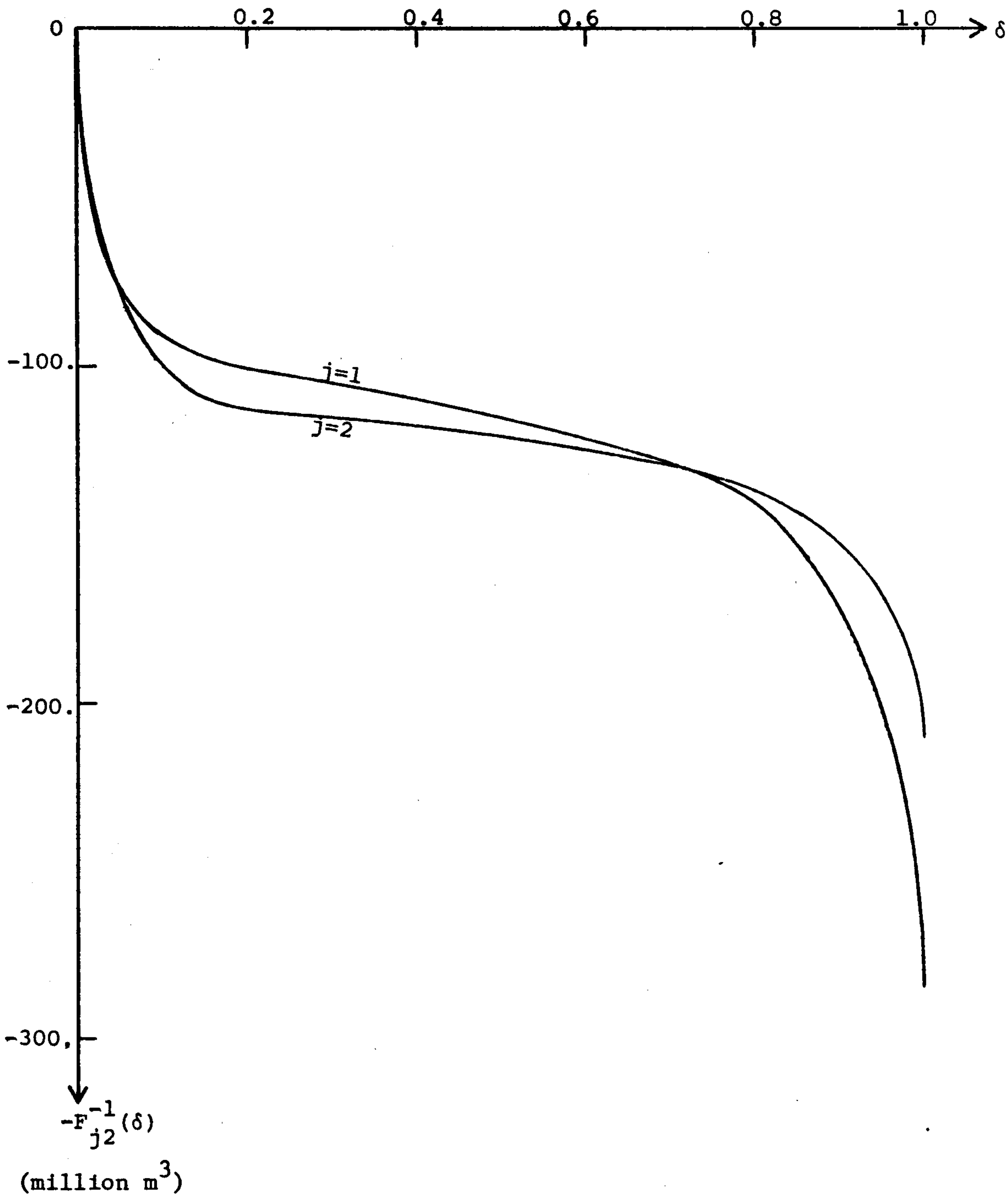


Figure 3.2. Conditional Cumulative Distribution Functions for Season 2.

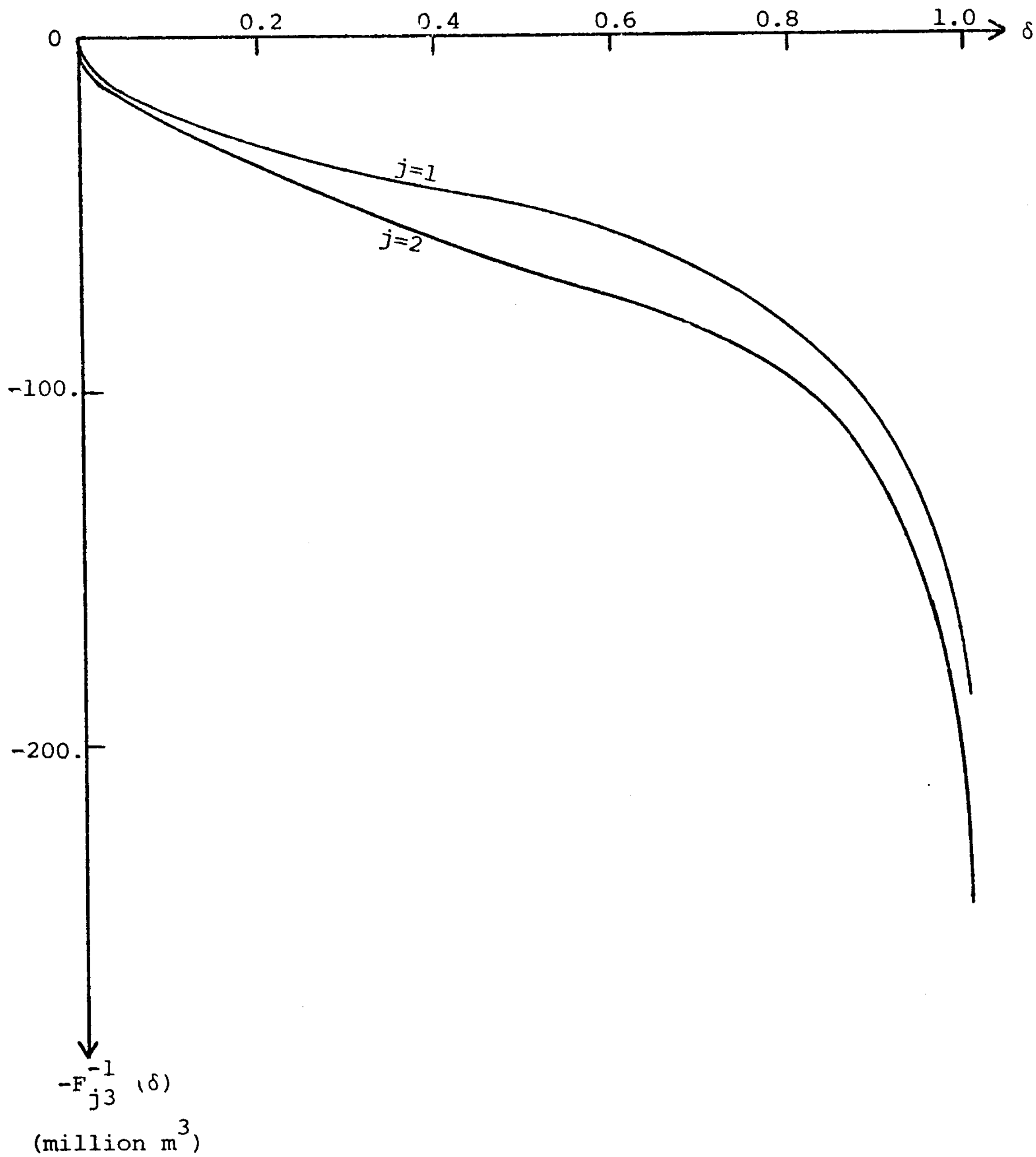


Figure 3.3. Conditional Cumulative Distribution Functions for Season 3.

Season, t	Months Included in Season t	Streamflow Exceeded in 50% of All Seasons t
1	October - January	$67.84 * 10^6$ m^3
2	February - May	$119.65 * 10^6$ m^3
3	June - September	$59.82 * 10^6$ m^3

Table 3.1 Relationship of Seasons to Months and Designation of Streamflow Intervals.

i	t	b_t^i
1	1	-9.905
2	1	-9.097
1	2	-38.462
2	2	-39.932
1	3	-0.698
2	3	+7.794

Table 3.2 Solution Values of b_t^i (10^6 meter^3)

$\underline{i}$	$\underline{j}$	$\underline{t}$	α_t^{ij}	P_t^{ij}
1	1	1	.881	.31
2	1	1	.881	.20
1	2	1	.964	.18
2	2	1	.964	.31
1	1	2	.997	.21
2	1	2	1.000	.27
1	2	2	.998	.29
2	2	2	1.000	.23
1	1	3	.964	.27
2	1	3	.960	.23
1	2	3	.961	.23
2	2	3	.940	.27

Table 3.3 Solution Values of α_t^{ij} and Coefficient Values of P_t^{ij}.

$\underline{j}$	$\underline{t}$	$\underline{\beta_t^j}$	$\underline{\sigma_t^j}$
1	1	.851	.920
2	1	.949	.880
1	2	.890	.939
2	2	.910	.880
1	3	.967	.873
2	3	.964	.933

Table 3.4 Solution Values of β_t^j and γ_t^j

1, 2, and 3 respectively. The probabilities of not exceeding the release specified by the linear decision rules,which also are the probabilities of storage volumes not exceeding the dam capacity, have values between .90 and .91 for each season.

In this exercise, the most time consuming operation was the processing of data. Historical monthly streamflow data for a fifty year period were readily available so that data collection was not a constraining factor. The data were used to calibrate a synthetic streamflow generator (after Valencia and Schaake (1973)) and either the original data or the synthetic data can be used to estimate the conditional cumulative distribution functions, to define the intervals designated by i and j, and to estimate the probabilities of streamflows occuring in those intervals. Following these time consuming tasks, the relatively simple preparing and solving of the linear program was undertaken.

3.4 Summary

A chance constrained linear programming model which employs multiple linear decision rules has been devleoped. The model incorporates explicitly the stochastic nature of the streamflow process, can be used in design and/or management situations, does not significantly restrict the operating policy prior to solution, and is economically and computationally feasible. A portion of the Yakima River system is modeled to demonstrate the use of the multiple linear decision rules.

4. SUMMARY

During the course of fifteen months of investigation, the major optimization models which have been discussed in the literature and which can be used as tools in water resource planning have been examined. Descriptions of three of these models -- Multi-Year Deterministic Linear Program, Multi-Year Deterministic Dynamic Program, and Explicitly Stochastic Linear Program -- are presented in chapter one. Each has major limitations. They all have very large (or prohibitively large) computational burdens when more than two or three reservoir planning problems are examined. The deterministic models rely on one or several streamflow sequences to represent the stochastic properties of the flows and do not generate general operating rules directly. And although the Explictly Stochastic Linear Program does not suffer from these two limitations, it is the most costly to solve: even single site planning situations can be too expensive to examine.

A fourth optimization model, The Explicitly Stochastic Dynamic Program, is described in chapter two. This model is superior to the previous three because the stochastic properties of the streamflows are explicitly incorporated in the model, a general operating policy is designated directly, most nonlinearities can be accomodated within the model, and the model is computationally feasible for one, two and three reservoir problems. It is still limited by a computational cost that increases as the reservoir system size increases but not to the extent that The Explicitly Stochastic Linear Program is. The dynamic program is also limited because it cannot be used to identify optimal system designs. However, chapter two includes a discussion of how multiobjective analysis can be used with the dynamic program to generate tradeoff curves which can be used to identify optimal designs.

Finally, a Chance Constrained Stochastic Linear Program is presented and discussed in chapter three. This is a new model which is based on previous work, principally by ReVelle. The model can be used to identify general optimal operating policies as well as designs. It explicitly incorporates streamflows' stochasticities, and it is much less computationally burdensome than any of the other models. Until additional testing can be performed, it appears to be superior in every way to all of the other optimization models.

REFERENCES

Armstrong, E. L. "Irrigation and Hydropower," EOS: Transactions of the American Geophysical Union, 57(11), 791-797, November, 1976.

Becker, L., and W. W-G. Yeh. "Optimization of Real Time Operation of a Multiple-Reservoir System," Water Resources Research, 10(6), 1107-1112, 1974.

Buras, N. "Dynamic Programming in Water Resources Development," in Advances in Hydroscience, 3. ed. V. T. Chow. New York: Academic Press, 1966.

Burges, S. J., D. P. Lettenmaier, and C. L. Bates. "Properties of the Three-Parameter Log Normal Probability Distribution," Water Resources Research, 11(2), 229-235, April, 1975.

Burges, S. J., and D. P. Lettenmaier. Operational Comparison of Stochastic Streamflow Generation Procedures, Technical Report Number 45. Seattle, Washington: Hydraulics Laboratory, Department of Civil Engineering, University of Washington, 1975.

Butcher, W. S. "Stochastic Dynamic Programming for Optimum Reservoir Operation," Water Resources Bulletin, 7(1), 115-123, February, 1971.

Caselton, W. F. and S. O. Russell. "Long-Term Operation of Storage Hydro Projects," Journal of the Water Resources Planning and Management Division, ASCE, 102(WRI), 163-176, April, 1976.

Cohon, J. L. and D. H. Marks. "A Review and Evaluation of Multiobjective Programming Techniques," Water Resources Research, 11(2), 208-220, April, 1975.

Doland, J. J. Hydro Power Engineering, New York, New York: The Roland Press, 1954.

Eastman, J., and C. S. ReVelle. "Linear Decision Rule in Reservoir Management and Design 3. Direct Capacity Determination and Intraseasonal Constraints," Water Resources Research, 9(1), 29-42, February, 1973.

Fiering, M. B., and B. B. Jackson. Synthetic Streamflows, Water Resources Monograph, Number 1, American Geophysical Union, 1971.

Gablinger, M., and D. P. Loucks. "Markov Models for Flow Regulation," Journal of the Hydraulics Division ASCE, 96(HY1), January, 1970.

Gundelach, J., and C. S. ReVelle. "Linear Decision Rule in Reservoir Management and Design, 5, A General Algorithm," Water Resources Research, 11(2), 204-207, April, 1975.

Hall, W. A. and J. A. Dracup. Water Resources Systems Engineering. New York, New York: McGraw-Hill Book Company, 1970.

Hoshi, K. and S. J. Burges. "A General Flow Disaggregation Model," submitted to Journal of the Hydraulics Division, ASCE, September, 1976.

Houck, M. H., Explicit Stochastic Linear Programming Models for Comprehensive River Basin Planning. Ph.D. Thesis, The Johns Hopkins University, Baltimore, Maryland, December, 1975.

Houck, M. H., and J. L. Cohon. Optimal Design and Operation of Multi-Reservoir Systems, 2 Volumes, The Johns Hopkins University, June, 1976. (NTIS Accession Numbers: PB-262 017/AS and PB-262 018/AS.)

Houck, M. H., and J. L. Cohon. "Sequential Explicitly Stochastic Linear Programming Methods: A Proposed Method for Design and Management of Multipurpose Reservoir Systems," Water Resources Research, forthcoming.

Houck, M. H., J. L. Cohon, and C. S. ReVelle. "Linear Decision Rule in Reservoir Design and Management 6. Incorporation of Economic Efficiency Benefits and Hydroelectric Energy Production," submitted for publication in Water Resources Research in April, 1977.

Howard, R. A. Dynamic Programming and Markov Processes. New York: Technology Press of the Massachusetts Institute of Technology and John Wiley and Sons, Inc., 1960.

Hufschmidt, M. M. and M. B. Fiering. Simulation Techniques for Design of Water-Resource Systems. Cambridge, Massachusetts: Harvard University Press, 1966.

Jacoby, H. D., and D. P. Loucks. "Combined Use of Optimization and Simulation Models in River Basin Planning," Water Resources Research, 8(6), 1401-1414, December, 1972.

Joeres, E. F., J. C. Liebman, and C. S. ReVelle. "Operating Rules for Joint Operation of Raw Water Sources," Water Resources Research, 7(2), 225-235, April, 1971.

Linsley, R. K., Jr., M. A. Kohler, and J. L. H. Paulhus. Hydrology for Engineers. New York: McGraw-Hill, 1958.

Loucks, D. P. Stochastic Methods for Analyzing River Basin Systems, Technical Report Number 16. Ithaca, New York: Cornell University Water Resources and Marine Sciences Center, August, 1969.

Loucks, D. P., and P. J. Dorfman. "An Evaluation of Some Linear Decision Rules in Chance-Constrained Models for Reservoir Planning and Operation," Water Resources Research, 11(6), 777-782, December, 1975.

Loucks, D. P. and L. M. Falkson. "A Comparison of Some Dynamic, Linear and Policy Iteration Methods for Reservoir Operation," Water Resources Bulletin, 6(3), 384-400, May-June, 1970.

Maass, A., et al. Design of Water-Resource Systems. Cambridge, Massachusetts: Harvard University Press, 1962.

Matalas, N. C., and J. R. Wallis. "Generation of Synthetic Flow Sequences," in A Systems Approach to Water Management. A. K. Biswas. New York: McGraw-Hill, 1976.

ReVelle, C. S., E. Joeres, and W. Kirby. "The Linear Decision Rule in Reservoir Management and Design. 1, Development of the Stochastic Model," Water Resources Research, 5(4), 767-777, August, 1969.

Science and Public Policy Program, University of Oklahoma. Energy Alternatives: A Comparative Analysis. Washington, D. C.: U.S. Government Printing Office, 1975.

Thomas, H. A., Jr., and P. Watermeyer. "Mathematical Models: A Stochastic Sequential Approach," in Design of Water-Resource Systems. Arthur Maass, et al. Cambridge, Massachusetts: Harvard University Press, 1962.

Torabi, M., and F. Mobasheri. "A Stochastic Dynamic Programming Model for the Optimum Operation of a Multi-Purpose Reservoir," Water Resources Bulletin, 9(6), 1089-1099, December, 1972.

Valencia, D. R. and J. C. Schaake. "Dissaggregation Processes in Stochastic Hydrology," Water Resources Research, 9(3), 580-585, 1973.

Young, G. K., Jr. "Finding Reservoir Operating Rules," Journal of the Hydraulics Division ASCE, 93(HY6), 297, 321, November, 1967.

Appendix - Data Analysis and Reduction

A crucial, yet often overlooked aspect of optimization modeling for determining reservoir operating rules is the data reduction requirements of the historical streamflow records. For both the M-YDLP and the M-YDDP models, sequences of inflows are needed which fully represent the stochasticity of the actual streamflows. For the stochastic models, the cumulative distribution functions, as well as joint and transition probabilities need to be estimated from the historical data through the use of some synthetic flow generation scheme. The success of these models is highly dependent on how accurately the synthesized flows reflect the actual statistical properties of the system which is modelled.

The first task undertaken was the analysis of both the arithmetic and the logarithmic statistics of the monthly and annual flow records at Lake Keechelus and Lake Kachess. The marginal distribution properties and correlation structures at both sites were studied and the series was examined for deterministic trends. The primary purpose of the exercise was twofold. First, it provided information used in the determination of the minimum time increment that was required for meaningful results from the optimization models. As stated previously, this analysis showed that there was a high degree of variability of system inflows throughout the year. Consequently, monthly time increments were used during the irrigation season. The second objective of this examination was to provide insight into which synthetic flow generation scheme could best preserve the statistical properties of the streamflow records. Because the output of the streamflow generator was used to form inflow scenarios for the deterministic models and to define the probability matrices for the stochastic models, it was necessary that the flow model selected retain the mean, the standard deviation, the coefficient of skewness, as well as the serial and cross correlations of the historical data.

The data provided by the Bureau of Reclamation expressed flows in the Upper Yakima River system in terms of the end of month contents at each reservoir site and in terms of the streamflows in the river immediately below the dams. It was assumed that evaporation and seepage were neglibible compared to the flow levels which passed through the system. Hence, knowing the change in reservoir storage and the outflow from the reservoir, the net inflow to each reservoir site could be calculated from the continuity equation, or

$$I_{st} = S_{st+1} - S_{st} + R_{st} \qquad \forall\ s,t \tag{A.1}$$

where I_{st} = inflow to reservoir s during time period t

S_{st+1} = reservoir storage at site s at the beginning of time period t+1

S_{st} = reservoir storage at site s at the beginning of time period t

R_{st} = flow in river immediately below reservoir s during time period t.

The records used were monthly flows covering the time period from 1925 through the end of water year 1974.

One minor adjustment of the raw data was required at Lake Kachess for August and September of 1940. Using Equation A.1 gave an inflow of 40,900 acre-feet for August and an inflow of -39,600 acre-feet for September which obviously was an error. However, it was noted that the combined inflow for these two months was 1300 acre-feet which fell within the range of flows observed during the other 49 years for the same period. Using the monthly averages from the remaining historical record for August and September (3490 acre-feet monthly average for August and 3070 acre-feet monthly average for September), the 1300 acre-feet was divided between August and September of 1940 based on the ratio of these monthly averages. This gave an August, 1940 net inflow of 700 acre-feet and a September, 1940 net inflow of 600 acre-feet.

After making the necessary adjustments to the raw data, the monthly statistics of the historical inflow records were examined to determine which of these statistics (e.g. mean, standard deviation, skewness, and serial and cross-correlation) warranted preserving in a synthetic flow generation scheme. The next task undertaken was an analysis of the effects of long-term hydrologic persistence at the reservoir sites. A statistic h, known as the Hurst coefficient, is a measure of long-term hydrologic persistence, i.e., the tendency for high flows to follow high flows and for low flows to follow low flows over long periods of time. The correlation coefficient, ρ is a measure of short-term persistence. If the generating process of streamflows belongs to the Brownian domain of attraction, which means that h is approximately equal to one-half, then ρ is a measure of the existing short term persistence and a Markov model will generally be an adequate flow generator. If $h \neq \frac{1}{2}$, say in the range of 0.7 or greater, then the use of a generating process characterized by $h = \frac{1}{2}$ will yield synthetic sequences such that over-year storage requirements will be underestimated. In this case, the process lies outside of the Brownian domain of attraction and a fractional noise model would represent a better approach to the generation of synthetic streamflows. Analysis of the historical records at Lake Keechelus and Lake Kachess revealed a

Hurst coefficient, h = 0.57. Hence, the use of a Markov process for the generation of synthetic streamflows appeared adequate.

In addition to the analyses described above, the use of histograms, correlagrams, and time series plots also provided valuable visual information that gave additional insight to the statistical nature of the flow patterns at the sites under investigation. Based on the results of the statistical analysis outlined and a review of the literature on synthetic flow generation methods, it was felt that the disaggregation of annual volumes to obtain seasonal flow volumes appeared to represent a practical approach which avoided many of the problems of seasonal generation models. Annual models which preserve both high frequency (low-lag) and low frequency (high-lag) behavior are well documented and provide the input to disaggregation techniques. One advantage of the disaggregation method is to bypass the issue of what generating mechanism to use for seasonal events. For example, the use of a simple Markov model as a monthly generator does not guarantee the preservation of the first two moments at the annual level, while the disaggregation method can preserve the simple relationships between monthly and annual events in terms of the first two moments. The flow generation model used in this study was a multi-site disaggregation model provided by Hoshi and Burges (1976) that was capable of preserving seasonal skew as well as season-season and site-site correlation structures. Skewed marginal distributions were represented by a three parameter log normal (3 PLN) distribution.

Burges and Lettermaier (1975) have demonstrated, particularly for systems experiencing low to medium water supply demand, that neglecting to model the skewed form of a marginal distribution can have significant consequences for situations where high or low frequency flow persistence structures are encountered. Smoothing outliers in raw data when building a model designed to preserve skewness is crucial. A quantile (q-a) plot was used as a visual aid to determine the adequacy of a particular 3 PLN fit to the smoothed empirical data from the upper Yakima system. Particular attention was paid to preserving the lower tail of the distribution because this represents the low flow situation that presents some of the greatest problems for reservoir operation.

As with the optimization models described, multi-site disaggregation is algebraically, relatively simple. Numerical computation is limited by storage considerations as the number of sites m, and seasons t increase, however. The model provided by Hoshi and Burges was a four season per year program. It was subsequently modified to generate monthly flows, a task that appeared relatively easy on paper, but became a demonstration of computer acrobatics when trying to fit the program within the storage limits of the machine. This same problem was encountered while developing the computer software to calculate the joint and transition probabilities required for the stochastic models. Hence, the computer limitations with respect to storage and execution time that were demonstrated when extending the optimization models to multiple sites, also pose a severe obstacle to the data reduction requirements of the models. Effective usage of these models for water resource systems may have to wait until the development of optimization models with reduced storage requirements or until the advent of more sophisticated computer technology which is both practical and economical.