5-2000

Stochastic optimazation models for planning and operation of a multipurpose water reservoir

Nasreddine Saadouli

To the Graduate Council:

I am submitting herewith a dissertation written by Nasreddine Saadouli entitled "Stochastic optimazation models for planning and operation of a multipurpose water reservoir." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Management Science.

Chanaka Edirisinghe, Major Professor

We have read this dissertation and recommend its acceptance:

Hamaparsum Bozdogan, Kenneth Gilbert, Mandyam Srinivasan

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Nasreddine Saadouli entitled "Stochastic Optimization Models for Planning and Operation of a Multipurpose Water Reservoir." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Management Science.

Chanaka Edirisinghe, Major Professor

We have read this dissertation
and recommend its acceptance:

Accepted for the Council:

Associate Vice Chancellor and
Dean of the Graduate School

Stochastic Optimization Models for Planning and Operation of a Multipurpose Water Reservoir

A Dissertation
Presented for the
Doctor of Philosophy Degree
The University of Tennessee, Knoxville

Nasreddine Saadouli

May 2000

Dedication

This thesis is dedicated to my mother and father whose guidance and constant prayers made me who I am. It is also dedicated to my wife Malika and my daughter Haneen who are the loves of my life.

Acknowledgments

When I decided to pursue a Ph.D. in operations research, I was only motivated by the opportunity to complete my education. I have always envisioned myself a Ph.D. degree holder since my early days in high school. However, I have never imagined that it would require the tremendous effort by myself and several other great personalities around me.

I would like to start with Dr. Edirisinghe. I am at a loss of words to describe my enormous gratitude for him for three great years. I truly admire his vision, his persistence, and most of all his patience. This thesis would have never been a reality without his continuous encouragement and challenges. I would be indebted for him for the rest of my life.

I would like to extend my thanks to my committee members: Dr. Hamaparsum Bozdogan, Dr. Kenneth Gilbert, and Dr. Mandyam Srinivasan. Dr. Srinivasan has always been a source of generous advice and fulfilling kindness. Dr. Bozdogan has always been of tremendous help in the course of my graduate studies. He has been a friend and a motivator. Dr. Gilbert has always been a great professor and a mere pleasure to know. I truly appreciate you taking the time off your busy schedules to

serve as members of my committee and to provide me with invaluable feedback.

I would like to thank Mr. Ike Patterson for his friendship and his tremendous willingness to help in dim situations. He shared with me the office and several moments of happiness, and sometimes even grief, so graciously. I would also take this opportunity to thank Dr. Fowler for his understanding and continuous financial support throughout my graduate studies. Finally, I would like to thankfully acknowledge the partial financial assistance provided by the University of Tennessee's SARIF grant in summer 1999.

Abstract

We consider the capacity determination problem of a hydro reservoir. The reservoir is to be used primarily for hydropower generation; however, commitments on release targets for irrigation as well as mitigation of downstream flood hazards are also secondary objectives. This thesis is concerned with studying the complex interaction among various system reliabilities (power, flood, irrigation, etc.) and to provide decision makers a planning tool for further investigation. The main tools are stochastic programming models that recognize the randomness in the streamflow. A chance constrained programming model and a stochastic programming model with recourse are formulated and solved. The models developed incorporate a special target-priority policy according to given system reliabilities. Optimized values are then used in a simulation model to investigate the system behavior. Detailed computational results are provided and analyzed.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Literature review

1.1 Introduction

Life cannot exist without water. This is a truth held by all humans regardless of their economic, social, or political status. Since the dawn of civilization, man has struggled to secure this life-permitting source. Therefore, it is just common wisdom to protect water resources and to manage water resources with utmost care. One method for managing the water resources since ancient times is to collect water in reservoirs which in turn can be regulated to feed consumption by various entities. For instance, water for irrigation, recreational use, maintaining livestock or even water for generation of hydroelectricity can be made avilable in a timely manner through such regulation. Furthermore, water reservoirs may also be used to mitigate the effects of natural disasters due to floods or droughts. This need to carefully manage water resources has enticed scholars and governing bodies alike to devise strategies for optimal management and operation of such facilities.

In the last three decades in particular, countless efforts have been expended on optimal sizing and operation of water reservoirs. There are many facets in the management of water reservoirs that can be optimized. The earliest water reservoirs in history have focused mainly on maintaining a supply of life-supporting water and mitigating flood hazards by building large controlling dams on water streams. With the development of social systems and the progress of science, the need to build water reservoirs to satisfy more than a single objective became more apparent. This led to the emergence of so-called multi-purpose water reservoirs. In such reservoirs, several objectives are considered simultaneously. The problem of planning the capacity of these reservoirs is no longer a trivial problem. With several and often conflicting objectives, the problem of planning the capacity and release policy of reservoirs becomes very complex. Consequently, the models of Operations Research/Management Science have been used extensively where one wishes to optimize the benefits of a multi-purpose water reservoir under multiple conflicting objectives and constraints on their operation.

Operations Research (OR) is a branch of applied mathematics. It is a multi-purpose discipline that utilizes mathematical models to represent real-life problems and uses mathematical tools to determine optimal solutions or *decisions*. In the context of the reservoir management problem, such a model attempts to determine an optimal tradeoff among many conflicting objectives. Among OR tools there are such modeling techniques as linear/nonlinear optimization, networks modeling, and queueing theory. Applications of such prescriptive OR tools are becoming increasingly valuable due to their powerful analytical capabilities and their abiblity to be used in conjunction with descriptive techniques such as simulation modeling. OR tools have been applied successfully in many areas, for instance, transportation and warehousing problems, computer networks analysis and configuration, healthcare management, financial planning, as well as in water reservoir management and operational problems.

To analyze water reservoir problems one needs a comprehensive understanding of the common terms used in the water reservoir context. We will divide this chapter into three sections. In section 1.2, the reservoir operation practices and procedures are defined and discussed. In section 1.3, several optimization models that have been proposed in the literature are presented. Section 1.4 presents the motivation for the problem setting discussed in this thesis.

1.2 Water reservoirs: purposes, practices, and operation

1.2.1 Water reservoirs defined

A reservoir can be considered to be an intermediate storage space, which acts as a buffer between an uncontrolled supply of water and dem nd for water. It serves the purpose of holding water in it when the supply of water was in excess of the demand, so that demand could better be satisfied in a subsequent period when supply is scarce.

A reservoir has its inputs as water inflows (which are subject to random variations) and output as water released from the reservoir for downstream use. One important question that may arise is that how large should the reservoir size be selected, in order

a.

to satisfy the demand. If there are no restrictions on the size of the reservoir, then the reservoir can be built larger, thereby the input water resource can be fully made use of for better satisfaction of the demand. However, construction costs of reservoir increases with increased dam heights, and hence building excess capacity becomes too costly. The demand for water is generally expressed as long term contracts or agreements that have been framed at the time of feasibility and acceptability of the water resource development project. Building a reservoir with an inadequate capacity often results in nonobservance of such agreements, which could lead to severe consequences, not only economic, but political and social as well. Thus, the optimal size determination becomes a problem of paramount importance. This problem of optimal sizing and allocation of the precious water resource, is complicated because of the uncertainty associated with supply.

Figure 1.1 depicts a typical water reservoir. A reservoir can be considered to be consisting of three major hypothetical storage volumes, as described below.

Dead Storage

This is usually the highest among the following:

1. level imposed for sediment impounding in the reservoir.
2. level imposed by the minimum operating head of turbines used for power generation.
3. level imposed for recreational development.

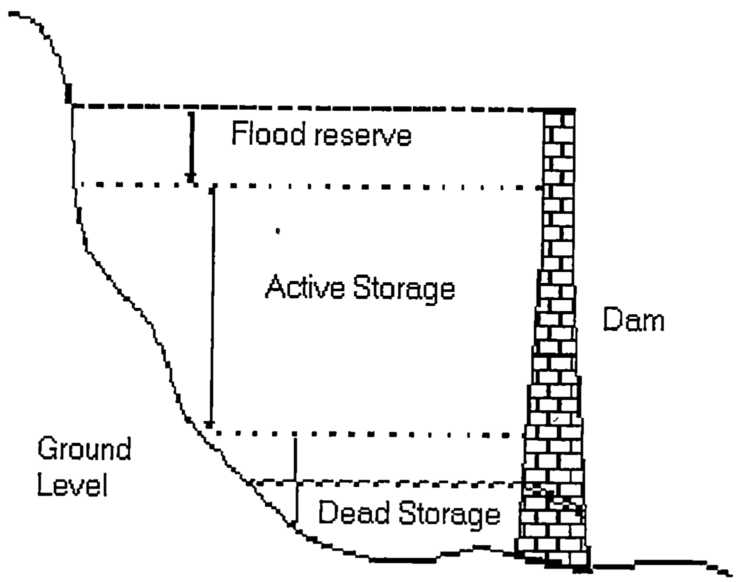

Figure 1.1: Typical water reservoir

Flood Storage

This storage is required to reduce potential downstream flood damage. This reserve is generally determined by the decison maker *a priori* based on the concept of a probable maximum flood occuring with a certain return period (say 50 years). Once the decision is made by the management on a flood occuring with a certain degree of severity, and when the amount of flood water coming into the reservoir is estimated, through flood routing calculations, the flood reserve volume required in the reservoir is generally determined.

Active Storage

This is the storage used for streamflow regulation for purposes of meeting downstream water requirements (arising from agriculture, municipal and industrial uses) and extracting energy as much as possible. If the runoff of the river impounding on the reservoir fluctuates very much according to the season, and also if there is considerable variation in annual runoff, it leads to the necessity of a reservoir with a larger active storage capacity to fully utilize the resources of the river by carrying the water from the wet periods over to dry periods.

1.2.2 Purposes of water reservoirs

Although the name can be misleading, water reservoirs are used for much more than just storing volumes of water or generating hydro-electricity. Water reservoirs are built for either one purpose or multiple purposes, in which case they are termed multipurpose water reservoirs. The purposes of building reservoirs include electricity generation, providing drinking water to human and animal populations, providing irrigation water for farming or industrial needs, and controlling flood hazards.

Irrigation

Most reservoirs are built on water streams or rivers, so they limit, in a sense, the downstream flow of water for agricultural needs. Almost every reservoir authority has to take this into consideration to prevent damage to crops that rely on the flow of water. Irrigation water demands have to be satisfied in general as a part of an agreement between the energy authority and farmer's groups or the government that oversees such agricultural activities. This purpose does not, in general, present any tangible profit to the energy authority. Hence, there is a need to consider this issue without hampering the energy generation which provides the main source of income for reservoir management.

Hydro-energy generation

In addition to the downstream water demand for irrigation, the water released from a reservoir can be utilized to generate electricity (known as hydro-power generation). In such reservoirs, the difference in water altitude is exploited to produce electricity. This is done by releasing water and letting the pressure generated by the flow of water turn the electricity generating turbines. Before the invention of the nuclear power plants, hydro-electricity was the main source of power for most developed nations. In most under-developed nations, this is still the method of choice next to thermal power generation. Its appeal comes mainly from the availability of water in most nations at no cost. Fortunately, a reservoir does not have to release water through two separate ways to satisfy both demands, because the release made to satisfy water demand can be used to generate electricity as well, and vice versa. Generation of hydro electric energy is extremely useful as a substitute for electricity that is normally produced by burning fossil fuel. Therefore, operating these reservoirs in the most optimal manner is a necessity to sustain the operation and growth of industries, as well as providing needed electricity to civilian populations.

Controlling Flood hazards

Water reservoirs are also heavily used as a storage space for excess rain waters that would otherwise overflow and cause floods that are costly to civilian lives and agricultural objectives. The water inflows are therefore stored in multi-purpose reservoirs and released in a controlled manner to mitigate the flood hazard. That is, whenever a flood inflow occurs, that excess water can temporarily be stored in the reservoir, which could in a later period be released, thus eliminating possible flood damages. This practice, known as flood routing, often requires reservoirs be provided with a certain flood storage volume or flood release capacity. This is a very important aspect of the operation of water reservoirs especially in areas where the amount of rain water varies dramatically from one season to the next, for example in monsoon areas. This purpose of water reservoirs, similar to the irrigation purpose, does not provide the energy authority with a tangible revenue. However, governments usually impose such agreements on reservoir authorities because of its tremendous economical and political impact on the civilian population.

Commercial navigation

Since the early days of civilization, water streams have been a preferred means of navigation. This fact remains true even in the twentieth century. Therefore it is imperative that water reservoirs are built with this in mind. Typically, barges in water ways are heavily used in the transportation of bulk commodities such as coal due to the lower energy and cost requirements in transportation. Moreover, the development of reservoirs and water ways is also considered a part of the national defense, since these steams provide inland transportation routes during times of war.

Other purposes

The afore-mentioned purposes of water reservoirs are not exclusive. Depending on its location, a water reservoir can provide a good natural recreational facility for the population. In east Tennessee, for example, there are several public parks on the sides of water reservoirs, such as Melton Hill dam and Norris lake. Other purposes include the protection of wildlife by providing sanctuaries for several species. These will require a certain minimum level of flow to be maintained in the river, which is generally determined according to standards set by administrative authorities based on environmental considerations. Maintainance of these minimum levels depends not only on the reservoir size decision, but also on how the reservoir releases are times. Hall and Dracup(1970) pointed out that the optimum design of any system depends not only on objectives but on the operating policy as well.

1.2.3 Operation of water reservoirs

The operating policy of the reservoir is usually geared towards satisfying the primal purpose of the project. A water release policy is defined as the amount and frequency of water release from the reservoir during a given period of time. A release policy that is primarily designed to generate electricity will be geared towards maximizing the benefits of energy generation. In a reservoir where the downstream water targets are primal, the focus of the policy will be on minimizing the shortage of water supplied for drinking, industrial, or farming needs. Reservoir authorities generally have broad contractual agreements that constrain their releases, and the need for an optimal release policy for operating the reservoir becomes inevitable. In the next section, we present an overview of the literature in this domain, and indicate how researchers used reservoir characteristics to devise objective functions and constraints to optimally plan and control the operation of reservoirs.

Water reservoir models have been studied in a multitude of ways. The most common are descriptive simulation models, prescriptive optimization models, and hybrid combinations of optimization and simulation models (See Army Corps of Engineers, 1991). For the purposes of this thesis, we are mainly interested in the latter two, especially the use of optimization tools such as stochastic programming and validating them using simulation models. It is important to note though that much of the recent literature have focused on assessing the reliability rather than the design and/or planning and operation of reservoirs. For an excellent review of the literature relating to water reservoir models see Louck et al.(1981) and Yeh (1985).

Given the multiple objectives, the notion of optimizing a multi-purpose water reservoir would be somewhat misleading. What one should consider is an optimal trade-off of benefits and costs of the concerned objectives. Benefits could include such things as the revenue from the energy generated, and costs could include construction costs, costs of damages to farming crops due to floods or droughts. Such problems are best studied under an optimization framework of modeling will be the emphasized in the next section.

1.3 Optimization Models Applied to Water Reservoirs

1.3.1 Linear Programming

Linear programming (LP) is a powerful, yet simple, modeling tool that can be used to study a wide variety of application problems. Its ease of understanding and use made it the tool of choice for much of the early applications of optimization modeling in reservoir systems. Due to the complex nature of system constraints in reservoir management and operation, a significant number of researchers have focused on chance-constrained formulations, utilizing linear decision rules. Revelle, Joeres, and Kirby (1969) published one of the key papers titled *"the linear decision rule in reservoir management and design: development of the stochastic model."* Much of the literature that followed built upon and extended the basic concepts laid out in that pioneering paper, for example Loucks and Dorfman (1975), Nayak and Arora (1971), Curry and Helm (1973), Houck (1979), Houck and Datta (1981), and Sreenivasan and Vedula (1996). Several specific LP models have also been presented in the literature. Loucks (1968) developed a stochastic LP model for a single reservoir subject to random net inflows. Release rates were determined which minimized the sum of the expected squared deviations from target reservoir volumes. Windsor (1973) developed a LP model for analyzing multiple reservoir flood control operations. Sreenivasan and Vedula (1996) present a chance constrained LP formulation for a multipurpose reservoir. The model determines a release policy that maximizes the annual hydro power production while meeting the irrigation demand at a specified reliability level.

1.3.2 Dynamic programming

The dynamic programming (DP) approach involves decomposing a complex problem into a series of simpler sub-problems which are solved sequentially, while transmitting essential information from one stage of the decomposition to the next using state space concepts. Louck et. al (1981) state that the generalization of deterministic dynamic programming to the stochastic case is somewhat straightforward. Stochastic dynamic programming can be used to study very complex situations, as long as a few state variables are involved. However, the applicability of dynamic programming is hindered by what is called "the curse of dimensionality", which means that the problem at hand becomes computationally intensive as the number of state variables increases.

Hall et al. (1968) used DP to determine releases over time for a single reservoir that maximized revenues from the sale of water and energy. Giles and Wunderlich(1981) describe a model based on DP and simulation that determines weekly releases and end-of-week storage levels for a 19 reservoir system. Allen and Bridgeman (1986) applied dynamic programming to three case studies involving hydroelectric power scheduling. Martin (1987) incorporated a DP algorithm in a modeling procedure for determining an optimal capacity expansion project for a water supply system. For a method of transforming stochastic dynamic programming models into nonlinear equivalents, see Ziemba (1971).

1.3.3 Nonlinear programming

The use of linear programming , despite its ease and tractability, is restrictive in modeling reservoir operations in many cases. The main reason is that several objective functions or constraints of the model are nonlinear in nature and they have to be linearized to fit in an LP framework. Nonlinear programming (NLP) can be used in such instances to reduce the approximations introduced by linearizing functions, and to provide a more general formulation than LP would. However, the mathematics required is generally much more complex, which generally implies a more expensive computational effort. For that reason, NLP techniques have not been applied extensively to problems of optimizing reservoir system operations. The advances in computer hardware and software in the recent years has alleviated the computational burden, and could result in a greater use of NLP in the future. Duren and Beard (1972) incorporated a univariate gradient search algorithm, with the Newton-Raphson technique, into a reservoir simulation model to develop a method of determining the economically optimum flood control diagram for a single multipurpose reservoir. Rosenthal (1981) used a nonlinear network flow algorithm to maximize benefits in a hydroelectric power system. Diaz and Fontane (1989) used a quadratic programming approach for optimizing hydroelectric power releases from a multiple-reservoir system based on the objective of maximizing economic benefits.

1.3.4 Stochastic Programming

Although stochastic programming can be presented as a special case of linear, dynamic, or nonlinear programming, it is a field that has received much attention recently and deserves to be presented separately. In linear programming, we usually assume complete knowledge of future events, and build models using deterministic constraints with no possibility of recourse, i.e. changing the decision policy after some random event that was approximated is realized. This limitation is overcome by using multi-stage stochastic programming with recourse models. Decisions are made at the beginning of the planning period, however, the model takes into account the randomness in the parameters and it penalizes deviations from their true realizations by using suitable cost functions. Decisions are therefore more informed and provide the decision maker an opportunity for taking corrective action once the random events are realized. In the case of water resrvoirs, the main random event is the amount of water inflow. Other parameters that can be viewed as random are the irrigation water demands.

Under this framework, it is important to distinguish between the so-called "here-and-now", or *non-anticipative*, and "wait-and-see", or *anticipative*, situations. In the former, or non-anticipative, the decision maker takes a decision before observing the stochastic events. In the latter, or anticipative situation, the decision maker is allowed to make a decision after having observed the realization of the stochastic event, in which case the optimal value of the objective function as a result of this decision may itself be a random variable. We shall restrict our focus on *non-anticipative* decision policies. The motivation here is that the reservoir management has to make decisions on releases before observing the stochastic water inflows in our case. The *non-anticipative* situation can be modeled as either a chance-contrained stochastic problem, which we will present in Chapters 2 and 3, or as a stochastic programming model with recourse, which will be presented in Chapter 4.

A computational procedure was presented for a two-stage stochastic LP model by Dantzig (1955), in which the activity levels are determined in the first stage, then a corrective action is followed in the second stage. In his monograph, Prekopa (1995) presents a thorough overview of stochastic programming. A rich theoretical background is provided together with a wide array of applications in such diverse areas as water reservoirs scheduling, inventory control, and banking. Several other researchers have studied the general theory of stochastic programming, see for example, Edirisinghe and Ziemba (1992), Salinetti (1983), Sengupta (1972), Wets (1983a, 1983b), and Ziemba (1974). Edirisinghe (1999) considers bounding techniques within multistage stochastic programming models. Bound-based approximations are proposed as means of solving large problems efficiently.

In the context of water resources planning, Prekopa (1978a) used stochastic programming in a flood control reservoir system design. Neveda (1988) presented several applications of stochastic optimization methods to the electric power system optimization. Multi-stage stochastic models have been used in the power scheduling of hydro-thermal systems under uncertainty on electrical load, see Nowak and Romisch (1999). Several other reserachers have modeled the water reservoir problem using stochastic programming, see for example, Dorfman (1962), Dupacova (1980), Dupacova et al. (1991), Prekopa (1978b), and Rapcsak (1974).

1.4 Multipurpose Water Reservoir with Target-Priority

1.4.1 Motivation for the Problem Setting

We consider that a choice on a particular dam site has already been made. Then depending on the geological conditions, the type of the dam (rockfill, earthen, concrete types, etc.) can be decided, and with this information the actual dam construction cost for a given dam height can be worked out, Apart from the increased capital cost required for a larger dam construction, there is an environmental or social problem associated with a larger reservoir. That is, if the reservoir is to be built large then it requires that a larger area be inundated, which may include certain resiential areas. Thus it may be necessary to establish compensation and resettlement programs to mitigate adverse effects on the existing inhabitants affected in the proposed reservoir area. Therefore, in practice, reservoir sizing decision is much complicated by the issues of socio-economic impact, downstream river development, probable environmental effects, and resettlement and relocation aspects.

1.4.2 Target-Priority Operation

In this thesis, we consider a situation witnessed in a number of reservoir applications where an energy authority is responsible for the reservoir operation with its main objective being generation of hydropower. It is assumed that the problem of choosing a location for the hydro plant has already been made; however, its capacity has yet to be determined. There is also a demand for water arising from downstream irrigation requirements. Although satisfaction of these irrigation targets does not bring explicit benefits to the energy authority, it is required to ensure an acceptable level of reliability in meeting these water targets due to contractual agreements. In recognition of these dual purpose services, in actual operation of such reservoirs, operations managers are generally required to adopt a priority based operating policy, which could be expressed as follows: *If the target water quantity can be released while the total amount of released water is used for generation of energy, then at least the targeted amount of water must be released.* In other words, so long as hydro power can be generated, managers choose to satisfy water needs occurring in any given time period. In this thesis, we consider a capacity-planning model that can accommodate such a target-priority policy. The premise on which this policy is based is that if the water level in the reservoir falls below a certain minimum level, as determined by the hydro turbine minimum head requirement, no hydro-power can be generated.

The capacity planning model is formulated to maximize the monetary net benefit due to power generation over a projected operational horizon, subject to constraints on reliabilities of operation. Apart from satisfying irrigation and power generation needs, the reservoir is also operated to mitigate downstream flood hazards. During flood inflow seasons, the reservoir is used for temporary storage of flood water that can be released in a controlled manner in later periods, thereby eliminating potential floods. This practice often requires reservoirs to be provided with a minimum flood storage volume or flood reserve capacity, see Figure 1.1. The flood reserve capacity to be maintained in any time period is assumed given and it is typically determined outside the operational models, considering historic flood inflow patterns.

Since flood reserve and dead storage volumes are regarded as given, the problem of reservoir capacity determination becomes a matter of optimally sizing the active storage as shown in Figure 1.1.

1.4.3 Organization of The Thesis

This thesis considers a single multipurpose reservoir planning problem with target-priority operation. The proposed models are intended to provide the decision maker a good insight into the problem, so that further socio-economic analyses could be made use of before making a final decision. In chapter 2, a chance constrained approach to the capacity planning for a multipurpose water reservoir with target priority operation is presented in detail. In chapter 3, a real life data set of cumulative inflows is analyzed, and the model developed in chapter 2 is executed to validate the solutions provided by the model. In chapter 4, we introduce a multi-stage stochastic programming approach that can be applied to determine real-time releases. In chapter 5, section 5.2, we present a dynamic programming model with independent inflows. In section 5.3, we develop a dynamic programming model with restricted dependence on the monthly inflows. In chapter 6, we present an aggregated dynamic programming model that will capture the general case of dependent inflows. Finally, in the concluding chapter 7, we summarize the results of the different approaches we have developed and provide guidelines for using these complex models in planning and controlling multi-purpose water reservoirs.

Chapter 2

Chance constrained model

2.1 Chance Constraints

Chance constrained programming is a deterministic model except that certain or all of the constraints are required to be satisfied only in a probabilistic sense. The idea of chance constraints for LP optimization was first introduced by Charnes et al. (1958), for determining refinery rates for heating oils to meet stochastic weather dependent demands. In the context of reservoir system optimization, the idea of chance constraints was first proposed by ReVelle et al. (1969). A chance constrained programming problem has the following structure:

$$\min \quad CX \tag{2.1}$$

subject to

$$AX = b \tag{2.2}$$

$$\mathcal{P}(TX \geq \zeta) \geq p \tag{2.3}$$

$$X \geq 0 \tag{2.4}$$

where $\mathcal{P}$ denotes the probability operator, X is the decision vector, A and T are deterministic (fixed) technology matrices, and ζ is a random vector. The constraint (2.3) is a chance constraint because the event $TX \geq \zeta$ is required to be satisfied only with a given probability $p \quad (< 1)$. If T has only one row and ζ is a univariate random variable, then (2.3) is called an individual chance constraint; otherwise if T is a matrix and ζ is a vector then (2.3) is a "joint" chance constraint. In the latter event, the satisfaction of the set of constraints $TX \geq \zeta$ is given a single probability.

In the case of a single chance constraint, if the probability distribution of the random variable ζ is known, (2.3) can be converted to a deterministic equivalent by using the cumulative probability distribution function of the random variable ζ, denoted by $\mathcal{F}_\zeta$. That is,

$$\mathcal{P}(TX \geq \zeta) \geq p$$

holds if and only if

$$\mathcal{F}_\zeta(TX) \geq p$$

which thus implies that the resulting deterministic equivalent is

$$TX \geq \mathcal{F}_\zeta^{-1}(p) \tag{2.5}$$

where $\mathcal{F}_\zeta^{-1}(.)$ is the inverse cumulative density of the random variable ζ. For example, if p is chosen to be 0.9 then there will be, at most, 0.1 or 10% probability that the constraint represented by (2.3) will not be met.

2.2 Literature Review

Chance Constrained Programming, (CCP), models based on individual chance constraints have been introduced by Charnes, Cooper and Symonds (1958). The more general joint chance constrained models where the random right-hand-side vector is allowed to have stochastically dependent components was first introduced by Prekopa (1970). The issue of transforming joint chance constraints into their deterministic equivalents has been extensively researched by Prekopa, see for example Prekopa (1999) where he shows how probability bounds can be incorporated in CCP models in order to compute approximate solutions. In his monograph, Prekopa (1995) also presents a thorough overview of the theory and applications of stochastic programming including the case of joint chance constraints in the context of water resources management. Sengupta (1972) describes the idea of incorporating decisions on system reliability into a CCP model in a more general setting. ReVelle (1999) discusses several models for optimizing reservoir resources including models for reservoir reliability.

Dupacova et al. (1991) describe the case of a multi-purpose water reservoir. Three different formulations are described: a chance-constrained model, a stochastic program with penalties model, and a mixed model. The first model attempts to minimize the capacity of the reservoir subject to chance constraints on the reliability of meeting demand for water for irrigation and industrial purposes. The second and third models minimize the expcted total cost which includes losses due to failure of some reservoir functions. The treatment in this chapter differs from that in Dupacova et al.'s in at least three dimensions. Firstly, we present a less aggregated model where monthly time periods are used. Secondly, we consider the problem from a net benefit perspective, where the benefit from generating energy is considered explicitly in the objective function. Thirdly, we adopt a target-priority policy that requires the model to choose a release that is greater than the targeted amount of water provided the release will result in energy generation. These distinctions will become apparent as the problem and the model are formally presented in subsequent sections.

In the proposed chance constrained planning model, maintaining minimum flood reserve capacity and dead storage, satisfaction of water targets, and generation of electricity are all considered in probabilistic sense, i.e., using chance constraints. Moreover, some of these constraints are of reliability type since probabilities of satisfaction for these constraints are not available a priori, and thus, they need to be treated as model parameters. In the sequel, we take a two-phase approach in addressing the capacity-planning problem:

(i) in phase one, optimize the energy generation for given reliability and capacity parameters, and

(ii) in phase two, optimize the net benefit of system operation to determine optimal capacity and reliabilities.

We assume that the hydro reservoir is to be used for providing base energy or firm energy with guaranteed reliability. The approach in phase one, see (i) above, captures the stated target priority policy within a chance-constrained setting of system operation to maximize the firm energy output. This is the focus of discussion in Sections 2 and 3. The approach in phase two, see (ii), which determines the optimal capacity that maximizes the system monetary net benefit of power generation is considered in Section 4. The solution of the resulting nonlinear model requires the solution of the model in (i) iteratively, and thus, it is computationally tedious. We propose a Quasi-Newton line search based on gradient estimation for efficient solution of the model, utilizing an important sensitivity theorem in nonlinear programming, see Section 5. The case study that motivated the present work is in Section 6, which also includes a simulation analysis of the stochastic optimization model output to verify the validity of the approach. This optimization-simulation framework is used for further analysis and for providing insight into the capacity decision problem.

The notation required for subsequent development of the model would be introduced as it becomes necessary.

2.3 Decision Policy

Consider a time period t of reservoir operation. In this presentation, each period of operation corresponds to a month. The total number of months in the planning horizon is 12, thus repesenting a year. Denote the volume of water released in period t by R_t. The amount of inflow to the reservoir in period t is denoted by I_t while the reservoir storage at the beginning of period t is denoted by S_{t-1}. In general, R_t would be a function of the history $h_t \equiv \{S_0, R_1, ..., R_{t-1}, S_{t-1}, I_t\}$, i.e., $R_t \equiv f_t(h_t)$.

The set of releases $f_1(h_1), f_2(h_2), \dots, f_t(h_t), \dots$ constitute a release policy. Moreover, notice that such a release policy is non-anticipative, namely, release decisions depend only on the random quantities realized so far and not on the hindsight of future occurrences. Multi-period stochastic optimization models incorporating such non-anticipative decision policy are extremely difficult to solve as the problem size increases, see Dempster (1980). Therefore, it is customary to assume a convenient functional form for $f_t(h_t)$. While there can be a potentially large set of decision rules, the one that is frequently used in water reservoir studies is the linear class of decision rules, i.e. LDR, denoted by Δ_1.

On the other hand, there is the class of so called zero-order rules, denoted by Δ_0, which require that the decision to be made at any time period be independent of the observations of the random variables at all previous time periods. Gartska and Wets (1974) give a detailed theoretical account of the drawbacks of restricting the optimal policy to the class of Δ_0 or Δ_1. For example, there are problems for which there exist optimal decision rules, but no feasible Δ_0 rule, and vice versa. Also, the possibility exists that there are feasible solutions to the original problem but no feasible solution to the equivalent LDR problem, and vice versa. Although one cannot hope to find a 'true' optimal policy when the search is restricted to Δ_0 or Δ_1, for computational appeal, such have been used in the past literature. Prekopa (1995, pp. 231-252) presents stochastic models applied to the optimization of water resources where chance constraints are used together with a zero-order decision rule.

For criticisms on using LDR in reservoir studies, see Curry et al. (1973), LeClerc and Marks (1973), Loucks and Dorfman (1975), or Stedinger (1984).

In this formulation, we restrict the search of release policies to the class Δ_0. The primary motivation for this restriction is that we are dealing with a long-term planning problem. Moreover, in a chance-constrained model with a one-year horizon and 12 monthly periods, the computational advantage of using such a decision rule is that it makes the multi-stage program collapse to a single-stage program. For details, see Dirickx and Jennergren (1975), and for some generalizations, see Ziemba (1971). The simulation study, reported in Section 6, establishes that the fixed Δ_0-optimal policy from the chance-constrained model achieves the reservoir operational characteristics as intended. Indeed, the Δ_0-policy that we seek must satisfy the target-priority behavior.

2.3.1 Target-priority operating policy

As mentioned previously, the model has to accommodate the target-priority policy. However, the hydro-turbines require a certain minimum water-head for generating electricity, a level usually referred to as the dead storage level of a hydro-reservoir. We denote this fixed storage by SD. In this regard, a necessary condition to be satisfied, so that the target water amount released in period t can be used to generate energy, is

$$S_{t-1} + I_t - T_t \geq SD \tag{2.6}$$

where T_t is the given deterministic water target for period t, $t = 1, \ldots, 12$. In the event the condition in (2.6) is satisfied, the model must accomplish a release $R_t \geq T_t$. Now, since $\{R_t\} \in \Delta_0$ and targets T_t are deterministic values, the satisfaction of water targets can be controlled via the constraint in (2.6). Consequently, the probability of satisfaction of (2.6) will be used in the model as a surrogate for the "probability of target satisfaction", denoted by θ. Thus, the target-priority policy in the context of the model can be stated as follows:

(P1) If the condition in (2.6) is satisfied during some period t, then the model is required to choose a release R_t not less than T_t.

The proposed approach for modeling the policy (P1) is described in the next section under the development of chance constraints.

2.4 Chance Constrained Model

The energy authority measures its monetary benefit in terms of the firm energy it generates, as this reservoir will be used to provide base energy. Firm energy is defined as the monthly minimum guaranteed energy generation level, herein denoted by L. $B(L)$ denotes the discounted monetary benefit from the generation of a firm energy level L over a specified planning horizon. The discounted capacity construction cost is denoted by $C(K)$ associated with a reservoir of capacity K. For a chosen capacity K, the energy authority is concerned with maximizing the firm energy generation, provided the system constraints are satisfied, including that of target priority water demand satisfaction. The maximized firm energy level is therefore dependent on the chosen capacity K. The trade-off between capacity construction costs and energy generation benefits will be investigated through maximizing the net benefit function $NB(K, L)$ expressed as,

$$NB(K, L) = B(L) - C(K). \tag{2.7}$$

2.4.1 System Constraints

As mentioned before, the storage level S_t in a given month t must be at least SD, the dead storage level, for energy to be generated. Since S_t is a stochastic variable, the satisfaction of $S_t \geq SD$ can only be made with a certain probability, herein denoted by α. Typically, α may be chosen to be large, say 99%; however, in our case the water target satisfaction is closely associated with the event $S_t \geq SD$, and thus, it is impossible to pre-specify a value for α without knowing its impact on target satisfaction probability, θ. Therefore, we consider the constraints

$$\mathcal{P}(S_t \geq SD) \geq \alpha \,, \forall t = 1, \ldots, 12 \tag{2.8}$$

in our model as reliability constraints where α is treated as a variable that must be chosen optimally to maximize the net benefit in (2.7).

By allowing a hypothetical empty volume of V_t, known as the flood reserve, the reservoir is intended to provide flood protection during month t. The corresponding chance constraint is

$$\mathcal{P}(S_t \leq K - V_t) \geq \beta \,, \forall t = 1, \ldots, 12, \tag{2.9}$$

where β is the flood protection probability pre-specified by the management. The value of V_t is typically determined a priori based on historical data.

Since the reservoir operation spans several years, while the model represents a horizon of one year, it is necessary that the terminal storage of the reservoir should ensure similar operation in the subsequent years as long as the inflow patterns remain the same. This is achieved in the model by requiring that the over-year storage, S_{12}, will remain no less than the initial storage, denoted by S_0, with high probability, denoted by ρ. The resulting over-year storage constraint is

$$\mathcal{P}(S_{12} \geq S_0) \geq \rho. \tag{2.10}$$

As discussed in section 2.3.1, the satisfaction of water targets is accomplished through the surrogate constraint

$$\mathcal{P}(S_{t-1} + I_t - T_t \geq SD) \geq \theta \ , \forall t = 1, \ldots , 12, \tag{2.11}$$

where θ is a probability chosen by the reservoir management. However, a value of θ in (2.11) cannot be specified a priori without knowing its implication on net benefits for the energy authority. We will investigate the effect of θ on reservoir capacity K and other characteristics by analyzing the proposed model for various values of θ. Finally, the continuity equation is given by

$$S_t = S_{t-1} + I_t - R_t \ , \forall t = 1, \ldots , 12, \tag{2.12}$$

assuming no evaporation or other losses. Observe that (2.12) is a physical constraint and it must be satisfied for all random realizations (of scenarios up to time t) of inflow. For the ease of exposition, we collect the set of feasible pairs (S_t, R_t) for all $t = 1, \dots, 12$ satisfying the chance constraints (2.8)–(2.11) and the continuity equation (2.12), and define the feasible set $\mathcal{C}$ as follows:

$$\mathcal{C} := \left\{ \begin{array}{l} (R_t, S_t) \in \Re^2_+, (t = 1, \dots, 12) \ : \\ \\ (R_t, S_t) \quad satisfy \quad (2.8), (2.9), (2.10), (2.11), (2.12) \end{array} \right\}. \tag{2.13}$$

2.4.2 Target-priority model

The firm energy level L must be maximized, for a given capacity K, using a release policy that satisfies the constraint system presented in the preceding section. However, the satisfaction of constraint (2.11) must also ensure that $R_t \geq T_t$ due to our target-priority policy. While there may exist several ways of enforcing this requirement on the model, we choose the following sequential optimization framework: Define the target deficits D_t by

$$D_t := [T_t - R_t]^+ \,, \forall t = 1, \dots, 12,$$

where the notation $[x]^+ \equiv \max\{0, x\}$. Given a sequence of deficits incurred in months $1, \dots, t-1$, i.e., $D_1, \dots, D_{t-1}$, the deficit in month t is minimized subject to the constraints of system operation in the entire time horizon of one year. By doing so, the deficit D_t in period t is chosen such that it will not lead to infeasibility in the future periods $t+1, \dots, 12$ of operation.

The sequence of deficit minimization problems that probabilistically ensure the target priority policy, for $t = 1, \ldots, 12$, is given by:

$$(CCP)_t: \qquad D_t^* := \min \; D_t \qquad (2.14)$$
$$\text{s.t.} \quad R_\tau + D_\tau \geq T_\tau, \;\; \tau = 1, \ldots, t$$
$$D_\tau \leq D_\tau^*, \;\; \tau = 1, \ldots, t-1$$
$$\{R_\tau, S_\tau\}_{\tau=1}^{12} \in \mathcal{C}, D_\tau \geq 0, \;\; \tau = 1, \ldots, t.$$

For $t = 1$, the constraints $D_\tau \leq D_\tau^*$ are considered void in the above formulation.

2.4.3 Firm Energy Generation Model

Corresponding to the sequence of deficits and releases, as determined by the models $(CCP)_1$, ..., $(CCP)_{12}$, the energy generated by the reservoir is denoted by $EG_1, \ldots, EG_{12}$, respectively. Thus, the firm energy level so generated is

$$L_{fe} := \min\{EG_1, \ldots, EG_{12}\}. \qquad (2.15)$$

Noting that the firm energy level L_{fe} is a stochastic variable, the reliability of ensuring a (deterministic level) L for firm energy, denoted by $\bar{\gamma}$, is

$$\bar{\gamma} := \mathcal{P}(L_{fe} \geq L). \qquad (2.16)$$

However, determining the distribution function of L_{fe} is quite complicated. For instance, it requires determining correlations among all random variables EG_t, and it is a statistically onerous task in itself. Therefore, we consider the following set of chance constraints, ensuring that monthly energy production exceeds a certain level L of firm energy with a fixed reliability γ at all time periods, i.e.,

$$\mathcal{P}(EG_t \geq L) \geq \gamma \, , \forall t = 1, \ldots, 12, \qquad (2.17)$$

where γ is a prespecified probability. Given a value for γ in the chance-constrained optimization model, we simulate the optimal policy so obtained to estimate the corresponding $\bar{\gamma}$. These simulation results are discussed in Section 6.

Phase-I of the capacity planning model seeks to maximize this firm energy level L subject to the minimum deficits D_t^* prescribed by the sequential deficit minimization models $(CCP)_t$ and satisfying the constraints of system operation. Consequently, the resulting firm energy maximization model in Phase-I has the format:

$$(FEP): \quad L^* := \max \; L \qquad (2.18)$$
$$\text{s.t.} \quad R_t + D_t \geq T_t, \;\; t = 1, \ldots, 12$$
$$D_t \leq D_t^*, \;\; t = 1, \ldots, 12$$
$$\mathcal{P}(EG_t \geq L) \geq \gamma, \;\; t = 1, \ldots, 12$$
$$\{R_t, S_t\}_{t=1}^{12} \in \mathcal{C}, D_t \geq 0, \;\; t = 1, \ldots, 12.$$

2.4.4 Net Benefit Maximization Model

For chosen model parameters K and α, let the corresponding maximum firm energy level as determined by (FEP) be denoted by $L^*(K, \alpha)$. This yields a net monetary benefit of

$$NB(K, L^*) = B(L^*(K, \alpha)) - C(K). \qquad (2.19)$$

In Phase-II of the proposed capacity planning model, we wish to maximize the net benefit $NB(K, L^*)$ by optimally choosing the capacity K and dead storage reliability α. The explicit constraints on the choice of K and α are given by,

$$0 \leq K \leq K_{\max} \quad \text{and} \quad 0 \leq \alpha \leq 1, \qquad (2.20)$$

where $K_{\max}$ is the maximum possible capacity for the reservoir based on the geographical and morphological limitations of the dam site. However, there will also be implicit constraints on K and α in order to ensure feasibility in the firm energy maximization problem. These will be discussed in a subsequent section.

2.5 Equivalent Deterministic Models

As depicted in Figure 2.1, the proposed methodology involves a two-phase procedure to maximize the net benefit function by searching over reservoir capacity (K) and dead storage reliability (α), for a fixed target satisfaction reliability θ and given values of firm energy reliability γ, flood storage reliability β, over-year storage reliability ρ, and initial storage S_0. In phase-I, for a current iterate (K, α), the twelve sequential programs $(CCP)_1, \ldots, (CCP)_{12}$ are solved to minimize monthly target deficits. These deficits are in turn used within the firm energy maximization program (FEP). Phase-II then finds step-sizes $(\Delta K, \Delta\alpha)$ such that the new iterate of capacity and dead-storage reliability pair yields an improvement of the net benefit function, that is,

$$NB(K + \Delta K, L^*(K + \Delta K, \alpha + \Delta\alpha)) > NB(K, L^*(K, \alpha)). \qquad (2.21)$$

When such a step-size pair for improving the net benefit function can no longer be found, the solution procedure is terminated. In the sequel, we will describe each of these solution phases, along with the necessary algorithmic details.

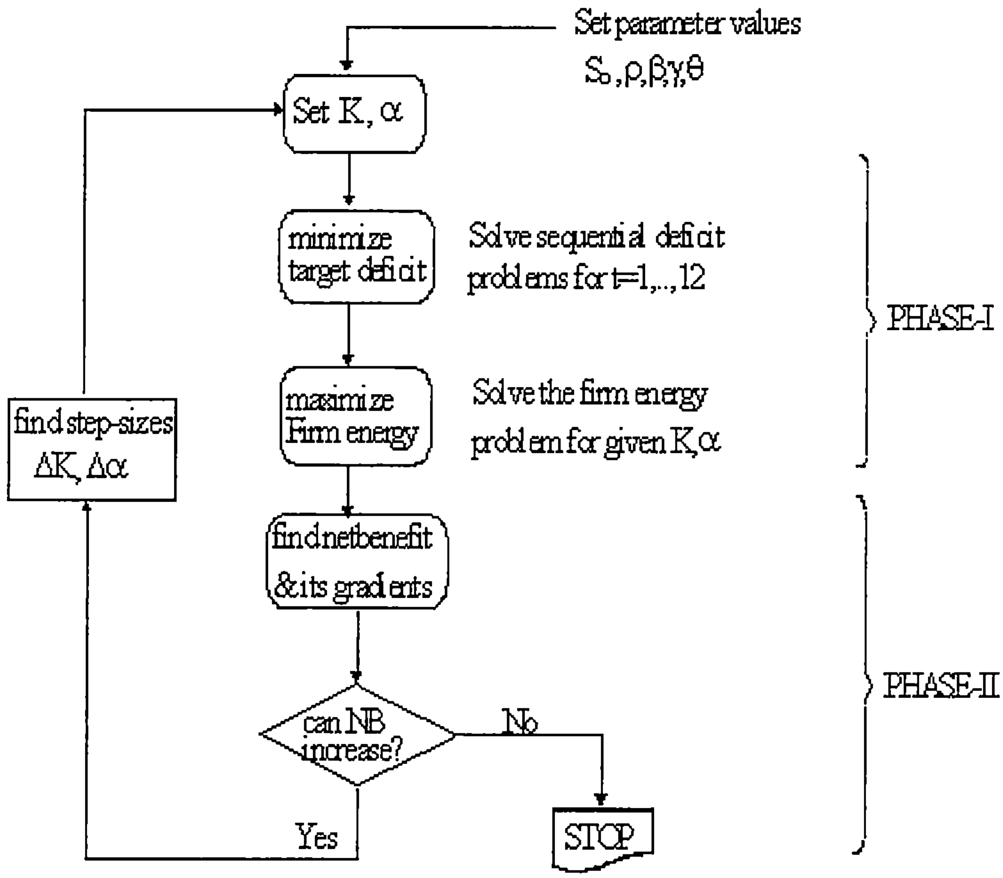

Figure 2.1: Proposed Methodology for the CCP Solution

2.5.1 Phase-I

In order to solve the chance-constrained models of target deficit minimization and firm energy maximization, the probabilistic constraints must first be transformed to their deterministic equivalent forms. This will be done in two steps: first, the target deficit minimization and then, the firm energy maximization.

Step one

The continuity equations in (2.12) imply that

$$\begin{aligned} S_1 &= S_0 + I_1 - R_1 \\ S_2 = S_1 + I_2 - R_2 &= S_0 + (I_1 + I_2) - (R_1 + R_2) \\ &\vdots \\ S_t &= S_0 + \sum_{\tau=1}^{t} I_\tau - \sum_{\tau=1}^{t} R_\tau \end{aligned}$$

or,

$$S_t = S_0 + Q_t - \sum_{\tau=1}^{t} R_\tau \tag{2.22}$$

where we have defined $Q_t := \sum_{\tau=1}^{t} I_\tau$ as the cumulative inflow from period 1 through t. Consider the dead storage chance constraint (2.8):

$$\mathcal{P}(S_t \geq SD) \geq \alpha, \forall t = 1, \ldots, 12$$

From the expression for S_t in (2.22), we can rewrite (2.8) as,

$$\mathcal{P}(S_0 + Q_t - \sum_{\tau=1}^{t} R_\tau \geq SD) \geq \alpha, \forall t = 1, \ldots, 12$$

$$\Longrightarrow \mathcal{P}(\sum_{\tau=1}^{t} R_\tau + SD - S_0 \leq Q_t) \geq \alpha, \forall t = 1, \ldots, 12$$

$$\Longrightarrow \mathcal{P}(Q_t \leq \sum_{\tau=1}^{t} R_\tau + SD - S_0) \leq (1-\alpha), \forall t = 1, \ldots, 12$$

$$\Longrightarrow F_{Q_t}(\sum_{\tau=1}^{t} R_\tau + SD - S_0) \leq (1-\alpha), \forall t = 1, \ldots, 12$$

$$\Longrightarrow \sum_{\tau=1}^{t} R_\tau + SD - S_0 \leq F_{Q_t}^{-1}(1-\alpha), \forall t = 1, \ldots, 12$$

$$\Longrightarrow \sum_{\tau=1}^{t} R_\tau \leq F_{Q_t}^{-1}(1-\alpha) + S_0 - SD, \forall t = 1, \ldots, 12$$

where $F_{Q_t}(.)$ is the cumulative density function of the random variable Q_t, $t = 1, \ldots, 12$. Similarly, we can obtain the deterministic equivalents of the chance constraints (2.8)–(2.11). Therefore, the deterministic equivalents of the chance constraints (2.8), (2.9), (2.10), and (2.11) of $(CCP)_t$ are easily obtained as, respectively,

$$\sum_{\tau=1}^{t} R_\tau \leq F_{Q_t}^{-1}(1-\alpha) + S_0 - SD\,, \quad \forall t = 1, \ldots, 12 \tag{2.23}$$

$$\sum_{\tau=1}^{t} R_\tau \geq F_{Q_t}^{-1}(\beta) + S_0 + V_t - K\,, \quad \forall t = 1, \ldots, 12 \tag{2.24}$$

$$\sum_{t=1}^{12} R_t \leq F_{Q_{12}}^{-1}(1-\rho) \tag{2.25}$$

$$\sum_{\tau=1}^{t-1} R_\tau \leq F_{Q_t}^{-1}(1-\theta) + S_0 - SD - T_t\,, \quad t = 1, \ldots, 12 \tag{2.26}$$

where $F_{Q_t}^{-1}(q)$ represents the q^{th} quantile of Q_t. For $t = 1$ in (2.26), we have set $\sum_{\tau=1}^{t-1} R_\tau = 0$, which thus yields:

$$F_{Q_1}^{-1}(1-\theta) + S_0 - SD - T_1 \geq 0. \tag{2.27}$$

The constraints in set $\mathcal{C}$, defined in (2.13), therefore, have the following equivalent deterministic description:

$$\mathcal{C}_d := \{R_1, \ldots, R_{12} \geq 0 \ : \ \{R_t\} \text{ satisfies } (2.23), (2.24), (2.25), (2.26)\}. \tag{2.28}$$

The deterministic equivalent linear program for deficit minimization problem $(CCP)_t$ in month-t is then given by,

$$\begin{array}{llll} (DP)_t: & D_t^* := & \min & D_t \\ & & \text{s.t.} & R_t + D_t \geq T_t \\ & & & R_\tau \geq T_\tau - D_\tau^*, \ \ \tau = 1, \ldots, t-1 \\ & & & \{R_\tau\}_{\tau=1}^{12} \in \mathcal{C}_d \,, D_t \geq 0, t = 1, \ldots, 12 \end{array} \tag{2.29}$$

Bounds on parameters

It must be noted that under certain combinations of values for parameters α, β, K, ρ and θ, it may turn out that the problem $(DP)_t$ in (2.29), for $t = 1, \ldots, 12$ are infeasible. Using the upper and lower limits on cumulative releases in constraints (2.23)–(2.26), and using $K \leq K_{\max}$, bounds on the latter system parameters can be determined to ensure feasibility of operation. Toward this, define the following thresholds on dead storage reliability:

$$\left.\begin{array}{l} \alpha_t := 1 - F_{Q_t}\left[F_{Q_{t+1}}^{-1}(1-\theta) - T_{t+1}\right], \ \ t = 1, \ldots, 11 \\ \alpha_{12} := 1 - F_{Q_{12}}\left[F_{Q_{12}}^{-1}(1-\rho) - S_0 + SD\right]. \end{array}\right\} \tag{2.30}$$

Furthermore, define the capacity constants for $t = 1, \ldots, 12$:

$$\bar{K}_t := \begin{cases} F_{Q_t}^{-1}(\beta) - F_{Q_{t+1}}^{-1}(1-\theta) + V_t + SD + T_{t+1} & \text{for } t = 1, \ldots, 11 \\ F_{Q_t}^{-1}(\beta) - F_{Q_t}^{-1}(1-\rho) + S_0 + V_t & \text{for } t = 12 \end{cases} \tag{2.31}$$

and consequently, for $t = 1, \ldots, 12$, let:

$$K_t(\alpha) := \begin{cases} \bar{K}_t & \text{if } \alpha \leq \alpha_t \\ F_{Q_t}^{-1}(\beta) - F_{Q_t}^{-1}(1-\alpha) + V_t + SD & \text{if } \alpha > \alpha_t. \end{cases} \tag{2.32}$$

Then, the following bounds on the parameters of the optimization model can be obtained:

Proposition 2.5.1 *For the feasibility of Phase-I problem, the following bounds must hold:*

$$\alpha \leq \alpha_{\max} := 1 - \max_{t=1,\ldots,12} \left\{ F_{Q_t}[F_{Q_t}^{-1}(\beta) + V_t + SD - K_{\max}] \right\}, \tag{2.33}$$

$$\theta \leq \theta_{\max} := 1 - \max_{t=1,\ldots,11} \left\{ F_{Q_{t+1}}[F_{Q_t}^{-1}(\beta) + V_t + SD + T_{t+1} - K_{\max}] \right\}, \tag{2.34}$$

and

$$K \geq K_{\min}(\alpha) := \max_{t=1,\ldots,12} \{K_t(\alpha)\} \tag{2.35}$$

Moreover, the bounds on the initial storage for the feasibility of the Phase-I problem are given by,

$$S_{0,\min} \leq S_0 \leq S_{0,\max} \tag{2.36}$$

where $S_{0,\min}$ and $S_{0,\max}$ are defined by:

$$S_{0,\min} := T_1 + SD - F_{Q_1}^{-1}(1-\theta) \tag{2.37}$$

$$S_{0,\max} := F_{Q_{12}}^{-1}(1-\rho) - F_{Q_{12}}^{-1}(\beta) - V_{12} + K_{\max}. \tag{2.38}$$

Proof. See Appendix A. ■

2.5.2 Step two

To determine the energy generation function EG_t, we make the following simplifying (but practical) approximations. First, we approximate the relationship between the water head (h) acting on the turbine and water storage (S) by a linear function as follows:

$$h = eS + f \tag{2.39}$$

where e and f are constants based on a typical operating range of the reservoir. Figure 2.2 depicts the range of linear approximation for the dam site under consideration in this chapter.

Since we are dealing with a capacity planning problem, with the system being observed only at discrete points in time, i.e., $t = 1, \ldots, 12$, the changes the system may undergo in any period are presumed to take place at the end of that period. Consequently, we approximate the effective water head acting on the hydro turbines during any month by the average water head on turbines in that period. Thus, given a transition of the system from S_{t-1} to S_t, associated with a release R_t, the amount of energy generation is

$$\omega R_t[\frac{e}{2}(S_t + S_{t-1}) + f] \tag{2.40}$$

where ω is a dimensional constant that reflects turbine efficiency. For a similar application of this approximation, see ReVelle (1999, pp. 46-48). However, notice that the

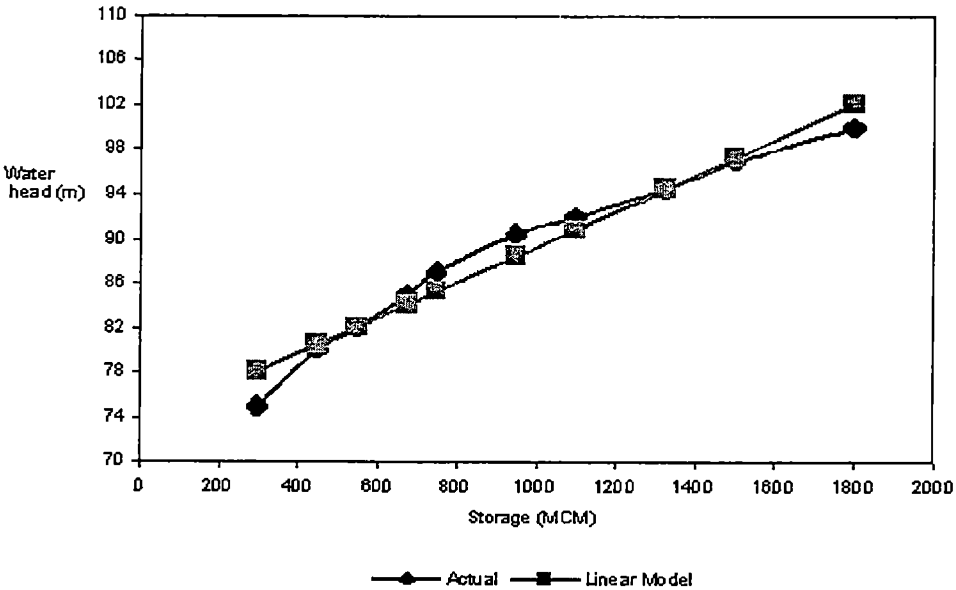

Figure 2.2: Water head and storage relationship

above expression is valid only if both S_{t-1} and S_t are not less than the dead storage, SD, to be maintained. Had one of these storage volumes been less than SD, then only a partial amount of the releases R_t can contribute towards generating electricity. To overcome this obscurity and to compute an average of the energy generated in any period t $(t = 1, \ldots, 12)$, we simplify the transitions of the system as follows. Transition,

(i) from $\{S_{t-1} \geq SD\}$ to $\{S_t \geq SD\}$ takes place with probability $(\alpha)^2$

(ii) from $\{S_{t-1} < SD\}$ to $\{S_t \geq SD\}$ with probability $\alpha(1-\alpha)$

(iii) from $\{S_{t-1} \geq SD\}$ to $\{S_t < SD\}$ with probability $\alpha(1-\alpha)$

(In this case, the amount of water released that can generate power is $R_t + S_t -$

SD ($< R_t$). However, since the storage S_t is not expected to be significantly below the dead storage SD, released amount in this case is approximated by R_t.)

(iv) From $\{S_{t-1} < SD\}$ to $\{S_t < SD\}$ with probability $(1-\alpha)^2$ does not produce any energy for period t.

In the case where $(S_{t-1} < SD)$ we approximate the average water head by $(SD + S_t)$, and in the case where $(S_t < SD)$ we approximate the average water head by $(S_{t-1} + SD)$. The expected energy production can then be expressed as follows:

$$
\begin{aligned}
EG_t &= (\alpha)^2\omega R_t[\frac{e}{2}(S_{t-1}+S_t)+f] + \alpha(1-\alpha)\omega R_t[\frac{e}{2}(SD+S_t)+f] \\
&\quad + \alpha(1-\alpha)\omega R_t[\frac{e}{2}(SD+S_{t-1})+f] \\
\Longrightarrow \; EG_t &= (\alpha)^2\omega R_t[\frac{e}{2}(S_{t-1}+S_t)+f] + \alpha\omega R_t[\frac{e}{2}(SD+S_t)+f] \\
&\quad - (\alpha)^2\omega R_t[\frac{e}{2}(SD+S_t)+f] + \alpha\omega R_t[\frac{e}{2}(S_{t-1}+SD)+f] \\
&\quad - (\alpha)^2\omega R_t[\frac{e}{2}(S_{t-1}+SD)+f] \\
\Longrightarrow \; EG_t &= (\alpha)^2\omega R_t[\frac{e}{2}(S_{t-1}+S_t)+f] + \alpha\omega R_t[\frac{e}{2}(2SD+S_{t-1}+S_t)+2f] \\
&\quad - (\alpha)^2\omega R_t[\frac{e}{2}(2SD+S_{t-1}+S_t)+2f] \\
\Longrightarrow \; EG_t &= -(\alpha)^2\omega R_t[\frac{e}{2}(2SD)+2f] + \alpha\omega R_t[\frac{e}{2}(2SD+S_{t-1}+S_t)+2f] \\
\Longrightarrow \; EG_t &= \frac{1}{2}\alpha\omega R_t(S_{t-1}+S_t) + \alpha\omega R_t[e(SD)+2f-\alpha e(SD)-\alpha f]
\end{aligned}
$$

Therefore, under these assumptions and approximations, we can express the expected

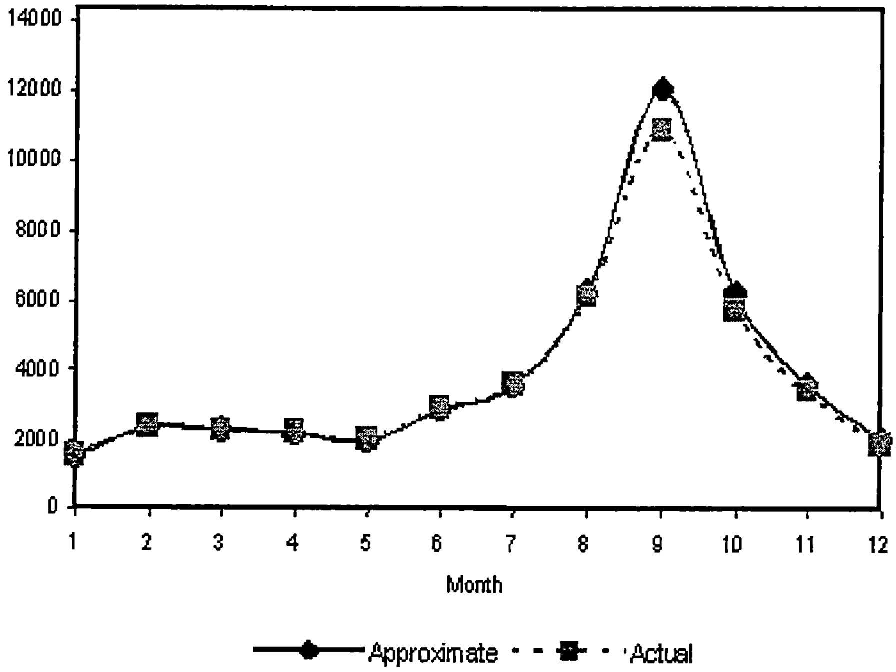

Figure 2.3: Monthly energy computed with approximated vs. actual head

energy production function in the following simplified form, for $t = 1, \ldots, 12$:

$$EG_t = \frac{1}{2}\alpha\omega R_t(S_{t-1} + S_t) + \alpha\omega[2f - \alpha f + e(1-\alpha)SD]R_t. \qquad (2.41)$$

To check the validity of the approximations in (2.39), (2.40), and (2.41), the energy computed by the model using (2.41) are compared with that obtained by simulating the optimal release policy using exact energy computations. One such comparison is depicted in Figure 2.3, which provides the necessary justification. Detailed simulation experiments are discussed in Chapter 3.

Noting that

$$S_{t-1} + S_t = 2S_0 + Q_{t-1} + Q_t - 2\sum_{\tau=1}^{t-1} R_\tau - R_t$$

and denoting the sum of 'adjacent' cumulative inflows by the random variable J_t (:=

$Q_{t-1} + Q_t$), the energy generation function in (2.41) can be restated as:

$$EG_t = \alpha\omega R_t \left[\frac{1}{2} J_t + (S_0 + 2f - \alpha f + e(1-\alpha)SD) - \left(\sum_{\tau=1}^{t-1} R_\tau + \frac{1}{2} R_t \right) \right]. \quad (2.42)$$

The chance constraint $\mathcal{P}(EG_t \geq L) \geq \gamma$ for firm-energy generation can then be written as follows:

$$\mathcal{P}\left(\alpha\omega R_t \left[\frac{1}{2} J_t + (S_0 + 2f - \alpha f + e(1-\alpha)SD) - \left(\sum_{\tau=1}^{t-1} R_\tau + \frac{1}{2} R_t \right) \right] \geq L \right) \geq \gamma$$

$$\Longrightarrow \mathcal{P}\left(\left[\frac{1}{2} J_t + (S_0 + 2f - \alpha f + e(1-\alpha)SD) - \left(\sum_{\tau=1}^{t-1} R_\tau + \frac{1}{2} R_t \right) \right] \geq \frac{L}{\alpha\omega R_t} \right) \geq \gamma$$

$$\Longrightarrow \mathcal{P}\left(J_t \geq \frac{2L}{\alpha\omega R_t} - 2\left(S_0 + 2f - \alpha f + e(1-\alpha)SD\right) - 2\left(\sum_{\tau=1}^{t-1} R_\tau + \frac{1}{2} R_t \right) \right) \geq \gamma$$

$$\Longrightarrow \mathcal{P}\left(J_t \leq \frac{2L}{\alpha\omega R_t} - 2\left(S_0 + 2f - \alpha f + e(1-\alpha)SD\right) - 2\left(\sum_{\tau=1}^{t-1} R_\tau + \frac{1}{2} R_t \right) \right)$$
$$\leq (1-\gamma)$$

$$\Longrightarrow F_{J_t}\left(\frac{2L}{\alpha\omega R_t} - 2\left(S_0 + 2f - \alpha f + e(1-\alpha)SD\right) - 2\left(\sum_{\tau=1}^{t-1} R_\tau + \frac{1}{2} R_t \right) \right) \leq (1-\gamma)$$

$$\Longrightarrow \frac{2L}{\alpha\omega R_t} - 2\left(S_0 + 2f - \alpha f + e(1-\alpha)SD\right) - 2\left(\sum_{\tau=1}^{t-1} R_\tau + \frac{1}{2} R_t \right) \leq F_{J_t}^{-1}(1-\gamma)$$

$$\Longrightarrow \frac{2L}{\alpha\omega R_t} \leq F_{J_t}^{-1}(1-\gamma) + 2\left(S_0 + 2f - \alpha f + e(1-\alpha)SD\right) - 2\left(\sum_{\tau=1}^{t-1} R_\tau + \frac{1}{2} R_t \right)$$

Multiplying both sides by $\frac{\alpha\omega R_t}{2}$, $\mathcal{P}(EG_T \geq L) \geq \gamma$ has the equivalent deterministic form:

$$L \leq \alpha\omega R_t \left[\frac{1}{2} F_{J_t}^{-1}(1-\gamma) + (S_0 + 2f - \alpha f + e(1-\alpha)SD) - \left(\sum_{\tau=1}^{t-1} R_\tau + \frac{1}{2} R_t \right) \right] \quad (2.43)$$

where $F_{J_t}^{-1}$ is the inverse cumulative density function of J_t. After algebraic manipulation, the deterministic equivalent of the firm energy chance constraint becomes:

$$L - c_t(\alpha)R_t - \psi_t(\alpha, R_1, \ldots, R_t) \leq 0\ ,\quad t = 1, \ldots, 12 \tag{2.44}$$

where,

$$c_t(\alpha) := \alpha\omega\left(\frac{1}{2}F_{J_t}^{-1}(1-\gamma) + S_0 + 2f - \alpha f + e(1-\alpha)SD\right) \tag{2.45}$$

and

$$\psi_t(\alpha, R_1, \ldots, R_t) := \alpha\omega\left(\sum_{\tau=1}^{t-1} R_\tau + \frac{1}{2}R_t\right)R_t. \tag{2.46}$$

Observe that the firm energy constraint (2.43) is nonlinear in the release quantities since ψ_t is nonlinear in R_t. Therefore, the firm-energy maximization problem is the following deterministic nonlinear program:

$$\begin{array}{llll} (FP): & L^* := & \max & L \\ & & \text{s.t.} & R_t \geq T_t - D_t^*\ ,\quad t = 1, \ldots, 12 \\ & & & L - c_t(\alpha)R_t - \psi_t(\alpha, R_1, \ldots, R_t) \leq 0\ ,\quad t = 1, \ldots, 12 \\ & & & \{R_t\}_{t=1}^{12} \in \mathcal{C}_d. \end{array} \tag{2.47}$$

2.5.3 Phase two

Having determined the maximum firm energy L^* in Phase-I, see (2.47), according to the current iterate (K, α), Phase-II seeks to maximize the net benefit function by optimally choosing K and α, subject to the bounds computed in (2.33) and (2.35). The resulting net benefit maximization problem is:

$$\begin{array}{lll} (NBP): & \max & NB(K,\alpha) = B(L^*(K,\alpha)) - C(K) \\ & \text{s.t.} & K_{\min}(\alpha) \leq K \leq K_{\max} \\ & & 0 \leq \alpha \leq \alpha_{\max} \end{array} \tag{2.48}$$

provided that the target satisfaction probability is chosen not to exceed $\theta_{\max}$, i.e., $0 \leq \theta \leq \theta_{\max}$, and the initial storage satisfies the bounds: $S_{0,\min} \leq S_0 \leq S_{0,\max}$. Note that the Phase-II problem, (NBP), is a bivariate nonlinear programming problem on variables K and α.

In general, the feasible region of the constraints of (NBP) may not be convex. The latter convexity may be useful in devising solution procedures for (NBP). First, we identify sufficient conditions under which the convexity can be ensured.

Proposition 2.5.2 *Let the p.d.f of Q_t be uni-modal with mode denoted by M_t, for $t = 1, \ldots, 12$. The feasible set of (NBP) is convex if*

$$\theta \geq \theta_{\min} := 1 - \min_{t=1,\ldots,11} \left\{ F_{Q_{t+1}}^{-1} \left[M_t + T_{t+1} \right] \right\} \tag{2.49}$$

and

$$\rho \geq \rho_{\min} := 1 - F_{Q_{12}}^{-1} \left[M_{12} + S_0 - SD \right]. \tag{2.50}$$

Proof. The unimodality of Q_t implies that $F_{Q_t}^{-1}(\alpha)$ is concave for $\alpha \leq F_{Q_t}(M_t)$ and convex for $\alpha \geq F_{Q_t}(M_t)$. Given the definition of $K_t(\alpha)$ in (2.32), requiring $\alpha_t \geq F_{Q_t}(M_t)$ yields that $K_{\min}(\alpha)$ is convex. Noting the definition of α_t in (2.30), the inequalities in (2.49) and (2.50) are obtained. ■

Nevertheless, $NB(K, \alpha)$ may still fail to be convex or concave in its arguments. The difficulty lies in that the optimal value function L^* of (FP) is neither convex nor concave in α, although it is concave in K. In the next section, we describe a procedure for solving (NBP).

2.6 Solution Procedure

The sequential deficit minimization problems in (2.29) of Phase-I are all linear programs which may be solved using standard LP solvers. However, the firm energy maximization problem in (2.47) is a nonlinear program due to the nonlinearity in the energy constraint. Furthermore, Phase-II involves a bivariate nonlinear program, the solution of which would require iterative solution of the programs in Phase-I. In this section, we first discuss solution on the firm energy problem and then the solution details of the Phase-II problem.

2.6.1 Solution of (FP)

This nonlinear program is solved by the following iterative methods. The primary solution procedure is credited to Griffith and Stewart (1961); also see, Bazaraa et al. (1993) for a detailed exposition of the basic algorithm. In this method, the nonlinear program (FP) is solved by successively approximating it by a sequence of linear programs. At the k^{th} iteration of this algorithm, the nonlinear constraint (2.44) is replaced by the first order Taylor's (linear) approximation, evaluated at $\hat{R}_t(k-1)$ where $\hat{R}_t(k-1)$ is the optimal solution of the approximating linear program at iteration $k-1$ for $k > 1$. For $k = 1$, the initial solution $\hat{R}_t(0)$ is set to be the releases provided by the optimal solution of $(DP)_{12}$. Although convergence of this algorithm is not generally guaranteed, it has been reported to be effective for solving many practical problems. Whenever the algorithm fails to converge, or when it cycles, two other secondary solution procedures were incorporated for solving (FP), the first of which is the Merit Function Successive Quadratic Programming (MSQP) algorithm given in Bazaraa et al. (1993). In this case, a Linear Complementary Problem (LCP) is solved using Lemke's Complementary Pivoting Algorithm (CPA). Whenever this procedure also fails to find a solution, the Penalty Successive Linear Programming (PSLP) algorithm is employed as the second secondary solution procedure.

As a necessary condition for optimality, resulting solutions are verified to be Karush-Kuhn-Tucker (KKT) points. Furthermore, the second order sufficient conditions (SOSC) for optimality, see Fiacco (1983, Lemma 3.2.1), are also checked. If SOSC of (2.47) are violated at the current iterate (K, α), then a direct search grid procedure is implemented within a specified neighborhood of (K, α) to seek an improved iterate at which SOSC can be satisfied up to a prescribed tolerance. Having such an iterate in hand, we proceed to solving the (NBP) using a gradient-based technique, as given in the following section.

2.6.2 Solution of (NBP)

A two dimensional search technique based on a quasi-Newton line search algorithm is used to solve the problem (NBP) for a given value of $\theta \in [0, \theta_{\max}]$. For the exposition here, consider the minimization version of (NBP). The iterative search procedure is made more efficient by identifying the gradients of the function $NB(K, \alpha)$ with respect to the variables K and α. Suppose at some iteration k of solving (NBP), the gradient vector of $NB(K, \alpha)$ at the point $x^k := [K^k, \alpha^k]'$ is given by g^k. Then, a descent search direction p^k is set by

$$p^k = -H^k g^k$$

where H^k is a secant approximation of the Hessian matrix of NB at iteration k. Then, a positive step length λ^k is chosen so that the changes to capacity and dead storage reliability, represented by $\lambda^k p^k$, yields a sufficient reduction of the objective value, i.e., negative of the net benefit.

First, a maximum step length $\lambda^k_{\max}$ along p^k is specified, satisfying the constraints of (NBP). Then the backtracking line search strategy (i.e., the reduction of step size from its initial value $\lambda^k_{\max}$ by quadratic and cubic interpolation) is implemented until an acceptable value of λ^k is found (see Dennis and Schnabel, 1983, p.126), which leads to the next iterate $x^{k+1} = x^k + \lambda^k p^k$.

Having determined x^{k+1}, the Hessian matrix update H^{k+1} is obtained by the "Broyden-Fletcher-Goldfarb-Shanno" (BFGS) update procedure, (Dennis and Schnabel, 1983, Luenberger, 1984). The algorithm terminates at some iteration k whenever one of the following conditions is met:

(i) The line search fails to produce a significant change in capacity and dead storage reliability to improve the objective value (i.e., net benefit) along the direction p^k, or,

(ii) The gradient of the objective is zero. The gradient directions of the net benefit function are computed as given in the next section.

Gradients of the net benefit function

The gradients of the net benefit function are derived as follows, assuming that $C(.)$ is differentiable. Clearly,

$$\frac{\partial NB(K,\alpha)}{\partial K} = \frac{\partial B(L)}{\partial L}\frac{\partial L^*(K,\alpha)}{\partial K} - \frac{\partial C(K)}{\partial K} \tag{2.51}$$

and

$$\frac{\partial NB(K,\alpha)}{\partial \alpha} = \frac{\partial B(L)}{\partial L}\frac{\partial L^*(K,\alpha)}{\partial \alpha}. \tag{2.52}$$

Note that the optimal objective value $L^*(K,\alpha)$ of problem (FP) in (2.47) depends on the optimal objective value D_t^* of the problem $(DP)_t$ in (2.29), $t = 1,\ldots,12$. The relevant gradient expressions are derived by applying a sensitivity theorem in nonlinear programming, given in Fiacco (1983, Theorem 3.4.1), under the assumptions stated below.

Assumption 2.6.1 *Given an iterate (K,α), at an optimal solution of the linear program $(DP)_t$ in (2.29), for $t = 1,\ldots,12$, and at an optimal solution of the nonlinear program (FP) in (2.47),*

1. *the gradients of the active constraints are linearly independent*
2. *strict complementary slackness holds, i.e., if a constraint is held as an equality, then the corresponding Lagrange multiplier is strictly positive.*

Proposition 2.6.2 *Given a current iterate (K,α), suppose the SOSC are satisfied for the problem in (2.47). At this iterate, let the optimal Lagrange multipliers associated* *with the constraints identified below of the problem $(DP)_t$ in (2.29) be denoted by:*

$u_{1\tau}^t$: *constraint (2.23) for* $\tau = 1, \ldots, 12$
$u_{2\tau}^t$: *constraint (2.24) for* $\tau = 1, \ldots, 12$
$u_{3\tau}^t$: *constraint* $R_\tau \geq T_\tau - D_\tau^*$, $\tau = 1, \ldots, t-1$

Moreover, let the optimal Lagrange multipliers be denoted by v_i^t for the constraints of the problem (FP) in (2.47):

v_1^t : *constraint (2.23) for* $t = 1, \ldots, 12$
v_2^t : *constraint (2.24) for* $t = 1, \ldots, 12$
v_3^t : *constraint* $R_t \geq T_t - D_t^*$, $t = 1, \ldots, 12$
v_4^t : *constraint (2.44) for* $t = 1, \ldots, 12$

Then, under Assumption 2.6.1, the derivatives $\frac{\partial L*}{\partial K}$ and $\frac{\partial L*}{\partial \alpha}$ are given by:

$$\frac{\partial L^*}{\partial K} = \sum_{t=1}^{12} v_2^t + \sum_{t=1}^{12} v_3^t \frac{\partial D_t^*}{\partial K} \tag{2.53}$$

and

$$\frac{\partial L^*}{\partial \alpha} = \sum_{t=1}^{12} v_3^t \frac{\partial D_t^*}{\partial \alpha} + \sum_{t=1}^{12} v_4^t \left(R_t \frac{\partial c_t(\alpha)}{\partial \alpha} + \frac{\partial \psi_t(.)}{\partial \alpha} \right) + \sum_{t=1}^{12} v_1^t \frac{\partial F_{Q_t}^{-1}(1-\alpha)}{\partial \alpha} \tag{2.54}$$

where $\frac{\partial D_t^*}{\partial K}$ and $\frac{\partial D_t^*}{\partial \alpha}$ are determined for $t = 1, \ldots, 12$ by the recursive equations:

$$\frac{\partial D_t^*}{\partial K} = -\sum_{\tau=1}^{12} u_{2\tau}^t - \sum_{\tau=1}^{t-1} u_{3\tau}^t \frac{\partial D_\tau^*}{\partial K} \tag{2.55}$$

and

$$\frac{\partial D_t^*}{\partial \alpha} = -\sum_{\tau=1}^{12} u_{1\tau}^t \frac{\partial F_{Q_\tau}^{-1}(1-\alpha)}{\partial \alpha} - \sum_{\tau=1}^{t-1} u_{3\tau}^t \frac{\partial D_\tau^*}{\partial \alpha}. \tag{2.56}$$

For $t = 1$, the above notation sets $\sum_{\tau=1}^{t-1}(.) = 0$.

Proof. See Appendix B. ■

The result in Proposition 2.6.2 is invoked only when the required SOSC conditions are satisfied, as indicated previously. Our computational experience with the model suggests that the use of the above gradient expressions can speed up the solution of (NBP) by a several orders of magnitude, as opposed to using a pure direct search technique.

2.7 Concluding Remarks

In this chapter, we proposed a model that captures the multipurpose objectives involved in the design of a water reservoir while recognizing the randomness in streamflow. This is done through a chance-constrained optimization model for which some of the reliabilities are specified a priori by the reservoir authority. The remaining system reliabilities are determined as part of the solution of the capacity determination model. Also, the model allowed the incorporation of power generation aspects of a reservoir to a sufficient degree of accuracy. A deterministic decision rule was assumed for the model which in turn transformed the multi-stage problem to a single stage problem. However, consideration of a special target-priority policy helped retain some of the dynamic aspects of the sequential decision problem. This leads to a multi-staged goal programming formulation which was solved by a two-phased iterative scheme.

In the next chapter, the model is applied to a real life situation. The data obtained for the case study is fed into the model to determine the optimal dam size and an optimal release policy. A simulation analysis of the release policy is carried out to validate the approximations and the modeling approach used in this chapter.

Chapter 3

Application of the CCP Model

3.1 Data Analysis

The two phase optimization model in the preceding chapter is applied to a real-life situation that involves planning the capacity (dam size) of a single hydro-reservoir. This case study comes from the Sai-Buri reservoir project of the Energy Generating Authority of Thailand (EGAT). The historical monthly inflow data is used to form the random variables Q_t (i.e., cumulative inflows), and their empirical distributions are generated. Tables 3.1, 3.2, and 3.3 provide the collection of cumulative monthly inflows from 20 years of observations. The cumulative inflow is shown with the corresponding cumulative probability.

Lognormal distributions are determined to closely approximate these empirical distributions. However, various other types of theoretical distributions have also been used in the literature, such as the uniform or multivariate normal distributions as in Dupacova et al. (1991). The parameters of the lognormal distributions are determined using the method of maximum likelihood estimation and the respective 'goodness-of-

Table 3.1: Cumulative Inflows and corresponding cumulative probabilities (Months 1, 2, 3, 4)

Q_1		Q_2		Q_3		Q_4	
CI(MCM)	CP[a]	CI(MCM)	CP	CI(MCM)	CP	CI(MCM)	CP
20	0.14	50	0.08	75	0.06	80	0.03
35	0.325	80	0.18	110	0.08	125	0.08
50	0.47	110	0.37	145	0.25	170	0.21
65	0.71	140	0.54	180	0.37	215	0.35
80	0.815	170	0.69	215	0.54	260	0.39
95	0.83	200	0.82	250	0.69	305	0.68
110	0.89	230	0.86	285	0.79	350	0.75
125	0.91	260	0.92	320	0.82	395	0.83
140	0.93	290	0.94	355	0.88	440	0.92
155	0.94	320	0.94	390	0.92	485	0.93
170	0.96	350	0.96	425	0.92	530	0.94
185	0.96	380	0.96	460	0.94	575	0.94
200	0.96	410	0.96	495	0.94	620	0.94
215	0.97	440	0.98	530	0.96	665	0.94
230	0.97	470	0.98	565	0.96	710	0.96
245	0.98	500	0.98	600	0.96	755	0.96
260	0.98	530	0.98	635	0.96	800	0.96
285	0.98	560	0.98	670	0.98	845	0.98
300	0.98	590	0.98	705	0.98	890	0.98
315	1	620	1	740	1	935	1

[a]CI: Cumulative Inflow; CP: Cumulative probability

Table 3.2: Cumulative Inflows and corresponding cumulative probabilities (Months 5, 6, 7, 8)

Q_5		Q_6		Q_7		Q_8	
CI(MCM)	CP[a]	CI(MCM)	CP	CI(MCM)	CP	CI(MCM)	CP
100	0.02	140	0.01	250	0.04	360	0.02
155	0.07	200	0.03	315	0.08	444	0.06
210	0.12	260	0.14	380	0.16	528	0.11
265	0.28	320	0.2	445	0.23	612	0.26
320	0.4	380	0.29	510	0.36	696	0.38
375	0.64	440	0.55	575	0.61	780	0.51
430	0.81	500	0.73	640	0.69	864	0.67
485	0.83	560	0.79	705	0.74	948	0.78
540	0.89	620	0.86	770	0.83	1032	0.85
595	0.92	680	0.9	835	0.88	1116	0.87
650	0.94	740	0.9	900	0.88	1200	0.88
705	0.94	800	0.94	965	0.9	1284	0.9
760	0.94	860	0.96	1030	0.91	1368	0.95
815	0.94	920	0.96	1095	0.93	1452	0.95
870	0.96	980	0.96	1160	0.95	1536	0.95
925	0.96	1040	0.96	1225	0.95	1620	0.95
980	0.98	1100	0.98	1290	0.95	1704	0.95
1035	0.98	1160	0.98	1355	0.97	1788	0.95
1090	0.98	1220	0.98	1420	0.97	1872	0.98
1145	1	1280	1	1485	1	1956	1

[a] CI: Cumulative Inflow; CP: Cumulative probability

Table 3.3: Cumulative Inflows and corresponding cumulative probabilities (Months 9, 10, 11, 12)

Q_9		Q_{10}		Q_{11}		Q_{12}	
CI(MCM)	CP[a]	CI(MCM)	CP	CI(MCM)	CP	CI(MCM)	CP
690	0.04	750	0.02	800	0.02	825	0.02
790	0.12	850	0.06	900	0.06	927	0.06
890	0.24	950	0.17	1000	0.14	1029	0.1
990	0.31	1050	0.19	1100	0.16	1131	0.17
1090	0.4	1150	0.35	1200	0.31	1233	0.29
1190	0.5	1250	0.39	1300	0.36	1335	0.37
1290	0.56	1350	0.44	1400	0.41	1437	0.39
1390	0.65	1450	0.54	1500	0.49	1539	0.44
1490	0.74	1550	0.62	1600	0.58	1641	0.54
1590	0.81	1650	0.64	1700	0.6	1743	0.56
1690	0.88	1750	0.74	1800	0.66	1845	0.65
1790	0.89	1850	0.83	1900	0.72	1947	0.72
1890	0.89	1950	0.88	2000	0.82	2049	0.8
1990	0.91	2050	0.88	2100	0.88	2151	0.86
2090	0.92	2150	0.9	2200	0.9	2253	0.91
2190	0.93	2250	0.9	2300	0.92	2355	0.91
2290	0.95	2350	0.94	2400	0.92	2457	0.91
2390	0.95	2450	0.95	2500	0.98	2559	0.95
2490	0.97	2550	0.96	2600	0.98	2661	0.98
2590	1	2650	1	2700	1	2763	1

[a]CI: Cumulative Inflow; CP: Cumulative probability

fit' χ^2-values are given in Table 3.5. The critical value is $\chi^2_{95\%,dof=9} = 16.919$.

An alternative to the traditional hypothesis testing, one could use the more recent information criteria measures, namely *Akaikes's AIC* and *Bozdogan's ICOMP*, see Bozdogan (1990). These criteria define the statistical model as follows:

$$\text{Statistical Model} = \text{Signal} + \text{Noise}.$$

ICOMP, for example is designed to estimate a loss function of the form:

$$\text{Loss} = \text{Lack of Fit} + \text{Lack of Parsimony} + \text{Profusion of Complexity}.$$

By deriving the expression for lack of fit, lack of parsimony, and profusion of complexity for several theoretical possible distributions, AIC and ICOMP will select the model that minimizes the loss function. For example, if we assume that the data is normally distributed, then:

$$AIC(Normal) = nln(2\pi) + nln(\hat{\sigma}^2) + n + 2(2),$$

where $\hat{\sigma}^2$ is the estimated variance. If we assume that data is lognormally distributed, then:

$$AIC(LogNormal) = nln(2\pi) + nln(\hat{\sigma}^2) + n + 2\sum_{i=1}^{n} ln(x_i) + 2(2),$$

where $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n} ln(x_i - \bar{x})^2$ is the estimated variance. If the data is assumed to be exponentially distributed, then:

$$AIC(Exponential) = 2n + 2nln(\bar{x}) + n + 2(1),$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the sample mean. Similarly,

Table 3.4: AIC Goodness of Fit Test

Month	AIC Values				
	Geometric	Poisson	Normal	Lognormal	Exponential
Q_1	246	∞	240	173	266
Q_2	274	∞	267	200	294
Q_3	282	∞	273	207	302
Q_4	291	∞	283	214	311
Q_5	299	∞	291	222	319
Q_6	304	∞	294	228	324
Q_7	312	∞	297	237	332
Q_8	324	∞	308	248	344
Q_9	338	∞	315	263	358
Q_{10}	339	∞	315	265	359
Q_{11}	340	∞	315	266	360
Q_{12}	341	∞	315	267	361

$$AIC(Poisson) = 2n\bar{x}ln(\bar{x}) + 2n\bar{x}\sum_{i=1}^{n} ln\Gamma(x_i + 1) + 2,$$

and

$$AIC(Geometric) = -2n\bar{x}ln(\bar{x}) + 2n(1+\bar{x}) + ln(1+\bar{x}) + 2.$$

Table 3.4 displays the goodness of fit results from using AIC. Note that the AIC criterion is minimized for the LogNormal Assumption indicating that that the lognormal distribution best describes the set of data we have. For more on informational model selection criterion, see Akaike (1973) and Bozdogan (1987, 1988).

The flood reserve capacities as well as the water targets are also reported in Table 3.5. Note that for the present analysis, V_t is constant at 124 MCM for all $t = 1, \ldots, 12$. The planning problem seeks to address three levels of water targets, labeled small, medium, and large targets. The remaining given fixed data is: dead storage SD=140 MCM and maximum reservoir capacity $K_{\max}$= 2,500 MCM.

Table 3.5: Inflow parameters and water targets

month	lognormal fit of Q_t			flood	target level		
	mean(MCM)	std.dev(MCM)	χ^2-value	reserve[a]	small	medium	large
1	66.82	44.49	4.378	124	11.05	38.95	66.85
2	154.01	93.36	1.536	124	33.34	60.25	87.15
3	237.46	128.83	4.025	124	29.91	56.68	83.45
4	320.22	170.66	4.389	124	23.38	53.07	82.76
5	392.29	192.62	7.643	124	23.07	47.57	72.07
6	493.24	219.71	7.151	124	30.52	65.86	101.2
7	631.44	246.65	4.123	124	22.15	80.05	137.95
8	867.83	284.38	3.992	124	30	133.2	236.4
9	1299.84	425.95	7.205	124	20	226.01	432.01
10	1507.19	481.03	10.478	124	15.41	111.76	208.1
11	1618.09	510.81	8.278	124	19.07	65.02	110.96
12	1692.56	534.32	11.244	124	8.45	41.48	74.51

Table 1: Inflow parameters and water targets

[a]For the present analysis, V_t is constant for all t.

Linear approximation of the dam elevation-storage relationship, see (2.39), yields $e = 0.016$m/MCM (t-statistic=15.139) and $f = 28.67$m (t-statistic=65.471). The dam cost-capacity relation was extrapolated from historical records and is determined to be linear with $C(K) = 380.90 + 0.136K$ m฿ (million Baht - Thai currency), see Table 3.6. Using simple linear regression, we get a Pearson correlation coefficient of 0.998 which indicates a strong positive relationship. The t-statistic was computed to be 24.3, and the p-value is 0.4 %. Note, however, that the latter linearity is really not a simplification, as the Phase-2 model can handle nonlinearity (preferably concavity) with equal ease. The constant for energy generation in (2.40) is $\omega = 1.962$, assuming an average efficiency of 90% for the turbine plant and making allowance for frictional head loss along the penstock length by a factor of 80%. The benefit function $B(.)$ is assumed to be linear based on a present worth factor of 0.4310m฿per mega watt-hour,

Table 3.6: Reservoir Construction Cost Data

Reservoir size	Cost (MB)[a]
450	445
675	468.5
950	510
1300	557.5

[a]MB: Million Bahts

i.e. $B(L) = 0.4310L$.

3.2 Model Application and Simulation Analysis

The two-phase optimization model was run for various parameter combinations for γ, ρ, θ, as well as for different initial conditions S_0. The output from the optimization model, i.e., the capacity and the release schedule, is used within a simulation model to investigate the system behavior. While optimization models utilize many simplifying assumptions, simulation models provide an alternative powerful tool for analyzing the system with its real-world complexity. Simulation systems have been studied extensively in the literature, see for example, Shannon (1975) and Law and Kelton (1991). In the context of water resources, ex-post simulation analysis of decisions obtained by optimization models has been recognized as a standard tool and has been researched extensively, see for example Dupacova et al. (1991). For an excellent review of simulation models, see US Army Corps (1991). In particular, simulation is an ideal tool for validating the results given by optimization models that have incorporated approximations for the purpose of solution tractability, as in our case. The specific objectives of the present simulation study are:

1. validating the solution given by the proposed optimization model, and

2. providing further insight into the problem and identifying complex interaction that exists among various system parameters.

3.2.1 Simulation Experiments

For a specified set of input parameters (reliabilities), the optimization model is first run to determine the corresponding optimal reservoir capacity and the associated optimal release schedule. Then, cumulative monthly inflows are simulated according to the specified lognormal parameters. For each random scenario of 12-monthly cumulative inflows generated, the simulation model tracks constraint violations, as well as the energy generation during each month. The simulation is carried out with 100,000 such scenarios generated randomly. Simulation procedure also tracks how often the specified target-priority policy is violated. Consequently, the simulated values of target satisfaction reliability (θ_s), dead storage reliability (α_s), flood storage reliability (β_s), and over-year storage reliability (ρ_s) are computed. These probabilities are evaluated by accumulating the number of times a given event is satisfied and then taking the average over the number of simulation runs (scenarios). A comparison of these simulated reliabilities with the reliabilities used within the optimization model is used as one yardstick for validating the proposed optimization model.

It may be noted that when energy production is computed, the simulation model applies the exact formulae rather than the linear approximations used in the optimization model. We also define the term *potential energy* as the (maximum) amount of energy that could be generated if the available storage can be released without bringing the storage below the dead storage. The simulation model computes both the monthly potential energy (PEG_t) and actual energy generated (EG_t) where the actual energy is computed following the release schedule given by the optimization model. The probability that actual monthly energy exceeds the firm energy level L^* determined by the optimization model, denoted by γ_s, is also computed.

For further validation of the proposed optimization model in terms of energy generation, we develop simulated probability (frequency) distributions of firm energy. Under a particular scenario of 12-monthly inflows, the realized value of the firm energy random variable L_{fe} is given by $L_{fe} = \min\{EG_t \ : \ t = 1, \dots , 12\}$. We focus on the expected firm energy $E[L_{fe}]$ of the distribution of L_{fe}, as well as the probability $\bar{\gamma}_s$ that the simulated firm energy exceeds L^*, where L^* is the firm energy determined by the optimization model corresponding to the input reliability level γ. Note that the simulated reliability $\bar{\gamma}_s$ provides an estimate of the 'true' firm energy generation reliability $\bar{\gamma} = \mathcal{P}\{L_{fe} \geq L^*\}$, see (2.16).

In addition to the validation of the output of the optimization model, sensitivity analysis of the output is also performed via simulation. This analysis focuses on such questions as:

1. the effect of varying water targets on system performance attributes,

2. the effect of initial storage on optimization model output and its simulation, and

3. the effect of varying the release schedule while fixing the reservoir capacity at a certain optimal size.

3.2.2 Computational Results

The optimization model was run for a broad range of input parameters. Generally, the simulation of the optimal capacity and releases confirms that resulting simulated system reliabilities exceed those specified in the optimization model. However, a particularly interesting feature needs to be highlighted here. Consider the case of $\theta = 0.8$ and setting the initial storage S_0 to $S_{0,\min} = 191.15$ - see (2.37) applying medium targets - along with $\rho = 0.95$, $\beta = 0.8$, and $\gamma = 0.9$ in the optimization model. Three different options are investigated: no targets, small targets, and medium targets. No targets case is modeled by setting $T_t = 0$ and $\theta = 0$ in the optimization model.

The simulated reliabilities of the resultant optimal capacity and releases are in Figures 3.1 and 3.2, respectively, for monthly energy production and dead storage reliabilities. In Figure 3.1, the simulated firm energy (monthly) reliability falls below $\gamma = 0.9$ for no targets and small targets, during months 2 through 8. Coincidentally, the optimal dead storage reliabilities as determined by the optimization model for the no target and small target cases remain at low levels of 0.60 and 0.68, respectively, while that for the medium target case is 0.88. This behavior characterizing simulated γ_s falling below the specified γ occurs almost consistently when the corresponding optimal α turns out to be rather small ($\alpha < 0.70$), while simulated β_s or ρ_s remain above the specified reliability levels. This leads us to the conclusion that at low levels

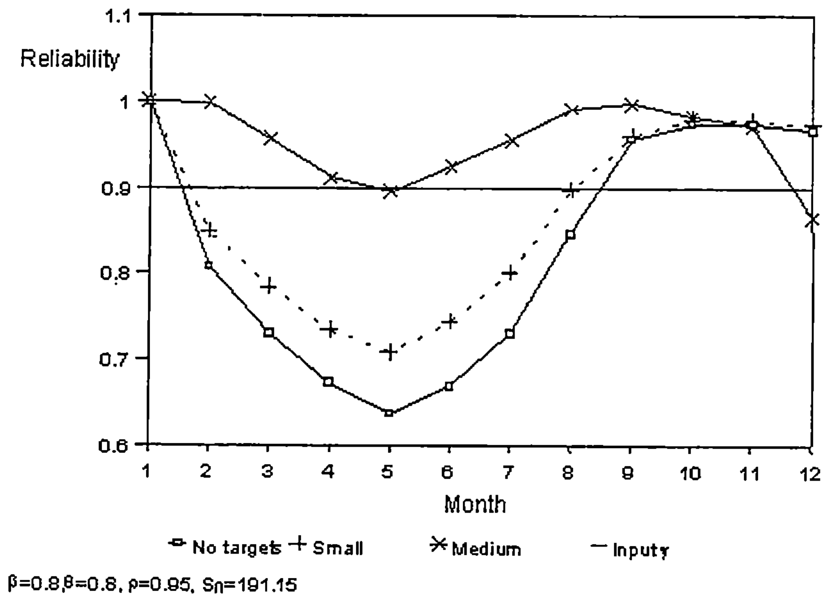

Figure 3.1: Simulated firm energy generation reliability for S_0=191.15

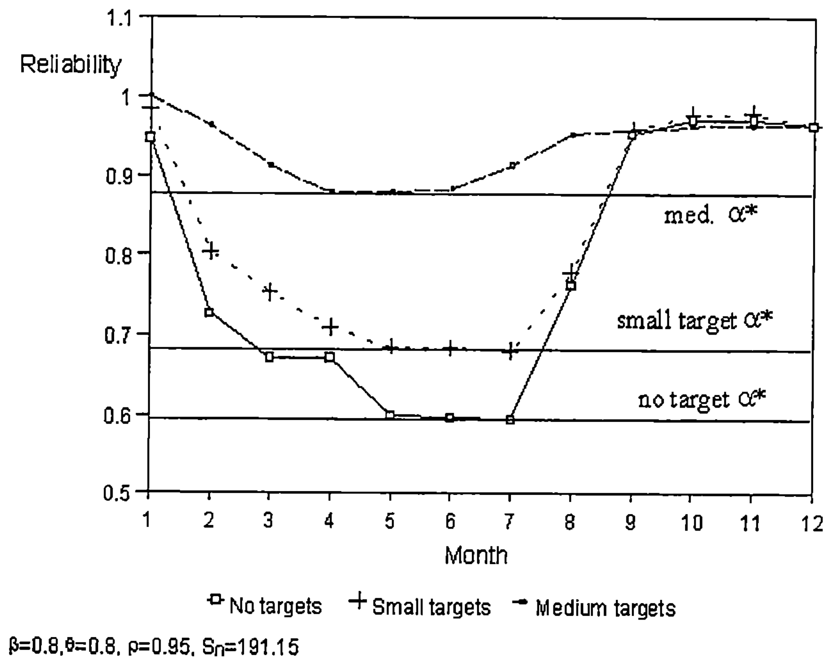

Figure 3.2: Simulated dead storage reliability

of α, the energy generation function EG_t in (2.41) over-estimates the actual energy production. This may be attributed to the simplifying approximation incorporated in the derivation of the EG_t expression. However, at higher values of α, this anomaly is not as prominent, see Figure 3.3, where the initial storage is set at $S_0 = 669$ and the corresponding optimal α at all target levels turns out to be 1.0. Consequently, in the remaining analysis, a minimum level for α, denoted $\alpha_{\min}$, is specified in (NBP) to ensure that no over-estimation of energy generation would occur in the optimization phase. Figures 3.4 and 3.5 show the sensitivity of L^* (firm energy from the optimization model) on the initial storage volume S_0 and the over-year storage reliability ρ, as targets vary from small to medium. These figures correspond to the case $\beta = 0.65, \gamma = 0.85$, and $\theta = 0.65$. Even at these moderate values of reliability, it appears that firm energy output is significantly affected by increased water targets. Figures 3.6 and 3.7 depict the latter sensitivity for the setting: $\beta = 0.95, \gamma = 0.95$, and $\theta = 0.95$. Evidently, with stronger requirement on performance reliability, presence of larger targets dramatically affects the optimized firm energy production level.

The simulated firm-energy production agrees closely with that produced by the optimization model. Figures 3.8 and 3.9 depict the sensitivity of the distribution of simulated L_{fe} on the over-year storage reliability ρ, for the case of medium targets and setting $S_0 = 669$, $\beta = 0.8$, $\gamma = 0.9$, and $\theta = 0.8$. As evident from these plots, as ρ increases, the variance of the distribution of firm energy diminishes while the expected firm energy also decreases. Thus, it appears that a higher level of over-year

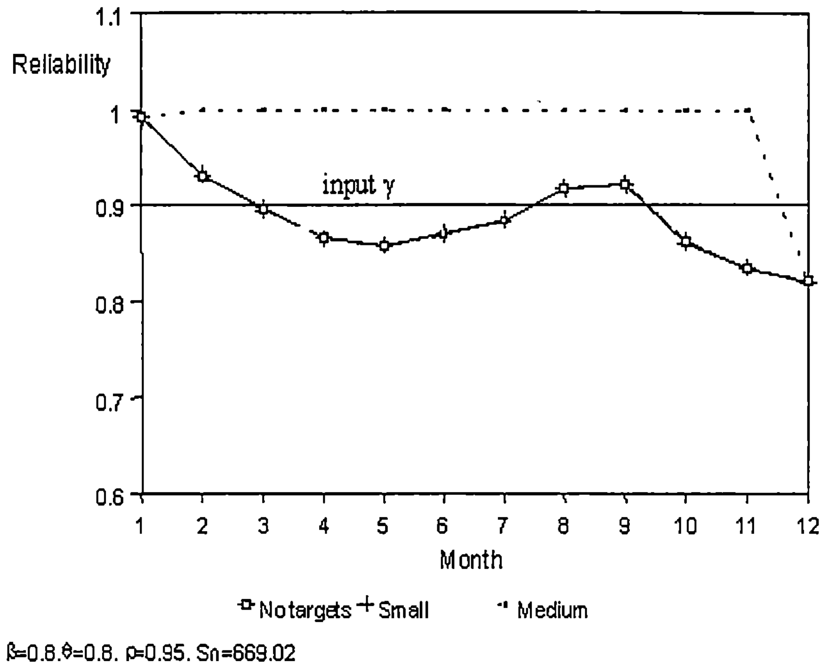

Figure 3.3: Simulated firm energy generation reliability for S_0=609.02

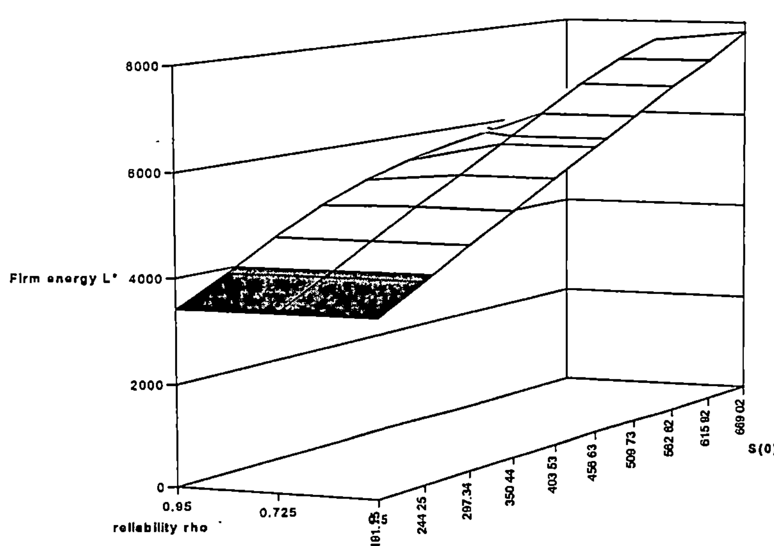

Figure 3.4: Sensitivity of L^* on ρ and S_0 for small targets (low reliabilities)

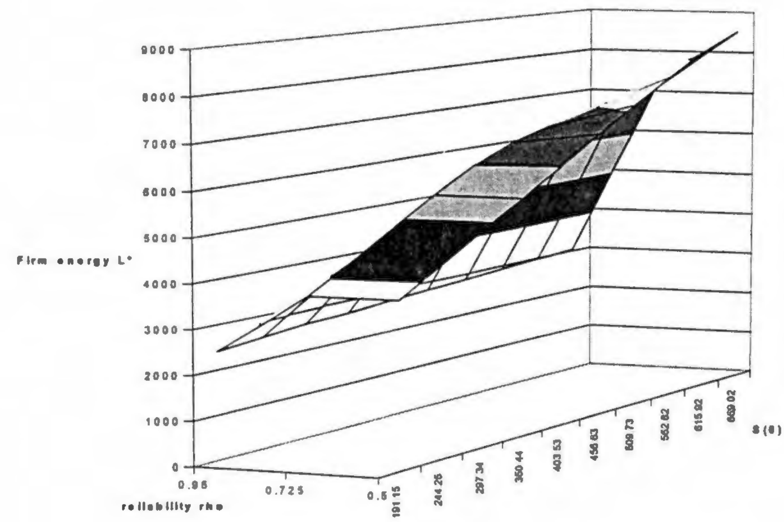

Figure 3.5: Sensitivity of L^* on ρ and S_0 for medium targets (low reliabilities)

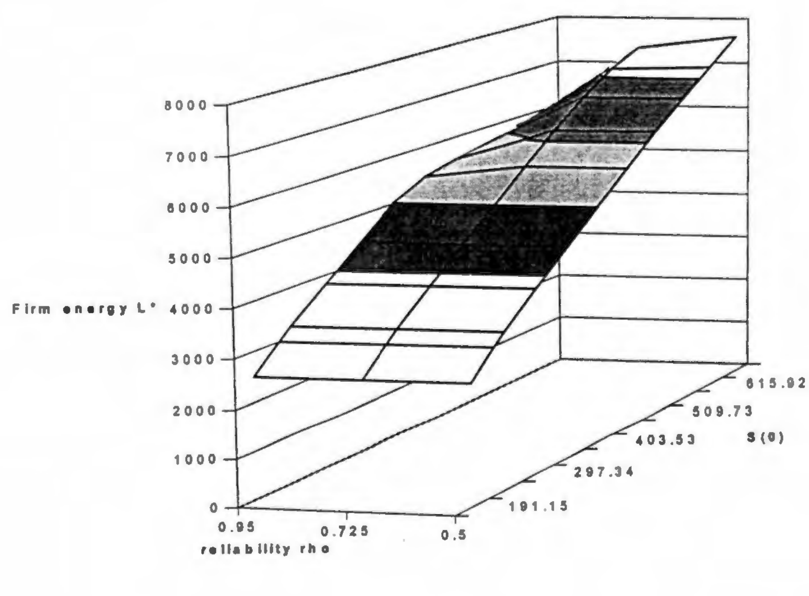

Figure 3.6: Sensitivity of L^* on ρ and S_0 for small targets (high reliabilities)

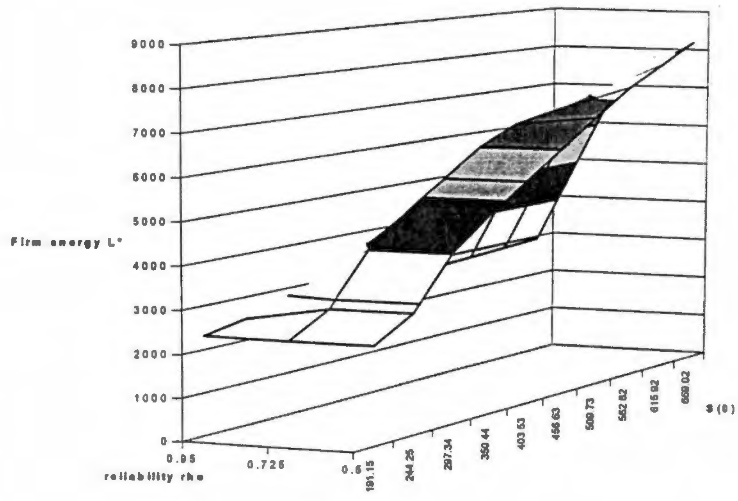

Figure 3.7: Sensitivity of L^* on ρ and S_0 for medium targets (high reliabilities)

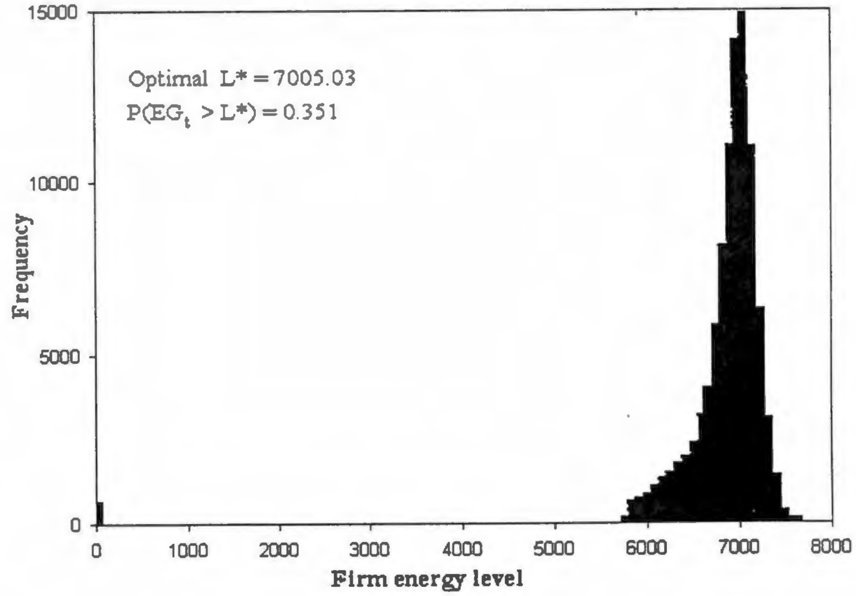

Figure 3.8: Simulated firm energy: ρ=0.725

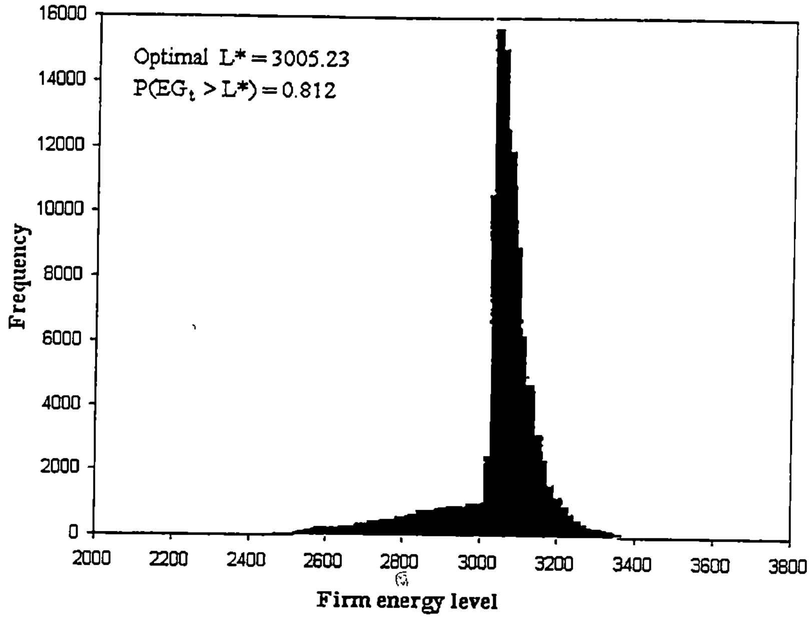

Figure 3.9: Simulated firm energy: ρ=0.95

storage reliability is preferable in the optimization model, which is also consistent with the intended long-term planning nature of the problem. Furthermore, as indicated on these figures, the simulated reliability $\bar{\gamma}_s$ of actually meeting the firm energy level L^* produced by the optimization model also improves with larger ρ.

To study the impact of varying reservoir releases from their optimal values, consider the case of medium targets with $\beta = 0.95$, $\theta = 0.9$, $\gamma = 0.9$, $\rho = 0.95$, and $S_0 = 450$. As releases are decreased from their optimal values, see Figure 3.10, the simulated firm energy reliability, $\bar{\gamma}_s$, drops drastically; in contrast, the over-year storage reliability (ρ) improves only slightly. Increasing releases, on the other hand, improves $\bar{\gamma}_s$ up to a certain level, but ρ is adversely affected. From a practical stand-

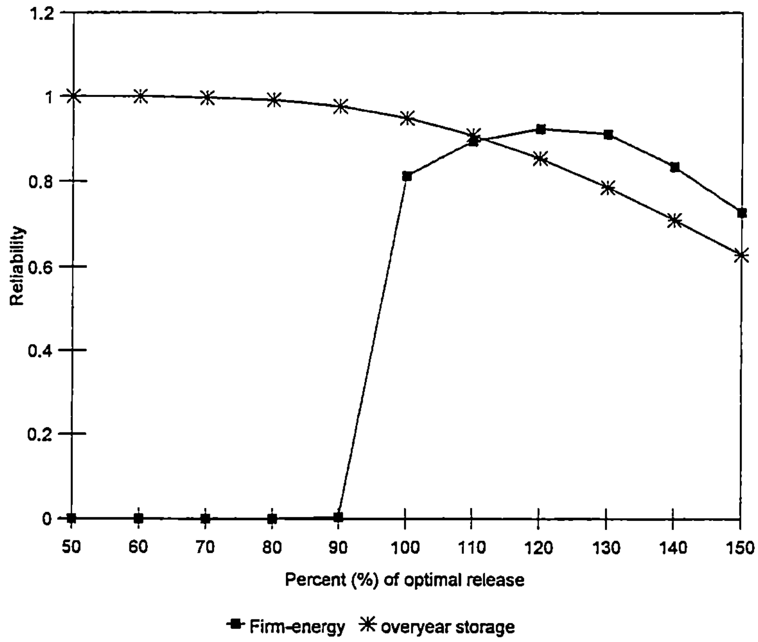

Figure 3.10: Sensitivity on optimal releases

point, it would be preferable to use a release schedule slightly larger than the optimal releases so that a stable $\bar{\gamma}_s$ is ensured at the expense of a slight degradation in ρ.

3.3 Concluding remarks

As demonstrated in this chapter, the methodology appears as an efficient and accurate procedure for determining the optimal reservoir capacity. The proposed model along with the simulation of its solution may be used by decision makers as a way of gaining further insight into the capacity planning problem at hand. Due to the longterm planning nature of the capacity determination problem, the use of the specific deterministic release policy within a monthly model to trade off longterm benefits versus costs may be appropriate. However, chance constrained formulations neither penalize explicitly the constraint violations nor provide recourse action to correct realized constraint violations as a penalty. For this reason, Hogan et al. (1981) warn that the practical usefulness of chance constrained programming as a modeling technique is limited and it should not be regarded as a substitution for stochastic programming with recourse. In the next chapter 4, we shall develop a multi-stage stochstic programming model where we take into account explicitly the magnitude of the deviations rather than only the frequency of violations as in the chance constrained formulation.

Chapter 4

Multi-stage stochastic programming model

4.1 Introduction

Stochastic programming with recourse provides a framework for modeling multiperiod sequential decision problems with uncertain data. Such a model corresponds to the real-life situation where, first, a decision is taken subject to the system constraints, and subsequently, upon observing realizations of random parameters, a second, *recourse*, decision is taken. This may be continued over many future decision epochs where a decision at some period t is taken contingent upon all observations made so far, but without hindsight of the future. This yields a multiperiod stochastic programming model with recourse. Dantzig (1955) pioneered this approach. He suggested a LP model with uncertain data for a two stage problem. Activity levels, or decisions, are determined in the first stage subject to the problem constraints. In the second stage, after some of the random parameters have been observed, a corrective action is made.

Stochastic Programming with recourse is conceptually straightforward but represents some difficulties in applications. In general, it is easier to solve linear simple recourse problems with random variables being discrete, uniform, or normal, see for instance Ziemba (1974). Stochastic linear programs with discrete random variables usually lead to deterministic equivalents of large size. Other types of random distributions either require discrete approximations, as in our case as we will show later, or resolve themselves into nonlinear deterministic equivalents, see Kall (1982), and Wets (1974). For computational approximations for such problems, see Edirisinghe and Ziemba (1992, 1996a, 1996b). Complete analysis by a stochastic programming with recourse model requires that consequences of recourse actions be modeled and computed for all possible realizations of the random variables. Recourse actions are evaluated by an adequate estimation of losses resulting from random variation, which is difficult to find in most cases. It is our intention to present the general multi-stage stochastic formulation in this chapter, and then in later chapters adapt dynamic programming type algorithms to solve this multi-stage stochastic programming model efficiently.

4.1.1 Literature Review

Stochastic LPs for Markov processes have been studied by Manne(1962) and Thomas and Watermeyer(1962). Loucks(1968) developed a stochastic LP for a single reservoir subject to random, serially correlated, net inflows that were described by a first order Markov chain and transition probabilities were estimated using historical inflows.

Houck and Cohon(1978) also assumed a discrete Markov structure for the streamflows. Dorfman(1962) and Dupacova(1980) applied the same idea to the problem of water resources management and planning. In this type of models, the decisions in the consequent periods may be represented by loss functions of not meeting some operational characteristics. Arunkumar and Yeh(1973) used the stochastic DP method to maximize firm power output. Turgeon(1981) developed stochastic DP models for the optimization of weekly operating policies of multireservoir hydroelectric power systems.

4.1.2 Problem Setting

In chapter 2, we developed a chance-constrained programming model that recognizes the randomness in monthly water inflow while allowing for target priority operation. The results obtained from the solution of the model were robust when validated through the simulation model. However, the CCP model allows only for uncertainty being described according to a probability density function, (pdf), and thus sequential realization of random inflow does not directly influence its release policy.

In this chapter, we develop a multi-stage stochastic programming with recourse model for the reservoir problem involving multiple periods representing 12 months of operation. However, we have to take into account the actual inflows and adapt the release policy to the sequential unfolding of the random parameters. The main source of randomness in the reservoir is the monthly water inflow to the reservoir. The downstream demand for irrigation water is prescribed *a priori* and thus it is not

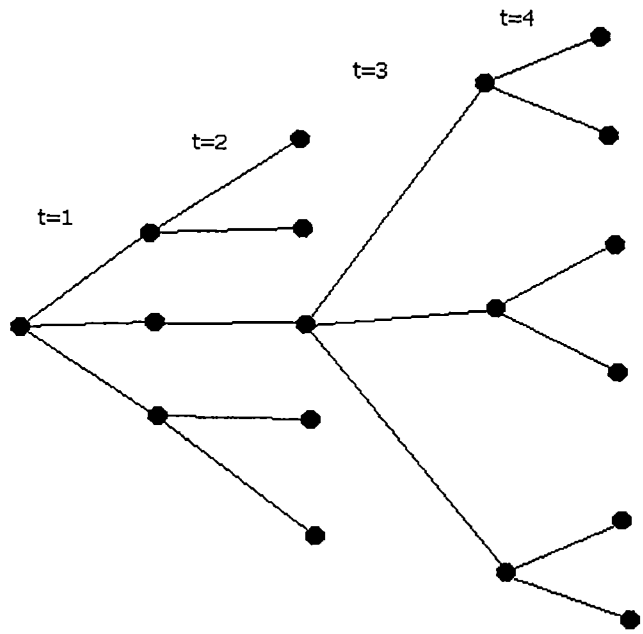

Figure 4.1: Scenario tree

random, see Chapter 2. During any month, the randomness of inflow will be modeled by a sample of discrete outcomes, generated randomly subject to the history of inflow up until that month. In the sequel, we will describe how such samples are generated to develop what is termed a *scenario tree*, see Figure 4.1, of potential future inflow patterns. With fairly dense scenario trees such models tend to become exponentially large as the number of stages and periods increase, and thus the computational cost to solve them also increases exponentially. Therefore, it would be imperative to either use approximation techniques such as Edirisinghe (1996, 1999) and Edirisinghe and You (1996), and/or exploit the structure of the problem and identify or devise decomposition techniques that render efficient solution of the multiperiod reservoir model with scenario trees.

4.2 Multistage Stochastic Model

When developing a stochastic programming model, it is important to understand the distinction between *anticipative* versus *nonanticipative* release policy. The former case corresponds to the situation where release decisions in a given month are made after observing, or forecasting, a particular sequence of inflows for future months. Needless to say that such a release policy suffers from the drawback of being tailor-made to a simple (sequential) forecast of inflow for future months. Among other things, such an anticipative release policy could lead to severe problems in meeting requirements such as energy generation, flood protection or irrigation demand and thus it is unimplementable. On the other hand, in a nonanticipative release policy, the model makes a decision at the present time while taking into account the different outcomes of the random event in a probabilistic sense. In the problem at hand, the reservoir manager must make a release decision before knowing what the inflows will be in the future. Therefore, the model we devise is *nonanticipative*, and it requires a "here-and-now" solution. As in the chance constrained model, the proposed model will maximize the net benefit from the energy generation less the costs associated with the reservoir construction as well as other operational or recourse costs. Operational or recourse costs are imposed on the model so as to penalize the system operation that would tend to violate the specified system constraints. These will be discussed next.

4.2.1 System Constraints

We consider the problem of planning and operation of a multi-purpose water reservoir as described in Chapter 2. In the CCP model, the system constraints are specified as chance constraints, where constraint violations are allowed and controlled via probabilities. In the model presented in this chapter, the degree of violation of a constraint is considered explicitly and controlled. First, the storage level at the beginning of month $(t+1)$, S_t, must be at least SD, the dead storage level, for energy to be generated. Therefore the deviation from SD, denoted by δ^{SD}, is modeled by the following equation:

$$S_t - SD = \delta_t^{SD}. \tag{4.1}$$

Note that δ_t^{SD} is a random variable and $\delta_t^{SD} \geq 0$ indicates the satisfaction of the dead storage constraint in month t. In the CCP model, we considered the chance constraint $\mathcal{P}(\delta_t^{SD} \geq 0) \geq \alpha$ requiring a 100% satisfaction of the constraint. In the present model the quantity $\max\{\delta_t^{SD}, 0\}$ is penalized directly under all possible realizations of inflows.

The reservoir is also used to mitigate flood hazards during high inflow seasons. The deviation from maintaining a specified flood reserve, V_t in month t, is given by the equation:

$$S_t - (K - V_t) = \delta_t^F. \tag{4.2}$$

In order to ensure continued operation of the reservoir in subsequent years provided that the inflow distribution remains unchanged, we require the terminal storage, S_T, be close in value to the initial storage, S_0. The deviation of S_T from the initial storage level S_0 is given by the equation:

$$S_T - S_0 = \delta_T^{S_0}. \tag{4.3}$$

In chapter 2, modeling the target priority directly with a linear constraint was not possible since the releases were considered to be deterministic. A *surrogate* constraint was used to ensure demand for water was satisfied with a certain probability. In the present model, releases are not constrained to be deterministic, i.e. releases conform to a nonanticipative policy. Therefore, the deviation of meeting water targets are given by:

$$R_t - T_t = \delta_t^D. \tag{4.4}$$

Note that flood reserve constraint violations correspond to $\delta_t^F > 0$, and water target constraint violations correspond to $\delta_t^D < 0$. However, violation of the overyear storage requirement indicates $\delta_T^{S_0} \neq 0$. Since the operation of a water reservoir is a continuous process in time, the ending storage and the beginning storage are related by the continuity equation,

$$S_t = S_{t-1} + I_t - R_t, \tag{4.5}$$

assuming no other loss of water is possible. In the next section, we consider the case of modeling the energy generation under the stochastic programming with recourse approach.

4.2.2 Energy generation

In chapter 2, the firm energy level, defined as the minimum guaranteed energy generated throughout a planning horizon, was maximized subject to the system constraints and that the target priority in the release policy is satisfied. In order to maintain the target priority nature and for computational convenience, a Δ_0 release policy was considered in the CCP model. However, in the present model, such a restriction is not needed and the releases are random functions that depend on the history of inflow realizations. In order to maximize the firm energy level, a certain firm energy level is specified to the model and the deviation of $\min(EG_t, t = 1, \ldots, T)$, is accounted for and minimized as will be explained in the next subsection.

Energy generation constraint

The energy generated is a function of the release and the average water head on the turbines. Given a transition of the system form S_{t-1} to S_t, as in (2.40), the energy generated at period t can be represented by the following

$$EG_t = \omega R_t[\frac{e}{2}(S_t + S_{t-1}) + f]. \tag{4.6}$$

Note that the release R_t will not contribute towards energy generation if both S_t and S_{t-1} are below the dead storage level. In order to compute the exact value of the energy generation, we define the variables x_t, y_t, h_t, and z_t as follows

$$x_t = \begin{cases} 1 & \text{if } S_t \geq \text{SD}, \\ 0 & \text{otherwise} \end{cases} \tag{4.7}$$

$$y_t = \begin{cases} 0 & \text{if } S_t \geq \text{SD}, \\ -\delta_t^{SD} & \text{otherwise} \end{cases} \tag{4.8}$$

$$h_t = \begin{cases} \delta_t^{SD} & \text{if } S_t \geq \text{SD}, \\ 0 & \text{otherwise} \end{cases} \tag{4.9}$$

$$z_t = \begin{cases} 1 & \text{if } x_{t-1}{=}1 \text{ OR } x_t{=}1, \\ 0 & \text{otherwise} \end{cases} \tag{4.10}$$

Note that $(S_t + S_{t-1})$ may be restated as:

$$S_t + S_{t-1} = 2SD + \delta_t^{SD} + \delta_{t-1}^{SD}. \tag{4.11}$$

For the case when both the beginning and ending storages in month t are above the dead storage,

$$S_t + S_{t-1} = 2SD + h_t + h_{t-1} \tag{4.12}$$

holds. In general, however, $\frac{1}{2}(2SD{+}h_t{+}h_{t-1})$ represent the average "effective storage" available for hydropower generation. We also want to determine the *effective release*, defined as the released amount of water that contributes towards energy generation. This can be done by subtracting the amount of water below SD from the release R_t as follows

$$R_{eff} = R_t - y_t - y_{t-1} \tag{4.13}$$

The energy generation function can therefore be written as follows

$$EG_t = \omega z_t (R_t - y_t - y_{t-1})[\frac{e}{2}(2SD + h_t + h_{t-1}) + f] \tag{4.14}$$

Observe now that only the amount of water released that contributes towards energy generation is taken into account. Likewise, the average water head on the turbines is not over-estimated by taking the simple average of S_{t-1} and S_t as in the CCP model. Now to ensure a given firm energy level, say η, any deviation from η, which we shall represent by δ^{EG}, is penalized. Therefore, the energy generation constarint violation can be written as follows:

$$\eta - EG_t = \delta_t^{EG}. \tag{4.15}$$

4.2.3 Multistage stochastic model

Rewriting the constraints

As we mentioned in the introduction, the energy authority has to make *here-and-now* decisions regarding the releases. This implies that the release in the first period, R_1, is independent of the inflow. The same holds true for the target violation variable δ_1^D since the target satisfaction depends only on the release. However, the storage at the end of period 1 depends on the inflows due to the continuity equation (4.5). This dependence implies that δ_1^{SD}, δ_1^F, and δ_1^{EG} also depend on the realization of the random event, i.e. which inflow occured. Note, however, that the release in period 2 depends on the ending storage of period 1, S_1, and does hence depend on the inflow in period 1, I_1. To reflect these dependencies on which of the inflows was manifested, let $\mathcal{H}_{t-1} := I_1, \ldots, I_{t-1}$ be the history of inflows up to a period t. The constraints (4.5),(4.4),(4.2), (4.1), and(4.3), are written as follows

$$S_{t,\mathcal{H}_{t-1}} = S_{t-1,\mathcal{H}_{t-1}} + I_{t,\mathcal{H}_{t-1}} - R_{t,\mathcal{H}_{t-1}} \quad ; t = 1, \ldots, T \tag{4.16}$$

$$R_{t,\mathcal{H}_{t-1}} - T_{t,\mathcal{H}_{t-1}} = \delta^D_{t,\mathcal{H}_{t-1}}; \quad t = 1, \ldots, T \tag{4.17}$$

$$S_{t,\mathcal{H}_{t-1}} - (K - V_t) = \delta^F_{t,\mathcal{H}_{t-1}}; \quad t = 1, \ldots, T \tag{4.18}$$

$$S_{t,\mathcal{H}_{t-1}} - SD = \delta^{SD}_{t,\mathcal{H}_{t-1}}; \quad t = 1, \ldots, T \tag{4.19}$$

$$S_{T,\mathcal{H}_{T-1}} - S_0 = \delta^{S_0}_{T,\mathcal{H}_{T-1}}; \quad t = T \tag{4.20}$$

The energy generation constraint is more cumbersome since we want to consider a firm-energy level, which is the same over a specified time horizon. Therefore, the penalty δ^{EG}_t is taken as the deviation of the firm energy level from the specified energy level across a scenario. This delicate dependence can be easily represented by

$$\eta - EG_{t,\mathcal{H}_{t-1}} = \delta^{EG}_{T,\mathcal{H}_{T-1}}; \quad t = 1, \ldots, T \tag{4.21}$$

The objective funtion

In the objective function, we want to minimize the penalty from operating the reservoir. The penalty being the cost of deviating from the reservoir operating characteristics. So it can be represented by the following

$$\min \quad F(\delta^{EG}_{t,\mathcal{H}_{t-1}}, \delta^{S_0}_{T,\mathcal{H}_{t-1}}, \delta^D_{t,\mathcal{H}_{t-1}}, \delta^F_{t,\mathcal{H}_{t-1}}, \delta^{SD}_{t,\mathcal{H}_{t-1}}); \quad t = 1, \ldots, T. \tag{4.22}$$

Note, however, that δ^{SD}_t need not be considered explicitly in the objective function since it's impact is implicitly penalized in the energy generation function EG_t. The cost function in (4.22) is nothing but the sum of the expected cost of each variable in each scenario. The complete objective function can be written therefore as

$$\begin{aligned} \min \quad & \lambda[\sum_{t=1}^{T} \mathcal{P}^{\mathcal{H}_{t-1}} \mathcal{F}_{\delta^{EG}}(\delta^{EG}_{t,\mathcal{H}_{t-1}}) + \sum_{\varphi} \mathcal{P}^{\mathcal{H}_{T-1}} \mathcal{F}_{\delta^{S_0}}(\delta^{S_0}_{\mathcal{H}_{T-1}}) \\ & + \sum_{t=1}^{T} (\mathcal{P}^{\mathcal{H}_{t-1}} \mathcal{F}_{\delta^{D}}(\delta^{D}_{t,\mathcal{H}_{t-1}})) + + \sum_{t=1}^{T} (\mathcal{P}^{\mathcal{H}_{t-1}} \mathcal{F}_{\delta^{F}}(\delta^{F}_{t,\mathcal{H}_{t-1}}))] \end{aligned} \tag{4.23}$$

where φ is the set of all possible scenarios.

Complete Multistage Stochastic Model

We have defined the variables x_t, y_t, h_t, and z_t earlier to define the energy generation constraint. However, we have introduced them as indicator function rather than constarints that can be included in the complete formulation of the model. To convert those to constraints, we proceed as follows. Let M denote a very large number.

$$M(x_{t,\mathcal{H}_{t-1}} - 1) \leq \delta^{SD}_{t,\mathcal{H}_{t-1}}; \quad t = 1, \dots, T \tag{4.24}$$

$$-Mx_{t,\mathcal{H}_{t-1}} - y_{t,\mathcal{H}_{t-1}} \leq \delta^{D}_{t,\mathcal{H}_{t-1}}; \quad t = 1, \dots, T \tag{4.25}$$

$$y_{t,\mathcal{H}_{t-1}} \geq 0; \quad t = 1, \dots, T \tag{4.26}$$

$$z_{t,\mathcal{H}_{t-1}} \leq x_{t,\mathcal{H}_{t-1}} + x_{t-1,\mathcal{H}_{t-1}}; \quad t = 1, \dots, T \tag{4.27}$$

$$0 \leq z_{t,\mathcal{H}_{t-1}} \leq 1; \quad t = 1, \dots, T \tag{4.28}$$

$$h_{t,\mathcal{H}_{t-1}} = x_{t,\mathcal{H}_{t-1}} \delta^{SD}_{t,\mathcal{H}_{t-1}}; \quad t = 1, \dots, T \tag{4.29}$$

With these definitions (constraints) in place, the multistage stochastic model is completely defined. The formulation is peresented in its general form as the following MultiStage Stochastic Program (MSSP).

$$
\begin{aligned}
(MSSP)_T: \ \min \quad & \lambda[\textstyle\sum_{t=1}^{T} \mathcal{P}^{\mathcal{H}_{t-1}} \mathcal{F}_{\delta^{EG}}(\delta^{EG}_{t,\mathcal{H}_{t-1}}) + \sum_{\varphi} \mathcal{P}^{\mathcal{H}_{T-1}} \mathcal{F}_{\delta^{S_0}}(\delta^{S_0}_{\mathcal{H}_{T-1}}) \\
& + \textstyle\sum_{t=1}^{T} (\mathcal{P}^{\mathcal{H}_{t-1}} \mathcal{F}_{\delta^{D}}(\delta^{D}_{t,\mathcal{H}_{t-1}})) + + \sum_{t=1}^{T} (\mathcal{P}^{\mathcal{H}_{t-1}} \mathcal{F}_{\delta^{F}}(\delta^{F}_{t,\mathcal{H}_{t-1}}))] \\
& \text{s.t.} \\
& S_{t,\mathcal{H}_{t-1}} = S_{t-1,\mathcal{H}_{t-1}} + I_{t,\mathcal{H}_{t-1}} - R_{t,\mathcal{H}_{t-1}} \quad ; t = 1, \dots, T \\
& R_{t,\mathcal{H}_{t-1}} - T_{t,\mathcal{H}_{t-1}} = \delta^{D}_{t,\mathcal{H}_{t-1}}; \quad t = 1, \dots, T \\
& S_{t,\mathcal{H}_{t-1}} - (K - V_t) = \delta^{F}_{t,\mathcal{H}_{t-1}}; \quad t = 1, \dots, T \\
& S_{t,\mathcal{H}_{t-1}} - SD = \delta^{SD}_{t,\mathcal{H}_{t-1}}; \quad t = 1, \dots, T \\
& S_{T,\mathcal{H}_{T-1}} - S_0 = \delta^{S_0}_{T,\mathcal{H}_{T-1}}; \quad t = T \\
& M(x_{t,\mathcal{H}_{t-1}} - 1) \leq \delta^{SD}_{t,\mathcal{H}_{t-1}}; \quad t = 1, \dots, T \\
& -Mx_{t,\mathcal{H}_{t-1}} - y_{t,\mathcal{H}_{t-1}} \leq \delta^{D}_{t,\mathcal{H}_{t-1}}; \quad t = 1, \dots, T \\
& y_{t,\mathcal{H}_{t-1}} \geq 0; \quad t = 1, \dots, T \\
& z_{t,\mathcal{H}_{t-1}} \leq x_{t,\mathcal{H}_{t-1}} + x_{t-1,\mathcal{H}_{t-1}}; \quad t = 1, \dots, T \\
& 0 \leq z_{t,\mathcal{H}_{t-1}} \leq 1; \quad t = 1, \dots, T \\
& h_{t,\mathcal{H}_{t-1}} = x_{t,\mathcal{H}_{t-1}} \delta^{SD}_{t,\mathcal{H}_{t-1}}; \quad t = 1, \dots, T \\
& \eta - EG_{t,\mathcal{H}_{t-1}} = \delta^{EG}_{T,\mathcal{H}_{T-1}}; \quad t = 1, \dots, T
\end{aligned}
\tag{4.30}
$$

This is a general formulation in all aspects. The time horizon is left as a parameter, T, and so are the penalty cost functions. Now that the model is complete, a solution procedure needs to be devised. In the next section, we propose stochastic dynamic programming as a solution technique to solve this model efficiently.

4.3 Solution procedure: Stochastic Dynamic Programming

4.3.1 Use of Stochastic Dynamic Programming

Dynamic programming (DP) is a solution procedure credited largely to Bellman(1957). The popularity and success of this technique can be attributed to the fact that most multistage optimization problems can be translated into a sequence of nested evaluation problems. In addition, it has the advantage of effectively solving highly complex problems with a large number of variables by decomposing it into a series of subproblems which are solved recursively, see Yeh (1985). To decompose a general problem into multiple stages with decisions required at each stage, the value of every stage should satisfy the separability condition and the monotonicity condition, see Nemhauser (1966). A linear objective function, for instance, is separable if the different decision variables, say x_j, appear in separate terms $c_j x_j$. The problem at hand has these separability and monotonicity properties. However, Yeh (1985) states that the usefulness of the technique is limited due to the computational complexity arising from *the curse of dimensionality.* The latter refers to the fact that a DP size increases exponentially with the number of periods/stages and outcomes considered in the model. He further states that stochastic DP is extremely well suited by its nature to handle stochastic problems for long-range operation, which is part of what our model accomplishes. By exploiting the special nature of the model, we intend to solve the model efficiently using stochastic DP.

4.3.2 Stochastic Dynamic Programming Model

By inspecting equations (4.1),(4.2),(4.3),(4.4), and (4.5), it becomes apparent that all these constraints depend on two variables, S_{t-1} and R_t. In other words, given the beginning storage and the release at period t, one can compute the deviations and the value of the objective function explicitly. This is a very important remark, it basically reduces the model (4.30) to an evaluation problem. One can, therefore, consider a grid of discretized S_{t-1} and R_t and compute all other values, including the energy generated, and find which combination minimizes the cost or maximizes the benfit depending on the objective.

To apply this idea to the model, we need first to introduce the notion of a "scenario tree". A tree is a connected graph that contains no cycles. A graph consists of a set of nodes that are connected together with a set of arcs. A graph is connected if every two nodes are connected in the sense that there is a path, or a walk without repetition of nodes, between the two nodes. A rooted tree is a tree with a specially designated node called the *root*. We often view arcs in a rooted tree as defining predecssor-successor relationships. The predecssor of a node is the next node in the unique path from that node to the root. Each node in a tree has one unique predecessor, but not necessarily one successor. A scenario tree is a tree where the nodes represent decision points, and the arcs represent possible realizations of a random parameter. In our model, the nodes coorespond to the beginning of each period following a specific scenario, and the arcs correspond to the possible realizations of the montly inflows.

Consider period $T-1$ and at each node in that period compute the reservoir characteristics for each combination S^i_{T-1} and R^j_T. For each discretized release, we compute the expected cost depending on the set of inflows I_T. For each beginning storage S^i_{T-1}, we find the release that minimizes the penalty cost. We store the information from each node as an array indexed on the discretized S^i_{T-1}. Once all nodes in period $T-1$ are done, we proceed backwards to period $T-2$. We repeat the same procedure with one exception, which is that for each combination S^i_{T-2} and R^j_{T-1}, the cost function for each inflow I^k_{T-1} needs to include the cost from period $T-1$. This is accomplished by computing S_{T-1} which is the ending storage at period $T-2$ and the beginning storage at period $T-1$, and finding the cost associated with that volume from the cost arrays that have been computed in the previous step for period $T-1$. The exact ending storage may not be in the cost array in which case we can use any of the interpolation techniques. We proceed iteratively backwards until we reach the starting point with given S_0 and search on R_1. This basic idea will be utilized in the coming chapters to devise algorithms depending on the different assumptions we make on the monthly water inflows.

4.4 Concluding Remarks

In this chapter, we have developed a mulit-stage stochastic program for the capacity planning of a multi-purpose water reservoir. This model is presented in its most general form, where all the model parameters are left as general. A solution approach using stochastic dynamic programming will be implemented. However, due to the size of the MSSP problem, it would not be possible to solve this problem in a reasonable amount of time. To solve this problem efficiently, we have to recourse to either restricting assumptions that will lead to approximate rather than optimal solutions, or aggregation which is in itself a decompoisition method that will yield a bound on the optimal solution. In other words, one cannot hope to solve a problem of this size to optimality using available methodologies and available computational tools.

In the ensuing chapters, we present three different stochastic dynamic programming models that would enable us to provide computational procedures that will also be validated using simulation. In chapter 5, we present a stochastic dynamic programming model for the case of independent monthly inflows, and the case of restricted dependent cumulative monthly inflows. We will show that under these two assumptions we devise a very efficient solution algorithm. Then, in chapter 6, we peresent an aggregated dynamic program where we conserve the dependence among monthly water inflows, but reseort to aggregation to reduce the problem size. We develop an efficient dynamic programming algorithm to solve the problem.

Chapter 5

Dynamic Programming Models of Independent and Dependent Inflows

In Chapter 4, we developed a multi-stage stochastic program in a general setting. However, the solution of such a model in its full generality would be computationally very difficult. This is owing to the fact that the deterministic equivalent of the stochastic program is a massively large scale nonlinear mixed integer programming formulation. For instance, with 12 monthly periods and 10 outcomes of random inflows per period, the model has 10^{12} diistinct scenarios and thus the size of the problem is of $O(10^{12})$. To overcome this seemingly impossible task, we propose a dynamic programming based solution technique. However, stochastic dynamic programming as a general solution technique suffers from the computational drawback known as the *curse of dimensionality.* The latter term refers to the fact that as the number of stages in a dynamic program increases, the computational complexity increases exponentially with the fineness of state space discretization. In a stochastic dynamic program, the complexity is even worse due to increases in the number of random outcomes per stage. In this chapter, we present two special cases of the model in Chapter 4 that can be solved rather efficiently, namely the case where monthly inflows are independent of each other, and the case when monthly inflows depend only on the cumulative inflow thus far. In Section 5.1, we present the general recursive formulae for the dynamic programming approach. In section 5.2, a "Dynamic Programming with Independant Inflows" model is presented. Then, in section 5.3, we present a "dynamic programming with resricted dependence" model.

5.1 Dynamic Programming General Recursion

In chapter 4, we proposed a multi stage stocahastic program with recourse, and we proposed dynamic programming as a solution approach. Let $\mathcal{H}_{t-1} = \{I_1, \ldots, I_{t-1}\}$, be defined as the history of the inflows realized up to a period t. Define the state space as being the pair $(\mathcal{H}_{t-1}, S_{t-1})$. The state of the system is completely defined by this state space definition. The history of inflows $\mathcal{H}_{t-1}$ provides the information on what scenario of inflows has realized up to the period t. The model is to select the release R_t that would minimize the expected penalty cost, less energy benefits, relative to the inflows. The value function, denoted by $\phi_t(.)$, at the node $\mathcal{H}_{t-1}$ of uncertainty resolution is then defined by the following, for a given value of S_{t-1}:

$$\phi_t(\mathcal{H}_{t-1}, S_{t-1}) = \min_{R_t} \quad E_{I_t|\mathcal{H}_{t-1}}\left[\mathcal{F}(\mathcal{H}_{t-1}, S_{t-1}, I_t, R_t) + \phi_{t+}(\mathcal{H}_t, S_t)\right]$$

$$\begin{aligned}
&\text{s.}t. \\
&S_t + R_t = S_{t-1} + I_t \\
&S_t - SD = \delta_t^{SD} \\
&R_t - T_t = \delta_t^{D} \\
&S_t - (k - V) = \delta_t^{F} \\
&\nu - EG_t = \delta_t^{EG} \\
&S_T - S_0 = \delta^{S_0}, \quad \text{if} t = T \\
&R_t \geq 0.
\end{aligned} \tag{5.1}$$

Where,

$\mathcal{F}(\mathcal{H}_{t-1}, S_{t-1}, I_t, R_t)$ is the net cost function due to the penalties δ_t^{SD}, δ_t^{D}, δ_t^{F}, δ_t^{EG}, and $\delta_T^{S_0}$ and benefit due to the firm energy ν. Also, $E_{I_t|\mathcal{H}_{t-1}}[.]$ denotes the conditional expectation with respect to the random inflow I_t given the resolution of uncertainty up to period t, i.e., $\mathcal{H}_{t-1}$.

5.2 Dynamic Programming: Independent Inflows Models

Solving the dynamic program (5.1) requires the solution of the nonlinear program at each node $\mathcal{H}_{t-1}$, for a specified S_{t-1}. This is an orenous task as the number of such nodes increases exponentially with the addition of periods and/or outcomes. Furthermore, the latter computation need to be performed for every possible S_{t-1}, s determined by a suitable grid of values for S_{t-1}. In this section we present an efficient version of the above DP under the assumption that the monthly inflows are independent.

Assumption 5.2.1 *The monthly inflow at period t, namely the random variable I_t, is assumed independent of the history of inflows, $\mathcal{H}_{t-1}$.*

5.2.1 Modeling Inflows in the DP model

Under assumption (5.2.1), note that random variables Q_{t-1} and I_t are stochastically independent. Therefore, the mean and variance of a distribution of the individual monthly inflows are determined as follows.

$$\begin{aligned} &Q_t = I_t + Q_{t-1}, \quad t = 1, ..., 12 \\ &E[Q_t] = E[I_t] + E[Q_{t-1}], \quad t = 1, ..., 12 \\ &Var[Q_t] = Var[I_t] + Var[Q_{t-1}], \quad t = 1, ..., 12. \end{aligned} \tag{5.2}$$

This yields the following mean and variance for I_t:

$$\begin{aligned} &E[I_t] = E[Q_t] - E[Q_{t-1}], \quad t = 1, ..., 12 \\ &Var[I_t] = Var[Q_t] - Var[Q_{t-1}], \quad t = 1, ..., 12. \end{aligned} \tag{5.3}$$

Note here that we don't have a description of the distribution of the monthly inflows, although lognormal distributions were fitted for Q_t in Chapter 3. There are two approaches by which one is able to generate outcomes for I_t. The first is to assume a distribution for the inflow with the mean and variance defined by (5.3). For example one can assume a uniform distribution with 3σ limits, i.e.

$$I_t := U(\max(0, E[I_t] - 3\sqrt{Var[I_t]}), E[I_t] + 3\sqrt{Var[I_t]},$$

or a lognormal distribution, i.e.

$$I_t := \mathcal{LN}(E[I_t], Var[I_t]).$$

A more convenient approach where we do not need an explicit functional form for the individual monthly inflows is to simulate the distribution of I_t from the distributions of Q_t and Q_{t-1} which we have already determined to be lognormally distributed. A scenario branch at a particular node in period t in this case is generated by considering the lower and upper limits,

$$U_t = E[I_t] + s(Var[I_t])^{1/2}$$

$$L_t = Max(0, E[I_t] - s(Var[I_t])^{1/2} \tag{5.4}$$

where s represents the number of standard deviations around the mean. With these limits on hand and the assumption that this is a uniform distribution, we can simply divide this interval $[L_t, U_t]$ into n subintervals Δ_t^n, where n is the number of outcomes in each period. The midpoint of each subinterval Δ_t^n would be the inflow at that particular outcome. If we assume a uniform distribution, then each outcome is equally likely with probability $1/n$. If the lognormal distribution is assumed, then we generate a large number of random variates and place them into the appropriate subinterval. A frequency count divided by the total number of random variates generated. The final approach is to simulate the individual monthly inflows from the cumulative monthly inflows distributions which have been determined to be lognormal. To generate a scenario branch at a particular node of a particular period, we generate a random variate q_{t-1} from Q_{t-1} and a random variate q_t from

Q_t, and define variate i_t (for I_t) as $q_t - q_{t-1}$ provided that $q_t > q_{t-1}$. The variate i_t thus generated is placed in the appropriate subinterval. A frequency count is made, and the probability of each outcome is computed as the frequency of variates in each subinterval divided by the total number of variates generated.

Model Description

Under Assumption (5.2.1), the outcomes at period t are determined independently of the historical scenario being followed until period t. Thus, the outcomes and their corresponding probabilities of occurence in each node in period t are identical. Therefore, the DP model in this can be presented as follows:

$$
\begin{aligned}
\phi_t(S_{t-1}) = \quad & \min_{R_t} \quad E_{I_t}\left[\mathcal{F}(S_{t-1}, I_t, R_t) + \phi_{t+1}(S_t)\right] \\
& \text{s.}t. \\
& S_t + R_t = S_{t-1} + I_t \\
& S_t - SD = \delta_t^{SD} \\
& R_t - T_t = \delta_t^{D} \\
& S_t - (k - V) = \delta_t^{F} \\
& \nu - EG_t = \delta_t^{EG} \\
& S_T - S_0 = \delta^{S_0}, \quad \text{for} \quad t = T \\
& R_t \geq 0.
\end{aligned}
\tag{5.5}
$$

Note that the value function does not depend on the history $\mathcal{H}_{t-1}$ of the inflows, and hence, for a given beginning storage S_{t-1}, the value function is identical for all nodes $\mathcal{H}_{t-1}$ in period t. This is a very important result since it implies that we only solve the model in (5.5) once in each period. For instance, for 12 periods with with 10 outcomes, instead of solving $\sum_{i=0}^{12} 10^i$ problems corresponding to the number of nodes, we only have to solve 12 problems of the format (5.5) for each value of S_{t-1}. We also propose to solve the minimization in (5.5) by a grid search on R_t, thereby avoiding a possibly difficult nonlinear programming procedure.

5.2.2 Pseudo Code

Given a set of discretized beginning storages S_{t-1}^n, for $n = 1, \ldots, N$, and a set of discretized releases R_t^j, for $j = 1, \ldots, J$, we compute the value function for one node in period 12 and copy it to all other nodes in the period.

Step 0:Initialization

1. Set reservoir size and maximum firm energy level

2. Obtain inflow data (mean, standard deviation, and number of outcomes per period; we also need probability of each outcome if a distribution other than the uniform distribution is used)

3. Set t= T-1, where T= number of periods to be considered in the model

Step 1:DO WHILE n $\leq$ N

1. Set $S_{t-1} = S_{t-1}^n$, Let $\phi^*(S_{t-1}) = M$, where M denotes a large positive number.

Step 2: DO WHILE j $\leq$ J

1. evaluate $\phi_t(S_{t-1}, R_t^j)$ as described in (5.5)

2. for $\phi^*(S_{t-1}) \quad < \quad \phi_t(S_{t-1}^i, R_t^j)$

   let $\phi^*(S_{t-1}) \quad = \quad \phi_t(S_{t-1}^i, R_t^j)$ and $R_t^*(S_{t-1}) = R_t^j$.

END DO

END DO

Step 3:

1. Set t=t-1.

2. if $t < 0$, STOP,

   else go to Step 1

5.2.3 Results Analysis and Simulation

Model Solution

As mentioned above, based on the assumptions we made, we only need to the value function at one node per period. This is a tremendous advantage since it reduces the problem to a simple evaluation problem that is solved in seconds even when we consider cases with large number of outcomes per period. This actually turns out to be very useful since we can make up for the loss of the dependence structure by considering more outcomes per period. It would seem fairly reasonable to assume that the larger the number of outcomes, the closer the model gets to resembling scenarios from a pure dependent inflows model. This remark will actually be made more clear in chapter 7. However, the accuracy of the solution is dependent largely on the fineness of the grids chosen for S_{t-1} and R_t.

The solution from this model is somewhat different from the CCGP model discussed in Chapter 2. The optimal solution from the CCGP model consisted of an optimal reservoir capacity, optimal dead storage reliability and a Δ_0-release policy. In the current model, however, the reservoir size is treated as input. The optimal solution is a release policy that is a nonanticipative policy that depends on the beginning storage since $R_t^* = \mathcal{F}(S_{t-1})$. An example of a release policy output by the model for 12 monthly periods is shown graphically in Figures C1-C11 in Apeendix C. Two important remarks need to made here. First, note that as we move from the early periods to the later periods, the function $R_t^* = \mathcal{F}(S_{t-1})$ becomes smoother. This can be explained by the fact that as we progress into the future months, more monhtly inflows are realized, and thus, we have less uncertainty to account for. Secondly, note that to a certain beginning storage level, the release policy fluctuates considerably. For larger values of the beginning storage, it becomes smooth in a direct relationship with the beginning storage. This is a very interesting behavior in its own right. It may be explained by first observing that the release amount is generally smaller than the beginning storage level. This indicates that up to a certain level of storage, the model chooses optimal releases with the knowledge of the expected value of the inflows in the coming periods, heavily affecting the choice of the optimal solution. When the beginning storage level becomes larger than a certain level, the information from the knowledge of the expected value of the inflows in the coming period has a lesser weight in the choice of the optimal release. Although, the releases are smaller than the beginning storage levels, in some cases rather drastically, this can be explained by the fact that this is a long-term planning problem. The model could for instance make the maximum release and generate a larger amount of energy, but due to the conflicting objectives, this might hamper the operation in subsequent periods.

Simulation Analysis

The simulation study of the results developed by this DP model is also different from that of the CCGP model. The input parameters input to the model are: reservoir size K, the firm energy level ν, the mean and standard deviation of the cumulative inflows, the targets, and the release policy $R_t^*(S_{t-1})$. The output is the simulated reliabilities, α_s, θ_s, β_s, γ_s, ρ_s. The simulation model also outputs the average energy generated each period, the average potential energy generated during each period, an average simulated firm energy level, and the firm energy reliability measure. We do 100,000 simulation runs, by changing the inflow sequence for 12 months. At each iteration, a release is selected from the optimal release policy depending on the period and the beginning storage level of that period, S_{t-1}. As evident from Figure 5.1, the simulated reliabilities α_s, θ_s, β_s, remain consistently above 0.9. This indicates that the model does take into account these secondary objectives and they are all satisfied with very high probabilities. Figures 5.2 and 5.3 depict the simulated $\gamma_{L_{fe}}$ for both the actual energy generation and the potential energy generation. The potential energy generation is defined as the energy generated if all available water above the dead storage volume is released. Both remain consistently high indicating

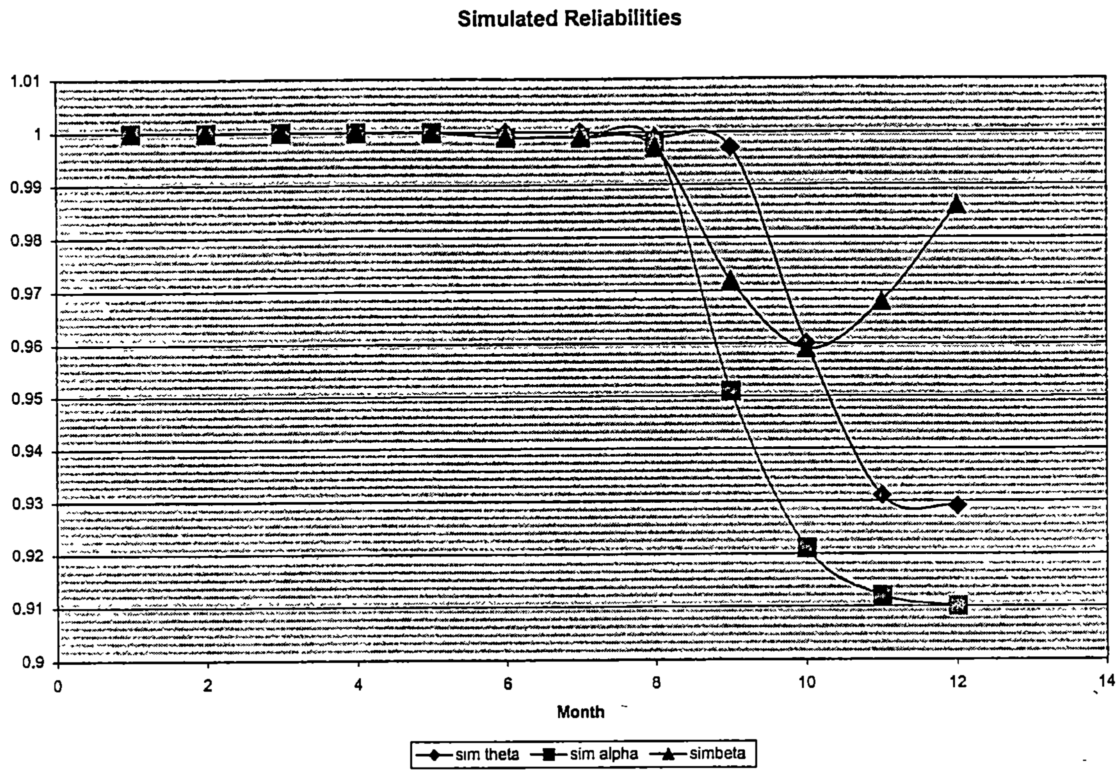

Figure 5.1: Simulated reliabilities α, β, and θ

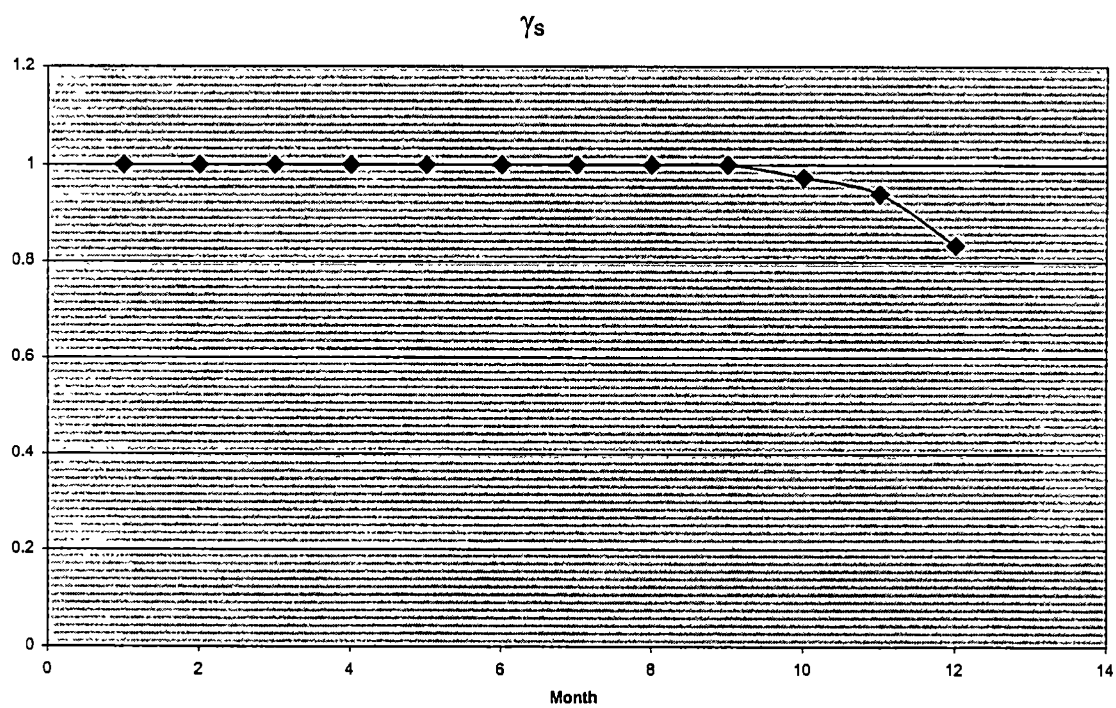

Figure 5.2: Simulated actual firm energy generation reliability γ_s

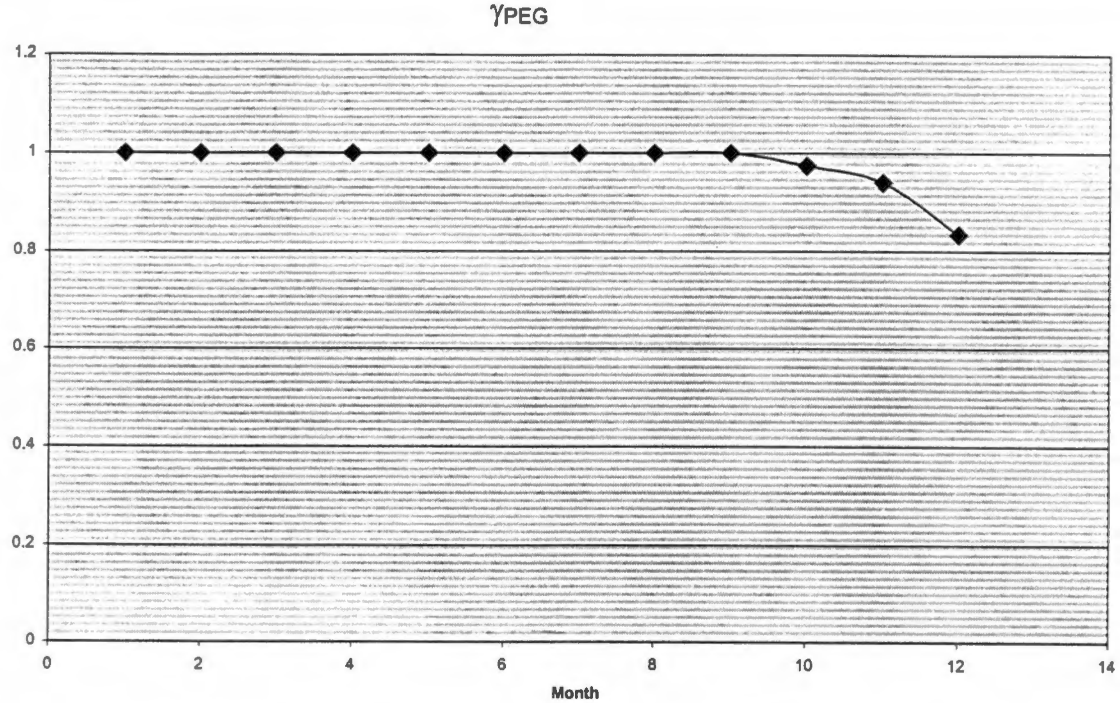

Figure 5.3: Simulated potential firm energy generation reliability

that the model is indeed maximizing returns for the energy authority through high firm energy levels. More importantly, the actual energy generation and the potential energy generation reliabilities are almost equal, indicating that the release policy chosen by the model maximizes firm energy benefits without bringing the level of water below the dead storage level. Figures 5.4 and 5.5 are of particular importance as they reveal the distribution of S_{12} and L_{fe}. Recall that one of the reasons we decided to use a reliability level on the firm energy constraint was that determination of its distribution was a complex matter as it requires the knowledge of the covariance terms among $EG_t, t = 1, ..., 12$.

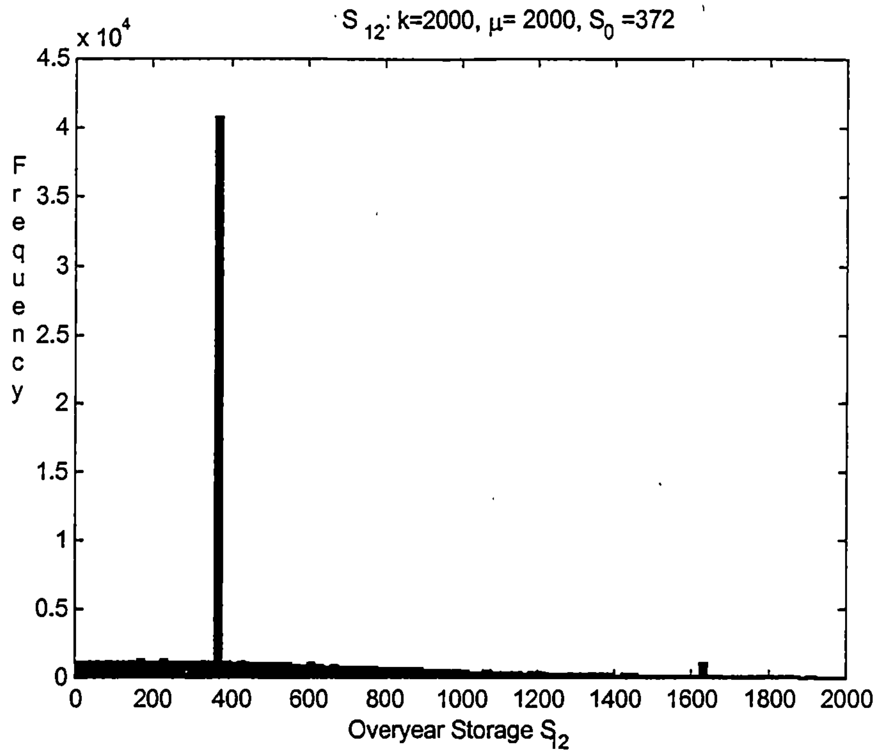

Figure 5.4: Frequency plot of S_{12}

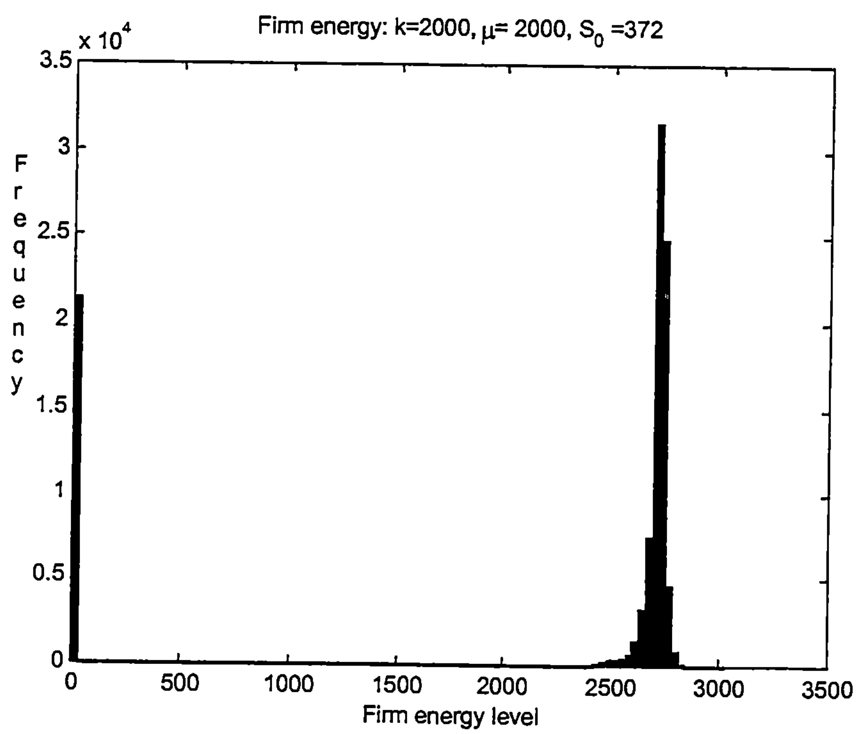

Figure 5.5: Frequency plot of L_{fe}

5.2.4 Summary

In this section, we have presented the first of this modeling effort using stochastic dynamic programming. We have made several assumptions that allowed us to devise a rather efficient alogorithm. The drawback of loosing the inherent dependence of the monthly inflows has been countered by the fact that we consider a large number of outcomes in each period which could essentially cover most of the scenario space even in the dependent inflows case. The results given by the model were validated using simulation. As evident from the discussion in the previous section, results of the model are fairly coherent. Through the simualtion experiments, it was shown that the optimal release policy maintained high reservoir reliabilities throughout most cases. Furthermore, the firm energy, which is the primary concern of the energy authority, remained consistently higher in simulation than that specified to the DP model. The firm energy relaibility also remained high indicating the robustness of the optimal release policy as validated by in simulation runs of 100,000 iterations. The main drawback still remain that the dependence structure in inflows has been ignored. In the next section, we present a model that would take into account the inherent dependence among the cumulative monthly inflow, yet taking advantage of a similar efficient algorithm as presented in this chapter.

5.3 Dynamic Programming: Restricted Dependent Inflows Models

5.3.1 Introduction

In Section 5.2, we presented a dynamic programming model with independent inflows. The results given by the model are satisfactory in general based on the simulation validation and analysis. We have shown, however, that the resulting optimal release policy is of the form $R_t^* = f(S_{t-1})$, which means that the release policy does not depend on the inflows realized so far explicitly. The continuity equation, $S_t = S_{t-1} + I_t - R_t$, may suggest that there has to be implicit dependence between R_t and I_t . In this section, we further strengthen the analysis by incorporating the monthly inflows information explicitly into the release policy, under the more realistic case that monthly inflows are dependent. This implies that the set of inflows and corresponding probabilities are not necessarily the same for each outcome in a given period. If we use dynamic programming in such a setting, we will run into major computational difficulties as the problem grows exponentially with the number of periods and scenarios. This is due to the fact that we would have to solve the DP model in (5.1) for each node in the scenario tree, which is $O(N^T)$ where T is the number of periods and N is the number of outcomes in each period. The challenge is to develop a model where the inflows do not depend on the history of the inflows $\mathcal{H}_{t-1}$, and yet make use of the cumulative inflow information. We will show in the coming section that we can accomplish this by considering the cumulative monthly inflows instead of the individual monthly inflows in constructing the scenario tree. In section 5.3.2, we formally present the dynamic programming model with *restricted* dependence. In Section 5.3.4 we present a suitable algorithm to solve the model efficiently. In Section 5.3.5, we present a simulation study of the results obtained using the model. And finally, in Section 5.3.6, we present some concluding remarks.

5.3.2 Dynamic Programming Model with Restricted dependence

In this section, cumulative monthly inflows will be explicitly considered in the decision policy. However, if we were to consider a full scale dependence, we would end up with a problem that is no easier to solve than the multi-stage stochastic program with recourse developed in Chpater 4. The restrictive nature of the dependence structure of inflows is described in the following assumption.

Assumption 5.3.1 *The individual monthly inflow at period t, random variable I_t, is assumed stochastically dependent only on the cumulative inflows up to period t, Q_{t-1}.*

5.3.3 Model Description

Under Assumption (5.3.1), for a given cumulative inflow Q_{t-1}, a given beginning storage level S_{t-1}, and a given release R_t, the dynamic programming recursive formulation can be rewritten as follows:

$$\begin{aligned}
\phi_t(Q_{t-1}, S_{t-1}) = \quad & \min_{R_t} E_{I_t|Q_{t-1}} \left[\mathcal{F}(q_{t-1}, S_{t-1}, I_t, R_t) + \phi(Q_t, S_t)\right] \\
& \text{s.}t. \\
& S_t + R_t = S_{t-1} + I_t \\
& S_t - SD = \delta_t^{SD} \\
& R_t - T_t = \delta_t^{D} \\
& S_t - (k - V) = \delta_t^{F} \\
& \nu - EG_t = \delta_t^{EG} \\
& S_T - S_0 = \delta^{S_0}, \quad if \quad t = T \\
& q_t = q_{t-1} + I_t \\
& R_t \geq 0.
\end{aligned} \tag{5.6}$$

Given a realization q_{t-1} of Q_{t-1}, we have

$$I_t|q_{t-1} = Q_t - q_{t-1},$$

$$E[I_t|q_{t-1}] = E[Q_t] - q_{t-1},$$

$$Var[I_t|q_{t-1}] = Var[Q_t].$$

Thus to generate a scenario tree for I_t in this case, one needs the distribution of $I_t|q_{t-1}$. This is due to the fact that for each instance of q_{t-1}, one needs to generate a set of outcomes from the distribution of $I_t|q_{t-1}$. This will, however, lead to a DP which is as complex and computationally tedious as that in the general setting . Note that we have a given distribution for Q_t and we need only generate one sample from the distribution of Q_t. Therefore, if we generate the scenario tree where the outcomes are cumulative monthly inflows, we get the same outcomes for each node in a particular period. However, the individual monthly inflows are dependent on the level of cumulative inflows of the previous period and are thus different in each outcome. This is also evident from the DP recursive formula in (5.6), as the value function is only dependent on the level of cumulative inflows realized thus far. From these observations, if we generate the senario tree where the outcomes in each period are based upon the cumulative inflows, then we need only to solve the problem for one node in each period. This is a tremendous advantage because it reduces the computational burden from solving $\sum_{i=0}^{12} 10^i$ problems to only solving 12 such problems.

5.3.4 Solution Algorithm

Given a set of discretized cumulative inflows Q_{t-1}^n, $n = 1, \ldots, N$, a set of discretized beginning storages S_{t-1}^j, $j = 1, \ldots, J$, and a set of discretized releases R_t^l, $l = 1, \ldots, L$, the algorithm can be described as follows.

Step 0:Initialization

1. Set reservoir size and maximum firm energy level

2. Obtain inflow data (mean, standard deviation, and number of outomes per period)

3. Generate outcomes and probabilities for each period from the provided distribution

4. Set t= T-1, where T= number of periods to be considered in the model

Step 1: DO WHILE n≤N and j≤J

1. Set $Q_{t-1} = Q^n_{t-1}$, and $S_{t-1} = S^j_{t-1}$, Let $\phi^*(Q_{t-1}, S_{t-1}) = M$, where M denotes a large positive number.

Step 2: DO WHILE l≤L

1. evaluate $\phi_t(, Q_{t-1}, S_{t-1}, R^l_t)$ as described in (5.6)

2. if $\phi^*(Q_{t-1}, S_{t-1}) \quad < \quad \phi_t(Q_{t-1}, S_{t-1}, R^l_t)$

   let $\phi^*(Q_{t-1}, S_{t-1}) \quad = \quad \phi_t(Qt-1, S_{t-1}, R^l_t)$ and $R^*_t(Q_{t-1}, S_{t-1}) = R^l_t$.

END DO

END DO

Step 3:

1. Set t=t-1.

2. if $t < 0$, STOP,

   else go to Step 1

5.3.5 Analysis and Simulation

In this section, we analyze the release policies generated by the preceding DP model. We first describe how random outcomes of different scenarios are generated and also how the discretized cumulative monthly inflows used in the model are generated. Then a simulation of the release policy is carried out to gain some insight into the complex structutre of the problem and to also validate the model's solution. The release policy of the current model is no longer a function of the beginning storage S_{t-1} only. The release in each period t depends on both the beginning storage S_{t-1} and the cumulative inflow of the previous period Q_{t-1}. Figures D.1 through D.11, in Appendix D, are a typical example of such a release policy. Note that the optimal release functions in Figures D.8 through D.11, which correspond to periods 9, 10, 11, and 12, are smoother than those in Figures D.1 through D.3 for instance, which correspond to periods 2, 3, and 4. This can be explained by the fact that uncertainty becomes smaller as monthly inflows are realized.

Furthermore, we notice a phenomenon that was not apparent in the independent inflows DP model of Section 5.2. Consider Figure D.7 for instance, which corresponds to the optimal release policy at the beginning of period 8, for a fixed beginning storage level, the optimal release level increases up to a certain critical cumulative inflow, Q^c_{t-1} and then starts decreasing as the cumulative inflow level Q_{t-1} increases. This is an interesting remark as it highlights one of the important tasks that the model is accomplishing. This is due to the fact that the model uses the knowledge of the value of the cumulative inflows in deciding which optimal release to choose. Before that critical point Q^c_{t-1}, the releases increase to maximize the energy generation. However, once that level is exceeded, the model realizes that the probability of large inflows in the coming periods decreases and so the releases are chosen to be smaller so as not to hamper energy generation in coming periods.

Simulation Analysis

[p] The simulation routine for this model is not too different form that of the previous model. The inputs consist of the reservoir size, the frim energy level desired, and the optimal release policy obtained from the model. The outputs include the simulated reservoir reliabilities, the simulated firm energy reliability, and the frequency files for the distribution of L_{fe} and S_{12}. Consider the case where the reservoir size is 2,000(MCM) and the frim energy level is required to be 4,000. In Figure 5.6, the simulated target satisfaction reliability θ_s is reported, and as evident form the graph, this reliability is very high (> 0.99). This proves that the model takes the target-priority operation in to consideration. Figure 5.7 depicts the simulated dead storage reliability α_s. The lowest level attained was 0.98.

Figure 5.8 shows the simulated flood storage reliability and we notice that it stays very high up to month 8 then drops a little before recovering in period 12. The lowest level attained was 0.65 which is still high for a flood storage reliability. Figures 5.9 and 5.10 depict the reliability of the potential energy generated and the reliability of the actual energy generated respectively. Two important observations in this regard. The first is that he simulated energy generation reliability $\gamma_{L_{fe}}$, in Figure 5.9, remains very high (> 0.9), in fact in 7 of the 12 monthly periods $\gamma_{L_{fe}} = 1$. This indicates that the model does cater to the primary concern of the reservoir management which is the generation of energy. The second remark is that the levels of the reliabilities of the potential energy and actual energy generation are close. This

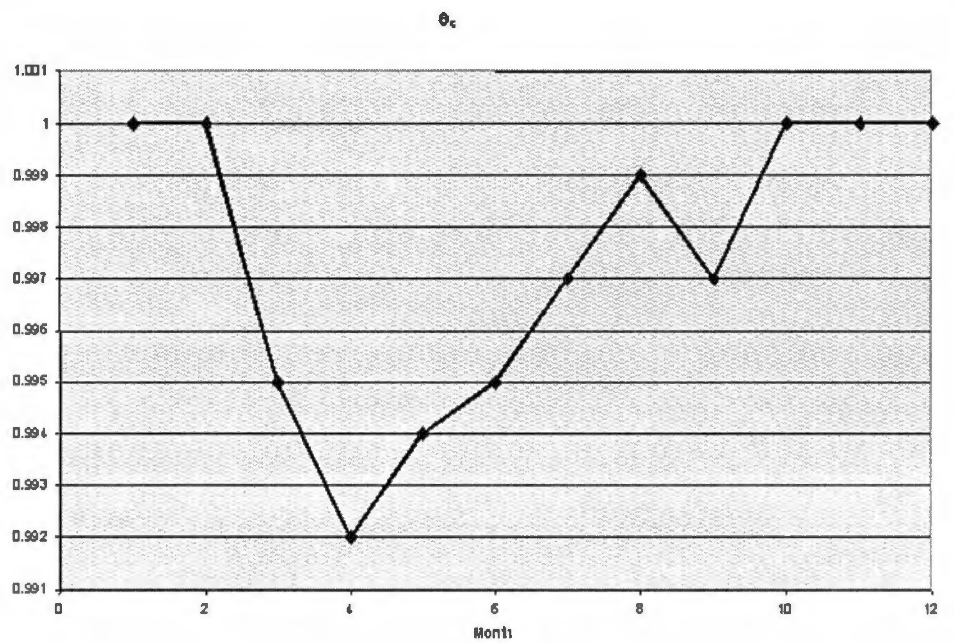

Figure 5.6: Simulated θ for Restricted Dependence Model

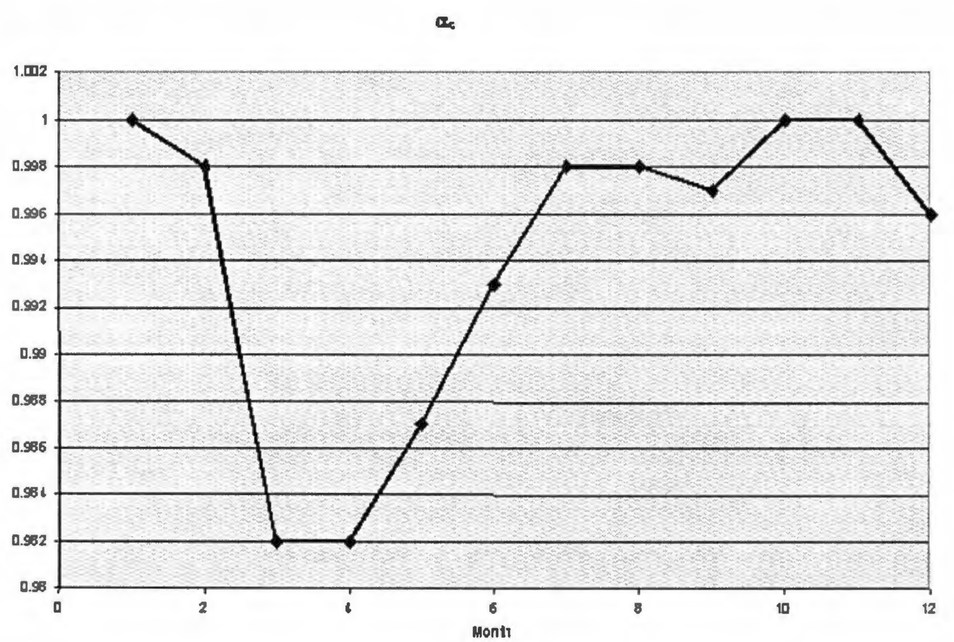

Figure 5.7: Simulated α for Restricted Dependence Model

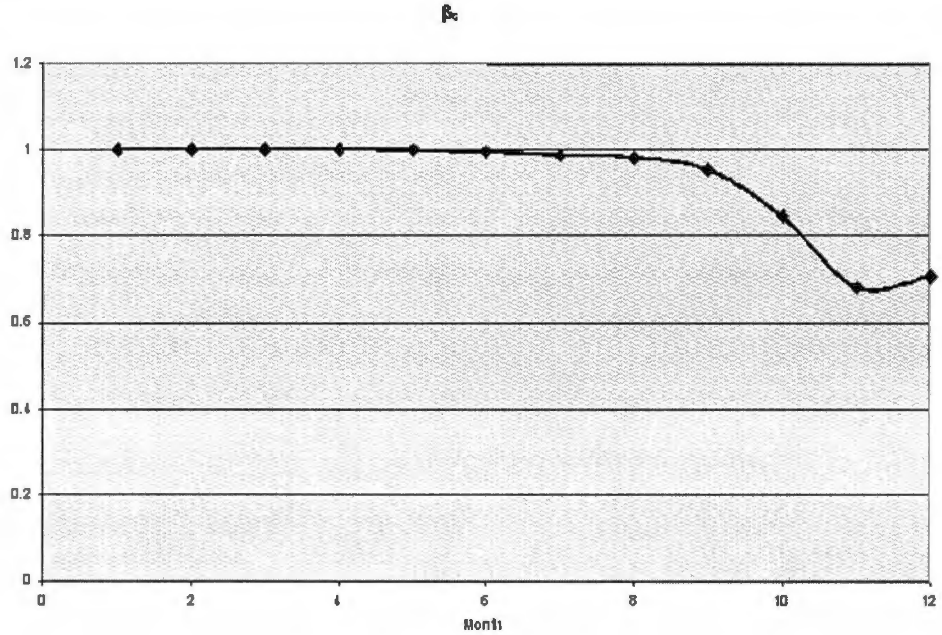

Figure 5.8: Simulated β for Restricted Dependence Model

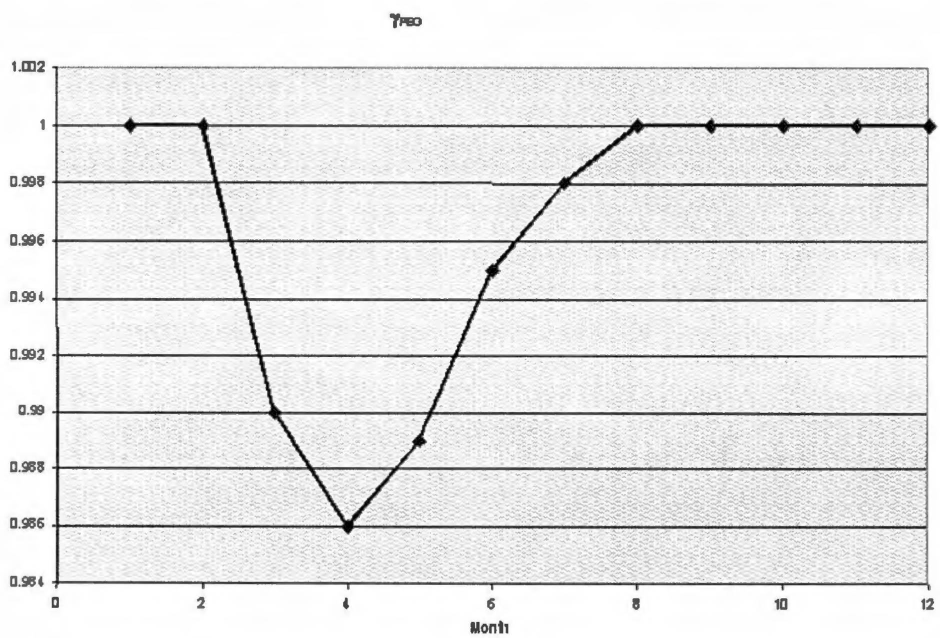

Figure 5.9: Simulated γ for Restricted Dependence Model

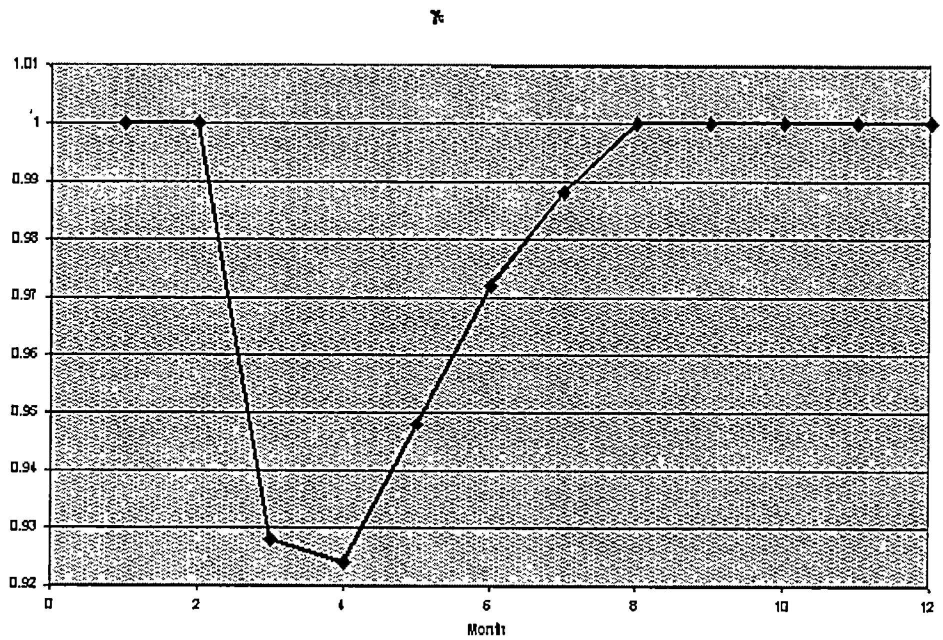

Figure 5.10: Simulated γ_{PEG}^{*} for Restricted Dependence Model

indicates that although the release policy is nonanticipative, the information provided by the cumulative inflow, the releases are fairly close to the potential release defined as the release of available water above the dead storage level SD. This validates the claim made in the preceding section that the knowledge of the level of the cumulative inflow helps the model in "estimating" the level of inflow in subsequent periods.

5.3.6 Summary of Results

To retain some of the inherent dependencies in the individual monthly inflows, and at the same time retain the computational efficiency of the independent inflows model, we have developed a model that is based on the cumulative inflow data. A scenario tree is generated from the lognormal distributions of the cumulative inflows. The model uses a dynamic programming algorithm that solves only one node per period. The model's solution seems to be very reliable when the results are analyzed. The simulation experiments further validate the model's solution. We have noted that although this is an independent inflows case, the restricted dependence formulation helped retain some of the dependence structure as was evident from the analysis of the results.

5.4 concluding Remarks

We have presented two efficient models that solve the multi-stage stochastic program presented in Chapter 4. The models solutions were analyzed and further validated using simulation models. However, neither model takes into account the inherent dependence among the individual monthly inflows. In a life supporting system like this one, it would be imperative to consider such dependencies explicitly. This will be the topic of the next chapter, where we develop a model that recognizes the dependence more rigorously. However, the model becomes extremely large, so we apply an aggregation technique as a means to reduce the size of the problem for efficient solution.

Chapter 6

Aggregated Dynamic Programming: Dependent Inflows Model

6.1 Introduction

In Chapter 5, we presented a dynamic programming model with independent inflows and a DP model for restricted dependence on the monthly inflows. In the former we have made the assumption that the individual monthly inflows are independent. We obtained an optimal release policy of the form $R_t^* = f(S_{t-1})$. In the DP model with restricted dependent inflows, we used the assumption that the monthly infows are only dependent on the level of cumulative inflows realized thus far and not on the particular history of the inflows. We derived an optimal release policy that is explicitly dependent on the beginning storage S_{t-1} as well the cumulative inflow Q_{t-1}, i.e. $R_t^* = f(S_{t-1}, Q_{t-1})$. While the latter model remained computationally tractable, it does not fully capture the dependence that exist among monthly inflows.

In this chapter, we consider a technique known as the "time-stage aggregation" whereby one combines or *aggregates* variables and constraints of several time periods into a grand/aggregated period. This concept was used in Edirisinghe and Ziemba (1992) to develop efficient bounds in the case of multistage stochastic convex models. We utilize this concept to develop an efficient DP-based solution technique to the multi-period reservoir problem presented in Chapter 4. In Section 6.2, we formally develop and present the proposed "aggregated dynamic programming" (ADP) model. Then, in Section 6.3, we develop a solution algorithm that solves the problem. In Section 6.4, we study the results obtained with the model and validate them using a simulation study. Finally, in Section 6.5, we conclude by summarizing the main aspects of the model and the main results obtained.

6.2 ADP Model With Dependent Inflows

6.2.1 Prelude

To solve the model developed inChapter 4 in its full dependent inflows formulation would be a formidable task. As menetioned previously, even a simple case of 12 periods with 5 outomes per period translates to solving $\sum_{i=0}^{11} 5^i$ subproblems, and thus even if each subproblem can be solved in $\frac{1}{1000}$th of a second, it would take about 68 hours to come to determine an optimal policy. Furthermore, the memory requirements in the solution process would be simply astronomical. The method of aggregation is a general decomposition method that has been used to reduce the probem size. The general idea of aggregation within mathematical programming has been studied extensively in the literature. The typical approach involves aggregating rows and/or columns to yield an approximate mathematical program that is easier to solve, see Zipkin (1980a, 1980b). In the context of stochastic mathematical programs, Birge (1985) and Edirisinghe and Ziemba (1992) developed approximations by aggregating constraints according to the underlying probability distributions. For a state fo the art survey on aggregation in optimization, see Rogers et al.(1991). Its applications span over a wide scope of applications. Aggregation can be applied at several levels. Some methods aggregate the search space by reducing it using some limiting assumptions. Other more commonly used methods of aggregation is to reduce the scenario space whereby several scenarios of the original problem are grouped together as one scenario, see Rockefeller and Wets (1991). The solution obtained would not be optimal to the original problem. It would, however, be feasible and would present a bound on the original solution. In a world puctuated by uncertainty, this limitation could be tolerated to achieve an implementable solution that can be obtained faster.

In our problem, the monthly inflows are highly uncertain random variables, i.e. they have high variance. For that matter, we use the aggregation principle to provide a *good* solution, although it might no be *optimal.* One way of aggregating the scenarios is to combine months or periods as stages and solve a smaller stochastic program. For example, if we divide the 12 periods into 4 stages where each stage has 3 periods, with 10 outcomes in each period, we solve 10^4 problems. Although the size of each subproblem may increase, the overall size of the DP model reduces to a more manageable size.

6.2.2 Model Description

Suppose we divide the number of periods, T, in the planning horizon into Π number of stages. This yields $\frac{T}{\Pi}$ periods in each stage assuming $\frac{T}{\Pi}$ is integral. Let $\mathcal{H}_{\pi-1}$ be defined as the history of inflows up to stage π, i.e.,

$$\mathcal{H}_{\pi-1} = \hat{I}_1, \ldots, \hat{I}_{\pi-1}$$

where $\hat{I}_\pi$ is the vector of inflows in stage π. Note here that the history of the inflows is a series of vectors of size $\frac{T}{\Pi}$. For instance, if $T = 12$, and $\Pi = 3$, then we get $\frac{12}{3} = 4$ periods per stage. In this case, the inflows vector in stage 1, $\hat{I}_1$, consists of the random inflows of periods 1, 2, 3, and 4, i.e $\hat{I}_1 = (I_1, I_2, I_3, I_4)'$. With these, we define the state of the system to be $(\mathcal{H}_{\pi-1}, \mathcal{S}_{\pi-1})$, where $\mathcal{S}_{\pi-1}$ is the storage level of the reservoir at the beginning of stage π. Also define $\hat{I}_\pi$ to be the vector of inflows and $\hat{R}_\pi$ to be the vector of releases in stage π. Note that in the following, we use the notation that $\hat{I}^i_\pi$ to denote the inflow in period i of stage π, and the same holds for the releases. Hence, we write the DP recursion formulae as the following:

$$
\begin{aligned}
\phi_\pi(\mathcal{H}_{\pi-1}, \mathcal{S}_{\pi-1}) = \min_{\hat{R}_\pi} \quad & E_{\hat{I}_\pi|\mathcal{H}_{\pi-1}} \left[\mathcal{F}(\mathcal{H}_{\pi-1}, S_{\pi-1}, \hat{I}_\pi, \hat{R}_\pi) + \phi(\mathcal{H}_\pi, \mathcal{S}_\pi) \right] \\
& \text{s.}t. \\
& \mathcal{S}_\pi^1 + \hat{R}_\pi^1 = \mathcal{S}_{\pi-1} + \hat{I}_\pi^1 \\
& \mathcal{S}_\pi^2 + \hat{R}_\pi^2 = \mathcal{S}_\pi^1 + \hat{I}_\pi^2 \\
& \vdots \\
& \mathcal{S}_\pi^p + \hat{R}_\pi^p = \mathcal{S}_\pi^{p-1} + \hat{I}_\pi^p \\
& \hat{R}_\pi^1 - T_\pi^1 = \delta_\pi^{D,1} \\
& \hat{R}_\pi^2 - T_\pi^2 = \delta_\pi^{D,2} \\
& \vdots \\
& \hat{R}_\pi^p - T_\pi^p = \delta_\pi^{D,p} \\
& \mathcal{S}_\pi^1 - (k - V) = \delta_\pi^{F,1} \\
& \mathcal{S}_\pi^2 - (k - V) = \delta_\pi^{F,2} \\
& \vdots \\
& \mathcal{S}_\pi^p - (k - V) = \delta_\pi^{F,p} \\
& \nu - EG_\pi^1 = \delta_\pi^{EG,1} \\
& \nu - EG_\pi^2 = \delta_\pi^{EG,2} \\
& \vdots \\
& \nu - EG_\pi^p = \delta_\pi^{EG,p} \\
& \mathcal{S}_\Pi^p - \mathcal{S}_0 = \delta^{\mathcal{S}_0}, \\
& \hat{R}_\pi \geq 0.
\end{aligned} \tag{6.1}
$$

Where p is the index of the last period in each stage π.

6.3 Solution Algorithm

At each node in the scenario tree, we compute the value function defined in (6.1). Note that each node in stage π is completely defined by the history of inflows, $\mathcal{H}_{\pi-1}$. Consider a set of discretized beginning storage levels $\mathcal{S}^n_{\pi-1}$, $n = 1, \ldots, N$, and a set of discretized releases $\hat{R}^j_\pi$, $j = 1, \ldots, J$. The following algorithm evaluates the value function at each node of each stage for a pre specified set of discretized beginning storage levels and releases. Note, however, that releases in this model are vectors, in which case we would have a p-dimensional search on the releases.

Step 0:Initialization

1. Set reservoir size and maximum firm energy level
2. Obtain inflow data (mean, standard deviation, and number of outomes per period; we also need probability of each outcome
3. Set $\pi = \Pi - 1$, where $\Pi =$ number of stagess to be considered in the model
4. The number of periods per stage is $p = \frac{T}{\Pi}$

Step 1: DO WHILE n≤N:

1. Set $\mathcal{S}_{\pi-1} = \mathcal{S}^n_{\pi-1}$, Let $\phi^*_\pi(\mathcal{H}_{\pi-1}, \mathcal{S}_{\pi-1}) = M$, M denotes a large number.

Step 2: DO WHILE j≤J:

1. evaluate $\phi_\pi(\mathcal{H}_{\pi-1}, \mathcal{S}_{\pi-1}, \hat{R}^j_\pi)$ as described in (6.1)

2. if $\phi^*_\pi(\mathcal{H}_{\pi-1}, \mathcal{S}_{\pi-1}) \quad < \quad \phi_\pi(\mathcal{H}_{\pi-1}, \mathcal{S}_{t-1}, \hat{R}^j_\pi)$

   Set $\phi^*_\pi(\mathcal{H}_{\pi-1}, \mathcal{S}_{\pi-1}) \quad = \quad \phi_\pi(\mathcal{H}_{\pi-1}, \mathcal{S}_{\pi-1}, \hat{R}^j_\pi)$

   and $\hat{R}^*_\pi(\mathcal{H}_{\pi-1}, \mathcal{S}_{\pi-1}) = \hat{R}^j_\pi$

END DO

END DO

Step 3:

1. Set $\pi = \pi - 1$.

2. if $\pi < 0$, STOP,

   else go to Step 1

6.4 Results Analysis and Simulation Study

6.4.1 Results Analysis

The model presented in thsi chapter is an attepmt to retain some of the dependence structure among the monthly inflows, and by using an aggregation technique, we reduce the overall problem size. The use of DP enabled us to decompose the general problem into smaller problems relevant to each node in the scenario tree. There is a potentially large number of experiments that could be run on the model to see the effect of changing the number of stages for a 12 monthly periods model. As expected, the computational cost increases with the addition of stages. Memory requirements also increase exponentially makning it almost impossible to solve the model for a reasonable number of outcomes (> 20 per stage). This restricted the scope of our experiments to the case where we either use a small number of stages, say 3, or use a small number of outcomes, say 10. Despite this drawback, the model's results seemed quite accurate. The model's solution in this case is a release policy for each stage. The release policy is a vector of releases for the number of periods in a particular stage.

6.4.2 Simulation Study

Finding The Appropriate Scenario

The simulation study for this model is different from that of the previous models. In the previous models, we have considered scenarios that are based on the individual monthly inflows. We have also assumed independence among the inflows, so it did not matter which inflows was realized. We have considered one node in each period, so the release policy in each node of a particular period is the same. In the present model, this is no longer the case. The release policy does depend explicitly on which inflows have been realized. Say we are at stage 0, and we generate a random vector of inflows for the number of periods in that stage, the release policy for the next stage is dependent on which inflow vector has realized. Recall that in generating the scenario tree, each branch corresponds to a vector of inflows, say $\hat{I}^G$. In the simulation, we generate a vector of equal size $\hat{I}^R$, and that decides which of the nodes of the coming stage will be the node of reference, i.e. which release policy to use for the next stage. Basically, the branch of the scenario tree that most closely resembles the randomly generated vector of inflows is the scenario we follow to get the release policy. There are potentially several methods of measuring the distance between vectors. The most common being the *Euclidean* distance, ED, computed as follows:

$$ED = [\sum_{i=t}^{t+p}(I_i^G - I_i^R)^2]^{\frac{1}{2}}.$$

where p is the number of periods in each stage. The scenario branch that minimizes the Euclidean distance is chosen as the scenario that is being manifested.

Simulation Experiments and Results

The simulation is designed to not only validate the solution given by the model, it is also designed to provide further insight into the problem. The setup for the simulation remains no different from the one used in the previous models. We feed into the simulation model the reservoir size, the firm energy level, and the release policy tree. The release policy is now a tree since it depends on which scenario manifested. In other words, the release policy that is a function of the beginning storage is different for each node in a period, and betwen periods. The outcome of the simulation is similar to the previous models, where the intention is to validate the model's solution through the observation of the frequency of deviating from the optimal reservoir characteristics. Since the energy generation is of primal concern to the energy authority, we also validate the solution by observing the simulated firm energy level and reliability. The simulated target satisfaction relaibility is depicted in Figure 6.1. As evident form the graph, θ_s remains very high. In 10 of the 12 periods, $\theta_s = 1$, indicating that the target priority operation is being carried out by the model.

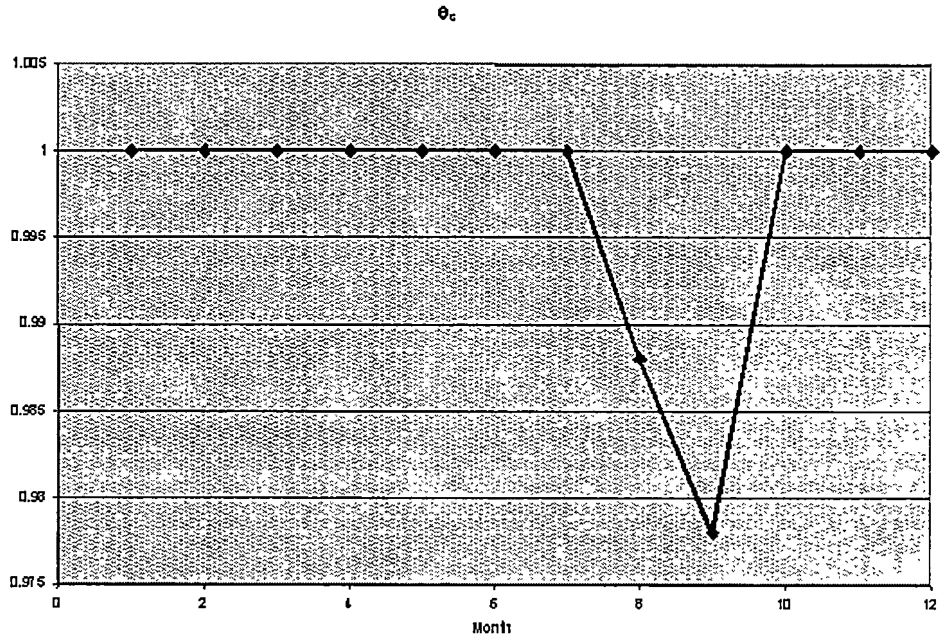

Figure 6.1: Simulated θ for Aggregated DP Model

We see a slight drop in the target satisfaction reliability in months 9 and 10, which correspond to the two months with the highest demand, see Figure 6.2. Even in such circumstances, the target satisfaction relaibility does not fall below 0.975. Another important aspect of the reservoir operation is the dead storage reliability α. Figure 6.3 depicts the simulated α_s, and its level is consistently higher than 0.9. To control flood hazards, the model provides a release policy that yields a high flood reserve relaibility β_s, as evident from Figure 6.4. Note the slight drop in β_s towards the later months. This may be explained by the fact that towards the end of the planning horizon, the overyear storage reliability becomes a concern. In order to satisfy the condtion that $S_T \geq S_0$, larger volumes of water are stored in the reservoir in the later periods.

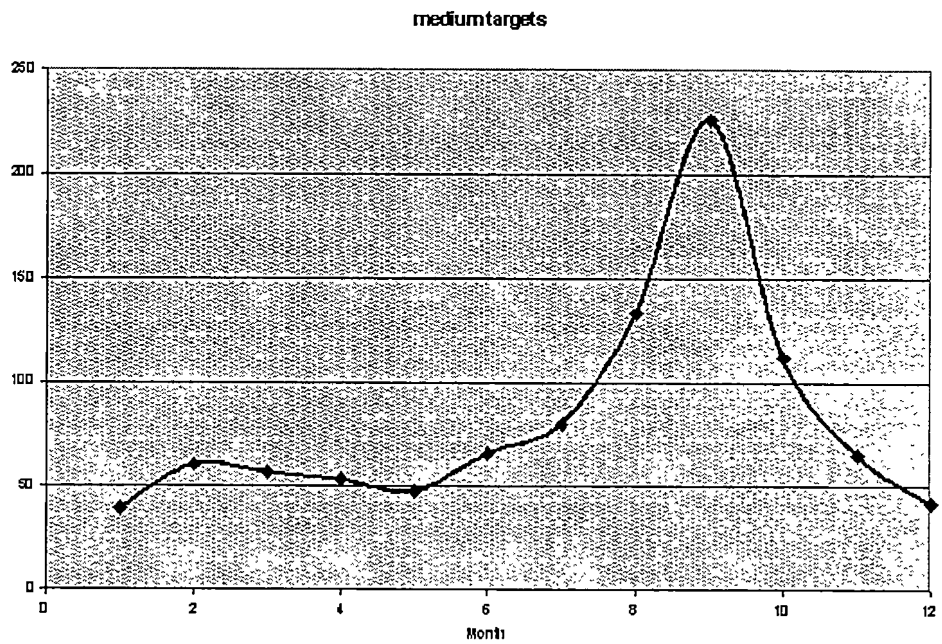

Figure 6.2: Medium monthly water targets

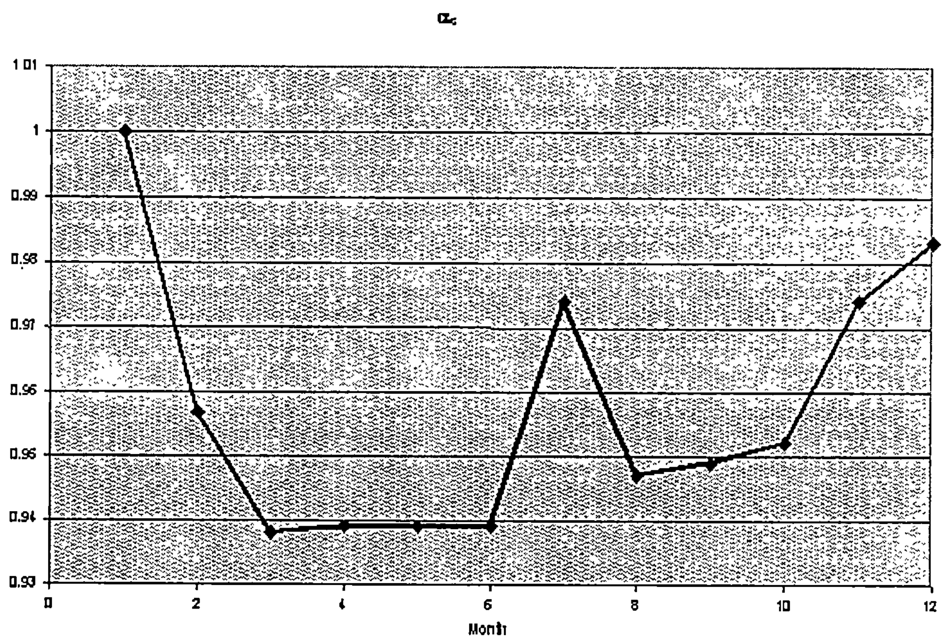

Figure 6.3: Simulated α for Aggregated DP Model

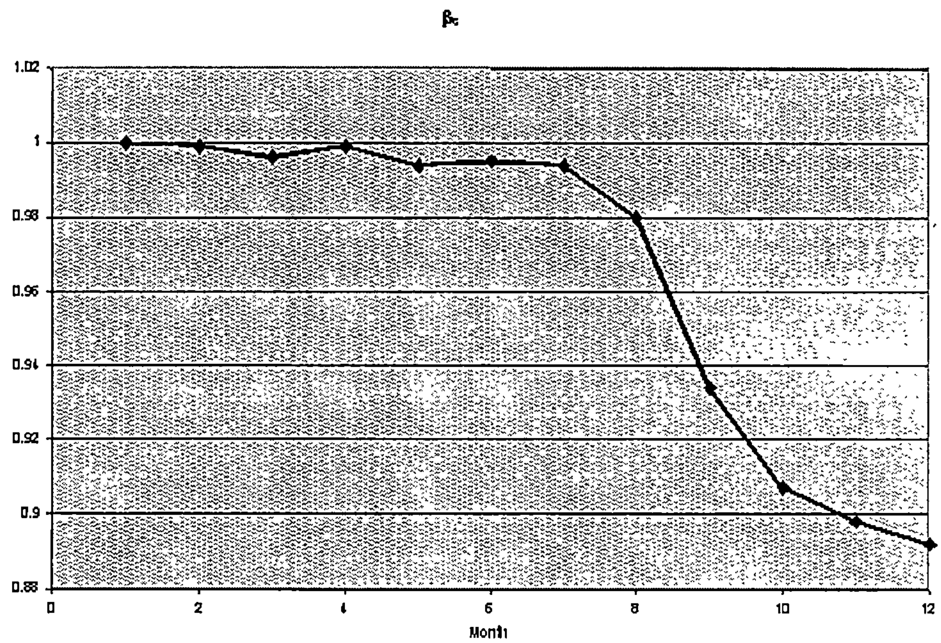

Figure 6.4: Simulated β for Aggregated DP Model

Firm energy is the primary concern of the reservoir authority as it provides the main source of income. Figure 6.5 depicts the reliability of the potential energy generation, and Figure 6.6 depicts the reliability of the actual energy generation. Recall that the actual energy generated is the energy generated by following the optimal release schedule provided by the model, and potential energy generated is the energy that could be generated if the volume of available water above the dead storage level is relesed. Two important remarks on the performance of the model in this regard have to be mentioned here. The first is that the actual energy generation reliability remains consistently higher than 0.9 as shown in Figure 6.5. This indicates that the model's release policy is maximizing the energy generation. The second remark is

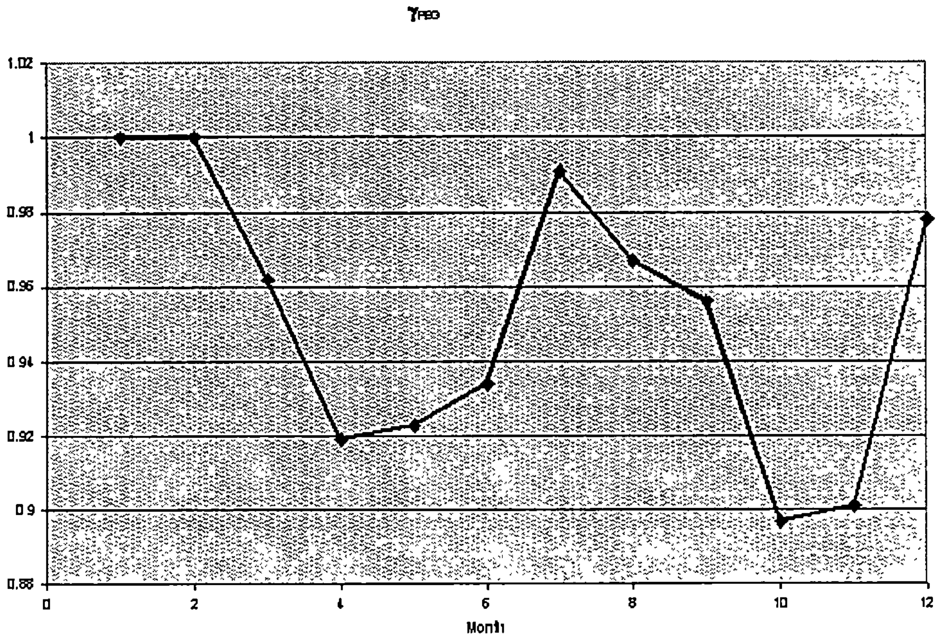

Figure 6.5: Simulated γ for Aggregated DP Model

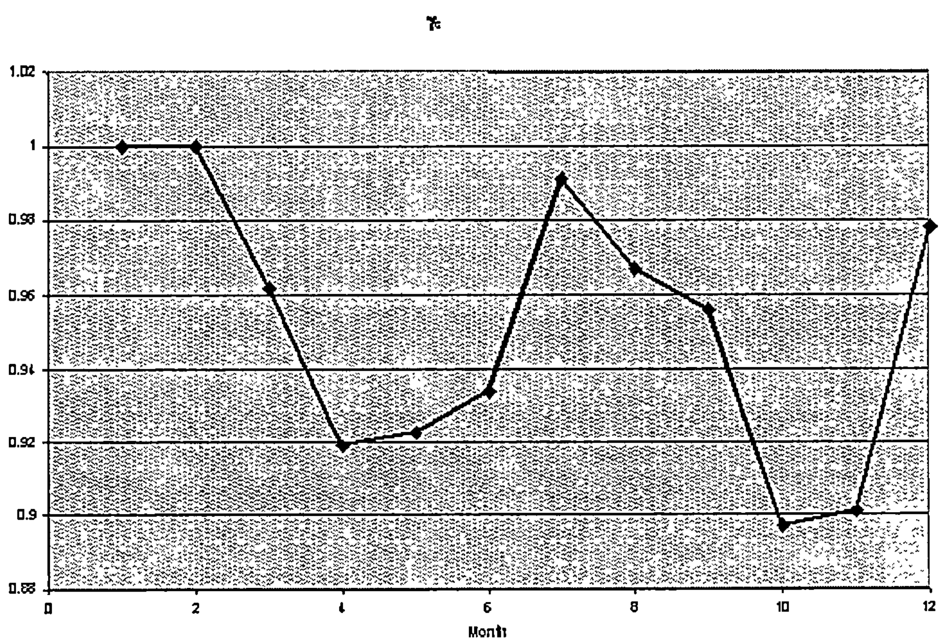

Figure 6.6: Simulated γ_{PEG} for Aggregated DP Model

that the levels and the behavior of the potential energy and actual energy generation relaibilities are close. This leads us to the conclusion that the release policy is in effect releasing as much water as possible without dropping the water level below the dead storage reliability, given that all other reservoir reliabilities are not degraded as eveident form the discussion of Figures 6.1, 6.3, and 6.4.

6.5 Concluding Remarks

In this chapter, we presented a dynamic programming solution for the multi-stage stochstic model we have developedin chapter 4. The model considers the case of dependent inflows. To alleviate the computational burden, a "time-stage" aggregation scheme was used as a means of reducing the problem size. An efficient solution technique was developed. The model's release policy was validated using a simulation study. The reservoir reliabilities, namely the dead storage, the target satisfaction, the flood reserve, and the firm energy relaibilities remained consistently high, indicating the accuracy of the approach developed.

Chapter 7

Conclusion

7.1 Comparison of the Different Models

We have presented several state of the art models to study the planning and operation problem for a single mulit-purpose water reservoir. For each model presented, we have carried out a detailed simulation analysis. What remains to be done is to compare the performance of each model relative to the other models. This is an important issue since there are several factors that influence the performance of a model, such as the planning horizon, the data available, and decision time frame. We have shown through the use of simulation that all of the models performed fairly well given the assumptions that were imposed. In this chapter, we will compare these models on several aspects: validity of model results, computational efficiency, as well as the model robustness. In section 7.2, the thesis is concluded with possible directions for future research.

7.1.1 Validity of Optimal Solution

We have shown through the results analysis and the simulation study of the different models that all of the models generated accurate results. The chance constrained model have the flexibility of providing an optimum reservoir size and an optimum dead storage reliability. The three stochastic programming models with recourse have the flexibility of accounting for inflow dependence to a certain degree. We have shown that these models, being planning models, have all suceeded in meeting the multiple and conflicting objectives of the reservoir operation and management. We have also shown, however, that the optimal solution and especially the firm energy level and reliability, are significantly affected by the beginning storage level S_0 and the over year storage reliability ρ. Overall, the stochastic programming models with recourse provided solutions that could better accomodate the reservoir operating characteristics. This is due to the fact that these models actually consider the randomness in the inflows explicitly in the search for a solution through a scenario approach, whereas the CCP model only uses the marginal distributions.

7.1.2 Computational Efficiency

The chance constrained model is generally efficient in computation; for instance, for the case of a specified reservoir relaibilites, the problem is solved in less than 60 seconds. The computational efficiency of the three stochastic programming with recourse models, however, depend on the number of random inflow outcomes considered in each period, and on the fineness of the discretization of beginning storage and beginning cumulative inflow vectors. This is understandable since these two parameters increase the search space exponentially. Despite this fact, solution of the independent inflows model, for the case of 50 outcomes per period and 50 discrete points for the search space arrays, takes about 12 seconds including generating the scenarios. The restricted dependent inflows model takes about 6 minutes, because we are adding a new search dimension, i.e., cumulative inflows. Computational time of the aggragated dependent inflows model depends not only on the above dimensions but also on the number of aggregations. For example, the case with 4 stages, i.e. 3 periods per stage in a 12 month planning horizon, and 30 outcomes per stage takes approximately 30 minutes. In a situation where a preliminary solution is desired, it can be concluded that the restricted dependent inflows model is the preferred choice because of its solution efficiency, and its ability to incorporate partial inflow dependence information by considering a large number of outcomes in each period.

7.2 Summary

In this thesis, we have considered the problem of optimally planning the capacity of a water reservoir under a special *target priority operation.* To account for the randomness in the monthly water inflow, we have used stochastic programming as a tool to represent the random event. In Chapters 2 and 3, we have presented a chance constrained goal programming model. We have shown via the model analysis and the simulation study that one obtains robust release policies. We have investigated, in particular, the effect that the beginning storage and the year ending storage reliability have on the firm energy and relaibility of operation. Realizing the fact that a chance constrained formulation fails to take into account the magnitude of the constraint violations, three different models based on stochastic programming with recourse were proposed. Modeling the problem with explicit dependent structure among monthly inflows is computationally intractable since it makes the problem very large. To alleviate this burden, we have made certain simple, but practical, assumptions on the structure of inflows. On one hand, we assumed that monthly inflows are independent of each other. This is not a very restrictive assumption as long as a large number of outcomes is generated. The resulting nonanticipative release policy was analyzed via simulation and the results are shown to be quite accurate. On the other hand, we assumed that the monthly inflows are dependent only on the cumulative inflow realized thus far, but not on the complete history of inflows. Under this assumption, an effcient solution algorithm was developed to evaluate the DP recursion. The optimal relase policy generated by the model depends on the beginnning storage and on the cumulative inflow realized so far. Thus, although the releases do not depend explicitly on the history of the inflows, some dependence is preserved in the information in the cumulative monthly inflows. In other words, although the release do not depend on which scenario was followed, they depend on the accumulation of inflow thus far. In the final modeling effort using stochastic dynamic programming with recourse, we propose a model that captures dependence among monthly inflows in a more general setting. However, an aggregation methodology was proposed to circumvent the exponential growth of the resulting DP model. The aggregated dynamic programming model so obtained was solved using an efficient algorithm and the resulting release policy, which now is dependent on a particualr scenario being foloowed, was validated using simulation. The results show that the latter release policy outperforms the optimal policies of the other models in the metrics of reservoir reliability and the firm energy production and reliability.

7.3 Future research

This thesis presents a comprehensive study of the optimal capacity planning problem of a multi purpose water reservoir under uncertainty. While the proposed models provided robust results, their focus is limited to monthly decision periods. An operational model would have to take into account a decision period much shorter than a month, and would need to have the flexibility of providing better solutions as random events unfold. One possible alternative is the *rolling horizon* approach, where the model is re-solved at the end of each operational period, the model being revised with the new data being observed. These avenues would certainly be worth considering and should be the subject of future research.

Bibliography

Bibliography

[1] Akaike, H., "Information Theory and an Extension of the Maximum Likelihood Principle," *Second International Symposium on Information Theory*, B.N. Petrov and F. Csaki Ed., Budapest: Academia Kiado, (1971), 267-281.

[2] Allen, R.B., and S.G. Bridgeman, "Dynamic Programming in Hydropower Scheduling," *Journal of Water Resources Planning and Management*, ASCE, 112(3), (1986).

[3] Arunkumar, S., and W. W-G Yeh, "Probabilistic Models in the Design and Operation of a Multi-purpose Reservoir System," *Contrib. 144*, Clifornia water Resources Center, U.C. Davis, December (1973).

[4] Bawa, V.S., "On Chance-Constrained Programming Problems with Joint Constraints," *Management Science*, 19 (1976), 1326-1331.

[5] Bazaraa, Mokhtar S., Hanif D. Sherali, and C.M. Shettys, *Nonlinear Programming: Theory and Algorithms*, Second Edition. John Wiley & Sons Ltd.. New York, N.Y. (1993).

[6] Bellman, R.E., *Dynamic Programming,* Princeton University Press, Princeton, N.J., 1957.

[7] Birge, J. R., "Aggregation Bounds in Stochastic Linear Programming," *Mathematical Programming*, 31 (1985), 25-42.

[8] Bozdogan, H., "Model Selection and Akaike's Information Criterion(AIC): The General Theory and its Analytical Extensions," *Psychometrica*, 52(3) (1987), 345-370.

[9] Bozdogan, H., "ICOMP: A New Model Selection Criterion," *Classification and Related Methods of Data Analysis*, Hans H. Bock Ed., Amsterdam: North-Holland, (1988), 599-608.

[10] Charnes, A., W.W. Cooper, and G.H. Symonds, "Cost Horizons and Certainty Equivalents: An Approach to Stochastic Programming of Heating Oil," *Management Science*, 4(14) (1958), 235-263.

[11] Curry, G. L., Helm, J. C., and Clark, R. A., "Chance constrained model of a system of reservoirs," *Journal of the Hydraulic Division*, ASCE, 99(HY12)(1973), 2353-2366.

[12] Dantzig, G.B.,"Linear Programming Under Uncertainty," *Management Science*, 1(3/4), (1955), 197-206.

[13] Dempster, M. A. H., *Introduction to Stochastic Programming*, in Stochastic Programming, M.A.H. Dempstare (ed.), Academic Press, London, (1980), 3-59.

[14] Dennis, J.E. and Shnabel, R.B.,*Numerical methods for Unconstrained Optimization and Nonlinear Equations*, Princeton University Press, Princeton, New Jersey, (1983).

[15] Diaz, G.E., and D.G. Fontane, "Hydropower Optimization via Sequential Quadratic Programming," *Journal of Water Resources Planning and Management*, ASCE, 115(6), (1989).

[16] Dirickx, Y. and Jennergren, L., "On the Optimality of Myopic Policies in Sequential Decision Problems," *Management Science*, 21 (1975), 550-556.

[17] Duren, F.K., and L.R. Beard, "Optimization of Flood Control Allocation for a Multipurpose Reservoir," *Water Resources Bulletin*, AWRA, 8(4), (1972).

[18] Dorfman, R., "Mathematical Models: The Multi-structure Approach, in the Design of Water Resources Systems," Harvard University Press, Cambridge, Massachussetts, (1962).

[19] Dupacova, J., "Water Resources System modeling Using Stochastic Programming Models," *Recent Results in Stochastic Programming*, Springer-Verlag, New York (1980).

[20] Dupacova, J, Gaivoronski, A., Kos, Z., and Szantai, T., "Stochastic Programming in Water Management: A Case Study and a Comparison of Solution Techniques," *European Journal of Operational Research*, 52 (1991), 28-44.

[21] Edirisinghe, N. C. P., and W. T. Ziemba, "Tight Bounds for Stochastic Convex Programs," *Operations Research*, 40 (1992), 660-677.

[22] Edirisinghe, N. C. P., and W. T. Ziemba, "Bounds for Two-Stage Stochastic Programs with Fixed Recourse," *Mathematics of Operations Research*, 19 (1994a), 292-313.

[23] Edirisinghe, N. C. P., and W. T. Ziemba, "Bounding the Expectation of a Saddle Function, with Application to Stochastic Programming," *Mathematics of Operations Research*, 19 (1994b).

[24] Edirisinghe, N. C. P., and G. M. You, "Second-Order Approximation and Refinement in Optimization Under Uncertainty," *Annals of Operations Research*, 64 (1996), 143-178.

[25] Edirisinghe, N. C. P., and W. T. Ziemba, "Implementing Bound-Based Approximations in Convex-Concave Two-Stage Stochastic Programming," *Mathematical Programming*, 75 (1996), 295-325.

[26] Edirisinghe, N.C.P., "Bound-Based Approximation in MultiStage Stochastic Programming," *Annals of Operations Research* 85(1999) 103-127.

[27] Edirisinghe, N.C.P., Ike Patterson, and Nasreddine Saadouli, "Capacity Planning Model for a Multipurpose Water Reservoir with target-Priority Operation," submitted for publication in*Annals of Operations Research* (2000).

[28] Fiacco, A.V., Introduction to Sensitivity and Stability Analysis in Nonlinear Programming, Academic Press, New York, (1983)

[29] Garstka, S.J. and Wets, R.J-B., "On Decision Rules in Stochastic Programming," *Mathematical Programming*, 7(2) (1974), 117-143.

[30] Giles, J.E., and W.O. Wunderlich, "Weekly Multipurpose Planning Model for TVA Reservoir System," *Journal of Watre Resources Planning and Management*, ASCE, 107(WR2), (1981).

[31] Gleit, A., *Stochastic Linear Programming*, Mathematisk Institut, Aarhaus Universitat, Aarhaus, Denmark, (1977).

[32] Hall, W.A., W.S. Butcher, and A. Esogbue, "Optimization of the Operations of a Multipurpose Reservoir by Dynamic Programming," *Water Resources Research*, AGU, 4(3), (1968), 471-477.

[33] Hall, W. A., and J. A. Dracup, *Water Resources Systems Engineering*, McGraw Hill, New York, NY, (1970).

[34] Hogan, A. J., J. G. Morris, and H. E. Thompson, "Decision Problems Under Risk and Chance Constrained Programming: Dilemmas in the Transition," *Management Science*, 27(6), (1981), 698-716.

[35] Houck, M. H., "A Chance Constrained Optimization Model for Reservoir Design and Operation", *Water Resources Research*, 14(5) (1979), 1011-1016.

[36] Houck, M.H., and J.L. Cohon, "Sequential Explicitly Stochastic Linear Programming Models: A Proposed Method for Design and Management of Multi-purpose Reservoir Systems," *Water Resources Research*, 14(2), (1978), 161-168.

[37] Houck, M. H., Cohon, J. L., and ReVelle, C. S., "linear Decision Rules in Reservoir Design and Management 6: Incorporation of Economic Efficiency Benefits and Hydroelectric Energy Generation," *Water Resources Research*, 16(1) 91980), 196-200.

[38] Houck, M. H. and Datta, B., "Performance Evaluation of a Stochastic Optimization Model for reservoir Design and Management with Explicit Reliability criteria," *Water Resources Research*, 17(4) (1981), 827-832.

[39] Joeres, E.F., Seus, G.J., and Engelmann, H.M., "The Linear Decisin Rule (LDR) Reservoir Problem with Correlated Inflows : 1. Model Development," *Water Resources Research*, 17(1) (1981), 18-24.

[40] Kall, P., "Stochastic Programming", *European Journal of Operational Research*, 10(2) (1982), 125-130.

[41] Lamond, Bernard F. and Matthew J. Sobel, "Exact and Approximate Solutions of Affine Reservoir Models," *Operations Research*, Vol. 43, No. 5, 1995.

[42] Law, A.M and W D. Kelton, *Simulation Modeling and Analysis*, McGraw-Hill, New York, N.Y., (1991).

[43] LeClerc, G. and Marks, D.H., "Determination of the Discharge Policy for Existing Reservoir Networks under Differing Objectives," *Water Resources Research*, 9(5) (1973), 1155-1165.

[44] Lepp, R., "Approximations to Stochastic Programs with Complete Recourse," *SIAM Journal on Control and Optimization*, 28 (1990), 382-394.

[45] Loucks, D.P., "Some Comments on Linear Decision Rules and Chance Constraints," *Water Resources Research*, 6(2) (1970), 668-671.

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

[47] Loucks, D.P., Stedinger, J.R. and Haith, D.A., *Water Resources Planning and Analysis*, Prentice Hall, Englewood Cliffs, N.J., (1981).

[48] Loucks, D.P., "Computer Models for Reservoir Regulations," *Journal of th Sanitary Engineering Divisin*, ASCE, 94(SA4), (1968).

[49] Manne, A.S., "Product Mix Alternatives: Flood Control, Electric Power and Irrigation," *International Economics Review*, 8(1), (1962), 30-54.

[50] Martin, Q.W., "Optimal Daily Operation of Surface-Water Systems," *Journal of Water Resources Planning and Management*, ASCE, 113(4), (1987).

[51] Nayak, S.C. and Arora, S.R., "Optimal Capacities for a Reservoir System Using the Linear Decision Rule," *Water Resources Research*, 7(3) (1971), 485-498.

[52] Nemhauser, G.L., *Introduction to Dynamic Programming,* John Wiley, New York, 1966.

[53] Neveda, C., "Some Applications of Stochastic Optimization Methods to the Electric Power System," *Numerical techniques for Stochastic Optimization*, Springer-Verlag, New York, (1988), 455-464.

[54] Prekopa, A., "On Probabilistc Constrained Programming," *Mathematical programming Study*, 28 (1970), 113-138.

[55] Prekopa, A., and T. Szantai, "Flood Control Reservoir System design Using Stochastic Programming," *Mathematical programming Study*, 9 (1978a), 139-151.

[56] Prekopa, A., and T. Szantai, "On Optimal Regulation of a Storage Level with Application to the Water Level Regulation of a Lake," *European Journal of Operational research*, 3 (1978b), 175-189.

[57] Prekopa, A., and T. Szantai, "Multi-Stage Stochastic Programming Model for the Control of Reservoir Systems," *Communications in Hydrology*, 60 (1980), 7-14.

[58] Prekopa, A., *Stochastic Programming*, Kluwer Academic Publishers, Dordrecht, The Netherlands,(1995).

[59] Prekopa, A., "The use of Discrete Moment Bounds in Probabilistic Constrained Stochastic Programming Models," *Annals of Operations Research*,85 (1999), 21-38.

[60] Rapcsak, T., *On the Numerical Solution of a Rservoir Model*, Ph.D. Thesis, University of Debrecen, hungary, (1974).

[61] ReVelle, C.S. and Gundelach, J., "Linear Decision Rule in Reservoir Management and Design, Part 4: A Rule that Minimizes Output Variance," *Water Resources Research*, 11(2) (1975), 197-203.

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

[63] ReVelle, C. and Kirby, W., "Linear Decision Rule in Reservoir Management and Design, 2: Performance Optimization," *Water Resources Research*, 6(4) (1970), 1033-1044.

[64] ReVelle, C., *Optimizing Reservoir Resources, Including a New Model for Reservoir Reliability*, John Wiley and Sons, Inc., New York, N.Y., (1999).

[65] Rockafeller, R. T., and Roger J.-B. Wets, "Scenario and Policy Aggregation in Optimization Under Uncertainty," *Mathematics of Operations Research*, 16(1) (1991), 119-147.

[66] Roger, D. F., R. D. Plante, R. T. Wong, and J. R. Evans, "Aggregation Techniques and Methodology in Optimization," *Operations Research*, 39 (1991), 553-582.

[67] Rosenthal, R.E., "A Nonlinear Network Flow Algorithm for Maximization of Benefits in a Hydroelectric Power System," *Operations Research*, 29(4) (1981), 763-781.

[68] Salinetti, G., "Approximations for Chance Constrained Programming," *Stochastics*, 10 (1983), 32-49.

[69] Sengupta, J.K., *Stochastic Programming: Methods and Applications*, North-Holland Publishing Company, Amsterdam, The Netherlands, (1972).

[70] Shannon, R. E., *Systems Simulation: the Art and Science*, Prentice-Hall Inc., Englewood Cliffs, N.J., (1975).

[71] Somlyody, L., and R. J.-B. Wets, "Stochastic Optimization Models for Lake Eutrophication Management," *Operations research*, 36 (1988), 660-681.

[72] Sreenivasan K. R. and S. Vedula, "Reservoir Operation for Hydropower Optimization: A Chance-Constrained Approach," *Sadhana*, Vol. 21, Part 4, August 1996, pp. 503-510.

[73] Satncu-Minasian, I.M., *Stochastic Programming with Multiple Oobjective Functions*, D. Reidel, Boston, (1984).

[74] Stedinger, J.R., "The Performance of LDR Models for Preliminary Design and Reservoir Operation," *Water Resources Research*, 20(2) (1984), 215-224.

[75] Stedinger, J.R., B.F. Sule, and D. Pei, "Multiple Reservoir Screening Models," *Water Resources Research*, 19(6) (1983), 1383-1393.

[76] Thomas, H.A, and P. Watermeyer, "Mathematical Models: A Stochastic-Sequential Approach in Design of Water Resources Systems," Harvard University Press, Cambridge, Massachussetts, (1962), 540-564.

[77] Turgeon, A., "Optimal Short-term Hydro Scheduling from the Principle of Progressive Optimality," *Water Resources Research*, 17(3), (1981), 481-486.

[78] U.S. Army Corps of Engineers, Hydrologic Engineering Center, "Optimization of Multi-Purpose Reservoir System Operations: A Review of Modeling and Analysis Approaches," Davis, California, January 1991.

[79] Wets, R. J.-B., "Solving Stochastic Programs with Simple Recourse," *Stochastics*, 10 (1983a), 219-242.

[80] Wets, R. J.-B., "Stochastic Programming Solution techniques and Approximation Schemes," *Mathematical Programming: The state of the Art*, Springer-Verlag, Berlin, New York, (1983b), 566-603.

[81] Yeh, William W-G, "Reservoir Management and Operations Models: A State-of-the-Art Review," *Water Resources Reasearch*, 21(12) (1985), 1797-1818.

[82] Ziemba, W.T., "Transforming Stochastic Dynamic Programming Problems into Nonlinear Programs," *Management Science*, 17 (1971), 450-462.

[83] Ziemba, W.T., "Stochastic Programs with Simple Recourse," *Mathematical Programming: Theory and Practice*, P. L. Hammer and G. Zoutendijk Eds., North-Holland, Amsterdam, (1974), 213-273.

[84] Zipkin, P. H., "Bounds on the Effect of Aggregating Variables in Linear Programs," *Operations Research*, 28 (1980a), 403-418.

[85] Zipkin, P. H., "Bounds for Row Aggregation in Linear Programming," *Operations Research*, 28 (1980), 903-916.

Appendices

Appendix A

Bounds on Parameters

1. α_t:

   By taking the difference, (2.26)-(2.23), fot $t = 1, \ldots, 11$, we get:

   $0 \leq F^{-1}_{Q_{t+1}}(1-\theta) - F^{-1}_{Q_t}(1-\alpha) - T_{t+1}$

   $\Longrightarrow F^{-1}_{Q_t}(1-\alpha) \leq F^{-1}_{Q_{t+1}}(1-\theta) - T_{t+1}$

   $\Longrightarrow 1-\alpha \leq F_{Q_t}[F^{-1}_{Q_{t+1}}(1-\theta) - T_{t+1}]$

   thus $\alpha \geq 1 - F_{Q_t}[F^{-1}_{Q_{t+1}}(1-\theta) - T_{t+1}], t = 1, \ldots, 11$

2. α_{12}:

   By taking the difference (2.25) - (2.23) for $t = 12$, we get:

   $0 \leq F^{-1}_{Q_{12}}(1-\rho) - F^{-1}_{Q_{12}}(1-\alpha) - S_0 + SD$

   $\Longrightarrow F^{-1}_{Q_{12}}(1-\alpha) \leq F^{-1}_{Q_{12}}(1-\rho) - S_0 + SD$

   $\Longrightarrow 1-\alpha \leq F_{Q_{12}}[F^{-1}_{Q_{12}}(1-\rho) - S_0 + SD]$

   which yields, $\alpha \geq 1 - F_{Q_{12}}[F^{-1}_{Q_{12}}(1-\rho) - S_0 + SD]$

3. $\bar{K}_t$:

   By using the fact that (2.24) $\leq$ (2.26), for$t = 1, \ldots, 11$, we get:

$F_{Q_t}^{-1}(\beta) + V_t - K \leq F_{Q_{t+1}}^{-1}(1-\theta) - SD - T_{t+1}$

therefore $K \geq F_{Q_t}^{-1}(\beta) - F_{Q_{t+1}}^{-1}(1-\theta) + SD + T_{t+1}$

For $t = 12$, we use the fact that (2.24) $\leq$ (2.25), for $t = 12$, which yields the following:

$F_{Q_{12}}^{-1}(\beta) + S_0 + V_{12} - K \leq F_{Q_{12}}^{-1}(1-\rho)$

therefore, $K \geq F_{Q_{12}}^{-1}(\beta) + S_0 + V_{12} - F_{Q_{12}}^{-1}(1-\rho)$.

4. $K_t(\alpha)$ if $\alpha \geq \alpha_t$:

   By using the fact that (2.24) $\leq$ (2.23), we get:

   $F_{Q_t}^{-1}(\beta) + V_t - K \leq F_{Q_t}^{-1}(1-\alpha) - SD$

   Therefore, $k \geq F_{Q_t}^{-1}(\beta) - F_{Q_t}^{-1}(1-\alpha) + V_t + SD$.

5. $\alpha_{\max}$ in (2.33):

   We use the fact that (2.24) $\leq$ (2.23) and subsituting K_{max} for K, we get:

   $F_{Q_t}^{-1}(\beta) + V_t - K_{max} \leq F_{Q_t}^{-1}(1-\alpha) - SD$

   $\Longrightarrow F_{Q_t}^{-1}(\beta) + V_t - K_{max} + SD \leq F_{Q_t}^{-1}(1-\alpha)$

   $\Longrightarrow F_{Q_t}[F_{Q_t}^{-1}(\beta) + V_t - K_{max} + SD] \leq 1-\alpha$

   thus $\alpha \leq 1 - F_{Q_t}[F_{Q_t}^{-1}(\beta) + V_t - K_{max} + SD]$

   Therefore, by taking the maximum value for $t = 1, \ldots, 12$, we obtain:

   $\alpha_{\max} := 1 - \max_{t=1,\ldots,12}\left\{F_{Q_t}[F_{Q_t}^{-1}(\beta) + V_t + SD - K_{\max}]\right\}$

6. $\theta_{\max}$ in (2.34):

   We use the fact that (2.24) $\leq$ (2.26) for $t = 1, \ldots, 11$, and substituting K_{max}

in (2.24), we get:

$F_{Q_t}^{-1}(\beta) + V_t - K_{max} \leq F_{Q_{t+1}}^{-1}(1-\theta) - SD - T_{t+1}$

$\implies F_{Q_{t+1}}^{-1}(1-\theta) \geq F_{Q_t}^{-1}(\beta) + V_t - K_{max} + SD + T_{t+1}$

$\implies 1-\theta \geq F_{Q_{t+1}}[F_{Q_t}^{-1}(\beta) + V_t - K_{max} + SD + T_{t+1}]$

thus, $\theta \leq 1 - F_{Q_{t+1}}[F_{Q_t}^{-1}(\beta) + V_t - K_{max} + SD + T_{t+1}]$

By taking the max over the range $t = 1, \ldots, 11$ we get:

$\theta_{\max} := 1 - \max_{t=1,\ldots,11} \left\{F_{Q_{t+1}}[F_{Q_t}^{-1}(\beta) + V_t + SD + T_{t+1} - K_{\max}]\right\}.$

7. $S_{0,min}$ in (2.37):

   From (2.27), we get:

   $S_0 \geq T_1 + SD - F_{Q_1}^{-1}(1-\theta)$, therefore

   $S_{0,min} := T_1 + SD - F_{Q_1}^{-1}(1-\theta)$

8. $S_{0,max}$ in (2.38):

   By observing that for $t = 12$, (2.24) $\leq$ (2.25), and substituing K_{max} for k in (2.24), we get:

   $F_{Q_{12}}^{-1}(\beta) + V_t + S_0 - K_{max} \leq F_{Q_{12}}^{-1}(1-\rho)$

   $\implies S_0 \leq F_{Q_{12}}^{-1}(1-\rho) - F_{Q_{12}}^{-1}(\beta) - V_t + K_{max}.$

   Therefore,

   $S_{0,\max} := F_{Q_{12}}^{-1}(1-\rho) - F_{Q_{12}}^{-1}(\beta) - V_{12} + K_{\max}$. This concludes the derivation of the expressions in chapter 2.■

Appendix B

Proof of Proposition 2.6.2

Consider the following parametrized family of LP problems $P_j(\varepsilon)$ of $(DP)_j$ in (2.29), for $j = 1, \dots, 12$:

$$
\begin{aligned}
P_j(\varepsilon): \ \min \ \ & f_j(x,\varepsilon) := D_j \\
\text{s.t.} \ \ & g_1^t(x,\varepsilon) := -\sum_{\tau=1}^{t} R_\tau + F_{Q_t}^{-1}(1-(\alpha+\Delta\alpha)) + S_0 - SD \geq 0, \ t = 1\dots,12 \\
& g_2^t(x,\varepsilon) := \sum_{\tau=1}^{t} R_\tau - F_{Q_t}^{-1}(\beta) - S_0 - V_t + (K+\Delta K) \geq 0, \ t = 1,\dots,12 \\
& g_3(x,\varepsilon) := -\sum_{\tau=1}^{12} R_\tau + F_{Q_{12}}^{-1}(1-\rho) \geq 0 \\
& g_4^t(x,\varepsilon) := -\sum_{\tau=1}^{t} R_\tau + F_{Q_t}^{-1}(1-\theta) + S_0 - SD - T_t \geq 0, \ t = 1,\dots,12 \\
& g_5(x,\varepsilon) := R_j + D_j - T_j \geq 0 \\
& g_6^t(x,\varepsilon) := R_t - T_t + D_t^*(K+\Delta K, \alpha+\Delta\alpha) \geq 0, \ t = 1,\dots,j-1 \\
& g_7^t(x,\varepsilon) := R_t \geq 0, \ t = 1,\dots,12 \\
& g_8(x,\varepsilon) := D_j \geq 0
\end{aligned}
\tag{B.1}
$$

where $x = (R_1, \dots, R_{12}, D_j)'$ and the perturbations $\varepsilon = (\Delta K, \Delta\alpha)'$. Let $x^*(\varepsilon)$ denote an optimal solution of $P_j(\varepsilon)$, where we have suppressed the explicit dependence of the solution on index j. Also denote the Lagrange multipliers associated with constraints $g_1^t, \dots, g_8$ by $u_{1t}^j(\varepsilon)$, $u_{2t}^j(\varepsilon)$, $u_3^j(\varepsilon)$, $u_{4t}^j(\varepsilon)$, $u_5^j(\varepsilon)$, $u_{6t}^j(\varepsilon)$, $u_{7t}^j(\varepsilon)$, and $u_8^j(\varepsilon)$, respectively.

Define the optimal (Lagrangian) value function by $f_j^*(\varepsilon)$ i.e.,

$$f_j^*(\varepsilon) = \mathcal{L}\left(x^*(\varepsilon), u(\varepsilon), \varepsilon\right) = f_j(x^*(\varepsilon), \varepsilon) - < u^j, g > .$$

Notice that functions f_j and g_i are continuosly differentiable w.r.t. x. Furthermore, they are (sub) differtiable in ε, since D_t^* is an optimal value function of a linear program. Furthermore, $P_j(0)$ is a linear program, ensuring the second order optimality conditions. Under the Assumption 2.6.1, thus, one can verify the regularity conditions of Theorem 3.2.2 of Fiacco (1983). Consquently, applying the sensitivity Theorem 3.4.1 of Fiacco (1983, pp.83),

$$\begin{aligned}\nabla_\varepsilon f_j^*(\varepsilon) &= \nabla_\varepsilon f_j\left(x^*(\varepsilon), \varepsilon\right) - \sum_{i=1,2,4,7} \sum_{t=1}^{12} u_{it}^j(\varepsilon) \nabla_\varepsilon g_i^t\left(x^*(\varepsilon), \varepsilon\right) \\ &\quad - \sum_{i=3,5,8} u_i^j(\varepsilon) \nabla_\varepsilon g_i\left(x^*(\varepsilon), \varepsilon\right) - \sum_{t=1}^{j-1} u_{6t}^j(\varepsilon) \nabla_\varepsilon g_6^t\left(x^*(\varepsilon), \varepsilon\right). \qquad \text{(B.2)}\end{aligned}$$

Since $\nabla_\varepsilon f_j\left(x^*(\varepsilon), \varepsilon\right) = 0$, we have

$$\begin{aligned}\nabla_\varepsilon f_j^*(\varepsilon) &= - \sum_{i=1,2,4,7} \sum_{t=1}^{12} u_{it}^j(0) \nabla_\varepsilon g_i^t\left(x^*(\varepsilon), \varepsilon\right) |_{\varepsilon=0} \\ &\quad - \sum_{i=3,5,8} u_i^j(0) \nabla_\varepsilon g_i\left(x^*(\varepsilon), \varepsilon\right) |_{\varepsilon=0} - \sum_{t=1}^{j-1} u_{6t}^j(0) \nabla_\varepsilon g_6^t\left(x^*(\varepsilon), \varepsilon\right) |_{\varepsilon=0} \text{(B.3)}\end{aligned}$$

Observing $\nabla_\varepsilon f_j^*(\varepsilon) |_{\varepsilon=0} = \left(\frac{\partial D_j^*(K,\alpha)}{\partial K}, \frac{\partial D_j^*(K,\alpha)}{\partial \alpha}\right)'$, it follows that for $j = 2, \ldots, 12$:

$$\frac{\partial D_j^*}{\partial K} = -\sum_{t=1}^{12} u_{2t}^j(0) - \sum_{t=1}^{j-1} u_{6t}^j(0) \frac{\partial D_t^*}{\partial K} \qquad \text{(B.4)}$$

and

$$\frac{\partial D_j^*}{\partial \alpha} = -\sum_{t=1}^{12} u_{1t}^j(0) \frac{\partial F_{Q_t}^{-1}(1-\alpha)}{\partial \alpha} - \sum_{t=1}^{j-1} u_{6t}^j(0) \frac{\partial D_t^*}{\partial \alpha}. \qquad \text{(B.5)}$$

For $j = 1$, we have

$$\frac{\partial D_1^*}{\partial K} = -\sum_{t=1}^{12} u_{2t}^1(0) \quad \text{and} \quad \frac{\partial D_1^*}{\partial \alpha} = -\sum_{t=1}^{12} u_{1t}^1(0) \frac{\partial F_{Q_t}^{-1}(1-\alpha)}{\partial \alpha}. \tag{B.6}$$

Thus, using (B.6) and applying (B.4) and (B.5) recursively, the expressions for $\frac{\partial D_t^*}{\partial K}$ and $\frac{\partial D_t^*}{\partial \alpha}$ are obtained.

Next, consider the following parametrized family of programs $\bar{P}(\varepsilon)$ of the model (FP) given in (2.47):

$$\begin{aligned}
\bar{P}(\varepsilon): \quad \min \quad & f(x,\varepsilon) := -L \\
\text{s.t.} \quad & h_1^t(x,\varepsilon) := -\sum_{\tau=1}^{t} R_\tau + F_{Q_t}^{-1}(1-(\alpha+\Delta\alpha)) + S_0 - SD \geq 0, \ t = 1\dots,12 \\
& h_2^t(x,\varepsilon) := \sum_{\tau=1}^{t} R_\tau - F_{Q_t}^{-1}(\beta) - S_0 - V_t + (K+\Delta K) \geq 0, \ t = 1,\dots,12 \\
& h_3(x,\varepsilon) := -\sum_{\tau=1}^{12} R_\tau + F_{Q_{12}}^{-1}(1-\rho) \geq 0 \\
& h_4^t(x,\varepsilon) := -\sum_{\tau=1}^{t} R_\tau + F_{Q_t}^{-1}(1-\theta) + S_0 - SD - T_t \geq 0, \ t = 1,\dots,12 \\
& h_5^t(x,\varepsilon) := R_t - T_t + D_t^*(K+\Delta K, \alpha+\Delta\alpha) \geq 0, \ t = 1,\dots,12 \\
& h_6^t(x,\varepsilon) := R_t \geq 0, \ t = 1,\dots,12 \\
& h_7^t(x,\varepsilon) := -L + c_t(\alpha+\Delta\alpha)R_t + \psi_t(\alpha+\Delta\alpha, R_1,\dots,R_t) \geq 0, \ t = 1,\dots,12
\end{aligned} \tag{B.7}$$

where $x = (R_1,\dots,R_{12},L)'$ and $\varepsilon = (\Delta K, \Delta\alpha)'$. Let $\bar{x}(\varepsilon)$ be the optimal solution of $\bar{P}(\varepsilon)$ and denote the associated Lagrange multipliers of $h_i^t(.)$ by $v_i^t(\varepsilon)$ for $i = 1,\dots,8$ (for $t = 1,\dots,12$). Since functions c_t and ψ_t are (twice) differentiable w.r.t. their arguments, one can ensure the differentiability requirements of the conditions in Theorem 3.2.2 of Fiacco (1983). Moreover, SOSC conditions for (FP) in the proposition, along with Assumption 2.6.1, allow us to apply Theorem 3.4.1 of Fiacco (1983) on the optimal (Lagrangian) value function $\bar{f}(\varepsilon)$, defined by

$$\bar{f}(\varepsilon) := f(\bar{x}(\varepsilon), \varepsilon) - < v, h > .$$

Then, it follows that

$$\nabla_\varepsilon \bar{f}(\varepsilon) \mid_{\varepsilon=0} = -\sum_{i=1}^{7}\sum_{t=1}^{12} v_i^t(0) \nabla_\varepsilon h_i^t(\bar{x}(\varepsilon), \varepsilon) \mid_{\varepsilon=0} . \tag{B.8}$$

Noting that $\frac{\partial h_7^t}{\partial \alpha} \mid_{\varepsilon=0} = R_t \frac{\partial c_t(\alpha)}{\partial \alpha} + \frac{\partial \psi_t(.)}{\partial \alpha}$, we get

$$\frac{\partial L^*}{\partial K} = \sum_{t=1}^{12} v_2^t(0) + \sum_{t=1}^{12} v_5^t(0) \frac{\partial D_t^*}{\partial K} \tag{B.9}$$

and

$$\frac{\partial L^*}{\partial \alpha} = \sum_{t=1}^{12} v_5^t(0) \frac{\partial D_t^*}{\partial \alpha} + \sum_{t=1}^{12} v_7^t(0) \left(R_t \frac{\partial c_t(\alpha)}{\partial \alpha} + \frac{\partial \psi_t(.)}{\partial \alpha} \right) + \sum_{t=1}^{12} v_1^t(0) \frac{\partial F_{Q_t}^{-1}(1-\alpha)}{\partial \alpha}. \tag{B.10}$$

This completes the derivation of the gradient expressions. ■

Appendix C

Sample release policies from Independent SDP Model

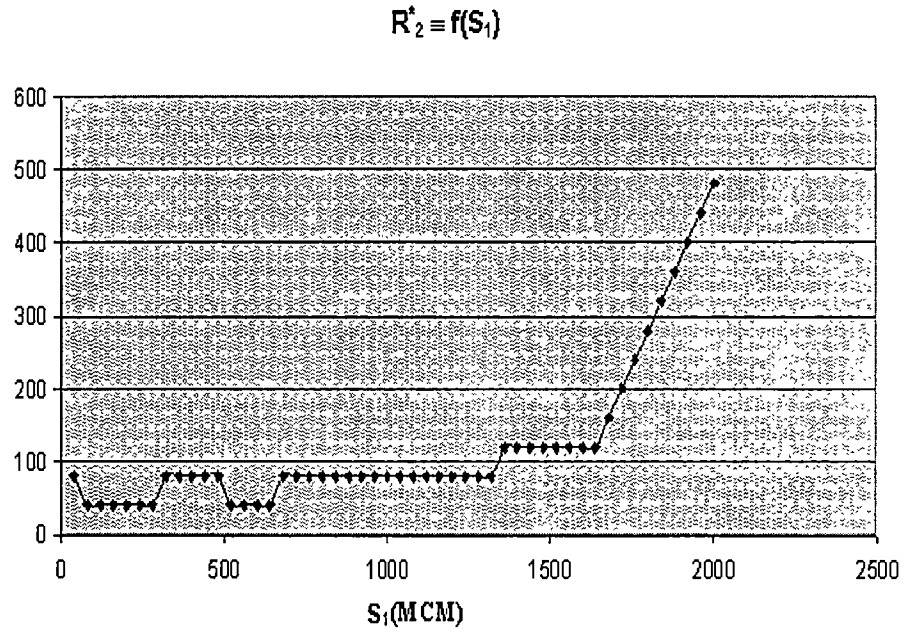

Figure C.1: Independent inflows DP Model: Optimal Release Policy for month 2

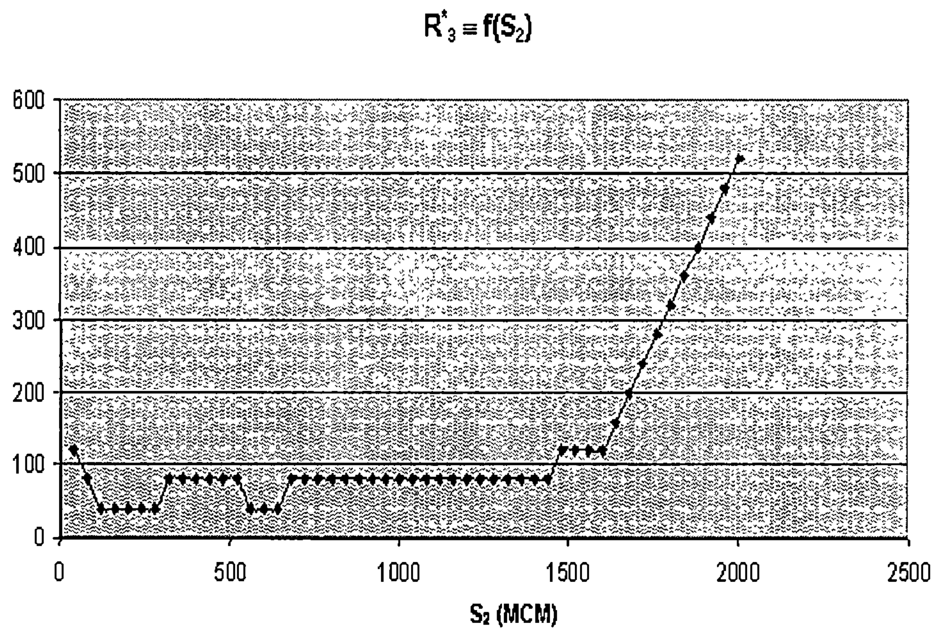

Figure C.2: Independent inflows DP Model: Optimal Release Policy for month 3

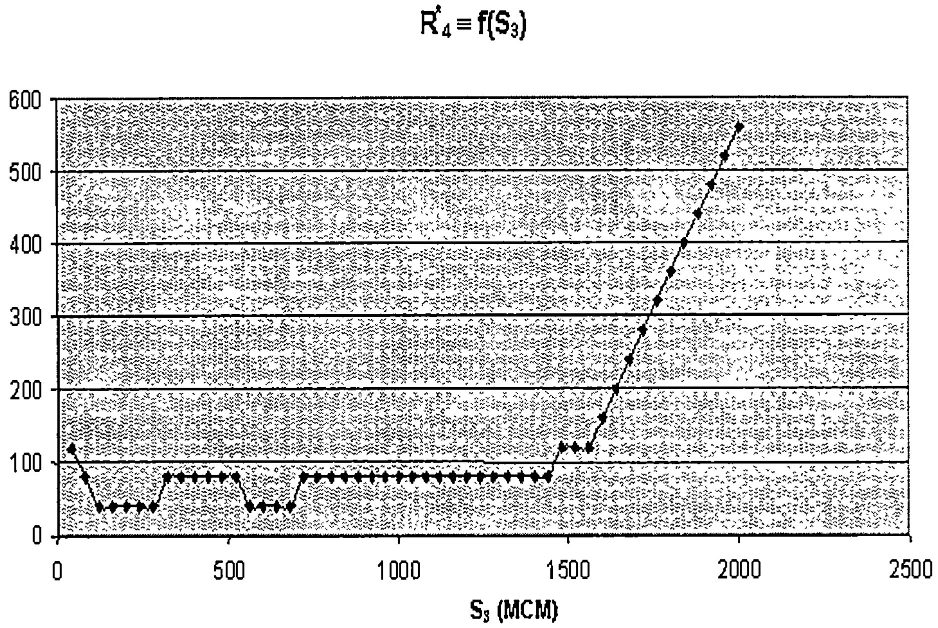

Figure C.3: Independent inflows DP Model: Optimal Release Policy for month 4

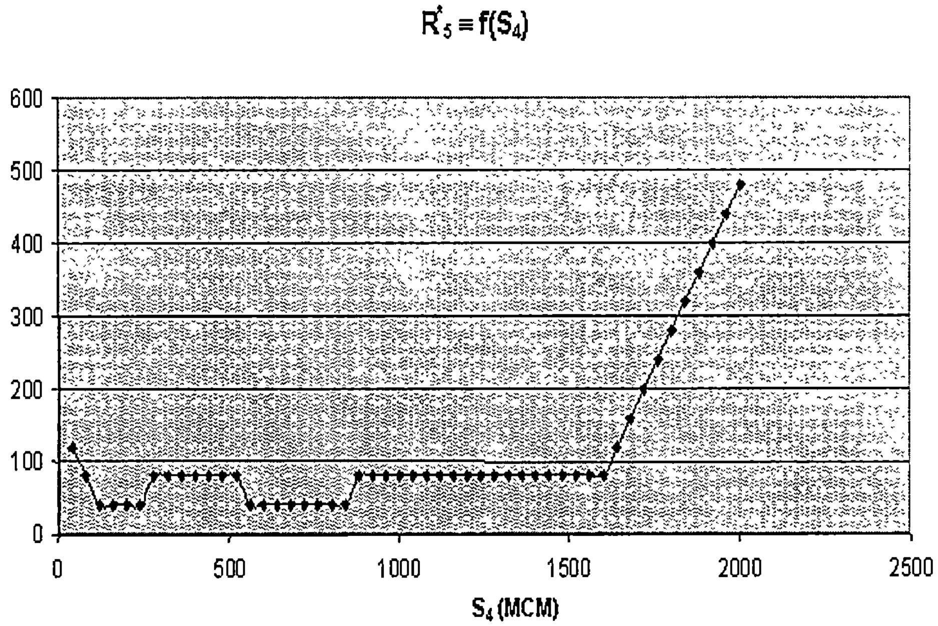

Figure C.4: Independent inflows DP Model: Optimal Release Policy for month 5

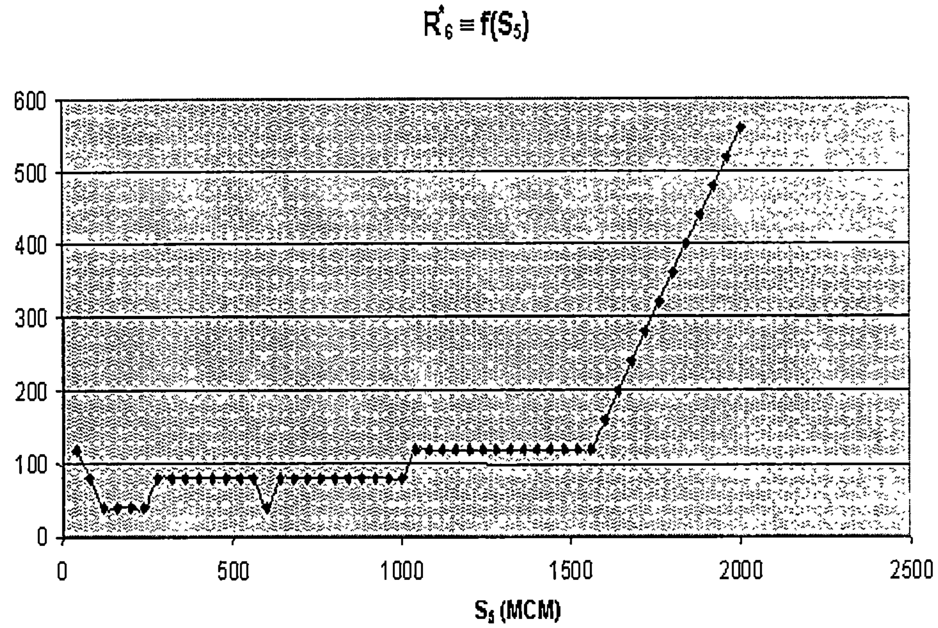

Figure C.5: Independent inflows DP Model: Optimal Release Policy for month 6

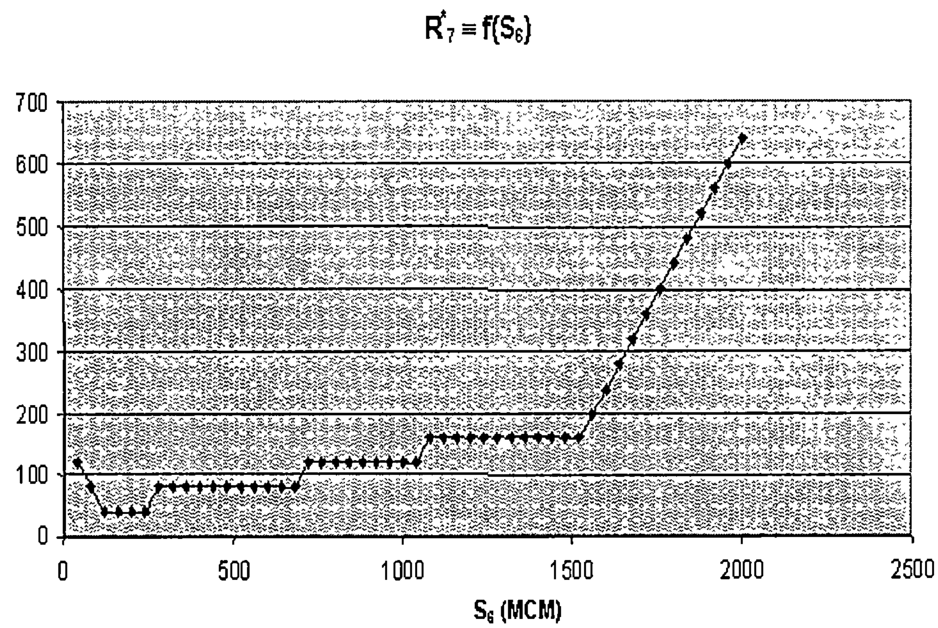

Figure C.6: Independent inflows DP Model: Optimal Release Policy for month 7

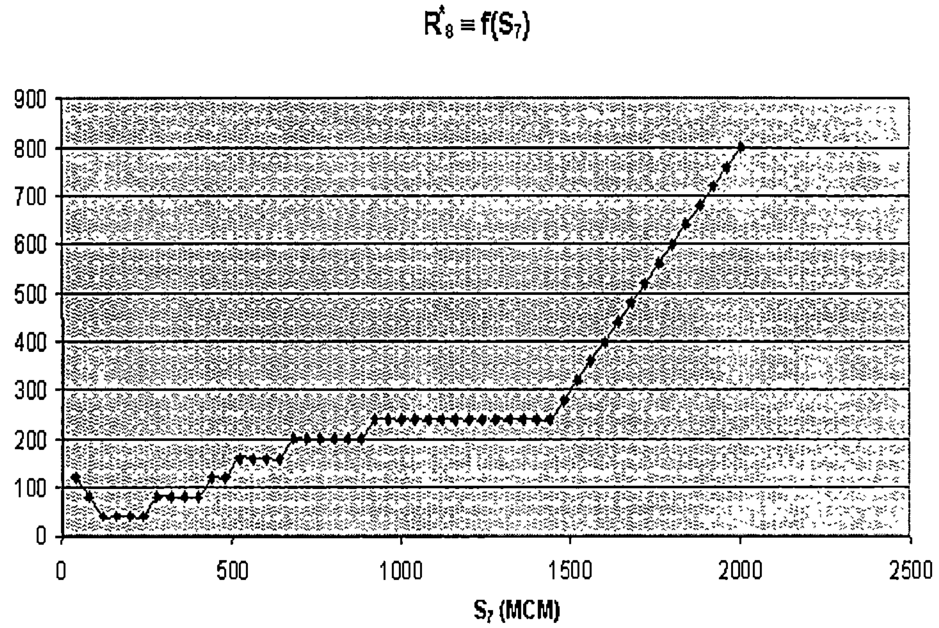

Figure C.7: Independent inflows DP Model: Optimal Release Policy for month 8

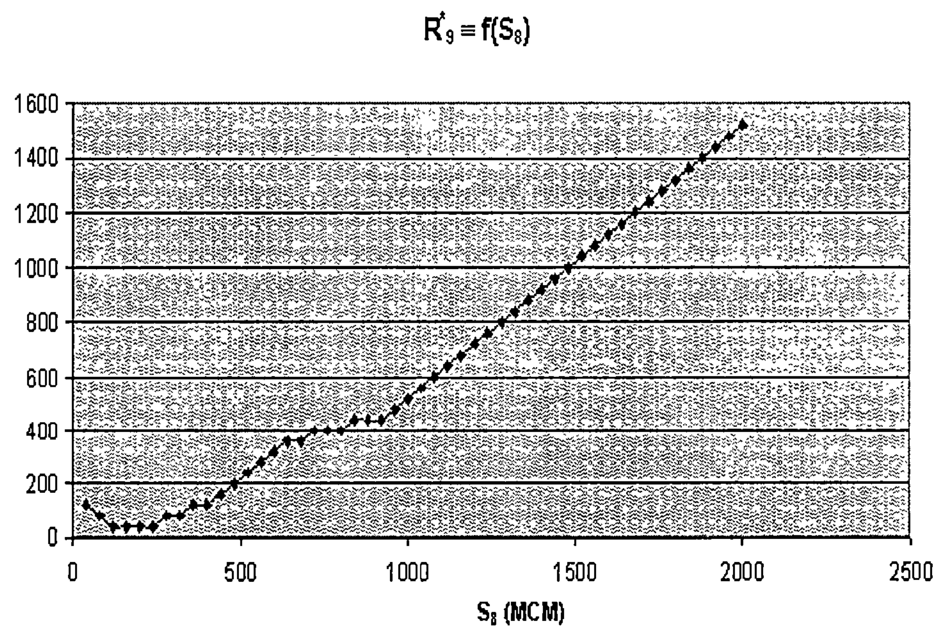

Figure C.8: Independent inflows DP Model: Optimal Release Policy for month 9

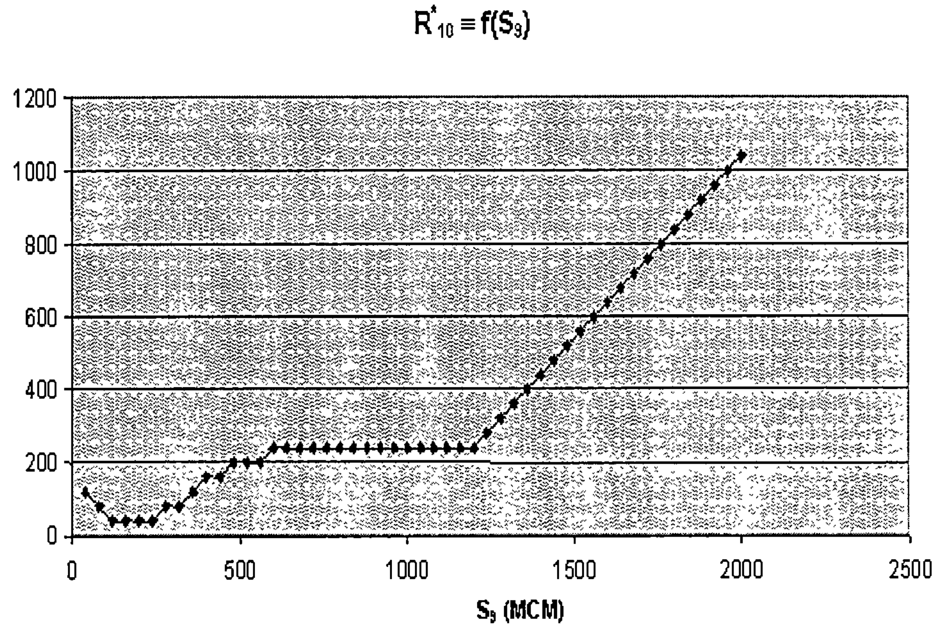

Figure C.9: Independent inflows DP Model: Optimal Release Policy for month 10

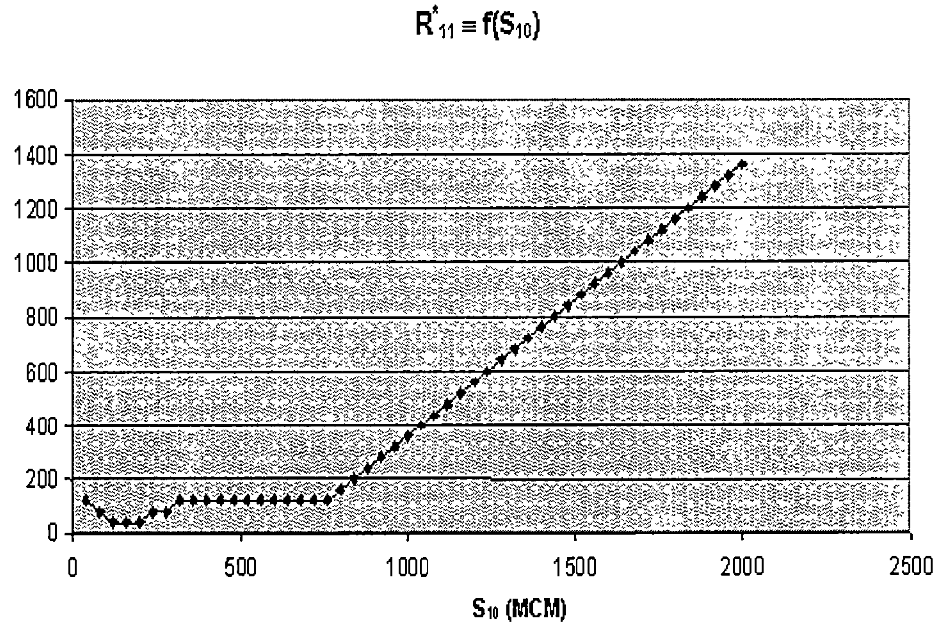

Figure C.10: Independent inflows DP Model: Optimal Release Policy for month 11

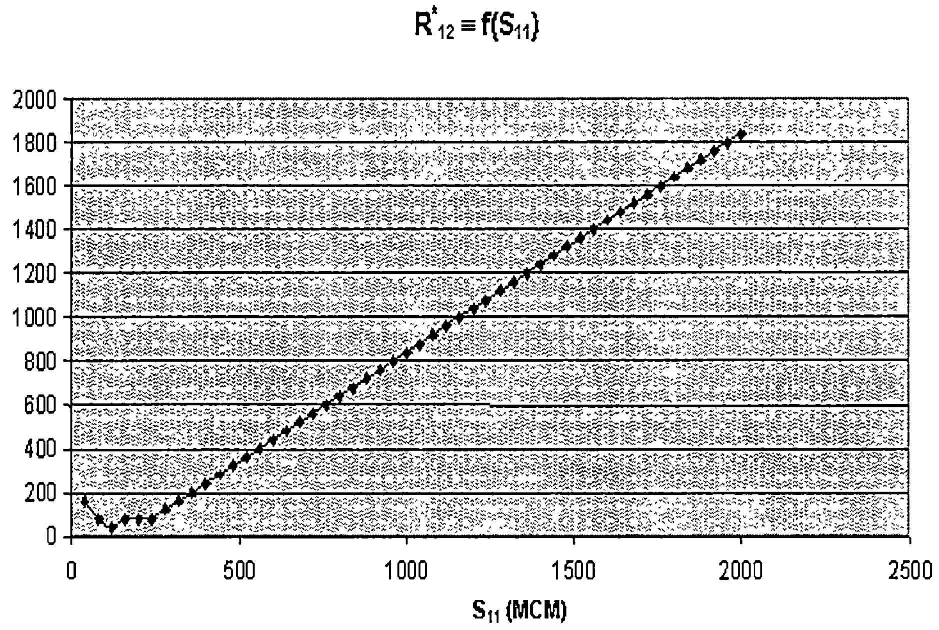

Figure C.11: Independent inflows DP Model: Optimal Release Policy for month 12

Appendix D

Sample optimal release policies from Restricted SDP Model

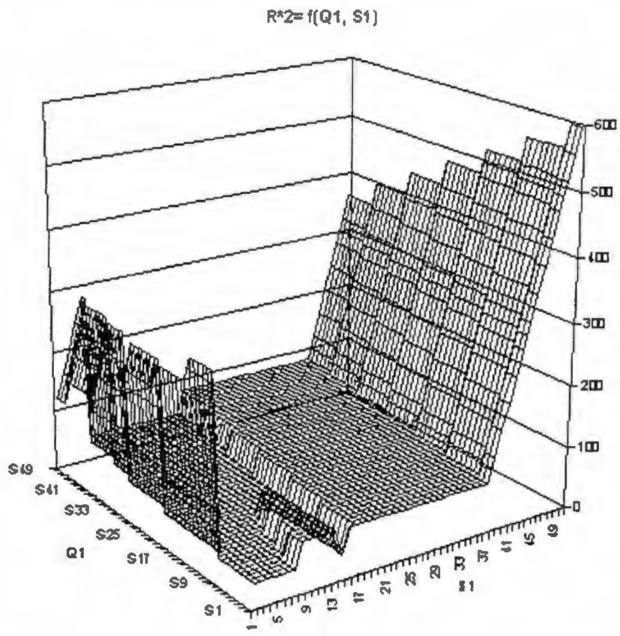

Figure D.1: Restrictive DP Model: Optimal Release Policy for month 2

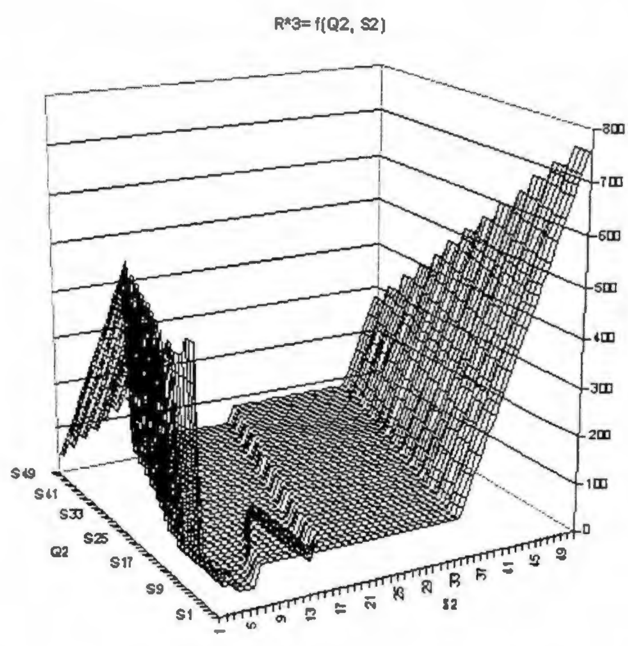

Figure D.2: Restrictive DP Model: Optimal Release Policy for month 3

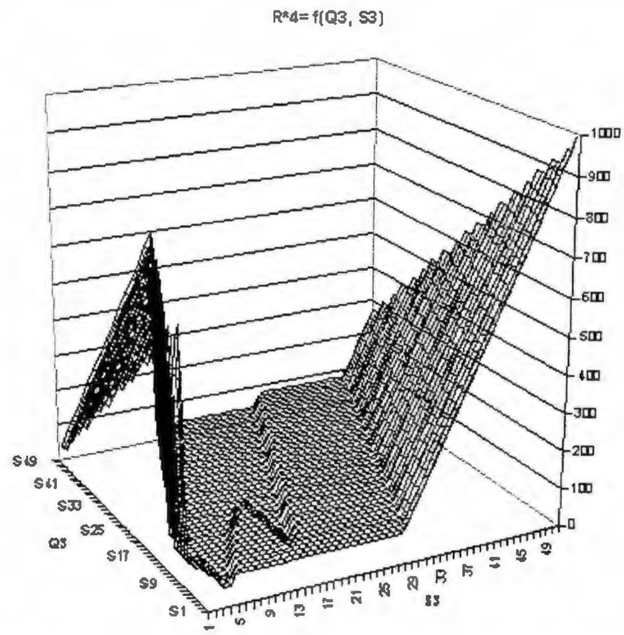

Figure D.3: Restrictive DP Model: Optimal Release Policy for month 4

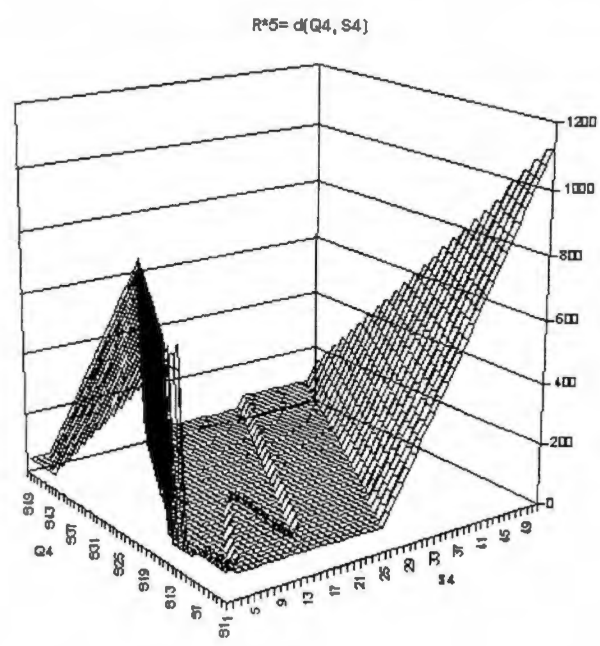

Figure D.4: Restrictive DP Model: Optimal Release Policy for month 5

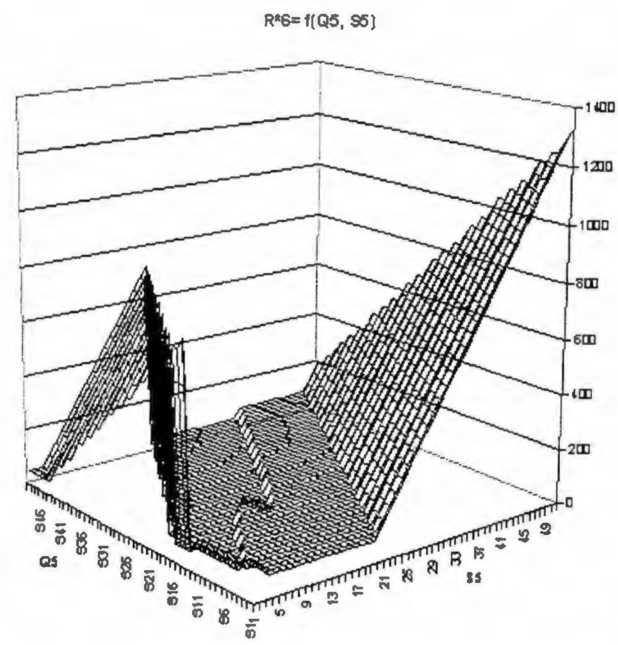

Figure D.5: Restrictive DP Model: Optimal Release Policy for month 6

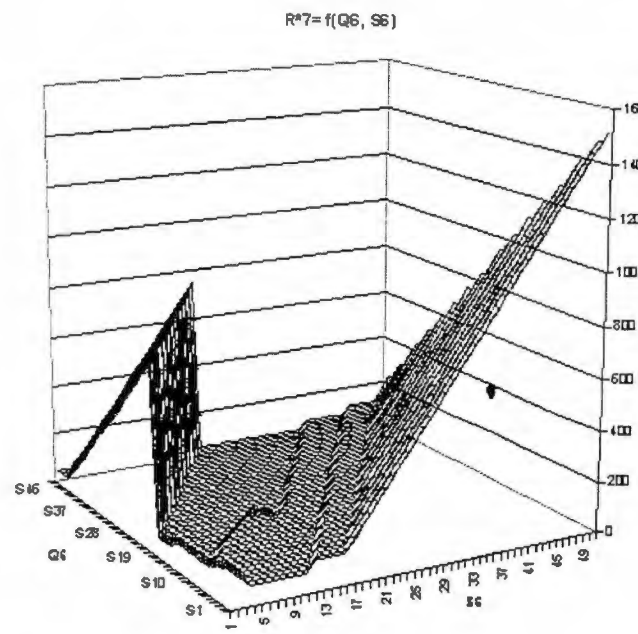

Figure D.6: Restrictive DP Model: Optimal Release Policy for month 7

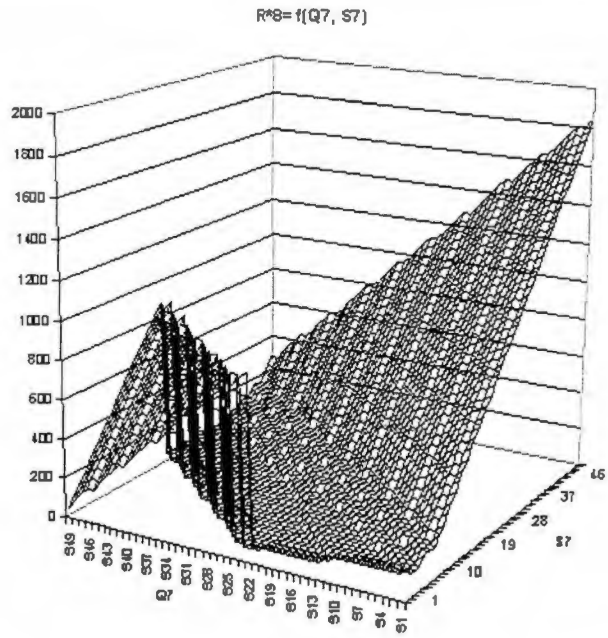

Figure D.7: Restrictive DP Model: Optimal Release Policy for month 8

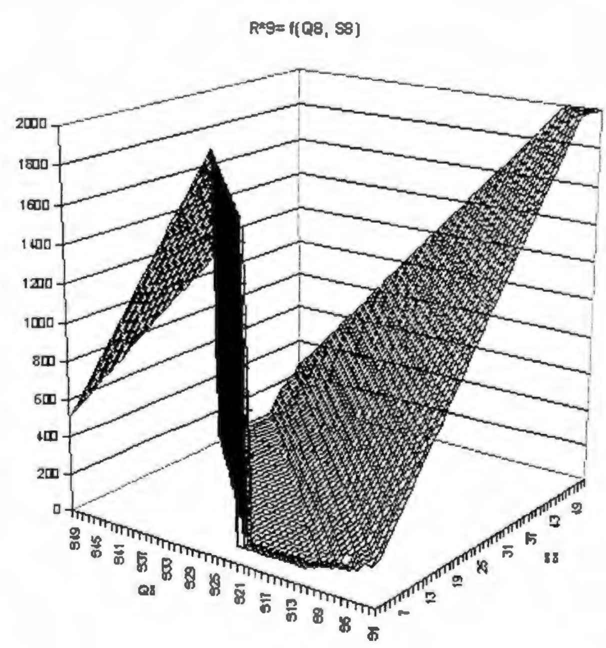

Figure D.8: Restrictive DP Model: Optimal Release Policy for month 9

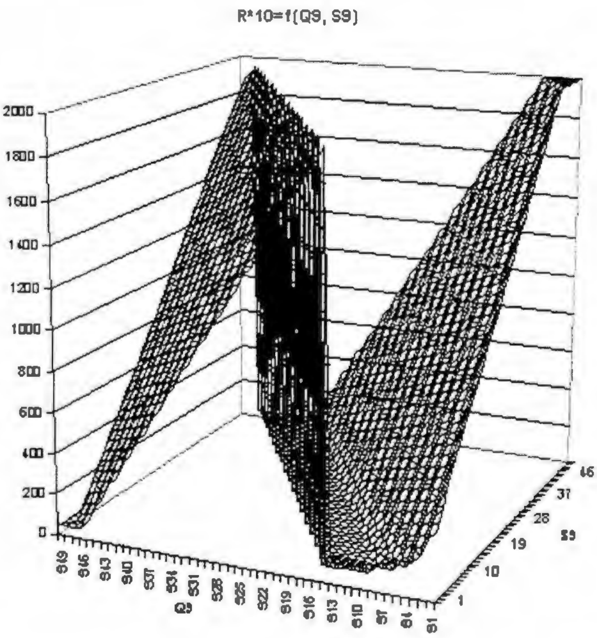

Figure D.9: Restrictive DP Model: Optimal Release Policy for month 10

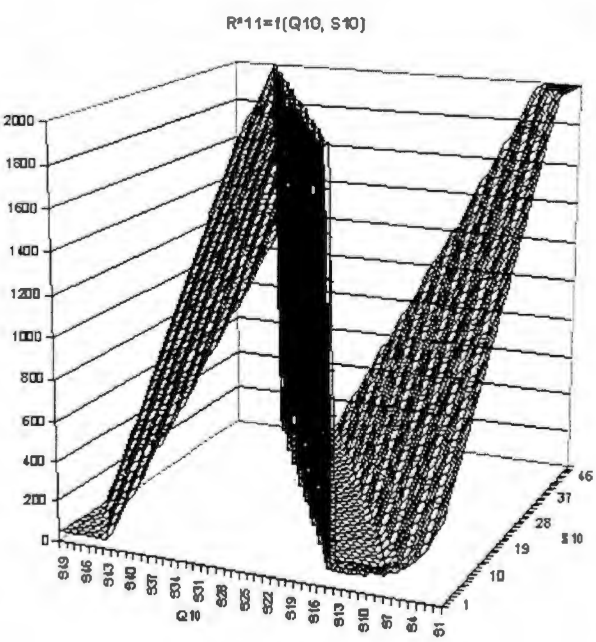

Figure D.10: Restrictive DP Model: Optimal Release Policy for month 11

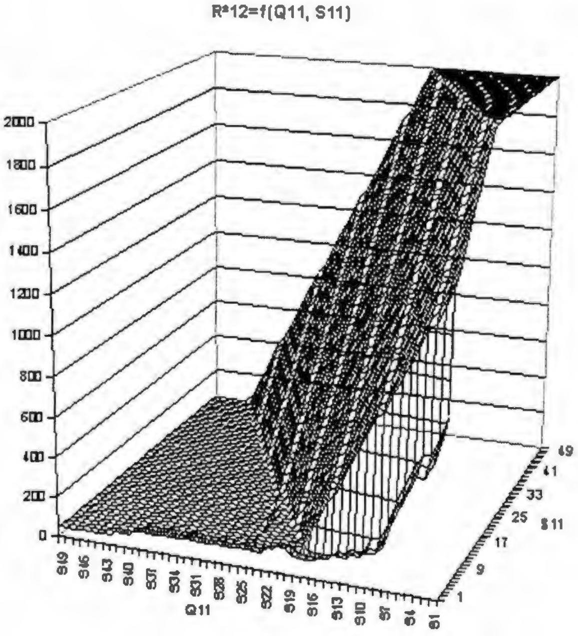

Figure D.11: Restrictive DP Model: Optimal Release Policy for month 12

Vita

The author was born in the North African Nation of Tunisia on September 1^{st}, 1971. He completed his elementary and high school studies in Tunisia, then moved to the United states to complete his University studies. He earned a Bachelors of science in Business management from the University of Tennessee, Knoxville, USA, with a minor in mathematics. He then earned a Masters of Science in Operations Research from the same institution, with a graduate minor in Satistics. This thesis is his last requirement to earn a Ph.D. degree in Operations Research. The author has one paper titled "Capacity planning Model for a Multi-purpose Water reservoir With Target-Priority Operation" to appear in *Annals of Operations research.* A second paper with the title "Multi-stage stochstic Model with Recourse for the Capacity Planning of a Multi-purpose water Reservoir with Target-Priority Operation" is currently a working paper.