APPLICATION OF PRINCIPAL COMPONENTS ANALYSIS TO LONG-TERM RESERVOIR MANAGEMENT

by

Maarouf Saad, B.SC., M.SC.A. (École Polytechnique de Montréal)

A thesis submitted to the Faculty of Graduate Studies and Research

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Electrical Engineering,

McGill University,

Montréal, Canada.

July 1988

APPLICATION OF PRINCIPAL COMPONENTS ANALYSIS

TO LONG-TERM RESERVOIR MANAGEMENT

Maarouf Saad, B.SC., M.SC.A. (École Polytechnique de Montréal)

ABSTRACT

Determining the optimal long-term operating policy of a multi-reservoir power system requires solution of a stochastic nonlinear programming problem. For small systems, the solution can be found by dynamic programming, but for large systems no direct solution method exists yet, so that one must resort to mathematical manipulations to solve the problem. This thesis presents a very efficient procedure for the case where high correlation exists between the state variables. It consists in performing principal components analysis on the trajectories to find a reduced model of the system. The reduced model is then substituted into the operating problem and the resulting problem is solved by stochastic dynamic programming. The reservoir trajectories on which principal components analysis are performed can be obtained by solving the operating problem deterministically for a large number of equally likely flow sequences. The results of applying the manipulation to Québec's La Grande river, which has four reservoirs, are reported. A comparison with the classical dynamic programming, that is without any reduction, is also studied and results are reported to show the efficiency of the principal components approach.

RESUME

Déterminer la règle optimale de gestion à long-terme d'un réseau hydroélectrique de grande taille revient à résoudre un problème d'optimisation stochastique nonlinéaire. Pour des systèmes de petite taille, ce problème peut être facilement résolu par la programmation dynamique, ce qui n'est pas le cas pour des systèmes de grande taille.

Dans cette thèse, une nouvelle approche est proposée pour les systèmes dont les états sont corrélés. Cette approche est basée sur l'analyse en composantes principales sur les états du système dans le but d'établir un modèle réduit. Cette réduction rend le problème résolvable par la programmation dynamique stochastique. Les états auxquels cette technique est appliquée sont obtenus à partir de la solution du problème déterministe appliqué à un grand nombre de séquences équiprobables d'apports naturels. Les résultats de cette approche seront illustrés pour les installations de la rivière La Grande dans la province de Québec. Une comparaison des résultats obtenus avec cette approche avec ceux obtenus sans aucune réduction est aussi faite pour illustrer l'efficacité de la méthode proposée.

ACKNOWLEDGEMENTS

The author is deeply indebted to Professor A. Turgeon for suggesting this study and for his guidance, continual encouragement and support of this work.

Grateful acknowledgement must go to Mr. A. Desrochers and Mr. H. Hamidizadeh. Their comments and suggestions are particularly appreciated.

Thanks are also due to the Ecole de technologie supérieure, IREQ and McGill University for providing the facilities.

Finally, I wish to thank my wife Marlen, my mother Zarifa and my father Abdul-Latif for their constant encouragement and support.

CLAIM OF ORIGINALITY

To the best of the author's knowledge, the following contributions are original:

1. The combination of the existing implicit and explicit stochastic methods. The implicit approach is used to solve deterministically the problem for a large number of flow sequences and the explicit approach to determine the optimal operating policy.

2. The application of principal components analysis and stochastic dynamic programming to long-term reservoir management. Principal components analysis is used to reduce the number of state variables in the problem and stochastic dynamic programming to find the optimal solution of the reduced problem.

3. The solution of large-scale problems taking into account the stochastic nature of river flows described by a discrete distribution.

4. The determination of a global feedback solution contrary to the decomposition and projection methods, which means that global constraints, such as satisfying the demand, will be met.

5. The determination of optimal rules taking into account all local constraints on the discharge and content of the reservoirs, so that the solution obtained is generally feasible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

Principal Symbols

C(He)	production cost
COV	covariance
de	demand of electric energy
E	expected value
h	water head
H(X,U)	production of hydroelectric energy
He	production of thermal energy
$\bar{He}$	capacity of the thermal plants
min F	minimum of function F
max F	maximum of function F
P	eigenvectors matrix of the inflows
Q	eigenvectors matrix of the contents
r(h,U)	plant efficiency
U	discharge
$\bar{U}$	capacity of the plants
V	amount of water spilled
VAR	variance
W	covariance matrix
X	content of the reservoir
$\underline{X}$, $\bar{X}$	lower and upper bounds of the reservoirs
Y	total natural inflow
z	penalty function

α	penalty cost
Φ	principal components related to the inflows
λ	Lagrange multiplier
Ψ	principal components related to the contents
$\underline{\Psi}$, $\bar{\Psi}$	lower and upper bounds of the principal components of the contents
μ	mean value
Ω	total outflows (U + V)
ρ	generation characteristics
e^T	transposed 1 by n unity vector =[1,...,1]
$\lvert\cdot\rvert$	absolute value

Subscripts

i=1,...,n	indexes the number of reservoirs
j=1,...,J	indexes the number probability branches (classes)
k=1,...,K	indexes the periods
m=1,...,M	indexes the line segments of the grid of the piecewise linearization

Superscripts

~	components selected
$\wedge$	optimal value
T	transpose
-1	inverse

CHAPTER I

INTRODUCTION

1.1 General Requirements

Electricity plays an important role in all modern societies. Over the years, power systems have been expanded to meet the growing demand for electrical energy. These systems rely on two important sources of energy:

a) The potential energy of fuels such as oil and coal. These types of energy are converted into electrical energy by conventional thermal power stations.

b) The potential energy of water, converted into electrical energy by hydro-electric plants.

The generation of electricity is a complex problem due to the following facts.

a) Electricity cannot be produced in advance and stored for future use.

b) Electricity cannot be produced with delay.

c) Under normal conditions, the demand must be satisfied.

In other words, if during the period of time k, the demand is de(k), the production must be also, de(k). Generally, the natural inflows are very low when the demand is very high as shown in Figure 1.1. Thus to

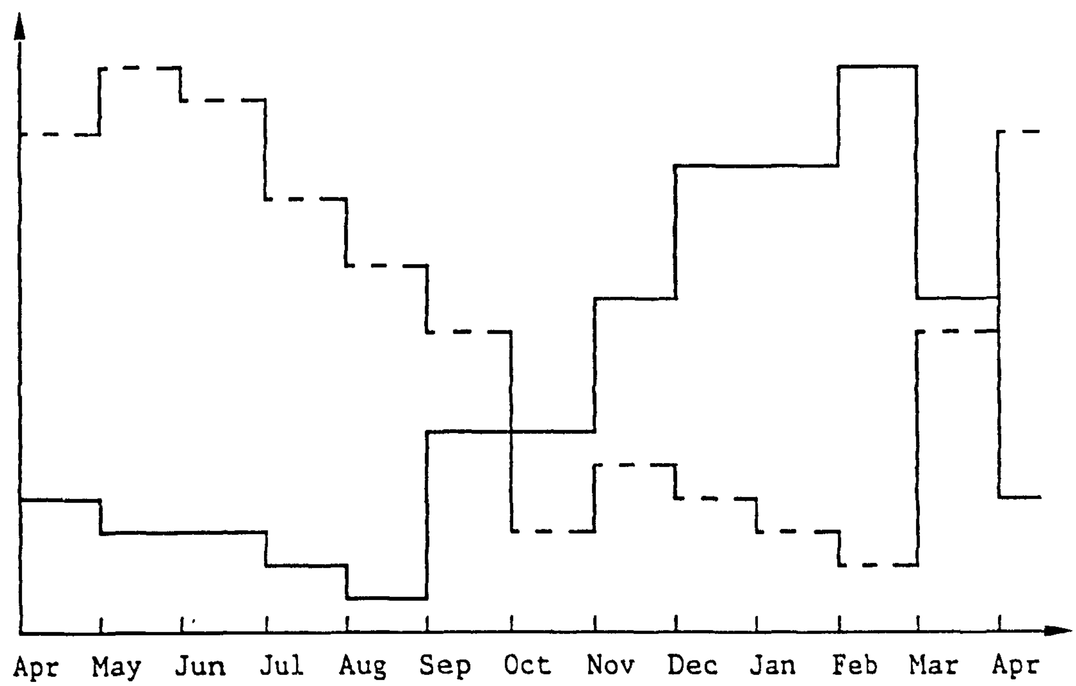


Figure 1.1 Demand and inflows pattern

meet demand at the lowest production cost, it becomes necessary to accumulate the excess water during wet periods into natural or artificial reservoirs. This water can be released during dry periods to produce electricity.

Usually, many reservoirs in series and/or in parallel form a hydro power system. They can be classified as reservoirs with daily, seasonal or yearly operation cycles. The reservoirs with daily cycles are used to store the water in off-peak periods during the night and used it in peak periods during the day. The seasonal reservoirs store the water surplus from spring for winter, and the reservoirs with yearly operation cycles, like Manic V in Québec, are used to save the excess water of rainy years for use in dry years.

The problem of determining the optimal operation of a multireservoir system is usually broken down into a deterministic short-term operating problem and a stochastic long-term operating problem. The long-term problem is a stochastic process since it is impossible to make exact prediction on the natural inflows. The stochasticity of the natural inflows plays an important role in the scheduling problem. For example, if the stocks of hydro-electric energy are depleted and low inflow volumes occur, it may be necessary to use expensive thermal generation in the future. On the other hand, if the reservoir levels are kept very high through a more intensive use of thermal generation, while high inflow volumes occur, there may be spillage or waste of energy in the system, which in turns increases operation costs.

Although the present thesis is mainly concerned with the operation of the stochastic long-term multireservoir power systems, this introductory chapter will continue further by describing the thermal and hydro power systems in Section 1.2 and then the power system optimization problems in Section 1.3. In Section 1.4 the solution methods for the stochastic long-term reservoir problem are described, whereas in Section 1.5, the special approaches to the large-scale problem are dealt with. The scope and contributions of this work are considered in Section 1.6, and finally Section 1.7 contains the outline of this thesis.

1.2 Hydrothermal Power System

A mixed hydrothermal power system, as illustrated in Figure 1.2, involves the coordination of thermal and hydroelectric energy to meet all system loads over a given planning horizon. Naturally given the expenses involved in the operation of the system, a great deal of effort is devoted to ensure that each subsystem (thermal and hydraulic) is as efficient and reliable as possible. In recent years the increasing demand of energy has focused attention on the problem of ensuring that maximum benefit is derived from the resources and equipment available. For this reason, in the next two subsections, the resources and equipment of thermal and hydro subsystems are briefly described.

1.2.1 Thermal Plants

Thermal plants can be operated by Gas, coal or oil. Gas turbines have very low start up costs but, on the other hand, they have high operating costs. As a result, gas turbines are suitable for short period operations to meet the demand during the peak periods. The thermal plants operated by coal or oil are known as "fossil fuel condensing plants" [Habibollahzadeh, 1984]. These plants have three different kinds

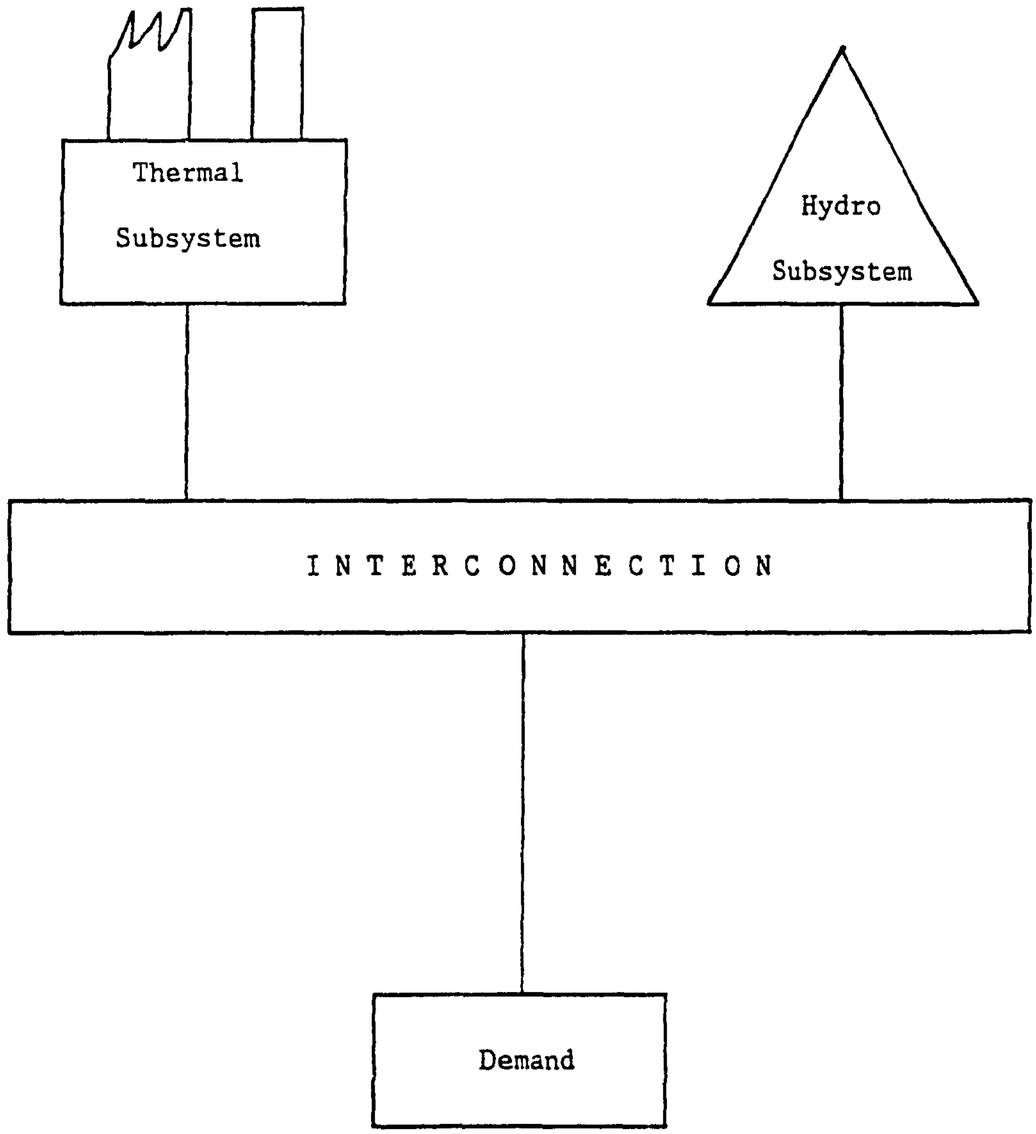


Figure 1.2 Hydrothermal power system

of production cost: start up, fixed and variable operation costs. The cost of starting a fossil fuel plant is very high especially if the boiler has been cooling [Turgeon, 1977]. Once this type of plants is started, it will have a fixed operation cost corresponding to its minimum output power. The variable operation cost of the plant increases as the power production is increased from the minimum to the maximum output level. According to an internal report of Hydro Québec [Report Hydro Québec, 1983], the various sources of thermal energy have also different operation costs, as shown in Table 1.1. In that table, the investment cost is also given. It can be seen that the investment cost increases as the operation cost decreases for different types of plants.

Type	Investment ($/KW)	Operation ($/KWh)
Gas Turbines	450	0.095
Oil Plant	1000	0.050
Coal Plant	1300	0.030

Table 1.1 Investment and operation costs
for different thermal plants

It should be mentioned that due to both scheduled and unscheduled maintenance (or maintenance outages), a thermal plant can only be used 80% of the time at the most [Report Hydro Québec 1983].

For the remainder of this thesis, the thermal type will refer to plants operated by gas, oil, coal or any combination of these fuels. For the sake of presentation, coal will be considered the energy source for thermal plants. Throughout the thesis, we will consider only the operation or production costs of thermal plants and will also assume that all the thermal plants under consideration are already in place and there will be no further investment in thermal generation.

1.2.2 Hydro Generation Plants

The potential energy available from rivers is converted into electric energy by hydro electric plants. A hydroplant installed on a river, as shown in Figure 1.3, consists of a reservoir, a dam and a plant. In general, the hydroplants are divided into three groups [EPRI EL-1659, 1981]:

a) Plants with small or no water storage, also called "run-of-river" plants.

b) Plants with moderate storage, usually used for short-term operation.

c) Plants with large storage, usually used for long-term operation.

For "run-of-river" plants, no reservoir exists, so that all the incoming water is used to produce electrical energy, whereas, for moderate and large storage plants, a reservoir is used to store water surplus to meet future requirements. As a matter of fact, when more water is stored in the reservoir the head increases, so that production of electricity increases also. In other words, the hydroelectric energy is a function of the gross head, that is the difference between the elevation of the surface of the reservoir and the elevation of the

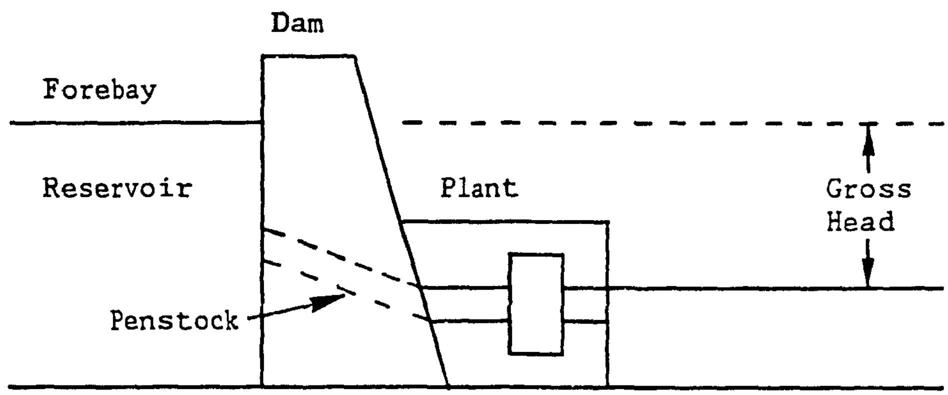


Figure 1.3 Hydro plant components

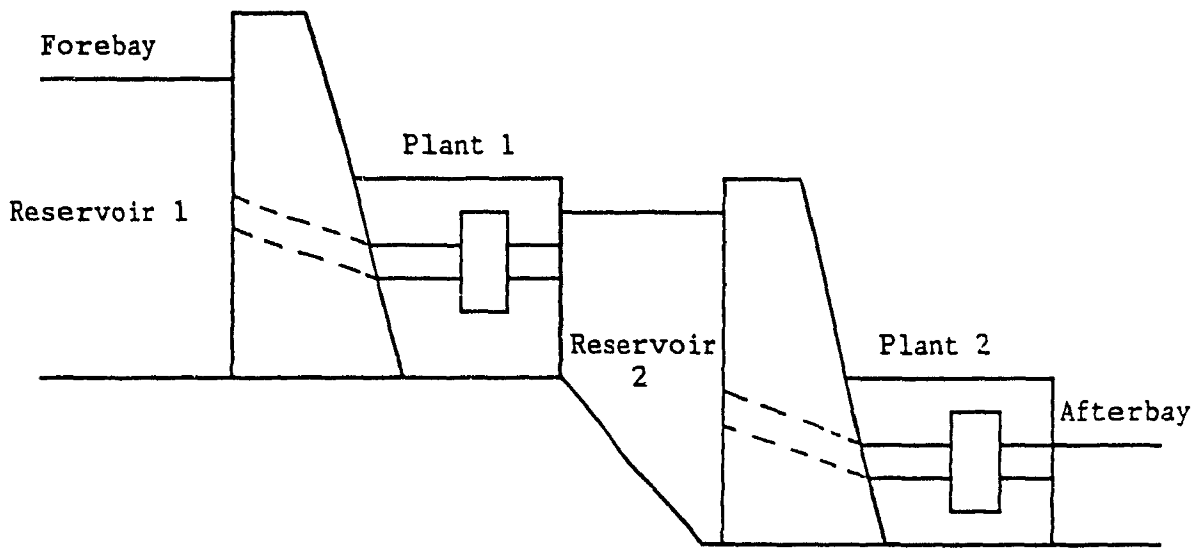


Figure 1.4 Two cascade hydro plants

afterbay, or downstream water level below the hydroelectric plant (Figure 1.4) [Wood and Wollenberg, 1984]. Therefore the potential energy of the water available for conversion into electrical energy is a function of both the turbined water and the gross head. In fact, the head at the turbine itself is slightly less than the gross head due to friction losses. That is why we define the net head which is equal to the gross head less the flow losses (measured in the same units as the gross head) [Wood and Wollenberg, 1984]. Figure 1.5 shows a typical curve where the hydraulic head is constant. This graph shows that the generation is a nonlinear function of the released water. Figure 1.6 shows the same nonlinear characteristic but for variable head. This type of curve is obtained whenever the variation in the forebay and/or the afterbay elevation (Figure 1.3) is a fairly large percentage of the overall hydraulic head.

Scheduling hydroelectric plants with variable characteristics is more difficult than scheduling hydroelectric plants with fixed heads. This is true not only because of the multiplicity of the characteristic curves that must be considered, but also because the maximum capability of the plant will also tend to vary with the hydraulic head [El- Hawary and Christensen, 1979; Wood and Wollenberg, 1984]. However head variation is not a major feature of the long-term reservoirs problem with which this thesis is concerned [Hanscom, 1976; Read, 1979].

For cascaded multireservoir systems, the hydroplants are coupled because the discharge from one reservoir constitutes part of the inflow to the next. Of course, these reservoirs have to be operated according to certain rules. These rules define the amount of energy produced by each plant along the river. Figure 1.7 illustrates the characteristics

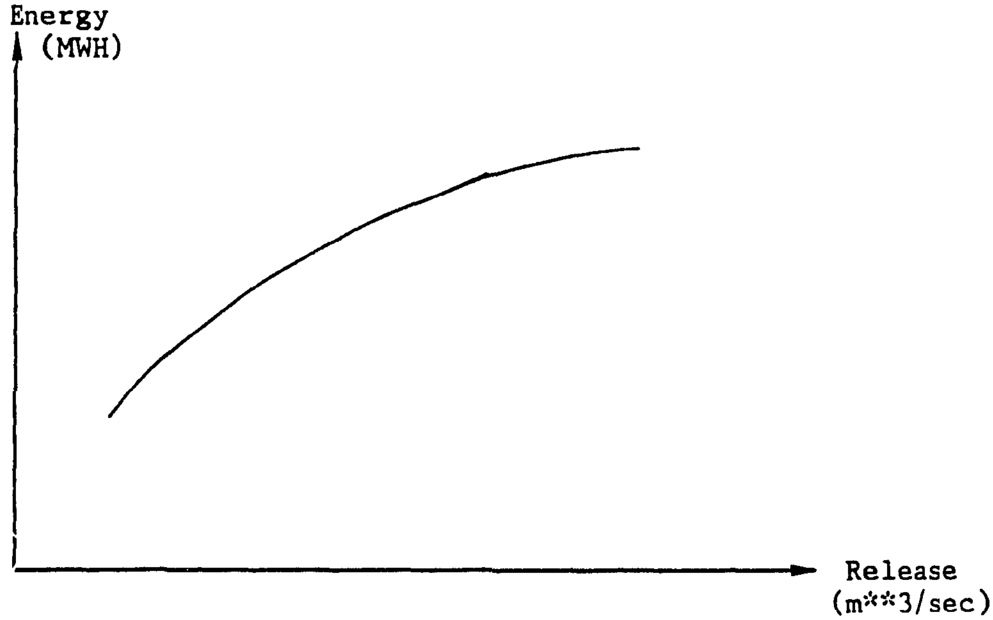


Figure 1.5 Hydro generation characteristic with constant head

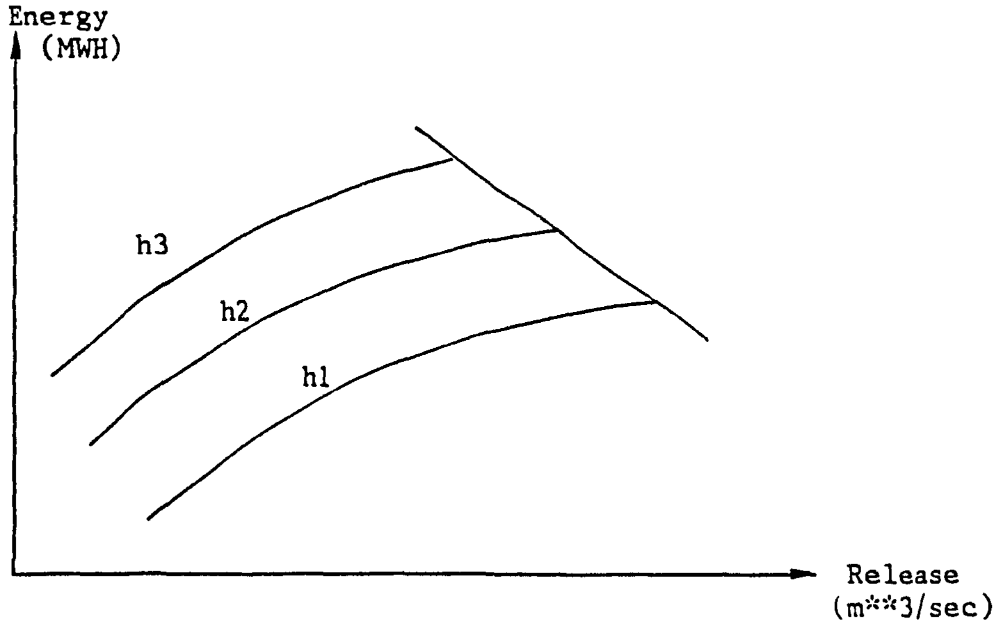


Figure 1.6 Hydro generation characteristic with variable head hi (h1< h2 < h3)

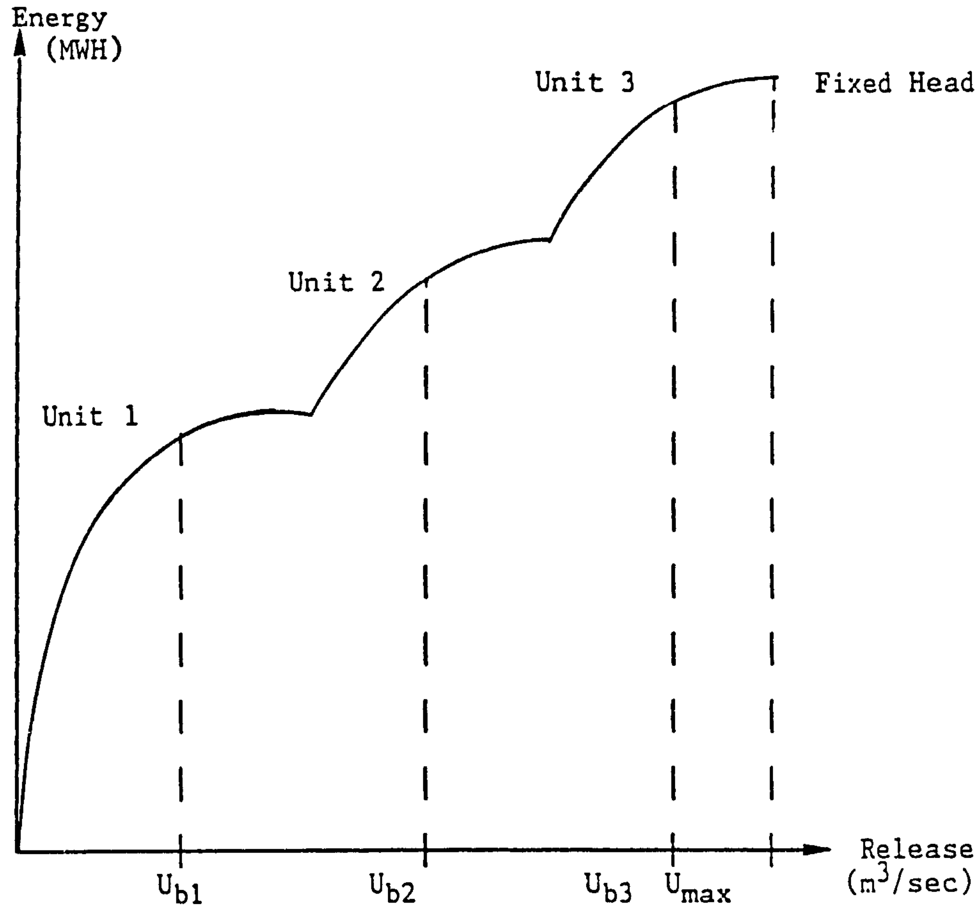


Figure 1.7 Hydro generation for multireservoir system at a fixed head

of such system for a given head. The curve consists of three segments which correspond to 1, 2 or 3 units in operation. The best operation point (the best efficiency points) for 1, 2 or 3 units in operation are also shown on the graph by U_{b1}, U_{b2} and U_{b3} respectively.

1.3 Power System Optimization Problems

The hydrothermal scheduling problem is usually decomposed into an energy problem and a power problem or into what Massé [1946] and Turgeon [1981b] call 'strategic' and 'tactical' problems, respectively. The first is related to the management of the available water according to the forecast of the natural inflows. The second problem, on the other hand, involves the hour by hour control and coordination of all generation units to satisfy a given demand with known inflows. Then it can be seen that the energy or strategic problem is stochastic and concerns the long-term scheduling, while the power or tactical problem is deterministic and concerns the short-term scheduling.

Although this thesis mainly deals with the energy or the long-term problem, a brief description of the nature of each problem is given in the next two subsections.

1.3.1 The Long-Term Scheduling Problem

In the long-term scheduling problem, the operating horizon spans from one to several years and is divided into weekly or monthly intervals. The objective is to determine a generation schedule that minimizes the expected cost during the planning horizon while meeting the demand at all times. In other words, the long-term problem consists in finding the optimal balance between the production of various plants for each of the intervals of the planning horizon. This problem is further complicated by the fact that the exact level of inflows in the future cannot be predicted. In this case, historic data and probability distribution can be considered due to the small variation of inflows from year to year. Then after having incorporated this probability distribution into the optimization model, the most successful solution technique has proven to be 'stochastic dynamic programming' [Pronovost and Boulva, 1978; Sachdeva, 1982; Turgeon 1980].

Since the time steps considered are weeks or months, the time delay between reservoirs can be neglected and sometimes reservoirs on the same stream can even be aggregated.

The long-term scheduling optimization involves other stochastic variables such as load and unit availabilities (thermal and hydro units). However, dealing with the random nature of these variables is outside the scope of this thesis. Only the natural water inflows will be considered as random.

1.3.2 The Short-Term Scheduling problem

The short-term scheduling problem distributes over the week or over the month the total discharge selected by the long-term problem for that period. This distribution is performed so that the total system production cost is minimized within the limits permitted by the hydraulic and thermal constraints. In such a scheduling problem, the load, hydraulic inflows and unit availabilities are assumed to be known. A set of starting conditions (e.g. reservoir levels) is given, and parts of the hydraulic constraints may involve meeting "end-point" conditions at the end of the scheduling interval so as to conform with a long-term water release schedule previously established as shown in Figure 1.8.

Due to the time needed to start the thermal units, some difficulties may be introduced, creating therefore, a need for "spinning reserve" to cover possible break-downs. Also if several hydro plants are located on the same stream, the time taken by the water to travel from one plant to the next may be of great importance.

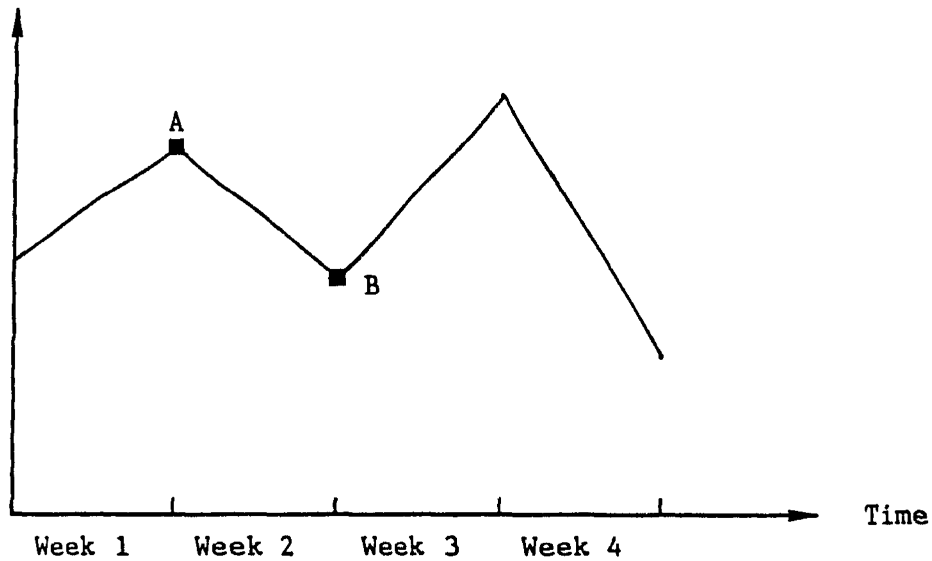
Note that the short-term problem must be re-optimized frequently to take into account the variations in the expected demand patterns, the natural inflows or the availability of the equipments.

For a thorough survey of this subject, the reader is referred to Chapter VII of Wood and Wollenberg [1984].

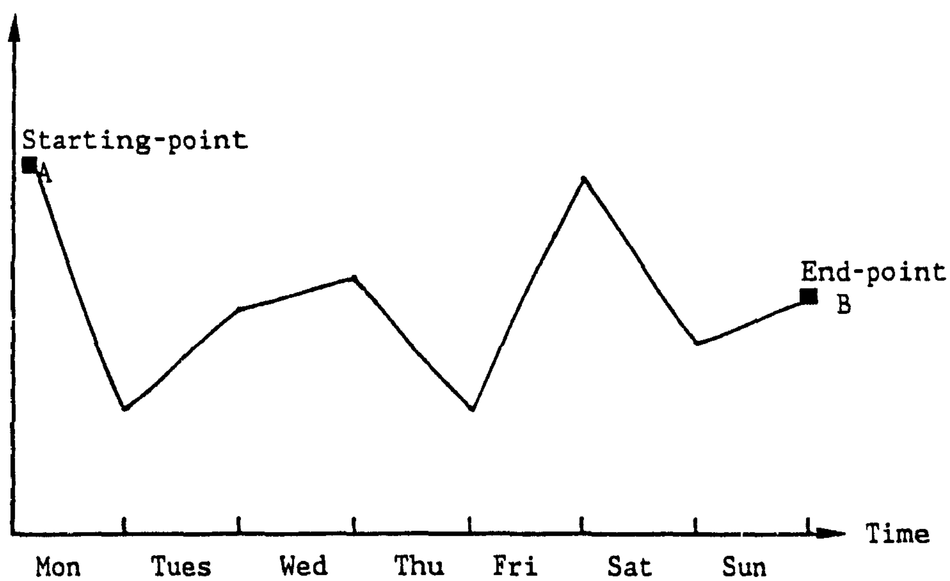
1.4 Solution Methods For the Stochastic Hydrothermal Scheduling Problem

Different solution methods have been proposed to solve the stochastic hydrothermal scheduling problem, using mainly linear or dynamic programming. In this section, only a cursory description of these two techniques is given and their application to a hydrothermal power system is presented. Then, in the next section, the approaches applied to a large-scale stochastic hydrothermal problem are studied more thoroughly.

Linear programming has been one of the most widely used techniques in water resources operations. It was designed for problems in which all relations among the variables are linear both in constraints and in the objective function to be optimized. Loucks [1968] and Houck and Cohon [1978] applied the stochastic linear programming to reservoir operation



(a)



(b)

Figure 1.8 Sequence of power system optimization problem

a) Long-term schedule with weekly intervals
b) Short-term schedule with daily intervals for week 2

assuming a Markov process model for the inflows. This technique, as pointed out in Loucks' paper [1968] leads to very high dimensional problems in real situations, that is to a problem where the number of constraints can easily exceed several thousands. Stochastic linear programming was also applied to reservoir operation by converting the probabilistic constraints into deterministic equivalences by using a cumulative probability distribution function of the random variable [Revelle et al., 1969; Hogan et al., 1981; Sobel 1975; Sniedovich, 1980; Houck et al., 1980; Houck and Datta, 1981]. This technique, known as "Chance-constrained linear programming" is severely hampered if cross-correlations exist among the inflows of a multireservoir system [Yeh, 1985].

Dynamic programming was first suggested as a solution technique for this type of problems by Little [1955]. The basic idea is to deal with multistage decision processes. As defined by Bellman [1957] multistage decision processing consists in separating a problem into a number of sequential steps, or stages which may be completed in one or more ways. The popularity and success of this technique in hydrothermal operation problems can be attributed to the fact that the nonlinear and stochastic features which characterize a large number of water resources systems can easily be incorporated into a dynamic programming problem. A large number of authors applied stochastic dynamic programming to find the optimal operating policy for hydro power systems. [Schweig and Cole, 1968; Gablinger and Loucks, 1970; Roefs and Bodin, 1970; Butcher, 1971; Askew, 1974a; 1974b; Su and Deiniger, 1974; Rossman, 1977; Turgeon, 1981a; 1981b]. In a review of mathematical models developed for reservoir operation, Yakowitz [1982] pointed out that the largest-sized stochastic dynamic programming problems found within or outside the water resources literature, involve no more than two or three state variables. Therefore, to solve large-scale stochastic problems, it becomes necessary to develop methods enabling us to approximate the solution of the operating problem at a reasonable computational cost. The next section will describe briefly these methods.

1.5 Review of Recent Approaches for Large-Scale Stochastic Reservoir Operation

The long-term scheduling problem is usually modelled as a stochastic nonlinear problem of very high dimension. The multiplicity of variables and constraints results from the large number of reservoirs and plants in the system.

Over the past thirty years, the determination of the optimal operating policy for this problem has been the subject of numerous publications. Yet a completely satisfactory solution has not been found since the problem has always been simplified in order to make it solvable. This is because the optimal feedback solution of large-scale stochastic optimization problems with bounds on the state and control variables is still unknown. Consequently, numerous approaches consist in transforming the large-scale problem into one or a series of small-scale problems. These approaches fall into four categories:

1) Aggregation
2) Decomposition
3) Aggregation / Decomposition
4) Projection

Aggregation methods reduce the number of state variables of a multireservoir system while conserving the stochastic nature of the problem. It consists in aggregating the whole system into one equivalent energy reservoir and calculating for that reservoir the optimal operation policy, as shown in Figure 1.9. Arvanitidis and Rosing [1970] were the first to propose this technique. They aggregated the whole Pacific Northwest system into a single reservoir and used stochastic dynamic programming with monthly time periods to solve the resulting one state variable problem. Many other papers, using this approach, have been published since. [Davis and Pronovost, 1972; Duran et al., 1985; Pereira and Pinto, 1984; Quintana and Chikhani, 1981; Sherkat et al., 1985; Turgeon, 1980, 1981b]. Turgeon [IREQ-2291, 1980] pointed out that the optimal operating rule obtained by using the aggregation method assumed simultaneous spilling from all the reservoirs or no spilling at all for a given period of time. Furthermore, it also assumed that all the reservoirs would become empty at the same time. Mathematically, if U_{ik} represents the "optimal" release from reservoir i in period k, $\bar{U}_i$ the capacity of the plant i, and X_{ik} the content of reservoir i at the beginning of period k then the following relations can be assumed:

$$U_{ik} \leq \bar{U}_i \quad \text{or} \quad U_{ik} > \bar{U}_i \tag{1.1}$$

and

$$X_{ik} > 0 \quad \text{or} \quad X_{ik} = 0 \tag{1.2}$$

$$i=1,2,\ldots,n$$

where n represents the number of reservoirs.

The decomposition approach consists in breaking down the huge

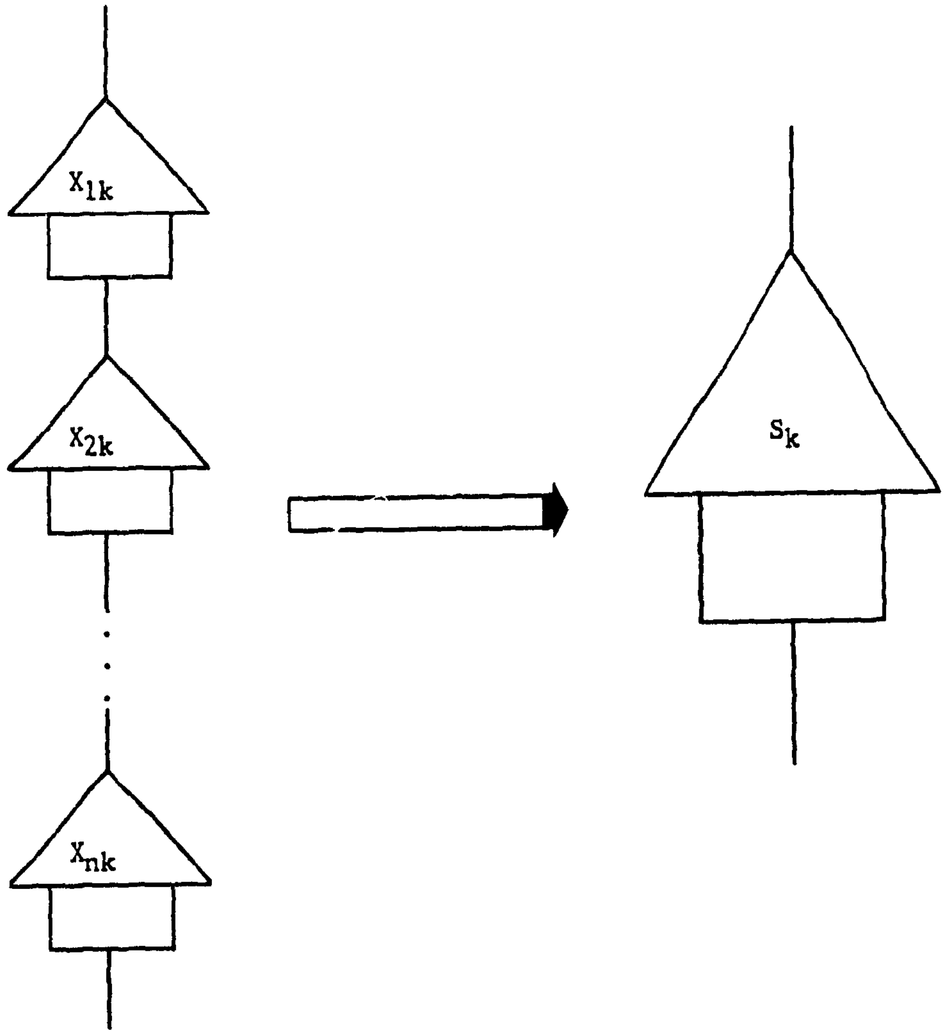


Figure 1.9 Aggregation Of n reservoirs into One equivalent energy content S_k.

operating problem into several smaller problems [Colleter and Lederer, 1981]. This technique has been applied to the French power system in order to form smaller problems, each corresponding to a particular valley. The link between the valleys and the entire generation system is established by using rules and relations based on the marginal production cost. More specifically, each valley is operated in order to maximize the expected amount of energy sold at the marginal production cost.

The aggregation/decomposition approach proposed by Turgeon [1981b] and used by Duran et al. [1985] and Lederer et al. [1983] was developed for a power system consisting of n hydroelectric power plants located in series on a river. The method consists in rewriting the stochastic optimization problem of n state variables as n-1 problems involving two state variables and using the stochastic dynamic programming formulation to obtain the solution. The release policy is then obtained for reservoir i as a function of the water content of that reservoir and the total amount of potential energy stored in the downstream reservoirs. Figure 1.10(a) illustrates the aggregation/decomposition of n reservoirs in series with content X_{ik}, i=1,...,n during a period k. Then, Figure 1.10(b) shows the reservoir 1 and reservoirs 2 to n combined with equivalent energy content S_{2k}. The advantage of this method, as pointed out by Turgeon [1981b], is that it is not iterative. In fact. the processing time increases only linearly with the number of reservoirs since only one additional dynamic programming problem of two state variables has to be solved for each new reservoir that is added to the system.

The projection methods proposed by Davis [1972] and used by

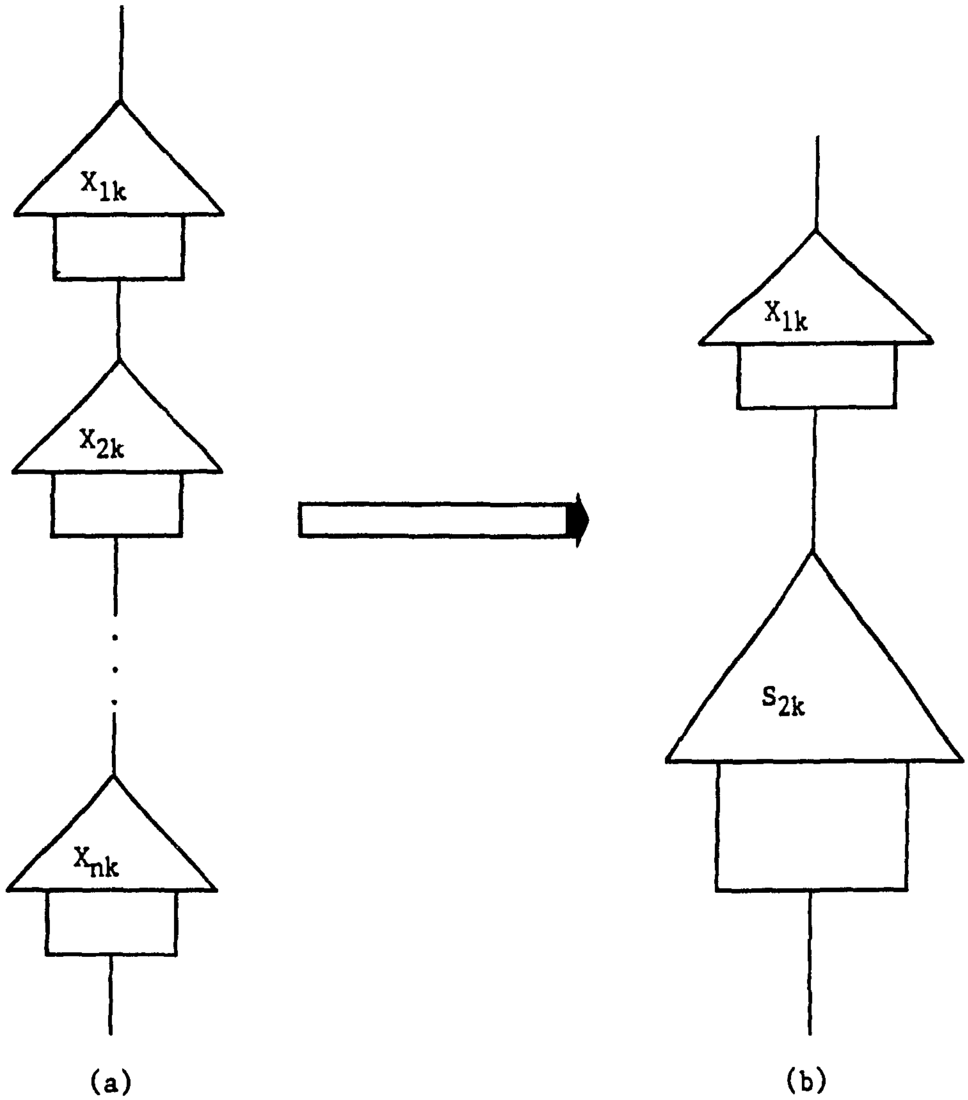


Figure 1.10 Aggregation / Decomposition illustration of

(a) The n reservoir power system
(b) reservoir 1 and reservoirs 2 to n combined

Delebecque and Quadrat [1978], Pronovost and Boulva [1978], and Sherkat et al. [1985] transforms the problem of n state variables into a series of n problems each with one state variable. This technique combined with dynamic programming and known as "Dynamic Programming Successive Approximation" involves a "one-at-a-time" stochastic optimization of each reservoir, and the procedure is repeated over all the reservoirs until convergence is attained. Figure 1.11 shows how this technique works for only two state variables.

The solution of the transformed problem obtained by applying one of the above manipulation techniques is never, unfortunately, the optimal global feedback solution sought. When aggregation is used, for instance, the solution may not even be feasible since the transformation cannot take into account all the local constraints on the reservoir content and the discharge of the power plants. Projection, on the other hand, always yields a feasible solution, albeit of local-feedback type. Moreover, as Turgeon [1980] has shown, such a solution can be very far from the global optimum when the states of the reservoirs are appreciably different from those expected. The price-decomposition approach of Collecter and Lederer [1981] gives a local-feedback solution similar to that obtained with projection. However, since the marginal production cost is not computed as an explicit function of the production of each valley, the solution thus obtained is not a global feedback.

The literature usually classifies all the methods mentioned above as "explicit", because they explicitly take into account the stochastic nature of the river flow. However, there are other methods that use synthetic flow sequences instead. The main features of those so-called "implicit" methods [Croley, 1974; McKerchar, 1975; Roefs and Bodin,

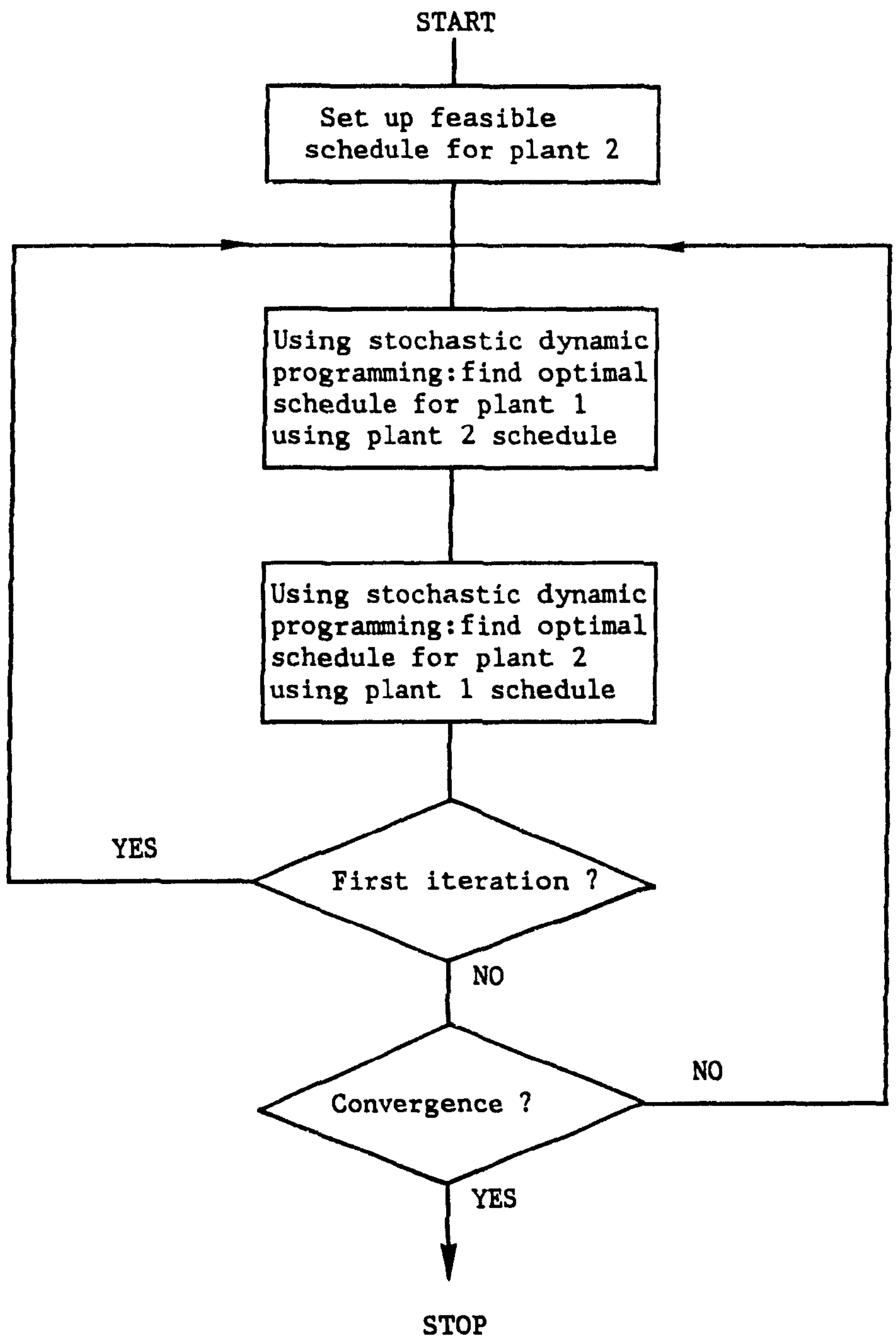


Figure 1.11 Projection flowchart

1970; Young, 1967] are streamflow synthesis, deterministic optimization and multivariate analysis. Streamflow synthesis is used to provide several equally likely future sequences of streamflows. For each streamflow sequence, deterministic optimization finds the optimal amount of water to release from each reservoir for each time period. If there are n streamflow sequences, the deterministic optimization therefore provides n different trajectories for each reservoir. Subsequently, multivariate analysis is applied to these trajectories to deduce an operating policy, that is a set of mappings from the state space (storage levels) into the decision space (reservoir releases). In other words, multivariate analysis is used to determine the functions $U_{ik}(X_{1k},X_{2k},\ldots X_{nk})$, $i=1,\ldots,n$; $k=1,\ldots K$, where U_{ik} represents the "optimal" release for reservoir i during period k and X_{ik} is the content of reservoir i at the beginning of period k.

However, the use of the word "optimal" to describe the functions $U_{ik}(X_{1k},X_{2k},\ldots X_{nk})$ determined by this process may be misleading. Actually, these functions are optimal for the implicit approach but their relation to the true optimal operating policy are not known and differ for each application. Unfortunately, determining the operating policy of a system by supposing the river flows to be perfectly known in advance rarely yields the same results as when explicit account is taken of the stochastic nature of the river flows. The difference between these two results can mainly be attributed to the limits on storage. Therefore, even if the deterministic problem were solved for a thousand different flow sequences to yield a thousand different trajectories for each reservoir, there is still no guarantee that the multivariate analysis of the results would yield the true optimal solution or even one that is just close to it.

The implicit stochastic approach was proposed principally because it is much easier to solve a series of deterministic large-scale optimization problems than only one stochastic large-scale problem. Roefs and Bodin [1970] reported that the growth in size of this approach is directly proportional to the number of reservoirs in the system rather than exponentially proportional as in the "Explicit Stochastic Approach". As mentioned before, the stochastic large-scale problem must be manipulated, decomposed or simplified before it can be solved. The question can be raised, however, as to whether or not the implicit approach has just simply substituted one difficult problem by another. Indeed, the implicit approach discarded the stochastic optimization problem but, as a result, must solve a very difficult multivariate analysis problem to determine the functions $U_{ik}(X_{1k},X_{2k},\ldots,X_{nk})$. The major difficulty here is to find the family of functions to which $U_{ik}(X_{1k},X_{2k},\ldots X_{nk})$ belongs. To assume that $U_{ik}(X_{1k},X_{2k},\ldots,X_{nk})$ is a linear function, as Young [1967] and McKerchar [1975] have done, is usually not acceptable in practice.

1.6 Scope and Contribution

The method proposed in this thesis to determine the optimal long-term operating policy of a multireservoir power system borrows from both the implicit and the explicit approaches explained in Section 1.5. As in the implicit approach, the problem is first solved deterministically for m different flow sequences. The results of those deterministic optimizations are then subjected to a principal component analysis (PCA) to find out whether the problem could have been modeled with fewer state variables. When this is found to be possible, the problem is transformed into a new one with fewer state variables, and stochastic dynamic programming is applied to determine the optimal operating policy. In other words, the implicit approach is used to reduce the number of state variables in the problem, and the explicit approach to find the optimal solution of the reduced problem.

Whether the number of state variables can be reduced or not depends on their degree of interdependency. If they are independent, reduction is of course impossible and the proposed method cannot solve the problem. However, if some dependency exists, reduction may be possible. Obviously, the higher the interdependency, the greater and better the reduction will be. Thus, our goal will be to reduce the number of variables sufficiently to allow straightforward application of stochastic dynamic programming.

The particular feature of the proposed approach is the search for linear dependencies among the variables which derives the required transformation for their reduction at the same time. It has many advantages over the explicit and implicit stochastic methods of the past [Saad and Turgeon, 1988a and 1988b]:

i) It solves larger problems (higher number of state variables) taking into account the stochastic nature of river flows, described by a discrete distribution as in the explicit stochastic optimization technique which was limited to two or three stochastic state variables [Yakowitz, 1982; Yeh, 1985].

ii) Unlike the aggregation methods [Arvantidis and Rosing 1970] it takes into account all local constraints on the discharge and content of the reservoirs, so that the solution obtained is generally feasible.

iii) Since linear combinations can easily be manipulated, the proposed technique decomposes the results to find the operating policy. In the previous works [Boggle and O'sullivan, 1979; Roefs and Bodin, 1970], the selection of the operating policy proved to be the most difficult problem to tackle.

iv) It gives a global feedback solution, contrary to the decomposition and projection methods. This means that global constraints, such as satisfying the demand, will be met.

v) It is easy to apply, which is certainly not the case for the methods involving the determination of the family of functions to which the discharge $U_{ik}(X_{1k},X_{2k},...,X_{nk})$ belongs.

The only major drawback is that the state variables must be interdependent in order to apply the method. In general, however, interdependency does exist among reservoirs located on the same river or on nearby rivers with similar flow patterns.

1.7 Outline and Methodology

The principal components approach, its theory, derivation, verification and justification are the focus of this thesis.

Firstly, in Chapter I, the hydrothermal power systems are outlined. A brief description of thermal and hydro generation plants is given. Then, the power systems optimization problems are described. These problems are usually decomposed into a long-term and short-term problems. The nature of each problem is briefly given. Some approaches to solve the stochastic long-term hydrothermal scheduling problem and a review of recent publications are also outlined.

The mathematical representation of the basic components of a power system is given in Chapter II. The objective function and the different constraints related to thermal and hydro plants are then considered. The electric demand and the corresponding constraints are also discussed. The model is given in its general deterministic form, as a nonlinear model, and a piecewise linearization is then proposed.

Chapter III presents different solution techniques for the deterministic model of Chapter II. In that chapter, dynamic, nonlinear and linear programming are discussed. The last technique is adopted and a modified linear model is suggested to eliminate the bang-bang solution usually associated with linear programming. The solution of the model for a four reservoir system is found with IBM's MPSX/370 package. Some optimization results are given.

Since the primary purpose of this work is to solve large-scale stochastic problems, a reduction of the number of state variables must be attempted. Chapter IV presents an efficient method to reduce a large-scale problem into a small-scale problem by using principal components analysis. First, the theoretical development to transform the original variables into a set of new components is formulated. These components have the two characteristics of being uncorrelated and in decreasing order of variance. An optimization problem built to obtain these characteristics shows that the new components are found by solving a simple eigenvalue problem. In this problem, the components and the percentage of variance are obtained from the eigenvectors and the eigenvalues of the covariance matrix respectively. In fact, the principal components analysis transforms a problem with n state variables into another having the same number of state variables. Therefore, reduction criteria are needed. Although the criterion selected is based on the component's percentage of the total variance, other criteria are also presented, namely: Scree test of Cattell, Kaiser's rule and Bartlett's rule.

This theory is then applied to the four reservoirs system of Chapter III and some results are given. These results are of paramount importance because they prove that the problem of eight state variables can be replaced by one of only four state variables.

Chapter V presents the stochastic model which is a generalization of the problem described in Chapter II, with the objective of minimizing the expected cost of the thermal energy.

The reduced problem is obtained by incorporating the principal components into the model, and is then solved by stochastic dynamic programming.

A comparison with the classical dynamic programming approach, that is without any reduction of the number of state variables, shows the efficiency of the proposed technique. The value of the objective function, the operating policies and the CPU time obtained from both methods confirm the advantage of the principal components approach when large-scale problems must be solved.

Finally, Chapter VI presents a summary of the results along with the conclusion. Further research recommendation are also outlined in that chapter.

CHAPTER II

PROBLEM FORMULATION

2.1 Introduction

The long-term operation problem of multireservoir systems is modelled in this chapter as a nonlinear deterministic problem. A linear version of the model is presented at the end of this chapter. The stochastic generalization of the problem is described in Chapter V.

The basic components of a power system are thermal and hydro plants. The mathematical representation of these components is given in different parts of this chapter. Section 2.2 considers the thermal plants, their production costs and the constraints introduced in the model. Section 2.3 studies the hydro plants and reservoir dynamics. The modelling of rivers is also discussed in that section and the constraints are explained there as well. The electric demand and the corresponding constraints are discussed in Section 2.4. Finally, in Section 2.5 a piecewise linearization of the model is proposed.

2.2 Thermal Plants Formulation

As mentioned in Chapter I, the most important economic consideration is to produce electricity at the lowest operation cost. Consequently, a minimum use of thermal energy over the whole time horizon is desirable.

For the remainder of this thesis, this horizon will be characterized by a finite, but possibly very large number of periods indexed by $k=1,\ldots,K$ (e.g. the periods may be months and the horizon be one year). The cost function $C_k(He_k)$ of thermal energy during period k is assumed to be convex as shown in Figure 2.1, which means that it can be written as

$$F = \sum_{k=1}^{K} C_k(He_k) \tag{2.1}$$

The electric production of a thermal plant is limited. For this reason, bounds are assumed on the thermal energy He_k and are expressed as

$$0 \leq He_k \leq \overline{He} \quad ; \qquad k=1,\ldots,K \tag{2.2}$$

where $\overline{He}$ is the capacity of the thermal plants (upper bound).

2.3 <u>Hydro Plants Formulation</u>

The hydro models described in this section consider plants in series only or plants in series and parallel. Two or more hydro plants on the same stream constitute a series arrangement as shown in Figure 2.2. A series and parallel arrangement is defined as two or more plants on different streams converging into another which may have several downstream plants as shown in Figure 2.3. In addition, it is assumed that a reservoir is associated with each plant.

2.3.1 <u>Water Continuity Equations</u>

Let X_{k+1} be the storage of the reservoir at the end of period k. This storage is equal to the volume X_k at the beginning of that period

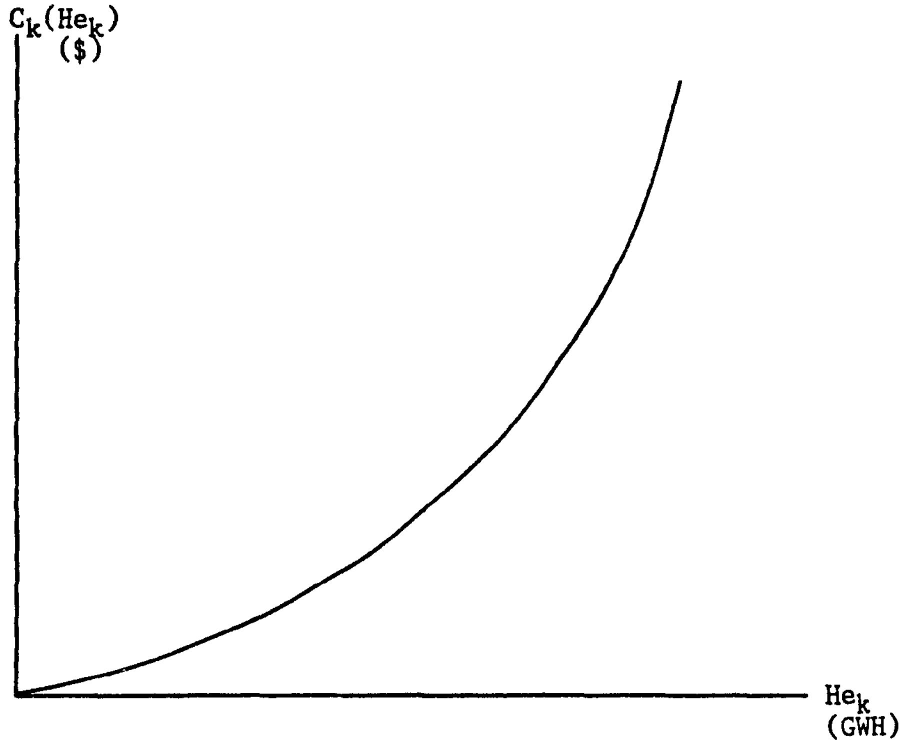


Figure 2.1 Production cost

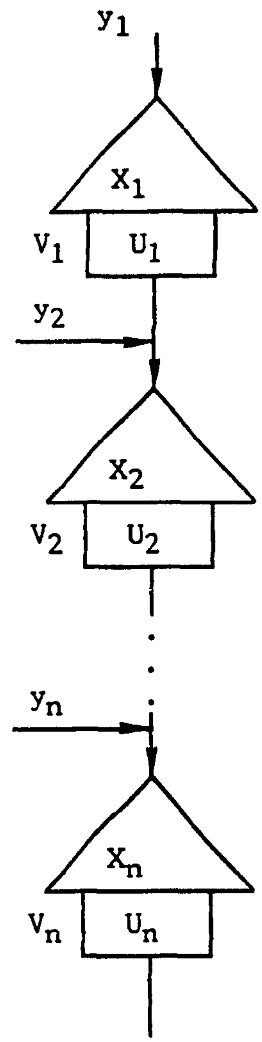


Figure 2.2 Hydro plants in series

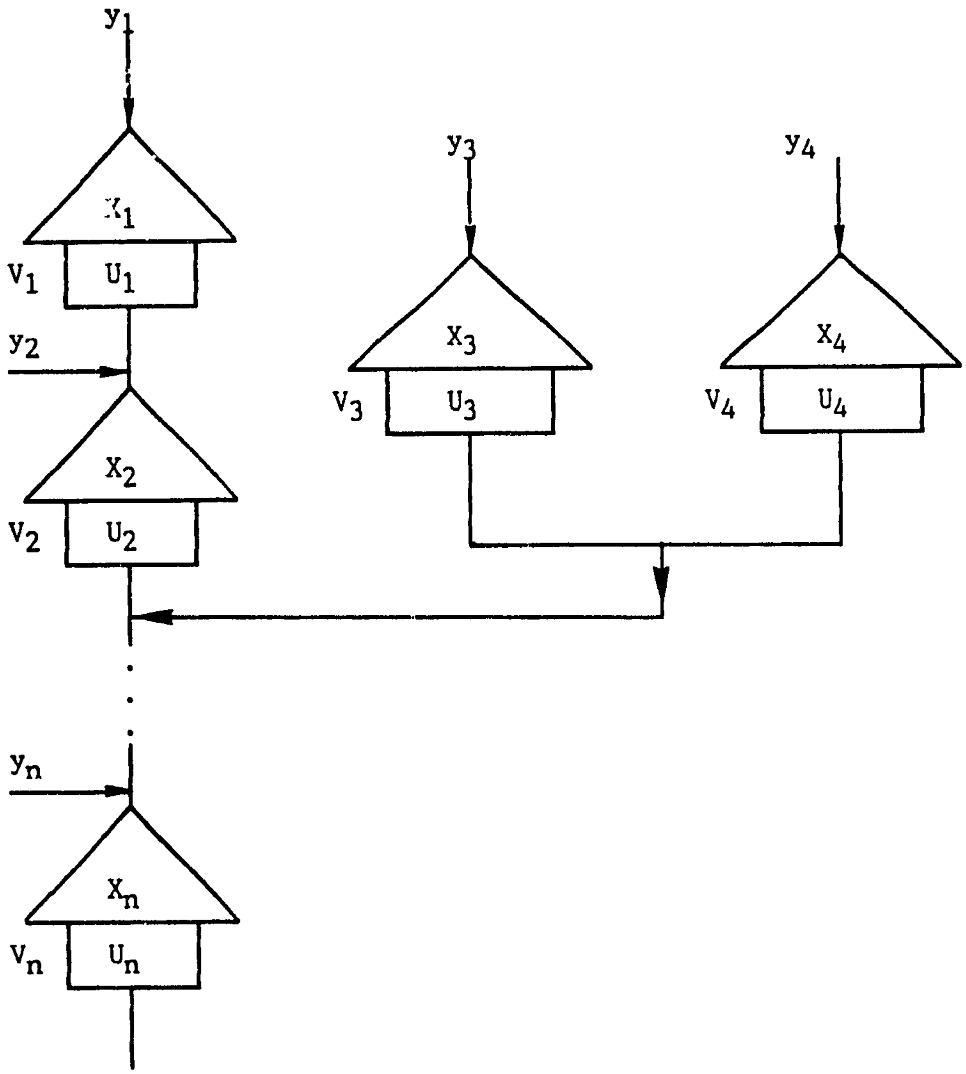


Figure 2.3 Hydro plants in series and parallel

plus the difference between the inflows Y_k and the outflows during that period. The inflows (rainfall, snow and melted snow) form the only input to the model, whereas the outflows consist of the discharge U_k from the reservoirs and the volume V_k discharged through the spillways. Reservoir levels vary throughout the year between their lower and upper bounds. Bounds also exist on U_k and V_k.

Taking all these considerations into account, the water continuity equations (or reservoir dynamics) can be represented by the following difference equations

$$X_{k+1} = X_k - \Gamma_1 U_k - \Gamma_2 V_k + Y_k \; ; \; k=1,\dots,K \tag{2.3}$$

and the bounds are represented by

$$\underline{X} \le X_{k+1} \le \bar{X} \quad ; \; k=1,\dots,K \tag{2.4}$$

$$0 \le U_k \le \bar{U} \quad ; \; k=1,\dots,K \tag{2.5}$$

$$V_k \ge 0 \quad ; \; k=1,\dots,K \tag{2.6}$$

where Γ_1 and Γ_2 are matrices dependant on the physical structure of the hydroelectric installation.

Example

Consider the four reservoir system shown in Figure 2.4. As assumed previously a reservoir is associated with each plant, therefore the water continuity equations become

$$X_{1k+1} = X_{1k} - U_{1k} - V_{1k} + Y_{1k} \tag{2.7}$$

$$X_{2k+1} = X_{2k} - U_{2k} - V_{2k} + U_{1k} + V_{1k} + Y_{2k} \tag{2.8}$$

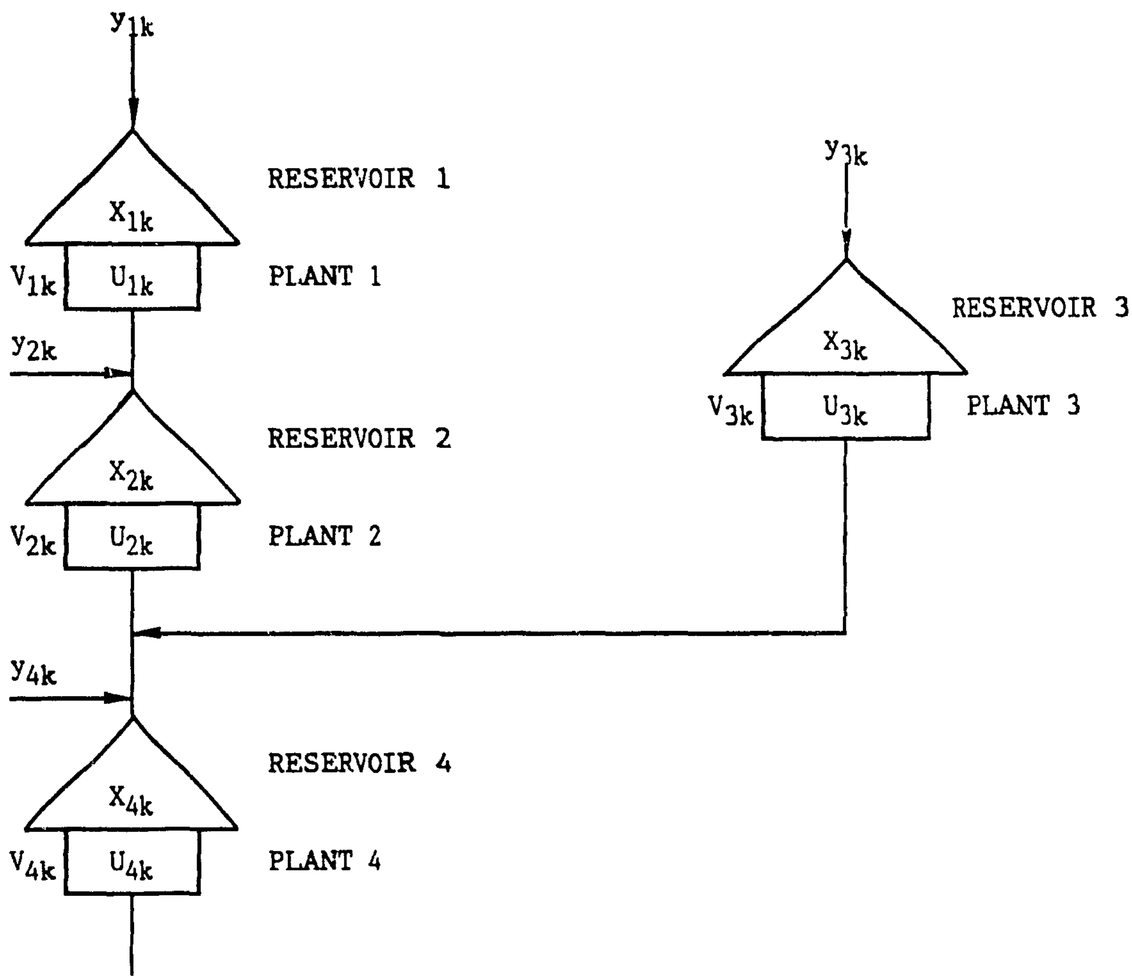


Figure 2.4 Schematic representation of the system

$$X_{3k+1} = X_{3k} - U_{3k} - V_{3k} + Y_{3k} \tag{2.9}$$

$$X_{4k+1} = X_{4k} - U_{4k} - V_{4k} + U_{2k} + V_{2k} + U_{3k} + V_{3k} + Y_{4k} \tag{2.10}$$

where the subsripts 1,...,4 denote the reservoir numbers.

Equations (2.7) to (2.10) can be rewritten in condensed form as equation (2.3) with

$$\Gamma_1 = \Gamma_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{bmatrix}$$

2.3.2 Hydroelectric Generation

The electrical energy that can be generated from the potential energy of the water stored in a reservoir is a function of both the released water volume U_{ik} and the hydraulic head h_{ik} of the reservoir i during the period k. The head h_{ik} is not an independent variable but a function of X_{ik}, the content cf the allied reservoir. It follows that the production $H_{ik}(X_{ik},U_{ik})$ of plant i in period k is usually given by the following equation

$$H_{ik} = \alpha \cdot r_i(h_{ik},U_{ik}) \cdot h_{ik} \cdot U_{ik} \text{ ; } i=1,\dots,n;\ k=1,\dots K; \tag{2.11}$$

where α is a constant and $r_i(h_{ik},U_{ik})$ denotes the plant efficiency.

The total generation by n hydroplants is then:

$$H_k(X_k,U_k) = \sum_{i=1}^{n} H_{ik}(X_{ik},U_{ik}) \text{ ; } k=1,\dots,K \tag{2.12}$$

It is important to recall that for long-term hydro scheduling, the variation of the head with respect to the storage can be assumed constant [Hanscom, 1976; Read, 1979].

2.4 The Electric Demand

Figure 2.5 shows how the load (the demand) pattern may look. In this figure d_p denotes the load in MW, and de_k the energy demand for period k in GWh. Usually, it is necessary to divide the load pattern into time intervals as shown in Figure 2.6 assuming of course that the load is constant during each interval.

The need to meet the demand during each period will result in the following energy balance constraints

$$H_k(X_k,U_k) + He_k = de_k \ ; \ k=1,...K \tag{2.13}$$

These constraints are the only coupling constraints between the hydro and thermal systems. It can be seen that de_k has to be satisfied in each period of time. In other words, if the hydroelectric generation is not sufficient the thermal source of energy should then be used in order to meet demand.

Finally, it is assumed that the demand de_k is known at the beginning of the operating horizon.

2.5 Linear Model

A complete linear model is obtained once the thermal production cost $C_k(He_k)$ and the hydroelectric generation function $H_k(X_k,U_k)$ are linearized. As shown in Figure 2.1, the cost $C_k(He_k)$ is a convex

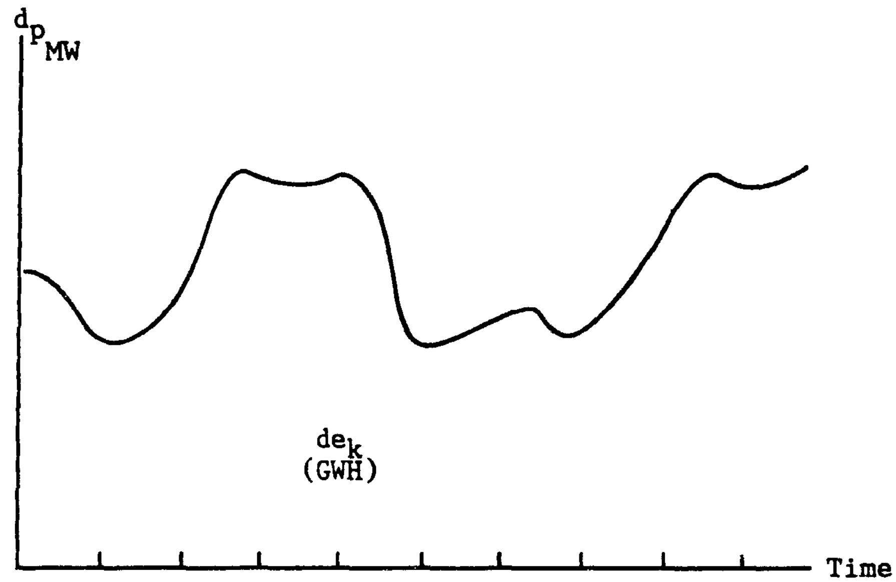


Figure 2.5 Load pattern

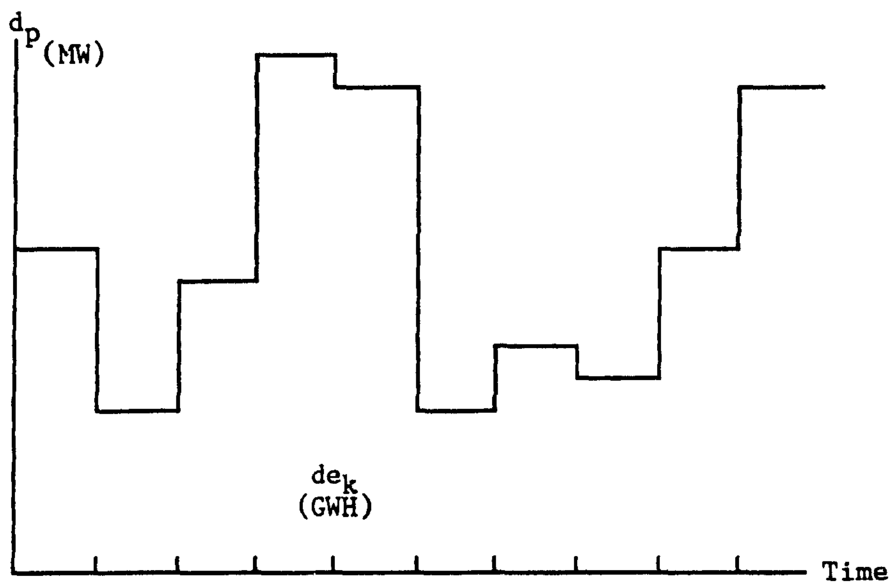


Figure 2.6 Discrete load pattern

function. This function can be approximated by the piecewise linear curve illustrated in Figure 2.7. This linearization introduces the following additional equations

$$C_k(He_k) = \sum_{m=1}^{M} a_m \cdot Hg_{mk} = a^T Hg_k \qquad ; k=1,\ldots K \qquad (2.14)$$

$$He_k \qquad = \sum_{m=1}^{M} b_m \cdot Hg_{mk} = b^T Hg_k \qquad ; k=1,\ldots,K \qquad (2.15)$$

$$0 \le Hg_k \le 1 \qquad ; k=1,\ldots,K \qquad (2.16)$$

where M is the number of line segments forming the grid of the piecewise linear cost function. $a^T = [\ a_1,\ldots,a_M]$ and $b^T = [\ b_1,\ldots,b_M]$ are the lengths of the resulting intervals on $C_k(He_k)$ and He_k axis, respectively (Figure 2.7). $Hg_k = [\ Hg_1,\ldots,Hg_m]^T$ is a vector of special variables or grid variables defined as follows

For He_k in interval m

$$Hg_{1k} = Hg_{2k} = \ldots = Hg_{m-1,k} = 1$$

$$0 \le Hg_{mk} \le 1$$

$$Hg_{m+1,k} = \ldots = Hg_{Mk} = 0$$

Appendix A contains more details on the piecewise linearization.

Assuming constant head for all reservoirs, the hydroelectric generation can be written as

$$H_k(X_k,U_k) = \sum_{i=1}^{n} \rho_i U_{ik} = \rho^T U_k \qquad ; k=1,\ldots,K \qquad (2.17)$$

where $\rho^T = [\ \rho_1,\ldots,\rho_n]$ is the vector of generation characteristics for

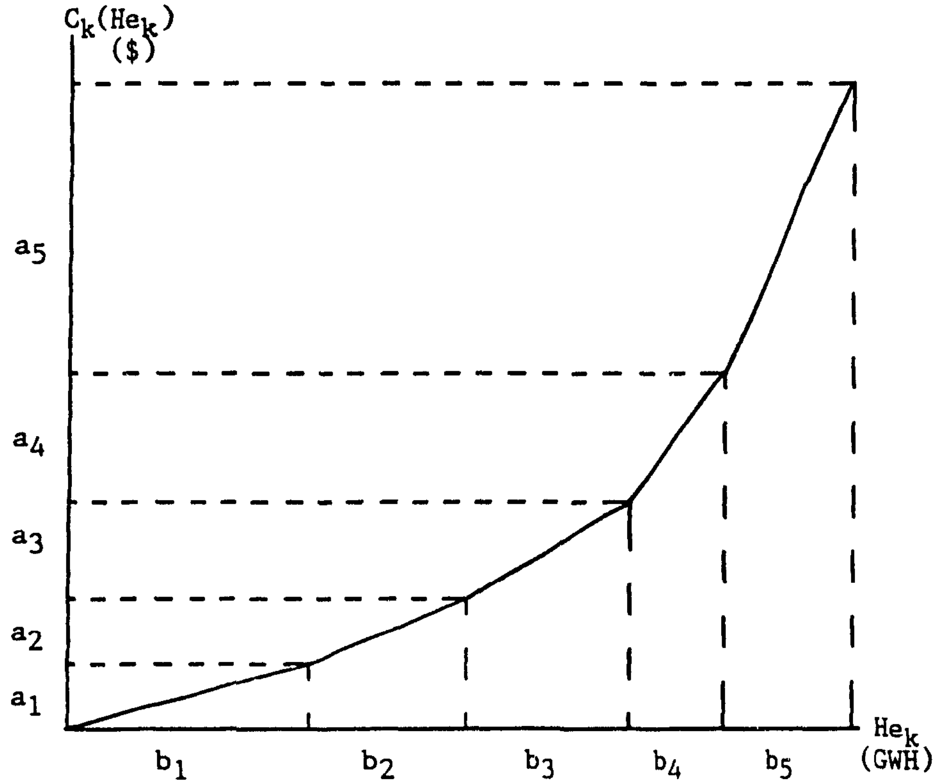


Figure 2.7 Piecewise linear approximation of thermal cost function

the hydro plants. This vector will be assumed constant over the operating time horizon.

In compact form, the complete linear model can be written as:

$$\min F = \sum_{k=1}^{K} \{C\ a^T\ Hg_k\} \tag{2.18}$$

Subject to:

$$X_{k+1} = X_k - \Gamma_1 U_k - \Gamma_2 V_k + Y_k \quad ; k=1,\ldots,K \tag{2.19}$$

$$\rho^T U_k + He_k = de_k \quad ; k=1,\ldots,K \tag{2.20}$$

$$He_k - b^T Hg_k = 0 \quad ; k=1,\ldots,K \tag{2.21}$$

$$\underline{X} \le X_{k+1} \le \bar{X} \quad ; k=1,\ldots,K \tag{2.22}$$

$$0 \le U_k \le \bar{U} \quad ; k=1,\ldots,K \tag{2.23}$$

$$V_k \ge 0 \quad ; k=1,\ldots,K \tag{2.24}$$

$$0 \le He_k \le \bar{He} \quad ; k=1,\ldots,K \tag{2.25}$$

$$0 \le Hg_k \le e \quad ; k=1,\ldots,K \tag{2.26}$$

where C is the cost of each GWh produced from thermal plants, and $e^T = [1,\ldots,1]$ is a 1 by n unit vector.

2.6 Conclusion

This chapter has presented the general model for the long-term multireservoir hydrothermal systems. The model variables are all continuous, whereas the three kinds of constraints are either on the thermal or on the hydroplants with only one coupling constraint between the hydro and thermal generation.

The model developed is nonlinear due to both the objective function (production costs) and the hydroelectric generation constraints. For these constraints, it is assumed that the water head remains constant and that the hydro generation is only a function of the released water.

Although this nonlinear model can be solved directly, a linearization was proposed to alleviate the computation of the solution. This technique will be discussed in the next chapter and some optimization results will be presented there.

CHAPTER III

SOLUTION TECHNIQUES FOR THE DETERMINISTIC PROBLEM

3.1 Introduction

The models developed in Chapter II can be solved using different methods. Nonlinear programming can be used for solving (2.1)-(2.13). The linear model of (2.18)-(2.26) can be solved by linear programming. Dynamic programming is applicable to both models.

This chapter describes the application of these methods to the solution of the multireservoir long-term scheduling problem. The dynamic programming formulation is discussed in Section 3.2. In Section 3.3, a nonlinear programming algorithm is presented, and a linear programming method is covered in Section 3.4. The last technique is selected to solve the deterministic linear model. In Section 3.5, a modified linear model is proposed and, in Section 3.6, an application to a multireservoir system is presented.

3.2 Dynamic Programming

As discussed previously, dynamic programming is a powerful tool for solving water resources systems problems. It has the advantage of effectively decomposing highly complex problems with a large number of variables into a series of subproblems which are solved recursively. In order to illustrate how dynamic programming can be applied to hydrothermal scheduling problem, the following nonlinear model is considered:

$$\min F = \sum_{k=1}^{K} C_k(de_k - H_k(X_k,U_k)) \tag{3.1}$$

subject to

$$X_{k+1} = X_k - \Gamma_1 U_k - \Gamma_2 V_k + Y_k \tag{3.2}$$

$$\underline{X} \le X_{k+1} \le \bar{X} \tag{3.3}$$

$$0 \le U_k \le \bar{U} \tag{3.4}$$

$$V_k \ge 0 \tag{3.5}$$

$$k = 1, \ldots, K$$

where $C_k(de_k - H_k(X_k,U_k))$ is the production cost for period k.

The model (3.1)-(3.5) is identical to the nonlinear one presented in Chapter II with the difference that the objective function $C_k(He_k)$ is written in terms of the demand de_k and hydroelectric production $H_k(X_k,U_k)$ in period k.

Let $J_k(X_k,U_k)$ be the production cost for period k with the vector X_k representing the storages (states) at the beginning of period k and the vector U_k the releases during the same period. Then (3.1) can be written as [Nemhauser, 1966]:

$$F_k(X_k) = \min_{U_k} [J_k(X_k,U_k) + F_{k+1}(X_{k+1})] \tag{3.6}$$

where $F_{k+1}(X_{k+1})$ is the optimal production cost from period k to period K+1. At the end of the horizon,

$$F_{K+1}(X_{K+1}) = 0 \tag{3.7}$$

For the long-term model, the state X_{i1}, $i=1,\ldots,n$ at the beginning of the time horizon is known.

Even if the state X_{ik} of reservoir i can take an infinite number of values between the lower and upper bounds, in practice the number of states is fixed at a finite value in order to reduce the computation time. Therefore, it will be assumed that the state X_k can take one of the following values only:

$$\underline{X} = \sigma_{1k} < \sigma_{2k} < \sigma_{3k} < \ldots < \sigma_{mk} = \bar{X} \tag{3.8}$$

Now, let us consider the case of a single reservoir. The state variable (storage) after being discretized into a number of feasible states is shown in Figure 3.1. Moreover, it is supposed that the inflow sequence is given and that the spilling term V_k is temporarily ignored, then the continuity equation (3.2) becomes

$$X_{k+1} = X_k - U_k + Y_k \tag{3.9}$$

If X_{k+1} and X_k are known, U_k can be directly computed from the above continuity equation. The optimization can thus be made with the proper values of U_k. The problem of interpolation is also avoided since the U_k's are computed by fixing the states X_{k+1} and X_k. Solutions are then imbedded in the discretized states. Moreover, the infeasible transitions are automatically discarded in the solution process. The procedure for determining the optimal schedule and the minimum operation cost is shown in the flowchart of Figure 3.2.

Next, when two reservoirs are in cascade, the recursive equation

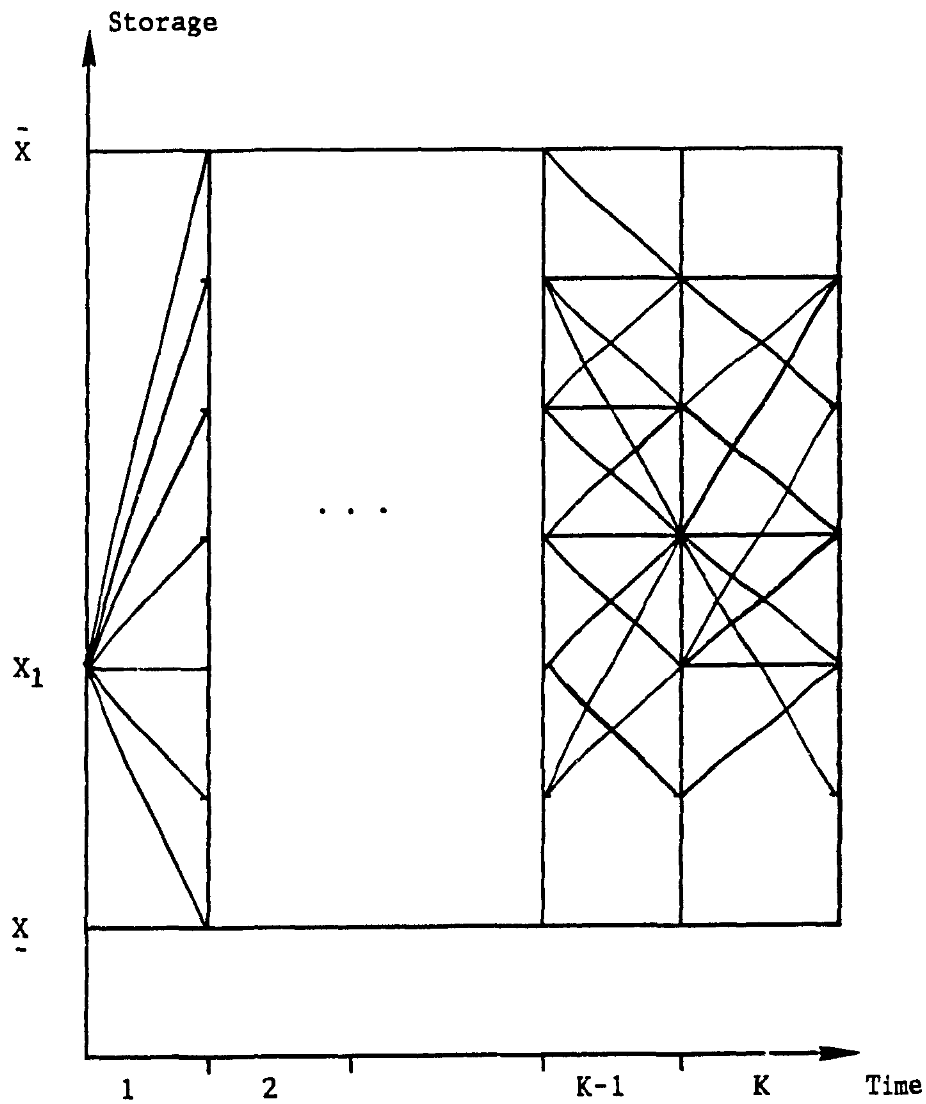


Figure 3.1 Dynamic programming state space

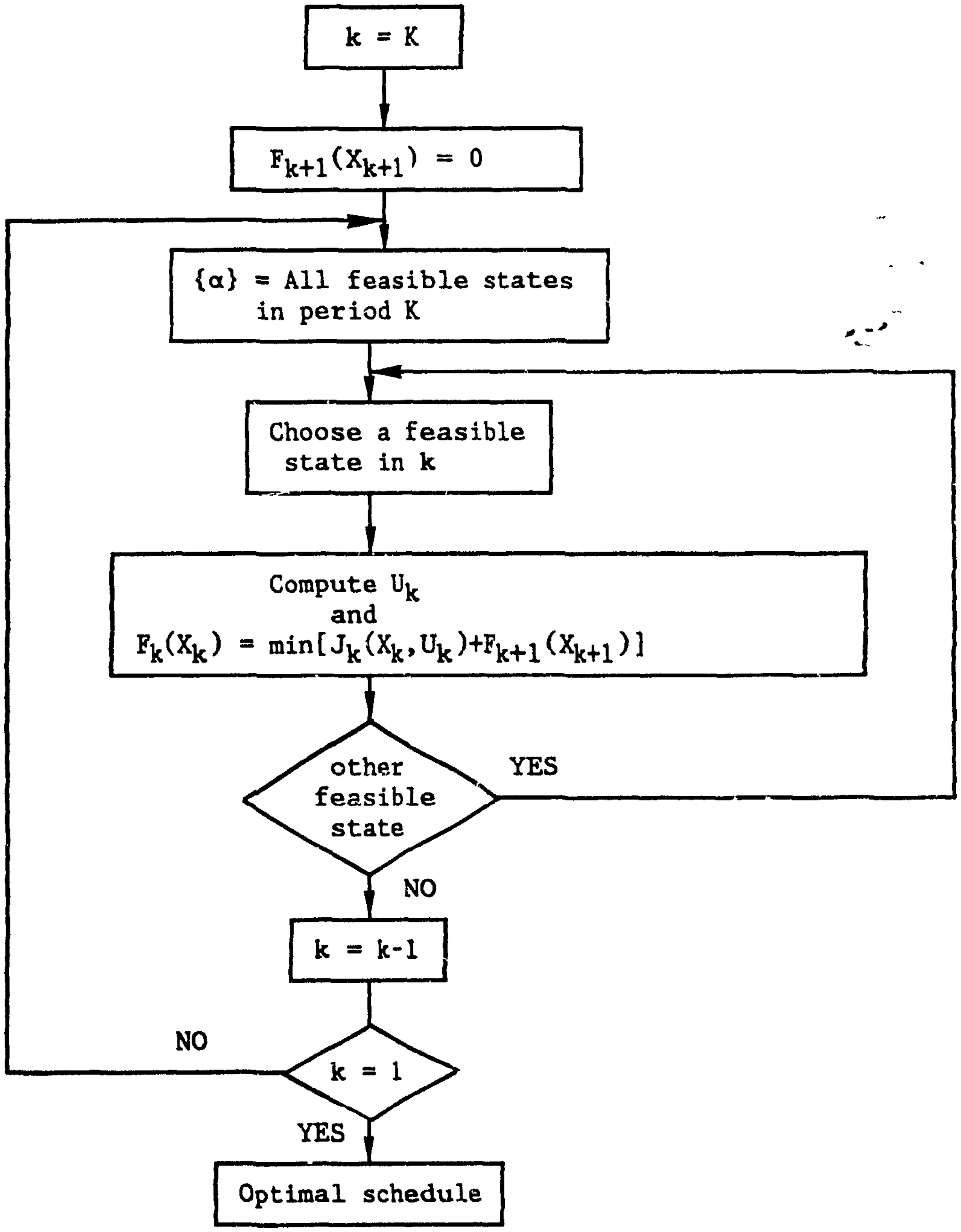


Figure 3.2 Backward dynamic programming flowchart

(3.6) can be rewritten as:

$$F_k(X_{1k},X_{2k})=\underset{U_{1k},U_{2k}}{\text{minimum}}\{J_k(X_{1k},X_{2k},U_{1k},U_{2k}) + F_{k+1}(X_{1,k+1},X_{2,k+1})\} \quad (3.10)$$

and in a similar way, for n cascaded reservoirs:

$$F_k(X_{1k},\ldots,X_{nk}) = \text{minimum}\{J_k(X_{1k},\ldots,X_{nk},U_{1k},\ldots,U_{nk}) + F_{k+1}(X_{1,k+1},\ldots,X_{n,k+1})\} \quad (3.11)$$

Unfortunately, equation (3.11) cannot be solved directly for large values of n because computation time and storage requirements become excessive then. According to Turgeon [IREQ-2291, 1980], problems should not have more than four state variables (four reservoirs) in order to be solvable by dynamic programming without difficulties.

3.3 Nonlinear Programming

Nonlinear programming has not gained as much popularity as dynamic programming in water resources system analysis. The solution of large-scale nonlinear problems are not generally easy to find within reasonable computation time. However, the long-term nonlinear model (2.1)-(2.13) can be solved using the "conjugate gradient" method [Gagnon et al., 1974; Hicks et al., 1974], or the "reduced gradient" approach [Hanscom et al., 1976], or even the "dual variables" or "Lagrange multipliers" approach [Haimes, 1977].

To illustrate how nonlinear programming can be applied to hydrothermal scheduling problems, the nonlinear model is considered for the particular case of a single reservoir. The solution method proposed here is the "dual variable" approach. For that purpose, the model can be written as

$$\min F = \sum_{k=1}^{K} C_k(He_k) \tag{3.12}$$

subject to

$$X_{k+1} = X_k - U_k - V_k + Y_k \tag{3.13}$$

$$\alpha.r(h_k,U_k).h_k.U_k + He_k = de_k \tag{3.14}$$

$$\underline{X} \le X_{k+1} \le \bar{X} \tag{3.15}$$

$$0 \le U_k \le \bar{U} \tag{3.16}$$

$$0 \le He_k \le \bar{He} \tag{3.17}$$

$$V_k \ge 0 \tag{3.18}$$

$$k = 1, \ldots, K$$

As previously mentioned, the production cost $C_k(He_k)$ and the hydroelectric generation $\alpha.r(h_k,U_k).h_k.U_k$ are nonlinear.

In its general form the nonlinear programming problem can be modeled as

$$\min_X f(X) \tag{3.19}$$

subject to

$$g(X) \ge 0 \tag{3.20}$$

in which X is a vector of decision variables, and f(X) and g(X) are real-valued and vector-valued given functions, respectively.

The dual problem is solved in two steps. First, the Lagrangian associated with the constrained problem is defined as

$$L(X,\lambda) = f(X) - \lambda g(X) \tag{3.21}$$

where λ (≥ 0) is the vector Lagrange multiplier.

Secondly, the dual function is defined by the equation

$$\mathcal{L}(\lambda) = \min_{X} L(X,\lambda) \tag{3.22}$$

and the solution for the dual problem associated with the original (primal) problem (3.19)-(3.20) requires the maximization of the function $\mathcal{L}(\lambda)$ over the set $\lambda \geq 0$; that is:

$$\max_{\lambda \geq 0} \mathcal{L}(\lambda) \tag{3.23}$$

The optimal solution is obtained when

$$\mathcal{L}(\hat{\lambda}) = f(\hat{X}) \tag{3.24}$$

in which $\hat{\lambda}$ and $\hat{X}$ are the optimal values sought [Lasdon, 1970; Luenberger, 1973]. The procedure to determine the optimal operation schedule is shown in Figure 3.3.

Now by applying this formulation to the model of (3.12)-(3.18) the Lagrangian is found to be:

$$L = \sum_{k=1}^{K} \{C_k(He_k) - \lambda_{1k}(X_{k+1} - X_k + U_k + V_k - Y_k) - \lambda_{2k}[\alpha . r(H_k, U_k) . H_k . U_k + He_k - de_k]$$
$$- \lambda_{3k}(X_{k+1} - \underline{X}) - \lambda_{4k}(\bar{X} - X_{k+1}) - \lambda_{5k}(\bar{U} - U_k) - \lambda_{6k}(\bar{He} - He_k)\} \tag{3.25}$$

The dual problem associated with this model is:

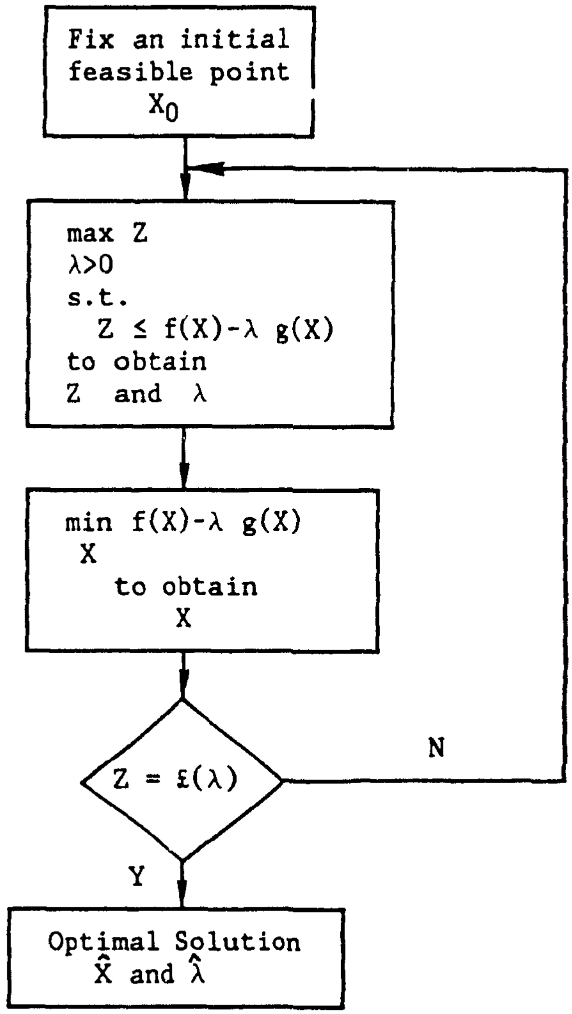


Figure 3.3 Dual variables flowchart

$$\pounds(\lambda)= \max_{\lambda} \min L \qquad (3.26)$$

subject to

$$\lambda \geq 0 \qquad (3.27)$$

Finally, it should be pointed out that nonlinear programming methods have been criticized for inconsistent convergence to acceptable results in real situations and long computing time requirements [Yeh 1985]. Furthermore, the mathematics involved in nonlinear models are usually very complex.

3.4 Linear Programming

In water resources applications, linear programming has been widely used for the optimization of complex reservoir systems with large number of variables and constraints. This approach is justified by the following facts:

1- The optimization process is usually fast and does not require large computer memory and time.

2- The dual formulation can be used to solve the problem if the number of constraints exceeds the number of decision variables. Indeed, it can be shown that every linear program has a dual formulation [Chvatal, 1983; Murtagh, 1981; Nazareth, 1987].

3- It is possible to solve a modified problem using the results obtained from the original problem. This property is very beneficial in solving the linear model of (2.18)- (2.26) for several inflow sequences.

4- Commercial computer programs, such as MPSX/370 by IBM, are widely available.

On the other hand, the linear model (2.18)-(2.26) has the following drawbacks:

1- Approximating the nonlinear objective function (2.1) by the piecewise linear function (2.18) introduces some inaccuracy. There is no fixed rule for selecting either the optimum grid size or the optimum number of grids to improve this approximation. However using large grid sizes may produce too inaccurate results, while specifying a large number of small grid points may prove to be unnecessary [MPSX, 1979].

2- The hydroelectric generation of plant i in period k is given by:

$$H_{ik}(X_{ik},U_{ik}) = \rho_i U_{ik} \tag{3.28}$$

Since the production is no longer a function of the water head when (3.28) is used, the advantage of emptying all reservoirs simultaneously in order to keep the head high at each plant disappears. As a result, a bang-bang solution, shown in Figure 3.4, in which a plant is run at maximum capacity in one period and shut down in the next is obtained. Such a solution is obviously unacceptable in practice. For this reason, a modi fied linear model is proposed in the next section.

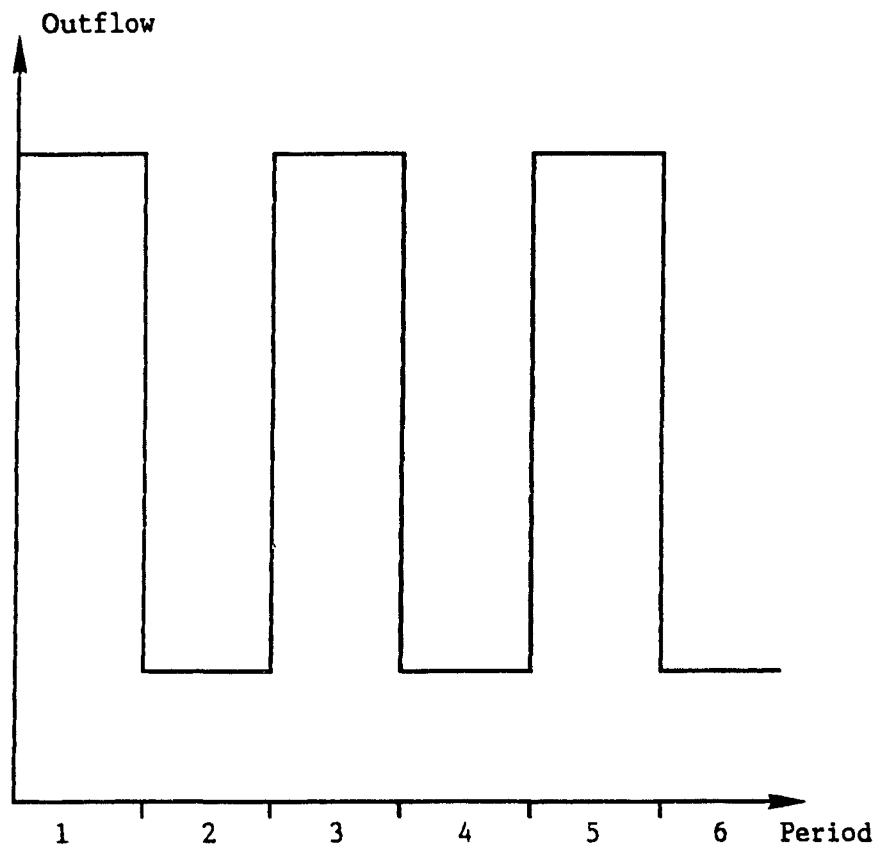


Figure 3.4 Bang-Bang output

3.5 Modified Linear Model

As mentioned previously, the bang-bang solution cannot be implemented in practice. Therefore, to eliminate this form of solution, penalty terms were added to the objective function in order to penalize any variation in the production of a plant from one period to the next. Thus, the objective (2.18) can be stated as:

$$\min \sum_{k=1}^{K} \{C.a^T Hg_k\} + \alpha \sum_{k=1}^{K-1} \sum_{i=1}^{n} |U_{ik+1} - U_{ik}| \tag{3.29}$$

where α is a constant having a small value compared to the production cost C (< 10%)

The function (3.29) can also be rewritten in the linear programming standard form as:

$$\min \sum_{k=1}^{K} \{C.a^T Hg_k\} + \alpha \sum_{k=1}^{K-1} \sum_{i=1}^{n} z_{ik} \tag{3.30}$$

Where z_{ik} is defined by:

$$U_{ik+1} - U_{ik} - z_{ik} \leq 0 \tag{3.31}$$

$$U_{ik+1} - U_{ik} + z_{ik} \geq 0 \tag{3.32}$$

$$z_{ik} \geq 0 \tag{3.33}$$

$$k=1,\ldots,K$$

Letting $e^T z_k = \sum_{i=1}^{n} z_{ik}$, where $e^T = [1,1,\ldots,1]$, then the complete linear model used to determine the optimal schedule becomes:

$$\min F = \sum_{k=1}^{K} \{C.a^T Hg_k\} + \alpha \sum_{k=1}^{K-1} e^T z_k$$

subject to

$$X_{k+1} = X_k - \Gamma_1 U_k - \Gamma_2 V_k + Y_k$$

$$\rho^T U_k - b^T Hg_k = de_k$$

$$U_{k+1} - U_k - z_k \leq 0$$

$$U_{k+1} - U_k + z_k \geq 0$$

$$\underline{X} \leq X_{k+1} \leq \bar{X}$$

$$0 \leq U_k \leq \bar{U} \tag{3.34}$$

$$0 \leq Hg_k \leq e$$

$$V_k \geq 0$$

$$z_k \geq 0$$

$$k=1,2,\ldots,K$$

The next section illustrates the solution of this model with IBM's MPSX/370 package for a four reservoirs system.

Installation	Reservoir Capacity (hm^3)	Reservoir Lower bound (hm^3)	Plant Capacity (m^3/sec)	Average Efficiency (KWh/m^3)
LG4	34000	8000	2581.9	0.2808
LG3	25200	6000	3432.9	0.1864
EOL	37500	10000	3560	0.1742
LG2	19370	5000	5954	0.3249

Table 3.1 Characteristics of the installations

3.6 Sample Application

The model of (3.34) is used to determine the optimal solution of Hydro Québec's installations on La Grande river. These installations are schematized in Figure 3.5 and their main characteristics are given in

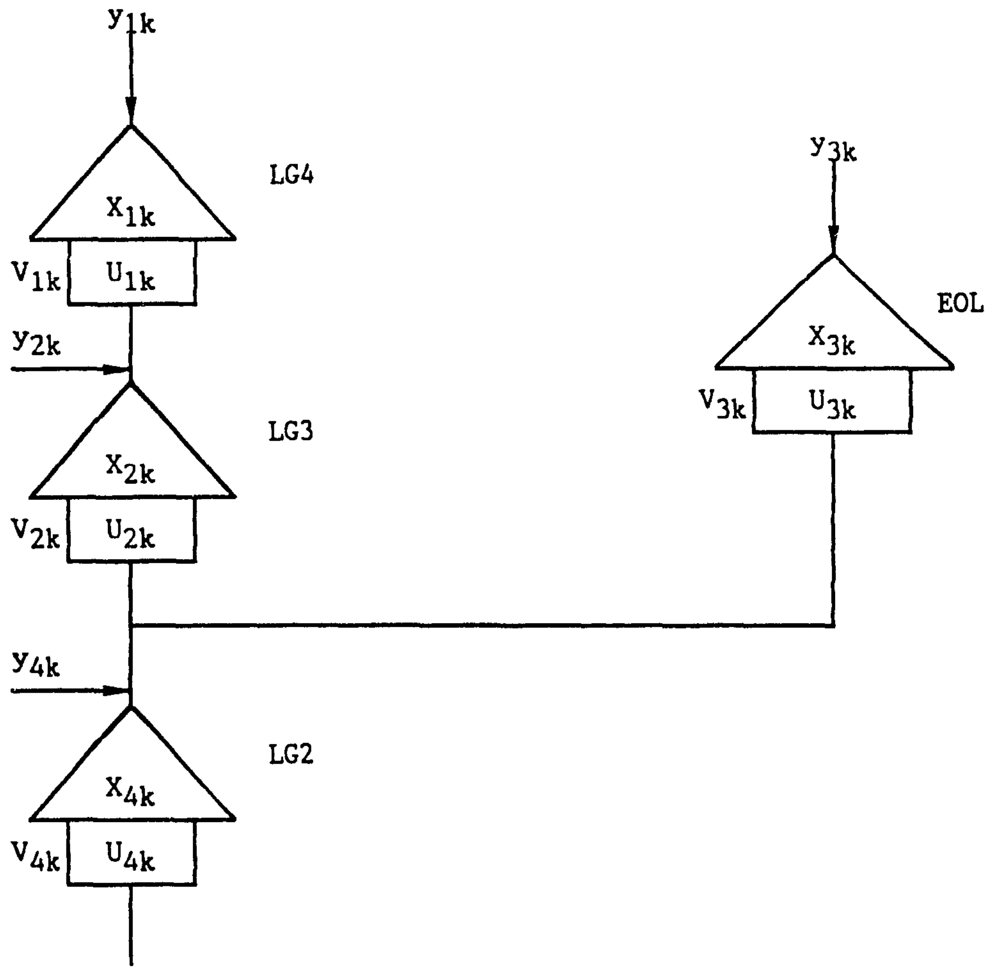


Figure 3.5 La Grande river intallations

Month	Inflow 1 (m^3/s)	Inflow 2 (m^3/s)	Inflow 3 (m^3/s)	Inflow 4 (m^3/s)	Demand (GWh)
May	1088	1095	2291	830	10275
June	1221	919	1348	800	8705
July	988	597	805	599	8563
August	993	686	1102	577	9316
September	1093	746	1196	630	9316
October	853	616	1005	560	11487
November	728	480	758	469	12327
December	553	368	513	323	16080
January	380	236	325	216	17530
February	289	179	242	162	14869
March	225	142	190	123	14869
April	164	119	209	89	11727

Table 3.2 Inflow and demand sequences.

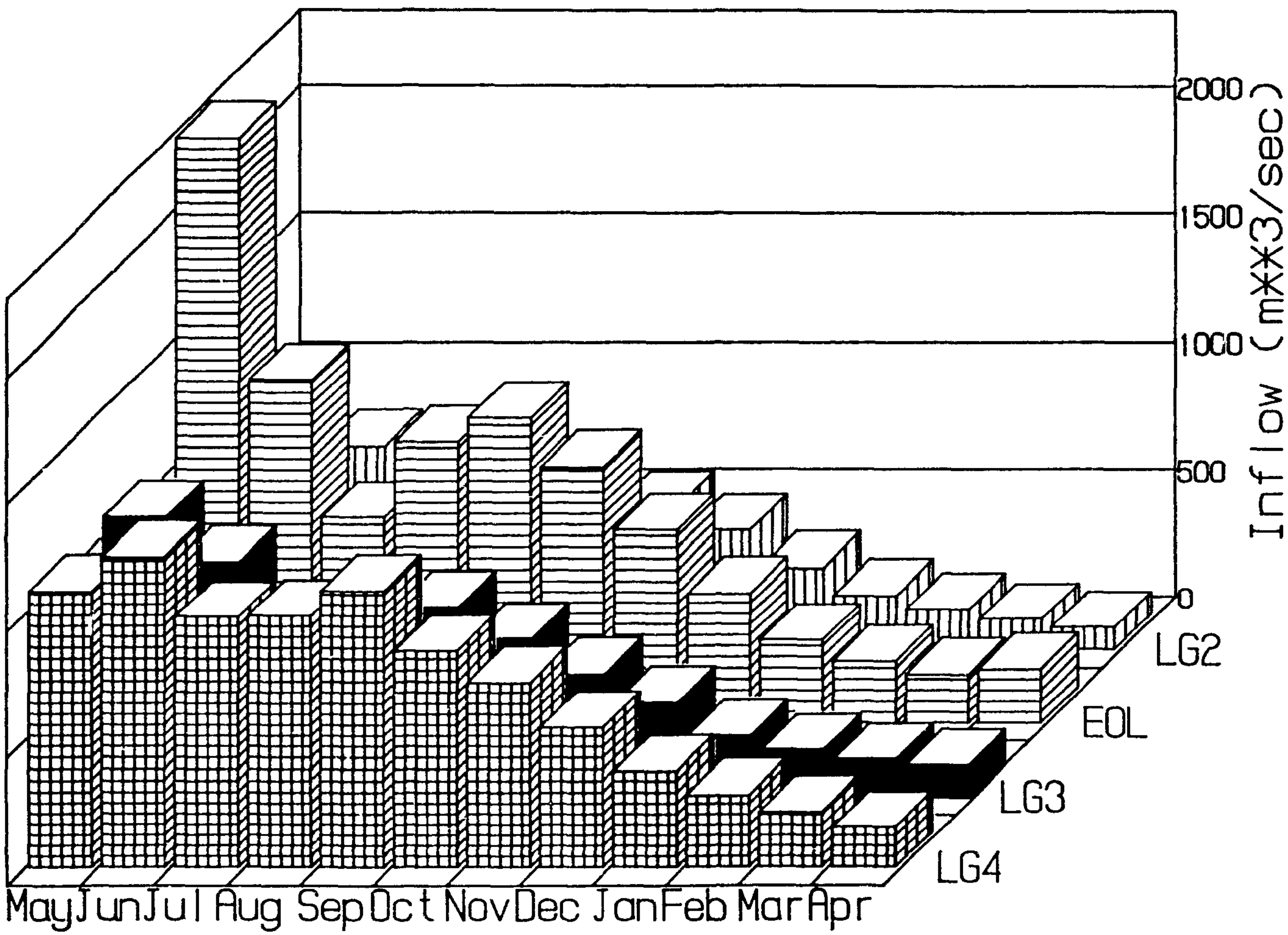


Figure 3.6 Inflow sequences

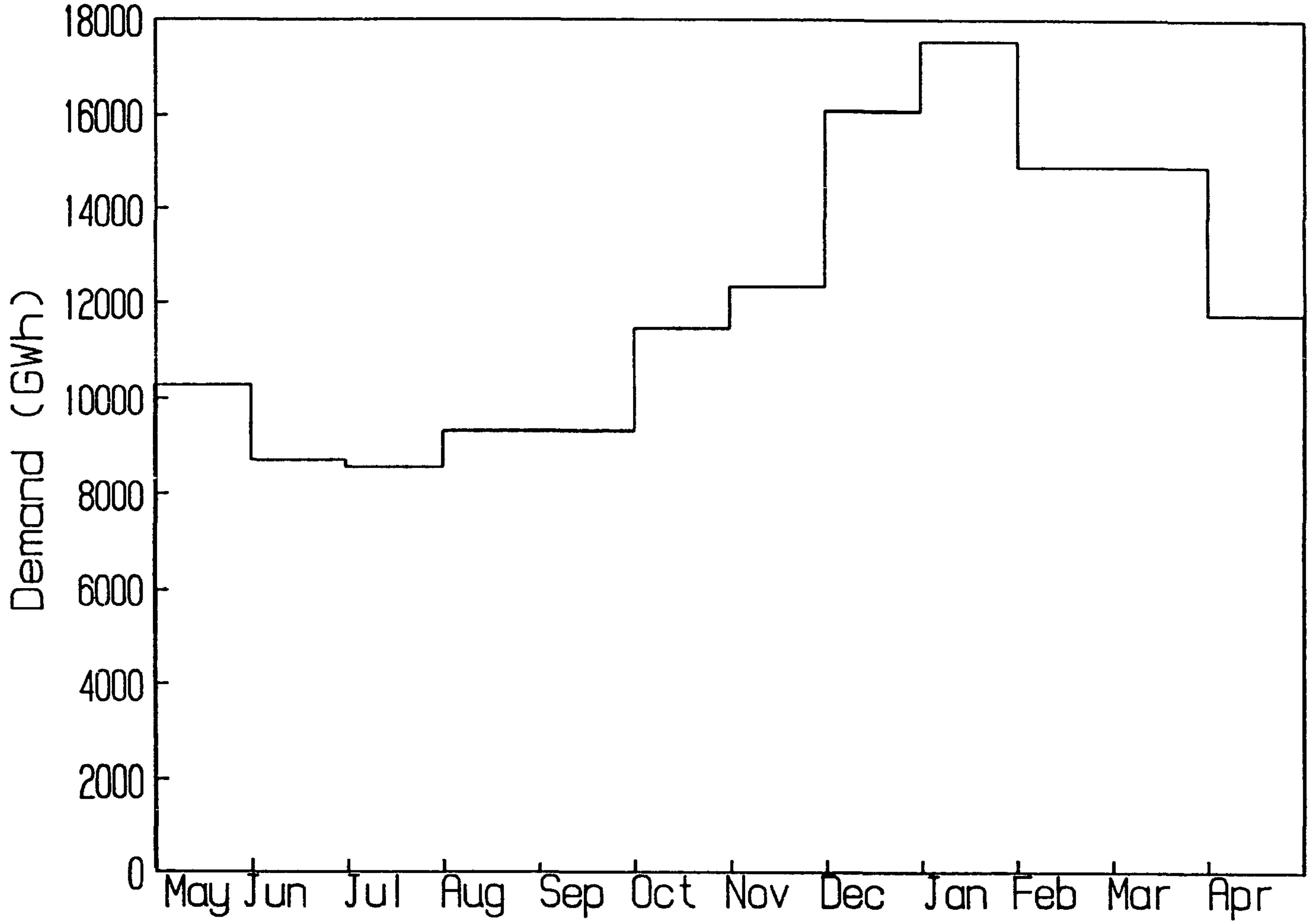


Figure 3.7 Demand pattern

Table 3.1.

The planning horizon considered here is a year with monthly periods starting in May. The reservoir levels at the beginning of May are set at their lower bounds shown in Table 3.1. The inflow sequences and the demand are given in Table 3.2 and illustrated in Figure 3.6 and Figure 3.7, respectively.

The optimization results will obviously depend on the choice of the constant α in the penalty term. There is no fixed rule to choose α but its value should remain small compared to the cost of the thermal energy. For this reason, different optimization results are shown for different values of α. Figures 3.8 and 3.9 illustrate the reservoir levels and the outflows when α is set equal to 0. It is clear that this case corresponds to the linear model of (2.18)-(2.26) in which no penalties are imposed. In this instance, big variations in storage levels and outflows can be observed from one period to the next. Nevertheless, these variations can be greatly reduced by choosing even just a small value for α. This is shown in Figures 3.10 and 3.11 where $\alpha = 0.02$. In that case, the water is stored from May to October, and then released from November to April. The situation is almost the same for $\alpha = 0.1$ as shown in Figures 3.12 and 3.13. However, larger values of α increase the value of the objective function. Consequently, for $\alpha = 0.02$, the thermal energy needed is 41373 GWh, whereas for $\alpha = 0.1$, it reaches 42318 GWh.

Figures 3.11 and 3.13 show high outflows (released water) from October to April. These periods correspond to the high energy demand, as illustrated in Figure 3.7. However, a very large portion of the natural inflows is received between May and October. That is the reason for the

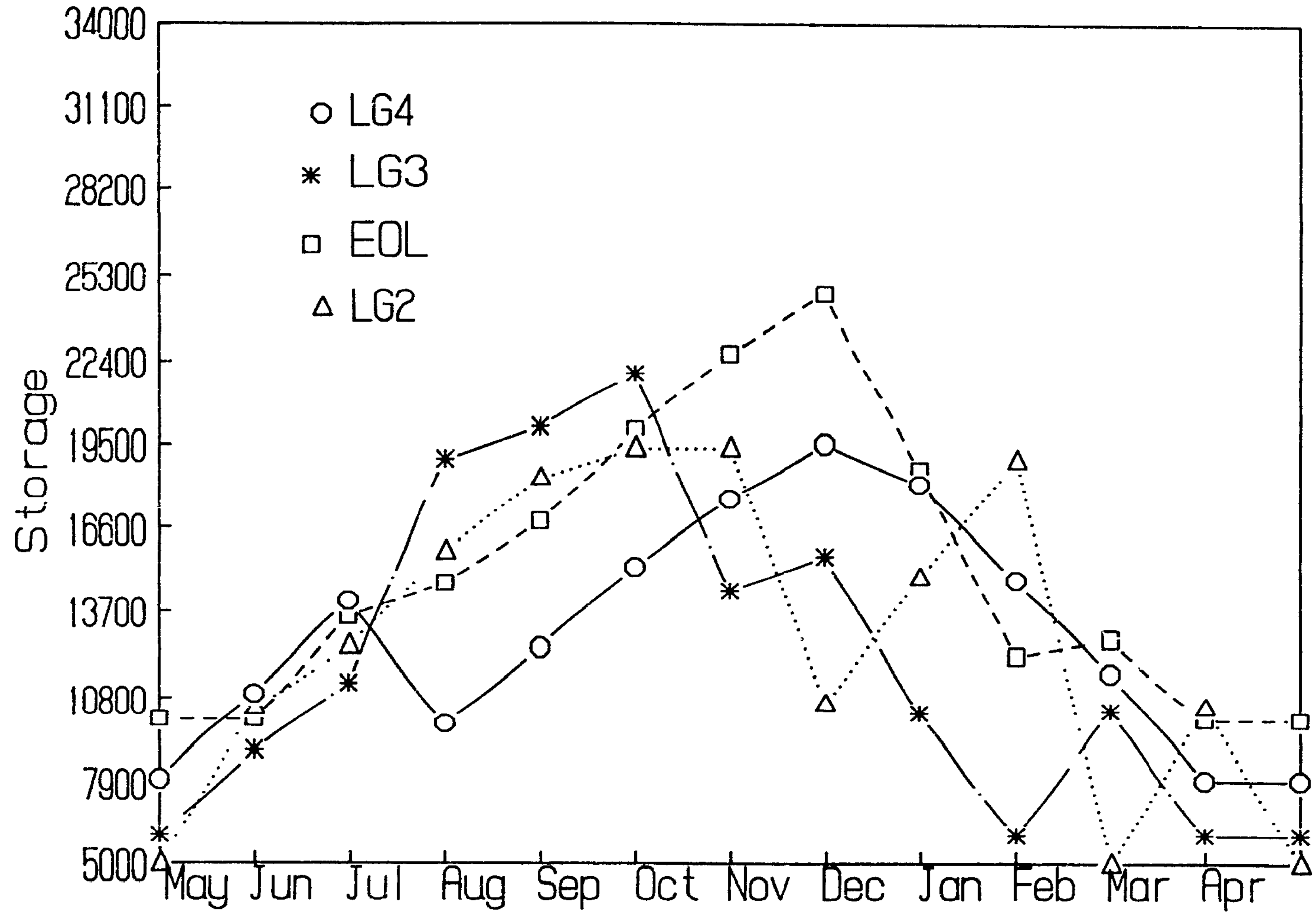


Figure 3.8 Storage for alpha = 0.0

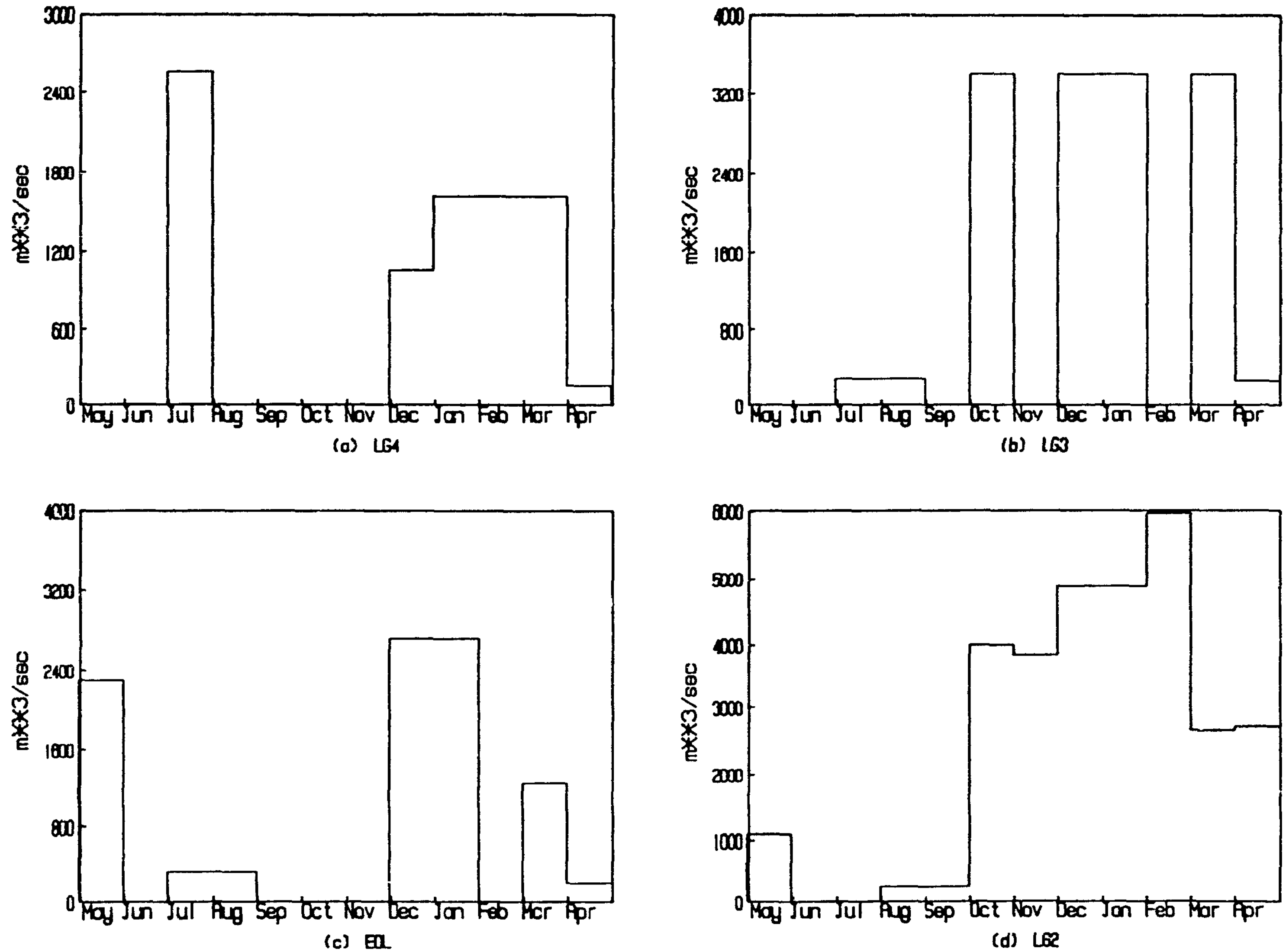


Figure 3.9 Outflow for alpha = 0.0

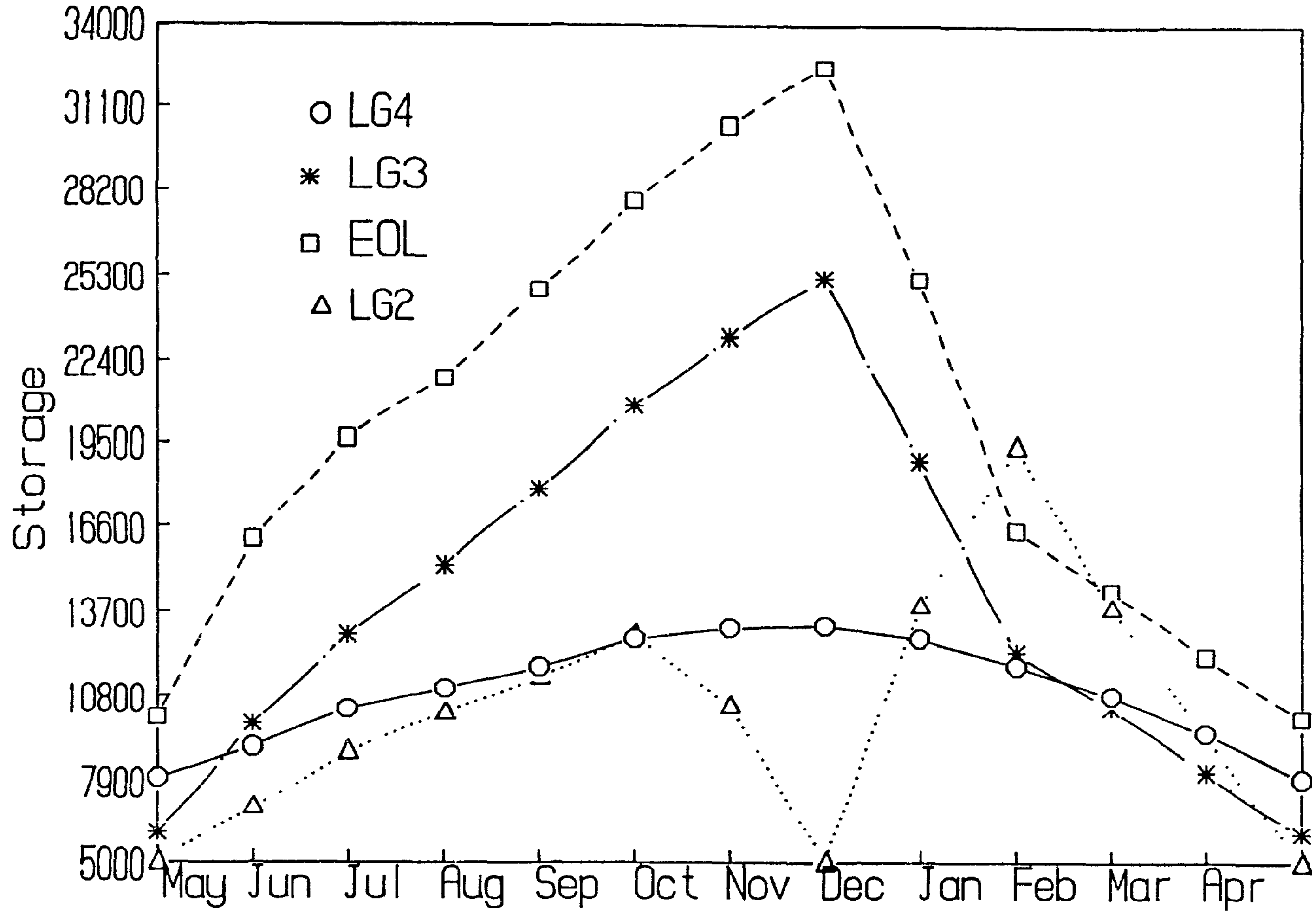


Figure 3.10 Storage for alpha = 0.02

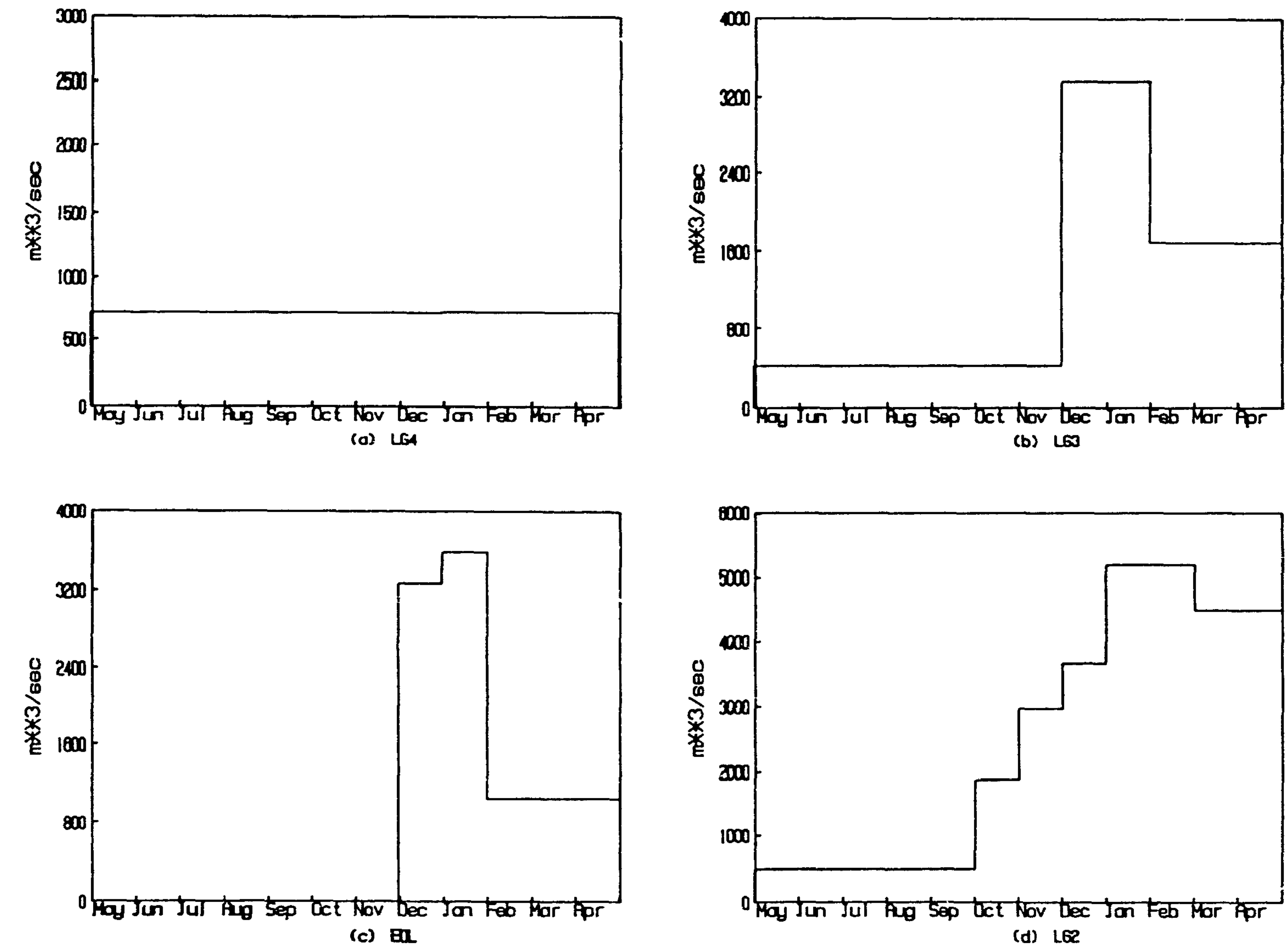


Figure 3.11 Outflow for alpha = 0.02

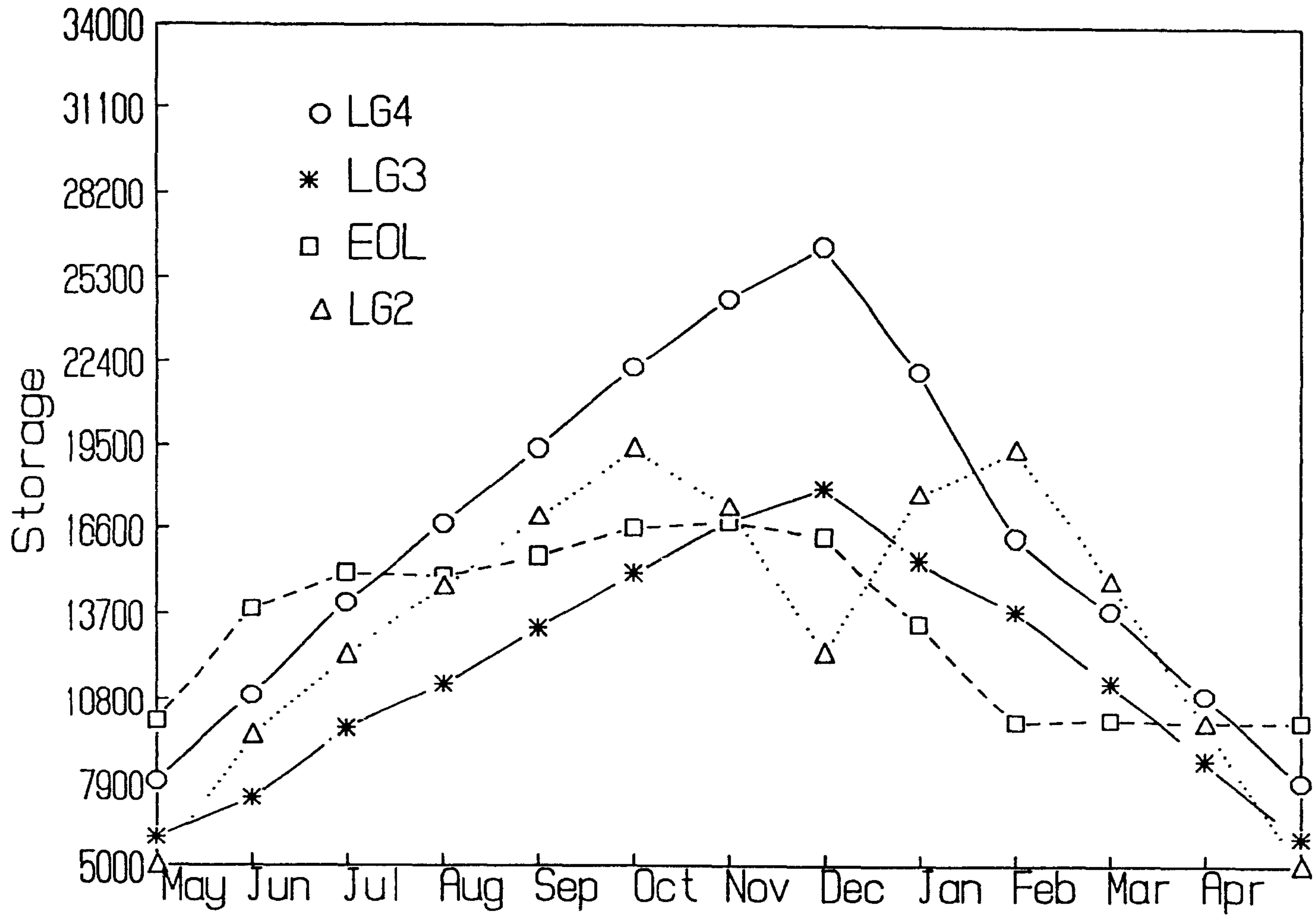


Figure 3.12 Storage for alpha = 0.1

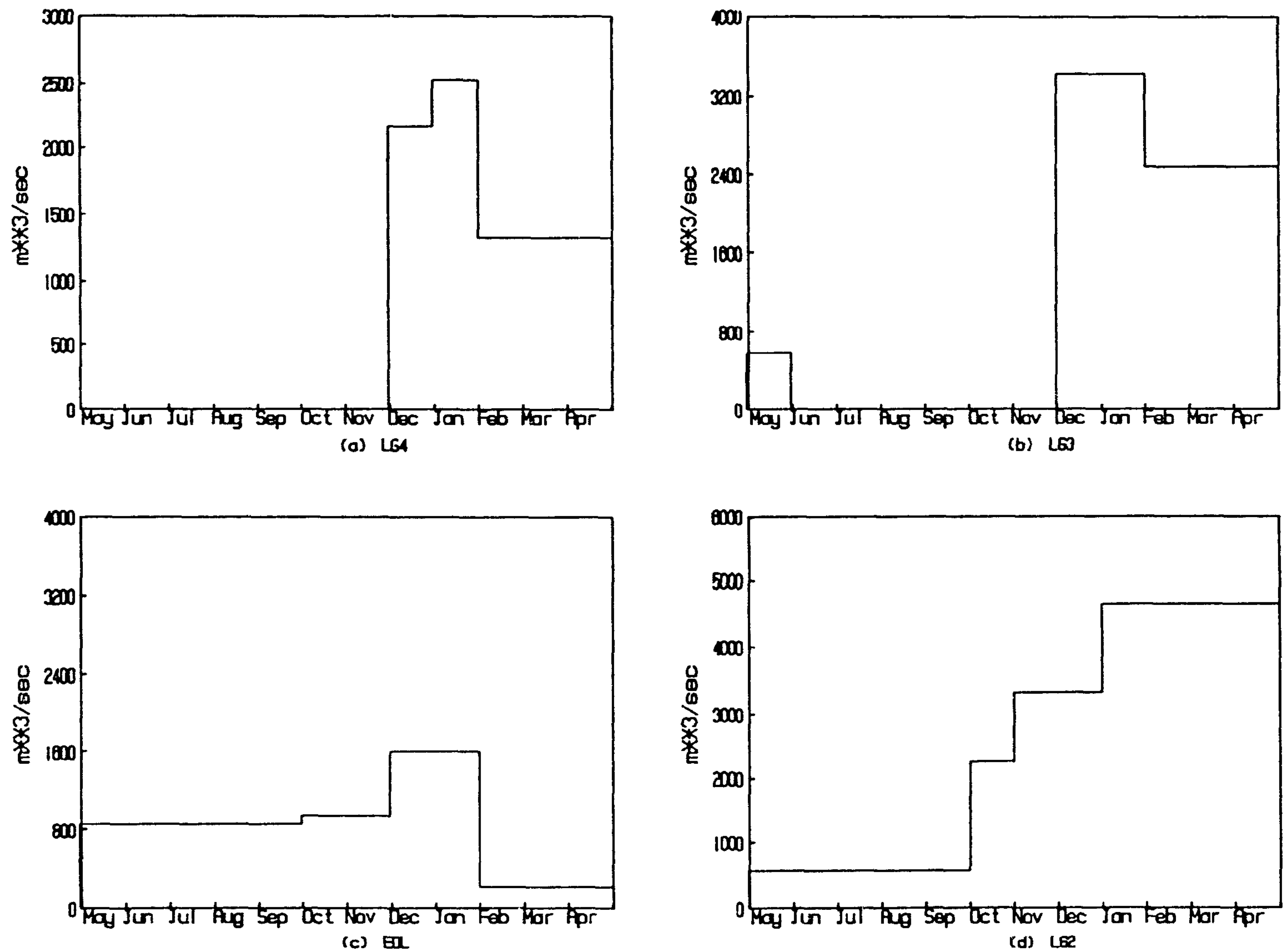


Figure 3.13 Outflow for alpha = 0.1

accumulation of water in the reservoirs during these periods, as shown in Figures 3.10 and 3.12.

It is not necessary to have the same α for all the reservoirs. Figure 3.14 illustrates a case where the α for reservoirs LG4, LG3, EOL and LG2 have been set equal to 0.001, 0.002, 0.001 and 0.003, respectively. It can be seen that the outflows are very high during the periods of peak demand, and low for the periods of high inflows. An important feature of Figure 3.14 is the smooth transition in operation characteristics between two successive periods, as compared to Figure 3.9 where α was set equal to 0 (that is without penalty terms) for each reservoir.

3.7 Conclusion

The solution methods for solving the deterministic reservoir problem, namely dynamic, nonlinear and linear programming, were explained in this chapter. The last technique was selected in order to be able to solve the problem with IBM's MPSX/370 package. The optimization process was found to be very fast. It took only 0.02 minutes of CPU time and 203 iterations on an IBM-3081 to solve the problem with an horizon of twelve periods.

A penalty factor α was introduced in the model to reduce the variations in the outflows. This constant cannot be determined according to fixed rules. For this reason, different optimization results for different values of α were shown. These results indicated that α should not exceed 10% of the cost of thermal energy, otherwise the objective function will be very far from its actual value with α set equal to zero.

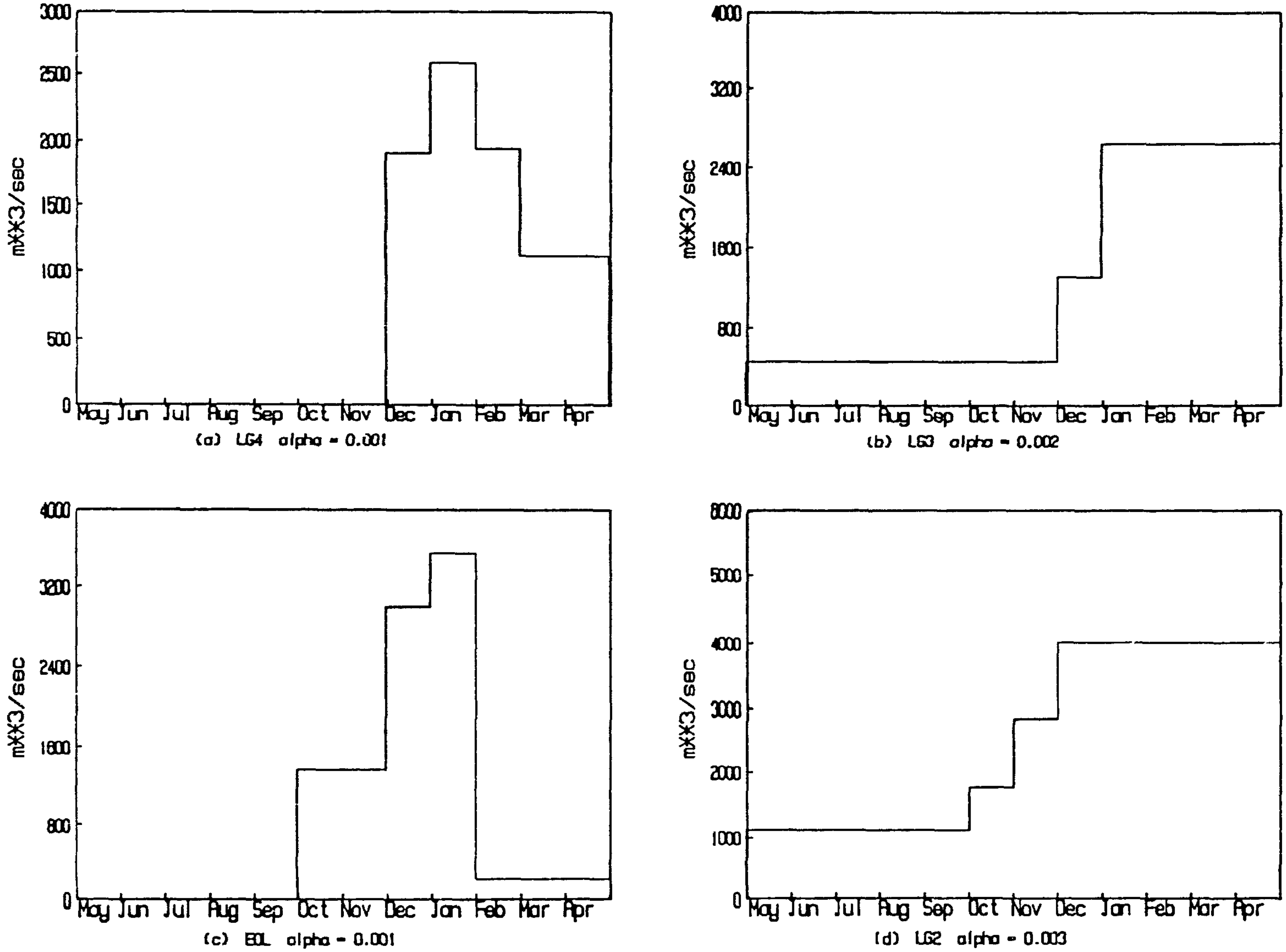


Figure 3.14 Outflow for different alpha

Finally, once α is fixed, this solution technique can be used to solve the problem for a large number of sequences of natural inflows and then the principal components analysis can be performed. This subject will be explained in the next chapter.

CHAPTER IV

PRINCIPAL COMPONENTS ANALYSIS (PCA)

4.1 Introduction

As in the implicit approach, the deterministic model developed in the previous chapter is solved for several inflow sequences. The results of these deterministic optimizations are then subjected to a Principal Components Analysis (PCA) to find out whether the problem could be modeled with a fewer number of variables.

Although n components are required to reproduce the total system variability, very often most of this variability can be accounted for by a small number p ($p<n$) of the principal components. In this case, the last n-p components, having a very small variance, can be replaced by their mean values without a significant effect on the solution of the problem.

The analysis of the principal components is often an intermediate step in much larger investigations. For example, principal components may be inputs to multiple regression [Johnson and Wichern, 1982; McCuen and Snyder, 1986], or, as in our case, inputs to stochastic dynamic programming.

This chapter presents the explanation and the application of the principal components analysis. In Section 4.2 a theoretical development is formulated to transform the original n var ables into a set of n new components. Section 4.3 describes the selection criteria for the number of components. Then Section 4.4 proposes an application of this theory to a four reservoir system. Finally, in Section 4.5, a discussion of the results is presented.

4.2 Principal Components Analysis

Assume that the deterministic optimizations for m flow sequences provide m values for the states X_{ik} ,$i=1,2,\ldots,n$, $k=1,2,\ldots,K$. Since PCA is applied to one period at a time, let us consider period k only, and denote by Z_i the optimal value of X_{ik},$i=1,2,\ldots,n$. Naturally, Z_i is a random variable since it depends on past river flows, which are random variables themselves.

Now let the random vector $Z^T=[Z_1, Z_2,\ldots,Z_n]$ have the expected vector $E(Z) = \mu$ and the covariance matrix $E(Z-\mu)(Z-\mu)^T = W$. This matrix is positive definite. Since principal components depend solely on the covariance matrix (or the correlation matrix) of Z, we can set, without loss of generality, $E(Z) = \mu = 0$.

The m values of Z_i are similar to those in table 4.1. Given those data, the goal of principal components analysis, as explained by Caillez [1984], Gnanadesikan [1977], Gendre [1976] and Kendall [1980], is to search for n linear combinations of the type:

$$\begin{aligned}
\xi_1 &= b_1^T Z = b_{11} Z_1 + b_{12} Z_2 + \ldots\ldots + b_{1n} Z_n \\
\xi_2 &= b_2^T Z = b_{21} Z_1 + b_{22} Z_2 + \ldots\ldots + b_{2n} Z_n \\
&\cdot \\
&\cdot \\
\xi_n &= b_n^T Z = b_{n1} Z_1 + b_{n2} Z_2 + \ldots\ldots + b_{nn} Z_n
\end{aligned} \tag{4.1}$$

or in a general form:

$$\xi_i = b_{i1}\ Z_1 + b_{i2}\ Z_2 + \ldots + b_{in}\ Z_n = \sum_{j=1}^{n} b_{ij}\ Z_j = b_i^T Z \qquad (4.2)$$

Variables

	Z_1	Z_2	....	Z_n
1	Z_{11}	Z_{21}		Z_{n1}
2	Z_{12}	Z_{22}		Z_{n2}
.	.	.		.
.	.	.		.
.	.	.		.
.	.	.		.
m	Z_{1m}	Z_{2m}		Z_{nm}

Table 4.1 Results of deterministic optimizations for period k.

These equations have the following characteristics:

i) ξ_1 has the largest possible variance;

ii) ξ_2 is orthogonal to ξ_1 (uncorrelated) and has the largest variance after ξ_1;

iii) ξ_3 is orthogonal to ξ_1 and ξ_2 and has the largest variance after ξ_1 and ξ_2;

iv) and so forth for the components $\xi_4, \xi_5, \ldots, \xi_n$.

In other words, the linear combinations presented above allow a set of variables: $Z_1, Z_2, \ldots, Z_n$, to be transformed into an equivalent set of variables: $\xi_1, \xi_2, \ldots, \xi_n$ that have two interesting attributes:

i) They are uncorrelated, and

ii) They are in decreasing order of variance.

Letting VAR denote the variance, and COV the covariance, it follows from equations (4.2) that:

$$VAR(\xi_i) = VAR(b_i^T Z) = E(b_i^T ZZ^T b_i) = b_i^T W b_i; \; i = 1,\dots,n \qquad (4.3)$$

$$COV(\xi_i,\xi_j) = COV(b_i^T Z, b_j^T Z) = E(b_i^T ZZ^T b_j) = b_i^T W b_j \quad i = 1,\dots,n \quad j = 1,\dots,n \qquad (4.4)$$

It is easy to show that the variances $VAR(\xi_i)$ are the diagonal elements of the covariance matrix.

The first principal component ξ_1 is equal to the linear combination with the maximum variance. That is, it maximizes $VAR(\xi_1) = b_1^T W b_1$. It is clear that VAR can be increased by multiplying any b_1 by some constant. To eliminate this indeterminacy, it is convenient to restrict attention to coefficient vectors of unit length. Thus, we define

i) First principal component = linear combination $b_1^T Z$ that maximizes $VAR(b_1^T Z)$ subject to the constraint $b_1^T b_1 = 1$.

ii) Second principal component = linear combination $b_2^T Z$ that maximizes $VAR(b_2^T Z)$ subject to $b_2^T b_2 = 1$ and $COV(b_1^T Z, b_2^T Z) = 0$.

iii) i^{th} principal component = linear combination $b_i^T Z$ that maximizes VAR $(b_i^T Z)$ subject to $b_i^T b_i = 1$ and COV $(b_i^T Z, b_j^T Z) = 0$ for $j < i$.

4.2.1 First Principal Component

The mathematical model for this component is obtained by choosing b_1 that maximizes

$$b_1^T W b_1 \tag{4.5}$$

Subject to

$$b_1^T b_1 = 1 \tag{4.6}$$

Using the Lagrangian method,

$$L(b_1) = b_1^T W b_1 - \lambda_1(b_1^T b_1 - 1) \tag{4.7}$$

$$\frac{\partial L}{\partial b_1} = 2 W b_1 - 2\lambda_1 b_1 = 0 \tag{4.8}$$

where λ_1 denotes the Lagrange multiplier unknown for the moment.

The solution to problem (4.5) - (4.6) is given by:

$$W b_1 = \lambda_1 b_1 \tag{4.9}$$

This last equation will no doubt be recognized as being the general form of an eigenvalue problem, in which b_1 is the eigenvector corresponding to the eigenvalue λ_1 of W. Since (4.6) must still be respected, (4.9) can be rewritten

$$\lambda_1 = b_1^T W b_1 = \mathrm{VAR}(b_1^T Z) = \mathrm{VAR}(\xi_1) \tag{4.10}$$

Choosing ξ_1 to maximize the variance means choosing the values of

$b_{11}, b_{12}, \ldots, b_{1n}$ composing the vector b_1 that corresponds to the largest eigenvalue λ_1 of W [Johnson and Wichern, 1982].

4.2.2 Second Principal Component

The value of b_2 will not be set equal to the value of the eigenvector corresponding to the second largest eigenvalue of W, because it is not known at this point whether the variable ξ_2 thus obtained will be independent of ξ_1. Instead, the following optimization problem must be solved.

$$\max b_2^T W b_2 \tag{4.11}$$

subject to

$$b_2^T b_2 = 1 \tag{4.12}$$

$$b_1^T W b_2 = 0 \tag{4.13}$$

Constraint (4.12) is added simply to ensure that the solution obtained is unique. Constraint (4.13), on the other hand, is set to guarantee independence between ξ_1 and ξ_2. It is important to note that

$$\mathrm{COV}\,(b_1^T Z,\ b_2^T Z) = 0$$

means that b_1 is orthogonal to b_2 or, in an equivalent form

$$b_1^T b_2 = 0 \tag{4.14}$$

Once again the Lagrangian method gives

$$L\,(b_2) = b_2^T\, W\, b_2 - \lambda_2\,(b_2^T\, b_2 - 1) - 2\mu\, b_1^T\, W\, b_2 \qquad (4.15)$$

where λ_2 and μ are the Lagrange multipliers.

$$\frac{\partial L}{\partial b_2} = 2W\, b_2 - 2\lambda_2\, b_2 - 2\mu\, W\, b_1 = 0 \qquad (4.16)$$

Multiplying (4.16) by b_1^T gives

$$b_1^T\, W\, b_2 - \lambda_2\, b_1^T\, b_2 - \mu\, b_1^T\, W\, b_1 = 0 \qquad (4.17)$$

Since (4.13) and (4.14) must hold, (4.17) can be rewritten:

$$\mu\, b_1^T\, W\, b_1 = 0 \qquad (4.18)$$

Since (4.10) must also remain true, (4.18) becomes

$$\mu\, \lambda_1 = 0 \qquad (4.19)$$

Remembering that λ_1 is the largest eigenvalue of W and is greater than zero, we can infer that

$$\mu = 0 \qquad (4.20)$$

Thus (4.16) can be rewritten as

$$W\, b_2 = \lambda_2\, b_2 \qquad (4.21)$$

It can also be noted that (4.21) has the same form as (4.9), so that the solution is

$$\lambda_2 = b_2^T W b_2 = VAR(b_2^T Z) = VAR(\xi_2) \qquad (4.22)$$

Therefore, b_2 can be set equal to the value of the eigenvector corresponding to the second largest eigenvalue of W.

It can be shown that b_3 can be set equal to the third largest eigenvalue of W, and so forth until the n^{th} component.

Therefore the matrix $B = [b_1, b_2, \ldots . b_n]$ will have the eigenvectors as its columns. An interesting property of matrix B is that it is orthonormal, which means that if

$$\xi = B^T Z$$

then $\qquad (4.23)$

$$Z = B \xi$$

Therefore, once the matrix B has been determined for every period k, the relations

$$X_{ik} = \sum_{j=1}^{n} b_{jik} \xi_{jk} \quad ; \quad i = 1,\ldots,n \ ; \ k = 2,\ldots K \qquad (4.24)$$

become known as well. The original problem with state variables X_{ik} can therefore be transformed into a problem with state variables ξ_{ik} using (4.24). Naturally, not much will be gained by doing so since this will simply transform a problem of n variables into another problem of n variables. Therefore, reduction criteria are needed. However, before establishing any of them, an important result should be emphasized.

4.2.3 Lemma

Suppose $Z^T = [Z_1, Z_2, \ldots, Z_n]$ has W as covariance matrix with the eigenvalue-eigenvector pairs $(\lambda_1, b_1), (\lambda_2, b_2), \ldots, (\lambda_n, b_n)$ where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0$. Let $\xi_1 = b_1^T Z$, $\xi_2 = b_2^T Z, \ldots, \xi_n = b_n^T Z$ be the principal components. It follows that

$$\sum_{i=1}^{n} VAR(Z_i) = \sigma_{11} + \sigma_{22} + \ldots + \sigma_{nn} = \sum_{i=1}^{n} VAR(\xi_i) = \lambda_1 + \lambda_2 + \ldots + \lambda_n$$

where σ_{ii}, $i=1,\ldots,n$ are the diagonal elements of the covariance matrix W.

The reader is referred to Appendix B for a demonstration of the lemma which states that

$$\text{total population variance} = \sigma_{11} + \sigma_{22} + \ldots + \sigma_{nn} = \lambda_1 + \lambda_2 + \ldots + \lambda_n \quad (4.25)$$

and consequently, the proportion η_i of the total variance due to the i^{th} principal component is

$$\eta_i = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \ldots + \lambda_n} \quad ; i=1,\ldots,n \quad (4.26)$$

If most (for instance, 80 to 90%) of the total population variance, for large n, can be attributed to the first p ($p<n$) components, then these p components can "replace" the original n variables without an appreciable loss of information.

4.2.4 Example

Suppose the random variables Z_1, Z_2, and Z_3 have the covariance matrix

$$W = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

It may be verified that the eigenvalue-eigenvector matrices are

$$\Lambda = \begin{bmatrix} 5.83 & 0 & 0 \\ 0 & 2.00 & 0 \\ 0 & 0 & 0.17 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.383 & 0 & 0.924 \\ -0.924 & 0 & 0.383 \\ 0 & 1 & 0 \end{bmatrix}$$

Therefore the principal components become

$$\xi_1 = b_1^T Z = 0.383\ Z_1 - 0.924\ Z_2$$
$$\xi_2 = b_2^T Z = Z_3$$
$$\xi_3 = b_3^T Z = 0.924\ Z_1 + 0.383\ Z_2$$

The variable Z_3 is one of the principal components because it is uncorrelated with the other two variables

Equation (4.10) can be demonstrated from

$$\begin{aligned} VAR(\xi_1) &= VAR(0.383\ Z_1 - 0.924\ Z_2) \\ &= (0.383)^2\ VAR(Z_1)+(-0.924)^2\ VAR(Z_2)+2(0.383)(-0.924)\ COV(Z_1,Z_2) \\ &=(0.147)(1) +(0.854)(5) - (0.708)(-2) = 5.83 = \lambda_1 \end{aligned}$$

and the independence among the components can also be shown by verifying that the covariance is equal to zero

$$\begin{aligned}
COV(\xi_1,\xi_3) &= COV(0.383\ Z_1-0.924\ Z_2, 0.924Z_1+0.383\ Z_2)\\
&= (0.383)\ (0.924)\ COV(Z_1,\ Z_1) + (0.383)^2\ COV(Z_1,\ Z_2) -\\
&\quad (0.924)^2\ COV(Z_1,\ Z_2) - (0.924)\ (0.383)\ COV(Z_2,\ Z_2)\\
&= (0.354)\ (1) + (0.147)(-2) - (0.853)(-2) - (0.354)(5) = 0
\end{aligned}$$

Moreover it is readily apparent that

$$\begin{aligned}
\sigma_{11} + \sigma_{22} + \sigma_{33} = 1 + 5 + 2 = \lambda_1 + \lambda_2 + \lambda_3 &= 5.83 + 2.000 + 0.17\\
&= 8
\end{aligned}$$

Finally, it can also be seen that the fraction of the total variance related to the first principal component is 73% (given by $\lambda_1/(\lambda_1+\lambda_2+\lambda_3) = 5.83/8 = 0.73$). Furthermore, the first two principal components account for a 98% ($(5.83+2)/8 = 0.98$) of the total variance. Since, in this case, the component ξ_3 has a very small variance, replacing it by its mean should not have a significant effect on the solution.

4.3 Selection of the Components

In the method proposed in this thesis, a criterion based on the component's percentage of the total variance is chosen, although other criteria were also proposed before.

1- Scree test of Cattell [Gendre, 1976]

This test starts by plotting the eigenvalues in descending order of magnitude, as shown in Figure 4.1. Then the components located to the left of the point where a significant change in slope occurs are

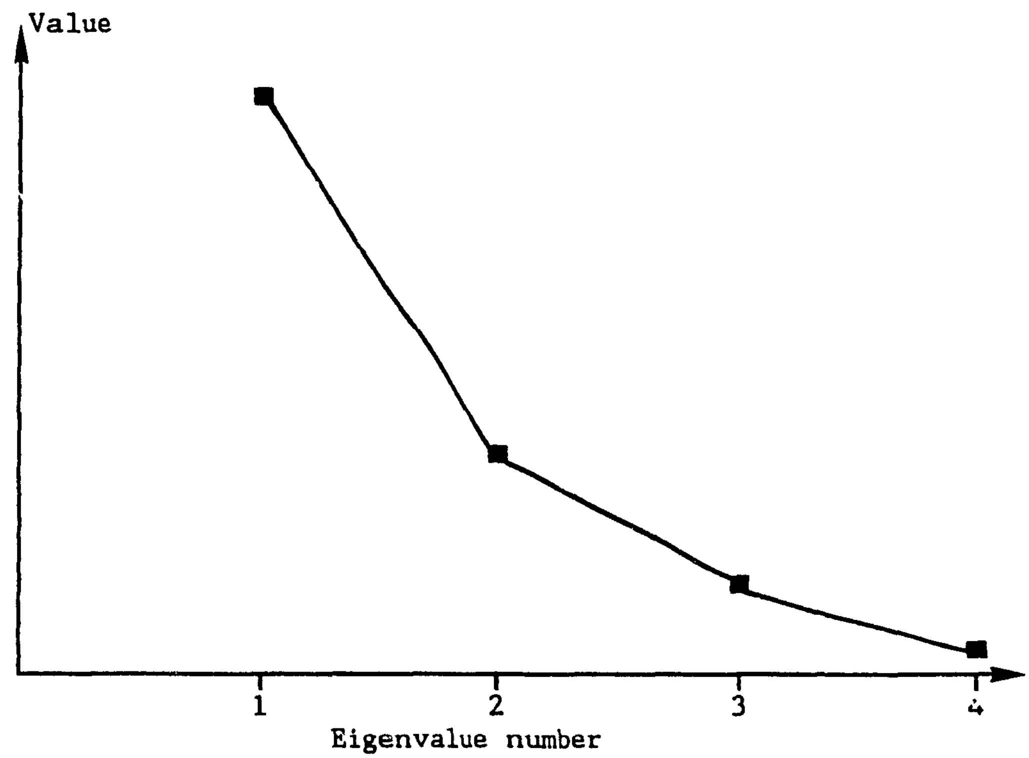


Figure 4.1 Scree test of Cattell

selected. In Figure 4.1, for instance, the first two components will be selected.

2- Kaiser's rule [Cooley and Lohnes, 1971]

Kaiser has stated a variety of compelling arguments for the selection of the p components corresponding to the larger than unity eigenvalues of the correlation matrix. This rule seems to work well when the number of sequences m is small or moderate. However for very large samples it may be worthwhile to take a value of p larger than the one prescribed by Kaiser's rule.

3- Bartlett's rule

Bartlett's rule provides a mean for verifying whether the determinant of the correlation matrix, after extraction of the p component, is zero or not. This, in turn, indicates when the factoring should stop. Then, after the components corresponding to the roots λ_1, $\lambda_2, \ldots, \lambda_p$ have been extracted, we have:

$$\chi^2_{0.5(n-p)(n-p-1)} = -[(m-1) - \frac{1}{6}(2n+5) - \frac{2}{3}p] \ln X_{n-p} \qquad (4.27)$$

where

$$X_{n-p} = \frac{|R|}{\{\prod_{j=1}^{p} \lambda_j [(n - \sum_{j=1}^{p} \lambda_j)/(n-p)]^{n-p}\}}$$

and $|R|$ is the determinant of the correlation matrix

n is the number of variables

p is the number of components selected

m is the total number of sequences (the samples of data).

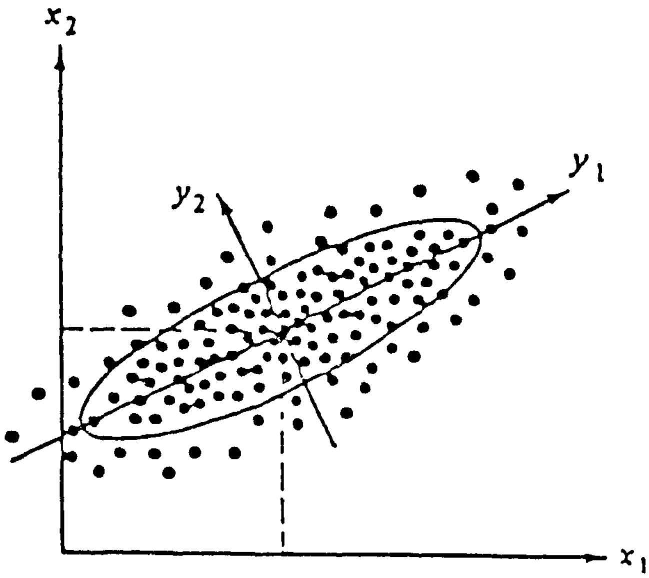
It should be mentioned that this rule is the extension of Bartlett's sphericity test.

$$\chi^2_{0.5(n^2-n)} = -[(m-1)- \frac{1}{6}(2n+5)] \ln |R|$$

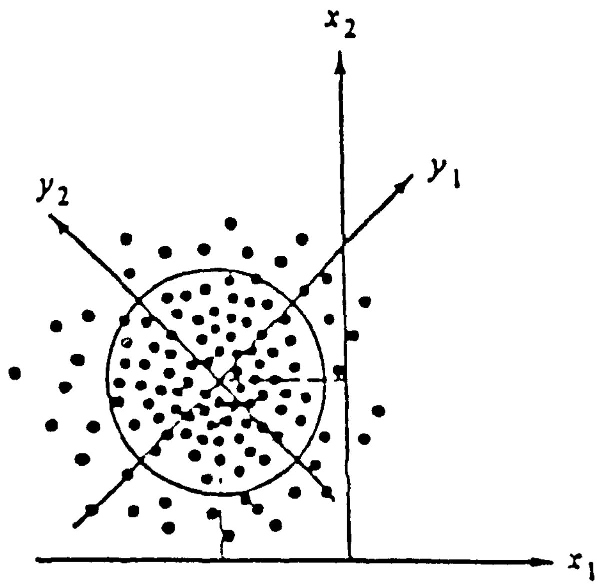
A nonsignificant chi-square χ^2 at some reasonable $0 \leq \alpha \leq 1$ level indicates that the matrix should not be factored since the vector variable may already be treated as a set of uncorrelated elements.

Geometrically, the data can be plotted as a set of m points in a n-dimensional space. Moreover, if the covariances or the correlations are very important, the data are then within an ellipsoid centered at the mean values E[X] of the n variables. Otherwise, they are within a sphere centered at E[X]. Figure 4.2(a) shows the case of a 2-dimensional space with high correlation. In this case the eigenvalue λ_1 is greater than λ_2. Therefore the principal components are well determined. They lie along the axes of the ellipse in directions perpendicular to those of maximum variance. Figure 4.2(b) shows the case of weak correlation. Here the eigenvalues λ_1 and λ_2 have almost the same values. Therefore, the axes of the ellipse, or of the circle in this case, are not uniquely defined and can lie in any perpendicular directions, including the directions of the original coordinate axes. Thus the principal components can lie in any two perpendicular directions, including those of the original coordinate axes. When the eigenvalues of the correlation or covariance matrix are nearly equal, the variation is homogeneous in all directions. It is not possible then to represent the data in fewer than n dimensions.

To illustrate the criterion for the selection of the components, let



(a)



(b)

Figure 4.2 Principal component
(a) $\lambda_1 > \lambda_2$
(b) $\lambda_1 = \lambda_2$

us assume that almost all the data are located within an ellipsoid. This means that the first few components of ξ_k account for a very large percentage of the variance of the observations, and that the number of variables can consequently be reduced. For example, suppose that ξ_{1k}, the first component of ξ_k, accounts for 85% of the total variance of the n variables, ξ_{2k} for 10%, and the remaining n-2 for 5% only. Then it can be inferred that each of the last n-2 components has a very small variance. Therefore replacing these components by their mean values should not have a significant effect on the solution. Thus equation (4.24) can be approximated as

$$X_{ik} = \sum_{j=1}^{2} b_{jik}\,\xi_{jk} + \sum_{j=3}^{n} b_{jik}\,\mu_{jk}$$

$$= \sum_{j=1}^{2} b_{jik}\,\xi_{jk} + b_{oik} \qquad (4.28)$$

where μ_{jk} is the mean of ξ_{jk} and b_{oik} a constant equal to

$\sum_{j=3}^{n} b_{jik}\,\mu_{jk}$.

Finally, the problem of dimension n is reduced to a two-dimensional problem.

4.4 Sample Application

Principal components analysis will now be applied to determine the reduced number of state variables for the installations shown in Figure 3.5. The first step is to solve the deterministic model developed in Chapter III for m different flow sequences. Since the historical flow record consists of thirty years of monthly inflows at each site, streamflow synthesis is used to provide several equally likely future sequences of streamflows (50 sequences). For each streamflow sequence, the problem is first solved deterministically to find the optimal operating policy. The results of these deterministic optimizations are then subjected to a principal component analysis to find out the number of components to retain.

4.4.1 Streamflow Synthesis

The statistical analysis of the historical flow record shows that:

i) The inflow at site i during month k is correlated to that of month k-1 but not to those of the previous months;

ii) The inflow at two different sites during month k are highly correlated.

As a result, a Markovian linear synthetic inflow generator of the form

$$Y_k = A_k + B_k Y_{k-1} + C_k W_k \qquad (4.29)$$

was used [Pronovost, 1974], where Y_k represents a column vector of random variables of inflows for month k, W_k a column vector of white noise and A_k a column vector of constants; B_k and C_k are square matrices. The coefficients A_k, B_k and C_k are determined from the historical record of inflows. Then it can be shown that (see Appendix B.2 for the demonstration)

$$A_k = M_k - P_k^T R_{k-1}^{-1} M_{k-1} \qquad (4.30)$$

$$B_k = P_k^T \; R_{k-1}^{-1} \tag{4.31}$$

$$C_k \, C_k^T = R_k - P_k^T \, R_{k-1}^{-1} \, P_k \tag{4.32}$$

where $M_k = E[Y_k]$ represents a column vector containing the mean values of Y_k, $R_k = E[(Y_k-M_k)(Y_k-M_k)^T]$ the correlation matrix for different sites for month k, and $P_k = E[(Y_{k-1}-M_{k-1})(Y_k-M_k)^T]$ the correlation matrix for the different sites, for months k-1 and k.

Unfortunately equation (4.32) does not give the matrix C_k directly. However it is easy to show that (see appendix B.2)

$$C_k = L_k \, \Lambda_k^{1/2} \tag{4.33}$$

where L_k is the matrix of eigenvectors for $C_k \, C^T_k$ and Λ_k is the diagonal matrix of eigenvalues for $C_k \, C^T_k$.

A flow chart of this process is shown in Figure 4.3 in which synthetic inflows are generated for one year on a monthly basis. The historic flow record used to determine the coefficients A_k, B_k and C_k spans from 1950 to 1979 (Table C.1 in Appendix C) whereas the synthetic inflow generated spans from year 1 to year 50 as shown in Table C.2

Finally, it is important to note that the statistical characteristics of the synthetic inflows are very close to those of the historic. Thus, even if the sequences themselves show big differences, their mean values are preserved as illustrated in Figures (4.4)-(4.7). In addition, it will be shown later in this chapter that the correlation among the sites is very high as expected.

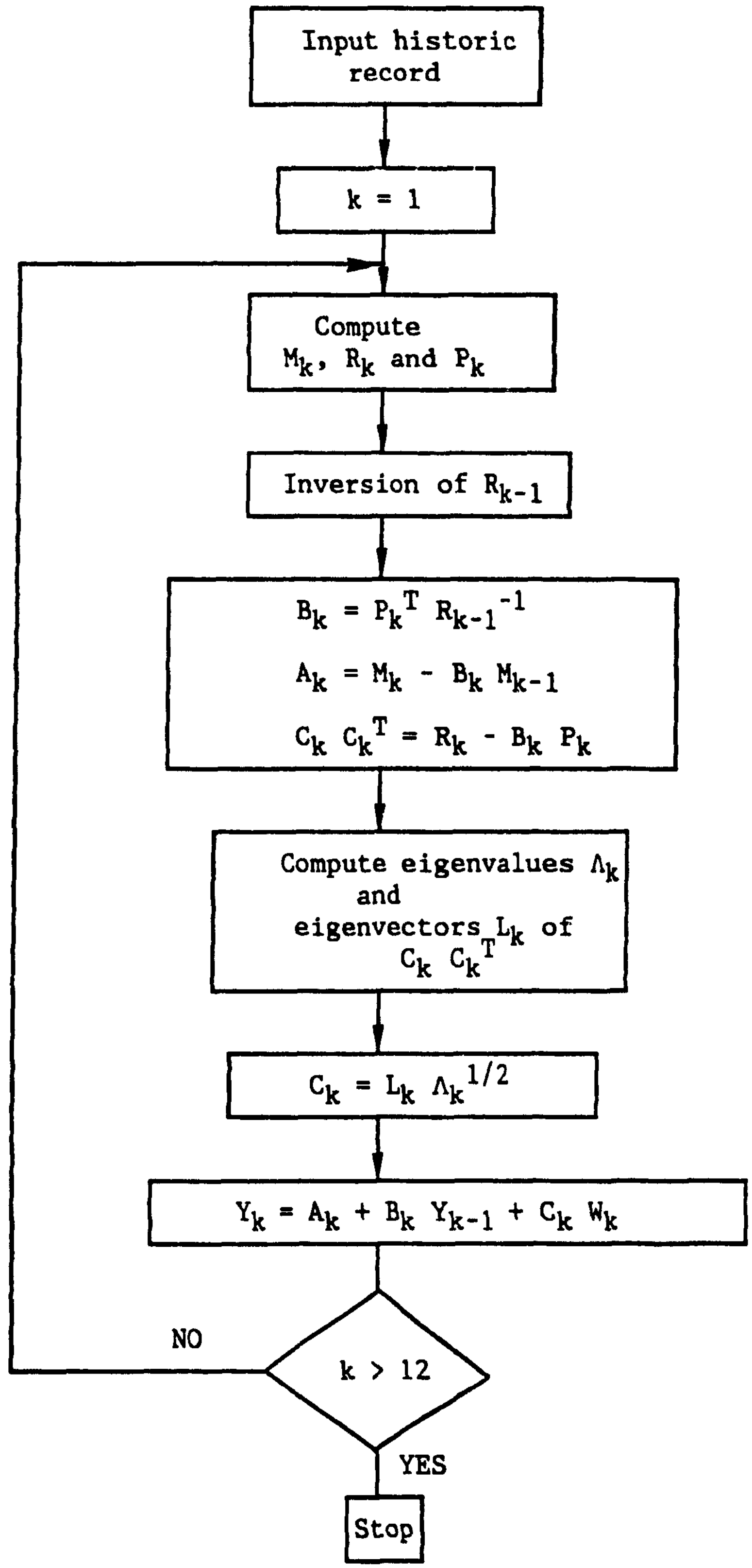


Figure 4.3 Flowchart of the synthesis generator

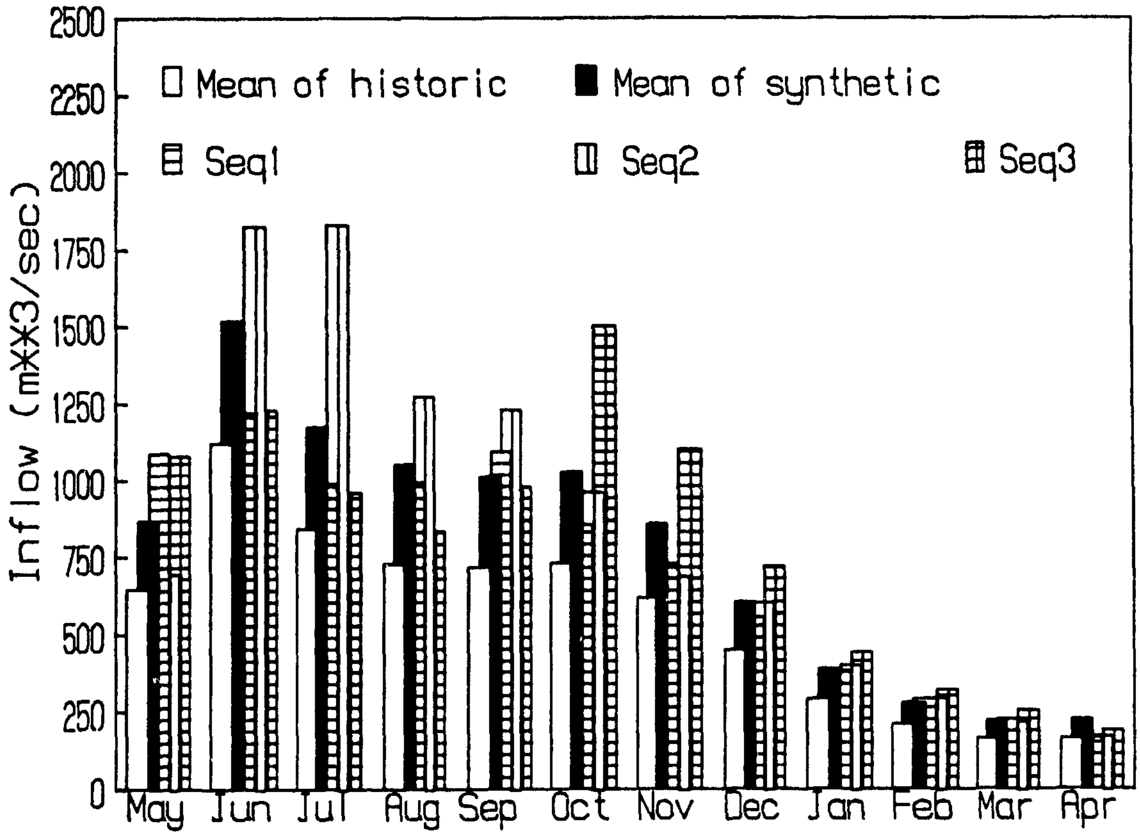


Figure 4.4 Historic and synthetic inflows for LG4

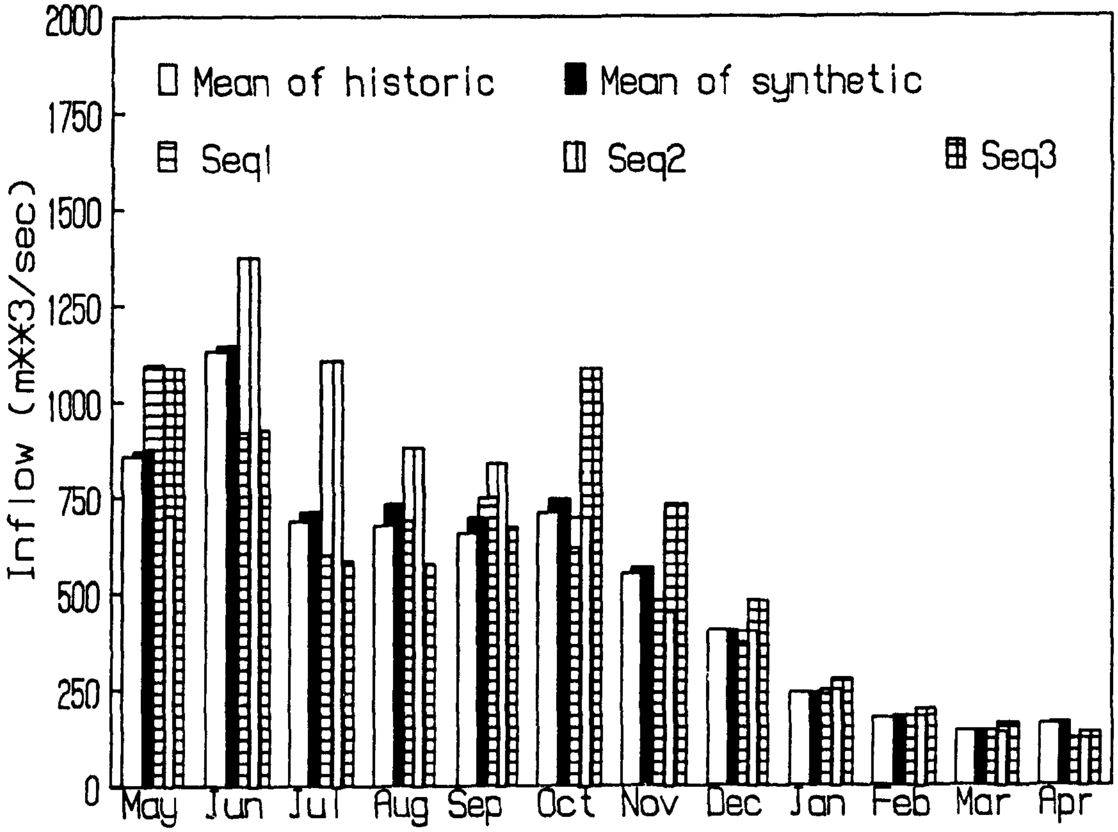


Figure 4.5 Historic and synthetic inflows for LG3

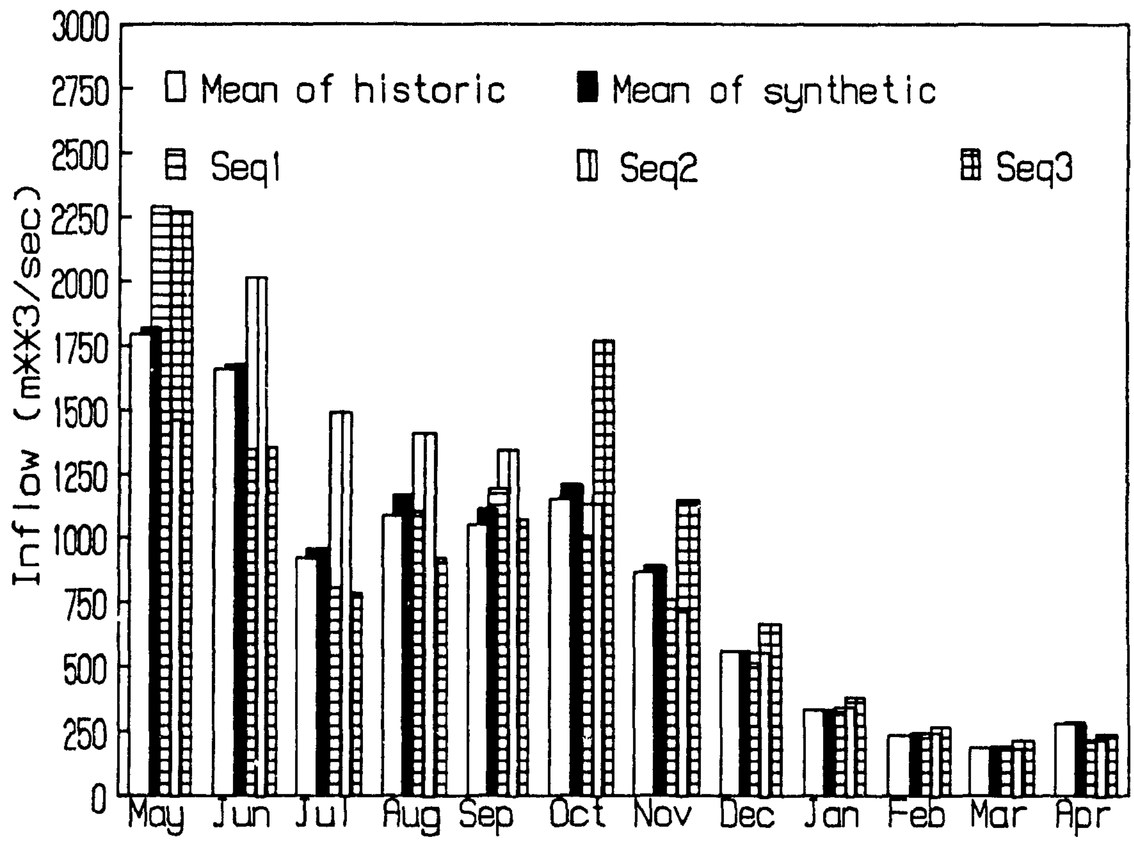


Figure 4.6 Historic and synthetic inflows for EOL

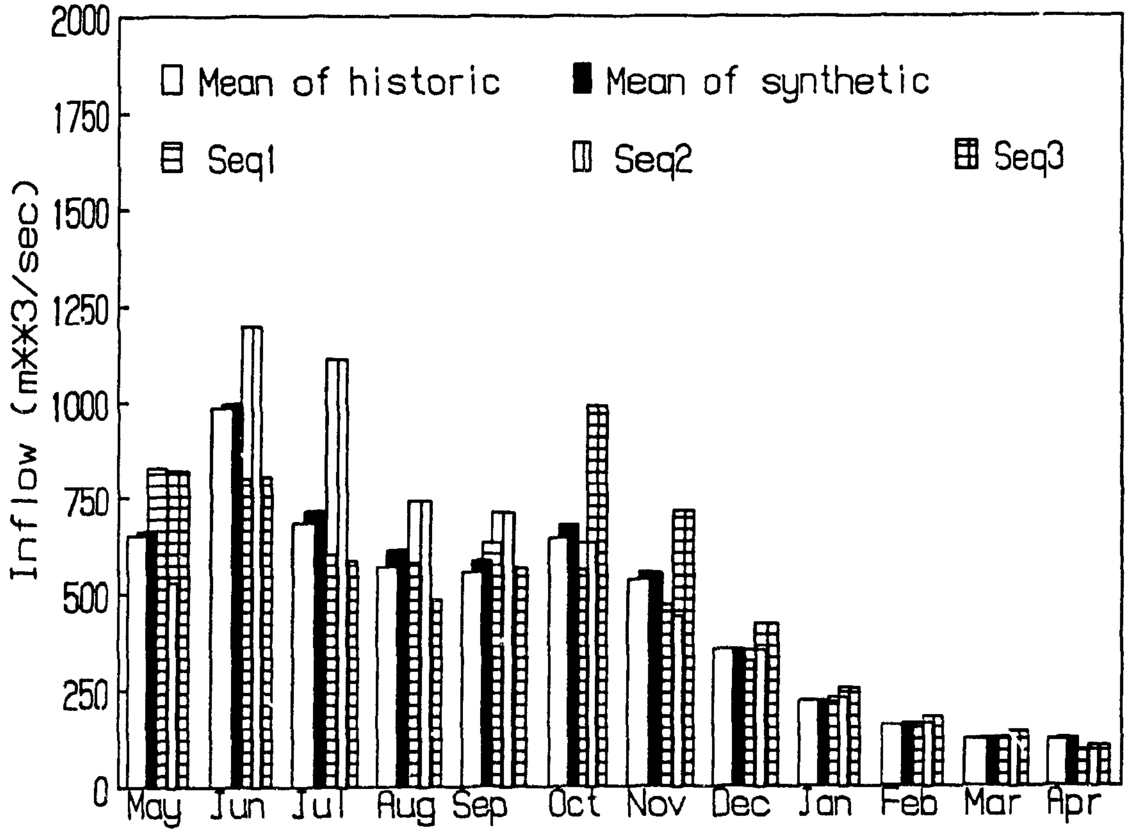


Figure 4.7 Historic and synthetic inflows for LG2

4.4.2 Deterministic Optimization for the 50 Inflow Sequences

The synthetic inflow sequences generated are used as inputs for the IBM's MPSX/370 package. The time horizon for the optimization process is one year on a monthly basis. Finally, for the fifty deterministic optimizations, the initial storage levels are set at their lower bounds as shown in Table 4.2.

Reservoir	Initial Storage (hm^3)
LG4	8000
LG3	6000
EOL	10000
LG2	5000

Table 4.2 Initial storages

The other characteristics (lower and upper storage levels, the efficiency of the plants and the demand) are shown in Tables 3.1 and 3.2. The penalty constant α is set at 0.01 or 1% of the cost attributed to the thermal energy generated. The flowchart for this optimization process is shown in Figure 4.8 while the optimal storage levels are represented in Table C.3. Figures 4.9-4.12 illustrate the first four optimal storage sequences of Table C.3. On those figures, the variations in the storage levels due to the differences in the corresponding inflow

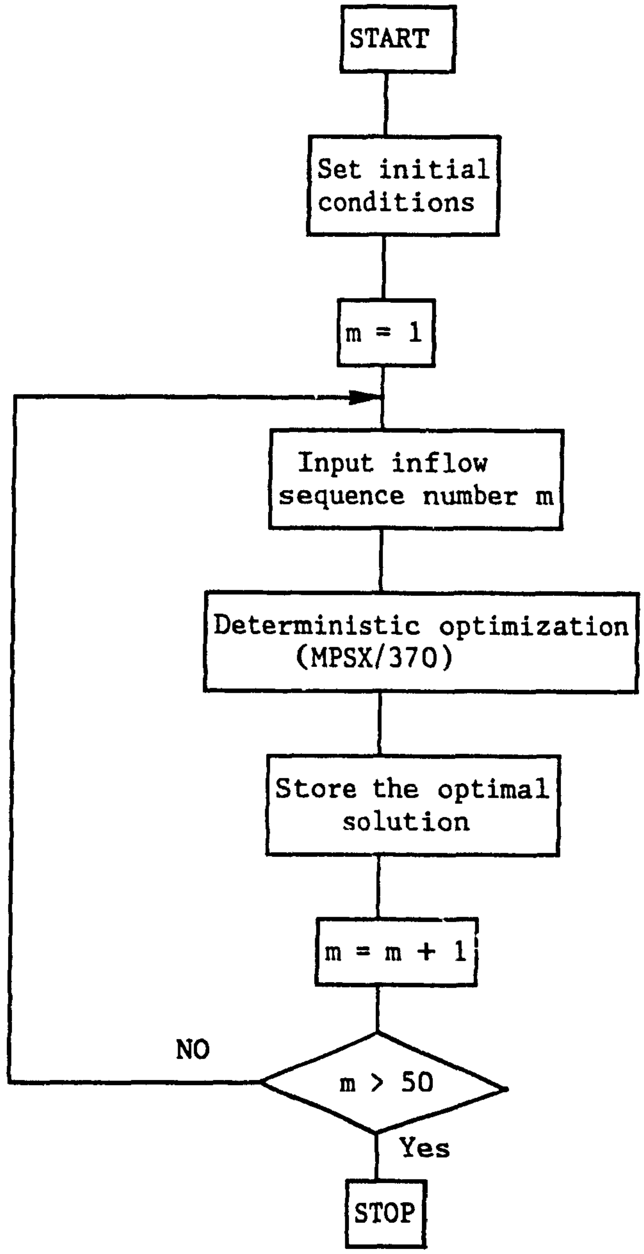


Figure 4.8 Flowchart of the 50 optimizations process

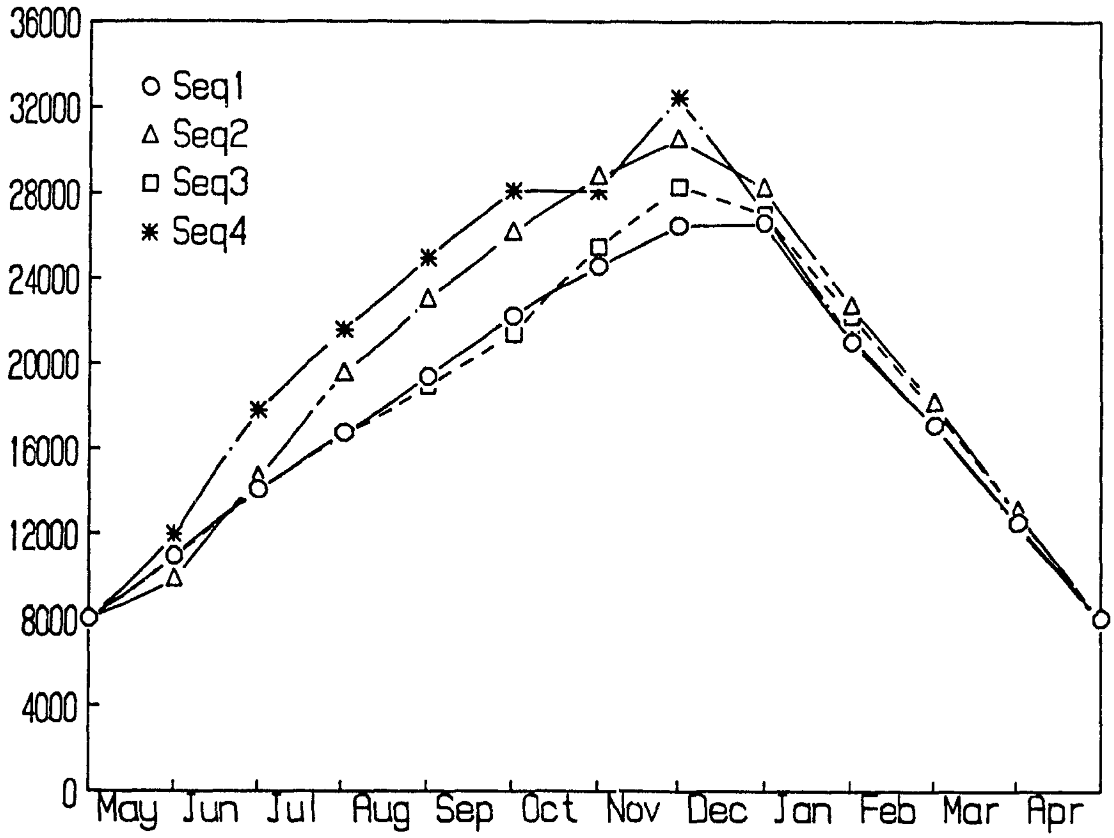


Figure 4.9 Optimal storages for LG4

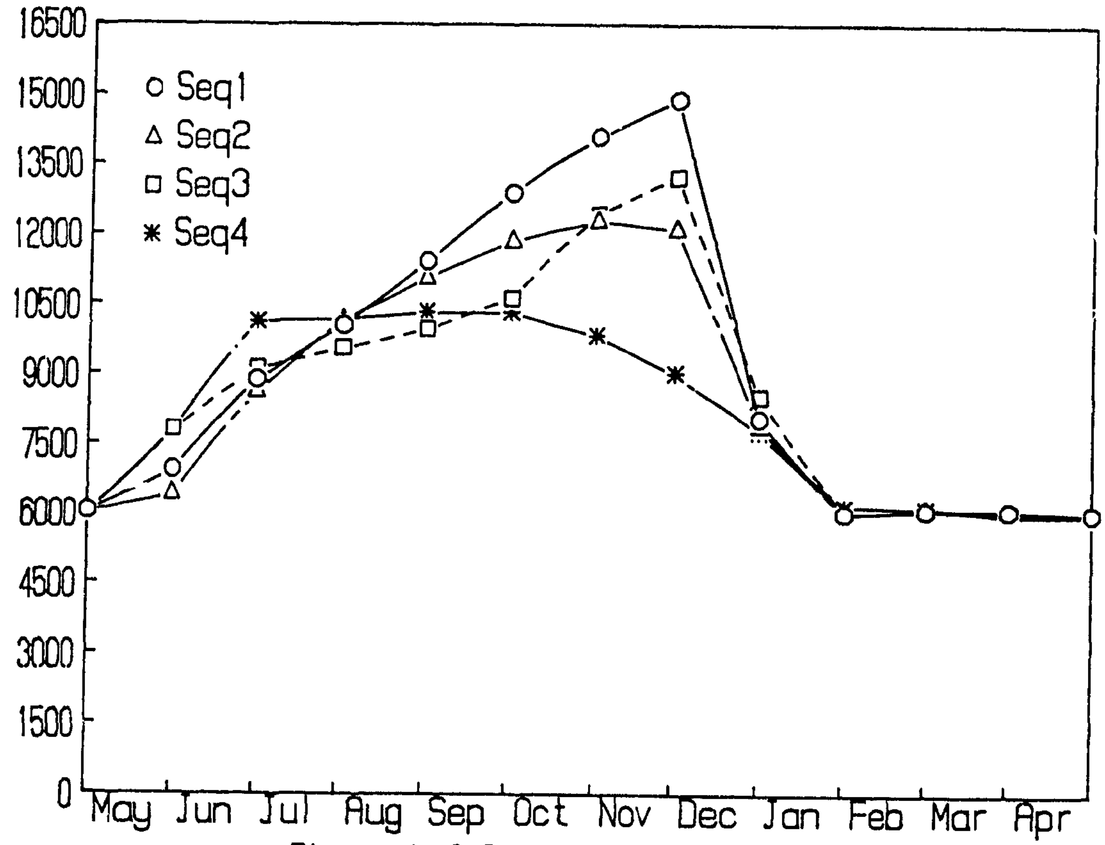


Figure 4.10 Optimal storages for LG3

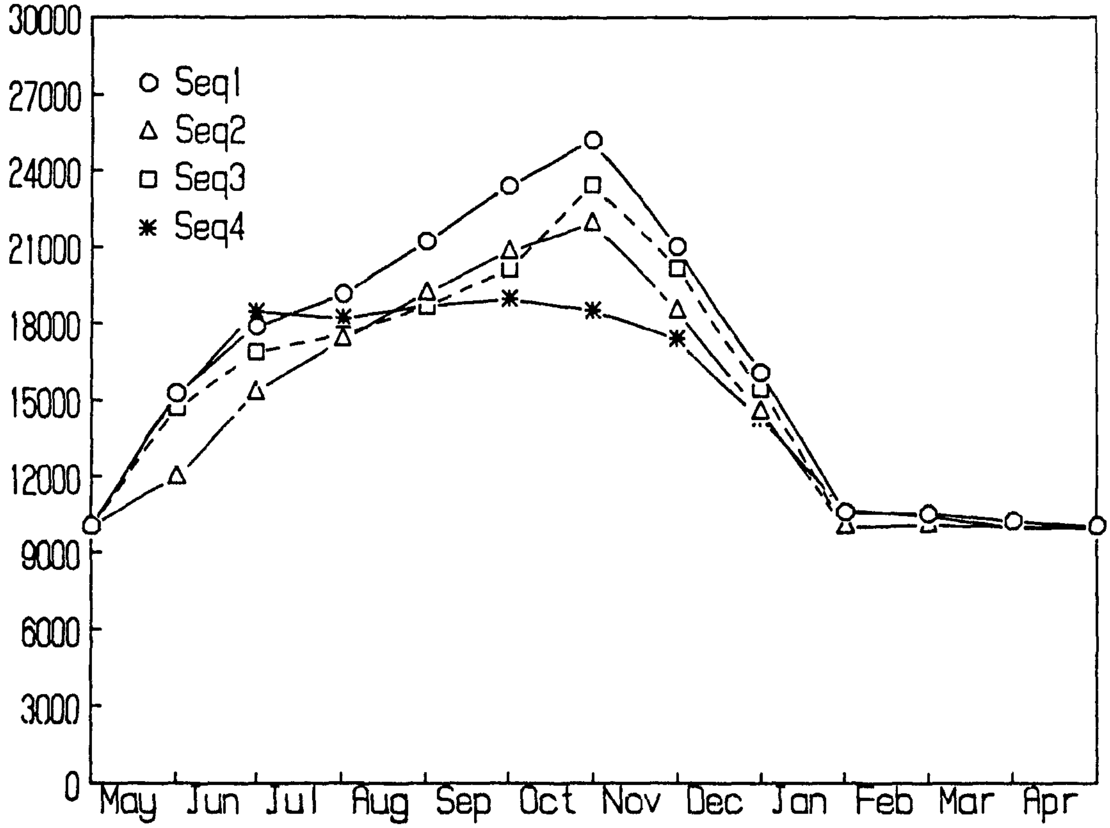


Figure 4.11 Optimal storages for EOL

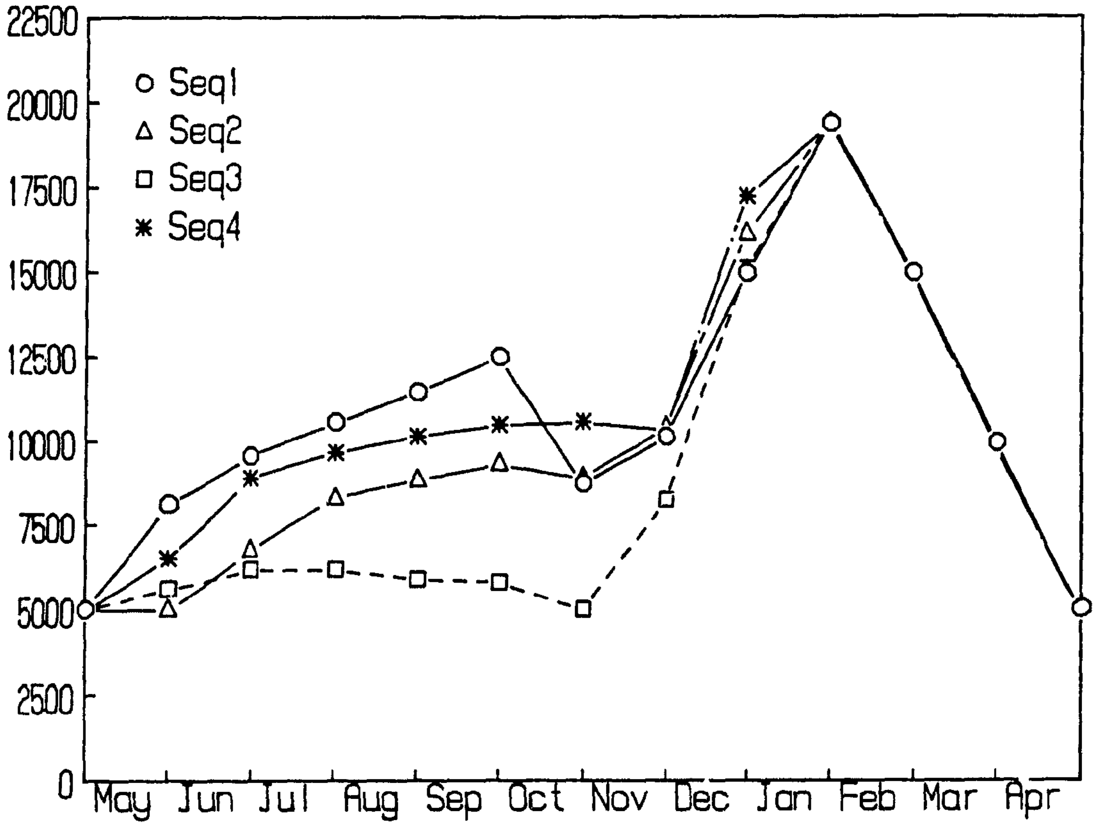


Figure 4.12 Optimal storages for LG2

sequences can easily be seen.

The number of iterations needed to obtain the optimal solution varies with the inflow sequences. For example, the optimal solution for sequence 1 is obtained after 205 iterations or 0.02 minutes, while for sequence 7, 249 iterations are needed with a cumulative time of 0.13 minutes. Finally, for the fifty inflow sequences, a cumulative CPU time of 0.97 minutes is needed to perform the required iterations.

4.4.3 Principal Components

In this example, principal components analysis was applied not only to the reservoir trajectories but also to the inflows since they are correlated in time. In other words, because of relation (4.29), the state variables for period k are $X_{1k},\ldots, X_{4k}$ and $Y_{1k-1},\ldots, Y_{4k-1}$. Hence, applying principal components analysis to the X's only will not reduce the number of state variables sufficiently to permit straightforward application of dynamic programming. Principal components analysis was therefore applied to the Y's also, albeit separately from the X's. For instance, for the month of May, the percentage contributions of the four components to the total variance of the sample of inflows are:

100.000   0.000   0.000   0.000

and the corresponding vector of eigenvalues and matrix of eigenvectors are:

$$\lambda = [\ 533351 \quad 0.096 \quad 0.078 \quad 0.062]$$

$$P = \begin{bmatrix} 0.3773 & -0.7371 & 0.0961 & -0.5524 \\ 0.3796 & 0.2869 & -0.8373 & -0.2693 \\ 0.7943 & 0.3852 & 0.4573 & 0.1081 \\ 0.2876 & -0.4754 & -0.2838 & 0.7814 \end{bmatrix}$$

In other words, let Z_{1k}, Z_{2k}, Z_{3k} and Z_{4k} be the four new components. For the first of these components, Z_{1k}, the percentage contribution to the total variance is 99.9988. Moreover, Z_{1k} is related to the previous variables, $Y_{1k},\ldots,Y_{4k}$, as follow

$$Z_{1k} = 0.3773\ Y_{1k} + 0.3796\ Y_{2k} + 0.7943\ Y_{3k} + 0.2876\ Y_{4k} \quad (4.34)$$

In the same way, the percentage contributions of Z_{2k}, Z_{3k} and Z_{4k} are, respectively, 0.000629, 0.000384 and 0.000187 and are related to the previous Y's as follow

$$Z_{2k} = -0.7371\ Y_{1k} + 0.2869\ Y_{2k} + 0.3852\ Y_{3k} - 0.4754\ Y_{4k} \quad (4.35)$$

$$Z_{3k} = 0.0961\ Y_{1k} - 0.8373\ Y_{2k} + 0.4573\ Y_{3k} - 0.2838\ Y_{4k} \quad (4.36)$$

$$Z_{4k} = -0.5524\ Y_{1k} - 0.2693\ Y_{2k} + 0.1081\ Y_{3k} + 0.7814\ Y_{4k} \quad (4.37)$$

In other form, (4.34) - (4.37) can be written as

$$Z_k = P^T\ Y_k \quad (4.38)$$

Then, if Y_k = [865 870 1821 660] m^3/sec

$$Z_k = [2292.9 \quad -0.303 \quad 0.110 \quad 0.462] \; m^3/sec$$

It is important to note that, due to the orthonormal property of P, Y_k can easily be obtained when Z_k is known from

$$Y_k = P \, Z_k \tag{4.39}$$

Therefore, in our example, Z_{1k} is equal to 2292.9 m^3/sec with a variance of 533351 (the first eigenvalue λ_1), Z_{2k} is equal to -0.303 m^3/sec with a variance of 0.000 (the second eigenvalue λ_2) and so forth for the remaining components.

Similarly, for every period k, the eigenvalues and the corresponding eigenvectors are determined. The percentage contributions of the new components to the total variance can then be deduced. Table 4.3 shows the percentage contribution of the four components of the sample of inflows, whereas Table 4.4 shows the percentage contributions of the sample of storages.

In Table 4.3, the last three components obviously have a very small variance. Therefore, the inflows to the four sites can be expressed as a function of the first component only. For the storages shown in Table 4.4, the percentage contributions of the first three components is at least 96% of the total variance. Hence, these components can be kept as random variables. As a result, principal components analysis has transformed the original problem of eight state variables into a problem of four state variables, thereby making it solvable by dynamic

	Component 1	2	3	4
May	100.000	0.00000	0.00000	0.00000
June	100.000	0.00000	0.00000	0.00000
July	99.9999	0.00005	0.000037	0.000027
August	99.9999	0.000044	0.000029	0.000026
September	99.9999	0.000038	0.000032	0.000027
October	99.9998	0.000065	0.000044	0.000042
November	99.9998	0.00010	0.000093	0.000063
December	99.9991	0.00043	0.00024	0.00018
January	99.9972	0.0010	0.00096	0.0008
February	99.9922	0.00358	0.00257	0.00168
March	99.9838	0.00674	0.00488	0.00458
April	99.9977	0.00099	0.00074	0.000569

Table 4.3 Contributions of the four components to the total variance of the sample of inflows.

	Component 1	2	3	4
May	64.225	32.173	2.513	1.090
June	61.742	28.662	6.086	3.51
July	59.272	27.07	9.777	3.878
August	63.46	20.364	12.488	3.689
September	64.40	18.079	13.503	4.01
October	65.349	22.309	10.015	2.325
November	84.033	11.173	3.940	0.854
December	88.696	8.14	2.306	0.858
January	88.363	8.190	3.027	0.419
February	92.351	5.533	1.749	0.366
March	88.116	7.616	3.806	0.463
April	25.00	25.00	25.00	25.00

Table 4.4 Contributions of the four components to the total variance of the sample of storages

programming.

However, two important remarks must be made. First the criterion for the selection of the number of components is not fixed. It would consequently have been possible to select only two components for the storages and to solve a problem of three rather than four state variables. The second remark concerns the storages at the beginning and the end of the horizon. Since the storages at the beginning of May and the end of April are the lowest, principal components analysis is not necessary because the state variables are known for these periods.

Finally, the computation process for this technique is very fast. The total CPU time required to find the principal components for the twelve period problem was less than one second on an IBM-3081.

4.5 Discussion

The sample of available data (number of sequences) is an important factor in principal components analysis. In other words, by using a population of data, the exact statistical characteristics can be extracted. Since it is impractical or uneconomical to observe a population as a whole, it is usually necessary to use a sample or a fraction of this population. Clearly, such a sample can be useful only if it is, in some way, "representative" of the population from which it has been derived. For this requirement to be met, two conditions must be satisfied. First, a set of observations $Z_{1j}, Z_{2j},\ldots, Z_{mj}$; $j = 1,\ldots,n$ constitutes a random sample of size m from a finite population of size M, only if each subset m of the population has the same probability of being selected. The second condition states that the results must be independent of the number of samples taken (20, 50 or even more). The first condition is easily satisfied since the synthetic generation provides several equally likely future sequences of streamflows. Satisfying the second condition however is more difficult and requires the following. First, the sample size m must be greater than the number of variables n to avoid any degeneracy [Johnson and Wichern, 1982, p. 108]. In fact, it is easy to show that if the sample size m is less than the number of variables n, the determinant of the covariance matrix is zero. The second requirement will be stated, without loss of generality, for unidimensional systems. If a random sample of size m is taken from a population having the mean μ and the variance σ^2, then

$$\mu_z = (\sum_{i=1}^{m} z_i)/m$$

is the value of a random variable whose distribution has the mean μ, and the variance σ^2/m (infinite population). [Miller and Freund, 1977, p. 165].

Although it is not very surprising that the mean of the theoretical sampling distribution of μ_z equals the mean of the population, the fact that its variance equals σ^2/m, for random samples from infinite populations, is interesting and important. To point out the implications of this rule, let us apply Chebyshev's theorem to the sampling distribution of μ_z. We thus obtain

$$\Pr\ (|\mu_z - \mu|) < \frac{L\sigma}{\sqrt{m}}\ ) \geq 1 - \frac{1}{L^2} \tag{4.40}$$

which states that the probability (Pr) of getting a value within L standard deviations ($\sigma/\sqrt{m}$) of the mean is at least

$$1 - \frac{1}{L^2}$$

Letting L $\sigma/\sqrt{m} = \varepsilon$, we get

$$Pr\ (|\mu_z - \mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{m\varepsilon^2} \tag{4.41}$$

Thus, for any given $\varepsilon > 0$, the probability that μ_z differs from μ by less than ε can be made arbitrarily close to 1 by choosing m sufficiently large. In less rigorous language, the larger the sample size, the closer will μ_z be to the mean of the population. In this sense we can say that μ_z becomes more and more reliable as an estimate of μ as the sample size is increased. The reliability of μ_z as an estimate of μ is often measured by the expression $\sigma/\sqrt{m}$, also called the standard error of the mean. Usually it does not pay to take excessively large samples since the extra labor and expense is not accompanied by a proportional gain in reliability. Table 4.5 illustrates the mean μ_z and the standard deviation σ_z of the inflows for different sample sizes m. Since it is impossible, in our example, to know exactly the standard deviation of the population, the reliability of μ_z is measured by $\sigma_z/\sqrt{m}$. Then if m equals 50, the standard error for LG4 is $276/\sqrt{50} = 39$ m^3/sec. For m = 30 the standard error is $268/\sqrt{30} = 48.9$ m^3/sec, and for m equals to 15, $\sigma_z/\sqrt{m}$ raises to $280/\sqrt{15} = 72.29$ m^3/sec. Therefore, it is suggested to use a sample size greater than 30 for LG4.

In addition to the sample size, another factor is of great importance in principal components analysis. This factor concerns the choice of the covariance matrix or the correlation matrix of the observations for the analysis. As Kendall [1980] pointed out, it is not an easy choice. For instance, in the case where some variables have a much larger variance than others, then the principal components will be

m	50		40		30		15	
	μ_z (m^3/sec)	σ_z (m^3/sec)	μ_z (m^3/sec)	σ_z (m^3/sec)	μ_z (m^3/sec)	σ_z (m^3/sec)	μ_z (m^3/sec)	σ_z (m_3/sec)
LG4	865	276	894	266	911	268	937	280
LG3	870	277	899	268	916	270	943	281
EOL	1821	580	1882	561	1918	565	1973	588
LG2	660	210	682	203	695	205	715	213

Table 4.5 Sample size effect on the statistical mean μ_z and standard deviation σ_z of the inflows during May

mainly a function of these variables if the covariance matrix is used. However, this will not happen if the correlation matrix is used since the elements of this matrix are all restricted to the range [-1, 1]. The covariance matrix was used in our approach because it is important to distinguish between large and small reservoirs and avoid giving them all the same weight, as the correlation matrix does.

To illustrate the difference between the covariance and the correlation matrices, principal components analysis was performed on the storages during the month of May. For the covariance matrix, the percentage contributions of the four components to the total variance are

64.22 32.17 2.51 1.09

and the eigenvectors matrix is:

$$P = \begin{bmatrix} 0.1346 & -0.3330 & 0.4993 & -0.7885 \\ 0.2158 & -0.2008 & -0.8576 & -0.4215 \\ 0.5012 & -0.7368 & 0.0790 & 0.4468 \\ -0.8270 & -0.5532 & -0.0946 & 0.0326 \end{bmatrix}$$

For the correlation matrix the percentage contribution are:

64.44 24.31 8.98 2.26

and the corresponding eigenvectors matrix is:

$$P = \begin{bmatrix} 0.5144 & -0.4562 & -0.5026 & 0.5242 \\ 0.5437 & 0.1441 & 0.7615 & 0.3220 \\ 0.5972 & -0.1557 & -0.0653 & -0.7841 \\ -0.2884 & -0.8642 & 0.404 & -0.0817 \end{bmatrix}$$

If the covariance matrix is used , the first three components will account for 98.9% of the sample variance, which is very close to the result obtained with the correlation matrix (97.7%). However, the eigenvectors are very different, and do not give the same components.

4.6 Conclusion

The principal components analysis, outlined above, has been used as an efficient tool to transform a large-scale problem into a smaller one having a fewer number of random variables. It was shown that the degree of the reduction is a function of the correlations among the variables. If the correlations (interdependencies) are very high, the data are within an ellipsoid and the reduction is significant. Otherwise, they are within a sphere and it is not possible to represent them adequately by fewer variables.

In addition, the sample size m is an important factor. It was shown that the sample is "representative" for large m. However, for economical considerations, a sample size greater than 30 can be considered as being sufficiently large.

Finally, it is very easy to implement the principal components analysis since it concerns only linear relations. Its combination with the optimization model can give interesting results since the number of random variables is reduced. This is the subject of Chapter V.

CHAPTER V

APPLICATION OF PRINCIPAL COMPONENTS ANALYSIS TO THE STOCHASTIC MODEL

5.1 Introduction

The model discussed in Chapter II is obviously inadequate as a representation of reality because it ignores the essential difficulty any power system manager must face: his uncertainty about the future. In general, there are three major areas of uncertainty:

1) The future natural inflows.
2) The future demand.
3) The future availability of plants.

The model described here is designed to cope with the first kind of uncertainty which, in a predominantly hydro system, is the most important. The model presented in Section 5.2 is, in fact, an extension of the deterministic one developed in Chapter II. Since the goal of this thesis is to solve large-scale stochastic systems, a reduction by principal components analysis will be helpful. The reduced model presented in Section 5.3 will therefore be solved by stochastic dynamic programming. The formulation of the problem using this solution technique is given in Section 5.4. Then, in Section 5.5, the approach is applied to a four reservoir system. The results of the optimization process are also presented in that section. To evaluate the performances of this technique, its results are compared to the ones obtained by classical stochastic dynamic programming. In Section 5.6 the model used for the classical technique and some results are given. Finally, in Section 5.7, a discussion and a comparison of the results obtained using both methods are presented.

5.2 The Stochastic Model

The deterministic model (3.34) can easily be extended to the stochastic case. The objective of the problem becomes to determine the optimal monthly operating policy of n hydroelectric powerplants with mutually correlated inflows. The policy is found by minimizing the expected cost of the thermal energy over the complete time horizon, that is

$$\text{Minimize } E\{\sum_{k=1}^{K} C.a^T Hg_k + \alpha \sum_{k=1}^{K-1} e^T z_k\} \tag{5.1}$$

under the following constraints

$$X_{k+1} = X_k - \Gamma\,\Omega_k + Y_k \tag{5.2}$$

$$\rho^T \min(\Omega_k, \bar{U}) + b^T Hg_k = de_k \tag{5.3}$$

$$\min(\Omega_{k+1}, \bar{U}) - \min(\Omega_k, \bar{U}) - z_k \le 0 \tag{5.4}$$

$$\min(\Omega_{k+1}, \bar{U}) - \min(\Omega_k, \bar{U}) + z_k \ge 0 \tag{5.5}$$

$$\underline{X} \le X_{k+1} \le \bar{X} \tag{5.6}$$

$$0 \le Hg_k \le e \tag{5.7}$$

$$\Omega_k \ge 0 \tag{5.8}$$

$$z_k \ge 0 \tag{5.9}$$

In (5.1) the symbol E stands for the expected value, and in (5.2) Γ replaces the matrices Γ_1 and Γ_2 in the model (3.34). Since it is assumed that a hydroplant is associated to each reservoir, then Γ_1 is equal to Γ_2. In addition, Ω_k replaces $U_k + V_k$ with the following assumptions:

$$U_k = \Omega_k \text{ and } V_k = 0 \qquad \text{if } \Omega_k \le \bar{U} \tag{5.10}$$

$$U_k = \bar{U} \text{ and } V_k = \Omega_k - U_k \quad \text{if } \Omega_k > \bar{U} \tag{5.11}$$

in another form, the vector of the released water U_k can be written as

$$U_k = \min(\Omega_k, \bar{U}) \tag{5.12}$$

This means that if the discharge U_{ik} from reservoir i in month k is greater than the capacity of the allied plant $\bar{U}_i$, then $U_{ik} - \bar{U}_i$ hcm (hcm $= 10^6$ cubic meter) of water are discharged through the spillways.

The only random variable in equations (5.1) to (5.9) is the vector of the natural inflows $Y_k = [Y_{1k}, Y_{2k}, \ldots, Y_{nk}]^T$ since the demand de_k and the unit availabilities are assumed to be known in advance. It is also assumed that a correlation exists between Y_{ik} and Y_{jk} for all reservoirs i and j which results in an interdependency between the optimal reservoir contents X_{ik} and X_{jk}.

5.3 The reduced Stochastic Model

As mentioned in Chapter IV, principal components analysis is applied to the storages $X_{1k}, X_{2k}, \ldots, X_{nk}$ and to the inflows $Y_{1,k-1}, Y_{2,k-1}, \ldots, Y_{n,k-1}$. Recall that the X's were obtained from deterministic optimizations for m different inflow sequences generated synthetically (see Section 4.4.1). Thus if, Φ and Ψ are the vectors of principal components related to the inflows and to the storages respectively, then, according to (4.23)

$$Y_k = P_k \Phi_k \tag{5.13}$$

$$X_k = Q_k \Psi_k \tag{5.14}$$

Now, assume that p ($p \leq n$) components of Φ_k and q ($q \leq n$) components of Ψ_k are selected. According to (4.28),

$$Y_{ik} = \sum_{j=1}^{p} P_{jik}\Phi_{jk} + \sum_{j=p+\cdot}^{n} P_{jik}\mu_{\Phi jk}$$

$$Y_{ik} = \sum_{j=1}^{p} P_{jik}\Phi_{jk} + P_{0ik} \; ; \; i=1,2,\ldots,n \tag{5.15}$$

and

$$X_{ik} = \sum_{j=1}^{q} Q_{jik}\Psi_{jk} + \sum_{j=q+1}^{n} Q_{jik}\mu_{\Psi jk}$$

$$X_{ik} = \sum_{j=1}^{q} Q_{jik}\Psi_{jk} + Q_{0ik} \; ; \; i=1,2,\ldots,n \tag{5.16}$$

Furthermore, define $\tilde{P}_k$ and $\tilde{Q}_k$ as being matrices of dimensions n*p and n*q respectively and $\tilde{\Phi}_k$ and $\tilde{\Psi}_k$ as being the vectors of the selected components of dimensions p*1 and q*1 respectively. P_{0k} and Q_{0k} will be constant vectors of dimensions n*1. Consequently, equations (5.15) and (5.16) can be written, for period k, as

$$Y_k = \tilde{P}_k\tilde{\Phi}_k + P_{0k} \tag{5.17}$$

and

$$X_k = \tilde{Q}_k\tilde{\Psi}_k + Q_{0k} \tag{5.18}$$

These last two equations are of paramount importance. They state that once the principal components are selected, the reduced model can be easily found by substituting Y_k and X_k in the model described by equations (5.1) to (5.9) with the new variables Φ_k and Ψ_k. Therefore the reduced stochastic problem is to minimize

$$F = E\{\sum_{k=1}^{K} C.a^T Hg_k + \alpha \sum_{k=1}^{K-1} e^T z_k\} \tag{5.19}$$

under the following constraints

$$\tilde{Q}_{k+1}\tilde{\Psi}_{k+1} = \tilde{Q}_k\tilde{\Psi}_k - \Gamma\,\Omega_k + \tilde{P}_k\tilde{\Phi}_k + P_{0k} + Q_{0k} - Q_{0,k+1} \tag{5.20}$$

$$\rho^T \min(\Omega_k, \bar{U}) + b^T Hg_k = de_k \tag{5.21}$$

$$\min(\Omega_{k+1}, \bar{U}) - \min(\Omega_k, \bar{U}) - z_k \le 0 \tag{5.22}$$

$$\min(\Omega_{k+1}, \bar{U}) - \min(\Omega_k, \bar{U}) + z_k \ge 0 \tag{5.23}$$

$$\underline{X} - Q_{0,k+1} \le \tilde{Q}_{k+1}\tilde{\Psi}_{k+1} \le \bar{X} - Q_{0,k+1} \tag{5.24}$$

$$0 \le Hg_k \le e \tag{5.25}$$

$$\Omega_k \ge 0 \tag{5.26}$$

$$z_k \ge 0 \tag{5.27}$$

One difficulty arises from this model. It concerns the determination of the lower and upper bounds of Ψ_{ik}, $i=1,2,\ldots,q$. It is clear that the bounds given by (5.24) are linear combinations of Ψ_{ik}. Therefore an infinity of values can be found for each Ψ_{ik}. To eliminate uncertainty, two methods can be used. First, the mean value and the variance of the components are already known. This information alone is sufficient to fix the interval of possible values for Ψ_{ik}. Secondly, since the original data (reservoir's storages) and the transformation matrix (eigenvector matrix) are known for each month k, it is very easy to compute all the Ψ_{ik}. Therefore, the lower and upper bounds can be chosen to be the minimum and the maximum of these transformed data. Obviously, these two methods, as well as relation (5.24), depend on the choice of the original data sample. Since it is assumed that the sample of inflows is sufficiently large and representative, either the first or the second method can be used, even though the second one has been chosen in our case. Consequently, relation (5.24) will be replaced by

$$\underline{\Psi} \le \tilde{\Psi}_{k+1} \le \bar{\Psi} \tag{5.28}$$

where $\underline{\Psi}$ and $\bar{\Psi}$ are the vectors of lower and upper bounds respectively. In addition, since stochastic dynamic programming will be used to solve model (5.19) to (5.27), the penalty term z_k can be found, contrary to linear programming, from

$$z_k = \left| \min(\Omega_{k+1}, \bar{U}) - \min(\Omega_k, \bar{U}) \right| \tag{5.29}$$

Therefore relation (5.29) can replace the constraints (5.22), (5.23) and (5.27). The reduced stochastic model becomes

$$F = E\{\sum_{k=1}^{K} C.a^T Hg_k + \alpha \sum_{k=1}^{K-1} e^T z_k\}$$

under the following constraints

$$\begin{aligned}
&\tilde{Q}_{k+1}\tilde{\Psi}_{k+1} = \tilde{Q}_k\tilde{\Psi}_k - \Gamma\ \Omega_k + \tilde{P}_k\tilde{\Phi}_k + P_{0k} + Q_{0k} - Q_{0,k+1} \\
&\rho^T \min(\Omega_k, \bar{U}) + b^T Hg_k = de_k \\
&z_k - \left| \min(\Omega_{k+1}, \bar{U}) - \min(\Omega_k, \bar{U}) \right| = 0 \\
&\underline{\Psi} \le \tilde{\Psi}_{k+1} \le \bar{\Psi} \\
&0 \le Hg_k \le e \\
&\quad \Omega_k \ \ge 0
\end{aligned} \tag{5.30}$$

5.4 Stochastic Dynamic Programming Formulation

Stochastic dynamic programming differs from deterministic dynamic programming in that the state at the end of period k is not completely determined by the state and the decision at the beginning of period k. Instead, there is a probability distribution for the next state. The resulting basic structure for stochastic dynamic programming is described schematically in Figure 5.1. In that figure, ℓ denotes the number of possible states at the end of period k, and $(p_{k1}, p_{k2}, \ldots, p_{k\ell})$ is the probability distribution of the inflows Y_k, or of what the state X_{k+1} will be, given X_k. When Figure 5.1 is expanded to include all possible states and decisions for all periods, it is sometimes referred to as a "decision tree" [Hiller and Lieberman, 1974]. For the sake of presentation, let Y_k takes the following values

Y_{k1} with the probability p_{k1}

Y_{k2} with the probability p_{k2}

Y_{k3} with the probability p_{k3}

then the decision tree will give three different possibilities for state X_{k+1} at the end of period k. Consequently, for Y_{k1} there exist an optimal discharge (decision) U_{k1} and an optimal operating cost C_{k1} with probability p_{k1}. For Y_{k2}, there exist a U_{k2} and a C_{k2} with probability p_{k2}, and similarly for Y_{k3}. Therefore, the optimal decision U_k is a function of X_k and Y_k and it can be written as $U_k(X_k, Y_k)$. In addition, the operating cost is

$$F_k = C_{k1}p_{k1} + C_{k2}p_{k2} + C_{k3}p_{k3}$$

$$= \sum_{j=1}^{3} C_{kj}p_{kj} \tag{5.31}$$

Using the backward recurrent equation [Bellman 1957], the operating

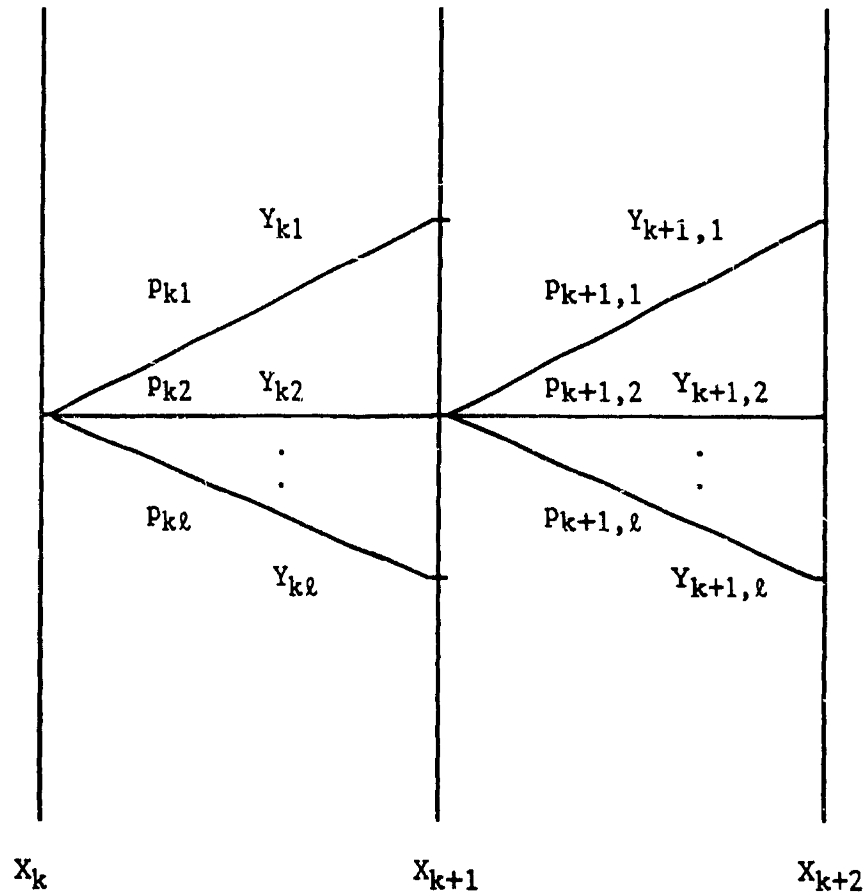


Figure 5.1 The basic structure for stochastic dynamic programming

cost can be written as

$$F_k(X_k) = \sum_{j=1}^{3} \{ \min[C_{kj} + F_{k+1}(X_{k+1})]\} p_{kj}$$

$$= \underset{Y_k}{E} \{ \underset{U_k}{\min}[C_k + F_{k+1}(X_{k+1})]\} \tag{5.32}$$

where $F_{k+1}(X_{k+1})$ is the optimal operating cost from period k+1 to period K.

Thus, the recurrent equation of model (5.30) is

$$F_k(\tilde{\Psi}_k, \tilde{\Phi}_{k-1}) = \underset{\Phi_k}{E} \underset{\Omega_k}{\min} [C.a^T Hg_k + \alpha.e^T z_k + F_{k+1}(\tilde{\Psi}_{k+1}, \tilde{\Phi}_k)] \tag{5.33}$$

Three remarks should be made here. First, the minimization is performed over the proper space of decisions. In this case the decisions are the variables representing the total discharge from the reservoirs and through the spillways ($\Omega_k = U_k + V_k$). Secondly, in (5.33), the random variables are the principal components Φ_k related to the natural inflows Y_k. Thus, the expected cost must be calculated for the transformed inflows Φ_k. The last remark concerns the penalty cost $\alpha.e^T z_k$. This cost is set equal to zero for the last period K, and computed according to (5.29) for k < K.

The complete procedure for the solution of the problem (5.30) using stochastic dynamic programming is given by the following algorithm.

Step1: Set k = K and the future cost $F_{k+1}(\tilde{\Psi}_{k+1}, \tilde{\Phi}_k) = 0$.

Set $\{\delta\}$ = all the feasible set of states at the end of period k.

Set $F_k(\tilde{\Psi}_k, \tilde{\Phi}_{k-1}) = \infty$.

Step2: Set $\{\mu\}$ = all the feasible set of states at the beginning of period k.

Step3: Choose a scenario of inflows $\tilde{\Phi}_{kj}$ for period k

Step4: Choose a feasible state. Compute the discharge Ω_k according to

$$\Omega_k = \Gamma^{-1}(\tilde{Q}_k\tilde{\Psi}_k - \tilde{Q}_{k+1}\tilde{\Psi}_{k+1} + \tilde{P}_k\tilde{\Phi}_{kj} + P_{0k} + Q_{0k} - Q_{0,k+1})$$

Step5: Find the thermal energy $b^T Hg_{kj}$ from

$$b^T Hg_{kj} = de_k - \rho^T \min(\Omega_k, U)$$

Step6: Find the penalty cost $\alpha . e^T z_k$ using

$$z_k = |\min(\Omega_{k+1}, \bar{U}) - \min(\Omega_k, \bar{U})|$$

Step7: Compute the cost function $C_k(\tilde{\Psi}_k, \tilde{\Phi}_{k-1})$ from

$$C_k(\tilde{\Psi}_k, \tilde{\Phi}_{k-1}) = [C.a^T Hg_k + \alpha . e^T z_k + F_{k+1}(\tilde{\Psi}_{k+1}, \tilde{\Phi}_k)]\, p_{kj}$$

if $C_k(\tilde{\Psi}_k, \tilde{\Phi}_{k-1})$ is less than $F_k(\tilde{\Psi}_k, \tilde{\Phi}_{k-1})$ then

$$F_k(\tilde{\Psi}_k, \tilde{\Phi}_{k-1}) = C_k(\tilde{\Psi}_k, \tilde{\Phi}_{k-1})$$

otherwise GO TO step8.

Step8: If there is another feasible state, GO TO step4.

Step9: If there is another possible scenario of inflows, GO TO step3; otherwise set k = k - 1

Step10: If k is different than 1, set {δ} = {μ} and GO TO step2; otherwise trace the optimal solution and STOP.

5.5 <u>Sample Application</u>

The four reservoir system, solved deterministically in Chapter III and analyzed by principal components in Chapter IV, will now be solved by stochastic dynamic programming. However, before doing so, it is important to recall the main steps of the solution. First, the deterministic solution for the 50 inflow sequences allowed us to perform principal components analysis. The contents of the reservoirs and the natural inflows were therefore analyzed. The components selected (one inflow component and three storages components) constitute the states of the reduced model. Consequently, explicit stochastic dynamic programming will be performed on a problem with 4 state variables instead of the 8 original ones. Moreover, the selection of the components was based on their percentage contribution to the total variance of the sample of data. So, to illustrate the effect of this selection, the solution of the same problem with only three state variables and with a variable number of components from period to period, will be presented later in this chapter.

In brief, in this application it is assumed that deterministic optimizations were performed, and that the optimal storages were obtained. It is also assumed that principal components analysis was performed on the inflows and storages. Thus three steps are required to find the optimal solution.

Step1: Find the explicit distribution of the single inflow component selected.

Step2: Solve the reduced model using stochastic dynamic programming.

Step3: Find the optimal solution for the original problem.

5.5.1 Distribution of the Inflows

The transformed data for the inflows are easily computed for each period k from the relation:

$$\Phi_k = P_k^T Y_k \tag{5.34}$$

Since only one component was selected, the inflows Y_k into the four reservoirs constitute a single random variable function. The 50 sequences for this random variable are illustrated in Table 5.1 for the twelve period problem. To construct a frequency distribution of these

Year	May	June	July	Aug	Sept	Oct	Nov	Dec	Jan	Feb	March	April
1	2884.61	2189.17	1529.58	1733.18	1891.69	1558.98	1246.94	899.29	593.63	447.64	349.30	304.40
2	1843.23	3276.21	2839.12	2226.49	2132.21	1763.56	1175.35	982.15	631.55	454.42	341.98	317.48
3	2858.06	2202.34	1488.04	1448.75	1692.84	2742.85	1887.69	1166.70	691.10	490.38	391.40	345.24
4	4008.50	4102.16	2195.28	2243.41	2111.58	1683.81	1333.67	945.22	621.38	457.91	364.65	518.70
5	2039.54	2617.96	1942.57	1719.56	1686.73	2062.09	1369.96	790.71	556.63	431.75	347.07	384.18
6	3482.81	2800.87	1299.26	589.94	1217.19	2299.31	1850.21	1142.74	692.68	479.57	393.34	682.06
7	2086.29	2582.21	1585.30	1812.07	1687.76	2102.05	1791.35	1123.91	697.22	478.39	370.14	437.41
8	1878.85	3814.21	2453.34	2120.67	1339.48	1291.23	1052.92	776.36	443.78	342.37	290.62	335.30
9	1864.13	946.57	981.05	1779.93	2276.17	1700.75	1231.65	769.68	453.59	328.43	303.03	350.97
10	2086.29	1558.27	1470.73	1178.28	844.48	1529.68	1123.54	779.50	483.92	365.98	279.82	267.38
11	2020.60	2373.62	1742.97	2302.81	2571.58	1278.47	1079.01	800.90	529.52	342.37	287.49	260.96
12	2524.22	3451.21	1791.01	2099.64	1559.36	1954.70	2019.33	1211.86	754.27	511.86	393.99	433.88
13	3079.85	2047.76	1628.79	2214.72	2230.72	1875.64	1250.08	916.53	612.67	407.13	309.16	550.68
14	3243.17	1291.10	1464.08	1314.97	1555.18	1435.23	1626.95	1086.98	594.55	409.73	328.59	407.73
15	1369.97	2325.48	1522.93	1263.40	2031.14	1269.61	1443.72	1020.42	637.28	456.37	353.20	410.33
16	2308.08	3768.84	1803.78	1148.38	1132.97	2279.79	1723.87	1206.77	720.43	486.60	373.27	499.89
17	2335.80	1393.26	995.76	1024.43	1040.19	1320.02	1373.86	867.87	551.54	407.13	330.13	416.98
18	2014.16	2086.64	1074.14	1382.27	1011.02	1169.53	1248.13	989.20	606.00	436.29	318.78	313.27
19	2290.22	3779.44	1984.49	1742.44	2117.87	1801.47	1098.44	988.23	661.25	474.85	358.92	428.44
20	2847.82	3454.34	1648.32	1798.06	2023.24	2004.46	1688.95	962.71	612.67	419.35	346.17	469.97
21	3662.82	2370.48	1131.03	1396.61	1121.20	2234.20	1377.01	848.42	490.96	352.28	294.17	457.83
22	2979.99	3422.14	1985.41	1311.43	1513.91	1732.03	1509.63	1069.49	637.28	433.70	351.25	505.62
23	1992.08	2567.50	1484.53	1511.70	1480.28	1712.60	1291.09	778.32	500.22	381.33	316.83	309.45
24	2793.83	3062.70	1799.87	1845.11	1962.87	2268.10	1834.31	1076.18	651.63	478.39	384.08	581.04
25	2542.07	1563.36	1659.51	1833.67	2092.48	1782.04	1479.64	922.64	572.16	379.38	307.56	376.14
26	967.00	2034.59	1386.35	2215.05	1877.35	2007.60	1917.31	1449.24	821.52	579.11	430.36	379.96
27	2132.29	1674.53	1532.18	1220.19	1192.37	1903.38	1314.04	734.96	476.61	355.42	267.06	383.49
28	3087.46	3398.17	2236.82	2671.79	1804.96	1147.51	1034.48	942.66	538.78	374.66	301.08	509.83
29	2053.59	2389.31	1632.95	2239.23	2154.46	2060.50	1354.67	984.11	659.30	464.04	364.65	515.95
30	1161.86	2888.26	2329.38	2339.00	2127.71	2204.32	1904.56	1227.19	755.21	511.86	414.47	476.63
31	1560.22	3245.55	1561.34	1580.93	1080.25	2478.34	1902.60	1243.09	702.95	496.51	366.59	230.06
32	1538.56	2221.78	1899.08	1994.78	2592.22	1625.83	1077.44	839.78	565.89	433.70	365.60	499.89
33	3371.92	3418.03	2288.38	2429.27	2031.14	2151.96	1397.02	714.90	478.20	360.50	297.94	397.25
34	2638.53	2870.16	1862.63	1796.51	1486.97	2255.82	1465.69	1076.18	641.18	461.45	379.00	451.71
35	3163.81	1380.08	1238.45	1160.79	1206.72	2468.08	1744.26	1111.52	672.06	472.26	376.41	421.20
36	1790.41	2752.73	2031.10	1796.12	1162.87	1454.11	1174.74	799.71	552.73	392.79	287.14	332.17
37	2872.91	3381.25	2277.18	2241.47	1992.43	1318.62	800.44	695.46	480.78	385.11	331.18	445.73
38	1620.27	2557.68	1335.71	1824.81	1922.80	1684.85	1468.83	952.84	564.30	398.92	347.07	300.19
39	1792.67	2786.52	1991.78	1811.04	1856.71	1860.10	1212.22	971.71	653.18	449.23	332.37	522.91
40	1976.86	3772.39	2780.27	2813.86	2305.67	1667.92	1642.24	1086.98	664.39	484.65	373.27	367.18
41	2485.45	3068.96	2109.87	2752.86	2301.56	2461.80	1559.84	1005.14	639.78	432.94	338.50	461.65
42	837.46	2766.05	1972.65	2434.99	2171.55	2206.45	1613.61	1052.62	631.55	445.15	344.57	225.85
43	2903.26	4083.89	3000.70	2699.17	2441.34	2803.43	2074.64	1286.45	720.43	490.38	397.53	684.66
44	1869.40	3834.63	1727.61	771.05	768.56	1721.47	1253.23	986.06	629.60	419.89	333.12	435.79
45	1942.70	3521.73	2346.84	2496.39	2399.67	968.48	906.36	534.26	346.86	267.99	251.12	400.77
46	2240.04	3259.29	2354.91	2384.85	1594.92	2333.33	1729.96	1164.13	665.94	439.43	351.25	448.57
47	2904.44	2608.33	1711.72	1784.11	1633.68	2020.00	1611.66	922.03	594.18	402.05	331.18	459.05
48	1236.24	2281.50	2003.37	1841.97	1844.30	1823.49	1673.05	975.26	551.54	376.25	311.75	301.81
49	2592.54	1835.79	1961.06	2014.21	1932.97	2264.53	1770.35	1086.98	700.36	489.20	371.32	326.05
50	859.54	3155.59	1838.27	1709.39	1553.64	2165.40	1606.56	1000.98	616.57	442.96	351.25	454.30

Table 5.1 Transformed inflows for the twelve period horizon

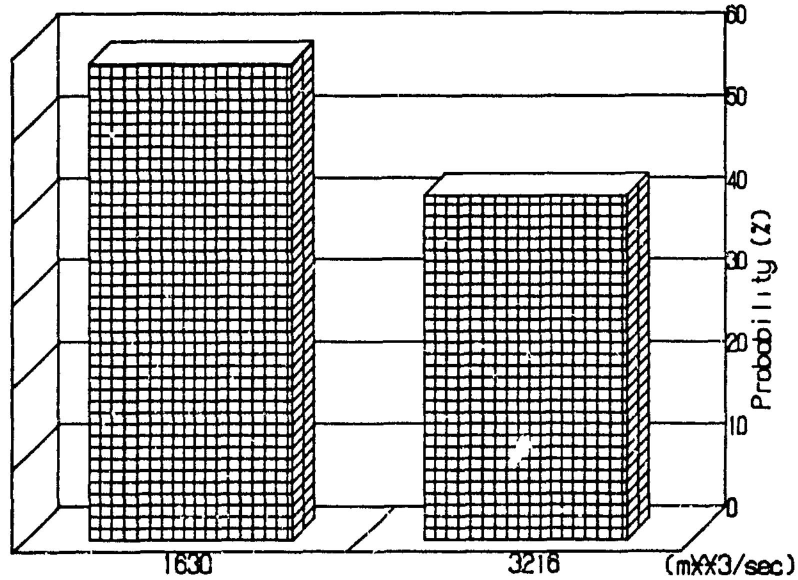
data, the first step consists in deciding on the number of classes (or categories) and their limits. Generally, the number of classes depends upon the number of observations. However, it is seldom profitable to use fewer than 5 or more than 15 classes [Miller and Freund, 1977]. Among other things, this decision is based on the range of the data, that is the difference between the largest and the smallest observation. Then we tally up the observations and, determine the class frequencies, namely, the number of observations in each class. Once data are grouped, each observation in a given class loses its own identity in the sense that its exact value is no longer known, but we get around this by representing each observation in a class by its midpoint, called the class mark. The histogram is then constructed by representing the class frequencies as a function of the successive class boundaries. The probability p_j for each class is computed as the number of observations ℓ_j in each class over the total number m of observations

$$p_j = \ell_j/m \; ; \; j=1,2,\ldots,J \tag{5.35}$$

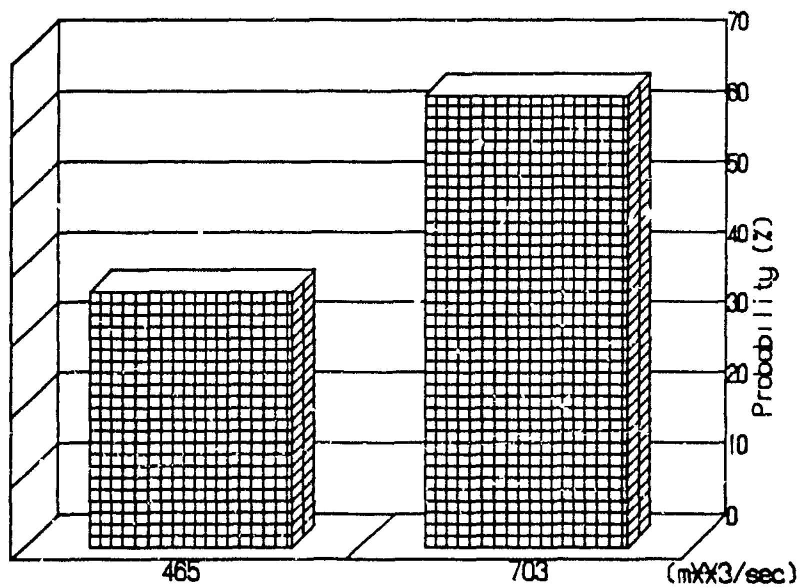
where J is the total number of classes.

For the sake of comparison with classical stochastic dynamic programming, let us set J equal to 2. Later in this chapter, the case of J = 5 will be considered.

A histogram of the random variable is illustrated in Figure 5.2 (a) for the month of May, in Figure 5.2 (b) for the month of January, while in Table 5.2 the different classes and the corresponding probabilities are shown for the twelve period problem.



(a) May



(b) January

Figure 5.2 Histogram of the inflows

	Class1	Prob1	Class2	Prob2
May	1629.72	0.580	3216.24	0.420
June	1734.97	0.380	3313.76	0.620
July	1485.46	0.700	2496.29	0.300
August	1145.42	0.300	2258.38	0.700
September	1223.97	0.400	2136.80	0.600
October	1426.72	0.520	2345.19	0.480
November	1118.49	0.480	1756.59	0.520
December	762.50	0.580	1220.99	0.420
January	465.02	0.360	703.35	0.640
February	345.27	0.420	501.83	0.580
March	295.43	0.440	386.05	0.560
April	340.05	0.680	570.46	0.320

Table 5.2 Distribution of the random variable

5.5.2 Stochastic Optimal Solution of the reduced Model

The algorithm developed in Section 5.4 is used to obtain the optimal solution. However, before presenting some results two remarks should be made. First, the principal components related to the storages were discretized into 5 states between the largest and the smallest values of these components. The program that was developed also offers the possibility to discretize these state variables according to the mean values and the standard deviations obtained from the principal components analysis program. Secondly, since the new bounds are obtained from the optimal solution of the deterministic problem, it is profitable to restrict the variations between the maximum and minimum obtained in each period. Consequently, the upper and lower bounds, contrary to the classical dynamic programming approach, vary from month to month. This point should be seen as an advantage because the number of discretizations is lower than the number needed in the classical dynamic programming method. Therefore, the computation time can be significantly reduced.

A typical output of the stochastic problem is shown in Tables 5.3 and 5.4 for period 6 and period 1 respectively. These tables illustrate the period number, the demand, the inflow and their corresponding probability. For example, for period 6 which corresponds to October, the demand is 11487 GWh and the inflows are equal to 1426.72 m^3/sec with a probability of 0.52 and to 2345.19 m^3/sec with a probability of 0.48. In addition, Table 5.3 shows the first 15 states of period 6. The first column indicates the state number while the second one shows the pointer for each state, that is the optimum state during period k+1. For example, the minimum cost for state 15 in period 6 is obtained from state 13 in period 7 with a probability of 0.52 and from state 33 with a probability of 0.48. Columns 3 to 10 indicate the storage levels and the outflows for each reservoir. Finally, column 11 shows the cumulative cost from period k to period K. Table 5.4, on the other hand, illustrates the single state at the beginning of period 1. The total cost is easily found from that table. It simply consists in adding up the costs of period one for all possible inflow branches (probabilities). In this case, the total cost is 24223.36 + 16098.54 = 40321.90. This cost, representing the value of the objective function of the optimization problem, is also given in Table 5.4.

5.5.3 Optimal Solution of the Original Problem

In this subsection, the optimal trajectories of the original model are sought. In order to achieve this goal, one must explore every branch, every trajectory of the decision tree. Since this is a very huge task for the problem dealt with, we will rather perform a simulation on

P E R I O D : 6

DEMAND : 11487. GWH

INFLOW : 1426.72 MCS   PROBAB. :0.520

State	Pt	Storage1	Outflow1	Storage2	Outflow2	Storage3	Outflow3	Storage4	Outflow4	Cost
1	21	18374.61	2128.81	12329.80	520.28	16917.68	76 3.97	18169.86	2230.88	16636.67
2	21	16795.68	1539.30	13138.27	232.62	16941.07	77 2.70	18169.86	1951.95	16922.57
3	21	15712.23	1134.79	13946.74	129.95	18418.32	132 4.24	17914.97	2305.66	16845.89
4	11	15712.23	1171.93	14755.20	146.84	19895.57	45 3.86	16674.62	2262.42	14414.75
5	17	15712.23	939.06	15563.67	5.10	21372.82	135 4.98	15434.27	1513.76	14087.22
6	47	19994.93	941.89	13378.23	310.18	17232.21	62 1.67	18169.86	2165.44	13628.89
7	47	18416.00	352.39	14186.70	22.52	18709.46	117 3.21	18169.86	2429.32	13664.07
8	17	16837.07	1359.03	14995.17	212.81	20186.71	91 2.13	18169.86	2299.98	13773.05
9	17	15712.23	939.06	15803.63	94.69	21663.96	146 3.67	17401.65	2446.58	13768.29
10	32	15712.23	193.94	16612.10	286.47	23141.21	55 8.34	16161.30	3505.11	13444.10
11	43	21615.25	1351.12	14426.66	578.03	19000.60	11 8.48	18169.86	2336.25	12647.43
12	43	20036.32	761.61	15235.13	290.37	20477.85	67 0.02	18169.86	2600.13	12681.19
13	43	18457.39	172.11	16043.59	2.71	21955.10	122 1.56	18169.86	2864.02	12716.37
14	18	16878.46	895.08	16852.06	70.38	23432.35	171 6.98	18128.68	2718.61	12764.43
15	13	15712.23	743.33	17660.53	59.43	24909.61	151 2.03	16888.33	2674.17	12827.02

INFLOW : 2345.19 MCS   PROBAB. :0.480

State	Pt	Storage1	Outflow1	Storage2	Outflow2	Storage3	Outflow3	Storage4	Outflow4	Cost
1	47	18374.61	839.43	12329.80	179.17	16917.68	109 6.28	18169.86	2838.86	12364.52
2	42	16795.68	533.60	13138.27	14.13	16941.07	34 8.51	18169.86	2651.60	12634.00
3	37	15712.23	412.76	13946.74	34.09	18418.32	14 3.56	17914 97	3096.98	12564.74
4	37	15712.23	412.76	14755.20	335.94	19895.57	69 5.10	16674.62	3487.28	12321.89
5	18	15712.23	962.15	15563.67	19.31	21372.82	154 0.09	15434.27	1814.50	12056.00
6	43	19994.93	1248.66	13378.23	447.02	17232.21	5 0.28	18169.86	2466.86	11693.51
7	43	18416.00	659.15	14186.70	159.36	18709.46	60 1.83	18169.86	2730.74	11725.41
8	18	16837.07	1382.12	14995.17	227.02	20186.71	109 7.25	18169.86	2600.72	11765.56
9	18	15712.23	962.15	15803.63	108.90	21663.96	164 8.79	17401.65	2747.33	11761.83
10	33	15712.23	217.03	16612.10	300.68	23141.21	74 3.46	16161.30	3714.87	11513.56
11	49	21615.25	1090.54	14426.66	469.61	19000.60	106 0.09	18169.86	2806.16	10822.02
12	43	20036.32	1264.11	15235.13	1155.75	20477.85	126 2.07	18169.86	4387.38	10854.17
13	24	18457.39	1224.00	16043.59	249.62	21955.10	210 7.06	18169.86	3633.13	10886.93
14	19	16878.46	918.17	16852.06	84.58	23432.35	190 2.10	18128.68	3247.76	10928.96
15	33	15712.23	217.03	17660.53	692.11	24909.61	140 3.79	16888.33	5037.98	10987.11

Table 5.3 Output of the optimization program for period 6

P E R I O D : 1

DEMAND : 10275. GWH

INFLOW : 1629.72 MCS PROBAB. :0.580

State	Pt	Storage1	Outflow1	Storage2	Outflow2	Storage3	Outflow3	Storage4	Outflow4	Cost
1	66	8000.00	384.32	6000.00	156.87	10000.00	109 5.44	5000.00	636.73	24223.36

INFLOW : 3216.24 MCS PROBAB. :0.420

State	Pt	Storage1	Outflow1	Storage2	Outflow2	Storage3	Outflow3	Storage4	Outflow4	Cost
1	84	8000.00	70.26	6000.00	773.18	10000.00	70 3.46	5000.00	1282.60	16098.54

O B J E C T I V E : .40322E+05

Table 5.4 Output of the optimization program for period 1 a certain number of the policies or trajectories obtained.

The simulation process starts by selecting a class of inflows. In our case, the number of possible classes is 2. Thus, for each period k, a class (1 or 2) will be selected and the optimal trajectory will be built accordingly. For example, suppose that for period 1 the class selected is 2, for period 2 it is equal to 1, for period 3 it is equal to 1, etc. Then, the trajectory will correspond to the second class of inflows for period 1, to the first class for period 2, to the first class for period 3, and so on. Table 5.5 shows a sequence for the twelve period problem. In addition, the corresponding trajectories of storages and outflows are given in the same table and illustrated in Figures 5.3 and 5.4. In fact, we can have a large number of these simulations ($2^{12} = 4096$). Moreover, we can take a sample of these trajectories and perform a Monte-Carlo simulation to see the probability that this approach gives an optimal solution outside the feasible region. For a sample of a hundred simulations, all these storages respect the upper and lower limits previously set. Unfortunately, it would be too unpractical and cumbersome to include all these results in this thesis. Therefore, only two other simulations will be presented here. Table 5.6 illustrates the optimal trajectory for the second simulation shown in Figures 5.5 and 5.6 while Table 5.7 and Figures 5.7 and 5.8 demonstrate the results obtained from the third simulation. It is important to remind that the outflows from the four reservoirs of Tables 5.5, 5.6 and 5.7 represent the discharge U_k and the spilled water V_k. Therefore, if the outflows Ω_k are greater than the plant capacity $\bar{U}$, then $V_k = \Omega_k - \bar{U}$ m^3/sec of water is discharged through the spillways. For example, LG3, represented by outflow 2 in Table 5.7, has a capacity

Month	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
Branch	2	1	1	2	2	1	2	2	1	1	2	2

SIMULATION #1

Month	State	Pt	Storage1	Outflow1	Storage2	Outflow2	Storage3	Outflow3	Storage4	Outflow4
May	1	84	8000.000	70.262	6000.000	773.178	10000.000	703.458	5000.000	1282.601
Jun	84	109	11061.715	195.516	7386.961	78.467	14958.531	139.962	7997.320	1783.314
Jul	109	86	13062.910	287.750	9577.977	471.650	17365.234	102.900	5584.320	497.116
Aug	86	64	14862.645	217.419	10637.984	390.140	19183.199	823.644	7350.746	516.831
Sep	64	39	17745.547	767.493	12571.008	440.147	20822.180	276.925	11231.500	22.459
Oct	39	38	18956.496	642.129	15603.527	147.381	23606.129	1081.490	14876.406	2364.516
Nov	38	33	19327.867	923.137	18438.434	791.151	23172.336	1303.671	13206.691	4132.031
Dec	33	43	19593.641	1797.822	20532.402	3368.762	22561.285	2849.862	9638.203	4210.777
Jan	43	15	16787.883	2222.000	17664.074	3212.737	16793.430	1121.111	16192.414	5316.211
Feb	15	47	11633.527	859.140	15505.148	3005.531	14473.043	1788.961	14014.840	3971.796
Mar	47	6	10094.762	1030.511	10647.121	2922.605	10596.187	431.642	16307.090	5676.250
Apr	6	1	8000.000	307.683	6000.000	530.274	10002.250	392.408	10453.078	3193.557

Table 5.5 Optimal trajectory for the first simulation

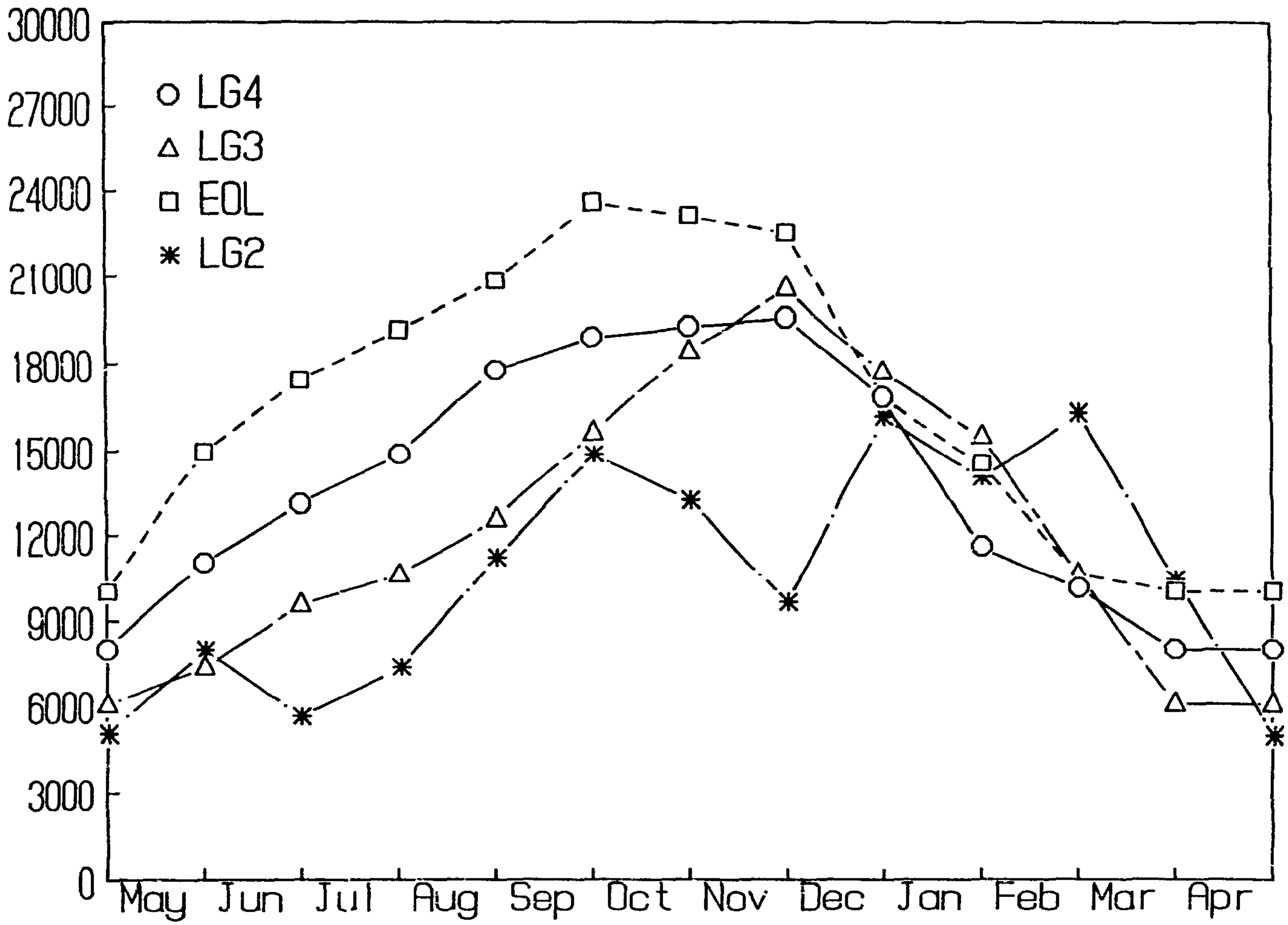


Figure 5.3 Optimal storages for simulation 1

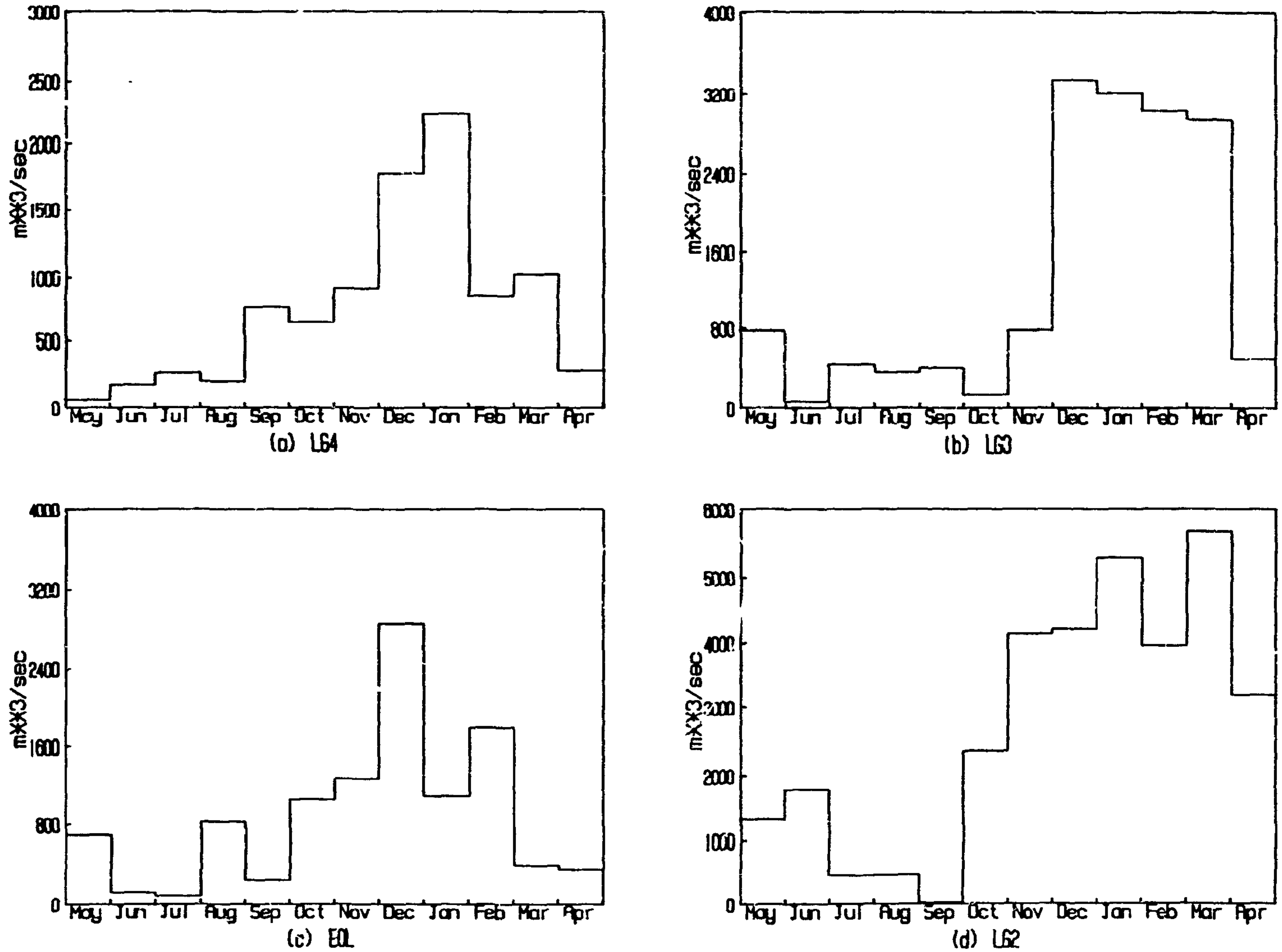


Figure 5.4 Optimal outflows for simulation 1

Month	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
Branch	1	1	2	2	1	2	2	1	2	2	2	1

SIMULATION #2

Month	State	Pt	Storage1	Outflow1	Storage2	Outflow2	Storage3	Outflow3	Storage4	Outflow4
May	1	66	8000.000	384.319	6000.000	156.866	10000.000	1095.437	5000.000	636.727
Jun	66	34	8617.270	59.433	8265.816	816.773	10533.406	72.977	7904.266	439.650
Jul	34	38	10971.191	22.063	8190.414	324.875	13113.734	257.579	10714.109	606.334
Aug	38	38	15231.258	371.784	9988.625	286.299	15942.316	259.122	13270.211	656.375
Sep	38	7	17700.707	431.195	12613.230	306.825	19093.312	921.980	14987.066	408.504
Oct	7	43	18415.996	559.150	14186.703	159.357	18709.457	601.825	18169.859	2730.744
Nov	43	43	20087.660	1047.416	18007.082	669.544	21146.133	1507.316	15149.977	3466.654
Dec	43	43	20031.305	1679.439	20738.387	3139.579	20007.234	1634.857	13518.793	4050.429
Jan	43	5	16787.883	2065.572	17664.074	3150.878	16793.430	537.003	16192.414	5777.484
Feb	5	72	12460.914	222.056	15505.148	3024.444	16387.520	2892.365	11281.691	4020.179
Mar	72	4	12708.320	2006.302	9211.973	3010.021	10045.531	206.274	16308.770	5878.238
Apr	4	1	8000.000	183.423	6944.277	680.482	10055.219	254.714	9544.270	2787.824

Table 5.6 Optimal trajectory for the second simulation

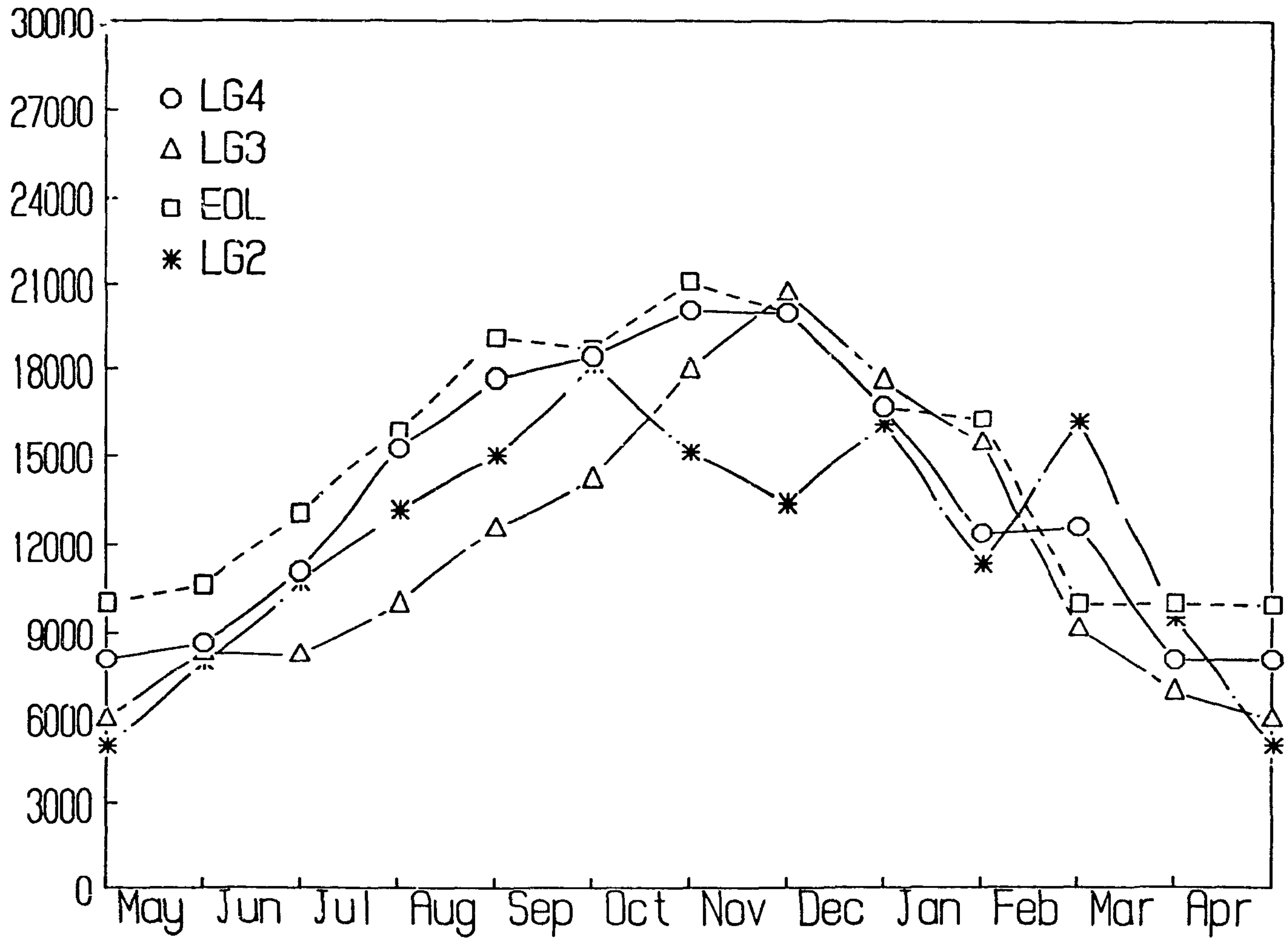


Figure 5.5 Optimal storages for simulation 2

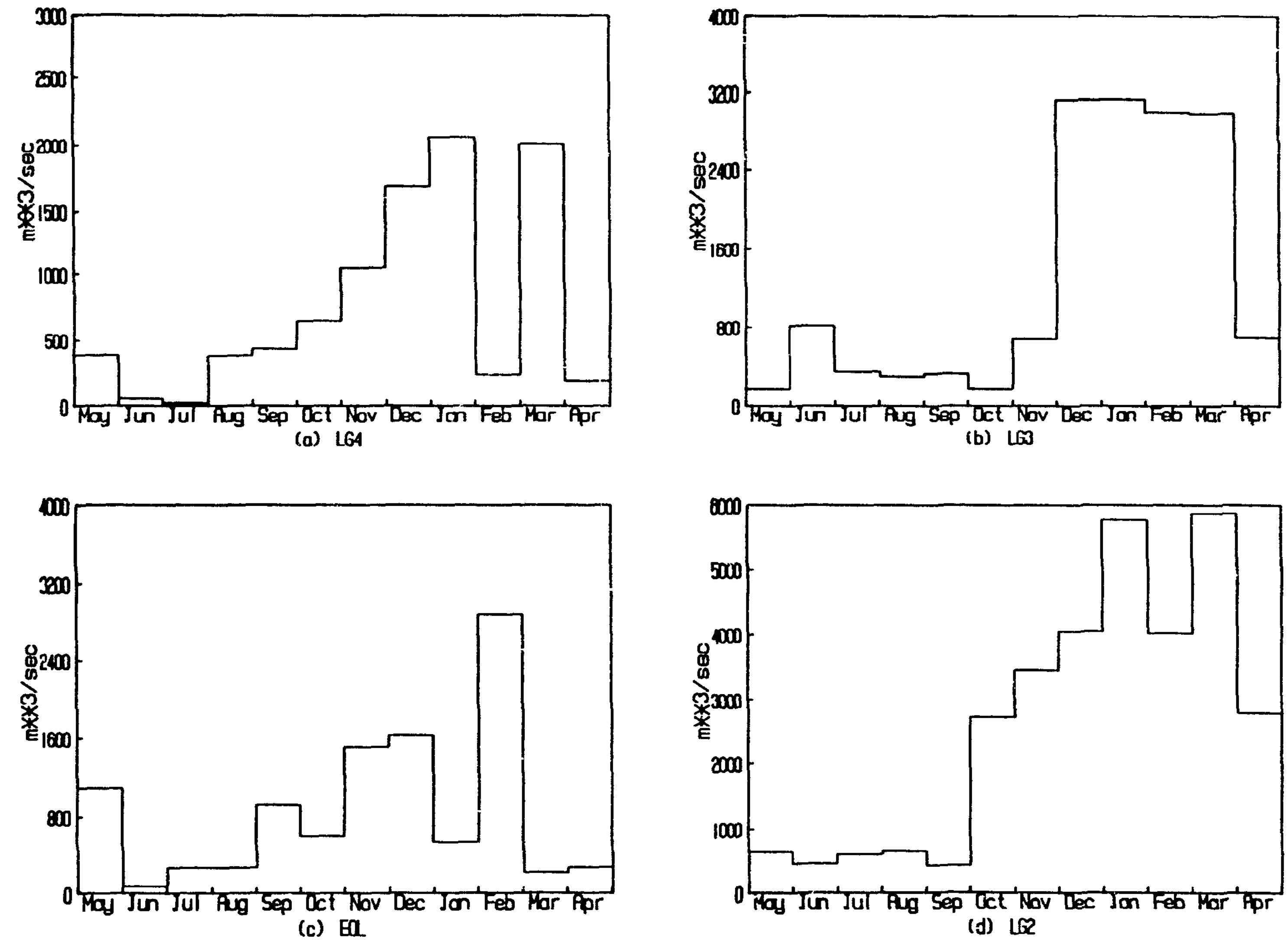


Figure 5.6 Optimal outflows for simulation 2

Month	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
Branch	2	1	1	1	2	1	1	1	1	1	2	2

SIMULATION #3

Month	State	Pt	Storage1	Outflow1	Storage2	Outflow2	Storage3	Outflow3	Storage4	Outflow4
May	1	84	8000.000	70.262	6000.000	773.178	10000.000	703.458	5000.000	1282.601
Jun	84	109	11061.715	195.516	7386.961	78.467	14958.531	139.962	7997.320	1783.314
Jul	109	86	13062.910	287.750	9577.977	471.650	17365.234	102.900	5584.320	497.116
Aug	86	59	14862.645	35.765	10637.984	194.170	19183.199	688.704	7350.746	79.330
Sep	59	33	16524.301	312.302	11428.895	260.719	19288.910	937.562	10524.039	32.688
Oct	33	37	18915.105	1106.079	13746.637	289.819	20360.488	276.637	15389.727	2213.142
Nov	37	46	18043.832	380.888	17442.684	74.295	22082.414	2563.809	12351.238	2152.528
Dec	46	46	18749.207	1890.800	19352.957	3536.337	17199.645	996.427	14699.930	3547.504
Jan	46	7	14939.676	2161.393	15782.078	3203.245	15695.812	88.656	18073.289	5205.391
Feb	7	11	9947.648	1028.156	13486.246	2407.440	16140.754	1707.156	13401.785	4123.844
Mar	11	6	8000.000	248.417	10484.004	2079.610	12461.801	1128.182	13681.391	4549.473
Apr	6	1	8000.000	307.683	6000.000	530.274	10002.250	392.408	10453.078	3193.557

Table 5.7 Optimal trajectory for the third simulation

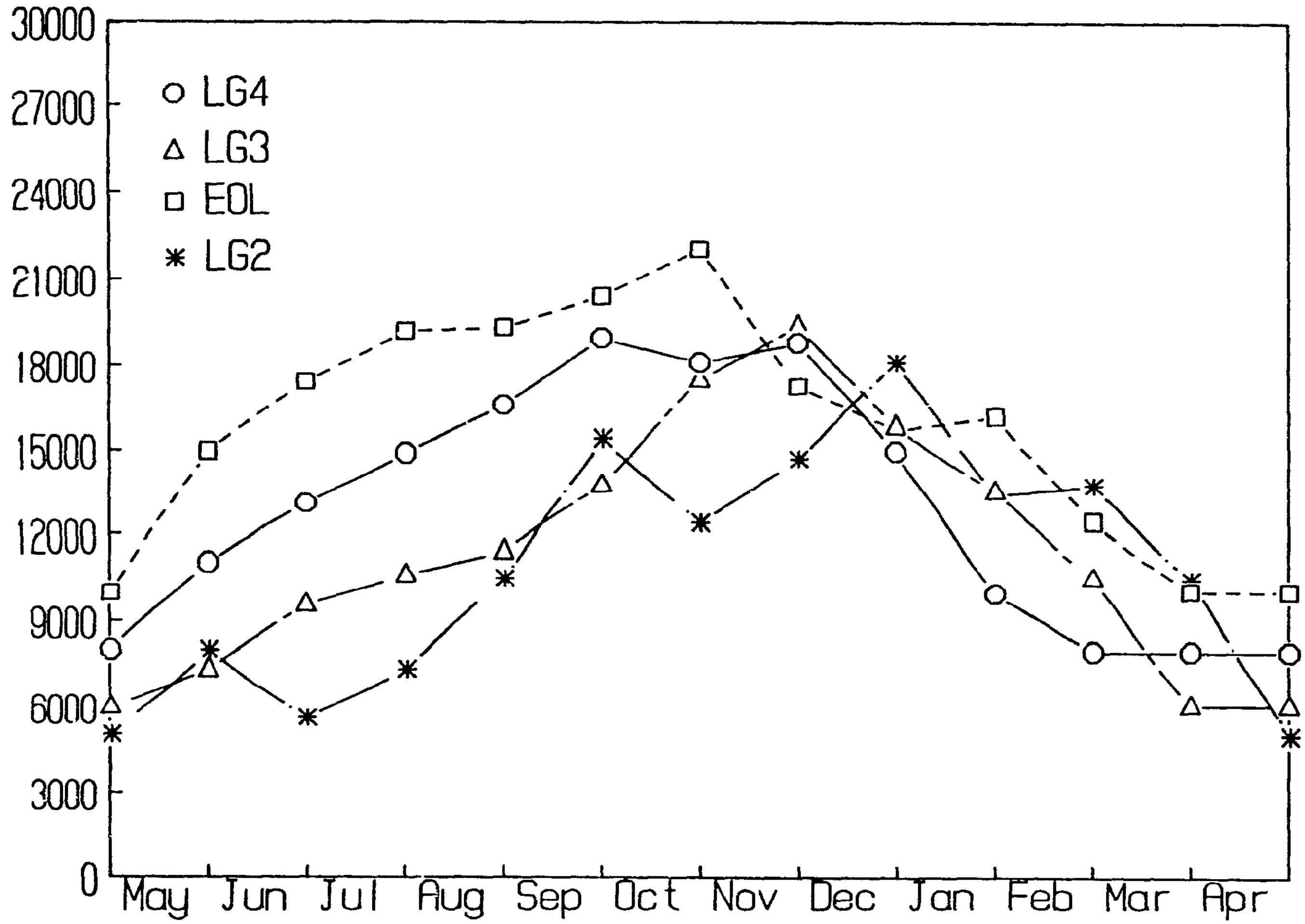


Figure 5.7 Optimal storages for simulation 3

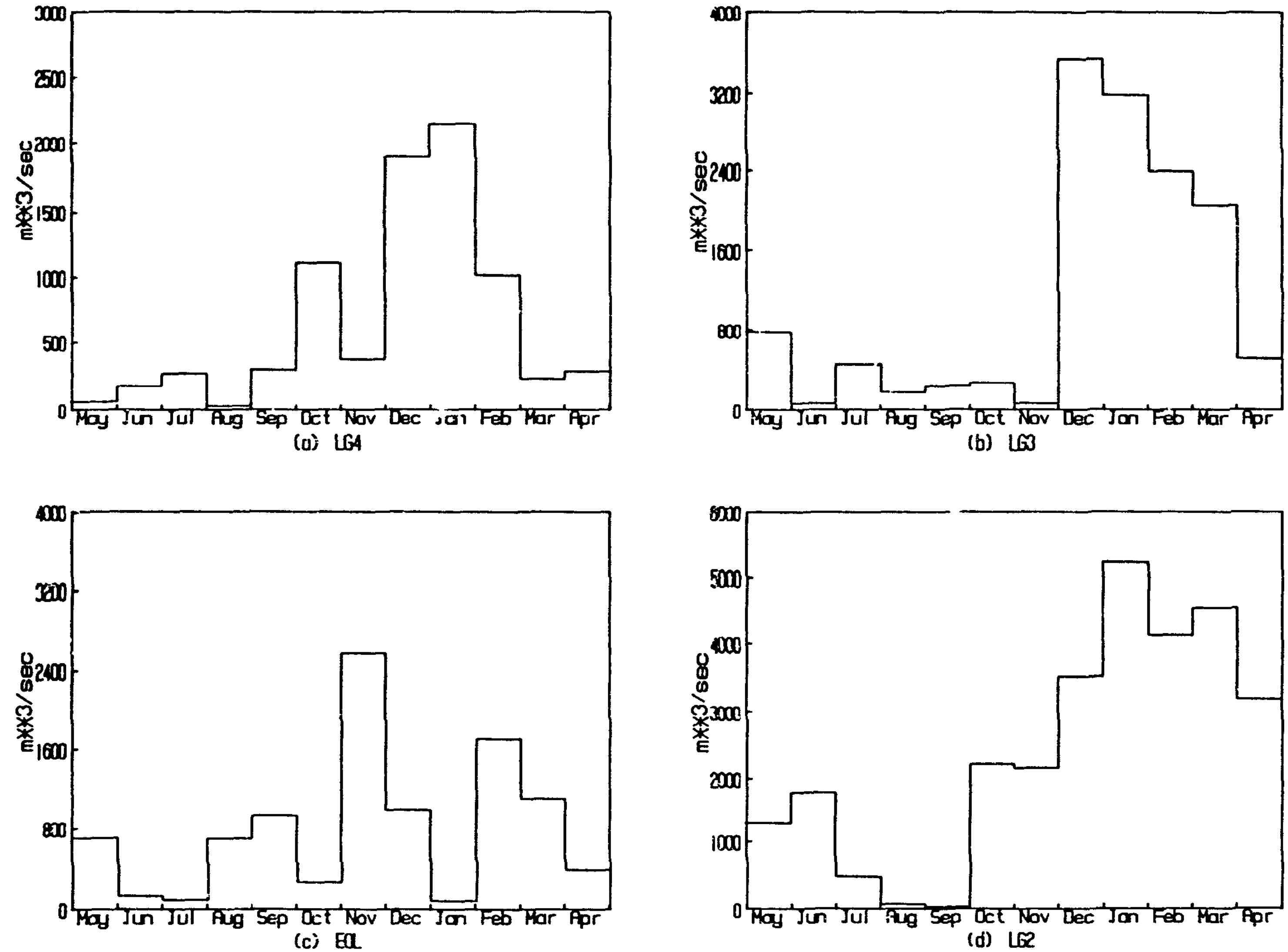


Figure 5.8 Optimal outflows for simulation 3

of 3433.5 m^3/sec. In December, the total discharge given by stochastic dynamic programming was 3536.337 m^3/sec. Thus 102.83 m^3/sec of water must be discharged through the spillways during that period.

Finally, in addition to being easy to implement, this approach is very fast to execute. Indeed, only 12 sec of CPU time on an IBM-3081 were required to solve the twelve period problem. Of course, to verify the efficiency of our approach, we must solve the same four reservoir system without any variable reduction by classical dynamic programming to be able to establish a basis of comparison between the two methods. This is the subject of the next section.

5.6 Classical Stochastic Dynamic Programming

The procedure for solving problem (5.1)-(5.9) is that given in Section 5.4. The only change that we made was to replace the selected components by the original variables.

The first step of the solution procedure is to determine the explicit distribution of the inflows for each period k. This distribution is found from the 50 sequences of inflows used to perform the principal components analysis. In addition, only two classes (branches) will be considered to obtain comparable results between the classical approach and the technique proposed in this thesis. A histogram of the inflows is illustrated in Figures 5.9 and 5.10 for the month of May and the month of January respectively, while in Table 5.8 the different classes and the corresponding probabilities are shown for the twelve period problem.

The next step is to discretize the storage levels between the maximum and minimum capacity of the reservoirs. A discretization into 8 states was chosen. The storage levels at the beginning of May were set

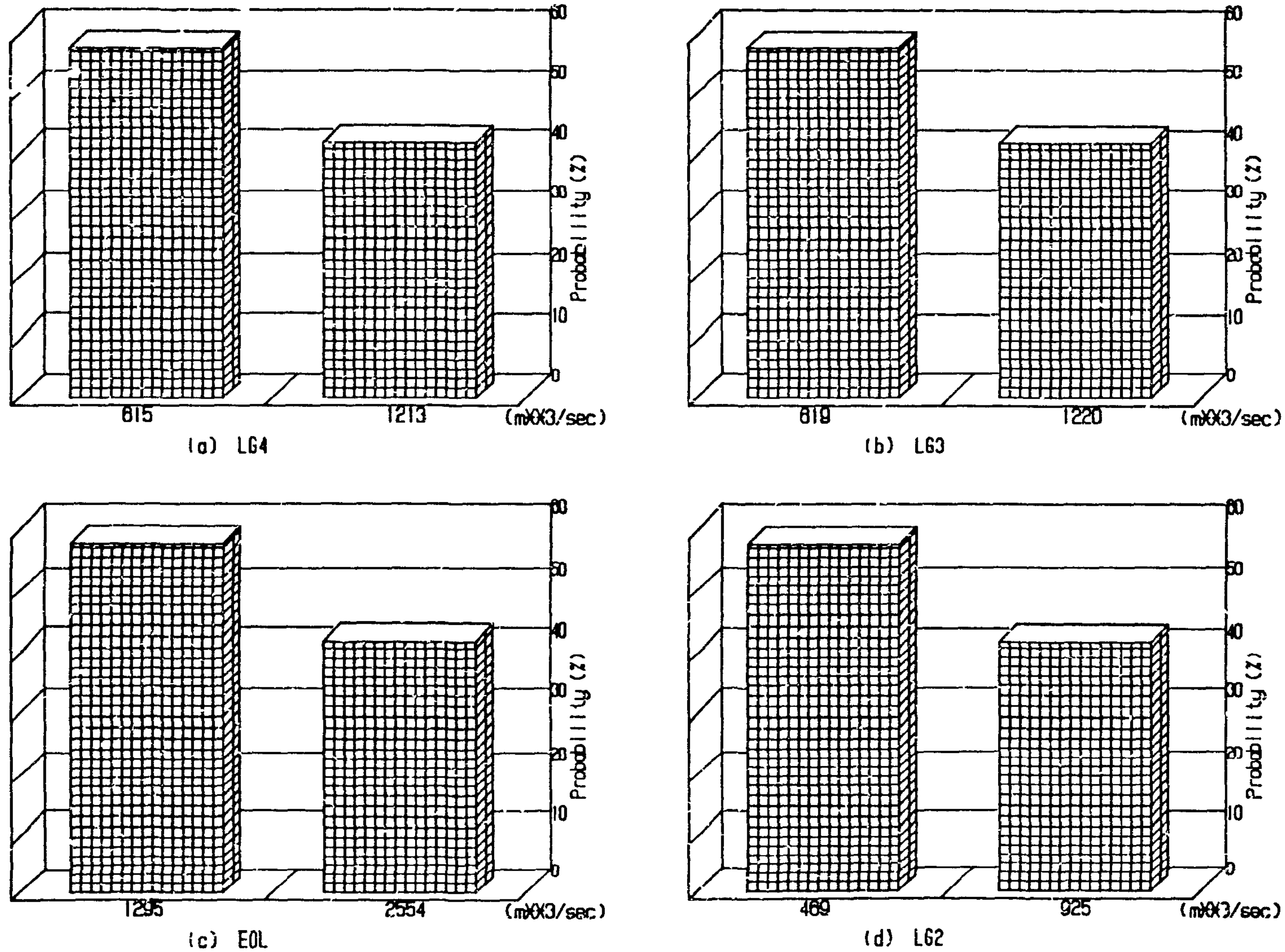


Figure 5.9 Histogram of the inflows in May

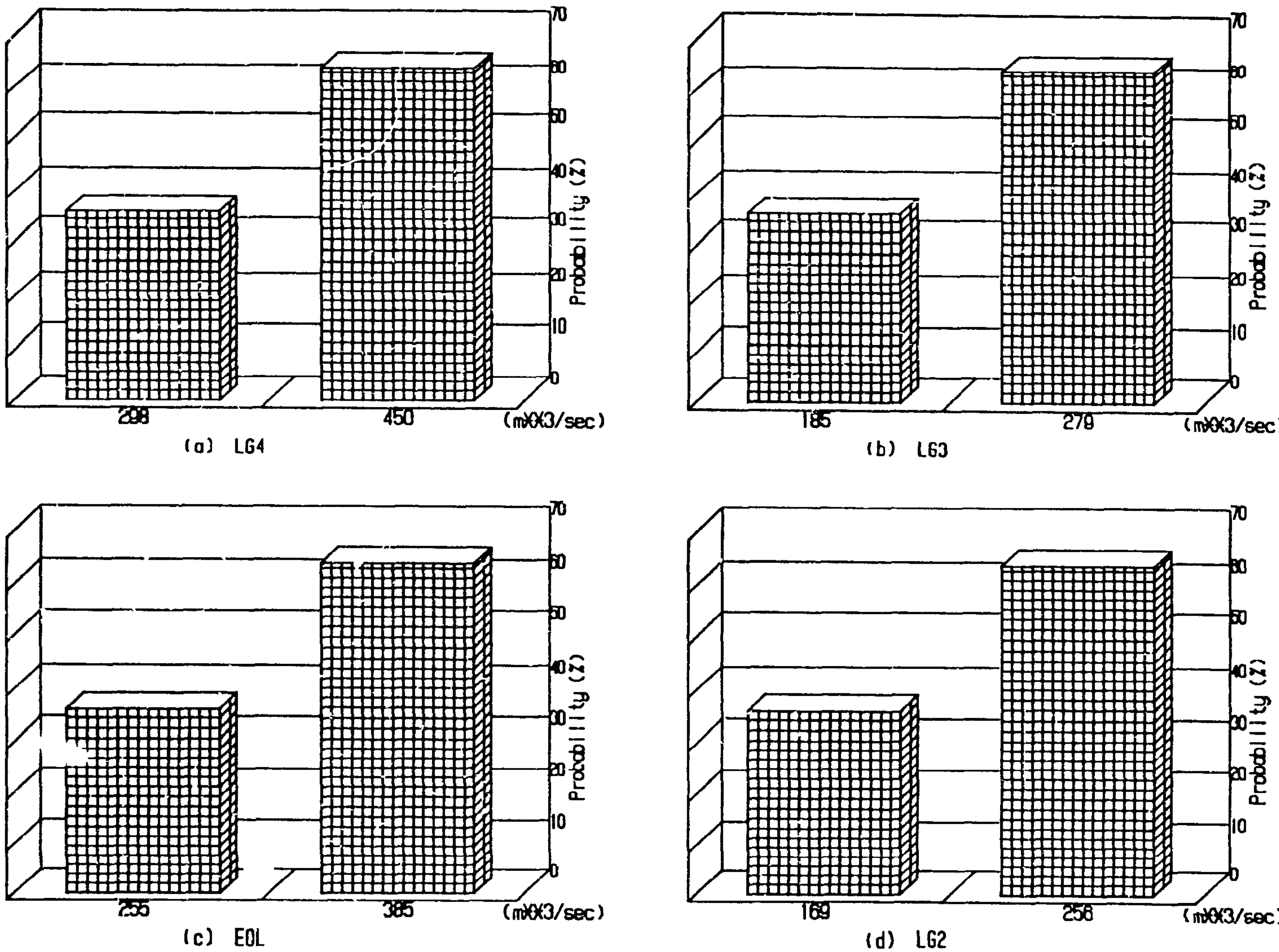


Figure 5.10 Histogram of the inflows in January

	Class1				Prob1	Class2				Prob2
	LG4	LG3	EOL	LG2		LG4	LG3	EO1	LG2	
May	615.00	618.75	1294.75	469.00	0.580	1213.00	1220.25	2554.25	925.00	0.420
Jun	968.00	728.25	1068.75	634.25	0.380	1848.00	1390.75	2040.25	1210.75	0.620
Jul	960.00	580.00	781.75	582.00	0.706	1612.00	974.00	1313.25	978.00	0.294
Aug	656.50	454.00	728.50	381.25	0.300	1293.50	894.00	1435.50	751.75	0.700
Sep	707.50	482.75	774.25	407.75	0.400	1234.50	842.25	1350.75	711.25	0.600
Oct	781.00	564.25	919.75	512.75	0.520	1283.00	926.75	1511.25	842.25	0.480
Nov	653.25	430.50	680.50	420.75	0.480	1025.75	675.50	1067.50	660.25	0.520
Dec	468.50	312.75	435.50	274.25	0.580	749.50	500.25	696.50	438.75	0.420
Jan	298.00	185.00	255.00	169.25	0.360	450.00	279.00	385.00	255.75	0.640
Feb	223.25	138.25	187.00	125.00	0.420	323.75	200.75	271.00	181.00	0.580
Mar	190.75	120.25	160.50	104.75	0.440	248.25	156.75	209.50	136.25	0.560
Apr	183.75	132.75	233.75	99.75	0.680	307.25	222.25	391.25	167.25	0.320

Table 5.8 Distribution of the inflows for the four reservoir system

at their lower bounds, and the penalty cost α was fixed to 0.01 as it was the case in the reduced stochastic model.

Once again, a large number of optimal trajectories was obtained. For this reason a simulation program was developed to study some of these policies. The simulation process generates random numbers between 1 and the number of possible classes of inflows. In this case, we have also 2 classes. Thus, the decision tree has a total of $2^{12} = 4096$ trajectories. From a sample of hundred simulations, only two will be given in this thesis. Table 5.9 illustrates the optimal trajectory of the first simulation shown in Figure 5.11 for the optimal storages, and in Figure 5.12 for the optimal outflows. The results of the second simulation, shown in Figures 5.13 and 5.14, are given in Table 5.10. The total thermal energy needed will therefore be 40279 GWh. Finally, the CPU time required to solve the problem is 44 minutes on an IBM-3081.

5.7 Discussion

From the results of the last 2 sections, the following 3 points will be discussed. First, a comparison of the results obtained from the classical and the reduced approaches will be made. Secondly, since the selection criteria is based on the percentage contribution of the components, the level at which this percentage will be satisfactory will be discussed. Finally, the explicit distribution of the principal component of the inflows was studied in 2 classes, which is not sufficient in practice. Therefore, as a third point, a discussion of the CPU time and of the optimal policies when 5 classes are considered will be presented.

Month	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
Branch	1	1	2	2	1	2	2	1	2	2	2	1

SIMULATION #1

Month	State	Pt	Storage1	Outflow1	Storage2	Outflow2	Storage3	Outflow3	Storage4	Outflow4
May	1	67	8000.000	615.000	6000.000	209.685	10000.000	1294.750	5000.000	440.539
Jun	67	133	8000.000	968.000	8742.855	638.050	10000.000	1068.750	9105.711	757.055
Jul	133	710	8000.000	225.245	11485.711	175.179	10000.000	1313.250	13211.426	1699.980
Aug	710	839	11714.285	1293.500	14228.570	139.370	10000.000	1435.500	15264.281	1560.171
Sep	839	904	11714.285	707.500	19714.281	132.048	10000.000	774.250	17317.141	522.052
Oct	904	976	11714.285	1283.000	22457.141	1185.685	10000.000	44.490	19369.996	2072.425
Nov	976	976	11714.285	1025.750	25199.996	1701.250	13928.570	1067.500	19369.996	3429.000
Dec	976	846	11714.285	468.500	25199.996	2829.381	13928.570	435.500	19369.996	5072.027
Jan	846	262	11714.285	1836.755	19714.281	3139.820	13928.570	1851.760	15264.281	5247.328
Feb	262	132	8000.000	323.750	16971.426	2792.074	10000.000	271.000	15264.281	4941.207
Mar	132	2	8000.000	248.250	11485.711	2453.130	10000.000	209.500	11158.570	4331.777
Apr	2	1	8000.000	183.750	6000.000	316.500	10000.000	233.750	7052.855	1441.997

Table 5.9 Optimal trajectory for the first simulation

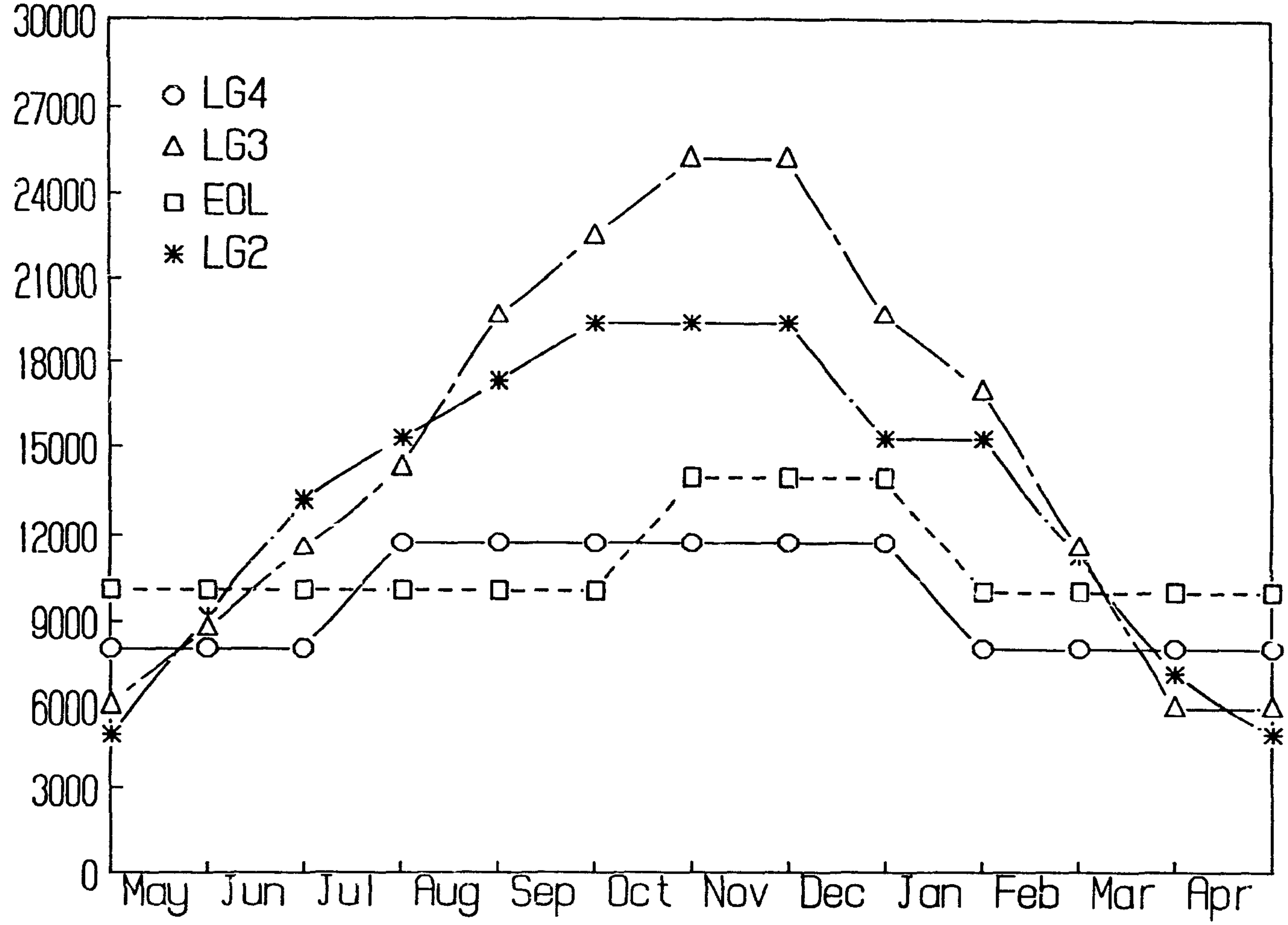


Figure 5.11 Optimal storages for simulation 1

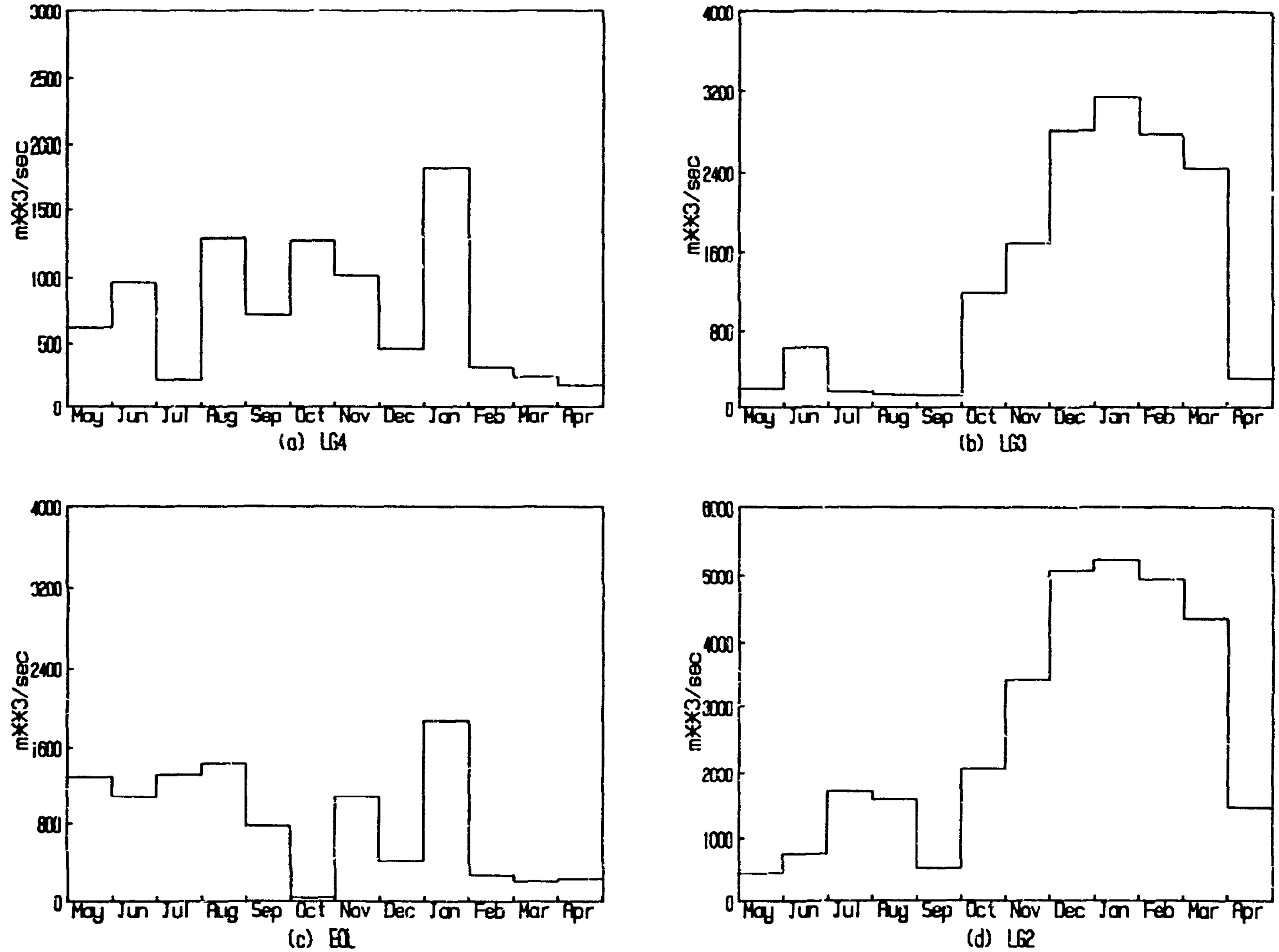


Figure 5.12 Optimal outflows for simulation 1

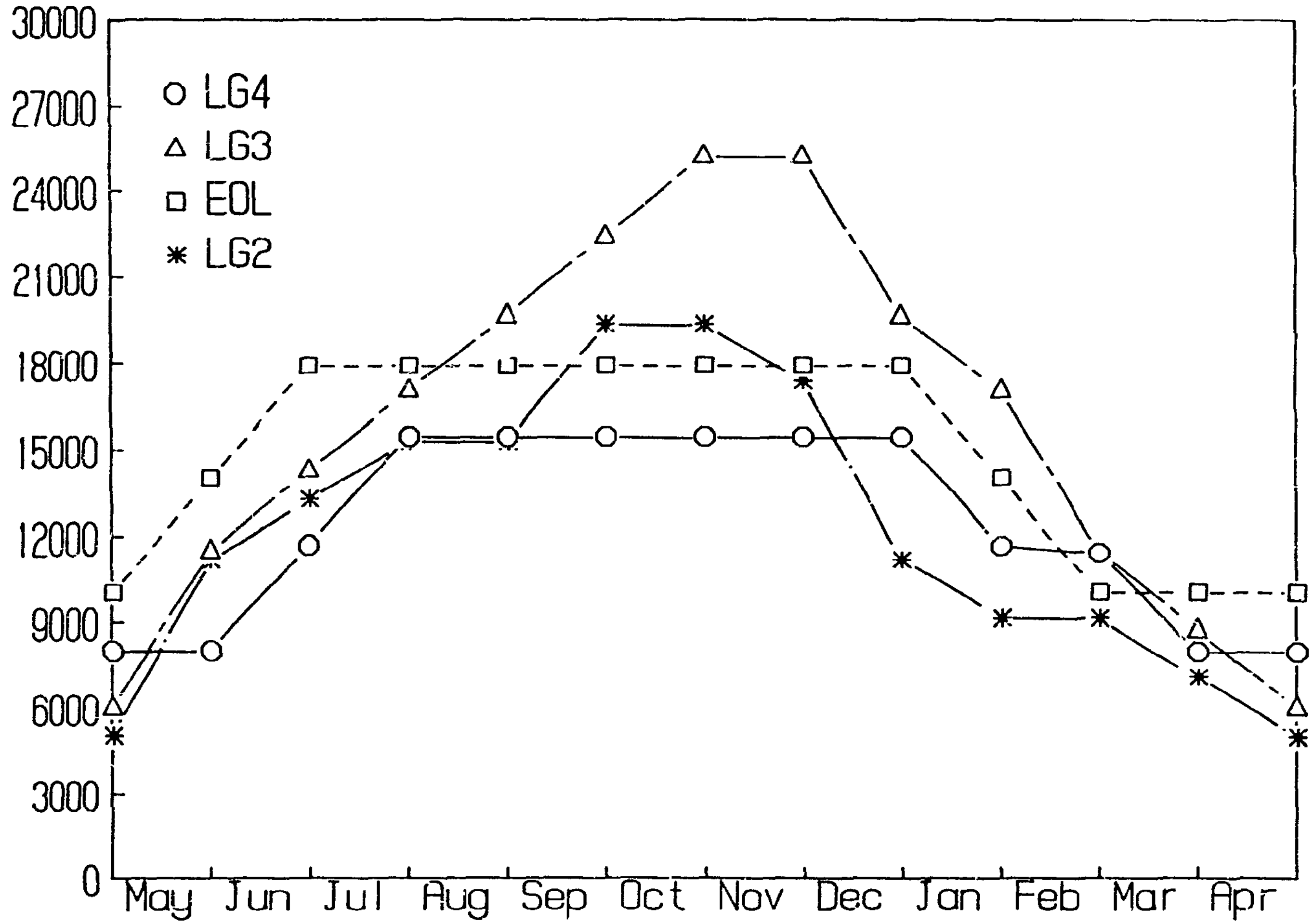


Figure 5.13 Optimal storages for simulation 2

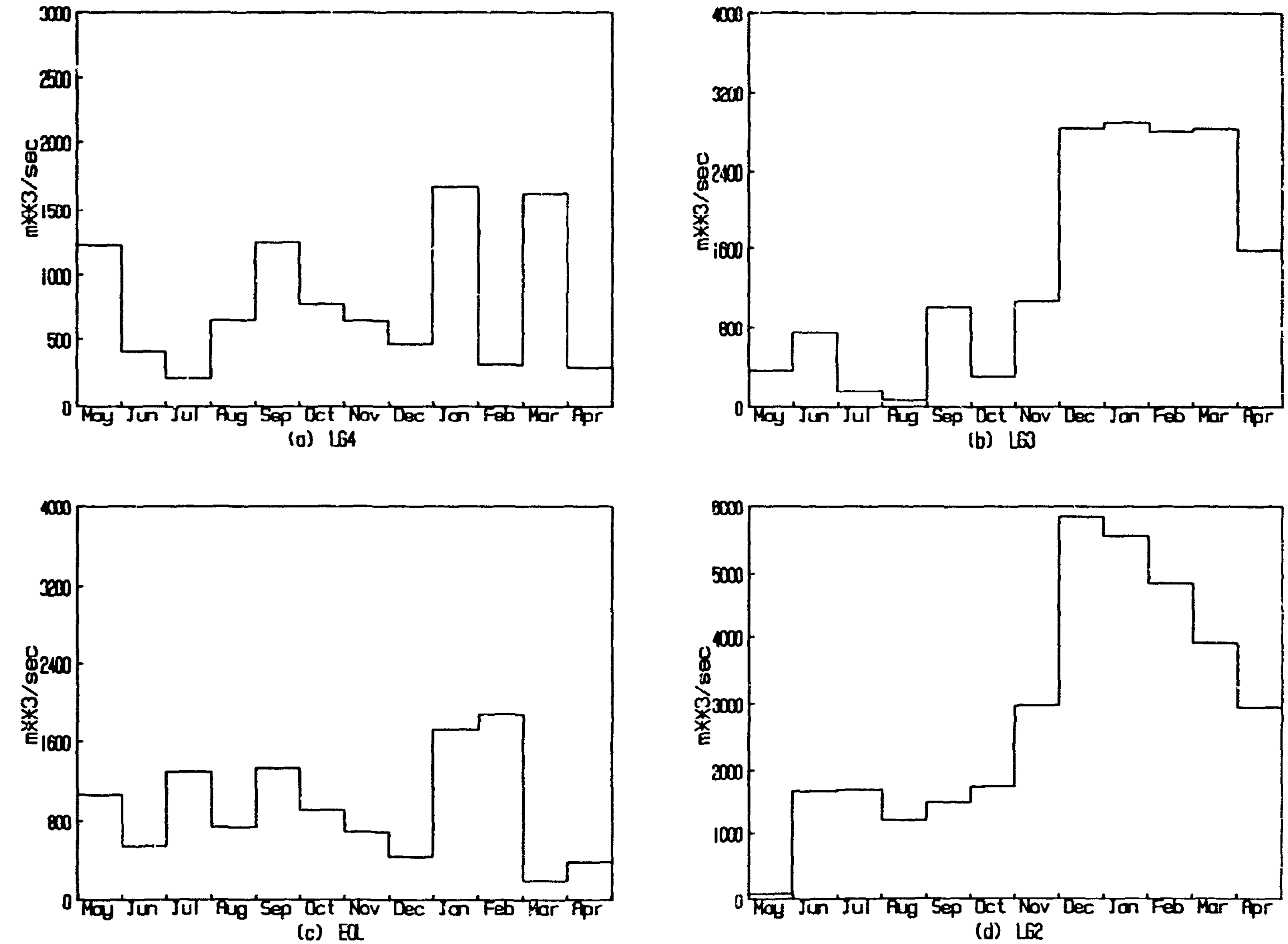


Figure 5.14 Optimal outflows for simulation 2

Month	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
Branch	2	2	2	1	2	1	1	1	1	2	2	2

SIMULATION #2

Month	State	Pt	Storage1	Outflow1	Storage2	Outflow2	Storage3	Outflow3	Storage4	Outflow4
May	1	140	8000.000	1213.000	6000.000	385.120	10000.000	1087.490	5000.000	98.264
Jun	140	725	8000.000	415.020	11485.711	747.568	13928.570	524.598	11158.570	1690.919
Jul	725	1302	11714.285	225.245	14228.570	175.180	17857.141	1313.250	13211.426	1699.982
Aug	1302	1366	15428.570	656.500	16971.426	86.435	17857.141	728.500	15264.281	1196.185
Sep	1366	1432	15428.570	1234.500	19714.281	1018.548	17857.141	1350.750	15264.281	1496.553
Oct	1432	1496	15428.570	781.000	22457.141	321.185	17857.141	919.750	19369.996	1753.685
Nov	1496	1495	15428.570	653.250	25199.996	1083.750	17857.141	680.500	19369.996	2976.997
Dec	1495	1364	15428.570	468.500	25199.996	2829.381	17857.141	435.500	17317.141	5838.477
Jan	1364	779	15428.570	1684.755	19714.281	2893.820	17857.141	1721.760	11158.570	5551.277
Feb	779	643	11714.285	323.750	16971.426	2792.074	13928.570	1894.913	9105.711	4867.984
Mar	643	66	11714.285	1635.005	11485.711	2815.820	10000.000	209.500	9105.711	3928.018
Apr	66	1	8000.000	307.250	8742.855	1587.700	10000.000	391.250	7052.855	2938.197

Table 5.10 Optimal trajectory for the second simulation

5.7.1 Comparison of the results

A major advantage of the principal component based approach is that the CPU time needed is small. The classical dynamic programming method requires 44 minutes on an IBM-3081 to solve the four reservoir system, whereas it takes only 12 seconds to solve the same problem on the same computer with the proposed approach. However, this huge difference in CPU time can be interesting only if the two methods give comparable results. In the classical approach, 40279 GWh of thermal energy are needed for the twelve period problem, whereas 40322 GWh are required for the reduced problem (see Section 5.5). This corresponds to a difference of 0.1% between the two solutions which is negligible, especially when considering the stochastic nature of the problem. Moreover, it should be mentioned that the discretization in more than 8 equally spaced values in the classical method may give a lower value for the objective function. However, this will not greatly affect the results and even a difference of 5% could be considered acceptable since the CPU time is enormously reduced.

In addition, the outflows, and especially the downstream reservoirs (LG3 and LG2) have similar patterns with both approaches (Figures 5.4 and 5.14). This is an important point because it means that the policies obtained in the reduced approach can be implemented as well as those given by the classical stochastic approach.

Finally, the proposed approach gives a global feedback just like the classical technique does since all the constraints are respected including those on storage.

5.7.2 Selection of Variable Number of Storage Components

The example given in Section 5.5 has four state variables: one inflow component and 3 storage components. According to Table 4.3, the first component of the inflows accounts for at least 99.98% of the total variance. For the storages, the contribution of the components vary from 96% if 3 components are selected to 82% if only 2 are retained (Table 4.4). Solving a stochastic dynamic programming problem with fewer state variables is always interesting. However, the degree of accuracy in every period should be respected when reduction is made. To see the difference with the case solved in Section 5.5, two additional selections were made. The first consists in solving the problem with one inflow and 2 storage components, that is a reduction from 8 to 3 state variables. The second consists in selecting a different number of variables in each period so as to respect a percentage of at least 88% of the total variance. In that case, and according to Tables 4.3 and 4.4, the number of components selected will be similar to that shown in Table 5.11.

The explicit distribution of the inflows is the same than the one given in Table 5.1. Thus, when 3 state variables (1 inflow and 2 storage components) are retained, the thermal production is 40916 GWh which is very close to that (40322 GWh) when considering 4 state variables (1 inflow and 3 storage components). Therefore, there is a difference of 1.47% between these two results. To establish a basis for comparison with the previous case, we will consider the first simulation. The corresponding trajectories for the storages and outflows are given in Table 5.12 and illustrated in Figure 5.15, for the optimal storages, and in Figure 5.16 for the optimal outflows.

O B J E C T I V E : .40916E+05

Month	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
Branch	2	1	1	2	2	1	2	2	1	1	2	2

SIMULATION #1

Month	State	Pt	Storage1	Outflow1	Storage2	Outflow2	Storage3	Outflow3	Storage4	Outflow4
May	1	17	8000.000	81.681	6000.000	764.981	10000.000	705.265	5000.000	1274.044
Jun	17	13	11031.129	53.545	7439.504	283.415	14953.691	923.584	8003.121	1057.317
Jul	13	8	13400.312	331.726	8731.305	425.167	15329.246	522.908	10034.281	338.600
Aug	8	8	15082.262	383.943	10033.598	288.123	16022.262	238.822	13225.727	656.744
Sep	8	8	17519.145	167.632	12685.883	147.028	19227.629	141.250	14892.109	603.327
Oct	8	9	20284.934	1020.888	14923.328	562.312	22363.246	1515.236	15919.973	2988.171
Nov	9	7	19641.836	1177.665	17661.352	878.163	20767.707	511.989	14852.957	4129.965
Dec	7	14	19247.875	452.912	20189.520	3347.624	22208.695	3019.944	9463.320	4094.474
Jan	14	15	20044.324	2117.075	13775.602	3283.077	15985.293	2166.232	16727.977	4632.246
Feb	15	10	15171.000	2228.338	11147.246	2420.922	10865.652	297.680	19370.000	4170.715
Mar	10	2	10319.871	1114.557	11015.867	3064.741	10596.504	426.666	16159.027	5766.047
Apr	2	1	8000.000	307.683	6213.160	612.512	10015.895	397.672	10431.871	3272.877

Table 5.12 Optimal trajectory for the first simulation with three state variables

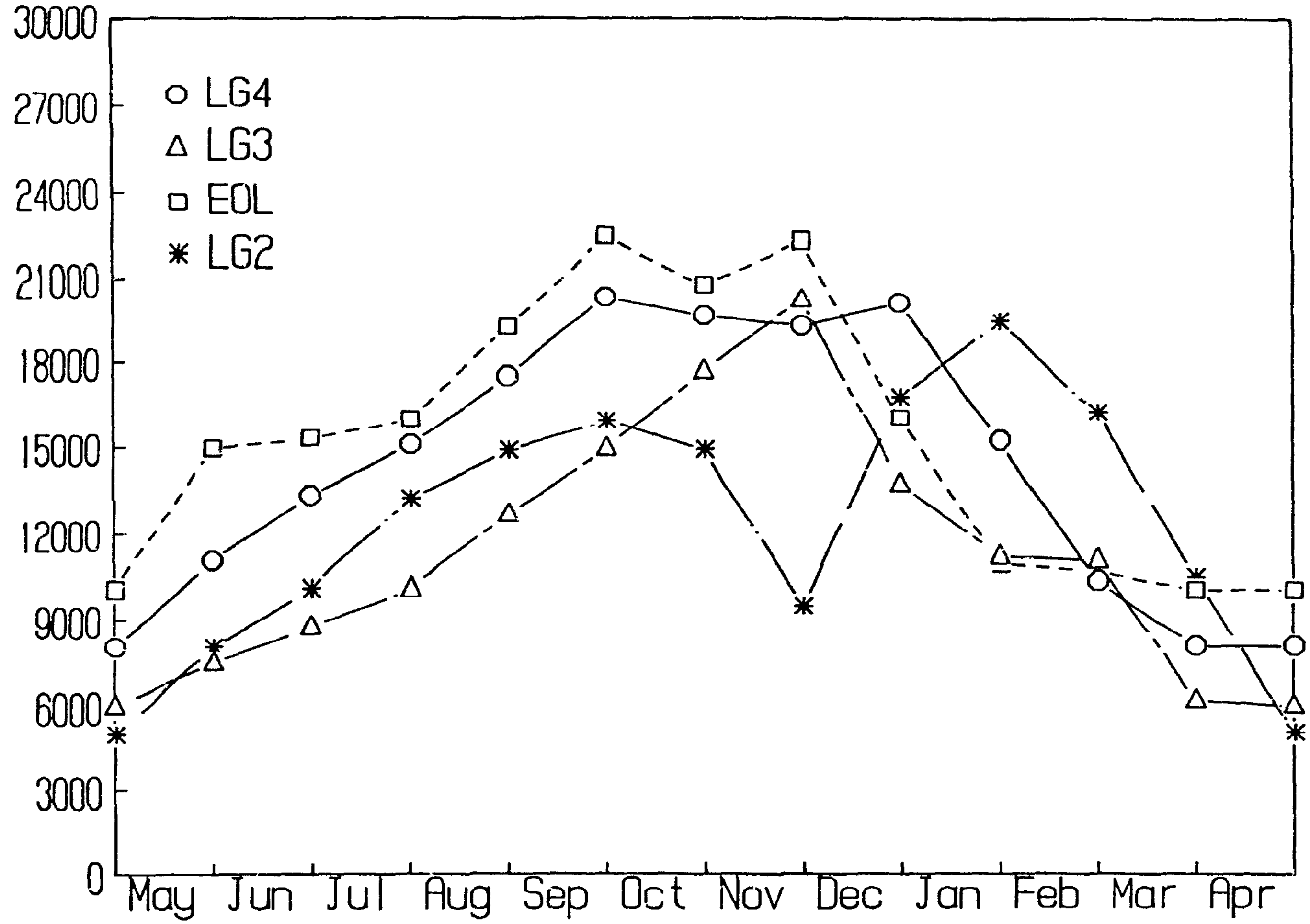


Figure 5.15 Optimal storages with 3 state variables

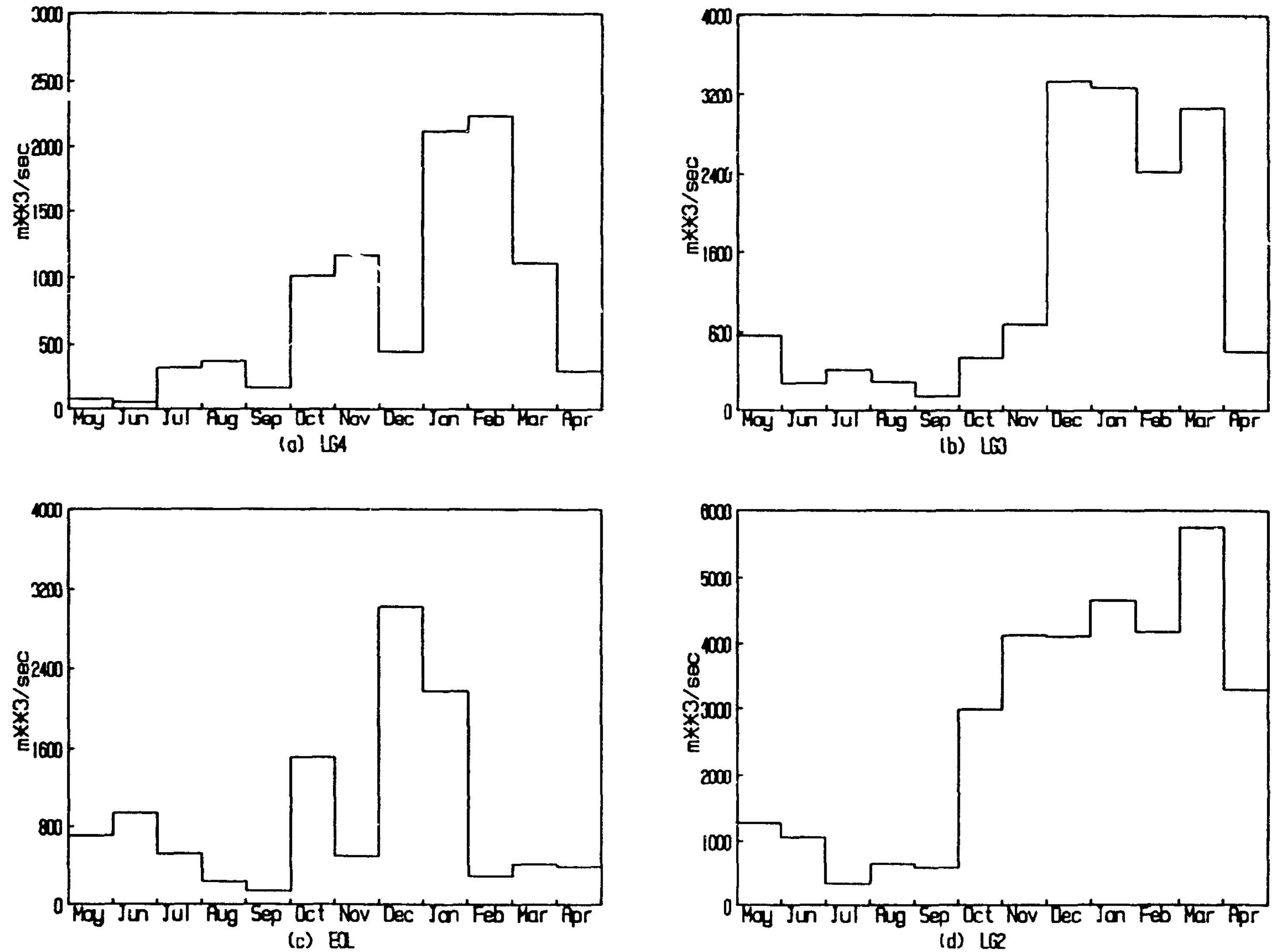


Figure 5.16 Optimal outflows with 3 state variables

Period	Inflows	Storages	Total
May	1	2	3
Jun	1	2	3
Jul	1	3	4
Aug	1	3	4
Sep	1	3	4
Oct	1	3	4
Nov	1	2	3
Dec	1	1	2
Jan	1	1	2
Feb	1	1	2
Mar	1	1	2
Apr	1	4	5

Table 5.11 Number of Components

The optimal policies for this case are very close to those obtained in Section 5.5. These policies can easily be compared in Figure 5.17 (a) and (b) representing, for LG4 and LG3 respectively, the results for 4 and 3 state variables. Although the operating rules are quite similar, the CPU time presents a big difference. With 3 state variables the computing time is 1.2 sec whereas for 4 state variables, it jumps to 11.58 sec. We should recall that the cost showed a difference of 1.47%, which is acceptable. However, it is still possible to minimize this difference by considering a variable selection adapted to each period. The computer program offers the possibility to perform stochastic dynamic programming with a variable number of states. Therefore, when considering the number of components shown in Table 5.11, the total thermal energy needed is 40747 GWh, a difference of almost 1% with the case presented in Section 5.5. The optimal policies are also very close. Table 5.13 summarizes the different storage levels and outflows shown in

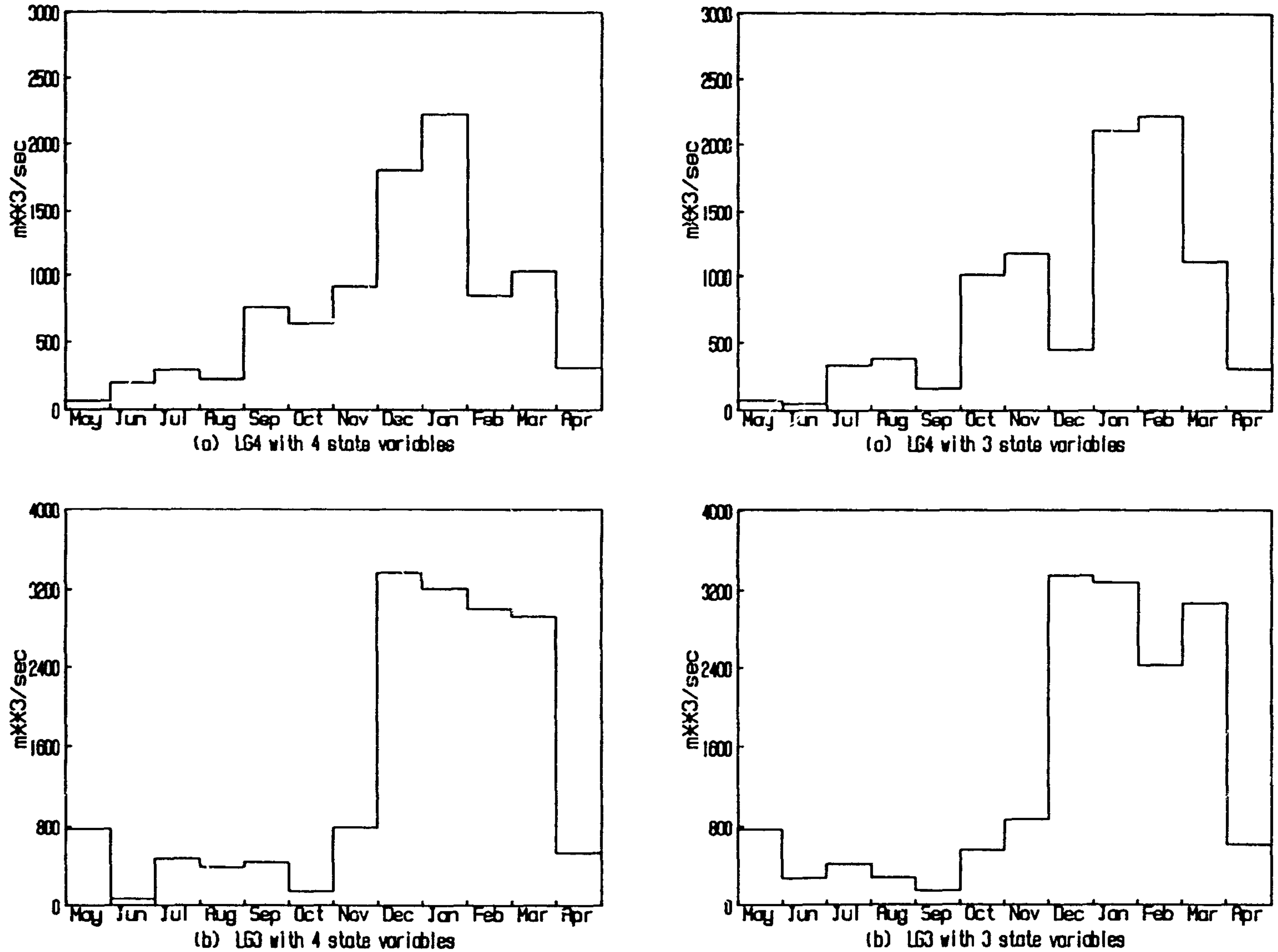


Figure 5.17 Comparison of the outflows

OBJECTIVE: .40747E+05

Month	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
Branch	2	1	1	2	2	1	2	2	1	1	2	2

SIMULATION # : 1

Month	State	Pt	Storage1	Outflow1	Storage2	Outflow2	Storage3	Outflow3	Storage4	Outflow4
May	1	17	8000.000	81.681	6000.000	764.981	10000.000	705.265	5000.000	1274.044
Jun	17	13	11031.129	53.545	7439.504	283.415	14953.691	923.584	8003.121	1057.317
Jul	13	38	13400.312	276.097	8731.305	386.329	15329.246	552.756	10034.281	313.001
Aug	38	38	15231.258	371.784	9988.625	286.299	15942.316	259.122	13270.211	656.375
Sep	38	38	17700.707	141.037	12613.230	141.889	19093.312	179.850	14987.066	597.504
Oct	38	43	20535.430	947.961	14795.062	312.416	22128.879	1286.445	16116.754	2472.059
Nov	43	7	20087.660	1349.665	18007.082	1183.547	21146.133	657.987	15149.977	4695.937
Dec	7	3	19247.875	430.520	20189.520	3460.312	22208.695	2702.099	9463.320	4526.023
Jan	3	3	20104.301	1894.396	13413.801	3311.083	16836.609	1917.009	15022.621	4364.258
Feb	3	2	15827.398	2467.851	10114.012	2626.364	12384.488	430.963	17789.914	4785.684
Mar	2	1	10396.840	1143.294	10065.059	2738.470	11792.902	873.353	13910.645	5046.762
Apr	1	1	8000.000	307.683	6213.203	612.528	10015.887	397.669	10432.555	3273.154

Table 5.13 Optimal trajectory with variable number of states

Figures 5.18 and 5.19, whereas Figure 5.20 shows a comparison of the outflows for LG2 with 4, 3 and variable number of states. The CPU time required for this last case is 3.89 sec, which is 3 times faster than the case with 4 state variables.

5.7.3 Explicit Distribution of 5 Classes

The distribution of the inflows was considered in 2 classes or branches. This was helpful for the comparison with the classical stochastic dynamic programming. However, the number of classes usually used varies from 5 to 15 classes [Miller and Freund, 1977]. In this section, the performance of the reduced stochastic dynamic programming approach will be studied with a distribution of 5 classes. However, before doing so, it is important to note that only 1 component of the inflows is selected, and we can consider 2, 3 or variable number of storage components as discussed in the Subsection 5.7.2. For the sake of presentation, the same selection as the one chosen in Section 5.5 will be considered: 1 inflow component and 3 storage components.

A histogram of the distribution is illustrated in Figure 5.21 for the month of May, while the different classes and the corresponding probabilities are shown for the twelve period problem in Table 5.14. Thus, the optimal thermal energy produced will be 40654 GWh as given in Table 5.15. The results of the first simulation are also given in the same table. In this case, the program starts by selecting a class of inflows. After that, the optimal trajectory is extracted for each month. Figure 5.22 illustrates the optimal storages, whereas Figure 5.23 shows the optimal outflows. The CPU time is very small: only 27 seconds are required to solve the 4 reservoir problem with 5 classes of inflows on

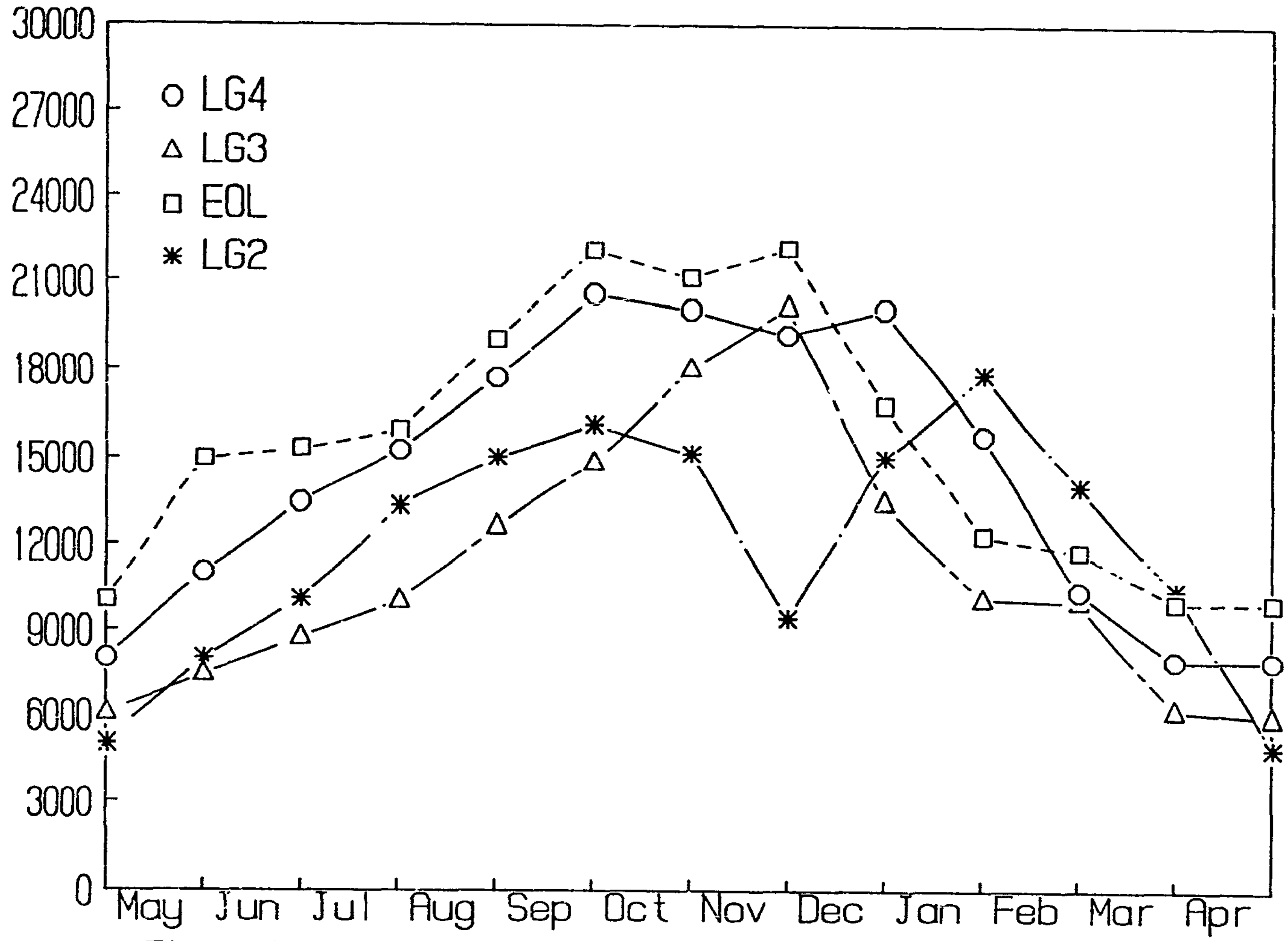


Figure 5.18 Optimal storages with variable number of states

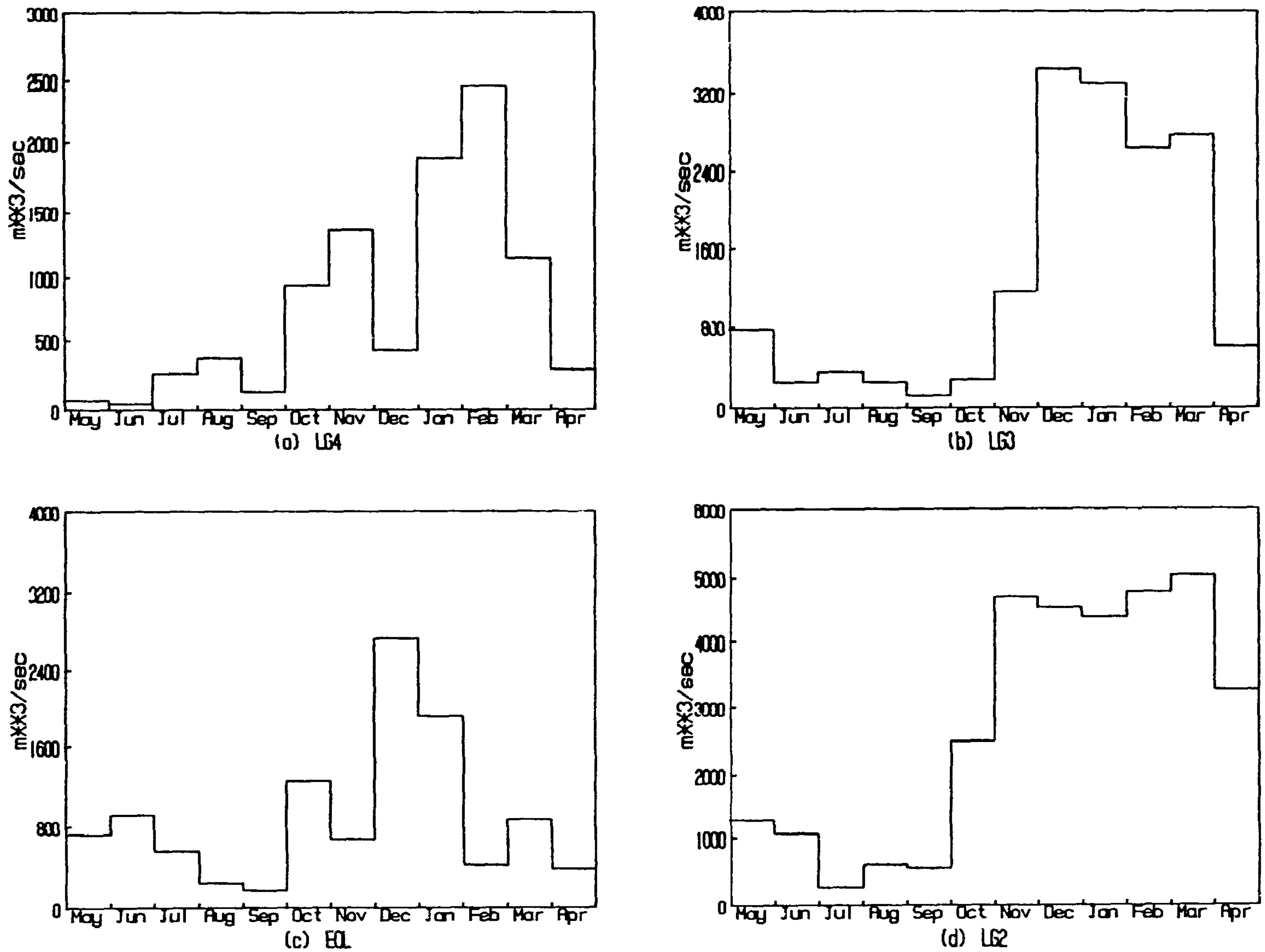


Figure 5.19 Optimal outflows with variable number of states

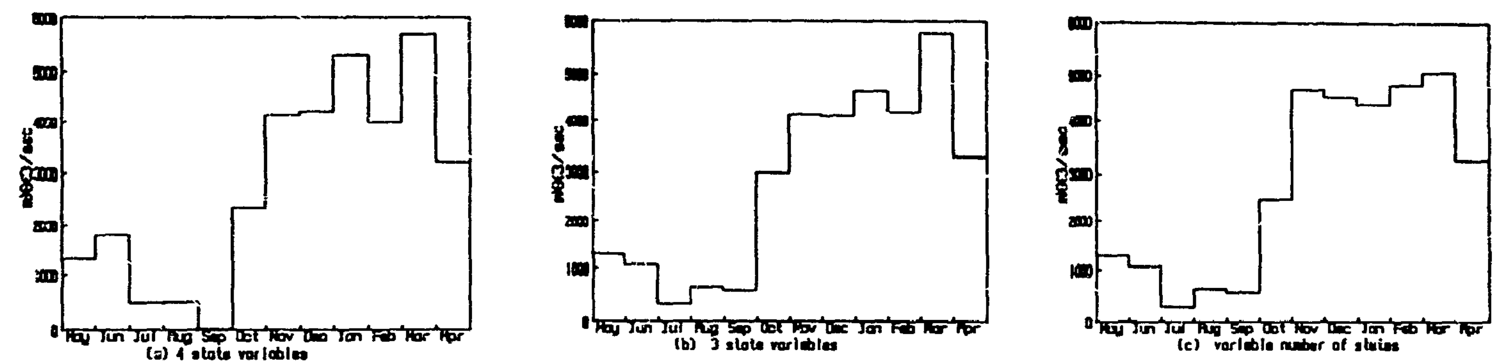


Figure 5.20 Optimal outflows for LG2

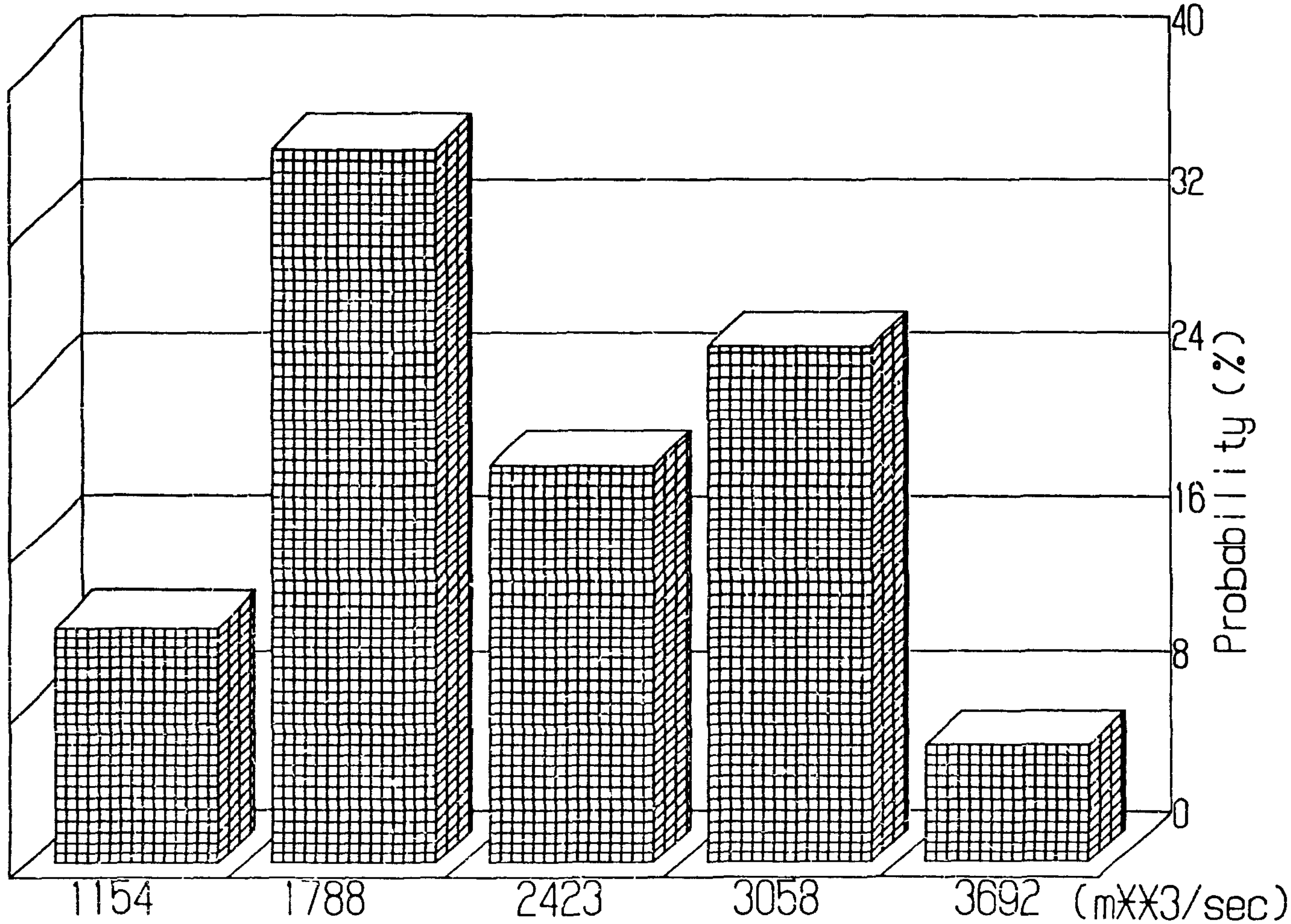


Figure 5.21 Histogram of the inflows in May

	Class1	Prob1	Class2	Prob2	Class3	Prob3	Class4	Prob4	Class5	Prob5
May	1153.76	0.120	1788.37	0.360	2422.98	0.200	3057.59	0.260	3692.20	0.060
Jun	1261.33	0.120	1892.85	0.140	2524.36	0.300	3155.88	0.280	3787.40	0.160
Jul	1182.21	0.140	1586.54	0.340	1990.87	0.300	2395.20	0.160	2799.53	0.060
Aug	811.53	0.060	1256.72	0.200	1701.90	0.340	2147.08	0.240	2592.27	0.160
Sep	950.13	0.120	1315.26	0.160	1680.39	0.240	2045.52	0.320	2410.65	0.160
Oct	1151.17	0.160	1518.56	0.180	1885.95	0.320	2253.34	0.240	2620.73	0.100
Nov	927.06	0.080	1182.30	0.260	1437.54	0.260	1692.78	0.240	1948.02	0.160
Dec	624.96	0.060	808.35	0.240	991.75	0.420	1175.15	0.240	1358.54	0.040
Jan	393.53	0.020	488.86	0.180	584.19	0.380	679.52	0.360	774.85	0.060
Feb	298.30	0.040	360.93	0.220	423.55	0.380	486.17	0.340	548.80	0.020
Mar	268.24	0.060	304.49	0.240	340.74	0.340	376.99	0.300	413.24	0.060
Apr	270.93	0.180	363.09	0.280	455.25	0.360	547.42	0.140	639.58	0.040

Table 5.14 Distribution of the random variable

OBJECTIVE: .40654E+05

Month	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
Branch	1	4	3	2	5	4	5	5	3	2	2	1

SIMULATION #1

Month	State	Pt	Storage1	Outflow1	Storage2	Outflow2	Storage3	Outflow3	Storage4	Outflow4
May	1	125	8000.000	6.628	6000.000	444.459	10000.000	726.020	5000.000	1502.346
Jun	125	83	9147.891	199.074	6000.000	522.886	10510.270	9.402	5000.000	1150.344
Jul	83	63	13193.871	76.605	8594.426	368.021	15523.375	563.978	6386.949	92.712
Aug	63	58	16433.484	86.405	9894.809	251.344	16818.836	320.261	10724.434	751.123
Sep	58	32	18130.293	481.000	10786.270	601.378	18100.859	1222.265	11363.992	594.770
Oct	32	43	20494.035	1384.750	12938.172	382.518	18883.234	607.500	16630.070	2351.706
Nov	43	38	20087.660	1243.637	18007.082	979.143	21146.133	1131.006	15149.977	4220.488
Dec	38	49	19812.473	1659.247	20635.395	2731.096	21284.262	2817.274	11578.496	3625.025
Jan	49	50	17604.301	1558.026	19254.660	3289.906	15813.758	1974.104	18038.980	5300.750
Feb	50	69	14164.707	534.727	15505.148	3031.222	11383.723	482.097	18510.426	5083.187
Mar	69	5	13435.262	2225.251	9815.336	3282.360	10688.859	396.264	15028.254	5833.707
Apr	5	1	8000.000	146.147	7315.863	759.614	10070.687	213.244	9544.270	2805.200

Table 5.15 Optimal trajectory for the first simulation

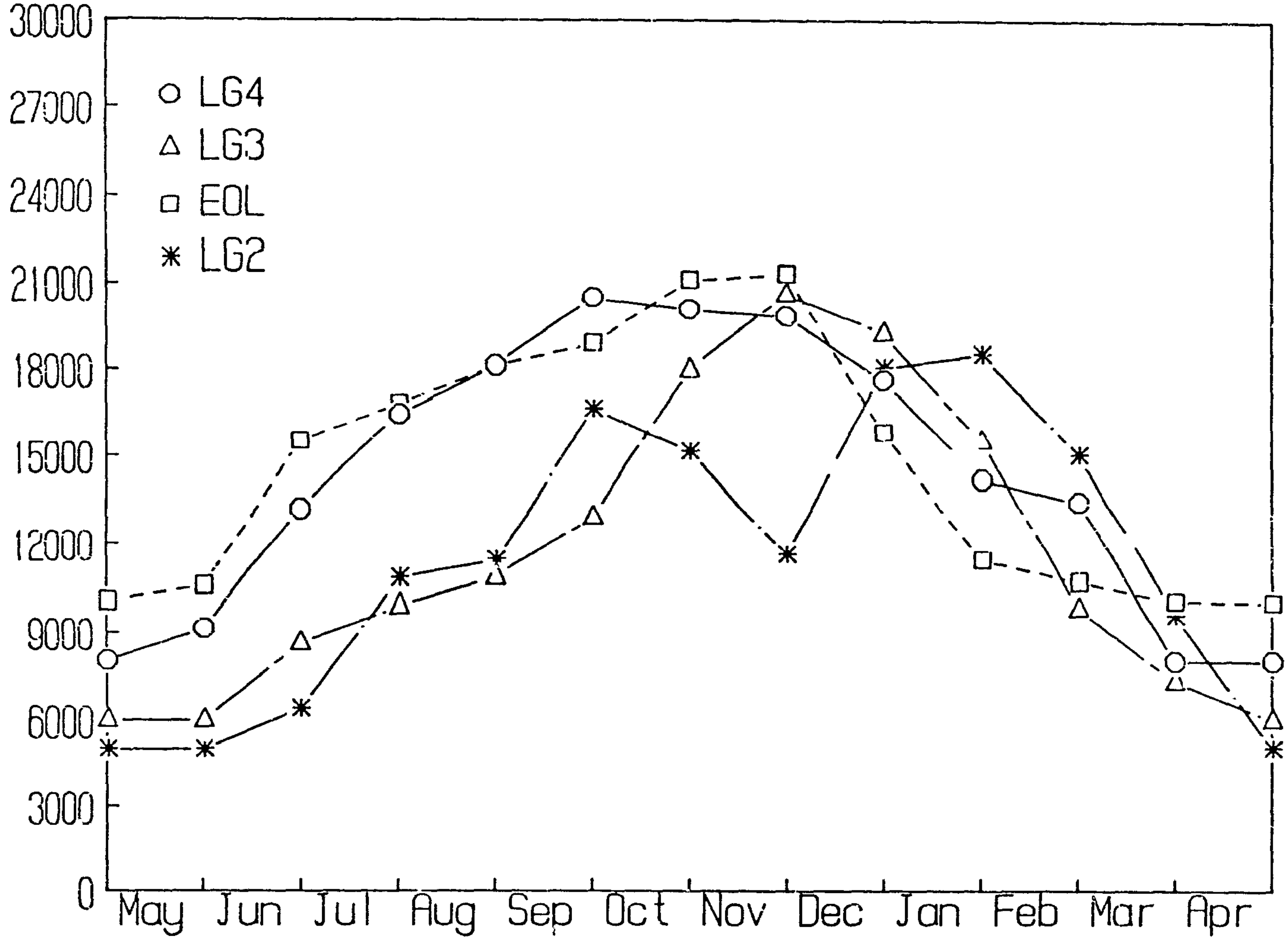


Figure 5.22 Optimal storages

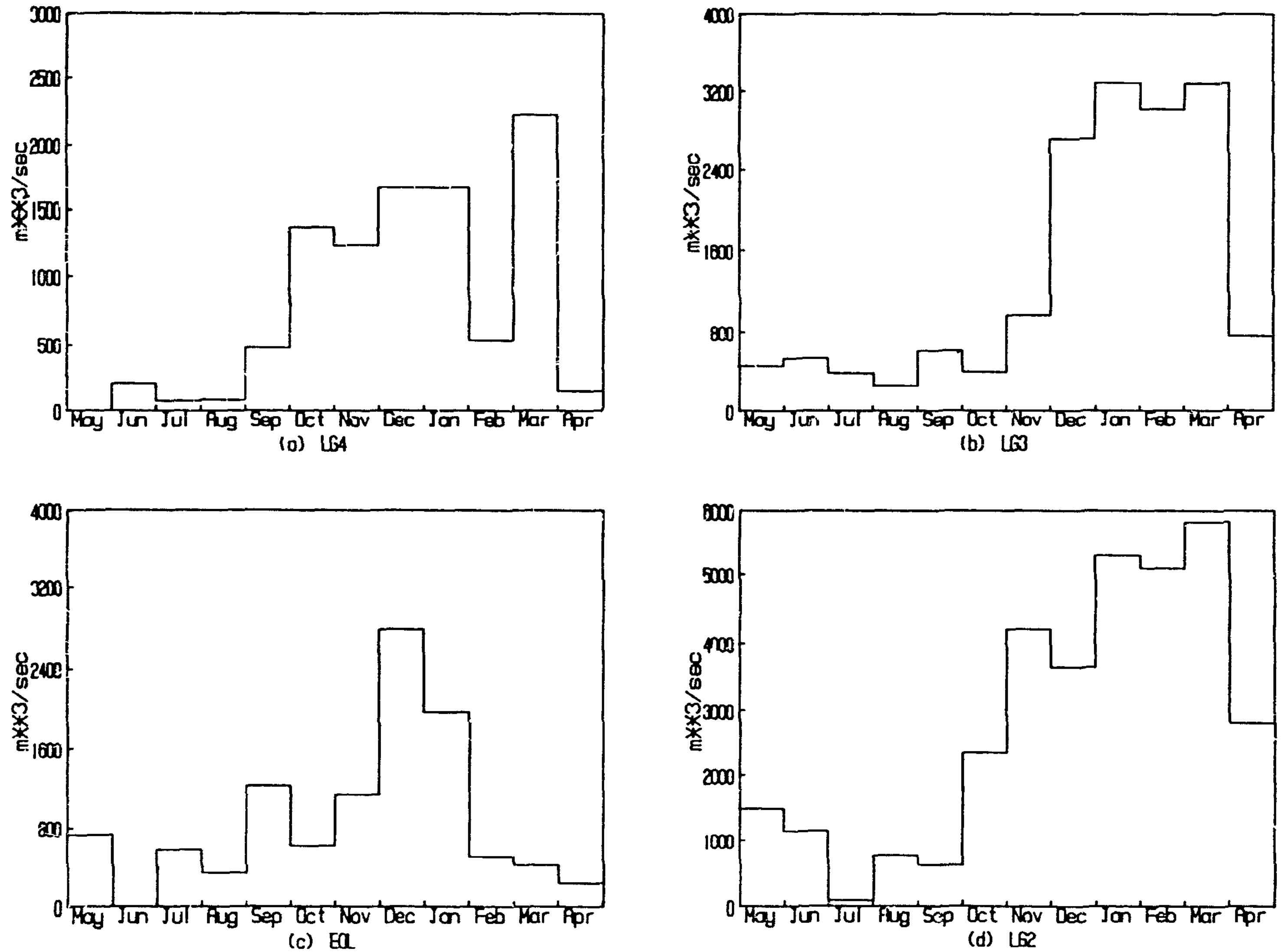


Figure 5.23 Optimal outflows

an IBM-3081.

5.8 Conclusion

We have presented, in this chapter, an application of the principal components approach to the determination of the monthly operating policy of a power system of n reservoirs. This method consists in reducing the number of variables in order to be able to use stochastic dynamic programming.

In addition to being simple and easy to program, this technique has the advantage of being very fast. In fact, in the sample application of four reservoir system, only 12 seconds of CPU time were required on an IBM-3081 to solve the problem for 2 classes of inflows, and 27 seconds for 5 classes. On the other hand, with the classical dynamic programming approach, 44 minutes were required to solve the 2 class problem and 1 hour of CPU time was not sufficient to obtain the results for the 5 class case.

The program that was written to solve this problem offers the possibility of selecting a different number of components in each period. This characteristic, studied in Section 5.7, determines the number of components according to a fixed percentage of the variance. Thus, it is possible to select n components in a period and only 1 component in the next one.

Finally, this approach gives global feedback solutions. In fact, the results of Section 5.5 prove that all the constraints are respected. In addition, the decisions obtained are functions of the explicit distribution of the inflows and consequently the explicit distribution of the states.

CHAPTER VI

SUMMARY AND CONCLUSIONS

6.1 Summary of Results

This work has presented a method for reducing the number of state variables in the stochastic long-term multireservoir operating problem. The approaches which have so far been proposed to deal with the problem were based mainly on the application of dynamic programming. These methods, briefly presented in Chapter I, have attempted to transform the large-scale problem into one or a series of small-scale problems using the aggregation, decomposition or projection techniques. The main draw-back of these methods is that they only give local feedback solutions.

Principal components analysis can be used to transform the large-scale problem into a small-scale one without neglecting the individual constraints on each reservoir, especially the storage level constraints. Thus, stochastic dynamic programming can be performed efficiently and global feedback solutions can therefore be found.

The proposed method was tested on a system of four reservoirs representing Québec's La Grande river installations. The problem was to minimize the expected cost of thermal energy.

The model developed was a nonlinear one. A piecewise linearization was used and penalty functions to limit high variations of the discharged water were added. This model is presented in Chapter II.

An important step of the principal components approach consists in the deterministic optimizations for several inflow sequences. This task was done in Chapter III with IBM's MPSX/370 package. The reduction process using principal components analysis was then performed on the optimal deterministic results as explained in Chapter IV. Finally, in Chapter V the reduced problem was solved using stochastic dynamic programming.

The performance of the proposed approach was thereafter compared to the classical dynamic programming method, that is without any reduction. Important results were obtained from that comparison. First, a great reduction was observed in the CPU time required to solve the problem. Secondly, the total cost and the operating rules were found to be extremely close. In fact, the difference in the objective function was just over 1% for the same four reservoir system, and the operating rules found showed similar patterns. Both techniques have the characteristics of giving low discharge in summer and very high discharge in winter, especially for the downstream reservoirs (LG3 and LG2). This is only logical because the demand gets very high in winter. Consequently, the water is stored during summer for future use in winter. Many graphics were also given in Chapter V to illustrate the feature of the reduced approach and to facilitate the comparison with the results obtained from the classical technique.

Taking advantage of the fast computing time and of the satisfactory operating rules, other cases were studied in Chapter V. For example, the reduced problem was solved with different number of state variables in each period, and then it was solved again but with a realistic explicit distribution of inflows in 5 classes.

6.2 Conclusions

The method proposed for reducing the number of state variables in the stochastic large scale problem can be applied wherever correlation exists between the inflows to two reservoirs or between the reservoir contents. Naturally, the higher the correlation, the greater and better the reduction will be.

The correlation between the inflows can vary throughout the year. In Québec, for instance, the correlation is very high in winter when everything is frozen and low in the spring during the thaw. The reduction of the state space will therefore be more important in winter months than in spring, which is totally acceptable. Dynamic programming does not require that the number of state variables be the same in every period. In fact it is more important to have the same degree of precision in every period than the same number of state variables. For instance, if the new set of variables is to account for 95% of the sample variance, then the reduction in every period should be made in order to respect that percentage.

Since linear programming was used to solve the deterministic model, the reservoir contents may have weak correlation due to the bang-bang solution, in which a plant is run at maximum capacity one month and shut down the next. Therefore, introducing a penalty function on the variation of the outflows from period to period can insure a better correlation. However, a problem arises when determining that penalty factor because it can not be found according to fixed rules. Moreover, its value must not exceed 10% of the cost of the thermal energy.

The deterministic optimizations allow us to find a set of reservoir contents, for many inflow sequences. The size of the sample of inflows is very important. As a matter of fact, the principal components can be very different for two different samples of data. Nevertheless, it was shown that the sample size is "representative" for large number of sequences. However, for economical considerations, a sample size of 30 may be considered as being sufficiently large.

Principal components analysis does not have to be applied to all installations at the same time. If two rivers with several installations on each have different flow patterns, then principal components analysis can be applied to each river separately so as to find a good reduced model of each.

Principal components analysis is very easy to implement and manipulate because it involves only linear relations. Its combination with the optimization model gives interesting results especially when compared with the existing classical technique which is limited to a small number of state variables. The proposed technique can therefore be applied without any problem to large-scale systems. The only condition is that the state variables must be interdependent. In general, interdependency does exist between reservoirs located on the same rivers or on nearby rivers with similar flows patterns. In this perspective, the proposed technique can be of great utility.

6.3 Further Research Recommendations

Further research work can proceed along several directions:

1- The piecewise linearization introduces some inaccuracy because there are no rules established for the selection of either the optimum grid size or the optimum number of grids. In this case, the possibility of considering nonlinear models can be investigated knowing that dynamic programming can deal with linear as well as nonlinear models. In addition, the penalty functions are then not necessary and can be neglected in nonlinear models. Therefore, more accurate results can be obtained.

2- The only random variable in the present work is the natural inflow in each period. However, there are other variables that can be regarded as stochastic. For example, it can be interesting to take into account the stochasticity of the demand and to solve a stochastic model having both the natural inflows and the demand as random variables.

3- Although the technique proposed in this thesis is applied to long-term problems, it is also possible to apply the method to short-term problems as well. In that case, the model should take into account the water head variations. For long-term operating problems water head can be considered constant whereas, this assumption is of no value for short-term problems.

If short-term problems are considered, care should be taken while using the covariance or the correlation matrices to perform the principal components analysis. That is because the correlation matrix imposes the same weight for large and small reservoirs. Therefore, this case should be discarded for long-term problems, while it can yield interesting results when compared with the covariance matrix in short-term problems.

4- Principal components analysis is performed only in space. In other words, the analysis is made only for reservoirs taken at a given time. However, since correlation does exist from period to period, it may be interesting to reduce the model in time in order to find a shorter operating horizon.

The reduction in space and in time together may add another dimension to the problem, and therefore can open an important research direction in the future.

5- Finally, the method proposed in this thesis can be used for reservoirs serving other purposes than the production of electricity. This would be a fruitful direction for future research.

APPENDIX A

PIECEWISE LINEARIZATION

A.1 General Description

The nonlinear function must be assumed to be convex in order to reach a "global" optimum. Any nonconvex function may result in a "local" optimum. For example, Figure A.1 depicts a two-dimensional model that contains two linear constraints, one convex nonlinear constraint, and a linear objective function to be maximized. The feasible region is shaded, and the optimal solution is indicated by the position of the objective function. Figure A.2 depicts a similar model except that the nonlinear constraint is not convex. Because of the nonconvexity of the nonlinear constraint, both solutions shown appear optimal to the program, and either one may be reached first. Figure A.3 shows that a nonlinear function can be approximated by a piecewise linear function, known as a polygonal approximation. The function and two possible polygonal approximations are shown in this figure. Each polygonal approximation is represented by linear equations, together with certain logical restrictions on the variables in the equations. The solution reached is an approximation of the true solution.

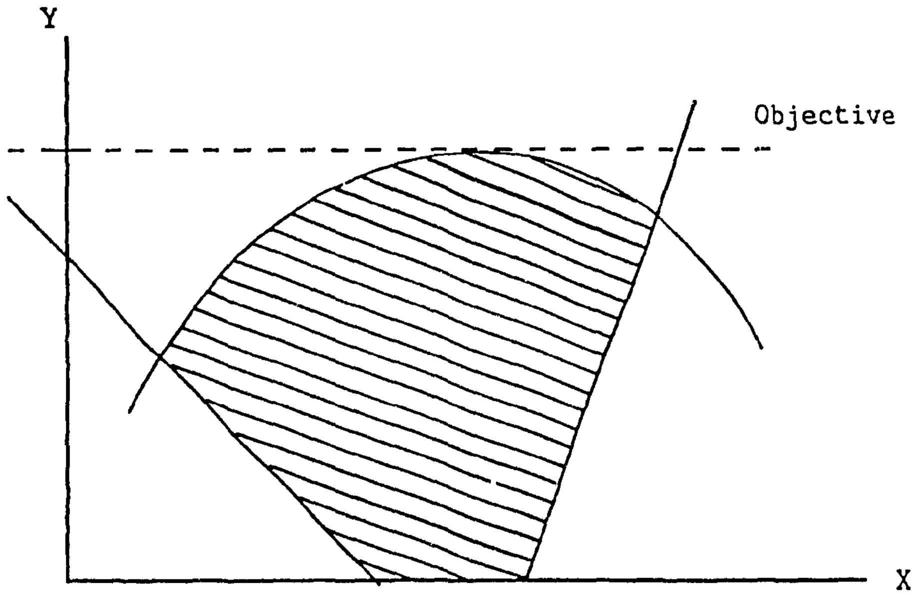


Figure A.1 Convex feasible region

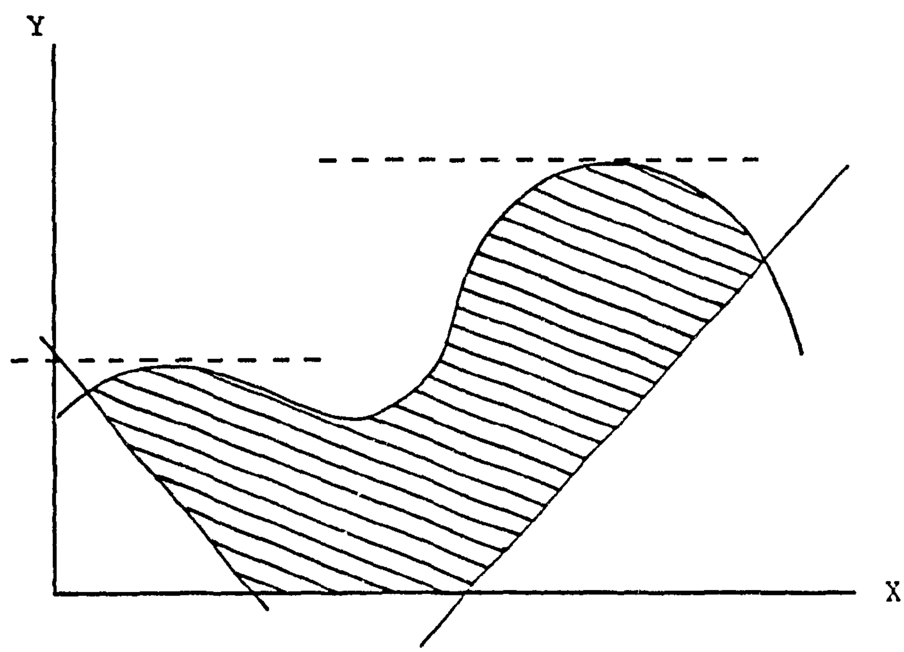


Figure A.2 Nonconvex feasible region

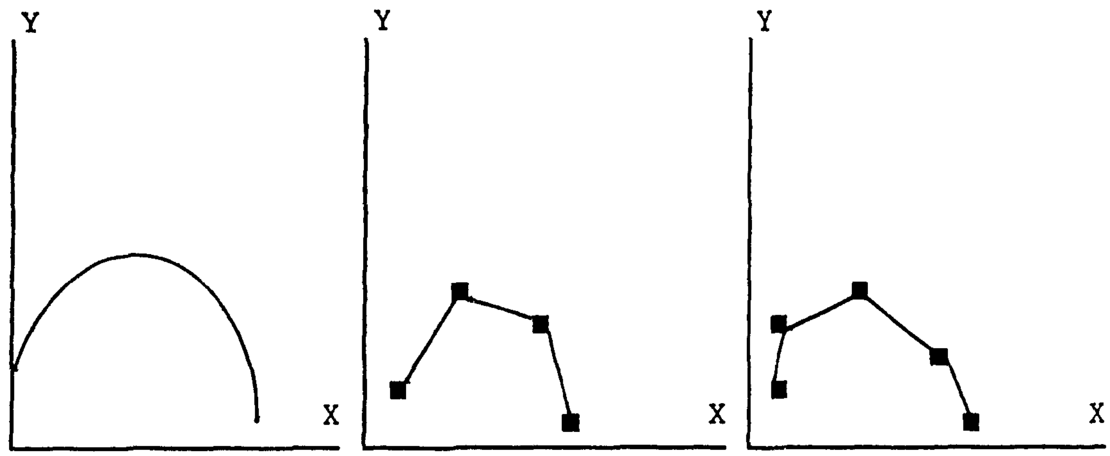


Figure A.3 Nonlinear function and polygonal approximation

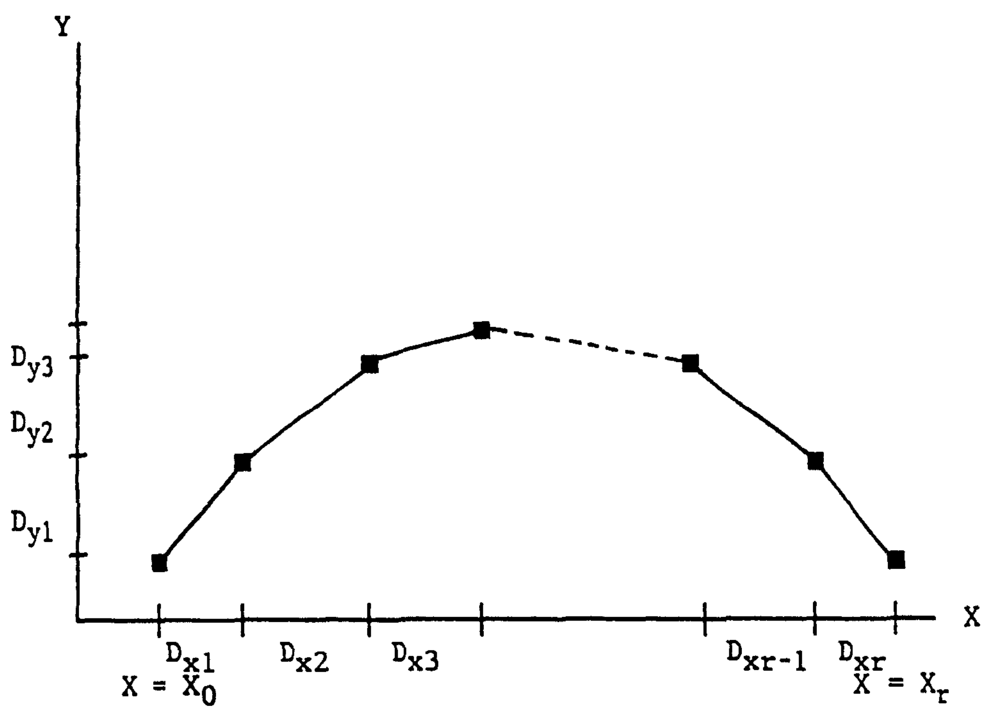


Figure A.4 Polygonal approximation of a nonlinear function

A.2 Development of Approximating Equations

Figure A.4 is a graph of a polygonal approximation of a nonlinear function $Y = F(X)$, defined in the interval $[X_0, X_r]$.

The grid is defined by a set of r+1 points on the X-axis. The lengths of the resulting intervals on the X-axis are D_{x1}, $D_{x2}, \ldots, D_{xr}$, and the lengths of the resulting intervals on the Y-axis are D_{y1}, $D_{y2}, \ldots,$ D_{yr}. The variable X can be developed as a function of "special" variables X_1, X_2, ..., X_r, where X_1 defines the first interval of length D_{x1}, X_2 the second interval of length D_{x2}, and so on.

Any value of X between X_0 and X_r can be expressed in terms of the equation

$$X = X_0 + D_{x1} \cdot X_1 + D_{x2} \cdot X_2 + \ldots + D_{xr} \cdot X_r \qquad (A.1)$$

if the special variables are defined as follows:

$$X_1 = X_2 = \ldots = X_{i-1} = 1 \qquad (A.2)$$

$$0 \leq X_i \leq 1 \qquad (A.3)$$

$$X_{i+1} = X_{i+2} = \ldots = X_r \qquad (A.4)$$

The value between 0 and 1 of X_i is that fraction of the interval i covered by the variable X. For example, if X has a value at the midpoint of interval i, the special variable X_i has the value 0.5. If X is at 3/4 of the length of interval i, the special variable X_i has the value 0.75. Equation (A.1) is referred to as the grid equation.

Similarly, the function Y can be expressed in terms of the same special variables and the lengths of the resultant intervals along the

Y-axis as follows:

$$Y = Y_0 + D_{y1} \cdot X_1 + D_{y2} \cdot X_2 + \ldots + D_{yr} \cdot X_r \tag{A.5}$$

Because of the linearity of the approximating functions, the values of X_1 to X_r that satisfy the grid equation also satisfy equation (A.5), known as the functional equation.

A.3 Example

The following example illustrates, for a specific nonlinear function, the equations relating the variable, the nonlinear function $Y = F(X)$ and the special variables.

The problem definition data are :

a) Nonlinear function : $Y = X^2$

b) Range of X value considered : $X = 0$ to $X = 1$

c) Defined set of special variables: X_1, X_2, X_3, X_4, X_5

d) Defined interval lengths : 0.3, 0.1, 0.2, 0.3, 0.1

The information required to develop the equations for this problem is contained in Table A.1.

Interval	k_1 = X at Beginning of Interval	k_2 = X at End of Interval	$D_{xi} = k_2 - k_1$	Y_1 = Y at Beginning of Interval	Y_2 = Y at End of Interval	$D_{yi} = Y_2 - Y_1$
1	0	0.3	0.3	0.00	0.09	0.09
2	0.3	0.4	0.1	0.09	0.16	0.07
3	0.4	0.6	0.2	0.16	0.36	0.20
4	0.6	0.9	0.3	0.36	0.81	0.45
5	0.9	1.0	0.1	0.81	1.00	0.19

Table A.1 Problem information

The general formulas for the grid and functional equations are:

$$X = X_0 + D_{x1} \cdot X_1 + D_{x2} \cdot X_2 + D_{x3} \cdot X_3 + D_{x4} \cdot X_4 + D_{x5} \cdot X_5 \quad (A.6)$$

$$Y = Y_0 + D_{y1} \cdot X_1 + D_{y2} \cdot X_2 + D_{y3} \cdot X_3 + D_{y4} \cdot X_4 + D_{y5} \cdot X_5 \quad (A.7)$$

The D_x and D_y columns give the required coefficients for the equations. We therefore have:

$$X = 0 + 0.3 \cdot X_1 + 0.1 \cdot X_2 + 0.2 \cdot X_3 + 0.3 \cdot X_4 + 0.1 \cdot X_5 \quad (A.8)$$

$$Y = 0 + 0.09 \cdot X_1 + 0.07 \cdot X_2 + 0.2 \cdot X_3 + 0.45 \cdot X_4 + 0.19 \cdot X_5 \quad (A.9)$$

APPENDIX B

DEMONSTRATION

B.1 Lemma and Proof of Section 4.2 [Johnson and Wichern, 1982]

B.1.1 Lemma

Let $Z^T = [Z_1, Z_2, \ldots, Z_n]$ have covariance matrix W with eigenvalue-eigenvector pairs $(\lambda_1, b_1), (\lambda_2, b_2), \ldots, (\lambda_n, b_n)$ where $\lambda_1 \geq \lambda_2 \geq, \ldots \geq \lambda_n > 0$.

Let $\xi_1 = b_1{}^T Z_1$, $\xi_2 = b_2{}^T Z$, ..., $\xi_n = b_n{}^T Z$ be the principal components. Then

$$\sum_{i=1}^{n} \mathrm{VAR}(Z_i) = \sigma_{11} + \sigma_{22} + \ldots + \sigma_{nn} = \sum_{i=1}^{n} \mathrm{VAR}(\xi_i) = \lambda_1 + \lambda_2 + \ldots + \lambda_n$$

B.1.2 Proof

By definition the trace (tr) of a matrix is the sum of the diagonal elements. Then

$$\mathrm{tr}\,(W) = \sigma_{11} + \sigma_{22} + \ldots + \sigma_{nn} \tag{B.1}$$

and

$$\mathrm{tr}\,(\Lambda) = \lambda_1 + \lambda_2 + \ldots + \lambda_n \tag{B.2}$$

where Λ is the diagonal matrix of eigenvalues.

We know from Section 4.2 that

$$\lambda_i = b^T{}_i \, W \, b_i \qquad i = 1,\ldots,n \tag{B.3}$$

or in another form,

$$W = B^T \Lambda B \tag{B.4}$$

where $B^T = [b_1, b_2,\ldots,b_n]$ and $B^T B = B B^T = I$ (unity matrix).

Then from (B.4)

$$tr\ (W) = tr\ (B^T \Lambda B) = tr\ (\Lambda B^T B) = tr\ (\Lambda)$$

$$= \lambda_1 + \lambda_2 + \ldots, \lambda_n. \tag{B.5}$$

Thus

$$\sum_{i=1}^{n} VAR\ (Z_i) = tr\ (W) = tr\ (\Lambda) = \sum_{i=1}^{n} VAR(\xi_i)$$

B.2 Synthetic Inflow Generator. [Pronovost, 1974]

The synthetic inflow generator used in this thesis has the following form

$$Y_k = A_k + B_k Y_{k-1} + C_k W_k \tag{B.6}$$

where Y_k represents a column vector of random variables of inflows for period k, W_k a column vector of white noise and A_k a column vector of constants. B_k and C_k are square matrices.

From the historic data, the mean values vector M_k, the correlation matrix R_k for different sites for period k and the correlation matrix P_k between periods k-1 and k are determined as follows.

$$M_k = E[Y_k] \tag{B.7}$$

$$R_k = E[(Y_k - M_k)(Y_k - M_k)^T] \tag{B.8}$$

$$P_k = E[(Y_{k-1} - M_{k-1})(Y_k - M_k)^T] \tag{B.9}$$

Replacing equation (B.6) in (B.7), the mean values vector M_k may then be written as

$$\begin{aligned} M_k &= E[A_k + B_k\, Y_{k-1} + C_k\, W_k] \\ &= E[A_k] + B_k\, E\,[Y_{k-1}] + C_k\, E[W_k] \\ &= A_k + B_k\, E[Y_{k-1}] \\ &= A_k + B_k\, M_{k-1} \end{aligned} \tag{B.10}$$

(B.10) can easily be obtained since A_k is a constant and W_k is a white noise with zero mean.

In the same way, by replacing equation (B.6) in (B.8) and (B.9) respectively, the following relations are obtained

$$R_k = B_k\, R_{k-1}\, B_k^T + C_k\, C_k^T \tag{B.11}$$

$$P_k = R_{k-1} B_k^T \tag{B.12}$$

Then to find A_k, B_k and C_k, the system of three equations (B.10), (B.11) and (B.12) has to be solved. From statement (B.12)

$$B_k = P_k^T R_{k-1}^{-1} \tag{B.13}$$

Substituting (B.13) in (B.10), A_k can be written as

$$A_k = M_k - P_k^T R^{-1}{}_{k-1} M_{k-1} \tag{B.14}$$

and finally, by substituting statement (B.13) in (B.11), the following expression is found

$$C_k C_k^T = R_k - (P_k^T R_{k-1}^{-1})(R_{k-1})(R_{k-1}^{-1} P_k)$$

$$= R_k - P_k^T R_{k-1}^{-1} P_k \tag{B.15}$$

To find C_k, it can be seen that

$$C_k C_k^T = L_k \Lambda_k L_k^{-1} \tag{B.16}$$

is the general form of an eigenvalue problem in which L_k is the matrix of eigenvectors for $C_k C_k^T$ and Λ_k is the diagonal matrix of eigenvalues for $C_k C^T{}_k$. L_k has the property of being orthonormal, or

$$L_k^{-1} = L_k^T \tag{B.17}$$

Then by replacing (B.17) in (B.16), $C_k\ C_k^T$ can be written as

$$C_k\ C_k^T = L_k\ \Lambda_k\ L_k^T = L_k\ \Lambda_k^{1/2}\ \Lambda_k^{1/2}\ L_k^T$$

$$= (L_k\ \Lambda_k^{1/2})(L_k\ \Lambda_k^{1/2})^T \tag{B.18}$$

Therefore

$$C_k = L_k\ \Lambda_k^{1/2} \tag{P.19}$$

APPENDIX C

DATA AND RESULTS OF THE DETERMINISTIC

OPTIMIZATIONS OF CHAPTER IV

Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
1950	773.	1965.	1555.	1000.	354.	790.	640.	487.	354.	241.	178.	218.
1951	479.	470.	459.	289.	640.	617.	702.	490.	328.	218.	176.	249.
1952	799.	954.	634.	934.	742.	923.	682.	515.	343.	232.	173.	184.
1953	484.	779.	663.	586.	728.	674.	476.	382.	261.	181.	142.	176.
1954	436.	824.	555.	345.	530.	575.	467.	360.	229.	164.	130.	133.
1955	555.	762.	501.	292.	433.	592.	547.	433.	306.	207.	167.	150.
1956	275.	1492.	1271.	1359.	1382.	694.	575.	382.	246.	178.	144.	136.
1957	501.	1161.	1008.	841.	946.	963.	612.	399.	244.	184.	144.	125.
1958	793.	1070.	759.	549.	838.	782.	773.	493.	326.	238.	176.	176.
1959	1014.	1334.	1031.	1019.	680.	818.	665.	411.	292.	210.	170.	201.
1960	555.	1031.	646.	892.	923.	864.	838.	595.	343.	246.	181.	147.
1961	456.	694.	572.	394.	391.	878.	765.	521.	328.	238.	176.	164.
1962	527.	951.	646.	436.	597.	547.	490.	413.	283.	190.	147.	142.
1963	651.	1036.	640.	507.	445.	631.	462.	331.	229.	161.	136.	142.
1964	1028.	1090.	663.	447.	580.	470.	433.	422.	266.	198.	164.	195.
1965	462.	1430.	1538.	1133.	1167.	813.	682.	479.	326.	249.	190.	144.
1966	583.	1504.	1158.	1280.	951.	991.	705.	498.	289.	198.	161.	204.
1967	521.	1133.	719.	719.	547.	889.	779.	549.	311.	221.	173.	144.
1968	866.	813.	685.	923.	847.	515.	541.	399.	227.	167.	136.	201.
1969	462.	1223.	1104.	773.	682.	816.	677.	442.	323.	232.	173.	139.
1970	504.	1155.	926.	722.	960.	1014.	906.	796.	530.	340.	249.	193.
1971	481.	564.	637.	637.	646.	677.	479.	377.	229.	153.	116.	142.
1972	317.	886.	779.	677.	813.	886.	711.	399.	241.	181.	142.	105.
1973	1087.	1104.	787.	572.	544.	753.	527.	297.	198.	164.	142.	139.
1974	464.	1283.	827.	445.	479.	711.	688.	408.	238.	193.	173.	139.
1975	549.	1422.	951.	917.	646.	804.	631.	399.	255.	193.	153.	136.
1976	1056.	1334.	654.	513.	416.	453.	445.	306.	204.	153.	130.	142.
1977	631.	1467.	903.	728.	725.	382.	323.	272.	221.	184.	156.	136.
1978	629.	1107.	1053.	1235.	1144.	631.	682.	654.	377.	263.	207.	167.
1979	1410.	1583.	866.	663.	728.	784.	620.	538.	343.	229.	184.	212.

Table C.1 (a) Historic inflow sequences for LG4

Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
1950	731.	1617.	838.	566.	311.	674.	501.	362.	224.	153.	127.	147.
1951	742.	597.	439.	439.	447.	855.	521.	518.	323.	221.	173.	278.
1952	997.	1195.	762.	779.	640.	1096.	487.	479.	289.	207.	161.	147.
1953	824.	697.	484.	680.	566.	578.	501.	340.	218.	150.	125.	246.
1954	719.	1042.	637.	479.	513.	555.	456.	343.	210.	147.	119.	167.
1955	810.	767.	507.	309.	374.	631.	456.	481.	297.	204.	164.	184.
1956	467.	1640.	770.	903.	1042.	816.	535.	501.	303.	229.	173.	122.
1957	699.	1045.	767.	646.	722.	830.	597.	362.	215.	159.	125.	161.
1958	807.	1070.	592.	722.	813.	470.	600.	445.	275.	195.	147.	190.
1959	1022.	1501.	892.	1011.	654.	668.	532.	351.	232.	167.	142.	127.
1960	1184.	1085.	595.	708.	671.	1104.	816.	402.	229.	173.	139.	167.
1961	592.	620.	447.	326.	300.	787.	637.	447.	280.	193.	147.	159.
1962	722.	903.	535.	481.	572.	473.	464.	360.	221.	170.	136.	125.
1963	714.	1090.	572.	464.	348.	538.	411.	283.	176.	127.	108.	116.
1964	1410.	1116.	496.	408.	620.	377.	337.	365.	224.	156.	130.	178.
1965	716.	1416.	1158.	864.	889.	782.	583.	408.	244.	173.	133.	108.
1966	861.	1470.	974.	1167.	974.	946.	612.	377.	212.	164.	147.	218.
1967	583.	1294.	620.	861.	518.	881.	671.	521.	255.	173.	130.	116.
1968	1192.	663.	612.	1141.	917.	447.	547.	408.	232.	170.	136.	331.
1969	504.	1280.	1068.	617.	682.	750.	589.	351.	238.	184.	150.	133.
1970	767.	1195.	714.	889.	765.	1042.	776.	575.	379.	266.	212.	193.
1971	643.	748.	496.	719.	654.	416.	374.	351.	201.	139.	113.	164.
1972	487.	1076.	691.	569.	762.	816.	564.	331.	218.	167.	142.	116.
1973	1470.	971.	589.	416.	578.	767.	481.	297.	198.	147.	116.	125.
1974	564.	1665.	677.	405.	518.	779.	688.	360.	210.	159.	133.	110.
1975	932.	1266.	835.	827.	583.	920.	651.	396.	238.	164.	125.	88.
1976	1455.	1065.	456.	377.	354.	473.	487.	300.	173.	116.	93.	122.
1977	949.	1379.	694.	708.	977.	402.	323.	218.	159.	130.	113.	116.
1978	866.	1036.	366.	1186.	1062.	620.	623.	515.	266.	184.	150.	133.
1979	1297.	1410.	742.	578.	830.	711.	586.	597.	300.	193.	150.	147.

Table C.1 (b) Historic inflow sequences for LG3

Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
1950	1750.	2532.	765.	742.	487.	1104.	756.	479.	300.	212.	161.	283.
1951	1648.	668.	660.	1218.	900.	711.	861.	705.	391.	249.	210.	541.
1952	1775.	1606.	915.	2387.	1532.	830.	793.	671.	416.	297.	300.	283.
1953	1761.	1019.	1034.	943.	1019.	1705.	770.	453.	272.	215.	190.	371.
1954	1903.	1286.	1085.	966.	824.	1354.	878.	561.	328.	221.	181.	127.
1955	1688.	1172.	541.	671.	847.	821.	694.	592.	371.	275.	201.	229.
1956	1263.	2019.	827.	1059.	1107.	991.	909.	654.	354.	241.	156.	261.
1957	1419.	2005.	1640.	988.	784.	1611.	841.	430.	258.	187.	153.	153.
1958	1911.	1801.	790.	912.	1150.	1116.	940.	609.	354.	218.	142.	176.
1959	1965.	2118.	1051.	1628.	1099.	1286.	833.	428.	289.	212.	178.	377.
1960	2047.	1249.	677.	1189.	1017.	1161.	1175.	552.	328.	212.	161.	309.
1961	1532.	915.	648.	626.	637.	1280.	977.	654.	354.	244.	184.	379.
1962	1651.	1385.	541.	552.	816.	790.	835.	643.	365.	241.	181.	164.
1963	1308.	1523.	697.	765.	496.	677.	521.	348.	215.	161.	139.	224.
1964	2778.	1900.	711.	850.	1218.	578.	513.	459.	278.	198.	164.	294.
1965	1498.	1900.	1351.	1447.	1172.	1305.	767.	442.	266.	195.	161.	150.
1966	1693.	2172.	1300.	1572.	1249.	1283.	782.	456.	292.	210.	210.	515.
1967	1509.	2073.	866.	1688.	765.	1555.	1130.	801.	391.	252.	193.	181.
1968	2175.	997.	1008.	1940.	1764.	728.	968.	606.	354.	241.	204.	742.
1969	1215.	2404.	1858.	1022.	1002.	1540.	1002.	515.	331.	252.	195.	184.
1970	1679.	1798.	646.	895.	951.	1444.	1113.	728.	467.	334.	246.	221.
1971	1167.	779.	790.	1422.	1390.	1073.	705.	530.	317.	195.	127.	133.
1972	1014.	1654.	796.	688.	1107.	2149.	1002.	549.	314.	229.	181.	156.
1973	2750.	1441.	1031.	682.	954.	1246.	799.	456.	328.	235.	184.	255.
1974	1487.	3336.	852.	433.	748.	1198.	929.	527.	326.	232.	181.	161.
1975	2115.	1569.	1189.	1141.	850.	1495.	909.	629.	436.	317.	235.	241.
1976	2466.	1455.	535.	682.	739.	855.	1124.	629.	328.	224.	184.	501.
1977	2059.	1693.	745.	971.	2104.	810.	595.	374.	255.	201.	181.	289.
1978	1659.	1438.	1181.	1535.	1371.	917.	957.	716.	416.	283.	212.	178.
1979	2959.	1849.	940.	886.	1424.	963.	844.	578.	292.	195.	159.	249.

Table C.1 (c) Historic inflow sequences for EOL

Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
1950	569.	1764.	762.	535.	379.	597.	501.	374.	229.	153.	122.	113.
1951	510.	484.	493.	314.	459.	600.	524.	317.	198.	142.	113.	110.
1952	654.	954.	719.	515.	620.	804.	643.	328.	283.	207.	153.	113.
1953	547.	643.	487.	408.	524.	711.	527.	289.	198.	144.	116.	113.
1954	847.	699.	580.	513.	442.	674.	402.	289.	170.	127.	102.	144.
1955	433.	779.	583.	343.	479.	552.	419.	340.	235.	167.	125.	102.
1956	326.	1390.	824.	949.	770.	496.	561.	294.	187.	136.	116.	113.
1957	504.	960.	790.	807.	515.	711.	476.	314.	198.	139.	113.	110.
1958	603.	1223.	688.	541.	527.	498.	473.	286.	170.	127.	99.	133.
1959	583.	1155.	756.	617.	405.	595.	564.	402.	246.	164.	122.	119.
1960	637.	835.	527.	343.	685.	909.	799.	411.	238.	164.	133.	125.
1961	640.	470.	634.	269.	297.	660.	620.	453.	272.	190.	139.	122.
1962	595.	875.	484.	323.	425.	507.	496.	374.	224.	147.	119.	108.
1963	861.	881.	524.	464.	351.	473.	428.	340.	224.	156.	122.	139.
1964	949.	864.	462.	360.	595.	331.	337.	345.	212.	147.	122.	127.
1965	351.	1382.	1206.	830.	920.	835.	620.	391.	241.	167.	122.	96.
1966	651.	1184.	801.	796.	697.	765.	541.	354.	210.	144.	110.	153.
1967	637.	1407.	682.	564.	419.	833.	555.	394.	244.	161.	127.	116.
1968	725.	705.	481.	697.	677.	473.	504.	320.	193.	142.	110.	88.
1969	708.	1119.	915.	547.	578.	892.	660.	360.	210.	161.	108.	119.
1970	422.	977.	665.	549.	493.	912.	756.	561.	354.	229.	156.	125.
1971	462.	439.	654.	883.	733.	510.	382.	258.	156.	116.	96.	136.
1972	402.	892.	697.	767.	739.	833.	612.	306.	178.	136.	110.	91.
1973	1371.	1028.	818.	388.	433.	722.	473.	292.	210.	167.	144.	144.
1974	538.	1144.	623.	323.	453.	665.	702.	445.	266.	193.	147.	116.
1975	719.	1079.	869.	1062.	640.	651.	575.	371.	238.	170.	122.	99.
1976	776.	1048.	595.	464.	314.	510.	631.	399.	229.	153.	119.	110.
1977	816.	1034.	974.	668.	623.	320.	275.	204.	153.	125.	110.	130.
1978	569.	790.	626.	714.	793.	515.	555.	496.	244.	153.	113.	119.
1979	1107.	1322.	688.	467.	617.	705.	419.	269.	235.	144.	99.	116.

Table C.1 (d) Historic inflow sequences for LG2

Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
1	1088.	1221.	988.	993.	1093.	853.	728.	553.	380.	289.	225.	164.
2	695.	1827.	1834.	1276.	1232.	965.	686.	604.	404.	294.	220.	171.
3	1078.	1228.	961.	830.	978.	1501.	1102.	717.	442.	317.	252.	186.
4	1512.	2288.	1418.	1285.	1220.	922.	779.	581.	398.	296.	235.	280.
5	769.	1460.	1255.	985.	975.	1129.	800.	486.	356.	279.	223.	207.
6	1314.	1562.	839.	338.	703.	1258.	1080.	702.	443.	310.	253.	368.
7	787.	1440.	1024.	1038.	975.	1150.	1046.	691.	446.	309.	238.	236.
8	709.	2127.	1585.	1215.	774.	707.	615.	477.	284.	221.	187.	181.
9	703.	528.	634.	1020.	1315.	931.	719.	473.	290.	212.	195.	189.
10	787.	869.	950.	675.	488.	837.	656.	479.	310.	236.	180.	144.
11	762.	1324.	1126.	1319.	1486.	700.	630.	492.	339.	221.	185.	141.
12	952.	1925.	1157.	1203.	901.	1070.	1179.	745.	483.	331.	254.	234.
13	1162.	1142.	1052.	1269.	1289.	1027.	730.	563.	392.	263.	199.	297.
14	1223.	720.	946.	753.	899.	786.	950.	668.	380.	265.	211.	220.
15	517.	1297.	984.	724.	1174.	695.	843.	627.	408.	295.	227.	221.
16	871.	2102.	1165.	658.	655.	1248.	1007.	742.	461.	314.	240.	270.
17	881.	777.	643.	587.	601.	722.	802.	533.	353.	263.	212.	225.
18	760.	1164.	694.	792.	584.	640.	729.	608.	388.	282.	205.	169.
19	864.	2108.	1282.	998.	1224.	986.	641.	607.	423.	307.	231.	231.
20	1074.	1927.	1065.	1030.	1169.	1097.	986.	592.	392.	271.	223.	253.
21	1382.	1322.	731.	800.	648.	1223.	804.	521.	314.	228.	189.	247.
22	1124.	1909.	1282.	751.	875.	948.	881.	657.	408.	280.	226.	273.
23	752.	1432.	959.	866.	855.	937.	754.	478.	320.	246.	204.	167.
24	1054.	1708.	1163.	1057.	1134.	1252.	1071.	661.	417.	309.	247.	313.
25	959.	872.	1072.	1050.	1209.	975.	864.	567.	366.	245.	198.	203.
26	365.	1135.	896.	1269.	1085.	1099.	1119.	890.	526.	374.	277.	205.
27	804.	934.	990.	699.	689.	1042.	767.	452.	305.	230.	172.	207.
28	1165.	1895.	1445.	1531.	1043.	628.	604.	579.	345.	242.	194.	275.
29	775.	1333.	1055.	1283.	1245.	1128.	791.	605.	422.	300.	235.	278.
30	438.	1611.	1505.	1340.	1230.	1206.	1112.	754.	483.	331.	267.	257.
31	589.	1810.	1009.	906.	624.	1356.	1111.	764.	450.	321.	236.	124.
32	580.	1239.	1227.	1143.	1498.	890.	629.	516.	362.	280.	235.	270.
33	1272.	1906.	1478.	1392.	1174.	1178.	816.	439.	306.	233.	192.	214.
34	995.	1601.	1203.	1029.	859.	1235.	856.	661.	410.	298.	244.	244.
35	1193.	770.	800.	665.	697.	1351.	1019.	683.	430.	305.	242.	227.
36	675.	1535.	1312.	1029.	672.	796.	686.	491.	354.	254.	185.	179.
37	1084.	1886.	1471.	1284.	1151.	722.	467.	427.	308.	249.	213.	240.
38	611.	1426.	863.	1045.	1111.	922.	858.	585.	361	258.	223.	162.
39	676.	1554.	1287.	1038.	1073.	1018.	708.	597.	418.	290.	214.	282.
40	746.	2104.	1796.	1612.	1332.	913.	959.	668.	425.	313.	240.	198.
41	938.	1712.	1363.	1577.	1330.	1347.	911.	618.	409.	280.	218.	249.
42	316.	1543.	1274.	1395.	1255.	1208.	942.	647.	404.	288.	222.	122.
43	1095.	2278.	1938.	1546.	1411.	1534.	1212.	790.	461.	317.	256.	369.
44	705.	2139.	1116.	442.	444.	942.	732.	606.	403.	271.	214.	235.
45	733.	1964.	1516.	1430.	1387.	530.	529.	328.	222.	173.	162.	216.
46	845.	1818.	1521.	1366.	922.	1277.	1010.	715.	426.	284.	226.	242.
47	1096.	1455.	1106.	1022.	944.	1106.	941.	566.	380.	260.	213.	248.
48	466.	1272.	1294.	1055.	1066.	998.	977.	599.	353.	243.	201.	163.
49	978.	1024.	1267.	1154.	1117.	1239.	1034.	668.	448.	316.	239.	176.
50	324.	1760.	1188.	979.	898.	1185.	938.	615.	394.	286.	226.	245.

Table C.2 (a) Synthetic inflow sequences for LG4

Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
1	1095.	919.	597.	686.	746.	616.	480.	368.	236.	179.	142.	119.
2	700.	1375.	1108.	882.	841.	697.	452.	402.	251.	182.	139.	124.
3	1085.	925.	581.	574.	668.	1084.	726.	478.	274.	196.	159.	135.
4	1521.	1722.	857.	889.	833.	665.	513.	387.	247.	183.	148.	202.
5	774.	1099.	758.	681.	665.	815.	527.	324.	221.	173.	141.	150.
6	1322.	1176.	507.	234.	480.	909.	712.	468.	275.	192.	160.	266.
7	792.	1084.	619.	718.	666.	831.	689.	460.	277.	192.	151.	171.
8	713.	1601.	957.	840.	528.	510.	405.	318.	176.	137.	118.	131.
9	707.	397.	383.	705.	898.	672.	474.	315.	180.	132.	123.	137.
10	792.	654.	574.	467.	333.	605.	432.	319.	192.	147.	114.	104.
11	767.	996.	680.	912.	1014.	505.	415.	328.	210.	137.	117.	102.
12	958.	1449.	699.	831.	615.	772.	777.	496.	299.	205.	160.	169.
13	1169.	860.	636.	877.	880.	741.	481.	375.	243.	163.	126.	215.
14	1231.	542.	571.	521.	613.	567.	626.	445.	236.	164.	134.	159.
15	520.	976.	594.	500.	801.	502.	555.	418.	253.	183.	144.	160.
16	876.	1582.	704.	455.	447.	901.	663.	494.	286.	195.	152.	195.
17	887.	585.	389.	406.	410.	522.	529.	356.	219.	163.	134.	163.
18	764.	876.	419.	547.	399.	462.	480.	405.	240.	175.	130.	122.
19	869.	1587.	774.	690.	835.	712.	423.	405.	263.	190.	146.	167.
20	1081.	1450.	643.	712.	798.	792.	650.	394.	243.	168.	141.	183.
21	1390.	995.	441.	553.	442.	883.	530.	348.	195.	141.	120.	179.
22	1131.	1437.	775.	519.	597.	685.	581.	438.	253.	174.	143.	197.
23	756.	1078.	579.	599.	584.	677.	497.	319.	199.	153.	129.	121.
24	1060.	1286.	702.	731.	774.	904.	706.	441.	259.	192.	156.	227.
25	965.	656.	648.	726.	825.	704.	569.	378.	227.	152.	125.	147.
26	367.	854.	541.	877.	740.	793.	738.	594.	326.	232.	175.	148.
27	809.	703.	598.	483.	470.	752.	506.	301.	189.	142.	109.	150.
28	1172.	1426.	873.	1058.	712.	453.	398.	386.	214.	150.	122.	199.
29	779.	1003.	637.	887.	850.	814.	521.	403.	262.	186.	148.	202.
30	441.	1212.	909.	926.	839.	871.	733.	503.	300.	205.	169.	186.
31	592.	1362.	609.	626.	426.	979.	732.	509.	279.	199.	149.	90.
32	584.	933.	741.	790.	1022.	642.	415.	344.	225.	174.	149.	195.
33	1280.	1435.	893.	962.	801.	850.	538.	293.	190.	144.	121.	155.
34	1001.	1205.	727.	712.	587.	891.	564.	441.	255.	185.	154.	176.
35	1201.	579.	483.	460.	476.	975.	671.	455.	267.	189.	153.	165.
36	680.	1156.	792.	711.	459.	575.	452.	328.	219.	157.	117.	130.
37	1090.	1419.	889.	888.	786.	521.	308.	285.	191.	154.	135.	174.
38	615.	1074.	521.	723.	758.	666.	565.	390.	224.	160.	141.	117.
39	680.	1170.	777.	717.	732.	735.	466.	398.	259.	180.	135.	204.
40	750.	1584.	1085.	1114.	909.	659.	632.	445.	264.	194.	152.	143.
41	943.	1288.	823.	1090.	908.	973.	600.	412.	254.	173.	138.	180.
42	318.	1161.	770.	964.	856.	872.	621.	431.	251.	178.	140.	88.
43	1102.	1714.	1171.	1069.	963.	1108.	798.	527.	286.	196.	162.	267.
44	709.	1610.	674.	305.	303.	680.	482.	404.	250.	168.	136.	170.
45	737.	1478.	916.	989.	946.	383.	349.	219.	138.	107.	102.	157.
46	850.	1368.	919.	945.	629.	922.	666.	477.	264.	176.	143.	175.
47	1102.	1095.	668.	707.	644.	798.	620.	378.	236.	161.	135.	179.
48	469.	958.	782.	730.	727.	721.	644.	400.	219.	151.	127.	118.
49	984.	771.	765.	798.	762.	895.	681.	445.	278.	196.	151.	127.
50	326.	1325.	717.	677.	613.	856.	618.	410.	245.	178.	143.	177.

Table C.2 (b) Synthetic inflow sequences for LG3

Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
1	2291.	1348.	805.	1102.	1196.	1005.	758.	513.	325.	242.	190.	209.
2	1464.	2018.	1494.	1415.	1348.	1137.	715.	560.	346.	245.	186.	218.
3	2270.	1356.	783.	921.	1070.	1768.	1148.	665.	379.	265.	213.	237.
4	3184.	2526.	1155.	1426.	1335.	1085.	811.	539.	340.	247.	198.	356.
5	1620.	1612.	1022.	1093.	1066.	1329.	833.	451.	305.	233.	189.	264.
6	2766.	1725.	684.	375.	770.	1482.	1125.	652.	380.	259.	214.	468.
7	1657.	1590.	834.	1152.	1067.	1355.	1089.	641.	382.	258.	201.	300.
8	1492.	2349.	1291.	1348.	847.	832.	640.	443.	243.	185.	158.	230.
9	1481.	583.	516.	1131.	1439.	1096.	749.	439.	249.	177.	165.	241.
10	1657.	960.	774.	749.	534.	986.	683.	445.	265.	198.	152.	184.
11	1605.	1462.	917.	1464.	1626.	824.	656.	457.	290.	185.	156.	179.
12	2005.	2125.	942.	1335.	986.	1260.	1228.	691.	413.	276.	214.	298.
13	2446.	1261.	857.	1408.	1410.	1209.	760.	523.	336.	220.	168.	378.
14	2576.	795.	770.	836.	983.	925.	989.	620.	326.	221.	179.	280.
15	1088.	1432.	801.	803.	1284.	818.	878.	582.	349.	246.	192.	282.
16	1833.	2321.	949.	730.	716.	1469.	1048.	688.	395.	263.	203.	343.
17	1855.	858.	524.	651.	658.	851.	835.	495.	302.	220.	180.	286.
18	1600.	1285.	565.	879.	639.	754.	759.	564.	332.	235.	173.	215.
19	1819.	2327.	1044.	1108.	1339.	1161.	668.	564.	362.	256.	195.	294.
20	2262.	2127.	867.	1143.	1279.	1292.	1027.	549.	336.	226.	188.	323.
21	2909.	1460.	595.	888.	709.	1440.	837.	484.	269.	190.	160.	314.
22	2367.	2107.	1045.	834.	957.	1116.	918.	610.	349.	234.	191.	347.
23	1582.	1581.	781.	961.	936.	1104.	785.	444.	274.	206.	172.	212.
24	2219.	1886.	947.	1173.	1241.	1475.	1115.	614.	357.	258.	209.	399.
25	2019.	963.	873.	1166.	1323.	1149.	900.	526.	314.	205.	167.	258.
26	768.	1253.	729.	1408.	1187.	1294.	1166.	827.	450.	313.	234.	261.
27	1694.	1031.	806.	776.	754.	1227.	799.	419.	261.	192.	145.	263.
28	2452.	2093.	1177.	1698.	1141.	740.	629.	538.	295.	202.	164.	350.
29	1631.	1471.	859.	1423.	1362.	1328.	824.	561.	361.	250.	198.	354.
30	923.	1779.	1225.	1487.	1345.	1421.	1158.	700.	414.	276.	225.	327.
31	1239.	1999.	821.	1005.	683.	1598.	1157.	709.	385.	268.	199.	158.
32	1222.	1368.	999.	1268.	1639.	1048.	655.	479.	310.	234.	199.	343.
33	2678.	2105.	1204.	1544.	1284.	1387.	849.	408.	262.	195.	162.	273.
34	2096.	1767.	980.	1142.	940.	1454.	891.	614.	351.	249.	206.	310.
35	2513.	850.	652.	738.	763.	1591.	1060.	634.	368.	255.	205.	289.
36	1422.	1695.	1069.	1142.	735.	937.	714.	456.	303.	212.	156.	228.
37	2282.	2082.	1198.	1425.	1260.	850.	487.	397.	263.	208.	180.	306.
38	1287.	1575.	703.	1160.	1216.	1086.	893.	544.	309.	215.	189.	206.
39	1424.	1716.	1048.	1151.	1174.	1199.	737.	554.	358.	243.	181.	359.
40	1570.	2323.	1463.	1789.	1458.	1075.	998.	620.	364.	262.	203.	252.
41	1974.	1890.	1110.	1750.	1455.	1587.	948.	573.	351.	234.	184.	317.
42	665.	1703.	1038.	1548.	1373.	1422.	981.	600.	346.	240.	187.	155.
43	2306.	2515.	1579.	1716.	1543.	1807.	1261.	734.	395.	265.	216.	470.
44	1485.	2361.	909.	490.	486.	1110.	762.	562.	345.	227.	181.	299.
45	1543.	2169.	1235.	1587.	1517.	624.	551.	305.	190.	145.	136.	275.
46	1779.	2007.	1239.	1516.	1008.	1504.	1052.	664.	365.	237.	191.	308.
47	2307.	1606.	900.	1134.	1033.	1302.	980.	526.	326.	217.	180.	315.
48	982.	1405.	1054.	1171.	1166.	1175.	1017.	556.	302.	203.	169.	207.
49	2059.	1130.	1032.	1280.	1222.	1460.	1076.	620.	384.	264.	202.	224.
50	683.	1943.	967.	1087.	982.	1396.	977.	571.	338.	239.	191.	312.

Table C.2 (c) Synthetic inflow sequences for EOL

Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
1	830.	800.	599.	577.	630.	560.	469.	323.	216.	162.	123.	89.
2	530.	1197.	1112.	741.	710.	633.	442.	353.	230.	164.	121.	93.
3	822.	805.	583.	482.	564.	985.	710.	420.	252.	177.	138.	101.
4	1153.	1499.	860.	747.	703.	605.	501.	340.	226.	166.	129.	152.
5	587.	957.	761.	573.	562.	740.	515.	284.	203.	156.	123.	112.
6	1002.	1023.	509.	196.	405.	826.	696.	411.	252.	173.	139.	200.
7	600.	944.	621.	603.	562.	755.	674.	404.	254.	173.	131.	128.
8	541.	1394.	961.	706.	446.	464.	396.	279.	162.	124.	103.	98.
9	536.	346.	384.	593.	758.	611.	463.	277.	165.	119.	107.	103.
10	600.	569.	576.	392.	281.	549.	423.	280.	176.	132.	99.	78.
11	581.	867.	683.	767.	856.	459.	406.	288.	193.	124.	102.	76.
12	726.	1261.	702.	699.	519.	702.	759.	436.	275.	185.	139.	127.
13	886.	748.	638.	737.	743.	673.	470.	330.	223.	147.	109.	161.
14	937.	472.	574.	438.	518.	515.	612.	391.	217.	148.	116.	119.
15	394.	850.	597.	421.	676.	456.	543.	367.	232.	165.	125.	120.
16	664.	1377.	707.	382.	377.	819.	648.	434.	262.	176.	132.	146.
17	672.	509.	390.	341.	346.	474.	517.	312.	201.	147.	117.	122.
18	579.	762.	421.	460.	337.	420.	469.	356.	221.	158.	113.	92.
19	659.	1381.	778.	580.	705.	647.	413.	355.	241.	172.	127.	126.
20	819.	1262.	646.	599.	674.	720.	635.	346.	223.	152.	122.	138.
21	1054.	866.	443.	465.	373.	802.	518.	305.	179.	127.	104.	134.
22	857.	1250.	778.	437.	504.	622.	568.	385.	232.	157.	124.	148.
23	573.	938.	582.	503.	493.	615.	485.	280.	182.	138.	112.	91.
24	804.	1119.	705.	614.	654.	822.	690.	387.	237.	173.	136.	170.
25	731.	571.	650.	611.	697.	640.	556.	332.	208.	137.	109.	110.
26	278.	743.	543.	738.	625.	721.	721.	521.	299.	209.	152.	111.
27	613.	612.	600.	406.	397.	683.	494.	264.	174.	128.	94.	112.
28	888.	1242.	876.	890.	601.	412.	389.	339.	196.	136.	106.	149.
29	591.	873.	640.	746.	717.	740.	509.	354.	240.	168.	129.	151.
30	334.	1055.	913.	779.	708.	792.	716.	441.	275.	185.	146.	140.
31	449.	1186.	612.	526.	360.	890.	715.	447.	256.	179.	130.	67.
32	443.	812.	744.	664.	863.	584.	405.	302.	206.	157.	129.	146.
33	970.	1249.	897.	809.	676.	773.	525.	257.	174.	130.	105.	116.
34	759.	1049.	730.	598.	495.	810.	551.	387.	234.	167.	134.	132.
35	910.	504.	485.	386.	402.	886.	656.	400.	245.	171.	133.	123.
36	515.	1006.	796.	598.	387.	522.	442.	288.	201.	142.	101.	97.
37	826.	1236.	892.	746.	663.	473.	301.	250.	175.	139.	117.	131.
38	466.	935.	523.	608.	640.	605.	552.	343.	206.	144.	123.	88.
39	516.	1018.	780.	603.	618.	668.	456.	350.	238.	162.	117.	153.
40	569.	1378.	1089.	937.	768.	599.	618.	391.	242.	175.	132.	108.
41	715.	1121.	827.	917.	766.	884.	587.	361.	233.	156.	119.	135.
42	241.	1011.	773.	811.	723.	792.	607.	379.	230.	161.	122.	66.
43	835.	1492.	1176.	899.	813.	1007.	780.	463.	262.	177.	140.	201.
44	538.	1401.	677.	257.	256.	618.	471.	355.	229.	152.	118.	128.
45	559.	1287.	919.	831.	799.	348.	341.	192.	126.	97.	89.	117.
46	645.	1191.	923.	794.	531.	838.	650.	419.	243.	159.	124.	131.
47	835.	953.	671.	594.	544.	725.	606.	332.	216.	145.	117.	134.
48	356.	834.	785.	613.	614.	655.	629.	351.	201.	136.	110.	88.
49	746.	671.	768.	671.	644.	813.	666.	391.	255.	177.	131.	95.
50	247.	1153.	720.	569.	517.	777.	604.	360.	225.	160.	124.	133.

Table C.2 (d) Synthetic inflow sequences for LG2

Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
1	10914.10	14078.93	16725.18	19384.84	22217.89	24502.56	26389.54	26515.30	21013.91	17075.50	12543.71	8000.00
2	9861.48	14597.06	19509.25	22926.89	26120.23	28704.88	30483.00	28169.13	22626.34	18187.45	13074.77	8000.00
3	10887.31	14070.29	16644.23	18867.30	21402.27	25422.55	28278.93	26977.30	22236.95	17919.51	12965.38	8000.00
4	11990.68	17864.01	21602.92	24985.60	28090.68	30501.11	32463.12	27103.91	21254.55	17166.55	12448.86	8000.00
5	10059.69	13844.00	17205.39	19843.62	22370.82	25394.73	27468.33	26550.95	20994.45	17012.20	12453.31	8000.00
6	11519.41	15568.11	17815.29	18720.59	20542.77	23912.19	26711.55	24386.05	20046.53	16257.34	11909.50	8000.00
7	10107.89	13840.38	16583.05	19363.23	21890.43	24970.59	27681.82	26436.85	21802.14	17579.25	12713.74	8000.00
8	9898.98	15412.16	19657.43	22911.68	24917.89	26811.51	28405.59	27918.80	22411.87	17952.16	12902.41	8000.00
9	9602.61	10971.19	12669.29	14632.04	17296.10	18669.86	19449.80	14943.80	9947.65	8000.00	8000.00	8000.00
10	10107.89	12360.34	14904.82	15943.52	16464.00	15678.28	14448.76	12704.17	10506.94	8343.32	8000.00	8000.00
11	10040.94	13472.74	16488.62	20021.43	23873.14	25748.02	27380.98	27241.48	21528.40	17367.85	12665.10	8000.00
12	10549.83	15539.43	18638.34	21860.45	24195.84	27061.73	30117.70	26689.47	21623.99	17501.38	12699.15	8000.00
13	11112.30	14072.36	16890.04	20288.92	23630.01	26380.72	28272.88	25884.23	21276.33	17127.08	12359.64	8000.00
14	11275.68	13141.92	15675.68	17692.52	20022.72	22127.95	24590.34	25022.51	19519.51	15977.59	11911.55	8000.00
15	8919.99	12281.81	14917.36	16856.52	19899.52	21761.01	23946.07	22110.73	16288.16	12435.20	10189.15	8000.00
16	10332.88	15781.26	18901.59	20663.98	22361.74	25704.38	28314.52	25961.52	21312.51	17241.14	12505.30	8000.00
17	10359.67	12373.65	14095.86	14898.86	15712.23	14664.58	13858.09	12304.22	10268.24	8211.56	8000.00	8000.00
18	9935.04	12952.13	14810.93	16163.00	16932.32	15472.07	14289.63	12743.67	10608.47	8423.47	8000.00	8000.00
19	10314.13	15778.07	19211.78	21884.82	25057.43	27698.32	29359.80	26756.34	21895.13	17649.64	12745.73	8000.00
20	10876.60	15871.38	18723.87	21482.62	24512.67	27450.88	30006.59	26686.50	21710.48	17468.24	12616.99	8000.00
21	11701.55	15128.17	17086.08	19228.80	20908.41	24184.09	26268.06	26344.05	20765.62	16766.84	12235.16	8000.00
22	11010.52	15958.64	19392.35	21403.83	23671.83	26210.95	28494.50	26148.52	21454.06	17272.54	12498.35	8000.00
23	10014.15	13725.89	16294.48	18613.97	20830.13	23339.79	25294.16	24822.85	18764.58	14443.11	11217.48	8000.00
24	10794.45	15193.91	18280.30	21082.79	23994.45	27319.22	30067.59	25729.48	20737.84	16797.71	12239.80	8000.00
25	10568.58	12828.80	15700.05	18512.37	21646.09	24257.53	26497.02	26634.04	21068.92	17031.20	12434.99	8000.00
26	8977.61	11919.53	14319.38	17718.27	20530.58	23474.14	26374.59	25966.69	21116.93	17209.80	12624.25	8000.00
27	9553.11	11393.09	13444.39	14716.28	15921.22	18111.80	19518.91	14969.75	10026.87	8093.74	8000.00	8000.00
28	11120.33	16032.16	19902.45	24003.08	26706.54	28388.57	29954.14	26555.43	21615.88	17360.10	12496.63	8000.00
29	10075.75	13530.89	16356.60	19792.99	23020.03	26041.26	28091.53	25893.95	21290.28	17181.78	12458.99	8000.00
30	8671.64	12362.04	15891.54	18979.11	21681.96	24410.62	26807.61	26911.66	21370.12	17278.06	12576.15	8000.00
31	9577.57	14269.09	16971.59	19398.22	21015.63	24647.54	27527.25	27510.92	22709.76	18314.24	13220.11	8000.00
32	9553.47	12764.95	16051.35	19112.76	22995.57	25379.35	27009.71	25611.84	20137.32	16381.49	12075.94	8000.00
33	11406.92	16347.27	20305.94	24034.27	27077.28	30232.43	32347.50	28156.43	22460.46	17961.14	12869.95	8000.00
34	10665.00	14814.79	18036.91	20792.98	23019.50	26327.33	28546.08	26438.94	21717.09	17491.13	12667.77	8000.00
35	11195.32	13191.16	15333.88	17115.02	18921.64	22540.16	25181.41	25724.91	20427.01	16634.08	12266.03	8000.00
36	9807.92	13786.63	17300.70	20056.77	21798.59	23930.59	25708.71	25436.72	19634.01	15883.47	11885.02	8000.00
37	10903.38	15791.89	19731.81	23170.88	26154.27	28088.07	29298.54	26952.57	22005.86	17637.19	12704.04	8000.00
38	9636.50	13332.69	15644.15	18443.07	21322.79	23792.27	26016.20	26119.62	20459.79	16647.29	12333.12	8000.00
39	9810.60	13838.56	17285.66	20065.84	22847.05	25573.66	27408.80	25938.32	21154.22	17086.38	12379.14	8000.00
40	9893.75	15246.34	19952.40	24165.64	27517.22	29858.26	32243.02	27116.83	21339.79	17316.82	12637.37	8000.00
41	10512.33	14949.83	18600.49	22824.32	26271.68	29879.48	32240.80	27526.25	21787.74	17527.55	12644.84	8000.00
42	8704.36	12566.38	15836.65	19431.00	22546.53	25640.03	27944.26	28366.04	22973.19	18453.50	13272.79	8000.00
43	8617.27	12280.96	15156.12	16801.67	18044.21	19657.62	20384.36	18775.36	16285.16	13687.59	10648.33	8000.00
44	9888.27	15432.55	18421.64	19605.50	20756.35	23279.40	25176.74	25375.57	19786.90	16172.10	11991.40	8000.00
45	9963.26	15053.95	19114.40	22944.51	26539.61	27959.16	29330.33	28100.21	22710.43	18056.63	12874.75	8000.00
46	10263.25	14975.50	19049.34	22708.04	25097.86	28518.18	31136.09	27403.89	21838.99	17568.27	12684.63	8000.00
47	10935.52	14706.88	17669.19	20406.51	22853.36	25815.67	28254.74	26638.38	21880.94	17563.29	12657.15	8000.00
48	9248.13	12545.15	16011.00	18836.71	21599.79	24272.82	26805.21	26861.31	21094.75	17090.07	12520.58	8000.00
49	10619.47	13273.68	16667.21	19758.08	22653.34	25971.88	28652.01	27121.49	22382.71	18036.86	13019.15	8000.00
50	8867.80	13429.72	16611.66	19233.81	21561.42	24735.32	27166.62	25979.46	20497.35	16649.59	12228.87	8000.00

Table C.3 (a) Optimal storage levels for LG4

Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
1	6939.29	8881.14	10025.28	11407.80	12901.24	14096.27	14900.24	8045.00	6000.00	6080.10	6069.69	6000.00
2	6440.75	8616.89	10150.43	11078.65	11870.66	12303.38	12087.10	7899.15	6000.00	6083.53	6060.84	6000.00
3	7793.30	9114.02	9557.41	9982.04	10636.6f	12427.25	13232.17	8538.20	6000.00	6080.92	6071.41	6000.00
4	7836.46	10134.66	10192.67	10336.38	10330.30	9874.05	9038.53	7794.14	6174.79	6144.63	6000.00	6000.00
5	7213.22	9244.64	10430.43	11410.00	12316.49	13654.95	14203.75	8094.35	6000.00	6045.72	6010.63	6000.00
6	9531.98	10603.85	10952.94	10570.82	10838.67	12264.47	13133.66	9396.61	6462.93	6283.91	6000.00	6000.00
7	7142.96	9005.92	9685.52	10630.29	11409.79	12657.22	13496.34	8627.86	6002.76	6053.57	6000.00	6000.00
8	7096.89	10460.10	12210.52	13647.57	14229.57	14782.75	15045.93	8465.76	6008.46	6034.82	6000.00	6000.00
9	6000.00	7029.02	8054.85	10712.34	13784.37	16704.09	19016.40	16436.70	13495.40	10994.65	6000.00	6000.00
10	8121.29	9816.45	11353.86	13373.89	14981.44	19629.41	23679.03	18975.66	13932.14	11254.75	6000.00	6000.00
11	6107.29	8174.79	9464.82	11376.25	13490.41	14311.72	14873.27	8012.77	6000.00	6045.39	6042.08	6000.00
12	6883.36	9010.90	9200.55	9743.76	9709.57	10094.75	10480.46	8036.30	6000.00	6079.41	6027.20	6000.00
13	8062.57	9257.68	9892.67	11173.16	12420.11	13336.33	13549.07	9253.78	6364.17	6238.38	6000.00	6000.00
14	6000.00	7404.86	8934.23	10329.67	11918.57	13437.22	15059.81	8412.42	6369.02	6282.28	6105.89	6000.00
15	6000.00	8529.79	10120.76	11459.96	13536.15	14880.70	16319.27	11757.24	10153.95	9915.68	7345.57	6000.00
16	7091.23	9977.22	10607.77	10571.41	10515.48	11673.68	12177.62	8644.84	6098.32	6115.17	6000.00	6000.00
17	8375.74	9892.05	10933.95	12790.61	14597.74	18977.32	23233.77	18507.63	13414.54	10977.69	6000.00	6000.00
18	6000.00	8270.59	9392.84	11627.15	13405.77	17817.61	22133.78	17196.66	11817.61	10158.86	6000.00	6000.00
19	7002.15	9833.03	10580.73	11103.45	11985.15	12566.79	12380.59	8498.30	6000.60	6056.25	6000.00	6000.00
20	7407.86	9726.76	9961.48	10381.02	11009.94	11643.74	11889.04	8653.74	6134.26	6112.49	6000.00	6000.00
21	7893.13	10050.91	10796.78	11842.63	12567.03	14496.75	15449.25	8504.51	6249.95	6158.02	6000.00	6000.00
22	7832.90	10399.84	11279.23	11472.96	11862.62	12500.96	12849.14	8931.68	6200.27	6144.63	6000.00	6000.00
23	6000.00	8794.17	10344.96	11949.33	13463.05	15276.33	16564.55	9974.25	8226.33	8606.90	7292.62	6000.00
24	7142.92	8834.77	9018.82	9280.55	9645.30	10370.39	10558.88	8652.30	6258.25	6190.16	6000.00	6000.00
25	6342.57	7660.96	9001.87	10551.70	12308.13	13799.04	14891.92	8089.70	6046.83	6058.92	6000.00	6000.00
26	6103.48	7465.93	8035.46	9504.93	10571.90	11816.39	12878.16	8064.52	6000.00	6116.53	6092.88	6000.00
27	6000.00	8403.12	10605.12	12499.10	14298.29	16912.76	18805.26	16174.96	13244.69	10893.43	6000.00	6000.00
28	7717.51	10037.98	10954.65	12366.82	12836.61	12628.35	12284.25	9071.32	6311.82	6206.23	6000.00	6000.00
29	7034.91	8617.04	9271.63	10595.80	11781.36	12910.02	13242.81	8943.93	6183.34	6144.63	6000.00	6000.00
30	6000.00	7998.43	9251.93	10550.95	11582.57	12734.28	13491.14	7557.56	6000.00	6045.83	6000.14	6000.00
31	6740.42	9452.80	10238.77	11070.26	11356.54	13133.50	14212.93	8442.57	6000.00	6131.02	6142.15	6000.00
32	6000.00	7736.42	9016.46	10427.75	12394.86	13409.75	13803.51	8308.52	6158.97	6123.20	6000.00	6000.00
33	7613.33	9576.38	10153.17	10914.77	11234.49	11696.11	11334.13	8289.51	6117.68	6091.06	6000.00	6000.00
34	7498.95	9478.31	10243.38	10968.27	11345.78	12550.11	12868.00	8730.45	6037.15	6058.92	6000.00	6000.00
35	6562.01	7825.13	8873.22	9859.71	10855.86	13221.73	14723.31	8031.53	6000.00	6049.91	6008.73	6000.00
36	6000.00	8996.35	11117.64	13021.98	14211.71	15751.79	16923.37	10192.68	8333.83	7889.36	6912.07	6000.00
37	7821.57	10437.15	11720.37	13000.91	13975.75	14273.32	14009.19	9065.89	6152.84	6104.46	6000.00	6000.00
38	6000.00	8624.36	9855.04	11626.76	13432.05	15051.11	16356.14	9667.88	7698.78	7215.41	6629.35	6000.00
39	6948.27	9136.04	10344.13	11391.51	12443.98	13539.57	13902.57	8841.77	6242.87	6184.81	6000.00	6000.00
40	6000.00	8161.72	9058.99	10033.93	10446.06	10202.33	9896.47	8807.44	7233.61	6933.97	6471.18	6000.00
41	6502.27	7882.57	8063.44	8959.43	9354.78	9937.40	9534.41	7811.42	6129.43	6112.49	6000.00	6000.00
42	6000.00	8185.05	9395.69	11125.94	12520.43	14004.27	14789.64	8058.89	6009.80	6112.13	6123.64	6000.00
43	8877.47	13248.42	16310.70	19279.46	21877.71	24950.93	24270.32	20210.48	15505.15	11037.45	6000.00	6000.00
44	6000.00	10068.68	11766.01	12475.01	13155.96	14869.36	16014.27	9364.33	7465.71	7039.87	6466.63	6000.00
45	7151.44	10186.41	11817.30	13643.70	15299.73	15503.02	15611.63	9110.55	6268.27	6147.31	6000.00	6000.00
46	6465.43	8258.51	8908.76	9628.64	9506.23	10164.52	10138.01	7866.58	6083.29	6085.71	6000.00	6000.00
47	7954.32	9827.45	10619.35	11515.70	12219.84	13359.93	14001.86	8950.34	6161.39	6117.85	6000.00	6000.00
48	6000.00	8047.90	9692.67	11198.17	12647.32	14128.71	15362.73	8786.06	6888.38	6664.96	6338.89	6000.00
49	7511.43	8422.00	9346.86	10360.11	11247.36	12520.41	13197.70	8512.99	6000.00	6094.18	6077.93	6000.00
50	6000.00	8589.41	9636.66	10576.78	11320.69	12668.07	13355.09	8091.33	6088.64	6091.06	6000.00	6000.00

Table C.3 (b) Optimal storage levels for LG3

Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
1	15236.99	17860.79	19117.69	21170.07	23399.89	25192.46	21022.02	16056.35	10587.14	10473.15	10207.67	10000.00
2	11973.18	15318.66	17372.20	19214.13	20822.98	21920.32	18497.57	14545.57	10020.39	10085.71	10000.00	10000.00
3	14697.62	16874.61	17589.45	18673.91	20109.60	23462.66	20166.15	15466.08	10000.00	10066.65	[illegible]	10000.00
4	15191.94	18510.86	18268.33	18751.65	18983.50	18553.49	17427.14	14312.84	10665.54	10423.18	10000.00	10000.00
5	13410.48	16690.22	18499.02	20498.00	22362.50	24993.57	20989.11	15828.02	10275.87	10200.88	10000.00	10000.00
6	16376.00	19848.05	20647.63	20619.59	21616.29	24553.23	21258.11	16586.28	11185.92	10680.31	10000.00	10000.00
7	13390.92	16498.80	17685.40	19723.73	21475.99	24058.04	20602.26	15831.37	10366.77	10265.16	10000.00	10000.00
8	13025.58	18174.91	20662.13	23302.03	24558.17	25816.01	21270.58	16045.99	10285.72	10192.84	10000.00	10000.00
9	10000.00	11511.13	12893.19	15922.46	19652.34	22587.87	20092.54	16683.72	12766.01	11468.58	10000.00	10000.00
10	14438.11	16926.42	18999.50	21005.63	22389.75	25030.65	22909.18	20079.53	16767.77	13614.42	10000.00	10000.00
11	13456.09	16430.04	18043.39	21121.82	24520.86	25885.12	21504.95	16445.76	10939.28	10673.11	10300.74	10000.00
12	13318.12	16840.25	17311.23	18834.83	19404.68	20727.39	18493.10	14746.03	10254.36	10224.98	10000.00	10000.00
13	15432.77	17618.78	18795.57	21448.17	24020.38	26139.98	21762.30	16603.93	10944.70	10562.46	10000.00	10000.00
14	14735.29	16460.45	18176.15	20068.63	22281.09	24411.95	21375.10	17248.70	12334.85	11587.57	10647.73	10000.00
15	10000.00	13570.46	15569.88	17574.64	20721.44	22725.00	19554.60	15253.41	10328.14	10241.05	10000.00	10000.00
16	13606.40	18361.37	19600.07	20252.20	20847.01	23478.48	20146.04	15738.30	10545.79	10374.97	10000.00	10000.00
17	14968.43	17192.36	18595.84	20339.48	22045.02	24324.34	22587.13	19881.36	16658.66	13549.46	10000.00	10000.00
18	10000.00	13330.72	14844.01	17198.32	18854.61	20874.13	19038.88	16620.18	13580.09	11959.87	10000.00	10000.00
19	13429.41	18064.94	19418.59	20943.67	23018.30	24685.33	20371.16	15634.64	10357.09	10265.16	10000.00	10000.00
20	14370.05	18249.23	18882.92	20255.86	21937.01	23709.02	20447.75	15797.48	10576.72	10361.58	10000.00	10000.00
21	16828.84	19681.59	20312.61	21728.41	22634.56	25528.84	21501.80	16395.05	10712.45	10412.47	10000.00	10000.00
22	15077.32	19316.93	20853.41	21824.74	23083.56	24810.20	20972.14	16181.20	10691.20	10417.83	10000.00	10000.00
23	10772.87	14708.61	16632.83	19039.16	21303.06	24092.40	20700.07	15849.61	10543.83	10396.50	10142.30	10000.00
24	14020.09	17047.37	17660.54	18879.02	20234.46	22261.82	19632.91	15574.48	10827.71	10508.89	10000.00	10000.00
25	14549.38	16214.85	17694.78	19959.49	22558.08	24777.25	21014.47	16124.55	10666.80	10445.82	10099.39	10000.00
26	11088.48	13398.98	14383.01	17185.68	19325.10	21822.43	18888.71	14949.23	10000.00	10108.97	10018.03	10000.00
27	11993.37	14665.72	16824.51	18902.95	20857.32	24143.71	21257.37	17257.03	12833.50	11548.53	10000.00	10000.00
28	15070.67	19047.24	20702.96	23754.12	25263.11	25748.36	21257.73	16373.68	10838.77	10498.18	10000.00	10000.00
29	13338.47	16154.54	17425.29	20206.66	22740.20	25267.12	21141.08	16173.09	10669.43	10417.83	10000.00	10000.00
30	10260.48	12731.31	13800.68	15571.78	16917.68	18512.01	16915.49	14039.04	10396.57	10273.20	10000.00	10000.00
31	12282.49	16461.28	17624.20	19279.96	20047.67	23291.71	20022.99	15445.40	10000.00	10147.84	10126.71	10000.00
32	12379.34	15179.70	17085.07	19710.93	23213.72	25250.34	20937.60	16009.71	10629.16	10385.69	10000.00	10000.00
33	14796.30	17952.68	18801.02	20560.03	21588.37	22926.86	19767.91	15322.48	10486.00	10297.30	10000.00	10000.00
34	14356.46	17719.62	19086.99	20888.27	22107.85	24744.78	20789.21	15959.03	10426.12	10278.55	10000.00	10000.00
35	15877.10	17254.12	18146.71	19269.65	20421.17	23828.78	20485.19	15889.15	10580.66	10412.38	10092.16	10000.00
36	10030.66	14169.13	16768.88	19564.15	21214.30	23460.50	19791.38	15308.93	10416.68	10332.28	10068.57	10000.00
37	14945.58	19213.22	21255.41	23905.61	26042.63	27152.74	22021.06	16477.26	10574.56	10337.48	10000.00	10000.00
38	10527.80	14109.31	15474.64	18064.00	20714.99	23106.15	19655.03	15154.11	10023.76	10045.53	10000.00	10000.00
39	12931.70	16525.70	18450.32	20650.83	22839.96	25169.03	20960.42	16121.38	10757.38	10476.75	10000.00	10000.00
40	11128.58	14172.54	15014.53	16729.68	17531.56	17334.34	15864.04	13332.28	10084.41	10131.24	10000.00	10000.00
41	12612.86	14923.70	15222.42	17235.31	18418.64	19994.96	17766.30	14458.96	10557.02	10356.23	10000.00	10000.00
42	10368.57	13415.75	14783.36	17516.96	19708.78	22104.90	19045.01	14862.66	10000.00	10113.97	10098.20	10000.00
43	13120.35	16681.77	17854.92	19395.01	20437.01	22220.83	19835.97	15960.10	11176.25	10680.31	10000.00	10000.00
44	10599.51	16414.66	18534.61	19532.32	20487.47	23145.79	19551.46	15301.64	10470.62	10316.05	10000.00	10000.00
45	13140.13	17801.54	20116.72	23374.70	26346.13	27024.81	22227.34	16611.07	10686.79	10372.30	10000.00	10000.00
46	12416.59	15346.21	16316.47	18028.65	18368.86	20048.90	17869.18	14577.57	10485.13	10313.37	10000.00	10000.00
47	15052.31	18124.65	19408.45	21319.00	22906.13	25266.65	21451.33	16292.84	10598.66	10361.58	10000.00	10000.00
48	10212.96	13281.91	15513.04	18057.55	20507.01	23062.23	19860.54	15317.38	10093.91	10101.78	10000.00	10000.00
49	14181.74	15820.62	17251.65	19346.92	21224.27	23801.65	20303.76	15467.93	10000.00	10084.86	10012.75	10000.00
50	10623.46	14953.70	16814.16	18996.03	20835.36	23844.86	20406.19	15765.47	10500.69	10324.09	10000.00	10000.00

Table C.3 (c) Optimal storage levels for EOL

Year	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
1	8125.31	9582.99	10550.90	11459.88	12476.92	8727.07	10127.29	14891.94	19370.00	14941.49	9934.04	5000.00
2	5000.00	6728.86	8287.69	8852.83	9319.39	8917.35	10419.13	16126.91	19370.00	14943.23	9927.00	5000.00
3	5627.60	6190.90	6178.37	5895.32	5833.95	5000.00	8237.19	15092.36	19370.00	14943.94	9939.20	5000.00
4	6547.22	8941.38	9703.83	10163.63	10494.55	10574.02	10381.35	17259.83	19370.00	14901.52	9839.30	5000.00
5	5735.28	7390.95	8576.85	9259.20	9891.04	7281.41	9437.79	14889.57	19370.00	14912.79	9889.63	5000.00
6	6404.74	7818.60	7902.89	7148.85	6960.86	5000.00	7948.93	15533.32	19370.00	14855.73	9766.73	5000.00
7	5367.93	6615.63	7039.80	7415.77	7673.33	5496.89	8445.77	15158.98	19370.00	14921.19	9883.22	5000.00
8	5602.64	8396.83	10124.41	11168.99	11505.96	8171.95	10034.25	15144.27	19370.00	14895.34	9873.14	5000.00
9	12576.25	13473.08	14501.59	16089.88	18033.13	13290.51	12309.99	18073.29	19370.00	14988.89	10106.24	5000.00
10	6607.04	8081.88	9624.64	10674.57	11381.45	8121.56	7795.84	9929.77	11785.15	11367.62	10816.96	5000.00
11	7298.42	8893.91	10049.77	11430.60	12997.58	9010.71	10280.04	14952.24	19370.00	14906.78	9906.43	5000.00
12	5000.00	6386.72	6322.44	6250.13	5713.58	5649.30	8568.04	16011.84	19370.00	14946.11	9907.34	5000.00
13	5858.65	6331.91	6526.33	6985.91	7446.21	5301.71	7965.29	15282.18	19370.00	14869.70	9785.45	5000.00
14	12396.04	13408.88	14728.66	15684.20	16816.26	11381.20	11637.73	15736.89	19370.00	14899.71	9864.76	5000.00
15	10240.34	12017.18	13175.63	13862.66	15228.55	9390.77	9231.64	15243.53	19370.00	15265.62	10614.35	5000.00
16	5144.33	7132.10	7391.60	6780.63	6176.40	5000.00	7950.17	15454.96	19370.00	14921.06	9860.77	5000.00
17	6799.88	8119.21	9163.79	10077.12	10952.47	7451.88	7346.19	9462.59	11281.69	10495.28	9544.27	5000.00
18	12983.07	14958.17	16085.78	17317.84	18169.86	14691.41	14470.69	17193.07	18438.08	14692.86	10425.84	5000.00
19	5000.00	6871.42	7190.16	6978.56	7097.80	5785.70	8228.90	15420.90	19370.00	14924.53	9882.23	5000.00
20	5154.78	6452.82	6144.24	5709.78	5483.73	5000.00	8077.55	15635.87	19370.00	14894.80	9845.21	5000.00
21	8565.02	10178.20	10712.19	11305.10	11640.42	8605.36	10197.05	14979.84	19370.00	14862.77	9811.02	5000.00
22	5626.41	7251.28	7666.11	7167.60	6858.83	5000.00	7912.54	15350.98	19370.00	14884.25	9829.49	5000.00
23	11426.80	13442.41	14571.71	15489.41	16351.59	10909.32	10570.48	15101.48	19370.00	14909.34	9901.11	5000.00
24	5070.18	5954.57	5759.59	5320.88	5000.00	5118.39	8102.75	16168.74	19370.00	14892.97	9820.61	5000.00
25	8188.42	9071.48	10195.57	11215.20	12424.86	8731.74	10204.99	14953.55	19370.00	14894.71	9864.93	5000.00
26	5000.00	6205.28	6915.05	8147.12	9046.54	6502.64	8927.43	15179.30	19370.00	14977.05	9960.75	5000.00
27	11806.84	13251.86	14712.91	15654.36	16520.63	11521.83	10986.88	17530.45	19370.00	14920.86	9903.95	5000.00
28	5316.08	6539.54	6823.49	7144.93	6706.91	5000.00	7643.30	15534.61	19370.00	14871.83	9798.36	5000.00
29	5211.68	6147.48	6490.40	7117.23	7648.68	5520.91	8128.11	15305.25	19370.00	14895.82	9837.82	5000.00
30	5000.00	6868.83	8419.63	9611.52	10580.93	11807.64	13776.09	16900.81	19370.00	14918.66	9885.93	5000.00
31	5199.09	7302.07	7937.74	8343.07	8305.06	6263.55	9099.34	15017.65	19370.00	14981.72	9992.03	5000.00
32	6404.49	7768.98	8996.86	10010.40	11507.09	8016.66	9439.85	15181.53	19370.00	14890.96	9842.01	5000.00
33	5536.27	6778.40	7119.14	7224.19	6981.11	6989.73	9415.04	16238.77	19370.00	14881.58	9845.30	5000.00
34	5370.14	6480.04	6772.51	6711.44	6385.36	5000.00	7919.80	15241.88	19370.00	14905.45	9874.16	5000.00
35	9310.25	10097.64	10860.40	11358.00	11881.01	8605.05	10103.53	14944.34	19370.00	14920.26	9891.99	5000.00
36	11432.95	13767.31	15617.02	16936.41	17666.32	12122.20	11814.40	15708.71	19370.00	14928.75	9888.26	5000.00
37	5613.59	7270.11	8060.47	8459.79	8631.10	6070.83	8412.43	15236.99	19370.00	14874.16	9837.70	5000.00
38	9708.55	11721.99	12699.04	13903.76	15152.56	10480.42	11086.51	15411.72	19370.00	14912.46	9921.08	5000.00
39	5494.90	7275.02	8477.02	9204.95	9948.27	7119.86	9098.22	15390.81	19370.00	14894.36	9818.66	5000.00
40	5000.00	7096.93	8489.70	9475.35	9991.16	10071.51	11278.38	17731.83	19370.00	14950.33	9925.22	5000.00
41	5105.90	6260.73	6666.62	7313.55	7548.23	8106.79	9975.33	16874.97	19370.00	14897.40	9846.50	5000.00
42	5000.00	6995.84	8420.75	9947.44	11196.79	9827.74	11707.02	15738.05	19370.00	14959.43	9971.84	5000.00
43	7342.29	11311.96	14567.59	17081.29	16788.91	17006.40	15988.79	17348.57	19370.00	16308.77	12820.46	5000.00
44	11002.14	14349.89	15870.06	16265.31	16645.21	11606.77	11614.70	15661.09	19370.00	14906.32	9858.79	5000.00
45	5581.40	8031.03	9576.65	10886.58	12071.31	8475.80	10243.16	15451.58	19370.00	14839.53	9802.23	5000.00
46	5000.00	6415.23	7159.83	7558.91	7263.42	7780.35	9827.88	16608.64	19370.00	14901.43	9860.33	5000.00
47	5856.87	6991.96	7409.57	7620.95	7695.91	5591.18	8510.97	15185.21	19370.00	14881.66	9837.44	5000.00
48	8027.23	9648.92	11193.43	12277.25	13328.70	9353.14	10703.37	15237.56	19370.00	14918.20	9906.79	5000.00
49	5395.74	5584.32	6038.99	6233.86	6352.46	5000.00	8163.09	15048.38	19370.00	14954.73	9943.19	5000.00
50	5271.64	7421.91	8484.10	9141.86	9643.61	6986.21	9231.16	15112.63	19370.00	14901.45	9857.71	5000.00

Table C.3 (d) Optimal storage levels for LG2

REFERENCES

Arvanitidis, N.V., and J. Rosing, "Composite Representation of Multi-Reservoir Hydro-Electric Power System," IEEE Trans. Power Appar. Syst., PAS-89 (2), 319-335, 1970.

Askew, A.J., "Optimum Reservoir Operating Policies and the Imposition of Reliability Constraints," Water Resour. Res., 10 (1), 51-56, 1974a.

Askew, A.J., "Chance-Constrained Dynamic Programming and The Optimization of Water Resource Systems," Water Resour. Res., 10 (6), 1099-1106, 1974b.

Bellman, R.E., "Dynamic Programming," Princeton University Press, Princeton, N.J., 1957.

Bogle, M., and M. O'Sullivan, "Stochastic Optimization of a Water Supply System," Water Resour. Res., 15 (4), 778-786, 1779.

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

Cailliez, F., "Analyse des Données," Séminaire de Mathématiques Supérieures, Université de Montréal, Montréal, 1984.

Chvatal, V., "Linear Programming", McGill University, Freeman, New York, 1983.

Colleter, P., and P. Lederer, "Cptimal Operation Feedbacks for French Hydro-Power System," CORS-TIMS-ORSA Joint National Meeting, Toronto, May 3-6, 1981.

Cooley, W., and P.R. Lohnes, "Multivariate Data Analysis," John Willey, New York, 1971.

Croley, T.E., "Sequential Stochastic Optimization for Reservoir System," Journal of the Hydraulics Division, ASCE, 100 (HY1), 201-218, 1974.

Davis, R.E., "Stochastic Dynamic Programming for Multi-Reservoir Hydro Optimization," Technical Memorandum No. 15, Systems Control Inc., Palo Alto, Ca., 1972.

Davis, R.E., and R. Pronovost, "Two Stochastic Dynamic Programming Procedures for Long-Term Reservoir Management," IEEE PES Summer Meeting, San Francisco, July 9-14, 1972.

Delebecque, F., and J.P. Quadrat, "Contribution of Stochastic Control Singular Perturbation Averaging and Team Theories to an Example of Large-Scale Systems: Management of Hydro-Power Production," IEEE Trans. Automat. Contr., AC-23 (6), 209-221, 1978.

Duran, H., C. Pueh, J. Diaz, and G. Sanches, "Optimal Operation of Multireservoir Systems Using an Aggregation-Decomposition Approach," IEEE Tans. Power Appar. Syst., PAS-104 (8), 2086-2092, 1985.

El-Hawary, M., and G.S. Christensen, "Optimal Economic Operation of Electric Power Systems," Academic Press, New York, 1979.

EPRI EL-1659, "Pool daily Fuel Scheduling, Volume 1: Technical Manuel," Electric Power Research Institute, 1981.

Gablinger, M., and D.P. Loucks, "Markov Models for Flow Regulation," J. Hydraul. Div., ASCE, 96 (HY1), 165-181, 1970.

Gagnon, C.R., R.H. Hicks, S.L.S. Jacoby, and J.S. Kowalik, "A Nonlinear Programming Approach to a Very Large Hydroelectric System Optimization," Math. Programm., 6 (1), 28-41, 1974.

Gendre, F., "L'Analyse Statistique Multivariée," Librairie Droz, Genève, 1976.

Gnanadesikan, R., "Methods for Statistical Data Analysis of Multivariate Observations," John Willey, New York, 1977.

Habibollahzadeh, H., "Application of Mathematical Programming to Short-Term Operation Planning of Hydrothermal Power System," The Royal Institute of Technology, Stockholm, 1984.

Haimes, Y.Y., "Hierarchical Analysis of Water Resources Systems," McGraw Hill, New York, 1977.

Hanscom, M., and L. Lafond, "Modelling and Resolution of the Deterministic Mid-Term Energy Production Problem for The Hydro-Québec System," Hydro-Québec Internal Report, IREQ-1453, July 1976.

Hillier, F.S., and G.J. Lieberman, "Operations Research," Holden-Day, San Francisco, 1974.

Hogan, A.J., J.G. Morris, and H.E. Thompson, "Decision Problems under Risk and Chance Constrained Programming: Dilemmas in the Transition," Manage. Sci., 27 (6), 698-716, 1981.

Houck, M.H., and J.L. Cohon, "Sequential Explicitly Stochastic Linear Programming Models: A proposed Method for Design and Management of Multi-Purpose Reservoir system," Water Resour. Res., 14 (2), 161-168, 1978.

Houck, M.H., and B. Datta, "Performance Evaluation of a Stochastic Optimization Model for Reservoir Design and Management with Explicit Reliability Criteria," Water Resour. Res., 17 (4), 827-832, 1981.

Houck, M.H., J.L. Cohon, and C. Revelle, "Linear Decision in Reservoir Design and Management, 6, Incorporation of Economic Efficiency Benefits and Hydroelectric Energy Generation," Water Resour. Res., 16 (1), 196-200, 1980.

Hydro-Québec, "Principes de Planification des Equipements de Pointe", Service Plan des Installations, April 1983.

Johnson, R.A., and D.N. Wichern, "Applied Multivariable Statistical Analysis," Prentice-Hall, New Jersey, 1982.

Kendall, M., "Multivariate Analysis," Macmillan, New York, 1980.

Lasdon, L.S., "Optimization Theory for Large Systems," Macmillan, New York, 1970.

Lederer, P., P. Torrion, and J.P. Bouttes, "Overall Control of an Electricity Supply and Demand System: A Global Feedback for the French System," 11th IFIP Conference on System Modelling and Optimization, Copenhagen, July 1983.

Little, J.D.C., "The Use of Storage Water in a Hydroelectric System," Oper. Res., 3 (2), 187-197, 1955.

Loucks, D.P., "Computer Models for Reservoir Regulation," ASCE, 94 (SA4), 657-669, 1968.

Luenberger, D.G., "Introduction to Linear and Nonlinear Programming," Addison-Wesley, Reading, Mass., 1973.

Masse, P.B., "Les réserves et la Régulation de l'Avenir dans la Vie

Economique," Hermann and Cie, Paris, 1946.

McCuen, R.H., and W.M. Snyder, "Hydrologic Modeling: Statistical Methods and Applications," Prentice-Hall, New Jersey, 1986.

McKerchar, A.I., "Optimal Monthly Operation of Interconnected Hydro-Electric Power Storages," Journal of Hydrology, 25, 137-158, 1975.

Miller, I., and J.E. Freund, "Probability and Statistics for Engineers," Prentice-Hall, New Jersey, 1977.

(MPSX), "IBM Mathematical Programming System Extended/370 (MPSX/370), Program Reference Manual, IBM, Order Number SH19-1095-3, 1979.

Murtagh, B.A., "Advanced Linear Programming: Computation and Practice," McGraw-Hill, New York, 1981.

Nazareth, J.L., "Computer Solution of Linear Programs," Oxford University Press, 1987.

Nemhauser, G.L., "Introduction to Dynamic Programming," John Wiley, New York, 1966.

Pereira, M.V.F., and L.M.V.G. Pinto, "Operation Planning of Large-Scale Hydro-Electrical Systems," TIMS/ORSA Meeting, San Francisco, May 14-16, 1984.

Pronovost, R., "Analyse et Synthèse des Apports Hydraulyques," Hydro Québec Internal Report, IREQ-1034, July 1974.

Pronovost, R., J. and Boulva, "Long-Range Operation Planning of a Hydro-Thermal System Modelling and Optimization," CEA Meeting, Toronto, March 13-17, 1978.

Quintana, V.H., and A.Y. Chikhani, "A Stochastic Model for Mid-Term Operation Planning of Hydro-Thermal Systems with Reservoir Inflows," IEEE Trans. Power Appar. Syst., PAS-100 (3), 1119-1127, 1981.

Read, E.G., "Optimal Operation of Power Systems," University of Canterbury, 1979.

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

Roefs, T.G., and L.D. Bodin, "Multi-Reservoir Operation Studies," Water Resour. Res., 6 (2), 410-420, 1970.

Rossman, L., "Reliability-Constrained Dynamic Programming and Randomized Release Rules in Reservoir Management," Water Resour. Res., 13 (2), 247-255, 1977.

Saad, M., and A. Turgeon, "Application of Principal-Component Analysis to Long-Term Reservoir Management," Water Resour. Res., 24 (7), 907-912, 1988a.

Saad, M., and A. Turgeon, "Application de l'Analyse en Composantes Principales pour la Gestion à Long-terme des Installations Hydro-électriques," Optimization Days 1988, Montréal, May 2-4, 1988b.

Sachdeva, S.S., "Bibliography of Optimal Reservoir Drawdown for the Hydro-Thermal Power System Operation," IEEE Trans. Power Appar. Syst., PAS-101 (6), 1487-1496, 1982.

Schweig, Z., and J.A. Cole, "Optimal Control of Linked Reservoir," Water Resour. Res., 4 (3), 479-497, 1968.

Sherkat, V.R., R. Campo, K. Moslehi, and E.O. Lo, "Stochastic Long-Term Hydro-Thermal Optimization for a Multireservoir System," IEEE Trans. Power Appar. Syst., PAS-104 (8), 2040-2050, 1985.

Sniedovich, M., "A Variance-Constrained Reservoir Control Problem," Water Resour. Res., 16 (2), 271-274, 1980.

Sobel, M.J., "Reservoir Management Models," Water Resour. Res., 11 (6), 767-777, 1975.

Su, S.Y., and R.D. Deiniger, "Modeling the Regulation of Lake Superior under Uncertainty of Future Water Supplies," Water Resour. Res., 10 (1), 11-25, 1974.

Turgeon, A., "Optimal Unit Commitment," IEEE Transactions on Automatic Control, AC-22 (2), 223-227, April 1977.

Turgeon, A., "Optimal Operation of Multireservoir Power Systems with Stochastic Inflows", Water Resources Research, 16 (2), 275-283, 1980.

Turgeon, A., "Optimal Short-Term Hydro Scheduling from the Principle of Progressive Optimality,", Water Resources Research, 17 (3), 481-486, 1981a.

Turgeon, A., "A Decomposition Method for the Long-Term Scheduling of Reservoirs in Series," Water Resources Research, 17 (6), 1565-1570, 1981b.

Turgeon, A., "Choix des Sites et Dimensionnement des Installations Hydroélectriques," Hydro-Québec Internal Report, IREQ-2291, December 1980.

Wood, A.J., and B.F. Wollenberg, " Power Generation, Operation, and Control," John Wiley, New York, 1984.

Yakowitz, S., "Dynamic Programming Applications in Water Resources,", Water Resour. Res., 18 (4), 673-696, 1982.

Yeh, W.W.G., "Reservoir Management and Operation Models: A State-of-the-Art Review," Water Resour. Res., 21 (2). 1797-1818, 1985.

Young, G.K, "Finding Reservoir Operation Rules," Journal of the Hydrau. Div., ASCE, 93 (HY6), 297-320, 1967.