PREFERENTIAL RESERVOIR CONTROL UNDER UNCERTAINTY

by

Roman Krzysztofowicz

Reports on Natural Resource Systems

No. 31

November 1978

Collaborative effort between the following Departments:

Hydrology and Water Resources

Systems and Industrial Engineering

The University of Arizona

Tucson, Arizona 85721

PREFERENTIAL RESERVOIR CONTROL UNDER UNCERTAINTY

by

Roman Krzysztofowicz

Reports on Natural Resource Systems

No. 31

November 1978

Collaborative effort between the following Departments:

Hydrology and Water Resources

Systems and Industrial Engineering

The University of Arizona
Tucson, Arizona 85721

PREFACE

This report is the second one of two reports devoted to the modeling of a decision process in real-time reservoir control under conditions of uncertainty. The investigation presented in Report 30, "Preference Criterion and Group Utility Model for Reservoir Control Under Uncertainty" concerns the derivation and assessment of a preferential control criterion. This preference criterion is then used in a stochastic control model which is the subject of the present report. Both reports originate from the doctoral dissertation completed by the author in March 1978 and accepted by the Faculty of the Department of Hydrology and Water Resources.

This work was supported in part by funds provided by the National Science Foundation under Grant GF-38183, "Decision Making Under Uncertainties in Water Resources Systems," and Grant ENG-7809365, "Bayesian Methodology for Rainflood Forecasting and Reservoir Control."

This report series constitutes an effort to communicate to practitioners and researchers the complete research results, including computer programs and more detailed theoretical developments, that cannot be reproduced in professional journals. These reports are not intended to serve as a substitute for the review and referee process exerted by the scientific and professional community in their journals.

ACKNOWLEDGEMENTS

I wish to give my thanks to Professors Donald Davis, William R. Ferrell, Jean Weber, and Martin Fogel with whom I was able to master my research capabilities in a never-to-be-forgotten atmosphere of stimulation and sincerity. I am grateful to Professor Sidney Yakowitz for indicating to me avenues of challenge in theoretical research.

My greatest debt of gratitude extends to Professor Lucien Duckstein, my advisor, who created for me an opportunity to study at the University of Arizona and who introduced me to the subject of decision theory and multiobjective decision-making. His enlightening advice stimulated and elevated all my research endeavors.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

ABSTRACT

A model for real-time control of a multipurpose reservoir under the conditions of uncertainty is developed. The control model is formulated as a multistage decision process. It is conceptualized in the form of two sub-processes. The first level process is a *Forecast-Strategy Process* which performs as an open-loop feedback controller. It is defined by a sequence of forecasts and optimal release strategies against these forecasts. At each forecast time (time of issuing the forecast), the optimal release strategy is computed for the time period equal to the lead time of the forecast, and it remains in execution until the next forecast time. The second level process, defined for each forecast time, is a *Control Process* which for the given forecast generates the release strategy satisfying the preference criterion (minimization of expected disutility). This process is formulated as a truncated Markovian adaptive controller performing on a finite set of discrete times--the same set which indexes the forecast inflow process.

To evaluate the past performance of the control, a set of measures of effectiveness is proposed. Computational aspects of the control model are analyzed. Structural properties of the reservoir control process are explored in the main theorem which assures the monotonicity of the optimal strategy with respect to one of the state variables. Also, the properties of the optimal strategy for the case of a categorical forecast are proven. Next, two suboptimal strategies are derived: (1) partial open-loop strategy and (2) naive/partial

open-loop strategy. Finally, a discretization procedure which guarantees convergence of the numerical solution is discussed, and the computational requirements of the optimal and two suboptimal strategies are compared.

CHAPTER 1

INTRODUCTION

1.1. Objective of the Study

From the standpoint of real-time operation of a reservoir, the multipurpose control problem may be reduced to a dual purpose problem of (1) *Flood Control* (FCO) under uncertain inflow and (2) *Conservation Control* (CCO) after the flood has receded (purposes of water supply, power generation, low flow augmentation, recreation, etc.).

The objective of this investigation is to develop a model for real-time reservoir control under the conditions of uncertainty. Three postulates form a foundation for this development:

1. The input to the control model is a stochastic real-time *forecast* of the reservoir inflow process over a finite time period.

2. The control process is guided by a *preference criterion* which reflects the decision maker's value judgments, including strength of preferences over operating attributes, trade-offs between the purpose of FCO and the purposes of CCO, and attitude toward risk.

3. The CCO is *imbedded* into FCO through (a) the attribute space of the preference criterion, which allows for explicit consideration of the trade-offs between reservoir purposes, and (b) the state space and time domain of the control process, which allows for maintaining the continuity of the control.

1.2. Organization and Perspective of the Study

Chapter 2 is devoted to the modeling of the reservoir control process. It begins with the definition of the time domain in which FCO is dovetailed with CCO. Next, the FCO is conceptualized in the form of two sub-processes: (1) Forecast-Strategy Process, which is modeled as an open-loop feedback controller, and (2) Control Process, which is modeled as a truncated Markovian adaptive controller. To evaluate the past performance of the FCO, a set of measures of effectiveness is proposed. Chapter 3 contains structural analysis of the model. Structural properties of the reservoir control process are explored in the main theorem which assures the monotonicity of the optimal strategy with respect to one of the state variables. Also, the properties of the optimal strategy for the case of a categorical forecast are proven. In Chapter 4, two suboptimal strategies are derived: (1) partial open-loop strategy and (2) naive/partial open-loop strategy. Computational aspects of the control model are analyzed in Chapter 5. A discretization procedure which guarantees convergence of the numerical solution is discussed, and the computational requirements of the optimal and two suboptimal strategies are compared.

1.3. Summary of the Preference Criterion Model

The reservoir control model proposed herein employs a *preference criterion* whose derivation and testing are described in Report 30 (Krzysztofowicz, 1978). A summary of the preference criterion model is given below.

The preference criterion for real-time flood control is developed within the framework of multiattribute utility theory. Toward this aim, the decision problem in reservoir control under the conditions of uncertainty is analyzed. It is shown that this decision problem may be modeled naturally as a game against nature. Consequently, the von Neumann-Morgenstern (1947) utility theory provides an appropriate axiomatization of the decision maker's value judgments. These value judgments underly the release decisions, and they reflect the decision maker's (1) strength of preferences over operating attributes, (2) trade-offs between reservoir purposes, and (3) attitude toward risk. For the selected two attributes, a rationale is given to support the utility independence assumption which results in the multiplicative disutility function. It is argued that *minimization of expected disutility* is a plausible and well motivated criterion for multi-purpose real-time reservoir control under uncertainty.

The most significant implications of the preference criterion may be summarized as follows.

1. It provides a valuation of both tangible and intangible consequences of the control decisions.

2. It encodes strength of preferences, trade-off judgments, and risk attitude of an individual or a group decision maker.

3. It guarantees consistently optimal (with respect to the underlying value system) decisions. The importance of this fact is particularly clear in the light of ample evidence about limited human

information-processing capabilities and inconsistencies of intuitive decisions.

4. It enables the optimization of the trade-offs between the purpose of FCO and the purposes of CCO to be performed in real-time. Consequently, the traditional concept of the fixed flood space reservation can be abandoned. Instead, flood storage space can be treated as an implicit function of the reservoir release whose optimal magnitude at any instant of time is determined through minimization of the expected disutility under the *current forecast* distribution of the inflow process. Operationally, then, if needed, the whole reservoir can be emptied in preparation for a large flood, or it can be refilled to the dam crest in the face of a severe drought. The superiority of the proposed concept over the traditional one is apparent, for it enhances the efficient utilization of the reservoir.

5. It allows the multipurpose (or multiobjective) control problem under conditions of uncertainty to be conveniently modelled as a multistage decision process and efficiently solved by single-criterion optimization methods such as dynamic programming.

CHAPTER 2

MODELING OF THE CONTROL PROCESS

In this chapter a stochastic model for real-time reservoir control is developed. The proposed approach breaks with hydrologic tradition in reservoir control studies in three aspects:

1. The input to the control model is a stochastic real-time forecast of the reservoir inflow process over a finite time period.

2. The trade-off between the purpose of flood control (FCO) and the purposes of conservation control (CCO) is optimized in real-time on the basis of the *current uncertainties* (encoded in the probabilistic inflow forecast) and the *current preferences* of the decision maker (encoded in the disutility criterion). Specifically, the concept of a fixed flood storage space, determined from the historical annual pattern of the flood potential, is abandoned. Instead, flood storage space is treated as an implicit function of the reservoir release whose optimal magnitude at any instant of time is determined through minimization of the *current risk*. This risk is defined for the current forecast distribution by expected disutility.

3. The FCO and CCO are dovetailed in the state space and time domain so that continuity of reservoir control in space-time can be maintained.

2.1. Time Domain of FCO and CCO

The FCO process is staged in time by the events END-OF-FLOOD defined as follows: Let $\{h(t) : t \varepsilon T\}$ be a realization of a process where h is the river stage measured above an initial damage level at a target location below the reservoir, and t is a parameter of a continuous time space T. Define an event

$$\text{END-OF-FLOOD} \equiv \{h(t - \varepsilon) > 0 \text{ and } h(t + \varepsilon) < 0, \text{ some } t \varepsilon T, \forall \varepsilon > 0\}.$$

Let t_B and t_E be epochs of two consecutive events END-OF-FLOOD. A *realization* of the FCO process is defined on the time interval $B = [t_B, t_E]$. Specifically, the FCO process begins at t_B and ends at t_E whereafter t_E becomes the starting time for the next realization, and so on.

Let $MM = \{m : m = 1, \ldots, M\}$ be a set of *forecast times* such that $t_m \varepsilon B$ for every $m \varepsilon MM$, and $\Delta t_m = t_{m+1} - t_m$. At each forecast time $m \varepsilon MM$, a *forecast* Φ_m of the rate of inflow X to the reservoir over a finite time period $[t_m, t_m + \lambda_m]$ is made on the basis of information available up to the time t_m. It is assumed that the forecast *lead time* $\lambda_m > \Delta t_m$. The reservoir remains in operation under the control model at any $t \varepsilon [t_m, t_{m+1}]$ $(m = 1, \ldots, M)$ if Φ_m specifies that

$$P\{X(t) > \text{THRESHOLD; some } t \varepsilon [t_m, t_m + \lambda_m]\} > 0.$$

Otherwise the reservoir remains in the operation under CCO requirements alone. An example of the time domain is shown in Figure 2.1.

2.2. Forecast-Strategy Process

The FCO is composed of a sequence of forecasts and strategies. This forecast-strategy mechanism is illustrated in Figure 2.2. Given the forecast Φ_m for the period $[t_m, t_m + \lambda_m]$, where $\lambda_m > \Delta t_m$, and the state of the system ω_m at time m, the control process (CP) problem is solved to yield an optimal release strategy S_m for the period $[t_m, t_m + \lambda_m]$. The risk associated with the initial state-time (ω_m, m), forecast Φ_m, and strategy S_m is $R_m = R_m(\omega_m, \Phi_m, S_m)$. The reservoir is operated according to S_m over the period $[t_m, t_{m+1}]$ at the end of which the system is in the state ω_{m+1}. At the time m+1 the next forecast Φ_{m+1} is issued for the period $[t_{m+1}, t_{m+1} + \lambda_{m+1}]$, where $\lambda_{m+1} > \Delta t_{m+1}$, and the CP problem is solved again for the current initial state ω_{m+1}. The solution is an optimal strategy S_{m+1} for the period $[t_{m+1}, t_{m+1} + \lambda_{m+1}]$; the corresponding risk is R_{m+1}.

Forecast mechanism: The forecast generating mechanism is assumed to be a Bayesian information processor where for any $m \in MM$, the forecast Φ_m constitutes prior information and the forecast Φ_{m+1} constitutes posterior information. Furthermore, it is assumed that the reservoir inflow X is *predictable* in the following sense (Lorenz, 1973):

Definition 2.1: Let $X(t)$, $t > t_{m+1}$, be a predictand of concern whose prior distribution function F_m is specified by Φ_m, and the posterior distribution function F_{m+1} is specified by Φ_{m+1}.

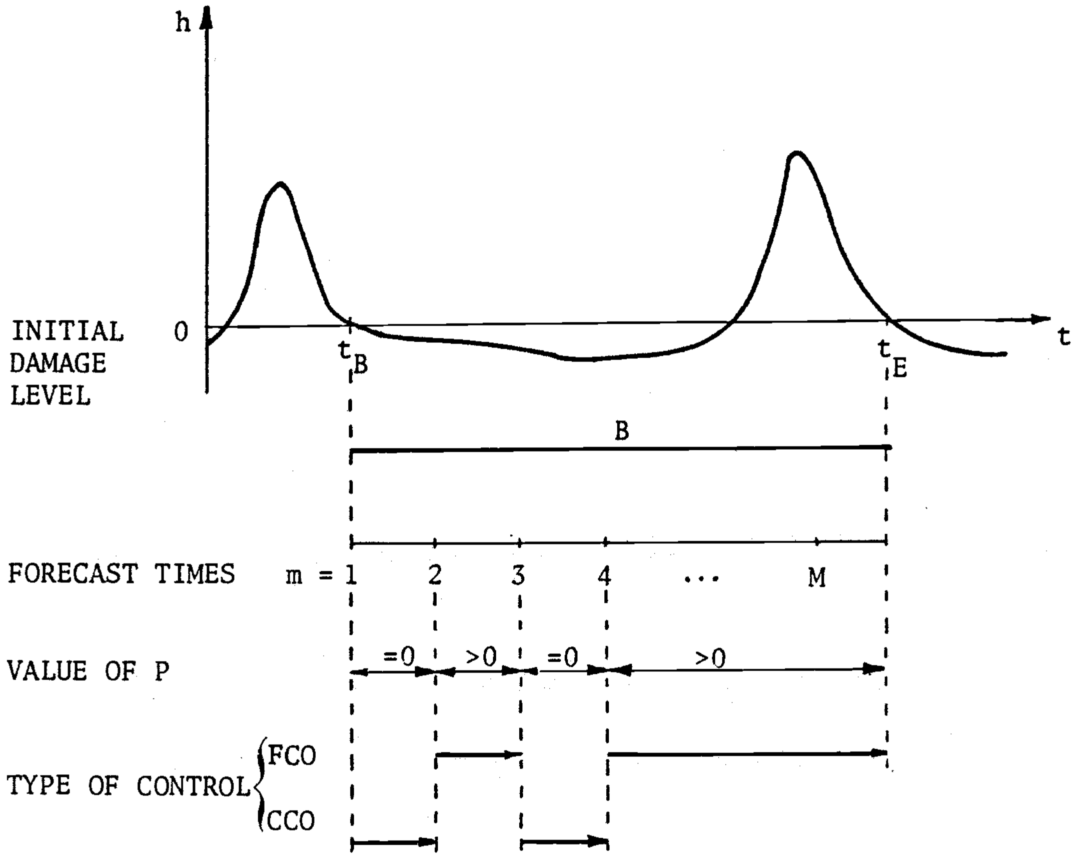

Figure 2.1. Time Domain of FCO and CCO

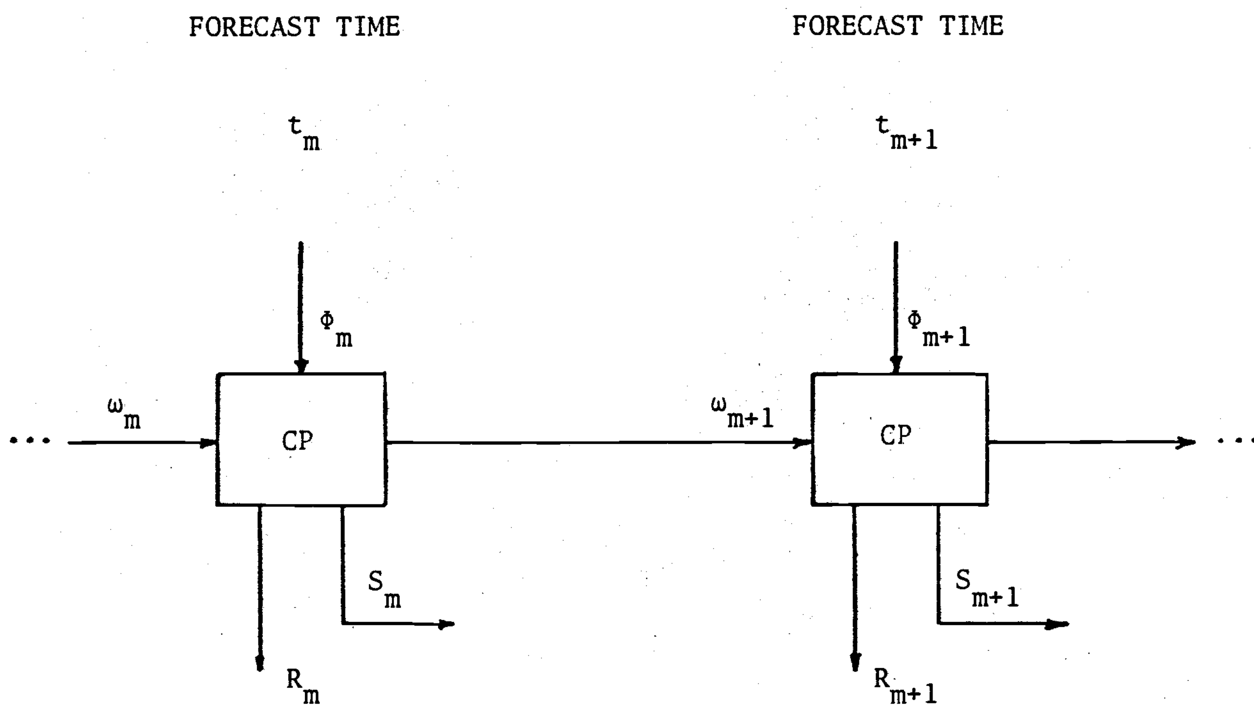

Figure 2.2. Forecast-Strategy Process

If $F_m(x(t)) = F_{m+1}(x(t))$ for all $x(t)$, then $X(t)$ is *unpredictable.*

If $F_m(x(t)) \neq F_{m+1}(x(t))$ for some $x(t)$ and $\mathrm{Var}[F_{m+1}] < \mathrm{Var}[F_m]$ then $X(t)$ is at least *partially predictable.*

If $F_{m+1}(x(t))$ = (step function) then $X(t)$ is *perfectly predictable.*

Strategy generating mechanism: The strategy generating mechanism has been described as an *open-loop feedback control* (OLFC). At each time $m \in MM$ the new information is used to update the prior forecast Φ_{m-1}. As soon as the posterior forecast Φ_m becomes available, the optimal strategy is recomputed and applied to the current state ω_m which resulted from the operation of the reservoir up to the time t_m. However, the optimal strategy S_m is selected as if no further forecasts will be received in the future.

The OLFC is *quasi-adaptive* (Bertsekas, 1976, p. 199), that is it satisfies

$$R^* \leq R \leq R_0^*$$

where

R = optimal risk of the OLFC,

R^* = optimal risk of a *closed-loop feedback control* (CLFC),

R_0^* = optimal risk of an open-loop control, i.e., the control that does not use updated forecasts.

An *adaptive control*, that is one which satisfies

$$R^* \leq R < R_0^*,$$

could be obtained with a closed-loop feedback, wherein at any time $m \in MM$ the optimal strategy is computed for the remaining part of the FCO process in a manner that accounts for the further forecasts which will be available at times m+1, ..., M. Although the OLFC is only suboptimal, it has been chosen here in order to keep the model relatively simple. In contrast to the OLFC, the CLFC would require a detailed modeling of the forecasting process--a formidable task, which is much beyond the scope of the present research.

An application of the OLFC to the FCO requires solution of M identical control problems in real time. One may think then of FCO as being composed of two hierarchically related processes: The first level process is a *Forecast-Strategy Process* (FSP) defined by a sequence $\{\omega_m, \Phi_m, S_m, R_m\}_{m=1}^{M}$ on the time domain B. The second level process is a *Control Process* (CP) defined for every $m \in MM$ on the time domain $[t_m, t_m + \lambda_m]$. The CP takes ω_m as an initial state, and for the given Φ_m generates (S_m, R_m); S_m is then applied at discrete control times $\{n : t_n \in [t_m, t_{m+1}], \forall n = 1, \ldots, K\}$. Figure 2.3 illustrates the dependence between the time scales of FSP and CP.

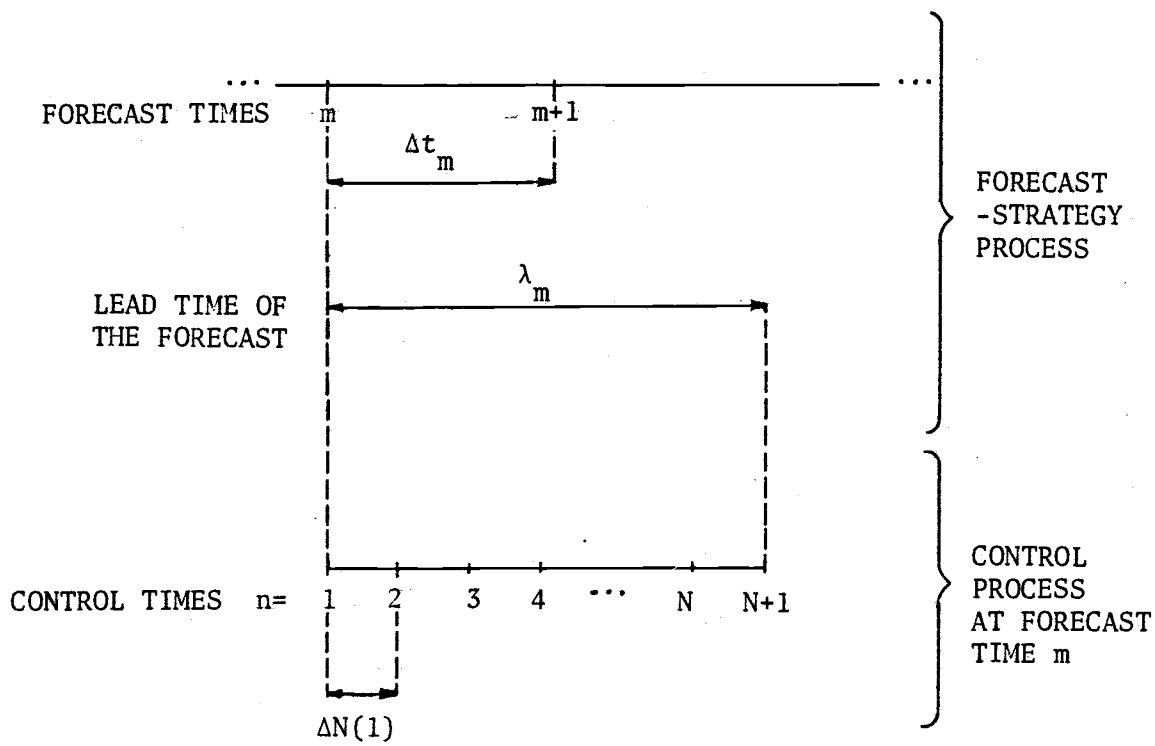

Figure 2.3. Time Scale of Forecast-Strategy Process and Control Process

2.3. Control Process

2.3.1. Preliminaries

The reservoir CP is formulated as an adaptive control process (or CLFC) as illustrated in Figure 2.4. The controls are applied at discrete times; the first control time, $n = 1$, matches the forecast time t_m. Given the state $\omega(n)$ at the control time n, a control element $a(n)$ is selected from the strategy S. The system moves then to the state $\omega(n+1)$ at the control time n+1 according to the law of motion (Φ, T, Ψ). When the process reaches the terminal state $\omega(N+1)$, the disutility $u(\omega(N+1))$ is incurred. At any time n, the state $\omega(i)$ for $i > n$ is a random variable whose distribution function can be obtained from the forecast Φ via a flood crest operator T and a storage operator Ψ. Hence, at any time $n < N+1$ only the expected disutility can be known. In particular, at $n = 1$ ($\equiv$ forecast time t_m) an optimal control strategy S_m can be selected so as to minimize the expected disutility (risk) R_m for the period $[t_m, t_m + \lambda_m]$.

Henceforth, upper case letters will be used to denote both sets and random variables under the expected value operation; the precise meaning should be clear from the context. A variable indexed by the set of control times, e.g., x(n), will often be written as x with omission of the time index, whereas x(n+1) will be denoted by x'. Since an identical CP is repeated for every forecast time $m \in MM$, the elements of the CP will not be indexed by m.

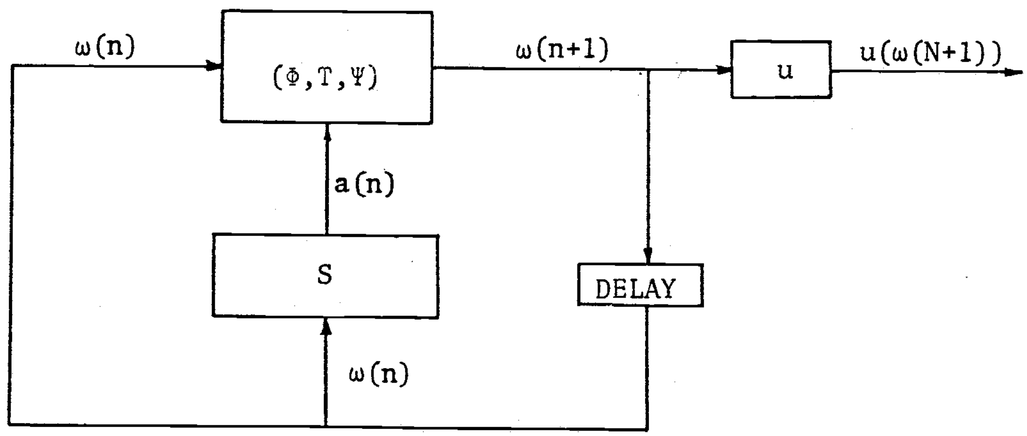

Figure 2.4. Control Process

2.3.2. The Model

Definition 2.2: The set of *control times*, NN, is an initial segment of the set of positive integers:

$$NN = \{n : n = 1, \ldots, N\} .$$

The real time interval between control times i and i+1 is $\Delta N(i)$. The first control time, n = 1, matches the forecast time t_m, and $\Sigma_{i=1}^{N} \Delta N(i) = \lambda_m$.

Definition 2.3: The *state space*, Ω, is a Cartesian product $\Omega = X \times Z \times Y$, a subset of the nonnegative segment of R^3 space. An element $\omega \in \Omega$ is a *state* $\omega = (x,z,y)$, where

x = the rate of inflow to the reservoir,

z = the maximum flood level measured above initial damage level at a target location below the reservoir,

y = the storage space in the reservoir.

For example, $\omega(n) = (x(n),z(n),y(n))$ $(1 < n \leq N)$ is the state of the system at the control time n, where x(n) is the average rate of inflow *during* the time interval $\Delta N(n-1)$, z(n) is the maximum flood level *up to* the control time n, and y(n) is the storage space *at* the control time n. It is assumed that both Z and Y are closed and bounded, i.e., $Z = [z_o, z_e]$ and $Y = [y_o, y_e]$ where (z_o, y_o) is *the most desirable* consequence of FCO and (z_e, y_e) is *the least desirable* consequence of FCO. For instance, z_o = level above which flood damage begins, z_e= level above which any increase in the depth of flooding causes no increase in flood damage, y_o = full reservoir, and y_e = empty reservoir.

Definition 2.4: A *condition on state space* is a set $\theta = \{\theta(n) : n = 1, \ldots N+1\}$ where the elements $\theta(n)$ are subsets of Z x Y.

Definition 2.5: The *control set*, A, is a set-valued mapping defined on Y x NN:

$$A = \{A(y,n) : y \varepsilon Y, n \varepsilon NN\}$$

where A(y,n) is the set of admissible controls available to the decision maker at the control time n when the third state coordinate is y. For example, A(•,n) may reflect the relationship between the head over spillway crest and the discharge from the spillway. A control element a ε A is the rate of release from the reservoir.

Definition 2.6: The *flood forecast*, Φ, (or shortly the forecast) is a family of conditional distribution functions on X:

$$\Phi = \{F(\cdot|x,n) : x \varepsilon X, n \varepsilon NN\}$$

where F(•|x,n) is the distribution function of the rate of inflow x(n+1) at the control time n+1 conditioned by the rate of inflow x(n) at the control time n.

Definition 2.7: The *flood crest operator* is a mapping $T : X \times Z \times Y \times A \times NN \rightarrow Z$ whose equation has the form

$$z(n+1) = T(x(n+1), z(n), y(n), a(n), n).$$

This study assumes that

$$z(n+1) = \max \{z(n),\ zz(x(n+1),y(n),a(n),n)\}$$

where z(n) is the maximum flood level up to the control time n, and zz(x(n + 1),y(n),a(n),n) is the maximum flood level during time period ΔN(n) induced by the inflow rate x(n+1), storage space y(n), and control a(n).

<u>Definition 2.8</u>: The *storage operator* is a mapping $\Psi : X \times Y \times A \times NN \to Y$ whose equation has the form

$$y(n+1) = \Psi(x(n+1),y(n),a(n),n).$$

This study assumes that

$$y(n+1) = \begin{cases} y_o & \text{if } yy < y_o \\ yy & \text{if } y_o \le yy \le y_e \\ y_e & \text{if } yy > y_e \end{cases}$$

where

$$yy = y(n) + [a(n) - x(n+1)]\ \Delta N(n).$$

<u>Definition 2.9</u>: The *law of motion* for the reservoir control process is a 3-tuple (Φ,T,Ψ).

Note that z(n+1) and y(n+1) are random variables because x(n+1) is random. Furthermore, T and Ψ may themselves be stochastic operators. For examples, the basic purpose of T is routing of the reservoir outflow to the target location. This routing may be affected by flows in the tributaries which empty to the main channel between the dam and the target location. If these flows are known only probabilistically then T is stochastic. Further development in this chapter is based on the assumption that both T and Ψ are deterministic.

The mechanism controlling the rate of release from the reservoir may now be summarized as follows. Given the state (x(n),z(n),y(n)) at time $n \in NN$, a control a(n) is applied whereafter the input x(n+1) is determined according to $F(\cdot|x(n),n)$, and the system moves to the state (x(n+1),z(n+1),y(n+1)). The actual rate of release during period $\Delta N(n)$ is constant and equal to the control a(n) *only if* $y_o \leq yy \leq y_e$. If at some epoch *during* the period $\Delta N(n)$ the reservoir reaches the state of fullness ($yy < y_o$) or the state of emptiness ($yy > y_e$), then obviously the release rate a(n) cannot be further maintained for it would lead to an infeasible state $y < y_o$ or $y > y_e$. Assuming that for the rest of the period $\Delta N(n)$ the reservoir remains in the extreme state it has reached, the release rate *cannot be further controlled* in the sense that it depends only on the input x(n+1) and characteristics of the dam and reservoir. This fact explains the appearance of y(n) and x(n+1) in the operator T.

Definition 2.10: A *policy*, $\bar{a}$, is a sequence of controls $\{a(i)\}_{n \leq i \leq N}$, where a(i) is the rate of release at the i-th control time and n is the initial time.

Definition 2.11: A *trajectory*, $\bar{\omega}$, is a sequence of states $\{\omega(i)\}_{n\leq i\leq N+1}$ associated with a policy $\bar{a}$ and an initial state-time $(\omega(n),n)$. The last element $\omega(N+1)$ of the trajectory is the terminal state resulting from the last control $a(N)$.

Definition 2.12: A *strategy* is a function $S : \Omega \times NN \to A$. $S(\omega,n)$ is the control element at the time n when the state of the system is ω.

Definition 2.13: The *set of feasible strategies*, σ, is a set of strategies such that $S \in \sigma$ implies

$$P[(z',y') \in \theta(n+1) \mid x,z,y,S,n] = 1 \quad \forall(z,y) \in \theta(n),\ n \in NN.$$

Definition 2.14: A *disutility function* is a mapping $u : Z \times Y \to R$. Disutility of a realization of the FCO process with the terminal states $z(N+1)$ and $y(N+1)$ is $u(z(N+1),y(N+1))$.

Definition 2.15: The *initial condition* is a 3-tuple $(x(1),z(1),y(1))$ where $x(1)$ is the inflow rate *prior* to the current forecast time, $z(1)$ is the maximum flood level since the beginning of the FSP *up to* the current forecast time, and $y(1)$ is the storage space *at* the current forecast time. Because at $n = 1$ the antecedent inflow rate $x(1)$ is already known, the forecast for $n = 1$ may be conditioned implicitly on the value of $x(1)$, which in this case does not have to be included in the initial condition.

<u>Definition 2.16</u>: A *reservoir Control Process (CP)* is a 7-tuple (NN, Ω, A, Φ, T, Ψ, u) where NN is a set of control times, Ω is a state space, A is a control set, Φ is a flood forecast, T is a flood crest operator, Ψ is a storage operator, and u is a disutility function.

<u>Definition 2.17</u>: A strategy S (S ε σ) is a *solution* to a reservoir CP problem with an initial state-time (ω,n) if for any S' ε σ

$$E[u(Z,Y) \mid \omega,S,n] \leq E[u(Z,Y) \mid \omega,S',n]$$

where Z = Z(N+1) and Y = Y(N+1). The *expected disutility* associated with an optimal strategy S and initial state-time (ω,n) will be denoted by R(ω,n).

The elements of the reservoir CP model are summarized in Table 2.1.

Table 2.1. Summary of the Elements of the Control Model

Element	Symbol
Input Elements	
1. Set of control times	$NN = \{n : n = 1, \ldots, N\}$
2. Condition on state space	$\theta = \{\theta(n) : n = 1, \ldots, N+1\}$
3. Control set	$A = \{A(y,n) : y \varepsilon Y, n \varepsilon NN\}$
4. Flood forecast	$\Phi = \{F(\cdot\|x,n) : x \varepsilon X, n \varepsilon NN\}$
5. Disutility function	u
6. Initial condition	$(x(1),z(1),y(1))$
Internal Elements	
7. State space	$\Omega = X \times Z \times Y$ $\omega \varepsilon \Omega, \omega = (x,z,y)$
8. Flood crest operator	T
9. Storage operator	Ψ
10. Policy	$\bar{a} = \{a(i) : n \le i \le N\}$
11. Trajectory	$\bar{\omega} = \{\omega(i) : n \le i \le N+1\}$
12. Set of feasible strategies	σ
Output Elements	
13. Optimal strategy	S
14. Expected disutility	R

2.3.3. Solution Algorithm

In a general formulation of an adaptive control process (Yakowitz, 1969), the loss function, L, is defined on a triple ($\bar{\omega}$, $\bar{a}$, n) where $\bar{\omega}$ and $\bar{a}$ are, respectively, trajectory and policy for times not less than n. A property required from the loss function by the dynamic programming theorem is separability (Yakowitz, 1969, p. 33; Bertsekas, 1976, p. 50). The loss function is said to be *separable* if it can be represented by

$$L(\bar{\omega},\bar{a},n) = \sum_{i=n}^{N} L(\omega(i),\omega(i+1),a(i),i).$$

The disutility function u (which is the loss function in the reservoir CP) is separable which is easily demonstrated by letting

$$L(\omega(i),\omega(i+1),a(i),i) = \begin{cases} u(z(N+1),y(N+1)) & \text{for } i = N, \\ 0 & \text{otherwise.} \end{cases}$$

From the structure of the law of motion for the reservoir CP, it is also clear that the trajectory $\bar{\omega}$ has *Markov* property, i.e., that the distribution function of $\bar{\omega}$ for times greater than n depends only on $\omega(n)$ and a(n), respectively, the state and control at time n.

An adaptive control process having a separable loss function and Markov type trajectory is called *Markovian*. The reservoir CP has thus been formulated as a truncated Markovian adaptive control process. Yakowitz (1969, p. 33) proved that a solution to such a process can be obtained by means of a dynamic programming algorithm. A specification of this algorithm for the reservoir CP follows.

Algorithm 2.1: (a) For every $n \varepsilon NN$, $x \varepsilon X$, and $(z,y) \varepsilon \theta(n)$, let $\Theta(x,z,y,n) = \{a : a \varepsilon A(y,n) \text{ and } P[(z',y') \varepsilon \theta(n+1) \mid x,z,y,a,n] = 1\}$.

(b) $S(x,z,y,N) = a^*$ where for every state (x,z,y), a^* is found as a solution to the equation

$$R(x,z,y,N) = \min_{a \varepsilon \Theta(x,z,y,N)} E[u(Z',Y')].$$

(c) $S(x,z,y,n) = a^*$, $n \varepsilon NN$, where for every state-time (x,z,y,n), a^* is found as a solution to the functional

$$R(x,z,y,n) = \min_{a \varepsilon \Theta(x,z,y,n)} E[R(X',Z',Y',n+1)],$$

where it is understood that

$$Z' = T(X',z,y,a,n),$$

$$Y' = \Psi(X',y,a,n),$$

and the expectation is taken with respect to X' having distribution $F(\cdot|x,n)$.

(d) With an initial condition $(x(1),z(1),y(1))$ at the forecast time $m \varepsilon MM$, the expected disutility associated with the strategy $S_m = S$ for the period $[t_m, t_m + \lambda_m]$ is $R_m = R(x(1),z(1),y(1),1)$.

The following theorem (Yakowitz, 1969, p. 41) gives a sufficient condition for the existence of a solution.

Theorem 2.1: Assume that:

(1) $\Theta(n)$ is compact for each n,

(2) $F(\cdot|x,n)$ is continuous for every fixed (x,n),

(3) u is continuous,

(4) $\Theta(\omega,n) = \Theta(n)$ is a compact set, independent of ω at all control times n. Then the reservoir CP has a solution for every initial state-time (ω,n), and the dynamic programming algorithm is effective.

2.3.4. Stochastic Operators

There may exist situations where it is appropriate to assume that the operators T or Ψ, or both, are stochastic or quasi-stochastic. Most likely each particular case will require an individual consideration as to the best choice of the structure for T and Ψ. Herein a general case is illustrated. The required modifications of the basic CP model include Definitions 2.7, 2.8, and 2.9.

Definition 2.7': The *flood crest operator* is a family of conditional distribution functions on Z:

$$T = \{F_Z(\cdot|x',z,y,a,n) \ : \ x' \varepsilon X, \ z \varepsilon Z, \ y \varepsilon Y, \ a \varepsilon A, \ n \varepsilon NN\}$$

where $F_Z(\cdot|x(n+1),z(n),y(n),a(n),n)$ is the conditional distribution function of the maximum flood level $z(n+1)$ up to the control time $n+1$.

Definition 2.8': The *storage operator* is a family of conditional distribution functions on Y:

$$\Psi = \{F_Y(\cdot|x',y,a,n) \ : \ x' \varepsilon X, \ y \varepsilon Y, \ a \varepsilon A, \ n \varepsilon NN\}$$

where $F_Y(\cdot|x(n+1),y(n),a(n),n)$ is the conditional distribution function of the storage space $y(n+1)$ at the control time $n+1$.

Definition 2.9': The *law of motion* is a 3-tuple (Φ, T, Ψ), or equivalently a family of conditional distribution functions on Z and Y:

$$\{\phi_Z(\cdot|x,z,y,a,n),\ \phi_Y(\cdot|x,y,a,n)\ :\ x \in X,\ z \in Z,\ y \in Y,\ a \in A,\ n \in NN\}$$

where $\phi_Z(\cdot|x(n),z(n),y(n),a(n),n)$ is the conditional distribution function of $z(n+1)$ and $\phi_Y(\cdot|x(n),y(n),a(n),n)$ is the conditional distribution function of $y(n+1)$. The family $\{\phi_Z,\phi_Y\}$ can be obtained from (Φ,T,Ψ) by elementary probability calculus. For every $z' \in Z$,

$$\phi_Z(z'|x,z,y,a,n) = \int_X F_Z(z'|x',z,y,a,n)\,dF(x'|x,n) \ ,$$

and for every $y' \in Y$,

$$\phi_Y(y'|x,y,a,n) = \int_X F_Y(y'|x',y,a,n)\,dF(x'|x,n) \ .$$

2.4. Measures of Effectiveness

To evaluate the past performance of the FCO, a set of measures of effectiveness is proposed. The measures are defined for a single realization of the FCO process. For a sequence of realizations the measures can easily be derived under the assumption of additivity for disutilities. Let

RA = disutility actually incurred as a result of the forecast-strategy process $\{\omega_m, \Phi_m, S_m, R_m\}$,

RC = disutility that would have been incurred in the case of certainty, that is if at the beginning of the FCO process $\{x(t)\}$ were known for every $t \in B$,

RF = disutility that would have been incurred under a fixed control rule which does not make use of the forecasts.

Under an assumption that the inflow process $\{x(t)\}$ is predictable (Definition 2.1), the following relation holds

$$RC \leq RA \leq RF .$$

Definition 2.18: The *efficiency* of the FCO is defined by either of the two equations:

$$e_1 = \frac{RC}{RA} ,$$

$$e_2 = \frac{RF - RA}{RF - RC} ,$$

where $0 < e_1, e_2 \leq 1$. Both types of efficiency are conditioned by the initial state of the system and given forecasting procedure. In addition, e_2 is conditioned by the fixed control rule which has to be assumed. This could be a rule under which the reservoir was operating prior to the implementation of the forecasting. RF - RA may be viewed as the utility of the actual forecasting procedure while RF - RC as the utility of a perfect forecasting procedure.

Definition 2.19: To measure the performance of the forecasting procedure in any particular realization of the FCO process, we define an *error sequence*

$$\rho_m = \frac{\sum_{i=m}^{M} |R_i - RA|}{(M - m + 1)\ RA} \qquad m = 1, \ldots, M$$

where $\rho_m = 0$ implies an unbiased forecast sequence for times not less than m. In general, one may expect that $\rho_m \to 0$ as $m \to M$.

Definition 2.20: Under an assumption that $\{\rho_m\}$ is monotonically decreasing, the *efficient lead time* provided by the forecasting procedure at the error level α is defined by

$$\Lambda_\alpha = \sum_{i=k}^{M} \Delta t(i)$$

where k is the smallest integer such that $\rho_k < \alpha$.

CHAPTER 3

MONOTONICITY OF THE OPTIMAL STRATEGY

In this chapter, structural properties of the reservoir CP model are explored. It is revealed that the model structure itself (supplemented with a few mild assumptions) is sufficient to guarantee monotonicity of the optimal strategy $S(x,z,\cdot,n)$ with respect to the state variable y. A special case of the CP problem with flood forecasts degenerated to categorical statements is also examined.

3.1. Main Theorem

We begin this section by stating two lemmas which provide a key tool in proving the main theorem on monotonicity of the optimal strategy. The following notation is used throughout this section. For any fixed $x' \in X$, $z \in Z$, and $n \in NN$,

$$(z'|y,a) = T(x',z,y,a,n),$$

$$(y'|y,a) = \Psi(x',y,a,n),$$

$$(yy|y,a) = y+(a-x')\Delta N,$$

$$(u|y,a) = u((z'|y,a),(y'|y,a)),$$

$$(R'|y,a) = R(x',(z'|y,a),(y'|y,a),n+1),$$

$$E[u|y,a] = E[u((z'|y,a),(y'|y,a))].$$

Lemma 3.1: (1) For any $n \in NN$ and some $\Delta a > 0$ define: $y_1 = y_1$, $y_2 = y_1 + \Delta a \cdot \Delta N$, $y_3 = y_1 + 2\Delta a \cdot \Delta N$, and $a_1 = a_1$, $a_2 = a_1 - \Delta a$, $a_3 = a_1 + \Delta a$ such that $y_i \varepsilon Y$ and $a_j \varepsilon A(y_i,n)$ $i,j = 1, 2, 3$.

(2) For every $x' \varepsilon X$, define:

$$\alpha = (y'|y_1,a_3) - (y'|y_1,a_1),$$
$$\beta = (y'|y_2,a_1) - (y'|y_1,a_1),$$
$$\gamma = (y'|y_2,a_3) - (y'|y_1,a_3).$$

Then $\alpha = \beta$ and $\gamma \geq \alpha$ for every $x' \varepsilon X$.

Proof: By Definition 2.8 of Ψ:

$$\alpha = \{y_1 + (a_1+\Delta a-x')\Delta N\} - \{y_1 + (a_1-x')\Delta N\}$$
$$= \{y_1 + \Delta a\cdot\Delta N + (a_1-x')\Delta N\} - \{y_1 + (a_1-x')\Delta N\} = \Delta a\cdot\Delta N,$$
$$\beta = \{y_1 + \Delta a\cdot\Delta N + (a_1-x')\Delta N\} - \{y_1 + (a_1-x')\Delta N\} = \Delta a\cdot\Delta N,$$

which clearly indicates that $\alpha = \beta$ for every $x' \varepsilon X$. Also by Definition 2.8,

$$\gamma = \{y_1 + \Delta a\cdot\Delta N + (a_1+\Delta a-x')\Delta N\} - \{y_1 + (a_1+\Delta a-x')\Delta N\}$$
$$= \{y_1 + 2\Delta a\cdot\Delta N + (a_1-x')\Delta N\} - \{y_1 + \Delta a\cdot\Delta N + (a_1-x')\Delta N\} = \Delta a\cdot\Delta N.$$

Now consider an $x' \varepsilon X$ such that $(yy|y_1,a_1) < y_o$ and $(y|y_1,a_3) > y_o$. By Definition 2.8, we set $(y'|y_1,a_1) \equiv y_o$ which leads to $\alpha < \Delta a\cdot\Delta N$. Furthermore, $(y'|y_1,a_3) > y_o$ implies $(y'|y_2,a_3) > y_o$, hence $\gamma = \Delta a\cdot\Delta N$, and consequently $\gamma > \alpha$.

Now suppose that for some $x' \varepsilon X$, $(yy|y_1,a_1) < y_o$, $(yy|y_1,a_3) < y_o$, but $(y'|y_2,a_3) > y_o$. Again by Definition 2.8, we set $(y'|y_1,a_1) \equiv y_o$ and

$(y'|y_1,a_3) \equiv y_o$. This leads to $\alpha = 0$, $\gamma > 0$, and again $\gamma > \alpha$. It is apparent that for all other choices of $x' \varepsilon X$, $\gamma = \alpha$. Q.E.D.

As an addition to Definition 2.7 of T, let us introduce an auxiliary function.

Definition 3.1: An auxiliary function of the operator T is a mapping $\hat{zz} : A \times NN \to Z$, and such that for every $a(n) \varepsilon A(n)$ and $n \varepsilon NN$,

$$\hat{zz}(a(n),n) = \lim_{y\to\infty} zz(x(n+1),y,a(n),n).$$

In other words, $\hat{zz}(a(n),n)$ is the maximum flood level during time period $\Delta N(n)$ induced by the rate of release $a(n)$ from an *infinite reservoir*. Since the infinite reservoir can store inflow of any magnitude without being overtopped, $\hat{zz}$ does not depend on the inflow rate $x(n+1)$.

Lemma 3.2: (1) Assume that for every (x',z,a,n), $T(x',z,\cdot,a,n)$ is a continuous function of y, $zz(x',\cdot,a,n)$ is a decreasing and convex function of y, and that for every n, $\hat{zz}(\cdot,n)$ is a strictly increasing function of a. (2) For any $n \varepsilon NN$, $z \varepsilon Z$, and some $\Delta a > 0$ define: $y_1 = y_1$, $y_2 = y_1 + \Delta a\cdot\Delta N$, $y_3 = y_1 + 2\Delta a\cdot\Delta N$, and $a_1 = a_1$, $a_2 = a_1 + \Delta a$ such that $y_i \varepsilon Y$ and $a_j \varepsilon A(y_i,n)$ $i = 1, 2, 3$; $j = 1, 2$. (3) For every $x' \varepsilon X$ define:

$$\alpha = (z'|y_2,a_1) - (z'|y_1,a_1),$$
$$\beta = (z'|y_2,a_2) - (z'|y_1,a_2).$$

Then $E[\beta(X')] \geq E[\alpha(X')]$.

Proof: Fix x, z, and n. Let

$$W_j = \{x' : (z'|y_1,a_j) = \zeta_j,\ \zeta_j = \max\,[z,\hat{zz}(a_j,n)]\},\ j = 1,2$$

where ζ_j is the value of z' given the release rate a_j from an infinite reservoir. Thus, ζ_j is the infimum of z' for a reservoir with a finite storage space given state z, control a_j and control time n. The monotonicity property of $\hat{zz}$ and the fact $a_1 < a_2$ imply $\zeta_1 \le \zeta_2$ and $W_1 \subset W_2$. Define the following sets: $\hat{X}_1 = W_1$, $\hat{X}_2 = W_2 - W_1$, and $\hat{X}_3 = X - \hat{X}_1 - \hat{X}_2$. Observe that $\alpha(x') = 0$ for every $x' \in \hat{X}_1$, and $\beta(x') = 0$ for every $x' \in \hat{X}_1 \cup \hat{X}_2$. Consequently,

$$E[\beta(X') - \alpha(X')] = \sum_{k=1}^{3} \int_{\hat{X}_k} [\beta(x') - \alpha(x')]dF(x'|x,n)$$

can be reduced to

$$E[\beta(X') - \alpha(X')] = -\int_{\hat{X}_2} \alpha(x')dF(x'|x,n) + \int_{\hat{X}_3} [\beta(x') - \alpha(x')]dF(x'|x,n).$$

Due to Assumption (1) of the lemma, T is decreasing in y, which means that $\alpha(x') \le 0$ and $\beta(x') \le 0$ for every $x' \in X$. Hence, in the last equation, the first term is nonnegative.

Now it will be demonstrated that the second term is also nonnegative. Suppose that for every $x' \in \hat{X}_3$, $(z'|y_2,a_2) > \zeta_2$ which, along with the monotonicity property of T and the fact $\zeta_1 \le \zeta_2$, implies that also $(z'|y_1,a_2) > \zeta_1$. In this case, the release rate a_2 may be "substituted" by the release rate a_1 and an additional storage space equivalent to the release rate $\Delta a = a_2 - a_1$ during time period ΔN, without affecting the state of the system in the next control time. The equivalence relations are:

$$(z'|y_1,a_2) \equiv (z'|y_2,a_1),$$
$$(z'|y_2,a_2) \equiv (z'|y_3,a_1).$$

Consequently,

$$\alpha = (z'|y_2,a_1) - (z'|y_1,a_1)$$
$$\beta = (z'|y_3,a_1) - (z'|y_2,a_1) \quad \forall\ x' \in \hat{X}_3$$

where $y_2 - y_1 = y_3 - y_2 = \Delta a \cdot \Delta N$. By definition of $\hat{X}_3$, for every $x' \in \hat{X}_3$, $(z'|y_i,a_1) = zz(x',y_i,a_1,n)$, $i = 1, 2, 3$. Since $zz(x', \cdot, a_1, n)$ is decreasing and convex in y, $\beta \geq \alpha$. Hence, $\beta(x') - \alpha(x') \geq 0$ for every $x' \in \hat{X}_3$, and so the second term is also nonnegative.

Now suppose that for some $x' \in \hat{X}_3$, $(z'|y_2,a_2) = \zeta_2$. Then $(z'|y_2,a_2) \geq (z'|y_3,a_1)$ and

$$(z'|y_3,a_1) - (z'|y_2,a_1) \leq (z'|y_2,a_2) - (z'|y_2,a_1).$$

Consequently, the relationship $\beta \geq \alpha$ must also be true. Finally, observe that it is not necessary to consider the case $(z'|y_2,a_1) = \zeta_1$ because $\hat{X}_3$ has been defined such that $(z'|y_1,a_2) > \zeta_2$. By the equivalence statement introduced earlier, it must then be true $(z'|y_2,a_1) > \zeta_2 \geq \zeta_1$. Q.E.D.

<u>Theorem 3.1</u>: Assume the following:

(1) For every $y \in Y$ and $n \in NN$, $A(y,n) = [a_\alpha(n), a_\gamma(y,n)] \subset R^+$ where the function $a_\gamma : Y \times NN \to R^+$ is constant or decreasing on Y.

(2) $T(x',z,\cdot,a,n)$ and $\Psi(x',\cdot,a,n)$ are continuous functions of y for every $x' \in X$, $z \in Z$, $a \in A(y,n)$ and $n \in NN$.

(3) For every (x',a,n), $zz(x',\cdot,a,n)$ is a decreasing and convex function of y having further the property that

$$zz(x',y,a,n) \geq zz(x',y+\Delta a\cdot\Delta N,a-\Delta a,n)$$

for every $y \in Y$ and every positive increment Δa of a.

(4) u is increasing and twice differentiable on a compact set $Z \times Y$.

Then:

(1) For every (x,z,n), $S(x,z,\cdot,n)$ is decreasing on Y_α:

$$Y_\alpha = \{y : y \in Y,\ S(x,z,y,n) > a_\alpha\},$$

$$Y_\alpha^c = Y - Y_\alpha = \{y : y \in Y,\ S(x,z,y,n) = a_\alpha\}.$$

(2) For every (x,z,n), $R(x,z,\cdot,n)$ is decreasing on Y_β and increasing on Y_β^c:

$$Y_\beta^c = Y - Y_\beta,$$

$$Y_\beta = \{y : y_o \leq y \leq \min(y_e, y_\beta),\ y_\beta \in Y_\alpha^c,\ y_\beta \text{ minimizes } E[\frac{\partial u}{\partial z}\Delta z + \frac{\partial u}{\partial y}\Delta y]\},$$

where

$$\Delta y = (y'|y,a_\alpha) - (y'|y_\alpha,a_\alpha),$$

$$\Delta z = (z'|y,a_\alpha) - (z'|y_\alpha,a_\alpha),$$

and y_α is the least element of Y_α^c.

(3) For every (x,y,n), $R(x,\cdot,y,n)$ is increasing on Z.

Proof: The proof will proceed by induction on i, the number of control times remaining at time n.

(a) If $i = 1$ then $n = N$. For any fixed (x,z,N), let $a_1^* > a_\alpha(N)$ be the optimal control corresponding to some $y_1 \in Y$. To prove the conclusion (1) of the theorem we have to show that if $y_2 > y_1$ then the optimal control a_2^* corresponding to $y_2 \in Y$ satisfies $a_2^* \leq a_1^*$.

Let Δa be a small positive increment of a such that $\Delta a \leq a_1^* - a_\alpha(N)$. Define

$$a_2 = a_1^* - \Delta a,$$

$$y_2 = y_1 + \Delta a \cdot \Delta N,$$

so that $a_2 < a_1^*$ and $y_2 > y_1$, $y_2 \in Y$. Consider a transition $(y_1,a_1^*) \rightarrow (y_2,a_2)$. Definition 2.7 of the operator T and Assumption (3) of the theorem imply

$$\Delta z = (z'|y_2,a_2) - (z'|y_1,a_1^*) \leq 0 \quad \forall\ x' \in X,$$

while Definition 2.8 of the operator Ψ implies

$$\Delta y = (y'|y_2,a_2) - (y'|y_1,a_1^*) = 0 \quad \forall\ x' \in X.$$

Consequently, by Assumption (4) of the theorem

$$(u|y_2,a_2) \leq (u|y_1,a_1^*) \quad \forall\ x' \in X,$$

and we can assert that

$$E[u|y_2,a_2^*] \le E[u|y_2,a_2] \le E[u|y_1,a_1^*]$$

where the first inequality reflects optimality of the control a_2^* for the state y_2. This proves that $R(x,z,\cdot,N) = E[u|\cdot,a^*]$ is decreasing on Y_α.

To assert that $a_2^* \le a_1^*$ we must exclude the possibility $a_2^* > a_1^*$. We shall do this by showing that $E[u|y_2,\cdot]$ is increasing on $[a_1^*,a_\gamma(y_2,N)]$. Then, because $E[u|y_2,\cdot]$ is a continuous function of a (Yakowitz, 1969, p. 41), it must attain its minimum on $[a_\alpha,a_1^*]$. Let

$$\hat{a}_2 = a_1^* + \Delta a,$$

and such that $y_2 - \hat{a}_2 \cdot \Delta N \ \varepsilon\ Y$. Consider a transition $(y_1,a_1^*) \to (y_1,\hat{a}_2)$. From the Taylor's formula

$$(u|y_1,\hat{a}_2) = (u|y_1,a_1^*) + \frac{\partial u}{\partial z}\Delta z_1 + \frac{\partial u}{\partial y}\Delta y_1 \quad \forall\ x' \ \varepsilon\ X,$$

and

$$E[u|y_1,\hat{a}_2] = E[u|y_1,a_1^*] + E[\frac{\partial u}{\partial z}\Delta z_1 + \frac{\partial u}{\partial y}\Delta y_1]$$

where

$$\Delta z_1 = (z'|y_1,\hat{a}_2) - (z'|y_1,a_1^*),$$

$$\Delta y_1 = (y'|y_1,\hat{a}_2) - (y'|y_1,a_1^*) \ge 0 \quad \forall\ x' \ \varepsilon\ X.$$

Optimality of a_1^* for y_1 implies

$$E[\frac{\partial u}{\partial z}\,\Delta z_1] > -\,E[\frac{\partial u}{\partial y}\,\Delta y_1].$$

In a similar fashion, consider a transition $(y_1,a_1^*) \rightarrow (y_2,a_1^*)$:

$$E[u|y_2,a_1^*] = E[u|y_1,a_1^*] + E[\frac{\partial u}{\partial z}\,\Delta z_2 + \frac{\partial u}{\partial y}\,\Delta y_2]$$

where

$$\Delta z_2 = (z'|y_2,a_1^*) - (z'|y_1,a_1^*) \leq 0,$$

$$\Delta y_2 = (y'|y_2,a_1^*) - (y'|y_1,a_1^*) \geq 0 \quad \forall\ x' \in X.$$

Finally, for the transition $(y_1,a_1^*) \rightarrow (y_2,\hat{a}_2) \equiv (y_1,a_1^*) \rightarrow (y_1,\hat{a}_2) \rightarrow (y_2,\hat{a}_2)$:

$$E[u|y_2,\hat{a}_2] = E[u|y_1,a_1^*] + E[\frac{\partial u}{\partial z}(\Delta z_1+\Delta z_3) + \frac{\partial u}{\partial y}(\Delta y_1+\Delta y_3)]$$

where

$$\Delta z_3 = (z'|y_2,\hat{a}_2) - (z'|y_1,\hat{a}_2) \leq 0,$$

$$\Delta y_3 = (y'|y_2,\hat{a}_2) - (y'|y_1,\hat{a}_2) \geq 0 \quad \forall\ x' \in X.$$

Observe that according to Lemma 3.1, $\Delta y_1 = \Delta y_2$, and $\Delta y_3 \geq \Delta y_2$ for every $x' \in X$, while according to Lemma 3.2, $E[\Delta z_3] \geq E[\Delta z_2]$, but because $\frac{\partial u}{\partial z} \geq 0$ at every point of $Z \times Y$ also $E[\frac{\partial u}{\partial z}\,\Delta z_3] \geq E[\frac{\partial u}{\partial z}\Delta z_2]$. Using these results and the optimality condition for a_1^*, we can write

$$E[u|y_2,\hat{a}_2] > E[u|y_1,a_1^*] - E[\frac{\partial u}{\partial y} \Delta y_1] + E[\frac{\partial u}{\partial z} \Delta z_2] + E[\frac{\partial u}{\partial y} \Delta y_1] + E[\frac{\partial u}{\partial y} \Delta y_2],$$

and finally

$$E[u|y_2,\hat{a}_2] > E[u|y_1,a_1^*] + E[\frac{\partial u}{\partial z} \Delta z_2 + \frac{\partial u}{\partial y} \Delta y_2] = E[u|y_2,a_1^*]$$

which proves the conclusion (1) of the theorem for the control time N.

To prove that $R(x,z,\cdot,N)$ is increasing on Y_β^c suppose that $y_\alpha \in Y$ is the least value of y for which the optimal control is a_α. According to the result obtained above, for any $y > y_\alpha$, $a^* \leq a_\alpha$. However, due to the constraint imposed by the control set $A(y,N)$, a^* must not be less than a_α. Hence, the following inequality must be true

$$E[u|y,a_\alpha] \geq E[u|y,a^*] \quad \forall\ y > y_\alpha.$$

Now we must establish the relationship between $E[u|y_\alpha,a_\alpha]$ (which is the infimum for $R(x,z,\cdot,N)$ on Y_α) and $E[u|y,a_\alpha]$ for every $y \in Y_\alpha^c$. Consider a transition $(y_\alpha,a_\alpha) \to (y,a_\alpha)$:

$$E[u|y,a_\alpha] = E[u|y_\alpha,a_\alpha] + E[\frac{\partial u}{\partial z} \Delta z + \frac{\partial u}{\partial y} \Delta y] \quad \forall\ y \in Y_\alpha^c$$

where

$$\Delta z = (z'|y,a_\alpha) - (z'|y_\alpha,a_\alpha) \leq 0,$$
$$\Delta y = (y'|y,a_\alpha) - (y'|y_\alpha,a_\alpha) > 0 \quad \forall\ x' \in X.$$

Due to Assumption (2) of the theorem, Δz and Δy are continuous functions of y. Therefore, $E[\frac{\partial u}{\partial z} \Delta z + \frac{\partial u}{\partial y} \Delta y]$ is also continuous in y; hence it attains its minimum at some point y_β in the compact set Y_α^c. Furthermore, if $y \to \infty$ then $\Delta y \to \infty$ and $\Delta z \to -C$ $(C > 0)$. Thus, $E[\frac{\partial u}{\partial z} \Delta z + \frac{\partial u}{\partial y} \Delta y] \to \infty$ starting from 0 (for $y = y_\alpha$) through negative and positive values, in which case $y_\beta > y_\alpha$, or only through positive values, in which case $y_\beta = y_\alpha$. Due to the boundedness of Y, let $y_{\beta\beta} = \min\ (y_e, y_\beta)$. Accordingly, $R(x,z,\cdot,N) = E[u|\cdot,a_\alpha]$ is decreasing on $[y_\alpha, y_{\beta\beta}]$ and increasing on $(y_{\beta\beta}, y_e)$ which along with the result obtained earlier proves conclusion (2) of the theorem for the control time N.

To prove the third conclusion of the theorem, fix (x,y,N) and assume $z_2 > z_1$ $z_1, z_2 \in Z$. For an arbitrary $a \in A(y,N)$ define a set

$$\hat{X}_i = \{x' : z' = z_i,\ x' \in X\} \qquad i = 1, 2$$

whose complement

$$\hat{X}_i^c = X - \hat{X}_i = \{x' : z' = zz(x',y,a,N),\ z{=}z_i,\ x' \in X\} \quad i = 1,2.$$

Note that for every $a \in A(y,N)$, $\hat{X}_1 \subset \hat{X}_2$, and that for every $x' \in X$, $(y'|z_1,a) = (y'|z_2,a)$. Now, for any $a \in A(y,N)$

$$E[u|z_1,a] = \int_X [u(z',y')|z_1,a]dF(x'|x,N)$$

$$= \int_{\hat{X}_1} [u(z_1,y')|a]dF(x'|x,N) + \int_{\hat{X}_1^c} [u(z',y')|a]dF(x'|x,N),$$

and similarly

$$E[u|z_2,a] = \int_{\hat{X}_2} [u(z_2,y')|a]dF(x'|x,N) + \int_{\hat{X}_2^c} [u(z',y')|a]dF(x'|x,N).$$

Observe that in both equations the integrands of the second component are equal for every $x' \in X$ whereas the integrands of the first component, due to Assumption (4) of the theorem, satisfy

$$[u(z_2,y')|a] \geq [u(z_1,y')|a] \quad \forall\ x' \in X.$$

This along with the fact $\hat{X}_1 \subset \hat{X}_2$ assures that

$$E[u|z_2,a] \geq E[u|z_1,a].$$

Because this relation is true for an arbitrary $a \in A(y,N)$, it must also be true for the optimal control $a_2{}^*$ associated with the state z_2. Consequently,

$$E[u|z_2,a_2{}^*] \geq E[u|z_1,a_2{}^*] \geq E[u|z_1,a_1{}^*]$$

where the second inequality reflects optimality of the control $a_1{}^*$ for the state z_1. Hence $R(x,\cdot,y,N) = E[u|\cdot,a^*]$ is increasing on Z.

(b) Suppose that the theorem is true for $i < k$ and $n = N - k$. As proven in Yakowitz (1969, p. 41), the cost functional in the dynamic programming is continuous in all its variables. Thus, R is a continuous function of (x,z,y,n). From the fundamental theorem of calculus (see Hoffman, 1975, p. 140) and definition of R, it follows that if u is twice differentiable on Z x Y, then R is also twice differentiable on Z x Y for every $n \in NN$.

For any fixed (x,z,n), let $a_2 < a_1{}^*$ and $y_2 > y_1$ be defined as before, and let $(R'|y,a) = R[x',(z'|y,a),\ (y'|y,a),n+1]$. Because R' is increasing in z,

$$(R'|y_2,a_2) \leq (R'|y_1,a_1^*) \quad \forall\ x' \in X,$$

and we can assert that

$$E[R'|y_2,a_2^*] \leq E[R'|y_2,a_2] \leq E[R'|y_1,a_1^*],$$

which indicates that $R(x,z,\cdot,n) = E[R'|\cdot,a^*]$ is decreasing on Y_α.

To assert that $a_2^* \leq a_1^*$ we must exclude the possibility $a_2^* > a_1^*$. This can be done by showing that $E[R'|y_2,\cdot]$ is increasing on $[a_1^*,a_\gamma(y_2,n)]$. Then, because $E[R'|y_2,\cdot]$ is continuous function of a, it must attain its minimum on $[a_\alpha,a_1^*]$. A close analogous-reasoning to that of Part (a) of the proof leads to the statement

$$E[R'|y_2,\hat{a}_2] > E[R'|y_1,a_1^*] + E[\frac{\partial R'}{\partial z}\ \Delta z_2 + \frac{\partial R'}{\partial y}\ \Delta y_2] = E[R'|y_2,a_1^*]$$

which proves the conclusion (1) of the theorem.

To prove that $R(x,z,\cdot,n)$ is increasing on Y_β^c, as before, we must first establish the relationship between $E[R'|y_\alpha,a_\alpha]$ (which is the infimum for $R(x,z,\cdot,n)$ on Y_α) and $E[R'|y,a_\alpha]$ for every $y \in Y_\alpha^c$. The Taylor's formula for the transition $(y_\alpha,a_\alpha) \rightarrow (y,a_\alpha)$ gives:

$$E[R'|y,a_\alpha] = E[R'|y_\alpha,a_\alpha] + E[\frac{\partial R'}{\partial z}\ \Delta z + \frac{\partial R'}{\partial y}\ \Delta y] \quad \forall\ y \in Y_\alpha^c\ .$$

Because R' is twice differentiable on Z x Y, $\frac{\partial R'}{\partial z}$ and $\frac{\partial R'}{\partial y}$ are continuous functions of y; also Δz and Δy are continuous functions of y due to Assumption (2) of the theorem. Therefore, $E[\frac{\partial R'}{\partial z}\ \Delta z + \frac{\partial R'}{\partial y}\ \Delta y]$ is a continuous function of y and it attains its minimum at some point y_β in the compact set Y_α^c. From this point, an argument exactly the same as in Part (a) completes the proof of the conclusion (2).

The proof of the third conclusion proceeds also in exactly the same manner as in Part (a) of the proof. Fix (x,y,n) and assume $z_2 > z_1$ $z_1, z_2 \varepsilon Z$. With sets $\hat{X}_i$ and $\hat{X}_i^c$ (i = 1, 2) defined now for time n, we have for any a ε A(y,n)

$$E[R'|z_1,a] = \int_{\hat{X}_1} [R(x',z_1,y',n+1)|a]dF(x'|x,n) + \int_{\hat{X}_1^c} [R(x',z',y',n+1)|a] dF(x'|x,n)$$

$$E[R'|z_2,a] = \int_{\hat{X}_2} [R(x',z_2,y',n+1)|a]dF(x'|x,n) + \int_{\hat{X}_2^c} [R(x',z',y',n+1)|a] dF(x'|x,n).$$

A reasoning, the same as in Part (a), leads to the statement

$$E[R'|z_2,a_2*] \geq E[R'|z_1,a_2*] \geq E[R'|z_1,a_1*],$$

which is the definition of an increasing function. Ergo $R(x,\cdot,y,n) = E[R'|\cdot,a*]$ is increasing on Z. Q.E.D.

3.2. Optimization Algorithm

A direct consequence of Theorem 3.1 is an efficient optimization scheme which may significantly reduce the computational requirements of Algorithm 2.1. Let

$\Delta y > 0$ be an arbitrary increment of y,

$\Delta a > 0$ be an increment of a such that $\Delta a << \Delta y / \Delta N$,

$Y = [y_o, y_e]$,

$A(y,n) = [a_\alpha(n), a_\gamma(y,n)] \quad \forall\ y \in Y,\ n \in NN$.

Algorithm 3.1: For any fixed (x,z,n) $x \in X$, $z \in Z$, $n \in NN$, the function $S(x,z,\cdot,n)$ of y can be found as follows:

(a) $y \leftarrow y_o$

find a*

$S(x,z,y,n) \leftarrow a^*$

(b) $y \leftarrow y + \Delta y$, if $y > y_e$ then STOP

$a_1 \leftarrow a^*$

$a_2 \leftarrow a^* - \Delta y / \Delta N$, if $a_2 < a_\alpha$ then $a_2 \leftarrow a_\alpha$

$a_3 \leftarrow a_2 + \Delta a$

(c) compute $R_2 = R(x,z,y,n) \mid a_2$, $R_3 = R(x,z,y,n) \mid a_3$

if $R_2 > R_3$ then continue search on $(a_2, a_1]$

if $R_2 \leq R_3$ then continue search on $[a_\alpha, a_2]$

(d) find a*

$S(x,z,y,n) \leftarrow a^*$

(e) if $a^* > a_\alpha$ then go to (b)

if $a^* = a_\alpha$ then $S(x,z,v,n) = a_\alpha \quad \forall\ v \in (y, y_o]$

STOP.

Another reduction of the computational requirements of Algorithm 2.1 can be achieved by noting that $\{z(i)\}_{i=1}^N$ is nondecreasing. Hence for every $n > 1$, the space Z can be bounded from below by the initial condition z(1).

3.3. Case of a Categorical Forecast

It is instructive to investigate a property of the optimal policy for the case of a categorical forecast, i.e., a forecast which for each $n \in NN$ specifies $x(n+1)$ with probability one.

Theorem 3.2: If in a reservoir CP

(1) $\theta(n) = ZxY$ for every n,

(2) $A(y,n) = A$ is independent of (y,n),

(3) $z(n+1) = \max\{z(n), zz(a(n))\}$ where zz is an increasing function of a,

(4) $\Phi = \{x(n+1) : n \in NN\}$,

(5) $u(\cdot,y)$ is increasing on a compact set Z,

then the optimal policy is stationary.

Proof: For n = N,

$$R(z,y,N|x(N+1)) = \min_{a(N) \in A} u[z(N+1),y(N+1)]$$

and by recursion, for any $n \in NN$,

$$R(z,y,n|x(N+1), \ldots, x(n+1)) = \min_{a(n) \in A} \ldots \min_{a(N) \in A} u[z(N+1),y(N+1)].$$

By Definition 2.7 of T

$$u[z(N+1),y(N+1)] = u[\max\{z(n),zz(n),\ldots, zz(N)\},y(N+1)]$$

and by Assumption (3) of the theorem

$$u[z(N+1),y(N+1)] = u[zz(\max\{a(n-1), a(n), \ldots, a(N)\}),y(N+1)].$$

From Definition 2.8 of Ψ

$$y(N+1) = y(n) + \sum_{i=n}^{N} [a(i) - x(i+1)] \Delta N(i),$$

which implies that y(N+1) depends on the policy $\bar{a}$ only through $\sum_{i=n}^{N} a(i)$. Thus the problem

$$\min_{a(n) \varepsilon A} \cdots \min_{a(N) \varepsilon A} u[zz(\max\{a(n-1), a(n), \ldots, a(N)\}), f(\sum_{i=n}^{N} a(i))],$$

for any fixed y(N+1) implying $\sum_{i=n}^{N} a(i) = C$, can be reduced by virtue of Assumption (5) to the problem

$$\min_{a(n) \varepsilon A} \cdots \min_{a(N) \varepsilon A} [\max\{a(n-1), a(n), \ldots, a(N)\}]$$

subject to

$$\sum_{i=n}^{N} a(i) = C$$

whose obvious solution is a(n) = ... = a(N) = a. Q.E.D.

CHAPTER 4

SUBOPTIMAL STRATEGIES

From a practical point of view, one may often wish to consider reduction of the computational requirements at the expense of optimality of the control. This approach motivates this chapter in which two suboptimal control strategies are presented. The first one is developed for a partial OLFC with the feedback mechanism defined on the state space reduced to two dimensions. The second suboptimal strategy is also based on a partial OLFC but with only one state variable remaining in the feedback; the second state variable is further eliminated through an application of a certainty equivalent (naive feedback).

4.1. Partial Open-Loop Strategy

In the triplet (x,z,y) the state variable whose measurements are usually the most likely always to be available to the controller is y. Therefore, the potential candidates for elimination from the feedback are x and z. Note that if the storage operator Ψ is deterministic and its inverse Ψ_x^{-1} with respect to x exists, then at any time n ε NN, x(n) can be uniquely determined from the measurements for y(n-1), y(n), and a(n-1) by Ψ_x^{-1}. Hence, the trajectory $\{y(i)\}_{i \leq n}$ and the policy $\{a(i)\}_{i \leq n}$ imply the trajectory $\{x(i)\}_{i \leq n}$ for every n ε NN. Furthermore, x(n) does not appear as an explicit variable in the operators T and Ψ. Under these circumstances, one can intuitively expect that elimination of x from the feedback will cause the least loss of information. A suboptimal strategy S will thus be developed for the partial OLFC (Bertsekas, 1976, Chapter 5) with the states remaining in the feedback being z and y (Figure 4.1). Specifically, S : Z x Y x NN → A, and it can be computed from the following algorithm.

Algorithm 4.1:

(a) For every n ε NN and (z,y) ε θ(n), let

$$\Theta(z,y,n) = \{a : a \in A(y,n) \text{ and } P[(z',y') \in \theta(n+1) \mid z,y,a,n] = 1, \ \forall\, x \in X\}.$$

(b) For every n = 2, ..., N, the marginal distribution function $F(x'|n) = F(x(n+1)|n)$ is determined from the forecast

$$\Phi = \{F(x(2)|1),\ F(x(3)|x(2),2),\ \ldots,\ F(x(N+1)|x(N),N)\}$$

by the recursive equation

$$F(x(n+1)|n) = \int_X F(x(n+1)|x(n),n)\,dF(x(n)|n-1).$$

(c) S(z,y,n) = a* where for every state-time (z,y,n), a* is a solution to the recursive equations:

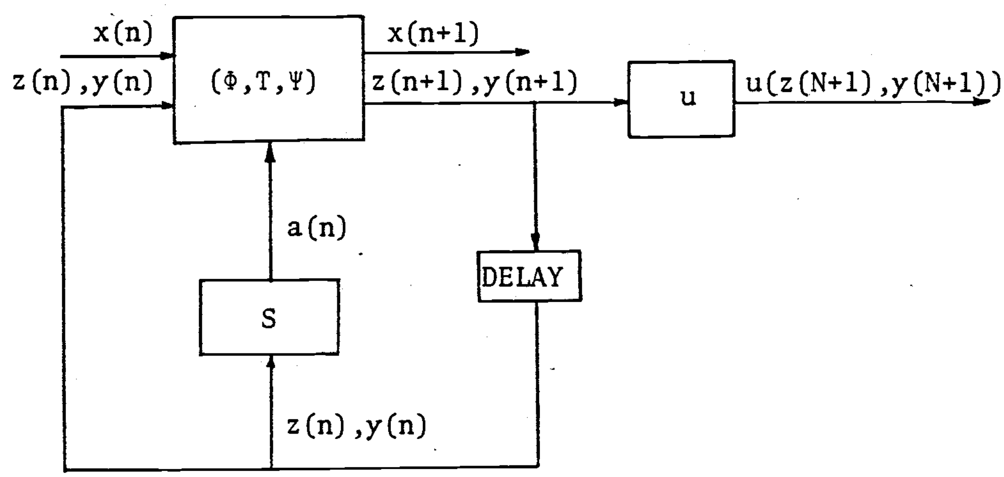

Figure 4.1. Partial Open-Loop Feedback Control

$$R(z,y,N) = \min_{a \varepsilon \Theta(z,y,N)} \int_X u(T(x',z,y,a,N),\Psi(x',y,a,N))\,dF(x'|N),$$

$$R(z,y,n) = \min_{a \varepsilon \Theta(z,y,n)} \int_X R(T(x',z,y,a,n),\Psi(x',y,a,n),n+1)\,dF(x'|n),$$

$$n = 1, \ldots, N-1.$$

(d) With an initial condition (z(1),y(1)) at the forecast time $m \varepsilon MM$, the expected disutility associated with the suboptimal strategy $S_m = S$ for the period $[t_m, t_m + \lambda_m]$ is $R_m = R(z(1),y(1),1)$.

4.2. Naive/Partial Open-Loop Strategy

The second suboptimal strategy S is developed for a partial OLFC wherein only y remains in the feedback, that is $S : Y \times NN \rightarrow A$. As has been shown in the previous section, elimination of x from the feedback does not present any problem. On the contrary, elimination of z is much more difficult. Recall that the objective of the control process is to find a strategy S which minimizes $E[u(Z(N+1),Y(N+1))]$. The terminal state z(N+1) can be expressed as a composite function $f = (T)^N$ such that

$$z(N+1) = f[\bar{x},\bar{y},S,z(1)].$$

It is this dependence of z(N+1) upon the trajectory $\bar{y}$ that causes the problem. Obviously, $\bar{y}$ cannot be further eliminated through the use of Ψ inasmuch as this step would negate our intention of keeping y as a partial feedback. We are thus facing a necessity of relating somehow

$z(N+1)$ to $\bar{y}$ and S. The problem will be resolved by exploiting some of the properties of the basic reservoir CP model and by developing a suitable *certainty equivalent* for the variate $Z(N+1)$. Substitution of a certainty equivalent for a state variable is also called the *naive feedback.*

<u>Certainty Equivalent for Z(N+1)</u>: The certainty equivalent for $Z(N+1)$ is sought in the form of a function $\zeta: Y \to Z$ associated with the suboptimal strategy S. Of course, prior to solving the control problem, S is not known. However, the structure of the reservoir CP allows for constructing ζ without explicitly referring to the strategy S.

Assume that $u(\cdot,y)$ is increasing on a compact set Z, Ψ is deterministic, and the inverse Ψ_a^{-1} with respect to a and the inverse Ψ_y^{-1} with respect to y exist so that

$$a(n) = \Psi_a^{-1}\,(x(n+1),y(n),y(n+1),n),$$

$$y(n) = \Psi_y^{-1}\,(x(n+1),y(n+1),a(n),n) \qquad \forall\ n \in NN.$$

Suppose $y(n)$ is the known terminal state of the CP which ends at time $n \in NN$. Now, due to the above assumptions and Definition 2.7 of T, minimization of $u(z(n),y(n))$ is equivalent to the direct minimization of $z(n)$. Inasmuch as $z(n)$ is a random variable, the problem is to find

$$\min E[u(Z(n),y(n))].$$

According to the conception of the naive feedback control (Bertsekas, 1976, p. 193), the above value is substituted by

$$u(\min E[Z(n)], y(n)).$$

Definition 4.1: The *certainty equivalent*, $\zeta(y,n)$ for $Z(n)$ ($n \in NN$) is the minimum expected value of $Z(n)$ when the terminal state at time n is y.

The function $\zeta : Y \times NN \to Z$ can be constructed recursively by a forward dynamic programming algorithm. The last step of this algorithm yields the desired certainty equivalent function $\zeta(\cdot, N+1)$ for $Z(N+1)$.

Algorithm 4.2:

(a) For every $n = 2, \ldots, N+1$ and $y(n) \in \theta(n)$ let

$\Theta(y,n) = \{a : a \in A(n-1) \text{ and } P[y(n-1) \in \theta(n-1) | y,a,n] = 1, \ \forall x \in X\}$.

(b) Given the initial condition $(z(1), y(1))$, for every $y(2) \in Y$, compute

$$\zeta(y(2),2) = \max \{E[zz(X(2),y(1),A(1),1)], z(1)\}$$

where

$$A(1) = \Psi_a^{-1} [X(2),y(1),y(2),2].$$

(c) For every state-time $(y(n),n)$, $n = 3, \ldots, N+1$, find

$$\zeta(y(n),n) = \min_{a(n-1) \in \Theta(y(n),n)} \max\{E[zz(X(n),Y(n-1),a(n-1),n-1)], E[\zeta(Y(n-1),n-1)]\}$$

where

$$Y(n-1) = \Psi_y^{-1}(X(n), y(n), a(n-1), n-1).$$

Everywhere the expectations are taken with respect to X(n) having distribution $F(\cdot|n)$ as defined in Step (b) of the Algorithm 4.1.

Once the certainty equivalent function $\zeta(\cdot, N+1)$ has been obtained, the suboptimal strategy S can be developed for the partial OLFC with the state variable y. Specifically, $S : Y \times NN \to A$, and it can be computed from the following algorithm.

Algorithm 4.3:

(a) For every $n \in NN$ and $(\cdot, y) \in \theta(n)$, let

$$\Theta(y,n) = \{a : a \in A(n) \text{ and } P[(\cdot,y') \in \theta(n+1) | (\cdot,y),a,n] = 1, \forall x \in X\}.$$

(b) $S(y,n) = a^*$ where for every state-time (y,n), a^* is a solution to the recursive equations:

$$R(y,N) = \min_{a \in \Theta(y,N)} \int_X u[\zeta(\Psi(x',y,a,N),N+1),\Psi(x',y,a,N)]\,dF(x'|N)$$

$$R(y,n) = \min_{a \in \Theta(y,n)} \int_X R[\Psi(x',y,a,n),n+1]\,dF(x'|n), \qquad n = 1, \ldots, N-1.$$

(c) With an initial condition y(1) at the forecast time $m \in MM$, the expected disutility associated with the suboptimal strategy $S_m = S$ for the period $[t_m, t_m + \lambda_m]$ is $R_m = R(y(1),1)$.

CHAPTER 5

COMPUTATIONAL ASPECTS

In general, the computational requirements of a dynamic programming algorithm grow exponentially with the dimension of the state space. High computational requirements may substantially reduce the extent of implementability of the control model, particularly in real-time control wherein the optimal strategy has to be computed repeatedly and always under pressure of shrinking time. There is, therefore, a definite need for a computationally efficient algorithm. In the preceding two chapters, the structure of the control model was explored to establish analytic properties of the optimal strategy. Stronger analytic results could be sought under more restrictive assumptions about the model components. This is a task for further research.

This chapter is designed to summarize the computational aspects of a dynamic programming algorithm. A discretization procedure which allows obtaining a numerical solution with desired error bounds is presented, and the computational requirements are derived.

5.1. Discretization Procedure

For numerical computations, the infinite state space and control set are partitioned into a finite number of subsets. Consequently, the functional, R, of the dynamic programming algorithm as well as the optimal strategy, S, are approximated by piecewise constant functions.

Let each of the sets X, Z, and Y be partitioned into finite mutually disjoint sets such that

$$X = \bigcup_{i=1}^{I_X} X^i, \quad x^i \in X^i, \quad i = 1, \ldots, I_X,$$

$$Z = \bigcup_{j=1}^{I_Z} Z^j, \quad z^j \in Z^j, \quad j = 1, \ldots, I_Z,$$

$$Y = \bigcup_{k=1}^{I_Y} Y^k, \quad y^k \in Y^k, \quad k = 1, \ldots, I_Y.$$

Algorithm 2.1 of the dynamic programming is approximated by an algorithm defined on the finite grids G_X, G_Z, G_Y and G_A, where

$$G_X = \{x^i : i = 1, \ldots, I_X\},$$
$$G_Z = \{z^j : j = 1, \ldots, I_Z\},$$
$$G_Y = \{y^k : k = 1, \ldots, I_Y\},$$
$$G_A = \{a^\ell : \ell = 1, \ldots, I_A\} \subset A.$$

It is assumed that

$$\Theta(x^i, z^j, y^k, n) \cap G_A \neq \emptyset \quad \forall\ i = 1, \ldots, I_X,\ j = 1, \ldots, I_Z,\ k = 1, \ldots, I_Y,\ n \in \mathbb{N}.$$

<u>Algorithm 5.1</u>: An approximation to Algorithm 2.1.

(a)

$$\hat{R}(x,z,y,N) = \begin{cases} \min\limits_{a\varepsilon\Theta(x,z,y,N)\cap G_A} E[u(Z',Y')] & \text{if } x \varepsilon G_X,\ z \varepsilon G_Z,\ y \varepsilon G_Y, \\ \hat{R}(x^i,z^j,y^k,N) & \text{if } x \varepsilon X^i,\ z \varepsilon Z^j,\ y \varepsilon Y^k, \\ & i = 1, \ldots, I_X,\ j = 1, \ldots, I_Z,\ k = 1, \ldots, I_Y. \end{cases}$$

(b)

$$\hat{R}(x,z,y,n) = \begin{cases} \min\limits_{a\ \varepsilon\Theta(x,z,y,n)\cap G_A} E[\hat{R}(X',Z',Y',n+1)] & \text{if } x \varepsilon G_X,\ z \varepsilon G_Z, \\ & y \varepsilon G_Y, \\ \hat{R}(x^i,z^j,y^k,n) & \text{if } x \varepsilon X^i,\ z \varepsilon Z^j,\ y \varepsilon Y^k,\ i = 1, \ldots, I_X, \\ & j = 1, \ldots, I_Z,\ k = 1, \ldots, I_Y,\ n = 1, \ldots, N-1. \end{cases}$$

(c) An approximation to the optimal strategy obtained from the above procedure is a function $\check{S} : G_X \times G_Z \times G_Y \times NN \to G_A$. The definition of $\check{S}$ is extended on the whole state space by constructing a piecewise constant function:

$$\hat{S}(x,z,y,n) = \begin{cases} \check{S}(x,z,y,n) & \text{if } x \varepsilon G_X,\ z \varepsilon G_Z,\ y \varepsilon G_Y, \\ \check{S}(x^i,z^j,y^k,n) & \text{if } x \varepsilon X^i,\ z \varepsilon Z^j,\ y \varepsilon Y^k,\ i = 1, \ldots, I_X, \\ & j = 1, \ldots, I_Z,\ k = 1, \ldots, I_Y,\ \forall\ n \varepsilon NN. \end{cases}$$

The convergence of $\hat{R}$ to R is proved in Bertsekas (1976, p. 188) under certain continuity, compactness, and Lipschitz assumptions. For the reservoir CP problem discretized as above, the theorem reads:

Theorem 5.1: There exist positive constants $\{C(n) : n \in NN\}$ (independent of grids G_X, G_Z, G_Y, G_A) such that

$$|R(x,z,y,n) - \hat{R}(x,z,y,n)| \leq C(n)[d_\Omega + d_A] \;\forall (x,z,y) \in \Omega,\; n \in NN$$

where

$$d_\Omega = \max_{\substack{i=1,\ldots,I_X \\ j=1,\ldots,I_Z \\ k=1,\ldots,I_Y}} \; \sup_{\substack{x \in X^i \\ z \in Z^j \\ y \in Y^k}} \; ||(x,z,y) - (x^i,z^j,y^k)||,$$

$$d_A = \max_{n \in NN} \; \max_{\substack{i=1,\ldots,I_X \\ j=1,\ldots,I_Z \\ k=1,\ldots,I_Y}} \; \max_{a \in \Theta(x^i,z^j,y^k,n)} \; \min_{a' \in \Theta(x^i,z^j,y^k,n) \cap G_A} ||a-a'||$$

and $||\cdot||$ denotes Euclidean norm.

5.2. Computational Requirements

The optimal control strategy S* constructed by Algorithm 2.1 can be compared with any of the suboptimal strategies S in terms of closeness of approximation of S* by S and in terms of computational requirements which are always higher for S* than for any S. The choice of the control mode to be implemented in real-time is then a matter of a trade-off between accuracy and cost (time) of computations. Inasmuch as the tight bounds on the performance of a suboptimal strategy are very hard to get (Bertsekas, 1976, p. 215), extensive computer experimentations are inevitable for each particular application case. Comparison of computational requirements is, however, straightforward and is presented below.

Assume that each strategy to be compared is constructed numerically according to the approximation given by Algorithm 5.1, and that monotonicity property revealed in Section 3.1 is not exploited during optimization. Let q denote the total number of comparisons required to construct a strategy, and let

$$I_\Theta(x^i, z^j, y^k, n) \equiv [\text{\# of elements in } \Theta(x^i, z^j, y^k, n) \cap G_A],$$

$$I_\Theta(z^j, y^k, n) \equiv [\text{\# of elements in } \Theta(z^j, y^k, n) \cap G_A],$$

$$I_\Theta(y^k, n) \equiv [\text{\# of elements in } \Theta(y^k, n) \cap G_A],$$

$$i = 1, \ldots, I_X, \; j = 1, \ldots, I_Z, \; k = 1, \ldots, I_Y.$$

For the optimal strategy generated by Algorithm 2.1

$$q^* = \sum_{n=1}^{N} \sum_{i=1}^{I_X} \sum_{j=1}^{I_Z} \sum_{k=1}^{I_Y} I_\Theta(x^i, z^j, y^k, n),$$

and noting that $I_\Theta(x^i, z^j, y^k, n) \leq I_A$ for every (i,j,k,n), we obtain upper bound on q^*:

$$q^* \leq N \cdot I_X \cdot I_Z \cdot I_Y \cdot I_A.$$

Similarly for the partial open-loop strategy (Algorithm 4.1)

$$q_p = \sum_{n=1}^{N} \sum_{j=1}^{I_Z} \sum_{k=1}^{I_Y} I_\Theta(z^j, y^k, n),$$

$$q_p \leq N \cdot I_Z \cdot I_Y \cdot I_A,$$

and for the naive/partial open-loop strategy (Algorithm 4.2 and 4.3)

$$q_n = I_Y + 2 \sum_{n=1}^{N} \sum_{k=1}^{I_Y} I_\Theta(y^k, n),$$

$$q_n \leq I_Y + 2NI_Y I_A.$$

As expected, $q^* > q_p > q_n$, and there is roughly an order of magnitude reduction in the value of q for each state variable eliminated from the feedback loop.

REFERENCES

Bertsekas, D.P., Dynamic Programming and Stochastic Control, Academic Press, New York, 1976.

Hoffman, K., Analysis in Euclidean Space, Prentice-Hall, Englewood Cliffs, 1975.

Krzysztofowicz, R., Preference Criterion and Group Utility Model for Reservoir Control Under Uncertainty, Technical Reports on Natural Resource Systems, No. 30, Department of Hydrology and Water Resources, University of Arizona, Tucson, Arizona, March, 1978.

Lorenz, E.N., Predictability and Periodicity: A Review and Extension, presented at the Third Conference on Probability and Statistics in Atmospheric Science, Boulder, Colorado, June 1973.

von Neumann, J. and Morgenstern, O., Theory of Games and Economic Behavior, 2nd edition, Princeton University Press, Princeton, New Jersey, 1947.

Yakowitz, S., Mathematics of Adaptive Control Processes, American Elsevier, New York, 1969.