OPTIMAL MIX OF ADJUSTMENTS TO FLOODS

by

T. L. Morin
W. L. Meier, Jr.
K. S. Nagaraj

June 1981

PURDUE UNIVERSITY
WATER RESOURCES RESEARCH CENTER
WEST LAFAYETTE, INDIANA
OPTIMAL MIX OF ADJUSTMENTS TO FLOODS

by

T.L. Morin
W.L. Meier, Jr.
K.S. Nagaraj

School of Industrial Engineering
Purdue University
West Lafayette, Indiana 47907

This is a completion report for the project entitled "Optimal Mix of Adjustments to Floods"

The work upon which this report is based was supported by funds provided by the United States Department of the Interior, Office of Water Resources Research, as authorized by the Water Resources Research Act of 1964 (P.L. 88-379) and is part of the Title I Project B-109-IND, Agreement No. 14-34-0001-9071.

Purdue University Water Resources Research Center
Technical Report Number 139
June, 1981
The contents of this publication do not necessarily reflect the view and policies of the Office of Water Research and Technology, U.S. Department of Interior, nor does mention of trade names or commercial products constitute their endorsement or recommendation for use by the U.S. government.
ACKNOWLEDGMENTS

This report includes the work done under Matching Grant Fund OWRT-B-'09-IND and was supported in part by the Office of Water Research and Technology, Department of the Interior, and in part by Purdue University. The authors wish to express their appreciation to Dr. Daniel Wiersma, Director of the Purdue University Water Resources Research Center and to Dr. Wilbur L. Meier, Jr., Head of the School of Industrial Engineering at Purdue University for their assistance in the administration of the project.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF FIGURES</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vii</td>
</tr>
<tr>
<td>CHAPTER I - INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER II - THE FLOOD CONTROL PROBLEM</td>
<td>6</td>
</tr>
<tr>
<td>The Framework</td>
<td>6</td>
</tr>
<tr>
<td>Review of Relevant Literature</td>
<td>11</td>
</tr>
<tr>
<td>CHAPTER III - MODEL AND SOLUTION APPROACH</td>
<td>19</td>
</tr>
<tr>
<td>Model Formulation</td>
<td>19</td>
</tr>
<tr>
<td>Applicability of Model</td>
<td>27</td>
</tr>
<tr>
<td>Solution Algorithm</td>
<td>29</td>
</tr>
<tr>
<td>CHAPTER IV - COMPUTATIONAL EFFICIENCY</td>
<td>47</td>
</tr>
<tr>
<td>Notation</td>
<td>50</td>
</tr>
<tr>
<td>The &quot;Sieve&quot;</td>
<td>53</td>
</tr>
<tr>
<td>A Posteriori results</td>
<td>56</td>
</tr>
<tr>
<td>Comparison of results on bounds</td>
<td>60</td>
</tr>
<tr>
<td>&quot;Sieve&quot; strategies</td>
<td>64</td>
</tr>
<tr>
<td>CHAPTER V - DISCUSSION</td>
<td>71</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>75</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>85</td>
</tr>
<tr>
<td>Appendix A</td>
<td>85</td>
</tr>
<tr>
<td>Appendix B</td>
<td>91</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>Trends and limits of adjustments to floods</td>
</tr>
<tr>
<td>2</td>
<td>Matrix of interaction of adjustments to floods</td>
</tr>
<tr>
<td>3</td>
<td>A &quot;Sieve&quot;</td>
</tr>
<tr>
<td>4</td>
<td>A hypothetical flood plain</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table                                                  Page

1  Comparison of the number of states required by
two different dynamic programming approaches for a
particular flood control problem ....................  87
ABSTRACT

Floods are the most widespread geophysical hazard in the United States and they account for greater average annual property losses than any other single geophysical hazard. Moreover, despite substantial expenditures on flood protection structures, flood damages continue to increase. The structural measures for flood control include levees, dams, and channel improvements. However, it has been observed that structural measures often provide a false sense of security to existing and potential floodplain occupants and, as such, have resulted in increased flood damages, contrary to their intended purpose. This realization among others has led to an increasing awareness and interest in the role of nonstructural measures such as floodplain zoning, land use allocation, insurance, and warning as an important and integral part of any overall flood damage mitigation program. However, determining an "optimal" mix of adjustments is very difficult as a consequence of both the interdependence between the structural and nonstructural measures and the multitude of feasible combinations of structural and nonstructural measures which must be considered over time and space. This report addresses the issue of developing a general methodology for the solution of the optimal mix of adjustments problem. We recognize that the overall problem has an underlying sequencing nature which can
be exploited in order to use a dynamic programming solution approach which bears distinct similarities to the sequencing approaches used in electric generation planning problems. Algorithmic details and computational refinements are also discussed. In particular, we develop an efficient approach for modifying the exact algorithm so as to provide solutions of any desired degree of optimality with reasonable computational effort using some newly developed results on a posteriori analysis. Data requirements and implementation details are also discussed and details of an application to a real-world problem are given.
CHAPTER I

INTRODUCTION

Floods are the most widespread geophysical hazard in the United States and they account for greater average annual property losses than any other single geophysical hazard [W5]. Moreover, despite substantial expenditures for flood control measures, flood damages continue to increase [U7]. Specifically, the total annual national flood damages have been increasing by about 4 percent annually in real dollars during this century and there are indications that this rate has accelerated to the 6 to 7 percent range during the last decade [N4]. The dollar value of these losses is truly staggering -- the total annual flood damages were $3.4 billion in 1975 and it has been estimated [N3] that even with improved flood stream management, damages will exceed $4.3 billion (measured in 1975 dollars) by the year 2000. Without such improvements in flood plain management, the damages could approach $6 billion in 1975 dollars by the year 2000. In Indiana alone the annual flood damages for 1980 were $128 million measured in 1978 dollars [S6].

The magnitude of problem prompted Congress to instruct the National Science Foundation to conduct a flood hazard mitigation study during fiscal year 1980 (House Report 96:91). NSF concluded that [N4, p.1],
"Innovative approaches and increased attention to flood problems nationwide are required if the United States is to arrest, much less reverse, rising flood losses and the social and economic burden they place on the people and the nation's tax-supported flood-relief institutions". This report discusses one such innovative approach -- the determination of an optimal mix of adjustments to floods.

The mix of adjustments to floods involves both structural and nonstructural measures. The structural (protective) measures for flood control typically include levees, floodwalls, channel improvements, and storage reservoirs. The nonstructural measures for flood control typically include land use control and management (i.e., floodplain zoning, outright purchase of portions of the floodplain, and land use conversion), flood proofing, warning and evacuation, relief and rehabilitation, and flood insurance. A recent analysis [W5, p. 103] of the net benefits of various measures as a function of the magnitude of the catastrophe is presented graphically in Figure 1 -- see also [W6].

Reliance has been historically placed primarily on structural measures for flood control. However, it has been observed that structural measures often provide a false sense of security to existing and potential floodplain occupants and, as such, may actually result in increased flood damages, contrary to their intended purpose. That is, potential benefits from structural flood control measures are often lost through subsequent unwise development in the areas presumed to be protected [N4]. These realizations among others, have led to an increasing awareness and interest in the role of nonstructural measures as an important and integral part of any overall flood damage mitigation
Figure 1. Trends and Limits of Adjustments to Floods. [W5]
program -- for example see [A5, H1, W6]. Furthermore, the age-old hope for complete protection from floods has given way to the realization that a more realistic and viable goal is the mitigation of flood damages [N4]. In fact, the NSF study concluded [N4, p. 214] that, "Flood hazard mitigation strategies can be effective only if they reflect mixes of alternative structural and nonstructural approaches appropriate to the circumstances. Much more work needs to be done on improving mitigation of planning for different aspects of flood hazard mitigation". This is precisely the issue addressed in this report.

The report is organized in the following manner. In Chapter II we review the relevant literature on previous approaches to flood control planning and related problems and outline our approach to the problem. The lack of viable planning methodologies is a consequence of the highly complicated nature of the problem which results both from the interdependence between the structural and nonstructural measures and from the multitude of feasible combinations of structural and nonstructural measures which must be considered over time and space. These factors have contributed to the fact that, although planning methods for problems involving only structural [M6] and nonstructural [H4] measures have been developed, the optimum mix of adjustment problem has yet to be satisfactorily resolved. We present a general formulation of the flood damage reduction problem in Chapter III. In Chapter III we also develop a suitable algorithmic approach to be used along with our general formulation methodology. Algorithmic details and computational refinements are discussed in Chapter IV. In particular, we develop a novel approach to modify the exact algorithm so as to efficiently obtain
solutions of any desired degree of optimality using a posteriori optimality analysis. The report concludes with a summary and recommendations for future research in Chapter V.
CHAPTER II

THE FLOOD CONTROL PROBLEM

In this Chapter, we will first present a framework for the optimal mix of adjustments problem and then review and critically evaluate previous attempts at modeling the Flood Control Problem (FCP) relative to our framework. The plan of this chapter is as follows. In Section 2.1 we present our framework for the Flood Control Problem. We follow this with an overview of various relevant models in Section 2.2 and observe that none of these models are directly applicable to the FCP. In Chapter III we consider modeling procedures for similar problems, evaluate their relevance to our framework, and, finally, recommend an adaptation of Erlenkotter and Rogers' [E6] sequencing approach for the FCP.

2.1 The Framework

Models of the Flood Control Problem cater to a variety of needs, have different emphases and considerations, and, hence are not comparable with regards to their comparative strengths and weaknesses. Furthermore, some of the models do not actually clarify precisely their audience and, thus, their applicability is in doubt. Moreover, most of the models have been derived with a particular application in mind and are therefore not readily adaptable to general models. Our approach is
to first present and motivate a model of the Flood Control Problem. Specifically, we first propose a general scenario of the FCP and then present a model which is suitable and adaptable to any particular application. The Decision Maker (DM) is typically a local government or legislative body who is interested in allocating funds so as to minimize the net annual damages incurred due to recurrent floods. Despite the fact that some of the funds may be allocated to measures which have benefits besides damage reduction, we shall restrict our considerations solely to damage reduction benefits of such measures. For situations in which this assumption is not reasonable, we refer to the multi-criteria formulations suggested in the recommendations of Chapter V. We essentially are envisioning those cases where there are significant annual flood damages due to flash floods etc. and are looking at ways to minimize these damages.

The various structural and nonstructural measures that the Decision Maker is considering allocating funds to are typically:

<table>
<thead>
<tr>
<th>Structural</th>
<th>Non Structural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leveses</td>
<td>Land Use Allocation</td>
</tr>
<tr>
<td>Reservoirs</td>
<td>Flood Plain Zoning</td>
</tr>
<tr>
<td>Channel Improvements</td>
<td>Flood Proofing</td>
</tr>
<tr>
<td>Floodwalls</td>
<td>Insurance</td>
</tr>
<tr>
<td></td>
<td>Warning</td>
</tr>
<tr>
<td></td>
<td>Temporary Evacuation</td>
</tr>
<tr>
<td></td>
<td>Emergency Procedures</td>
</tr>
<tr>
<td></td>
<td>Relief and Rehabilitation</td>
</tr>
</tbody>
</table>

For a good discussion of these measures, the interested reader is referred to White and Haas [W6].

An important consideration in the evaluation of any one measure is that the measure's damage reduction capability is dependent on the
presence and levels of the other measures and hence cannot be chosen independently. For example it is fairly obvious that the effectiveness of any nonstructural measure depends on the mix of structural measures used. We assume that the potential scales and locations of the structural alternatives have been specified and the DM then decides which of them are to be employed and when they shall be built. We also assume that the DM decides which of the nonstructural measures are to be used at any instant of time and their levels. Since it is not impossible for the optimal levels of the nonstructural measures to vary with time, (since damages could possibly change with time) the DM can take one of two approaches: (1) allow for the possibility of the nonstructural measures to vary with time, or perhaps more realistically (2) consider the possibility of the nonstructural measures changing the levels only with the construction of a new structural measure. In Chapter III we provide models and algorithmic approaches for both of these policies. We assume that associated with each structural measure is a capital cost and an annual maintenance cost, while the nonstructural measures have their annual cost as functions of their levels. We are considering the preliminary planning stage where the DM is primarily concerned with how the funds should be allocated rather than determining, say, the precise detail of some land use allocation scheme. In this context, we quote Booth [86] "there is never an 'optimal' solution to the expansion problem, and the best that can be expected is a plan that will consider the prime factors in a rational manner and provide the background pattern, within which the particular plants may be decided upon". 
Some additional considerations in modeling the Flood Control Problem include:

The ease of incorporating political and social constraints, and the emphasis to be given to:

- Population changes with time (which are in turn dependent on the mix of measures and their sequence).
- The flood routing aspects of the problem. (For example, should the variable costs of a routing scheme be considered?, or, should the effect of nonstructural measures on flood volume be considered?)
- Determining the exact timing of the projects. (For example, can the planning horizon be split into time blocks?)
- The methods of determining damages. (For example, should the change in economic rent be a criterion?, or, can simulation methods of estimating damages be incorporated?, or, in the case of closed form damage functions, can different forms of the function be allowed?)
- Damage reduction to private parties.
- Flood variability. (For example, can the Standard Project Flood Method be used?, or, can stage-frequency relationships be used?, or, can effects of flood intensity and duration on damage be provided for?, or, can the different effects of the floods along different reaches of the river be incorporated?)

We mention all the above considerations since different applications may require differing degrees of emphasis on these considerations. In particular, the damage reduction relationships of the
various measures would be unique to each application. Our goal is to develop a general modeling and solution approach which can encompass any combination of considerations given that they can be expressed in a certain reasonable format, rather than to develop specialized models for specific applications. In Appendix B we describe an approach for constructing comprehensive damage functions which consider many damage reduction interrelationships.

A very general version of the FCP can now be stated as follows: Find a combination \((x, y)\) of structural \((x)\) and nonstructural \((y)\) measures so as to

\[
\text{Max } f(x, y) \\
\text{subject to } (x, y) \in G(x, y) \tag{P}
\]

\[
x \in X \\
y \in Y
\]

in which \(x = (x_1, x_2, \ldots, x_N)\) is the vector of structural measures, where \(x_j = 1\) if structural measure \(j\) is selected and 0 if not, \(y = (y_1, y_2, \ldots, y_K)\) is the vector of nonstructural measures, where \(y_k\) is the level of the \(k\)th nonstructural measure selected, \(f(x, y)\) is the objective function, e.g., the discounted net reduction in flood damages resulting from plan \((x, y)\), \(G(x, y)\) is the set of feasible plans \((x, y)\) i.e., those satisfying the planning, financial, engineering, and social constraints, and \(X\) and \(Y\), respectively, are the sets of feasible structural and nonstructural measures.
2.2 Review of Relevant Literature

Prior to surveying the literature, we first clarify our interest in the various models. The major contribution of most of the models is in expressing analytically their particular emphasis and considerations (hydrologic or otherwise) rather than on the algorithms used in the solution. Since we are more interested in the general formulation, and since the relationships described in a particular model are more or less unique to it, when discussing a model we essentially will be listing the various considerations accounted for rather than going into the details of the relationships described. By doing so, we still can observe the structure of the general model without any loss of generality. However, we recognize an algorithm described or used in a model if it appears to be exploiting some kind of a structure—indeed the algorithm in [C5] bears similarities to the algorithm we will eventually develop.

Man has historically attempted to control floods through physical structures, such as dams, levees, and various channel improvements. Hence there have been numerous efforts, e.g., [B1, C4, D1, E4, F1, K2, M6, M16, W7] to model the FCP as one of choosing from some discrete set of structural measures so as to minimize expected flood damages. Hydrological considerations essentially reduced these models to evaluations of the flood routing effects of different combinations of structural measures and comparisons of the routing impacts with the help of damage functions. The routing problem has been handled differently depending upon individual emphasis. For example, Khavich [K2] performed a simple simulation, Windsor [W8] used simple linear relationships, while a systems approach was taken by Davis [D2]. Since the routing is
performed essentially to determine the flood damage for different mixes of measures, the most general viewpoint is to consider the structural measure problem as a Capacity Expansion Problem [M6, E4] with project interdependence and either budgetary or deterministic/stochastic demand constraints (refer Morin and Shin [M15] for details). This problem has been studied extensively and there are several exact and heuristic methods for its solution (see Akileswaran, Morin and Meier [A3, A4] for an excellent overview of such approaches).

Unfortunately, as observed in the Introduction, structural measures have not alleviated all of the flood damages and there have even been cases documenting increased flood damages because of more intensive land use accompanying structural contributions. With this there came the recognition, see [H2, K4, M3, S1] for example, that in order to reduce flood damages effectively the structural measures have to be complemented by nonstructural measures such as Land Use Planning, Flood Plain Zoning, etc., which essentially control undue encroachment into the flood plain. Since then, there have been significant efforts [B4, D3, D4, D5, H4, K1, S4] to model the damage reduction problem as essentially a land use allocation problem for a given routing scheme. As in the case of models for structural measures, the efforts can be classified based on the differing considerations and degrees of detail. Hopkins et al. [H5, H6] model the problem as a discrete dynamic program with the reaches corresponding to the stages of the program, while Bia!as et al. B5] use a linear program for determining the land use allocations. As might be expected there have also been detailed simulation models, see for example [A10], which are too cumbersome for
preliminary planning purposes. The objective in all of these models has been either to minimize damages or to maximize rents using the change in economic rents with land use.

All these land use planning models essentially take the routing characteristics as being fixed and, hence, are suitable for applications in which structural growth has been curtailed. Considering the combined problem of determining both the structural and nonstructural measures (of which land use allocation is only one) there have been a number of conceptual frameworks [01, R2, W5] which consider the interaction among various measures (Figure 2 [Fig. 3-3, W6] is an interactions matrix). As in models for earlier frameworks, a large number of combined models [B2 and B5, C5, D1, D6, D7, E2, H7, J2, J5, R1, SS, W1] have been developed which differ primarily on their individual emphasis and degree of detail. There have been also a number of different modeling efforts which consider only agricultural damage reduction [L1, P2, S3]. With the goals of distributing evenly the flood losses rather than just reducing them, several models [F2, L3, M2] have emphasized insurance as an important nonstructural measure.

One common feature of all the combined models is that invariably land use allocation is typically the only nonstructural measure considered. As a means of providing an overview of these combined models, we now briefly describe a sample of the major models in more detail below.

James [J2, J3, J4, J5] apparently provides the first published attempts at modeling the combined problem. Floodproofing is taken into account with the help of ‘reduction factors’, and the planning horizon
<table>
<thead>
<tr>
<th>Initial Adjustment</th>
<th>Control and Protection</th>
<th>Flood-Proofing</th>
<th>Land Use Planning</th>
<th>Warnings</th>
<th>Insurance</th>
<th>Relief and Rehabilitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control and Protection</td>
<td></td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
<tr>
<td>Flood-Proofing</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>?</td>
<td>〇</td>
<td>〇</td>
</tr>
<tr>
<td>Land Use Planning</td>
<td>〇</td>
<td>?</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
<tr>
<td>Warnings</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>?</td>
<td>〇</td>
<td>〇</td>
</tr>
<tr>
<td>Insurance</td>
<td>〇</td>
<td>?</td>
<td>?</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
<tr>
<td>Relief and Rehabilitation</td>
<td>☀</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
<td>〇</td>
</tr>
</tbody>
</table>

Stimulated by the initial adjustment:

〇 — High stimulation
● — Little or none
? — Doubtful

Figure 2. Matrix of Interaction of Adjustments to Floods. [W6]
is split into time blocks. The different structural possibilities are completely enumerated and the best one picked. This same approach of enumerating the structural measures has been employed by Day and Weisz [D6, W2, W4] where they determine the optimal land use allocation with the help of a linear program (LP). The parameters of the LP are inputted for different structural combinations and sequences and, hence, the choice of structural combinations and timings is fairly limited.

Aravanitidis et al. [A9, A10] present a very detailed simulation model (however land use is the only nonstructural measure considered) which is oriented towards detailed planning considerations. The model considers effects of population changes and integration of public policies, measures and projects flood damages. It determines benefits as a sum of damage reductions, land rents, and economic rent differences. Land use allocation is optimized independently. In [A10] Aravanitidis et al. describe an example of an application to a particular flood reach. The authors noted that for this application damage reduction accounted for 80% of the total benefits.

The approach of performing a simulation/optimization where the control benefits are determined using a simulation, has been used in a few other models, for example see [D1, E1, E7, J8, L2]. These models are similar to electrical generation planning models such as [A6, B6, J6, J7, K5, G1, M11], where the systems operating cost, as opposed to the flood control benefits, are determined by simulation. Unfortunately, most of these combined simulation/optimization models are much too detailed for preliminary flood control planning. Furthermore, since most of them determine damage using the economic rent concept,
whose evaluation could be cumbersome, they are usually unsuitable for preliminary planning purposes.

Ball et al. [B2, B5] provide a combined model with structural considerations in the first paper and land use allocation performed in the latter. Reference [B2] provides for a preliminary planning model and considers flood routing independent of the nonstructural measures. It does not consider sequencing of the projects nor any budget constraints but accounts for the variable cost of routing. The structural measures are selected using a Branch-and-Bound algorithm and in reference [B5] a linear program is used to determine the land use allocation. The probabilities of various maximum flood volumes are input to the linear program, which are determined after sizing the structural measures. In [B5] the economic rent net of expected flood losses and relocation costs is measured and the increase in flood volumes which result from various land use schemes are accounted for. Unfortunately, it does not appear that this land use allocation model can be easily adapted to account for other types of nonstructural measures.

Cortes-Rivera [C5] solves the combined problem using discrete dynamic programming (DP) with each stage corresponding to a time period and the land use allocation determined using parametric linear programming with the controlling parameter being a function of the probability of flood determined by the particular mix of structural measures. Furthermore, he solves the DP using an approximate algorithm called differential discrete DP.
A common characteristic of many analytical models for the combined problem is that the interaction is usually evaluated mainly through the use of damage functions. The approach of Debo and Day [D7] could be considered as a prototype of this methodology since they model the urban watershed from the damage reduction point of view. They do not have an optimization model but provide for the determination of the flood damages for any given mix of structural and nonstructural measures. There is no provision either for the interaction between the structural measures or for their sequencing. For further specific applications we refer to Brown [B8] which contains an extensive list of research studies on different aspects of flood control.

Thus, there are apparently no models which satisfactorily evaluate a comprehensive plan considering both structural and nonstructural measures. One reason for this is the lack of suitable data [G5, G6, S2, W9] on the interdependence between structural and nonstructural measures in terms of damage reductions. For example Shabman et al. [S2] state that the influence of flood plain zoning cannot be easily identified. However, White [W6] has made a significant effort to enumerate the various interaction possibilities qualitatively (refer to Figures 1 and 2). Some examples of the interrelationships include the use of warning complements floodproofing in damage reduction and engineering projects for controlling floods in one portion of the floodplain may be accompanied by local regulations preventing further encroachments into other sectors of the floodplain. As mentioned earlier, depth/stage-damage functions could play a critical role in specifying the interrelationships. Data for this type of function are
widely available \([B7, C1, C2, H3, K3, P1, U1-U6, W3]\). Unfortunately, many of these functions are descriptive rather than prescriptive and, hence, are unsuitable for our purposes. Furthermore, they are typically restricted solely to land use allocation schemes.

Another reason for there not being any model which considers comprehensively all the nonstructural possibilities, is the lack of solution procedures for such a difficult problem. To see this we only have to look at Problem \((P)\) which is a very difficult optimization problem for a number of reasons. Firstly, the reduction in flood damages for any \(x_j\) or \(y_k\) is dependent not only on the value of \(x_j\) or \(y_k\) but also on all of the measures which both precede it and those that come after it. That is, the benefits which accrue from the nonstructural measures are dependent on structural measures selected and vice versa and these vary over the time horizon. Secondly, the problem involves both discrete and continuous decision variables -- see \([G3]\) for a discussion of the difficulties involved in such problems. That is, whereas the decisions on the structural measures involve only which measures \((x)\) to select, the decisions on the nonstructural measures involve both which and how much \((y)\) of each measure to select. For example, if \(j\) corresponds to a storage reservoir of a certain predetermined size, a decision is made either to build it \((x_j = 1)\) or not \((x_j = 0)\), whereas if \(k\) corresponds to floodproofing, a decision has to be made whether to floodproof or not \((y_k = 0)\) and if so, at what level \((y_k > 0)\). Finally, the problem is confounded by the magnitude of the number of feasible combinations \((x,y)\) of structural and nonstructural measures which must be considered.
CHAPTER III

MODEL AND SOLUTION APPROACH

Since we have clarified the various modeling considerations in Chapter II, we are now in a position to discuss and present a suitable modeling approach for the Flood Control Problem (FCP). The plan of this chapter is as follows. In Section 3.1 we present our model. We follow this with a discussion of its practical applicability in Section 3.2. The chapter concludes with the presentation of the algorithm for solving the FCP in Section 3.3.

3.1 Model Formulation

We now describe a model which is general enough to encompass the important considerations and relationships involved in flood control planning and yet is specific enough to recognize the special structure of the FCP.

Capacity expansion problems which also involve operating level decisions which are continuous in nature are fairly common. A sampling of such discrete/continuous planning problems are: Simultaneous Investment Sequencing and Allocation Decisions for Water Resources Systems Design [A7, J1]; Multi-commodity Distribution Systems Design [G4]; Generation Expansion of Hydroelectric Systems [B6]; and Investment
Sequencing with Price Sensitive Dynamic Demand [E5].

The Flood Control Problem can be viewed as a capacity expansion problem involving a discrete set of structural measures accompanied by continuous set of nonstructural measures. Therefore, we first review the available solution methodologies for the general class of discrete/continuous planning problems in order to determine their applicability to the FCP. Solution approaches to the general class can be divided into two different types each of which is discussed below.

The first type of solution methodology consists of formulating the sequencing and selection of projects problem as an integer program (IP) using 0-1 variables. The continuous variable selection problem could in general be some nonlinear programming problem (NLP) and is conditioned upon the presence or absence of some key 0-1 variables. Thus, the overall model becomes a mixed 0-1 nonlinear integer programming problem. An example of this approach was presented by Armstrong and Willis [A8] who formulate a water resources planning problem as a mixed-integer quadratic programming problem. One of the more efficient algorithmic approaches for this type of formulation appears to be Generalized Benders Decomposition (GBD) [G2]. GBD exploits a special feature of the discrete/continuous formulation: when one of the two classes of variables (termed "complicating") are fixed, the problem of determining the optimal values of the variables in other class is much simpler than the overall problem. However, Generalized Benders Decomposition has the following characteristics when applied to the FCP:

1. The planning horizon has to be split into planning periods and capacity expansion decisions made period by period. This
restricts us to the finite horizon problem. More importantly
the size of the problem increases multiplicatively with the
number of planning periods. In fact, we [A2] have recently
proven that the structural problem alone is NP-complete.

(2) Since the project selection and sequencing is formulated as an
IP, when Benders cuts are added to this IP in the course
of the decomposition algorithm we are essentially losing the
special sequencing structure and, thus, are not able to most
effectively exploit this structure.

Both of these undesirable characteristics are absent from the
second solution methodology which recognizes that the sequencing problem
is the key component and considers the overall problem as essentially
being a sequencing problem subject to certain restrictions. Indeed, the
general framework of the sequencing problem defined by Erlenkotter and
Rogers [E6] subsumes versions of all the planning problems referred to
earlier as special cases. Their elegant and simple model exploits the
sequencing nature of the problem and, furthermore, considers timings of
projects as being variables thus precluding the need to divide the
planning horizon into planning periods.

The Flood Control Problem can also be described in Erlenkotter and
Rogers' framework and, hence, we shall adapt their algorithm to the FCP.
Though it might be preferable to have a continuous time sequencing of
the structural measures, as opposed to dividing the planning horizon
into planning periods, we need to consider the planning horizon as being
a multiple of some basic unit in order to meaningfully express the
expected flood damages. That is, in analyzing the effects of flood
events arising as a result of natural conditions which repeat in cycles, it would not be possible (except in trivial cases) to estimate flood damage over a partial cycle. This does not imply that we need to make a time unit by time unit analysis of the entire horizon, as done in [C5, J2] and others (indeed would not even be possible much less meaningful in the infinite horizon case), but while determining the timing of the structural measures, we need to consider time in multiples of a basic cycle unit (say a year) rather than as a continuous variable. For ease of formulation, we shall consider the time unit to be a year.

We first construct the model for the Flood Control Problem assuming that the nonstructural measures vary with each year. The selections of the nonstructural measures are related to the mix of the structural measures in a natural way. For any year, of the planning horizon and a given set of structural measures present during this year, the determination of the optimal levels, \( y^* \), of the nonstructural measures, which minimize the expected damages during that year is a relatively well defined problem. This is in accordance with the viewpoint that the nonstructural measures essentially complement a given set of structural measures in terms of damage reduction. We can express this problem as

\[
P(I, t) = \min_{y \in Y \cap G(I)} \{D(I, y, t) + C(y, t)\}
\]

where \( D(I, y, t) \) is the annual flood damage in the \( t^{th} \) year for a given combination 'I' of the structural measures when the level of nonstructural measures are 'y'; \( C(y, t) \) is the annual cost incurred in the \( t^{th} \) year with the levels of the nonstructural measures at 'y'; and \( G(I) \)
and \( Y \) are similar to those defined in problem (P). \( P(I, t) \) denotes the minimal net annual damage-nonstructural measure cost in the \( t \)-th year for the combination 'I' of the structural measures. We observe that the concept of considering the nonstructural measure determination as a subproblem conditioned on 'I' has previously been used in reference [C5]. In [C5], (1) is an LP parametrized by some function of \( I \). LP subproblems are also used in reference [W2].

We note that when the Decision Maker (DM) cannot differentiate between the years while constructing the damage function (as is normally the case), then \( P(I, t) \) could be reduced to \( P(I) \). However, in order to not lose the general structure, we shall allow \( P \) to be a function of \( t \). If we interpret \( P(I, t) \) as the annual operating cost in the \( t \)-th year of the set 'I' of structural measures, then we can use the very general (discrete time) sequencing framework developed by Erlenkotter and Rogers [E6]. Notice that the optimization problem in (1) need not be solved or expressed in closed form as long as the optimal levels of the nonstructural measures can be obtained in some fashion; for example, a simulation/optimization may be performed over the nonstructural measures. We caution the use of simulation though, since it usually requires modeling effort and has data requirements which may be excessive for preliminary planning purposes (refer to [D6] for similar comments). Though Erlenkotter and Rogers [E6] present their framework in continuous time, they observe that the approach can also be adopted for discrete time as we will do.

The following assumptions from [E6] are relevant to the formulation of the Flood Control Problem.
(a) A finite number, \( m \), of structural measures may be undertaken, each project is indexed by an \( i = 1, 2, \ldots, m \). The investment cost for project \( 'i' \) is given by \( c_i > 0 \). Investment costs include an allowance for the present value of maintenance, fixed routing operations, and other fixed operating costs. Hence, capacity once established may be considered to have infinite life.

(b) Let \( I^* \) denote the set of all project indices, \( I \) denote an arbitrary subset of project indices, and \( \emptyset \) denote the set of all subsets \( I \). The variable operating cost rate (annual net flood damage-nonstructural measure cost) in year \( 't' \) for the project set \( 'I' \) is expressed by \( P(I, t) > 0 \). \( P(\emptyset, t) \) denotes the annual damage-nonstructural measure cost for the current set of structural measures.

For each \( I \), some project \( i \in I \) must be established and added to \( I \) no later than the time \( T(I) > 0 \). If adding additional projects to \( I \) is not required \( T(I) \) is defined to be infinite. Furthermore, \( T(I) \leq T(I \cup i) \) for \( i \in I \) and \( T(I^*) = + \infty \).

(c) Costs are continuously discounted at a constant rate, \( r > 0 \), leading to a discount factor of \( e^{-rt} \) from time \( t \) to the initial time \( 0 \). Note that other forms of discounting could also be easily incorporated.

(d) Sequencing and timing decisions for the projects are to be selected so as to minimize the total net discounted damages over an infinite horizon. (The same approach applies to the
finite horizon case).

The minimum damage sequence is selected from the \( m! \) possible orderings for the projects. To express the selection problem, we introduce the following additional notation from [E6]: \( i[k] \) is the project index assigned to the \( k \)th position in a sequence; \( \{i[k]\} \) is the complete assignment of project indices for a particular sequence, where \( k = 1, 2, \ldots, m; \) \( S_I \) is the set of all permutations of project indices in \( I; \) \( I_k \) is the set of first \( k \) project indices for a particular sequence, where \( I_0 = \phi, I_{k+1} = I_k \cup i[k+1] \) for \( k = 0, 1, \ldots, m-1, \) and \( I_m = I^*; \) \( \tau_k \) is the establishment time for the \( k \)th project in a sequence where, \( \tau_0 = 0, \tau_k \leq \tau_{k+1}, \) and \( \tau_{m+1} = +\infty; \) \( c(I^*, \infty) \) equals the total net damage over the time interval \([0, \infty)\) discounted to time 0 for a minimum-damage sequencing.

The model of the Flood Damage Control Problem is

\[
C(I^*, \infty) = \min_{\{i[k]\} \in S_I} \min_{\{\tau_k\} \in I^*} \left\{ \sum_{k=0}^{m} \left( \sum_{t=\tau_k}^{\tau_{k+1}-1} p(I_{k+1}, t) e^{-rt} \right) \right\} + \sum_{k=1}^{m} c_i[k] e^{-r\tau_k}
\]

with,

\[
0 = \tau_0 \leq \tau_1 \leq \ldots \leq \tau_m < \tau_{m+1} \text{ and } \tau_{k+1} = T(I_k), \text{ } k = 0, 1, \ldots, m-1.
\]

Note that the model (2) allows for the possibility of not establishing some projects since an establishment timing of a very large value implies indefinite postponement, which is equivalent to
eliminating the project from consideration.

We now consider the model of the FCP when we allow for the possibility of the nonstructural measures to change their level only with the construction of a new structural measure. Let \( y_k \) be the new level of nonstructural measures accompanying the \( k \)th construction of a structural measure, where \( y_0 \) is the initial level of the measures. In the format of (2) this becomes

\[
C(I, \infty) = \min_{i \in I} \min_{k \in K} \sum_{k=0}^{m} \sum_{t=\tau_k}^{\tau_{k+1}-1} [D(I, y_k, t) + C(y_k, t)] e^{-rt} + \sum_{k=1}^{m} c_i[k] e^{-r\tau_k}
\]

(3)

with

\[
\tau_0 \leq \tau_1 \leq \cdots \leq \tau_m < \tau_{m+1} \quad \text{and} \quad \tau_k \leq T(I_{k-1}), \ y_k \in G(I_k) \cap Y, \quad k=1, \ldots, m.
\]

The main contribution of this section is recognizing that selection of nonstructural measures in a year is conditioned upon the presence of the structural measures in a natural way and, hence, the resulting annual damage-nonstructural cost can be considered as an annual operating cost of the structural measures. Thus, we can meaningfully adapt Erlenkotter and Rogers' model [E6] for our flood control problem. We conclude this section by mentioning that our models (2) and (3) fall within the framework presented in Section 2.1 and are
general enough to account for (see Section 3.2) most of the considerations in the various models surveyed in Section 2.2.

3.2 Applicability of Model

Recall that the subproblem involving the optimization over the nonstructural measures is

\[ P(I,t) = \min_{y \in Y(I)} \{ D(I,y,t) + C(y,t) \} \]

in which \( P(I,t) \) is the net minimum annual flood damage-nonstructural measure cost in year 't' when the mix of structural measures is 'I'. The key to the applicability of our model lies in the elements of expression (4) and their meaningful evaluation. Since one would not normally expect the damage effects to change with time, we will consider a simpler version of (4), namely,

\[ P(I) = \min_{y \in Y(I)} \{ D(I,y) + C(y) \} \]

This function characterizes the effects of the nonstructural measures with respect to damage reduction. Various considerations can be incorporated through (5) are:

- **Social and political constraints** can be expressed through the set \( Y \) or the constraint set of \( G(I) \).

- **Population changes with time** can be considered as affecting the damage function \( D(I,y,t) \) where now 't' has been appended in order to incorporate the changing considerations with time. Obviously only those population changes which would occur no matter what mix of measures is used should be considered.
Explicit consideration of the interdependence between flood control schemes and population changes may be unwarranted in preliminary planning and, hence, we emphasize only one aspect of this relationship.

- **Flood Routing**: We can consider the damages which occur during the presence of different levels of structural measures to be a function of some hydrological parameters. Thus, the variable cost of routing for a given set of structural measures could be incorporated into the damage function. Flood routing could be done using a Standard Project Flood or a distribution of floods to estimate expected hydrological parameters.

- **Damage function**: A simulation could be performed to determine the damage if need be. Note that this function is crucial in applying the model. Our models (2) and (3) provide for ease of incorporating different types of damage functions.

- **Damage reduction to private parties**: If this is also a criterion, then this could be considered as another criterion in a multi-objective framework. The sum of the expected damages to both private parties and the local government does not appear to be a reasonable objective. Hence, both the models (2) and (3) emphasize damage reduction to the government since they incur the major cost of the measures.

The program for determining $P(I)$ should be fairly simple (for example, a parametised LP) since this program will be used frequently during the execution of the dynamic programming (DP) algorithm discussed.
in the following section.

Section 3.3 Solution Algorithm

We now present and discuss algorithmic approaches for the Flood Control Problem expressed in the form of models (2) and (3). The basis for our approach to model (2) is the dynamic programming algorithm presented by Erlenkotter and Rogers [E6]. For convenience, we rewrite model (2)

$$C(I^*, \omega) = \min_{\{i[k] \in \mathcal{I}^*_1\}} \min_{\{\tau_k\}} \left\{ \sum_{k=0}^{m} \sum_{t=\tau_k}^{\tau_{k+1}-1} P(I_k, t) e^{-rt} \right\}$$

$$+ \sum_{k=1}^{m} c[i[k]] e^{-r \tau_k}$$

$$0 = \tau_0 \leq \tau_1 \leq \ldots \leq \tau_m < \tau_{m+1} = \omega, \quad 0 \leq \tau_k \leq T(I_{k-1}) \quad k=1, \ldots, m.$$  

where $\tau_k$ is the beginning of the year in which project $i[k]$ is available for use. If $\tau_{k+1} - 1 < \tau_k$ we define the summation to be 0.

Note that (6) is a very difficult optimization problem involving both discrete and continuous variables. Fortunately, however, this problem reduces to a sequencing problem when we make some suitable transformations. Observe that (6) is of the form:

$$C(I^*, \omega) = \min_{\{i[k] \in \mathcal{I}^*_1\}} \{F(i[k])\}$$

That is, we are optimizing the function $F$ over all permutations of the i's. For a general $F$, (7) is a difficult combinatorial problem.
However, if $F$ is separable into functions $g(I_k, i[k]), k=1, \ldots, m$, we can exploit branch-and-bound strategies to solve (7). In particular, (7) would fit naturally into the discrete dynamic programming (DP) framework, and indeed, it is this approach that we shall use to solve (6).

Initially it may not be obvious how to separate the return function into a sum of $g(I_k, i[k])$'s. That is, when we consider each $i[k]$ to be associated with the determination of $\tau_k$, the limits, $\tau_k^{-1}$ and $\tau_k$ of the summation term complicate matters as regards the separation into functions of $I_k$ and $i[k]$. However, Erlenkotter and Rogers [E6] use a simple result to transform the return function. They use the fact that

$$\tau_k^{-1} \sum_{t=\tau_k}^{\tau_k+1} P(I_k, t)e^{-rt} = \sum_{t=\tau_0}^{\tau_k} P(I_k, t)e^{-rt} - \sum_{t=\tau_0}^{\tau_{k-1}} P(I_k, t)e^{-rt} \quad (8)$$

Thus, (6) can be rewritten as

$$C(I^*, \infty) = \min_{\{i[k]\} \in S} \sum_{k=1}^{m} \min_{I_k, k=1} \tau_k^{-1} \leq \tau_k \leq \tau(I_k^{-1}) \quad (9)$$

$$\left\{ \begin{array}{l}
\tau_k^{-1} \\
\sum_{t=\tau_0}^{\tau_k} [P(I_k, t) - P(I_k, t)]e^{-rt} \\
\sum_{t=\tau_0}^{\tau_{m+1}} P(I^*, t)e^{-rt}
\end{array} \right\} + c_{i[k]} e^{-rf_k}$$

We can now interpret
\[ \min_{\tau_k \leq \tau_k^*} \sum_{\tau_{k-1} \leq \tau_k \leq T(I_{k-1})} \left[ P(I_{k-1}, t) - P(I_k, t) \right] e^{-rt} + c_i e^{-r\tau_k} \] (10)

as the return associated with adding \( i[k] \) to obtain the combination \( I_k \).

We now consider using DP to solve (9), with \( 'I' \) as a state, \( |I| \) as the stage associated with \( 'I' \), and adding \( i \in I \) as the last project to obtain \( 'I' \) as decisions associated with \( 'I' \). In the notation of [M9] the return (local income) function, \( h \), is

\[ h(\epsilon, I = I_k, i = i[k] \in I) = \] (11)

\[ \epsilon + \min_{\tau_{k-1} \leq \tau_k \leq T(I-i)} \left\{ \sum_{\tau_{k-1} \leq \tau_k \leq T(I-i)} \left[ P(I-i, t) - P(I, t) \right] e^{-rt} \right\} + c_i e^{-r\tau_k} \]

This return function (11) does not necessarily satisfy the monotonicity condition (MA) [M10] required for the validity of the functional equations of DP. That is, let \( \epsilon_1, \) and \( \tau_{k-1} \) be the total return and project timing respectively associated with \( i_{k-1} \) as the last project added to construct \( I-i \), and let \( \epsilon_2 \) and \( \tau_{k-1} \) be those associated with \( i_{k-1} \) as the last project added to construct \( I-i \). Then if \( \epsilon_1 < \epsilon_2 \) and \( \tau_{k-1} > \tau_{k-1} \) it is possible that \( h(\epsilon_1, I, i) > h(\epsilon_2, I, i) \). Notice that the reason that this return function does not satisfy the MA is the constraint \( \tau_{k-1} \leq \tau_k \). Thus, consider a relaxation to (9) with the constraints \( \tau_{k-1} \leq \tau_k \) deleted, namely,
\[ C_R(I^*, \omega) = \min_{\{i[k]\} \in S^*} \min_{k=1}^m \left\{ \tau_k^{-1} \right\} \sum_{t=\tau_0}^{\tau_k - 1} [P(I_{k-1}, t) - P(I_k, t)] e^{-rt} + c_{i[k]} e^{-r \tau_k} \sum_{t=\tau_0}^{\tau_{m+1} - 1} P(I^*, t) e^{-rt} \] (12)

For (12), the local income function is

\[ g(I, i) = \min_{0 \leq \tau \leq \tau(I-i)} \{g(I, i, \tau)\} \] (13)

\[ = \min_{0 \leq \tau \leq \tau(I-i)} \sum_{t=\tau_0}^{\tau - 1} [P(I-i, t) - P(I, t)] e^{-rt} + c_i e^{-r \tau} \]

Notice that the return function associated with this local income function (13) satisfies the MA and hence we can use DP to solve (12). Note that when using DP we need efficient mechanisms to evaluate the local income function (13) since it will be frequently evaluated in the course of the algorithm.

Since we are only solving a relaxation to (9), we need a procedure to translate the solution of the relaxed problem into an optimal solution to the original problem. We do this by imposing some reasonable conditions on \( P \) which guarantee that the optimal solution to the relaxed problem is feasible and hence optimal to the original problem. For this, we need the following definition. If there are many \( \tau' \)'s which minimize (13), we define \( \tau^* = \max \tau' \), i.e., \( \tau^* \) is the latest minimum damage ('lmd') establishment time for (13). Combining (12) and (13), the relaxed formulation becomes
\[ C_R(I^{i,\infty}) = \min_{\{i[k]\}\in S_I^*} \left( \sum_{k=1}^{m} g(i_k, i[k]) + \sum_{t=\tau_0}^{\tau_{m+1}-1} P(I^*, t) e^{-rt} \right) \]  

(14)

Denoting \( n = |I| \), we define

\[ f_n(I) = \min_{\{i[k]\}\in S_I} \left( \sum_{k=1}^{n} g(i_k, i[k]) \right) \]  

(15)

Finding \( f_m(I^*) \) is equivalent to finding \( C_R(I^{i,\infty}) \) since

\[ C_R(I^{i,\infty}) = f_m(I^*) + \sum_{t=\tau_0}^{\tau_{m+1}-1} P(I^*, t) e^{-rt} \]  

(16)

It is easily seen that the sequencing of the projects in \( I \) must be the minimum net damage as defined in (15). Possibly several sequences of projects in some subset \( 'I' \) may provide equal minimum net damage values. If this occurs, we adopt the tie-breaking rule of selecting the minimum net damage sequence that has the latest 'lmd' establishment time \( \tau^*(I, i) \) for the last project 'i' added to complete the set 'I'. We shall call a sequence obtained by this rule a latest minimum damage solution for 'I' and shall denote 'lmin' the lexicographic solution process of comparing solutions first on costs and second on the 'lmd' establishment time for the last project, i. Notice that this definition of the latest minimum damage time for I is consistent with the definition for the 'lmd' establishment time for each 'i'.

We may determine a latest minimum-damage solution for the relaxed formulation (14) with the following dynamic programming recursion
\[ f_n(I) = \min_{i \in I} \{ g(I, i) + f_{n-1}(I-i) \} \]  

(17)

In order to guarantee that the optimal solution to the relaxed formulation is feasible for (9) and hence optimal to (6), \( P \) must satisfy certain conditions. One such condition is

\[ P(I, t) - P(I \cup i, t) \geq P(I \cup i', t) - P(I \cup i' \cup i, t) \]  

(18)

That is, the damage reductions obtained by adding project 'i' to the set 'I' are nondecreasing as additional projects are added to 'I'. Another way of viewing this is

\[ P(I, t) - P(I \cup i \cup i', t) \leq [P(I, t) - P(I \cup i, t)] \]  

\[ + [P(I, t) - P(I \cup i', t)] \]  

(19)

That is, the projects have a sub additive relationship in terms of damage reduction; there is no synergism. For example, if we are considering constructing a levee and performing channel improvement, the sum of the damage reductions associated with constructing a levee only and the damage reductions associated with performing channel improvement only, will be greater than the damage savings associated the implementation of both these structural measures.

The use of this condition (18) in guaranteeing optimality of solution given by (17) to the original problem (9) is easily established.

Proposition 3-1. Given that (18) is valid, in the solution to the relaxed formulation (12) defined through (17), the timings for
successive expansions satisfy the constraints $\tau_k \leq \tau_{k+1}$ and, thus, the solution also solves the original formulation (6).

Proof: Follows from the proof of Proposition 2 of [E6], mutatis mutandis and, hence, is omitted.||

Thus, when the minimum net damage function $P$ satisfies (18), the solution procedure defined by (17), is valid and $C(I^*_{k-1}) = C_R(I^*_{k-1})$. Erlenkotter and Rogers [E6] observe that this forward dynamic programming relationship (17) generalizes the forward formulation for the simple expansion sequencing problem of [E3, E4].

We now discuss solution procedures for evaluating the local income function, $g(I^*_k, i[k])$, as given in (13). Notice that this evaluation involves 1) determining the 'lmd' establishment time $\tau^*_k$ and 2) calculating the value $g(I^*_k, i[k]) = g(I^*_k, i[k], \tau^*_k)$. Note that it is possible that $g(I^*_k, i[k], \tau^*_k)$ is obtained in the process of determining $\tau^*_k$, though this need not necessarily be the case. We first discuss obtaining the latest minimum damage establishment time $\tau^*_k$. If $T(I^*_{k-1})$ is finite, the most primitive approach would be to enumerate $g(I^*_k, i[k], t)$ for $t = 0, 1, \ldots, T(I^*_{k-1})$ and to select the minimum $g(I^*_k, i[k], t)$. Furthermore, if only one such 't' corresponds to the minimum value of $g$, we can automatically set it to be $\tau^*_k$. If more than one such 't' minimizes $g$, then we set the latest such $t$ to be $\tau^*_k$. Also, with the requirement that condition (18) be satisfied, we need to enumerate only for $t = \tau^*_k$ to $\tau(I^*_k-1)$ where $\tau^*_k$ corresponds to the optimal timing associated with the set $I^*_{k-1}$. Note that in the enumeration process,
\[ g(I_k, i[k], t) = \sum_{\tau=\tau_0}^{t-1} [P(I_k-i[k], \tau) - P(I_k, \tau)] e^{-\tau} + c_i e^{-\tau}, \]

has to be calculated for each \( \tau = \tau_0, \tau_0+1, \ldots \). This is best implemented as follows. For convenience, we first define

\[ j(I_k, i[k], t) = P(I_{k-1}, t) - P(I_k, t) + c_i[k] (1 - e^t) \quad (20) \]

Then we have

\[ g(I_k, i[k], t) = g(I_k, i[k], t-1) + j(I_k, i[k], t-1) e^{-\tau} \quad (21) \]

Thus, determining \( g(I_k, i[k], t) \) in the year \( 't' \) would essentially involve determining the damage reductions \( P(I_k-i[k], t) - P(I_k, t) \) in the year \( 't' \) only. However, this explicit enumeration may be laborious or impractical when \( T(I_k-1) \) is large or infinite. If the damage reduction function \( P(I_{k-1}, t) - P(I_k, t) \) satisfies some reasonable properties, we can avoid explicit enumeration. In particular, let

\[ P(I, t) - P(I i, t) \quad (22) \]

be nondecreasing in \( t \) for each \( I \in \mathcal{I} \) and \( i \in I \).

That is, let the damage reductions associated with adding \( 'i' \) to the set \( 'I' \) not decrease with time. If condition (22) is satisfied, we have

**Proposition 3.2.** Define

\[ \tau^* (I, i) = \tau : j(I, i, \tau) \leq 0 \text{ and } j(I, i, \tau+1) > 0, \text{ if } j(I, i, T(I-i)) > 0 \]

\[ T(I-i) \text{ otherwise.} \]

Then if (22) is satisfied, \( \tau^* (I, i) \) is the 'lmd' establishment timing
for (13).

Proof: From (21) observe that

\[ g(I, i, t) - g(I, i, t-1) = j(I, i, t) e^{-rt} \]

Also from condition (22) we have \( j(I, i, t) \) nondecreasing in 't' for \( t = 0, 1, \ldots, T(I, i) \). Thus, eliminating the extreme cases of \( j(I, i, 0) > 0, T(I-i) = 0, \) and \( j(I, i, T(I-i)) < 0 \), observe that \( g(I, i, t) \) monotonically decreases at each \( t=0,1, \ldots, \tau^*(I, i) \), takes the minimum value at \( \tau^*(I, i) \) and monotonically increases for \( t=\tau^*(I, i) +1, \ldots, T(I-i) \). Thus, \( \tau^*(I, i) \) corresponds to the latest minimum cost establishment time for (13).||

Proposition 3-2 essentially provides a stopping criterion for the enumeration process of \( g(I, i, t) \) and thus we need to determine \( g(I, i, t) \) for \( t=0,1, \ldots \) using (21) until we reach a 't' such that \( j(I, i, t) > 0 \). Then \( \tau^*(I, i) = t-1 \). Since with condition (22), \( g(I, i, t) \) is unimodal in 't' binary search techniques may be employed for determining \( \tau^*(I, i) \). Given any \( t_1 \) and \( t_2, t_2 > t_1 \) such that \( j(I, i, t_1) \leq 0 \) and \( j(I, i, t_2) \geq 0 \), we use the property \( t_1 \leq \tau^*(I, i) \leq t_2 \) to perform the binary search.

We now discuss procedures for determining \( g(I, i, \tau^*(I, i)) \) given that we have determined \( \tau^*(I, i) \). Notice that if the explicit enumeration scheme was employed to determine \( \tau^*(I, i) \), \( g(I, i, \tau^*(I, i)) \) is also determined in the process. However, this is not true if binary search is used to determine \( \tau^*(I, i) \).

The suitability of procedures for determining \( g(I, i, \tau^*(I, i)) \) depends on whether it is easier and natural to evaluate the net annual damages \( P(I, t) \) or to evaluate the net annual damages reductions associated with the addition of 'i' to the set 'I-i', \( P(I-i, t) - P(I, t) \).
If the damage reductions are easier to evaluate, it is reasonable to determine \( g(I, i, \tau^*(I, i)) \) directly from the relation

\[
g(I, i, \tau^*(I, i)) = \sum_{t=\tau^*(I, i)-1}^{n} [P(I, i, t) - P(I, t) + c_i (1-e^r)] e^{-rt}
\]  

(24)

This basically involves for each \( I \) and \( i \in I \) determining the damage reductions associated with adding \( i \) for \( t=0, 1, \ldots, \tau^*(I, i)-1 \).

If the net annual damages, \( P(I, t) \) are easier to evaluate, we can use a modified formulation [E6] that avoids having to determine \( g(I, i, \tau^*(I, i)) \). The modified formulation involves rearranging terms so that the sum of damage reductions expressed in (9) need not be computed directly. Define \( \tau^*(I, *) \) as the timing of the last project added in the solution obtained from (17) for the projects in \( I \). Let \( C(I, \tau^*(I, *)) \) denote the total damages for the project set \( I \) over the interval from time 0 until the year before the last project is added at \( \tau^*(I, *) \).

Comparing (6), (12), (14), and (15), we see that with \( n=|I| \)

\[
C(I, \tau^*(I, *)) = f_n(I) + \sum_{t=\tau^*(I, *)-1}^{n} P(I, t) e^{-rt}
\]  

(25)

Also, from (16) and (25) we have

\[
C(I, \tau^*(I, *)) = C(I, \tau^*(I, *)) + \sum_{t=\tau^*(I, *)}^{\infty} P(I, t) e^{-rt}
\]  

(26)

The values of \( C(I, \tau^*(I, *)) \) cannot be found directly in a formulation such as (17) since \( \tau^*(I, *) \) is a result of the procedure. We shall devise an indirect procedure that also removes the necessity of
determining damage reductions over the entire interval from time 0. Suppose we designate $i^*$ as the candidate expansion that has the latest expansion timing, among all candidates $i \in I$. Define any $\tau^+$ in the range $\tau^*(I, i^*)$ to $T(I)$. Adding $\sum_{\tau=0}^{\tau^+} P(I, \tau) e^{-r \tau}$ to both sides of (17) and substituting for $f_n(I)$ from (25) yield

$$f_n(I) + \sum_{\tau=0}^{\tau^+} P(I, \tau) e^{-r \tau} = \min_{i \in I} \left\{ C(i-i, \tau^*(I-i, *)) \right\}$$  \hspace{1cm} (27)

$$+ \sum_{\tau=\tau^*(I-i, *)}^{\tau^+} P(I, \tau) e^{-r \tau}$$

$$+ \sum_{\tau=\tau^+(I, i)}^{\tau^+} P(I, \tau) e^{-r \tau}$$

for all non-empty $I \in \mathcal{I}$.

For any given set 'I' with $n = |I|$, given that we know $C(I-i, \tau^*(I-i, *))$ for the project sets, 'I-i' in stage $n-1$ and that $\tau^*(I, i)$ has been determined for each $i \in I$, we can perform the minimization in (27) and determine the next suitable project $i^*$ to be added to obtain 'I' and it associated optimal timing $\tau^*(I, i^*)$. Notice that in (27) neither $f_n(I)$ nor the summation on the left-hand side is found explicitly since all that is needed is the sum of the two, provided by the minimization on the right-hand side. When this minimization is completed, we can find $C(I, \tau^*(I, *))$ for the next stage from the relation
\[ C(I, \tau^*(I, *)) = [f_n(I) + \sum_{t=0}^{\tau^+} P(I,t) e^{-rt}] - \sum_{t=\tau^*(I, *)}^{\tau^+} P(I,t) e^{-rt} \] (28)

Both of the quantities on the right-hand side of (28) have been determined during the solution of (27). When the solution process is completed, \( C(I^*, \infty) \) may be obtained through (26).

We now discuss an interpretation for (27). Notice that the summation on the left-hand side is the minimum total damages among possible sequences in the set 'I' over a sufficiently long planning horizon, \( \tau^+ \). Thus (27) says that, no matter what timings are assigned to the projects of the latter set \( I^*-I \), the total damages associated with the projects of the initial set 'I', can be obtained independently, by considering the sequencing problem associated with the minimum of the total damages using the projects of set 'I' over a sufficiently long planning horizon. In particular, if the planning horizon is not sufficiently long, for example if \( \tau^+ < \tau^*(I, i) \) for some 'i', then we cannot validly use (27) to determine \( i^* \) and corresponding \( \tau^*(I, i) \).

In solving (27), we see that the sums of the damage reductions are no longer included. It is only necessary, for each \( i \in I \), to find \( P(I-i, t) \) for \( t = \tau^*(I-i, *) \) to \( \tau^*(I, i) - 1 \), and \( P(I, t) \) for \( t = \tau^*(I, i) \) to \( \tau^+ \). There are many possible strategies for implementing these computations in (27), and the selection among them will depend on the effort involved in determining \( P(I, t) \). One such strategy is to determine \( P(I, t) \) over a sufficiently long interval, and to store this information for later use in finding the timing decisions \( \tau^*(I, i) \) as well as in performing project selection computations as in (27).
Considering the implementation of dynamic programming procedures for (17) and (27), Morin [M7] discusses that it is more useful to employ the alternate label-setting technique of reaching. In particular, reaching can often be accelerated as in the case of dynamic programming with bounds. Indeed, since later in this section we will be considering bounding approaches to eliminate some states and transitions, it is more appropriate to use reaching for (17) and (19) as opposed to the "pulling" of DP. Furthermore, even if no bounding approaches are used, reaching can take no more effort than the usual recursive computation of DP. For a discussion of the particular advantages of reaching refer to [M7].

We now consider some exclusion tests which help in reducing the number of potential expansions 'i' to be considered while solving (27). Reaching involves updating the temporary label for a state i as potential expansions 'i' are progressively considered. If lower bounds on $P(I,t)$ can be determined with relative ease, then when solving (27), it is possible to obtain fairly easily lower bounds on the local income function associated with potential expansions. Thus, if for some $i \in I$, the lower bound on the return function, is greater than the current temporary label for the state $I$, then the potential expansion on 'i' can be excluded and thus, we can avoid determining the $P(I,t)$'s associated with the exact computation of the local income function. One instance where the lower-bounds on $P(I,t)$ can be obtained fairly easily is when $P(I,t)$ for a given $I$, is an LP parametrized on the RHS by some function of $t$. In this case, the dual variables obtained for any time 't' can be used to calculate lower-bounds on $P(I,t)$ for any $t$. 
Another initial test for screening potential expansions, $i$, is if
\[ P(I-i,\tau^*(I-i_*)) - P(I,\tau^*(I-i_*)) + c_i(1-e^r) \geq 0 \] with $\tau^*(I-i_*) > 0$, then from Proposition 3-2, we know that $\tau^*(I_*^*) < \tau^*(I-i_*)$ which is infeasible and hence 'i' cannot correspond to the latest minimum damage solution and can be excluded.

A third type of exclusion test [E6] involves performing a pair wise interchange comparison between projects adjacent in a sequence. Let $i^* \in I-i$ be the last project added in the solution for projects in $I-i$ obtained through (17) or (27). Project 'i' cannot be the optimal last project added to 'I' if interchanging the order of expansions $i^*$ and $i$ will provide a solution with lower damages. Thus project 'i' can be excluded from consideration if
\[ g(I-i_i) + g(I,i) > g(I-i_i^*) + g(I,i^*) \] (29)

It has been shown [E6] that the following is a sufficient condition for (29)
\[ P(I-i,t) - P(I-i_i^*,t) + rc_i^* - rc_i > 0 \] (30)

for all $\tau = \tau^*(I-i_*)$ to $\tau^*(I_i)-1$.

It is apparent that this exclusion test is useful when damage reductions are more easily computed than the direct damages. Erlenkotter and Rogers [E6] claim that exclusion tests such as (30) often yield considerable insight into the relative superiority of various capacity types when added to different base systems.

All the above exclusion tests help in reducing the number of reaches to be performed from any state. We now consider some bounding
tests, where the state itself can be eliminated from further consideration. If for some project set \( I, \tau^*(I-i,*):=\infty \) (or a large value) for each \( i \in I \), then it is not necessary to undertake all the projects in \( I \) and make a detailed examination. Indeed we have,

\[
C(I,*):= \min_{i \in I} \{C(I-i,*))\} \quad (31)
\]

Thus, if at any stage of the reaching algorithm, the minimum damages for all remaining states, \( I \), can be defined as in (31), then the algorithm may be terminated and the state having the least minimal damages selected to reconstruct the optimal sequence.

We can also use branch-and-bound strategies in the reaching algorithm (refer Morin and Marsten [M12, M13]) to eliminate states. If the sum of the lower bound on future costs and an intermediate value of \( f_n(\cdot) \) or \( C(I, \tau^*(I,*)) \) exceeds an upper bound (given by some feasible sequence), the state 'I' can be eliminated from further consideration.

Thus, in summary, for model (2) given the assumptions (18) and (22), we can use most effectively the reaching implementation of a hybrid DP branch-and-bound approach so as to incorporate various tests helpful towards increasing the computational efficiency of the algorithm.

We now discuss algorithmic approaches for model (3). For convenience we rewrite it
\[ C(I^*, \omega) = \min_{\{i[k]\} \in \mathcal{S}^*} \min_{\{\tau_k, Y_k\}} \sum_{k=0}^{m} \sum_{\tau=\tau_k}^{\tau_{k+1}} P(I_{k-1}, Y_{k-1}, t) e^{-rt} \]

\[ + \sum_{k=1}^{m} c_i[k] e^{-rt} \]

(32)

with \( \{\tau_k\} \) and \( \{Y_k\} \) subject to feasibility constraints, where \( P(I_k, Y_k, t) = D(I_k, Y_k, t) + C(Y_k, t) \).

Similar to the approach for model (3), we can rearrange the terms of (32) so that we have

\[ C_R(I^*, \omega) = \min_{\{i[k], Y_k\}} \sum_{k=1}^{m} \sum_{t=\tau_0}^{\tau_{k+1}} g(I_k, Y_k, i[k]) + \sum_{t=\tau_0}^{\tau_{k+1}} P(I^*, Y_m, t) e^{-rt} \]

(33)

where

\[ g(I_k, Y_k, i[k]) = \min_{\tau_k \leq t(I_{k-1})} \sum_{t=\tau_0}^{\tau_{k+1}} [P(I_{k-1}, Y_{k-1}, t) - P(I_k, Y_k, t)] e^{-rt} \]

(34)

\[ + c_i[k] e^{-rt_k} \]

This relaxation to the original problem (32) can be solved as a control problem in discrete time and a discrete-continuous state space. The basic approach is to obtain a closed form solution for the optimal \( \tau_k \) as a function of I and Y, but for our problem (34) since we do not expect the function \( P(, ,) \) to be of convenient closed forms, this approach does
not appear to be useful. We conclude that the model (3) requires a solution approach more complex and different from that for model (2).

We now discuss some features of model (2) and our algorithm for solving it. Model (2) requires that for each project, I, the absolute year T(I) in which new construction must be added to I, be supplied. But it appears that it is more reasonable to require that the year in which new construction must be added relative to the year of the worst case project in I be supplied. It is not clear how this information can be used to obtain T(I) for each I as required by model (2).

In general, the net minimum annual damage-nonstructural measure cost for any year 't' is not a closed form function of 't' and has to be derived by performing an optimization. Thus it appears that barring some special cases, assumptions (18) and (22) can be checked only in a qualitative sense. Topkis [T1, T2] provides a general theory that could provide verification of both these conditions for any application. Though it is true that the damage reductions \( P(I,t) - P(I^h,t) \) as well as the damages \( P(I,t) \) are positive for all 't', we do not need this property for our algorithm.

Our solution approach has a natural behavior. The basic problem is that of sequencing of the structural measures. The details of the nonstructural measures are absorbed in the \( P \) function. Thus there appears to be a natural hierarchy in the solution procedure as regards degree of detail.

Our model (2) allows for the nonstructural measures, their cost and utility, to influence the timing of the structural measures. This is in contrast to previous models (see for example Wiesz and Day [W2]) where
the timings are predefined, or where timing decisions are exhaustively enumerated in a discrete time framework (see for example Cortes-Rivera [CS]).
CHAPTER IV

Computational Efficiency

This chapter is addressed to the efficient implementation of our algorithm for model (2) of the flood control problem (FCP). In Section 3.3 we noted that our algorithm is best implemented as a 'reaching' procedure for a hybrid dynamic programming branch-and-bound (DP-B&B) algorithm. In particular, we discussed the use of elimination by bounds (fathoming) in order to reduce the state space of dynamic programming. Elimination by bounds is particularly important for sequencing problems since dynamic programming involves storage and computational requirements which increase exponentially with the size of the problem. In particular, when the number of structural measures, \( m \), is odd, the in-core storage required for implementing the algorithm will have to be at least \( \left( \frac{m}{(m-1)/2} \right) + \left( \frac{m}{(m+1)/2} \right) \) states. Indeed, it has been noted (for example, see [A1]) that for the project sequencing problem, the maximum size of problems, which can be reasonably solved using pure dynamic programming is in the range of 15 to 20 projects. This is consistent with the observation [A2] that the capacity expansion problem is NP-complete. Note that our model (2) of the flood control problem subsumes the project sequencing problem as a special case. Thus, it is
also important to reduce the state space of dynamic programming for the FCP by employing elimination by bounds.

An alternative to the exact DP-B&B approach is to employ heuristic procedures to solve the FCP. However, even if some heuristic algorithm results in feasible solutions for the FCP with objective values "close" to the optimal value, we have no way of guaranteeing the degree of optimality unless relaxations are also employed as in [A4]. However, if we employ the heuristic algorithm to generate a "good" incumbent feasible solution to the hybrid algorithm, then we can expect to reduce significantly the state space of the hybrid algorithm and at the same time generate an optimal solution.

Unfortunately, even the use of a "good" incumbent need not necessarily significantly improve the computational efficiency of the DP-B&B algorithm. For example, Akileswaran [A1] observed that despite the fact that excellent heuristic procedures were employed, the hybrid algorithm was only marginally superior to pure dynamic programming in terms of maximum size project sequencing problem which was solvable. The reason was that the fathoming test was not strong enough to significantly reduce the number of states considered in the algorithm. Since the incumbent solution had near optimal objective value, it is apparent that the relaxations to the subproblems associated with each state (a partial solution) must have vastly underestimated (in a minimization) the optimal completions to the partial solutions, so as to result in weak fathoming tests. In general, the fathoming test could also be weak if the incumbent solution is not necessarily a "good" solution. Note that the incumbent solution also could be updated in the
course of the DP-B&B algorithm.

In this chapter we describe a novel approach (called the "sieve" [M8]) to modify the hybrid DP-B&B algorithm, so as to efficiently generate feasible solutions with near optimal objective values and at the same time provide strong \textit{a posteriori} bounds on the optimal value. We also show how these bounds can be used to guarantee the optimality of the final solution \textit{a posteriori}. The basic idea of the sieve is quite simple. At the top of the enumeration tree (initial stages of the algorithm) many poor partial solutions (states, nodes) are not eliminated (fathomed) since the fathoming test is the weakest at the top of the tree. This, as explained earlier, is a consequence of two phenomena: firstly, the value, $u_n$, of the best solution (incumbent) found so far is the worst (highest for a minimization problem) at the top of the tree, i.e., small $n$, and secondly, the relaxations typically vastly underestimate the optimal completion to the partial solutions. The sum of these two effects is the total estimation error or "slack" in the fathoming test which allows poor partial solutions to pass the test. As we go down the tree, both of these conditions improve. Thus, the estimation errors decrease and the fathoming tests become "tighter" as we go down the tree. But, it is at the top of the tree where the algorithm's performance is affected the most since any unfathomed node (partial solution) in stage $n$, gives rise to $2^{m-n}-1$ possible descendant nodes in the FCP (where $m$ is the total number of structural measures considered). Therefore, if we could tighten the fathoming test by eliminating the slack or at least part of it at the top of the tree search, enormous portions of the tree could be pruned as shown in [M14].
This is precisely what the sieve does. We estimate the decreasing slack in the fathoming test at each level (stage) of the tree and use these estimates to tighten the tests (to "cheat" as it were). An obvious physical analogue is a sequence of \( m \) sieves (termed a sieve or sifter) whose decreasing mesh sizes estimate the slack as shown in Figure 3. We then verify the degree of optimality after the run to see if we "got away with it". We formalize the development after recalling our notation.

Section 4.1 Notation

Recall from Section 3.3, that the functional equation of the dynamic programming recursion for model (2) of the FCP is

\[
f_n(I) = \min_{i \in I} \left\{ f_{n-1}(I-i) + g(I,i) \right\}, \tag{35}
\]

where \( f_n(I) \) is the objective function value of an optimal sequencing and timing of the first \( I \subseteq I^* \) projects (structural measures) with \( |I| = n \). Thus, \( I \) is a node at stage \( n \) of the enumeration tree. The partial solution associated with node \( I \) is the optimal sequence(s) \( \{x[k]\} \in S_I \) of the projects in \( I \) and their optimal timings, \( \{t_k\}_{k=1}^n \).

Relating this to the sieve we have that the residual problem associated with node \( I \) is
Figure 3. A "Sieve"
\[ \overline{f}_{m-n}(I^*-I) = \min_{\{i[k]\} \in S} \min_{k=n}^{m} I \min_{k=n}^{m} \tau_k \leq \tau(I_{k-1}) \]

\[ \sum_{t=\tau_0}^{\tau_k-1} [P(I_{k-1}, t) - P(I_k, t)] e^{-r(t) \tau_k} + c_i[k] e^{-r\tau_k} \]

where \( \overline{f}_{m-n} \) is objective function value associated with the optimal sequencing and timings of the remaining \( m-n \) projects in \( I^*-I \).

An optimal completion of \( \{x[k]\} \in S \) \( I^*_{t_k} \) \( k=1 \) (hereafter also referred to as an optimal completion to node \( I \)) has objective functional value \( f_n(I) + f_{m-n}(I^*-I) \geq u^* \) (the value of the optimal solution to the original problem). Let \( \ell \) be a lower bound functional for \( \overline{f} \), i.e., \( \overline{f}_{m-n}(I^*-I) \geq \ell_{m-n}(I^*-I) \) for all \( I \subseteq I^* \). Typically, this lower bound functional vastly underestimates \( \overline{f} \). Let us denote this underestimation error or surplus as \( \delta^u_{m-n} \), where for any \( I \subseteq I^* \),

\[ \delta^u_{m-n}(I^*-I) = \overline{f}_{m-n}(I^*-I) - \ell_{m-n}(I^*-I). \]

Let \( u_n \) be the value of the best solution (incumbent) found to date at stage \( n \). Then, clearly

\[ u^* \leq u_n \quad n=1, \ldots, m. \]

The over estimation error \( \delta^o_n \) for any stage \( n \) is thus

\[ \delta^o_n = u_n - u^* \quad n=1, \ldots, m. \]

Note that we would expect that the underestimation slacks \( \delta^u_{m-n} \) and the
over estimation surplus \( \delta_n^0 \) both decrease (or at least do not increase) with \( n \). But the explosive growth in nodes could have devastated us by the time this total "slack" becomes small enough to allow the fathoming of "poor" solutions (nodes). This is precisely where the sieve comes in.

Section 4.2 The "Sieve"

Prior to formally introducing the sieve recall that our fathoming test is

Fathom node \( I \) at stage \( n \) if:

\[
f_n(I) + \lambda_{m-n}(I^*-I) > u_n - \delta_n^0
\]  
(37)

We have noted that this test is "loose" because both

\[
\lambda_{m-n}(I^*-I) < \bar{f}_{m-n}(I^*-I),
\]

and

\[
u_n > u^*.
\]

Adding the slack plus surplus in (37) we obtain

\[
f_n(I) + \lambda_{m-n}(I^*-I) + \delta_{m-n}(I^*-I) > u_n - \delta_n^0 = u^*
\]

or

\[
f_n(I) + \lambda_{m-n}(I^*-I) > u_n - (\delta_n^0 + \delta_{m-n}(I^*-I))
\]

\[
= u^* - \delta_{m-n}(I^*-I)
\]

Thus, if we knew or could estimate \( \delta_n^0 \) and \( \delta_{m-n}(I^*-I) \) for all \( n \) and \( I \) we could "tighten" the fathoming test as above by subtracting this total
slack from \( u_n \). This is precisely what we intend to do.

Let the \textit{total} slack \( s_n(I) \) in the fathoming test at node \( I \) in stage \( n \) be the sum of the over estimation surplus plus the under estimation slack. That is,

\[
s_n(I) = s^O_n + s^U_{m-n}(I^*-I)
\]

Furthermore, let \( s_n \) be a lower bound on the total slack at stage \( n \). That is,

\[
s_n \leq s_n(I) \text{ for all } I \subseteq I^* \text{ such that } |I| = n.
\]

Then, we could subtract \( s_n \) from \( u_n \) in all of the fathoming tests and still be assured of obtaining an optimal solution. That is, we could use the following \textit{sieve fathoming test}.

Fathom node \( I \) at stage \( n \) if:

\[
f_n(I) + s^U_{m-n}(I^*-I) \geq u_n - s_n.
\]

Note that \( \{s_n\} \) is a nonincreasing sequence. Thus, the slacks could be physically interpreted as mesh sizes in a sequence of sieves (also referred to as a sieve) as shown graphically in Figure 3.

The larger initial slacks allow poor solutions to pass through at the top of the sieve where they hurt us the most. However, if we subtract the slacks from the right-hand side of the fathoming test, we can eliminate not only these poor nodes but also all their descendants.

We will interchangeably refer to the same sieve as either a nondecreasing sequence of numbers \( \{s_n\} \) or a nondecreasing sequence of
percentages \( \{ \varepsilon_n \mid \varepsilon_n \in [0,1] \} \), where \( \varepsilon_n \) is expressed as a percentage of \( u_n \). That is, set \( s_n = \varepsilon_n u_n \).

Following the execution of a sieve, we need to guarantee the optimality of the resulting solution. If we could guarantee the optimality a priori with some sieve \( \{ \varepsilon_n \} \), we would essentially be guaranteeing the validity of the following fathoming test:

Fathom node \( I \) at stage \( n \) if:

\[
\bar{f}_n(I) + \lambda_{m-n}(I^*-I) \geq u_n (1-\varepsilon_n).
\]

That is,

\[
\bar{f}_n(I) + \lambda_{m-n}(I^*-I) + u_n \varepsilon_n \geq u_n
\]

is a valid test. Thus, by guaranteeing a priori, the validity of the sieve, we are essentially guaranteeing that \( \lambda_{m-n}(I^*-I) + u_n \varepsilon_n \) is a valid improved lower bound to \( \bar{f}_{m-n}(I^*-I) \). This could be interpreted as providing a stronger relation to the residual problem. However, our motivation for a given relaxation mechanism, i.e., a given lower bound functional, is to strengthen the lower bound test using the sieve and at the same time guarantee optimality of the final solution. Thus, it is apparent that the only alternative to do this is to use the sieve and guarantee a posteriori the optimality of the final solution. That is, if \( u \) is the optimal value using a \( \{ \varepsilon_n \} \) sieve, we would like to be able to tell if \( u = u^* \). If so, we are done. If not, then how close is \( u \) to \( u^* \), i.e., what is an upper bound on \( u - u^* \)? If this is less than some pre-specified acceptable tolerance \( \bar{\varepsilon} \), that is, \( \bar{\varepsilon} > \frac{u-u^*}{u} \), then we
are also done. If not, we would like to know how to modify our sieve \( \{ \epsilon_n \} \) so that the next run with a modified sieve \( \{ \epsilon'_n \} \) will be successful.

In the following analysis, we will make the convenient, but unnecessary, assumption that there is only one optimal solution with objective function value \( u^* \). This assumption lets us avoid the unenlightening consideration and complications in the analysis caused by the possibilities of multiple optima. Finally, we will let \( \{ \epsilon_n \in [0,1] \} \) be a non-increasing sequence.

Section 4.3 A Posteriori Results

First, we state that the obvious and not very useful a priori result.

**Proposition 4-1** \( \frac{u-u^*}{u} \leq \max \{ \epsilon_n \} \).

**Proof:** If the optimal solution is not eliminated at any stage, then \( u = u^* \) and we are done. Otherwise, some partial solution, \( I_k \) is fathomed at some stage \( k < m \). Thus, we have

\[
u^* = f_k(I_k) + f_{m-k}(I_k) \geq f_k(I_k) + \ell_{m-k}(I^*-I_k) \geq u_k(1-\epsilon_k).\]

Since \( k \) could be any of the stages from 1 to \( m \),

\[
u^* \geq \min \{ u_n(1-\epsilon_n) \} \\
\geq \min \{ u_n \} \cdot \min \{ 1-\epsilon_n \} \\
\geq u \cdot (1 - \max \{ \epsilon_n \})
\]

Thus \( \frac{u-u^*}{u} \leq \max \{ \epsilon_n \} \).
Note that this result does not use any of the desirable properties of the sieve. Furthermore, if this were the only available result, then we would always construct the trivial sieve \( \epsilon_n = \epsilon = \max\{\epsilon_n\} \). Any other choice would obviously be foolish. Fortunately, however, more powerful \textit{a posteriori} results can be obtained. For ease of notation, define \( u_k = u_n \) for \( k=n+1, \ldots, m \) and \( u = u_n \) for any sieve which terminates at some stage \( n < m \). Also we define the optimality error, \( e \), as

\[
e = \max\{0, \bar{e}\},
\]

where \( \bar{e} = 1 - (\min\{u_n (1-\epsilon_n)/u\}) \).

Firstly it is easy to establish

Proposition 4-2

\[
\frac{u-u^*}{u} \leq e.
\]

Proof: As shown in the proof of Proposition 4-1, whenever some optimal partial solution is fathomed, we have

\[
u^* \geq \min\{u_n (1-\epsilon_n)\}.
\]

Thus,

\[
\frac{u-u^*}{u} \leq 1 - \frac{\min\{u_n (1-\epsilon_n)\}}{u} = \bar{e}.
\]

Otherwise, if the optimal solution is not eliminated, \( \frac{u-u^*}{u} = 0 \). Thus,

\[
\frac{u-u^*}{u} \leq \max\{0, \bar{e}\}.
\]
We now see that the \textit{a posteriori} bound of Proposition 4-2 can be no worse (and, in fact, usually better) than the \textit{a priori} bound provided by Proposition 4-1. Whenever the optimal solution is fathomed, as seen earlier, we have

\[ u^* \geq \min_n \{ u_n (1 - \epsilon_n) \} \geq u(1 - \max_n \{ \epsilon_n \}). \]

Thus, \( \frac{u-u^*}{u} \leq \bar{e} \leq \max_n \{ \epsilon_n \} \) and hence, \( e = \max(0, \bar{e}) \leq \max_n \{ \epsilon_n \} \).

Also, observe that Proposition 4-2 provides for guaranteeing optimality \textit{a posteriori}.

**Corollary 4-1.** If \( u \leq \min_n \{ u_n (1 - \epsilon_n) \} \), the final solution is optimal.

**Proof:** By Proposition 4-2 we have

\[ u \leq \min_n \{ u_n (1 - \epsilon_n) \} \Rightarrow \bar{e} \leq 0 \Rightarrow e = 0 \]

However, we can provide stronger \textit{a posteriori} results.

**Proposition 4-3.** Denoting

\[ \bar{e} = 1 - \frac{\min_n \{ f_n(I) + t_{m-n}(I^* - I) \} | I \text{ fathomed}, n=1,\ldots,m \}}{u}, \quad \text{and} \]

\[ e = \max(0, \bar{e}), \quad \text{we have} \quad \frac{u-u^*}{u} \leq e. \]

**Proof:** If the final solution is optimal, we have error = 0 \( \leq \bar{e} \).

Otherwise, we have some optimal partial solution, \( I_k \), fathomed at some stage \( k \). Thus,
\[ u^* = f_k(I_k) + \gamma_{m-k}(I^*-I_k) \]
\[ \geq f_k(I_k) + \gamma_{m-k}(I^*-I_k) \]
\[ \geq \min\{f_n(I) + \gamma_{m-n}(I^*-I) \mid I \text{ fathomed}, \ n=1,\ldots,m\} \]

and

\[ \frac{u-u^*}{u} \leq u - \frac{\min\{f_n(I) + \gamma_{m-n}(I^*-I) \mid I \text{ fathomed}, \ n=1,\ldots,m\}}{u} = \bar{e} \leq e. \]

Also, similar to Corollary 4-1 we have

**Corollary 4-2:** If

\[ u \leq \min\{f_n(I) + \gamma_{m-n}(I^*-I) \mid I \text{ fathomed}, \ n=1,\ldots,m\}, \]

then final solution is optimal.

Also, since any fathomed state, \( I_k \), which has a lower bound equal to

\[ \zeta = \min\{f_n(I) + \gamma_{m-n}(I^*-I) \mid I \text{ fathomed}, \ n=1,\ldots,m\}, \]

was fathomed at same stage, say \( k \), we have

\[ \zeta \leq u_k(1-e_k) \leq \min\{u_n(1-e_n)\}, \]

and thus Proposition 4-3 and Corollary 4-2 provide stronger results than Proposition 4-3 and Corollary 4-1, respectively.

We note that a result similar to Proposition 4-3 has been used by Ibaraki [11] in the more general context of branch-and-bound algorithms.
Section 4.4 Comparison of Results on Bounds

We now critically compare the strengths of the bounding results presented above.

We first compare the error bounds as given by Propositions 4-1 and 4-2 (denoted as $e_1$ and $e_2$, respectively). As has been shown above, the a posteriori bound provided by the latter is no worse than the a priori bound of the former. We now discuss the magnitude of the superiority of the a posteriori bound. For convenience, we shall assume that the incumbent is not updated until the end of the last stage. That is, $u_n = u_0$, $n=1,\ldots,m$. Indeed, this procedure is employed in the hybrid algorithm for the project sequencing problem by Akileswaran [A1]. Thus, we have from Proposition 4-1 that

$$e_1 = \max_{n} \{e_n\},$$

and from Proposition 4-2 that

$$e_2 = \max(0, \bar{e}_2),$$

$$\bar{e}_2 = 1 - \frac{u_0 \cdot \min_n (1-e_n)}{u}.$$

Let $0 \leq \alpha \leq 1$ be the improvement of the final solution relative to the incumbent. That is, $u = \alpha \cdot u_0$. Thus

$$\bar{e}_2 = 1 - \frac{u_0 \cdot \min_n (1-e_n)}{\alpha \cdot u_0} = \frac{\max_n (e_n) - (1-\alpha)}{\alpha} = \frac{e_1 - (1-\alpha)}{\alpha}$$

and
\[ e_2 = \max(0, \frac{e_1 - (1-\alpha)}{\alpha}). \]

Consider the case where the final solution is not significantly better than the incumbent, i.e., \( \alpha \) is close to 1. Then we have

\[ e_2 \text{ is close } \max(e_1, 0) = e_1 \]

and thus the a posteriori bound of Proposition 4-2 is not significantly better than the a priori bound of Proposition 4-1. In particular, when the incumbent is near-optimal, the final solution cannot be significantly better than the incumbent and, thus, the a posteriori bound of Proposition 4-2 is not very different from the a priori bound of Proposition 4-1. However, when the sieve \( \{e_n\} \) is such that the final solution is significantly better than the incumbent, i.e.,

\[ \frac{e_1 - (1-\alpha)}{\alpha} \ll e_1, \]

then the a posteriori bound \( e_2 \) is much stronger than the a priori bound \( e_1 \). In particular, if \( \alpha < 1-e_1 \), we have \( e_2 = 0 \) and final solution \( u \) is guaranteed to be optimal.

We now compare the a posteriori bounds as given by Propositions 4-2 and 4-3 (denoted as \( e_2 \) and \( e_3 \) respectively). Again, as shown earlier, the bound given by Proposition 4-3 can be no worse than that given by Proposition 4-2, and we will now attempt to determine the magnitude of the improvement. We have
\[ e_2 = \max(0, \overline{e}_2), \]

where

\[ \overline{e}_2 = 1 - \frac{l_2}{u}, \quad l_2 = \min \{ u_n(1-\epsilon_n) \} \]

\[ e_3 = \max(0, \overline{e}_3), \]

where

\[ \overline{e}_3 = 1 - \frac{l_3}{u}, \quad l_3 = \min(f_n(I) + l_{m-n}(I^*-I) \mid \text{Fathomed } I, \ n=1, \ldots, n). \]

Therefore we essentially need to compare the lower bounds to the optimum value as given by the propositions, namely \( l_2 \) and \( l_3 \).

Define

\[ l_3^k = \min(f_k(I), l_{m-k}(I^*-I_k) \mid \text{Fathomed } I_k \text{ in stage } k) \]

and thus

\[ l_3 = \min \{ l_3^k \}. \]

Also define \( l_2^n = u_n(1-\epsilon_n) \). We first compare \( l_3^n \) and \( l_2^n \) for any stage \( n \).

Associated with each state, \( I^n_i \), in stage \( n \) (prior to the fathoming test), we have the lower bound, \( l^n_i \), for the state given by \( f_n(I^n_i) + l_{m-n}(I^*-I^n_i) \). Consider the sequence \( \{ l^n_i \} \) of the lower bounds arranged in the non-decreasing order. When the number of states in a stage (prior to fathoming) is sufficiently large (as is normally the case) then \( l^{i+1} - l^i \) will be fairly small for a majority of the \( i \)'s. When the
sieve estimate, $e_n$ is reasonable so that $1 \leq u_n(1-e_n) \leq i_{max}$, we have for some $i$, $i^i \leq u_n(1-e_n) \leq i^{i+1}$. Since $i_2^n = u_n(1-e_n)$ and $i_3^n = i^{i+1}$, and $i_3^n - i_2^n$ is fairly small. Also note that when $i_3^n - i_2^n$ is fairly small for each stage $n$, $i_3^n - i_2^n = \min\{x_2^n\} - \min\{x_3^n\}$ is close to $\min\{z^n_3 - z^n_2\}$ which is close to 0. Thus, $i_3^n$ is close $i_2^n$ and $e_2^n$ is close $e_3^n$. Thus, whenever the sieve estimates are reasonably within range of the lower-bounds associated with the states in each stage, we cannot expect the bound $e_3^n$ to be significantly superior to $e_2^n$. However, if for some stage, $e_n$ is such that $u_n(1-e_n) << i^i$ (i.e., all states in the stages are fathomed (prematurely terminated), we could expect $e_3^n$ to be significantly superior to $e_2^n$. Such a situation could happen with the initial sieve, when we have a 'poor' (significantly over) estimate of the slacks in the lower-bound tests in each stage. Otherwise, in general, we cannot expect the two a posteriori bounds to be significantly different from each other. Indeed this conclusion was confirmed by some preliminary computational experience.

In summary, Propositions 4-2 and 4-3 give similar bounds when the sieve terminates normally (except, of course, when no state is fathomed). The a posteriori bounds would be a significant improvement over the a priori bound (as given by Proposition 4-1) only when the final solution $u$ is a significant improvement over the incumbent $u_0$. We now apply these observations to construct suitable sieve strategies which exploit the particular characteristics of the different bound estimates.
Section 4.5 Sieve Strategies

All the above \textit{a posteriori} results allow for guaranteeing the validity of using any given sieve either through optimality or through bounds on the optimal value. However, we still need to answer the question of how a sieve should be defined so as to obtain satisfactory \textit{a posteriori} results. In particular, we have the following questions:

(1) How should the initial sieve be defined?

(2) If after the execution of a sieve the \textit{a posteriori} results do not satisfactorily demonstrate the validity of the sieve, then how should the sieve be modified so that the \textit{a posteriori} results will most likely guarantee the validity of the new sieve?

A sieve is defined to be valid if the error from optimal of \( u \) is within some prespecified tolerance, \( t \in [0,1] \).

In discussing the performance of a sieve, it is important to state the computation measure explicitly. Next we introduce and discuss a reasonable computation measure which we shall use for analyzing the performance of the sieve.

For a run of some sieve, \( S \), let \( T(S) \) be the number of reaches (state generations) performed and \( M(S) \) be the maximum number of states stored in-core in the course of the sieve. Naturally for a given relaxation \( T(S) \) is related to the total computation time of \( S \). Then define sieve, \( S \), to be \textit{feasible} if \( T(S) \leq T^* \) and \( M(S) \leq M^* \) where \( T^* \) and \( M^* \) are the appropriate upper limits for a given computing machine. In the more general context of branch-and-bound algorithms, Ibaraki [I1] considers that when two algorithms for the same problem are
feasible, the one which has less associated error from optimal is more efficient. However, there are certain difficulties associated with such a performance measure especially in the context of comparing pure hybrid DP-B&B with the sieve approach. Notice that Ibaraki's definition of computational efficiency does not consider the total computation time, \( T(A) \) related to efficiency, as long as \( T(A) \leq T^* \). We claim that the implementation, \( A \), of any algorithm (e.g., consider the hybrid DP-B&B algorithm) which has \( T(A) > T^* \) and, hence is infeasible, can be modified with relative ease, into a sequence of iterations, \( A_1, \ldots, A_n \), with \( T(A_i) \leq T^* \) for each 'i', and thus the implementation can be modified to become feasible. A justification for this is as follows. At any instant of the execution of some implementation, the information required for the remainder of the execution can be completely described by the current system state parameters (a Markovian-like property). For example, for the hybrid DP-B&B algorithm, the state parameters are the (1) stages currently reached from and to, (2) the states and their associated labels for both the stages, (3) the incumbent, and (4) the state currently being reached from. Thus during he implementation of the algorithm, whenever current time \( T \) becomes close to \( T^* \), we can dump the system description onto an external device and later continue with the implementation, after recovering the system description back onto in-core but now with \( T \) being reset to 0. The sequence of recovering until termination can be described by the sequence of iterations \( A_1, \ldots, A_n \). Notice also that the activity of dumping and recovering can be done with relative ease and minimal time. Given two algorithms \( A^1 \) and \( A^2 \) for the same problem, it thus is more reasonable to compare
the total computation time \( \Sigma T(A_1^1) \) and \( \Sigma T(A_1^2) \), rather than check for feasibility in terms of whether \( T(A_1^1), T(A_1^2) \leq T^* \). Also, since the modification involves minimal time, we have \( \Sigma T(A_1^1) \) close \( T(A) \) (the time required for original implementation). In summary, we define an algorithm to be feasible if \( M(A) \leq M^* \). Furthermore, when comparing efficiencies, two feasible algorithms \( A_1 \) and \( A_2 \) should be compared in terms of both \( T(A) \) and \( e(A) \), where \( e \) is the relative error as given by the algorithm \( A \). Also since normally one is satisfied as long as the relative error is less than some prespecified tolerance, \( t \), we redefine an algorithm, \( A \), as being feasible if \( M(A) \leq M^* \) and \( e(A) \leq t \). and if two algorithms are feasible, the one which has the lower associated computation time \( T(A) \) is more efficient.

For the flood control sequencing problem, we are comparing two fundamentally different approaches, for obtaining a given prespecified tolerance \( t \), namely, (1) using the trivial sieve, \( \{e_n = t\} \) and (2) using a sieve \( \{e_n\} \) with \( e_n > t \) for some stages \( n \). Notice that the a priori bound, \( t \), for the trivial sieve is superior to the a priori bound, \( \max \{e_n\} \), for the regular sieve. Thus, in order for a regular sieve to be effective, the a posteriori bound must be a significant improvement over the a priori bound. Consider a sieve which terminates prematurely, i.e., at some stage \( n < m \). For this premature termination to be satisfactory, we expect that the a posteriori bound is within the prespecified tolerance. That is, we have for some fathomed state, \( s^* \),

\[
1 - \frac{\varepsilon(s^*)}{u} \leq t
\]

where \( \varepsilon(s^*) \) is the total lower-bound associated with the state. Notice
that because of premature termination, \( u = u_0 \). Thus \( \xi(s^*) \geq u_0(1-t) \).

Also, since for all fathomed states, \( s \), we have (by definition) \( \xi(s) \geq \xi(s^*) \), these states would also be fathomed by the trivial sieve \( \xi_{n} = t \). Thus, whenever premature termination of a sieve is satisfactory, the trivial sieve can do no worse. Thus, in order for a regular sieve to be effective, i.e., do better than the trivial sieve, it must terminate normally, that is, at the end of stage \( m \). Furthermore, it was observed in the previous section that the \textit{a posteriori} bounds for the sieve are superior to the \textit{a priori} bound only when the final solution is a significant improvement over the incumbent. Thus, with the regular sieve, we need to exploit the \textit{a posteriori} bounds. That is, these bounds must be significantly superior to the \textit{a priori} bound, and hence, the sieve must result in a much improved solution and in particular, must terminate normally. Indeed, this was the original motivation behind the sieve.

Sieves can be essentially divided into two categories, static and adaptive, depending on the way they are defined. Static sieves are defined prior to the execution of the sieve while adaptive sieves are defined dynamically during the course of the algorithm. We shall restrict our attention to the construction of static sieves. See Marsten and Morin [M1] for computational results on the particular success of adaptive sieves in the more general context of Breadth First Search branch-and-bound algorithms.

We now address the question of how the initial sieve be defined. Our objective is to construct as "loose" a sieve as possible (i.e., large \( \varepsilon_n \)) and at the same time ensure normal termination. One way to
guarantee normal termination is by ensuring that some feasible solution or some solution superior to it does not have partial solutions fathomed at any stage.

We thus have the simple result,

**Proposition 4.4:** The sieve

\[ \epsilon_n = 1 - \frac{\varepsilon(P_n)}{u_0} \]

terminates normally, where \( P_n, n=1, \ldots, m \) is the sequence of partial solutions associated with the incumbent solution with objective value \( u_0 \), and \( \varepsilon \) is the lower-bounding function over the partial solutions.

Note that the sequences \( \{P_n\} \) and \( \{\varepsilon(P_n)\} \) can be determined prior to the execution of the sieve, and thus the sieve as given by Proposition 4.4 can be used as a suitable initial sieve. Also such a construction avoids using subjective slack estimates.

Following the termination of the initial sieve, if the final solution, \( u^1 \) is such that

\[ e_1 = \frac{\varepsilon}{\varepsilon} = 1 - \frac{\max \{\varepsilon(P)\}}{u^1} \leq t, \]

the prespecified tolerance, we are done. However, if \( e_1 > t \), we need to modify the sieve and rerun until the error from optimal is less than the prespecified tolerance.

As a special case, consider using for the second run the trivial sieve \( \{\epsilon_n = t\} \) with the incumbent initialized at \( u_0 = u^1 \) after the unsuccessful termination of the initial sieve. We now compare the two
approaches of (1) directly using the trivial sieve, and (2) using an unsuccessful initial sieve as given by Proposition 4-4 and later using the trivial sieve with an updated initial incumbent. It is possible that the initial incumbent is so weak that when directly using the trivial sieve, $S$, the number of unfathomed nodes increases drastically which results in this approach becoming infeasible because $M(S) > M^*$. However, when performing the two runs $S_1$ and $S_2$ of the second approach, the initial sieve slacks may allow for enough extra nodes to be fathomed so that $M(S_1) < M^*$, and furthermore, since for the second run, $S_2$, we have an improved initial incumbent, again enough extra nodes may be fathomed so that $M(S_2) < M^*$ and, hence, the second approach could be the only feasible procedure for obtaining the prespecified tolerance. In such cases, we could interpret the initial sieve as another algorithm used for generating a "good" incumbent in order that the trivial sieve can be feasibly implemented.

The above approach of using the trivial sieve following an unsuccessful initial sieve, can be interpreted as a particular implementation of the modified General Enumeration Method [N1, P3], where the a posteriori bound test following the termination of the initial sieve is equivalent to adding the set $W$ of withheld problems back to the set $A$ of active nodes and then performing a Selection and Elimination step.

Even if approach (1) is feasible, i.e., $M(S) \leq M^*$, the improvement in the incumbent with the initial sieve may be so significant that the number of unfathomed states may be such that $T(S) > T(S_1) + T(S_2)$, and hence the second approach could be more efficient than directly using
the trivial sieve.

As a generalization of using the trivial sieve for the second run, we can construct a sequence of sieves $S_1, \ldots, S_k$ where the solutions improve with each sieve in such a way that the final error $e(S_k)$ is within the prespecified tolerance. We now describe the considerations involved with modifying the sieves when constructing sieve $S_{i+1} = \{e_n^{i+1}\}$ from sieve $S_i = \{e_n^i\}$. For any stage $n$, where $e_n^i > t$, we need to have $e_n^{i+1} < e_n^i$ for the new sieve. If $e_n^i$ is reduced to $t$, for each such stage $n$, then we can a priori guarantee the validity of sieve $S_{i+1}$. However, such a reduction in the sieve could result in a significant increase in the number of nodes, $T(S_{i+1})$, generated. However, it is possible that with just a minor reduction in the slack estimates from sieve $S_i$ to $S_{i+1}$ (and thus with only a marginal increase in the number of nodes generated), the final solution improves significantly enough that the sieve $S_{i+1}$ becomes valid. Since such performance is strongly data dependent, it is not obvious how to develop a general modification scheme for constructing an efficient sequence of feasible sieves which terminate in a valid sieve. Furthermore, even for a particular instance, we eventually need to take calculated gambles when modifying the sieves. Formally, we need a modification scheme $: \{S\} \rightarrow \{S\}$, where $\{S\}$ is the set of all possible normal terminating sieves, so as to construct a sequence of sieves $S_0, S_1, \ldots, S_k$ with $S_{i+1} = \{S_i\}$, $i=0,1,\ldots,k-1$, and $M(S_i) \leq M^*$ for each $i$, so that $T(S_i)$ is minimal and $e(S_k) \leq t$. 
CHAPTER V

DISCUSSION

We have presented a very general model for the flood control problem and have shown that this model fits naturally into a dynamic programming framework. This model is general enough to include numerous different types of damage functions. Furthermore, we proved that if the damage functions satisfy some reasonable and fairly mild conditions, then dynamic programming provides an exact solution algorithm. We discussed different ways of implementing the algorithm to account for and exploit properties of different methods of estimating the damage functions. Finally, we addressed issues of computational efficiency and pointed out various improvements which could be incorporated into the solution algorithm.

We next consider applications. A specific application of our solution approach is discussed in detail in Appendix A. In particular we indicate how our approach can be easily implemented even using hand-calculations. That is, the computational requirements are modest compared to the alternate solution approach originally used by Cortez-Rivera [C5] to solve this problem. Although the effects of interdependence between the structural and nonstructural measures has
been addressed qualitatively [W5, W6] as demonstrated in Figures 1 and 2, and data exist for specific applications -- see [C5, W2] for example, to date there appears to be no readily available general quantitative methods for estimating the damage functions which comprehensively take into account different types of structural and nonstructural measures. Thus, we might critically ask if our model imposes excessive information requirements on the user when it requires such damage functions to be constructed. The answer is, of course, that if detailed quantitative decision making support is required then it is clearly essential that these damage functions be specified. However if less detail is required, then it may be more reasonable to aggregate the effects of the nonstructural measures either in the constraints or in the damage function and explicitly only consider the structural measures as decision variables. We could then use the capacity expansion approach [M6] to determine the sequencing of the structural measures. Once the sequencing has been determined, it may then be possible to assign different levels of the nonstructural measures using the decision maker's qualitative understanding of the effects of the nonstructural measures. However, not only would such an approach be heuristic, but it also would completely neglect the significant interdependencies and interactions which are inherent to the problem. We could, therefore, recommend our general model and solution algorithm for all except the most preliminary stages or purely qualitative analysis.

5.1. Recommendation for Future Research

In the framework for the flood control problem which was presented in Chapter II, we referred to the fact that flood control is in general
beneficial to more than one party or sector of the economy. In particular, damage reduction is beneficial to both the government and private parties who are directly affected by the flood control program. Furthermore, it is the private citizens who ultimately bear the burden of tax supported governmental flood control and relief programs. If the benefits of flood damage reduction are fairly significant to more than one sector of the economies, it may be more appropriate to model the problem as a multi-decision maker problem [N2] or a multi-objective problem [C3].

In Chapter II we also mentioned that since water resource development is typically multiple purpose in nature, damage reduction is only one of the many benefits of a flood control scheme (see also [G5, U1]). Our model would be most appropriate to those cases where damage reduction is the main purpose of the flood control scheme and other benefits are secondary or if the other benefits did not vary significantly under different flood control schemes. Indeed, Arvanitidis et al. [A10] describe a flood control application where damage reduction accounted for 80% of the total benefits. In instances in which flood damage reduction is not the sole or primary purpose, it would be more appropriate to model the problem as a multicriteria problem. We note that our model and solution approach could be extended to handle such multicriteria extensions.

We also note that our model does not explicitly account for the uncertainty involved in estimating the different flood control benefits. As previously mentioned, estimating flood damage functions could be a significant task -- one reason for this could be the uncertainty
involved in estimating benefits. For example, the actual damages could be different for different flood scenarios and it may be more reasonable to construct damage functions for these scenarios individually rather than an average expected scenario. In such cases, if probabilities could be assigned to each scenario, it may be appropriate to construct utility functions which thus account for uncertainty as well as measure damage effects.

An additional consideration in applying the model, results from the fact that it implicitly assumes that the decision maker can arrive at a suitable rate "r" to discount the different benefits in order to make a meaningful comparison. Just what is an appropriate discount rate for the public sector has been the topic of many research efforts [M4, M5, B3]. Furthermore, we note that whereas the solution to an all structural measure sequencing problem would not be dependent on this rate so long as it remained constant, the discount rate could have profound effects on the solution of the overall problem involving both structural and nonstructural measures.

In conclusion, we have formulated a very general model and developed a solution algorithm for the flood control problem. However, much meaningful work remains. This involves testing and evaluation on specific real problems as well as extensions of the approach to account for multiple conflicting objectives and risk and uncertainty. These are precisely the issues we are currently addressing.
REFERENCES


Belvoir, VA, February 1972.


M3. McNealy, J.G. and R.D. Lacewell, Flood Plain Management, The Texas Agricultural Experiment Station, Texas A&M University, College Station, TX, November 1976.


M8. Morin, T.L., "The Sieve or 'How to Cheat and Get Away with It'," Working Paper, School of Industrial Engineering, Purdue University, West Lafayette, IN, June 1979.

M9. Morin, T.L., IE 633 (Dynamic Programming) Lecture Notes, School of Industrial Engineering, Purdue University, West Lafayette, IN, Spring 1981.


APPENDIX A

Application to a Real-World Problem

This appendix presents details of the application of the solution methodology of Chapter III to a real-world problem. Specifically, we compare our solution approach to a conventional dynamic programming (DP) approach with time periods as stages used by Cortes-Rivera [C5] for the Embarras River basin located in South-Central Illinois. We first describe the structural and non-structural measures considered, damages specification, and the constraints involved. We then briefly describe Cortes-Rivera's DP solution approach. This is followed by a brief description of our DP approach. In particular, the state spaces and stages for both approaches are compared in Table 1. Hand-calculations are all that is required by our approach on this application as opposed to having to resort to the computer as required in the Cortes-Rivera approach.

Description of the Problem

The structural components are:

1(a) A Flood Detention Reservoir which can be one of seven different sizes.

(b) The land required to be purchased for constructing the
Table 1. Comparison of the number of states required by two different dynamic programming approaches for a particular flood control problem

<table>
<thead>
<tr>
<th>Stage</th>
<th>Our Approach</th>
<th>No. of states (worst case)</th>
<th>Cortes-Rivera's Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>255*</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>756</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1,039</td>
<td>Total</td>
</tr>
</tbody>
</table>

* 255 = 7 × 9 + 7 × 12 + 9 × 12

† 7125 = (7 + 1 + 7) × (9 + 1 + 9) × (12 + 1 + 12)
reservoir with different amounts for different sizes.

2(a) A levee (denoted Levee-3) which can be one of nine different heights.

(b) The purchasing of land that goes along with them.

3(a) Another levee (denoted Levee-4) (at a different location from the former) which can be one of 12 different heights.

(b) The purchasing of land that goes along with them.

The nonstructural measures are agriculture land-use distribution in terms of allocation of land for (1) corn, (2) soybeans and (3) pasture.

After incorporating various hydrological considerations and performing linear programming optimizations over the agricultural land-use measures, Cortes-Rivera provides the function $P(I,t)$ where $P$ is the expected annual net income in year $t$ for a given combination $I$ of the structural measures. Other considerations involve (1) a rate of Land-Value Differential Inflation different from the normal rate of discount, (2) budget constraints in the form of a given budget supply schedule, and (3) institutional constraints in the form of minimum time interval required between the construction of the two levees. The problem also provides for changes in levee height once a levee has been built. The objective is to determine which of the structural measures are to be built, and if they are built, when and at what rates should they be built. Furthermore, if there are to be changes in levee heights with time, when and how much are to be determined. The optimal land-use allocation in each year has already been determined in the process of determining the $P$ function. All the above sequencing and timings of
the structural measures are subject to the constraints referred to earlier.

Cortes-Rivera's Solution Approach

The planning horizon (50 years) is split into 10 subperiods, five years each, and thus construction decisions are to be made only on a subperiod basis. A DP approach is taken with each stage corresponding to a subperiod. Each state in any stage is a 6 dimensional vector with attributes:

1. Storage capacity of the detention reservoir.
2. Elevation of Levee-3.
4. Land required for the construction of the detention reservoir.
5. Land required for the construction of Levee-3.

The decision variables for each state are represented by the magnitude of the enlargements or expansions of each of the structural components and by the additional land which is necessary to acquire for the construction of each structural components. Thus there are six decision variables, in one-to-one correspondence with the six state variables.

Since the state space turns out to be large (see Table 1), the author has to resort to an approximate algorithm called differential discrete DP. This disadvantage is in particular absent in our proposed solution approach described below.

Our Solution Approach
For now, we shall add an additional condition that once a levee has been built to a particular height, its height cannot be changed. We assume this since it provides for a much simpler solution procedure. In our DP approach, a stage corresponds to the construction of a structural measure. Thus, we only have three stages with each state in stage 1 being some structural measure at some capacity, states in stage 2 being all possible (unordered) pairs of structural measures and finally the states in stage 3 being all possible (unordered) triplets of the structural measures. The decisions associated with each state in stages 1 and 2 are the construction of some additional structural measure.

Notice that the timing and the amount of land purchased for any structural measure can be incorporated into the cost of the measure in the form of

$$C_i(I, \tau) = C_i + C_{\text{land}} \min_{B(I, i) \leq t \leq \tau} \{(1+d)^t \cdot e^{-r t} \} e^{rt}.$$ 

where $I$ is the resulting set of structural measures, $i$ is the additional measure considered, $d$ is the value of Land-Value Differential Inflation, $C_{\text{land}}$ is the cost of the land required for measure $i$, $C_i$ is the cost of the measure, $r$ is the discount rate, and $B(I, i)$ is the minimum possible time for purchasing as determined by the budget supply schedule and the current budget consumption. Notice that the optimization can be trivially solved since if $r < \log_e(1+d)$, then $t^* = B(I, i)$ and if $r \geq \log_e(1+d)$, $t^* = \tau$.

The return function associated with the addition of $i$ to form $I$ is
\[ g(I,i) = \min_{i} \sum_{t=0}^{\tau-1} \left[ P(I,t) - P(I-i,t) \right] e^{-rt} + C_i(I,t) e^{-rt} \]

Since \( P \) has already been determined, it is fairly easy to obtain the return function as defined above.

The particular advantage of our approach to that of Cortes-Rivera is demonstrated in Table 1. Hence the number of states and stages required by both the approaches are compared. We also note that by exploiting properties as defined by equation (31), we could eliminate even more states both explicitly and implicitly.
APPENDIX B

An Approach for Constructing Comprehensive Damage Functions

As observed in Chapter II, there is a lack of general quantitative data on the interaction effects on damage reduction from measures such as Warnings, Relief, etc. The only available information on such damage reduction effects are the qualitative descriptions of some interactions by White [W5, W6]. We now describe an approach for constructing comprehensive damage functions from such qualitative descriptions by formulating a damage function for a hypothetical floodplain.

We use the following assumptions on the interdependence among the various measures.

- The nonstructural measures do not significantly alter the routing characteristics of a combination of structural measures.
- The hydrological influences of any combination, 'I' of structural measures can be effectively described by the resulting expected (1) depth of flooding, $d(I)$, (2) intensity of flood, $v(I)$, and (3) the duration of flooding, $u(I)$. We also assume that with the help of hydrological evaluations we can determine these three parameters for any combination, $I$, of the structural measures.
The floodplain we are considering is divided into three regions based on their distance from the river reach. This is shown in Figure 4.

- The hydrological effects of a flood are uniform over a particular region but could vary over the regions. In particular, we assume that the depth of flooding is different for each of the three regions and thus, let \( d(I,j) \) be the depth of flooding in region \( j, j = 1, 2, 3 \).

- The land use measures we are considering are 1) residential development, 2) agriculture, and 3) recreational use. We can allot areas of land in each of the regions for each of the activities. We shall assume that these land use measures do not interact with each other as regards damage reduction.

- Besides these land use measures, we also shall consider using a Warning system and thus have to determine an appropriate level for this nonstructural measure.

- The total damage reduction is a result of a combination of different damage reduction effects. We shall assume that keeping all other damage reduction effects constant, the total damage changes in proportion to the level of a particular damage reduction effect.

With this assumption on the behavior of the damage reduction effects, we now attempt to formulate an overall damage function which comprehensively takes into account various interdependence effects. Our formulation is similar in spirit to the approach of James [J3] where effects of flood-proofing are incorporated in structural measures planning. The interdependence effects are expressed in the form of
"reduction factors".

- Damages increase linearly as distance from the river diminishes.

Thus, let \( f(j) \) be a measure of the distance of region 'j' from the river react. Thus we have

\[
D(y_{ij}) \text{ a linear function of } (1 - \frac{f(j)}{f^*}) \quad \text{where}
\]

\( y_{ij} \) is the area allotted to land use activity 'i' in region 'j'; \( D(y_{ij}) \) is the damage incurred as a result of this activity; \( f^* \) is a suitably high value.

\[\text{Figure 4. A Hypothetical Flood Plain.}\]
• Damages as a result of the intensity of flood decrease linearly with the level of the Warning scheme.

Thus

\[ D(y_{ij}) \text{ a linear function of } v(I)\left(1 - \frac{y_4}{y^*_4}\right) \quad \text{where} \]

\[ y_4 \text{ is the level of the Warning scheme, and } y^*_4 \text{ is a suitably high value.} \]

• Warning schemes are more useful in terms of damage reduction to residential development than to agricultural use.

Let \( w(i) \) be a measure of any given level of Warning scheme's relative effectiveness for land use 'i'. \( 0 \leq w(i) \leq 1 \quad i=1,2,3. \)

Thus, we have \( w(1) > w(2) \) and

\[ D(y_{ij}) \text{ a linear function of } (1 - \frac{y_4}{\bar{y}_4})(1 - w(i)) \]

where \( \bar{y}_4 \) is a suitably high value.

• Damages vary in proportion to the area allotted for a land use activity.

Thus \( D(y_{ij}) \) is a linear function of \( y_{ij} \).

Thus, overall we have

\[ D(y_{ij}) \text{ a linear function of } y_{ij}\left(1 - \frac{f(j)}{f^*}\right)v(I)\left(1 - \frac{y_4}{y^*_4}\right) \]

\[ (1 - \frac{y_4}{\bar{y}_4})(1 - w(i)) \frac{d(I,j)}{d^*} \frac{u(I)}{u^*} \]

where the last two terms represent the effects of depth and duration of
flooding respectively and \(d^*\) and \(u^*\) are suitable constants. We can express this as

\[
D(y_{ij}) = K \cdot F(y_{ij}, Y_4, i, j, I)
\]

Thus the total damages are

\[
D(I, y) = K \sum_{i=1,2,3} \sum_{j=1,2,3} F(y_{ij}, Y_4, i, j, I).
\]

- Land unused in the floodplain has to be purchased at rates varying over the regions.

Thus let \(y_{5j}c_j\), \(j=1,2,3\) represent the area purchased and the unit cost of purchasing respectively for each one of the three regions.

Thus, the total cost of the nonstructural measures is of the form

\[
c(y) = \sum_{j=1,2,3} (y_{5j}c_j + \sum_{i=1,2,3} y_{ij}c_{ij}) + c_4Y_4
\]

where \(c_{ij}\) is the cost of implementing land use activity \('i'\) in region \('j'\); and \(c_4\) is the cost rate of implementing a Warning system.

Thus, \(P(I)\) is of the form

\[
P(I) = \min \left\{ \left[ K \sum_{i=1,2,3} \sum_{j=1,2,3} F(y_{ij}, Y_4, i, j, I) \right] + \right. \\
\left. \sum_{j=1,2,3} \sum_{i=1,2,3} (y_{5j}c_j + y_{ij}c_{ij}) + c_4Y_4 \right\}
\]

s.t.
\[ \sum_{i=1,2,3} y_{ij} + y_{5j} = A_j \quad j=1,2,3 \]

\[ 0 \leq y_4 \leq y_4^{\text{max}} \]

where \( A_j \) is the area available in region \( 'j' \) and \( y_4^{\text{max}} \) is the maximum possible level of the Warning scheme.

Note that in our assumptions that the damages are proportional to the level of some damage reduction effect may, in general, be unreasonable for an application. However our aim is to show how, with the help of some assumptions, we can meaningfully translate qualitative information on the interaction effects into a damage function. This approach was motivated in part to the availability of only qualitative information on the general interaction effects.