WRC RESEARCH REPORT NO. 89

OPTIMAL REGIONALIZATION OF WASTEWATER TREATMENT FOR WATER QUALITY MANAGEMENT

by

Lewis A. Rossman
and
Jon C. Liebman

University of Illinois at Urbana-Champaign
Urbana, Illinois

FINAL REPORT

Project No. A-063-ILL

The work upon which this publication is based was supported in part by funds provided by the U. S. Department of the Interior as authorized under the Water Resources Research Act of 1964, P. L. 88-379
Agreement No. 14-31-0001-4013

UNIVERSITY OF ILLINOIS
WATER RESOURCES CENTER
2535 Hydrosystems Laboratory
Urbana, Illinois 61801

July 1974

ABSTRACT

OPTIMAL REGIONALIZATION OF WASTEWATER TREATMENT
FOR WATER QUALITY MANAGEMENT

A mathematical decision model is developed which determines how a group of waste dischargers should regionalize their treatment facilities and the amount of treatment each facility should provide so that the cost of achieving a specified water quality goal is minimized. The waste dischargers are assumed to lie in a linear configuration along (or on both sides of) the river and several other regionalization restrictions are imposed. Treatment plant and piping costs as functions of wasteflow can be of any form and may include fixed costs. The model is solved by using a dual approach to nonlinear programming and is applied to data from the Delaware Estuary. The results compare favorably with previous regionalization schemes.

The model is extended to consider branched systems and the use of bypass piping. Two additional minimum cost, regional wastewater treatment models are developed; one which finds the regional treatment facility pattern when degree of treatment is fixed and another which finds the regional facility pattern and uniform level of treatment for all facilities so that a water quality goal is met.

Rossman, Lewis A. and Jon C. Liebman
OPTIMAL REGIONALIZATION OF WASTEWATER TREATMENT FOR WATER QUALITY MANAGEMENT
Final report to the Office of Water Resources Research, Department of the Interior on Annual Allotment Project A-063-ILL, June 1974.

KEYWORDS: wastewater treatment/regional systems/mathematical model/
cost minimization/water quality/nonlinear programming/optimization

ACKNOWLEDGEMENT

This report represents the thesis of Lewis A. Rossman, submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Civil Engineering at the University of Illinois at Urbana-Champaign, 1974.

The authors wish to thank Drs. Mohamed I. Dessouky, Richard S. Engelbrecht, Gary L. Hogg and John T. Pfeffer for their helpful comments. The input provided by Drs. E. Downey Brill, James W. Male, E. Earl Whitlatch and Mr. Makram T. Suidan is gratefully acknowledged.

The work upon which this report is based was supported, in part, by funds provided by the United States Department of the Interior as authorized under the Water Resources Research Act of 1964, Public Law 88-379, and in part by an award from the Granite City Steel Scholarship Fund, administered by the University of Illinois.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

1.1 Preliminaries

The commitment of our nation to improving the quality of our rivers, lakes, and streams is evident in the huge expenditures, both public and private, made for pollution control activities. To insure that our resources are directed most efficiently in this effort, a sound program of planning of pollution control projects is required. In its guidelines for water quality management, the U.S. Environmental Protection Agency (1971) requires that this planning be done on a regional, basin-wide basis with the goal of maximum cost effectiveness in meeting desired objectives.

This goal of cost effectiveness has led to the development of mathematical decision models to aid the basin planner in selecting the proper pollution control program from the vast number of alternatives available. These models make many simplifying physical and economic assumptions and are best viewed as information generators or screening devices in the planning process. Typically, they determine the amount of waste reduction each point discharger should provide so that a specified water quality can be maintained at minimum cost. Different effluent discharge control policies, such as uniform treatment, zoned uniform treatment, and effluent charges can be examined. A critical review of some of these models is given in Enviro Control, Inc. (1971).

As an additional alternative to pollution reduction at its source the use of regional treatment facilities can result in potential cost savings. This is because there are economies of scale in the costs of building and operating wastewater treatment facilities. However, there are two additional costs which may offset these savings. One is the cost of piping to the regional facility; the other is the possible cost of having to provide higher levels of treatment at the regional facility. This may be necessary because with regionalization more waste is concentrated and released at a given point on the river, imposing greater stress on water quality. The economies of scale also present problems in finding globally optimum solutions since they lead to nonconvex functions. If fixed costs for establishing treatment plants and pipelines are to be considered, the problem is made more difficult.

Conceptually there are two sets of decisions to be made in constructing a water quality management decision model which considers regionalization. The first, which may be termed the facility location decisions, determines where regional treatment facilities are to be built and the assignment of waste sources to them. Here the cost tradeoffs between economies of scale and piping come into play. The second, which can be called the degree of treatment decisions, determines what degree of waste reduction each facility (regional and nonregional) should provide so that a given water quality is maintained. Here the cost tradeoffs are between the economies of scale and the required treatment level. The optimal sets of decisions are the ones which minimize total cost. Note that the two sets are not independent of each other since the regionalization pattern from the facility location decisions provides the input for the degree of treatment decisions.

It is the purpose of this study to combine the above considerations into a mathematical decision model, the solution of which yields the minimum cost regionalized waste treatment plan for a river with a specified water quality goal. In addition, the model will be evaluated for its utility in river basin planning.

1.2 Literature Review

Although a large body of work in the mathematical modeling of water quality management has been produced in the past ten years, few quantitative studies have been made of the regionalization of wastewater treatment facilities. One type of formulation of the regionalization problem ignores the question of receiving stream quality and centers on the facility location decisions; that is, on finding the number and location of regional plants and the assignment of waste sources to the plants so that total cost is minimized, under an assumption about required degree of treatment. Even though the present work will go beyond this to consider water quality, this formulation is of interest because it serves as a subproblem which is solved repeatedly in the solution method to be described later on. What follows is a review of some of the approaches taken to solving this more restricted problem.

The assumptions made are that the location of sources and their waste flows are fixed in advance as are the regional treatment plant locations and the allowable pipeline routes. The costs of treatment plants and pipelines are functions only of waste flow handled and exhibit economies of scale. This implies concavity when the functions are continuous. Deininger and Su (1971) used the fact that since the cost functions are concave the optimum solution must occur at an extreme point of the constraint set (which is linear). Using a piecewise linear approximation of the cost functions and an algorithm that ranks the extreme points, they solved a hypothetical problem with seven waste sources located on a single branched network configuration. Solution time was about 5 seconds on an IBM 360/67.

Bhalla and Rikkers (1971) presented a heuristic technique for solving the regional plant location problem as part of an effort to plan the capacity expansion of regional systems. At each stage in their algorithm, the facility which can serve all unassigned sources most cheaply is identified. Then for each location where no facility currently exists, the subset of the unassigned sources which presents the most savings when served by a facility at this location is found. The location (and corresponding assigned sources) with the greatest savings is added to the solution and the process is repeated. Additional rules are given for dropping facilities from the solution and sending their sources elsewhere.

Another heuristic was proposed by McConagha and Converse (1973). Feasible solutions are successively improved by searching for the least costly option available to each source, one at a time (i.e., treat the source plus accumulated piped-in waste, send it to another location, or keep the existing solution). An optimal solution was actually obtained for all problems tested. The solution of the Deininger and Su example was found in 10 percent of the time required by the Deininger and Su algorithm.

Meier (1971) proposed a branch and bound procedure which he claims obtained optimal solutions. A branching occurs at some source k in the set of sources not yet assigned to treatment facilities. One regional facility will be added for the sources in this set. The branches correspond to the decisions to build the regional plant at k, build a nonregional plant at k or build no plant at k with the regional plant built somewhere else. The least cost assignment of sources for each case is found by enumeration. If the lowest cost decision involves building a plant at k, then k and sources assigned to it are added to the solution. The second lowest cost is used as a bound to cut off further branching. A 10 source problem was solved on an IBM 360/85 in about 10 seconds.

Converse (1972) treated a more simplified version of the regionalization problem in that the sources are assumed to lie in a linear configuration. A dynamic programming procedure was used where the stages are the number of plants built, the decisions are where to add an additional plant between a group of consecutive sources (found by enumeration) and the state is the number of source locations away from the last source. An additional computation optimizes the tradeoff between pumping head and sewer diameter so that velocity constraints are met in all pipelines. An 18 source example based on data from the Merrimack River basin was solved in 15.4 seconds (the computer used was not stated). In Section 4.4 of the present study an alternative dynamic programming formulation of this problem is given.

Wanielista and Bauer (1972) and Joeres et al. (1974) formulated a regionalization problem where a limit is placed on the quantity of flow (but not the mass of pollutant) which can be discharged at any plant location. A network of allowable piping connections between sources and facilities is established and all cost functions are made piecewise linear with fixed costs. The resulting mixed 0-1 integer program is solved to find the minimum cost pattern of treatment facilities and source assignments. In an application to data from the Little Econ

River basin, Wanielista and Bauer (1972) solved a 12 node, 19 arc network problem in 0.36 minutes (computer not specified). Jöeres et al. (1974) solved a problem with 12 possible treatment plants and 20 possible pipeline routes in 14.4 minutes on a Univac 1108.

A second type of research investigates the effects of prescribed regionalization patterns and treatment levels on water quality. Yao (1973) simulated the effect on stream dissolved oxygen due to several regionalization and BOD removal schemes for a portion of the Connecticut River basin. He noted that with secondary treatment a centralized facility on the main stem would result in better water quality than a decentralized system due to the larger dilution flow available along the main stem. Adams and Gemmell (1973) used a deterministic and stochastic simulation to observe the effects of decentralizing treatment plant BOD discharges on stream dissolved oxygen. A hypothetical river was used with "nominal" values for its parameters and the treatment plants were distributed uniformly over its length. The results from the deterministic analysis showed that the minimum dissolved oxygen in the stream improved with increasing disaggregation up to about 8 plants after which improvement was negligible. The stochastic analysis considered the important question of the effect of effluent variability from treatment plants on water quality. The variance in effluent quality was inversely related to plant size. As expected, the minimum oxygen levels were lower and occurred more frequently for the highly regionalized system.

The third type of study considers both the facility location decisions and the degree of treatment decisions in finding the minimum cost regionalized waste treatment plan to meet quality standards. Water quality is measured by dissolved oxygen, and waste reduction by

BOD removed. The present research falls in this category. Haimes et al. (1972) presented a model which considers a single regional plant location and determines which waste sources ship to the regional plant and the degree of BOD reduction required by all plants so that a specified dissolved oxygen standard is met at minimum cost. The actual problem context is one of finding the pollution taxation scheme which achieves the above. The formulation is a mixed 0,1 nonlinear program. It is solved by forming a Lagrangian with respect to the water quality constraints and then searching for a saddle point for this Lagrangian. The model was applied to data from the Miami River basin. There were 15 sources and 27 river reaches considered. An optimal solution which called for combining 12 of the sources at the regional plant was obtained in 33 seconds on a Univac 1108. This gave a savings of $2305 per day in comparison to the least cost at-source (nonregionalized) treatment solution. The possibility of no Lagrangian saddle point existing was not considered in the paper and is explored more thoroughly in the present work.

Whitlatch (1973) allows for more than one regional plant but assumes that the set of sources and regional plant locations (which may coincide) lies in a linear configuration along the river. Sources are allowed to pipe around a limited number of others to reach regional plants but cannot split their flow between different treatment plants. Treatment costs with respect to plant size and piping costs can be of any form while treatment costs with respect to BOD removed must be convex. A two-phase heuristic procedure is given which seeks to find the regionalization pattern and necessary degree of treatment to meet a dissolved oxygen goal at minimum cost. In Phase I, BOD removal efficiencies are fixed at 85 percent. Allowing piping in only one direction, two dynamic programs are solved (one for piping upstream and another for piping downstream) to obtain the optimum regionalization pattern under this restriction. Then by a set of heuristic rules these solutions are combined and successively modified so that reduction in costs is always obtained. In the Phase II calculations each regionalization modification is accompanied by a linear programming calculation of the minimum cost degree of treatment to meet the oxygen standards. The approach was applied to data from the Delaware Estuary. The 72 miles of river had 22 domestic waste discharges which were allowed to regionalize at 9 potential locations and in addition there were 22 industrial polluters whose optimal degree of BOD removal was also to be determined. To meet a dissolved oxygen goal of 3 mg/ℓ with required primary treatment, the results showed that the regionalized solution was 35 percent cheaper than the at-source solution. Total solution time was 15.9 minutes on an IBM 360/91.

The most comprehensive optimization models developed to date in this area are those of Graves et al. (1970, 1971) and Pingry and Whinston (1973, 1974). A potential network of piping between waste sources, regional treatment plants and river reaches is specified and the optimal waste flow assignment through this network plus the level of BOD removal at each treatment plant is determined so that a dissolved oxygen standard is met. Split flows and bypass piping of waste discharges to other reaches are allowed. However, all cost functions must be continuous. Treatment plant costs are nonlinear functions of the size of the plant and percent of BOD removed, while piping costs are functions of flow. Conservation of flow and BOD around each node of the piping network leads to a set of linear equality constraints.

In Graves et al. (1970) the Thomann (1972) BOD-DO model for segmented estuaries is used to construct quality constraints. Although the original model is linear in amount of BOD discharged to a given reach, expressing this quantity in terms of the network flow and BOD removal leads to a nonlinear set of inequalities. In Graves et al. (1971) the Streeter-Phelps equations are applied sequentially over each river reach to constrain dissolved oxygen deficit. Nonlinearities are introduced in expressing BOD discharged as described above and in considering time of flow and the reaeration coefficient to be nonlinear functions of river flow which is itself a function of regional plant waste flow discharges. This type of treatment allows flow augmentation to be considered as another decision alternative. In Pingry and Whinston (1973, 1974), the effect of heated effluent discharges on the river temperature and dissolved oxygen is added to the problem. The use of cooling towers is introduced as a decision alternative for controlling this effect.

All of the above nonlinear programming models are solved using a linearization algorithm based on the feasible directions approach. From a given solution a direction of search is found by solving a linear program derived from a first order Taylor Series expansion. The step size for this direction is then chosen from quadratic approximations so as to give the greatest improvement while maintaining feasibility.

Notable features of this algorithm include parametric adjustments of the error term in the Taylor series expansion to maintain consistency in the linear program and insure a gain in the optimization, and the use of priority classes of variables (a form of restriction strategy) to reduce computational effort. For nonconvex problems only local optimality can be guaranteed.

The model formulated in Graves et al. (1970) was applied to the Delaware Estuary problem described above. Piping was allowed between each of the 22 domestic waste sources and the 9 regional treatment plant locations except where the river would be crossed and between each of the 44 total sources and each of the 30 estuary reaches. The resulting program had over 2,000 variables and 80 constraints. The solution for an oxygen goal of 3 mg/ℓ and no requirement of primary treatment utilized 3 regional plants serving a total of 10 sources and bypass piping by 2 industrial sources. The optimal cost was $\$2.292 \times 10^6$ per year as compared with $\$4.1 \times 10^6$ per year for the least cost at-source treatment solution. Solution time was about 10 minutes (computer not specified).

The models in Graves et al. (1971) and Pingry and Whinston (1973, 1974) were applied to the West Fork White River. There were 11 waste sources and 46 river sections in the analysis made by Graves et al. (1971). Regional plant locations corresponded to each river section. Each source could pipe to any location within 25 river sections up or down river. There was a potential reservoir available to provide flow augmentation. A DO goal of 5 mg/ℓ in all reaches was sought with a minimum of 85 percent BOD removal required at all plants. The resulting program had 1880 variables and 138 constraints. The solution used 3 regional plants partly serving a total of 7 sources and employed 100 cfs of augmentation. Its cost is half a million dollars cheaper than a policy of uniform 98 percent removal, which would not be sufficient to meet the standard.

For the Pingry and Whinston (1973, 1974) study there were 13 BOD sources, 3 waste heat sources, and 62 river sections. Four different DO goals were used. All solutions employed one regional plant serving 2 sources, no flow augmentation and a cooling tower at one of the thermal pollution sources.

An alternate formulation given in Graves (1972) parallels the 0-1 mixed integer approach of Wanielista and Bauer (1972) and Joeres, et al. (1974) described above. Instead of constraints on the amount of flow which can be discharged from any location, the DO constraints as given in Graves et al. (1970) are used. All costs consist of a fixed portion and a linear variable portion. A solution algorithm is suggested for the resulting mixed integer program based on a generalized Bender's decomposition. No numerical applications are given.

1.3 Objectives

The above literature review indicates that the problem of determining the least cost regionalized waste treatment plan to meet a water quality goal for a river is not well solved. Of the two existing models actually applied to problems with more than one regional facility, one presents only a heuristic solution technique while the other provides a local optimum finding technique which requires continuous cost functions (i.e., no fixed costs). Both models require lengthy solution times for problems with only a small number of regional facility locations.

It is the objective of the present research to develop a new mathematical decision model for regionalizing waste treatment efforts in a river basin subject to water quality goals. The only restriction placed on the form of the cost functions is that they be convex with respect to degree of treatment. As shown above there is considerable difficulty in solving the facility location problem even when water quality is ignored. The approach to be taken in this research is to simplify some aspects of the location problem so that optimal solutions can be obtained in a manner similar to that in Converse (1972). The decisions on treatment levels for each facility, which are related to one another through the water quality constraints, are decoupled so they can be examined independently. The goal is to produce a solution algorithm to a problem not quite as general in scope as, say, the Pingry and Whinston model but one that still has the ability to consider a large number of regional arrangements and find the globally optimum one in a reasonable amount of time. In addition, as improved techniques are developed for solving the more general location problem the methodology developed herein can be readily applied so that a more comprehensive model can be constructed.

The model will be evaluated for its effectiveness as a tool in river basin planning. It will be applied to data from the Delaware Estuary and the results will be compared with previous work. Its behavior under varied inputs will also be examined. In addition it will be shown how the model can be extended so that several of its simplifying assumptions can be dropped and so that other decision problems in regional waste treatment facility planning can be solved.

CHAPTER 2. MODEL FORMULATION

2.1 Problem Statement and Assumptions

The problem under consideration is how a group of waste dischargers along a river should plan and construct a regionalized system of treatment facilities so that a water quality standard is met at minimum cost. To be decided are the number and location of treatment plants, the assignment of waste sources to each plant, and the level of treatment to be provided by each plant. The total cost of the system as well as the resulting water quality are related to these decisions.

There are two contexts in which the problem can be viewed. The first ignores all current treatment and assumes that facilities are built from the ground up. This context would probably be most suitable for long range studies. The second considers what additional level of treatment should be given above that already being provided. A different set of treatment costs would be associated with each context.

In order to convert this rather general problem statement into a precise mathematical programming model which can be efficiently solved several assumptions will be made and some restrictions added. Although water quality is measured by a number of physical, chemical, and biological characteristics, only the carbonaceous biochemical oxygen demand (BOD) of wastewater and its effect on stream dissolved oxygen (DO) will be considered. Dissolved oxygen has long been recognized as a measure of the overall "health" of a stream. Adequate levels of DO are necessary to support the natural aquatic life in a stream and to maintain aesthetically pleasing conditions. Mathematical models which predict the effect of BOD discharges on DO (and thus determine the assimilative capacity of the water) have been developed and used frequently in water quality management (Streeter and Phelps, 1925; Dobbins, 1964; O'Connor, 1960; and Thomann, 1972). Other wastewater constituents can be accounted for by fixing their allowable discharge levels in advance. The methodology presented could, of course, be used with any other quality measure, or, at least in theory, with any group of quality measures.

Figure 2.1 shows an idealized river with N waste dischargers, or sources. By convention they are numbered starting at the upstream source. It will be assumed that the sources lie in a linear configuration down the river or that a single pipeline could be drawn which links the first source with the last and with all those in between. This assumption helps simplify the plant location aspects of the problem and, in effect, restricts the model to unbranched rivers. For rivers with sharp bends some degree of approximation may have to go into the selection of the pipeline route. For wide rivers where piping across the river may be prohibitive the model can consider two linear source configurations, one for each side.

Each waste source j produces a waste flow of q_j mgd and currently discharges a BOD of s_j lb/day. Depending on whether s_j represents the source BOD or the BOD after some existing level of treatment determines what context the problem is viewed in. Each source is considered as a potential location for a regional plant. Additional locations can be added by using dummy sources with zero waste loadings. This allows for a large number of possible regional plant locations along the river so that chances of reducing costs are increased. The river is divided into M reaches. Physical parameters are assumed constant within each reach.

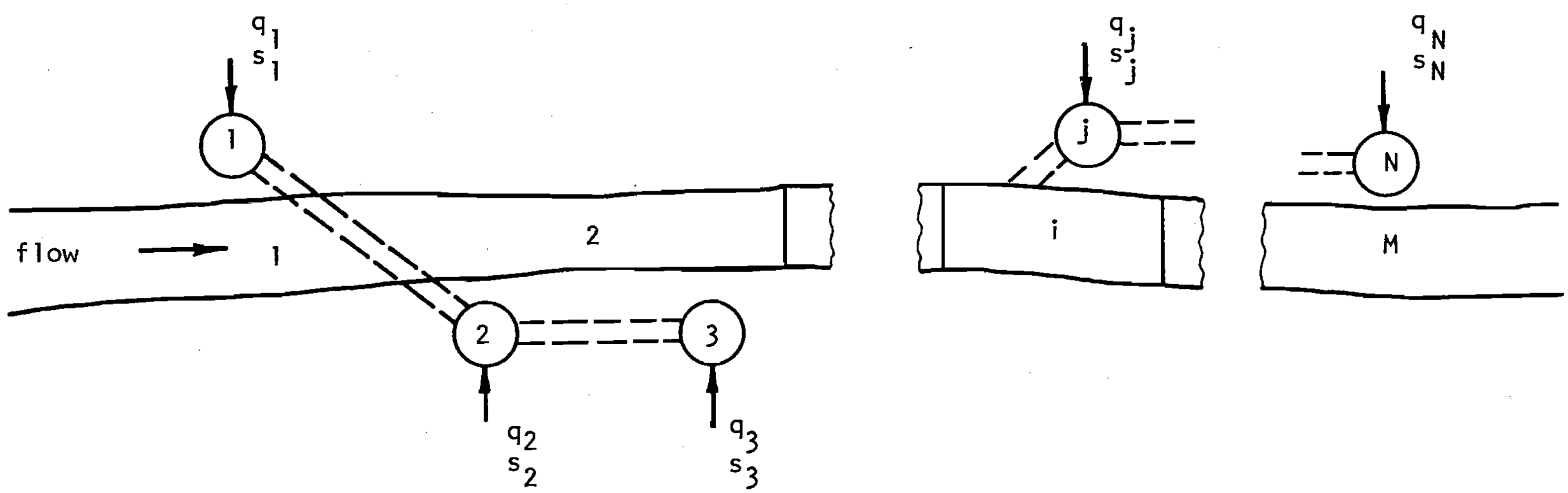
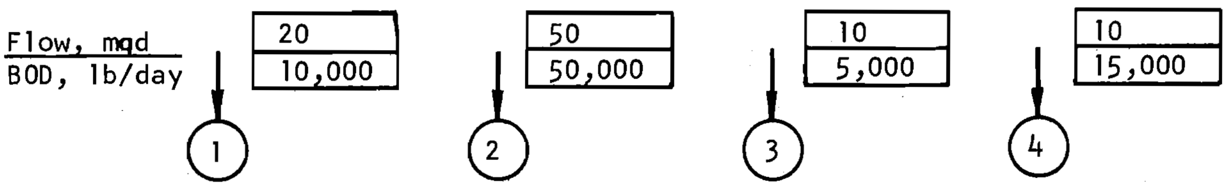


Figure 2.1 Idealized River

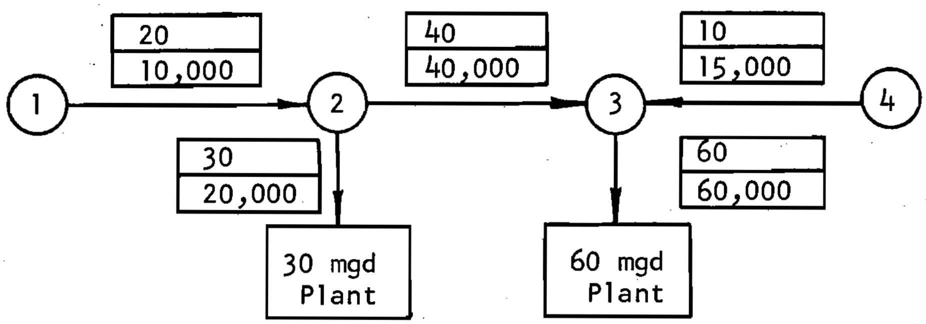
The minimum amount of DO improvement which must be attained for each reach, Δc_i (mg/ℓ), is the difference between the existing DO and the DO called for in the water quality standard. More will be said about this later. All waste loads and river parameters represent steady state values at some time.

To further simplify the locational aspects of the problem the following regionalization restrictions are made: (1) Bypassing of sources is not allowed, e.g., if source j is to pipe its waste to source j + 2 then it must pass through source j + 1 and the combined waste of both is piped to j + 2; (2) If a treatment plant is to be built at a given location it must be at least as large as the waste flow piped into it. To demonstrate the meaning of these restrictions Fig. 2.2a shows a problem with four waste sources whose flow and BOD loadings are as shown. Figure 2.2b shows a feasible regional configuration. Note that at location 2 the size of the plant is greater than the flow piped into it and that source 2 is able to split its flow between locations 2 and 3. Figures 2.2c and 2.2d show infeasible regional configurations since they violate restrictions (1) and (2) respectively.

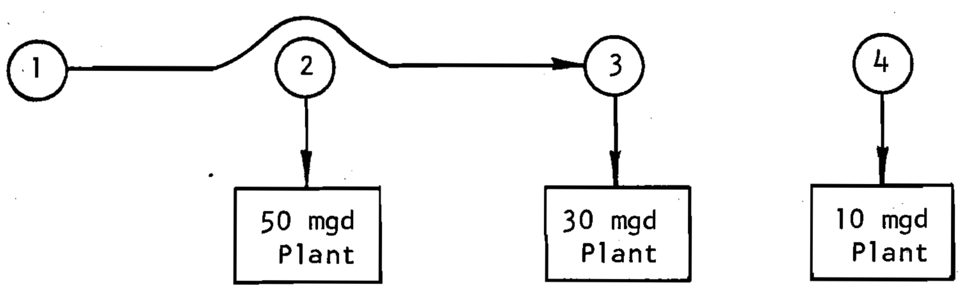
Notice that under these restrictions it is possible to get a regional treatment scheme where the flow piped into a location is treated there while the source flow of that location is piped somewhere else. For instance in reference to Fig. 2.2 it is possible for sources 1 and 2 to ship their wastes to source 3 where they would be treated and released to the river while source 3 would ship to source 4. If this seems to be an unrealistic situation it can be interpreted as having the waste of sources 1 and 2 actually treated at location 2 and then piping the



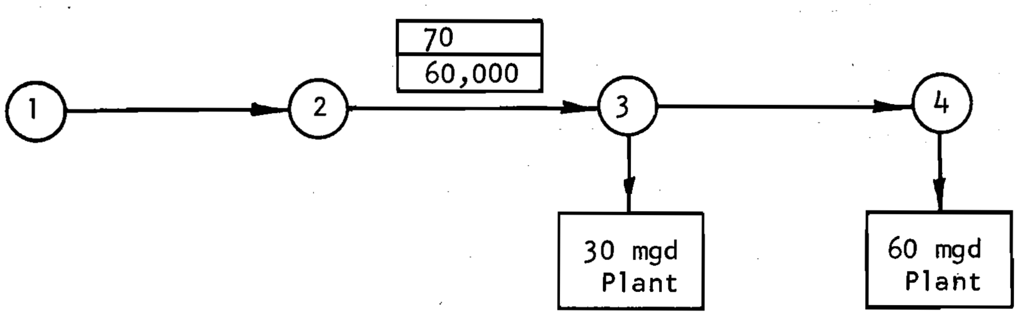
(a) Waste Generated



(b) Feasible Regional Configuration



(c) Infeasible Regional Configuration



(d) Infeasible Regional Configuration

Figure 2.2 A Four Source Example

effluent down to location 3 for discharge into the river.

These restrictions are made because they allow a simple determination of the BOD entering a given location based only on the knowledge of the flow treated at that location and the sum of flows treated upstream of it. Detailed information on the actual decisions made upstream is not required. As demonstrated later, this results in an efficient method for solving the facility location portion of the problem.

Obviously, these restrictions reduce the number of potential regional configurations considered to some subset of all possible configurations. However, in any kind of model building effort one must be aware of the tradeoffs between model specification and ease of solution. The present formulation sacrifices some degree of generality for greater ease of solution. Even so, the model can still examine a significant number of regionalization arrangements to make it a potentially valuable tool in water quality planning studies. A case in point is the analysis made for the Delaware Estuary in Chapter 5.

The regionalization restrictions made here are not without some foundation. As shown in Section 4.4 they are actually optimality conditions when degree of treatment is dropped as a decision variable, i.e., all costs are functions of waste flow only. It should be noted that the methodology presented later for solving the mathematical programming formulation of the model could, in theory, be applied to a problem statement which allowed more general regional arrangements providing that the facility location portion of the problem could be efficiently solved.

Even with these assumptions and restrictions the model remains a formidable one to solve. It allows for a large number of possible regional facility sites and, due to the no-bypassing restriction, automatically combines flows with common pipe routes in a single pipe. In addition, it seeks to find a global minimum to a problem where, because of the economies of scale, the cost functions are nonconvex and may include fixed charges. In the following sections the model will be converted into a mathematical program. Expressions will be derived for the objective function to be minimized, the water quality goals to be satisfied, and various other physical constraints on the problem.

2.2 Objective Function

The cost of the system is composed of the piping costs and the costs of building treatment plants. Since the location and length of the potential pipeline connecting any two adjacent source locations is specified, the cost of piping from location j to j + 1 is a function of the flow being piped, yp_j. This function, call it $P_j(yp_j)$, may be a simple power function as in Graves et al. (1970) or may include such details as meeting velocity constraints and pumping as in Converse (1972). If yp_j is positive then flow is piped from j to j + 1. If it is negative then flow is piped from j + 1 to j. The function $P_j(yp_j)$ will typically show economies of scale since as more flow is piped the allowable area for flow in the pipe increases in proportion to the diameter squared while the cost of the pipe increases nearly in proportion to the diameter (circumference) only.

The cost of a treatment plant at j is a function of the hydraulic size of the plant, y_j, and the percent BOD removal provided. Percent BOD removed can be expressed as $1 - z_j/w_j$ where w_j is the BOD of the influent waste in lb/day and z_j is the BOD of the effluent in lb/day. Hence the cost of a treatment plant is some function of y_j, z_j, and w_j, say $T_j(y_j, z_j, w_j)$. This function may assume a different form at location j depending on whether only source j's waste is being treated or a regional plant is built there. When treatment levels are held constant the treatment cost function typically shows economies of scale with respect to hydraulic size. This is reflected in a decreasing marginal cost as the amount of flow is increased. On the other hand, when the quantity of flow treated is held constant, the marginal cost of BOD removal increases as amount of BOD removal increases in the range from 30-50 percent removal on up. This behavior is demonstrated in the cost curves developed by Frankel (1965) shown in Fig. 2.3.

The total cost of the regional system can be expressed as

$$\text{Cost} = \sum_{j=1}^{N} P_j(yp_j) + T_j(y_j, z_j, w_j) \qquad (2.1)$$

For the moment no restrictions will be placed on the form of the piping and treatment cost functions. Later on it will be required that the treatment cost be a convex function of percent BOD removal.

2.3 Water Quality Constraints

There is a certain level of DO improvement, Δc_i, which is required in each reach i, i = 1,...,M. This required improvement has a different interpretation depending on what type of model is used to relate BOD

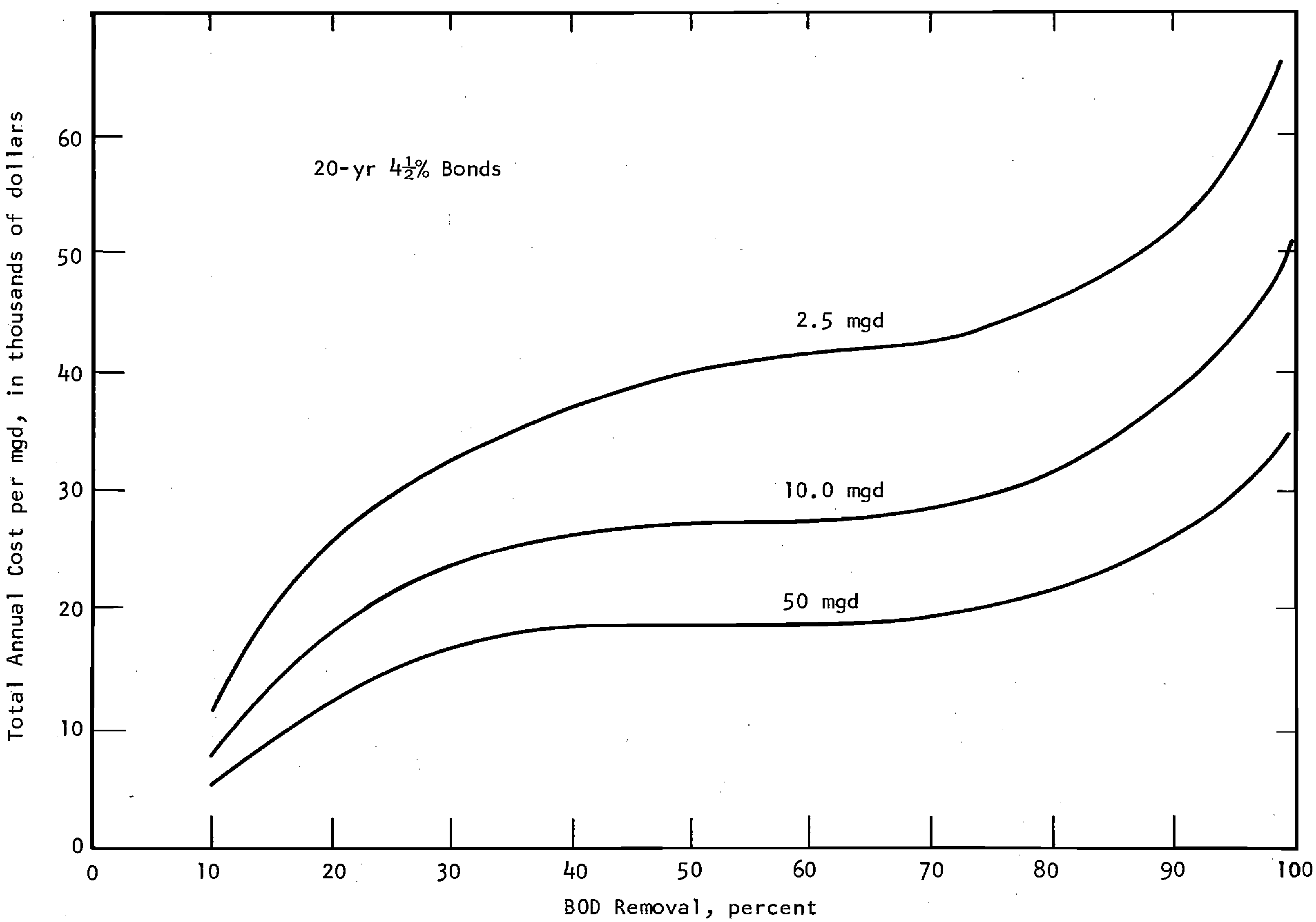


Figure 2.3 Cost of a Wastewater Treatment Plant

discharged to stream DO. For rivers which have no longitudinal mixing Δc_i is the required improvement at a specified point in reach i (Loucks et al., 1967). For rivers with longitudinal mixing (i.e., estuaries) each reach is assumed to be completely mixed and thus Δc_i is the improvement throughout the reach (Thomann, 1972). In either case the change in DO in reach i can be expressed as some linear combination of the changes in BOD discharged from each treatment plant location, $(s_j - z_j)$. Hence the constraint for reach i can be written as

$$\sum_{j=1}^{N} a_{ij}(s_j - z_j) \geq \Delta c_i \tag{2.2}$$

where a_{ij}, the DO transfer coefficient, is the unit change in DO for reach i associated with a unit change in BOD released from location j. Assuming constant levels of flow, BOD decay and reaeration rates in each reach, these coefficients can be computed from equations in Loucks et al. (1967) or Thomann (1972). However, with regionalization taking place the flow entering each reach is no longer constant. For rivers with no longitudinal mixing and M = N, if only the dilution effects of variable inflows are considered the constraint set can be written as

$$A_1 y + A_2 z + a \leq 0 \tag{2.3}$$

where A_1 and A_2 are N x N matrices of coefficients calculated from waste source and river parameters, including the Δc's, and a is an N x 1 vector of constants (see Appendix A for details). However, also affected are the time of flow to the end of the reach and the reaeration rate constant. When considered to be functions of inflow they destroy the linearity of the BOD-DO models and the use of a simple linear, separable relation as in Eq. (2.3) is not valid. For this reason the additional assumption will be made that any changes in the river flow due to regionalization will have small effect on the DO transfer coefficients. Hence the use of Eq. (2.2) will suffice. This assumption is obviously most valid for the case when the base flow in the river is large compared to the waste flow generated. For several approaches to relaxing this assumption see Appendix A.

2.4 Physical and Inventory Constraints

In this section constraints will be developed to express the following:

1. the flow piped out of (into) location j to (from) j + 1 as a function of the total flow allocated for treatment at all points between 1 and j;
2. the requirements that the flow from all sources must be allocated for treatment;
3. the regionalization restriction that if a plant is built it must be of size equal to or greater than the flow piped into it from other locations;
4. the influent BOD to location j as a function of the total flow treated between 1 and j - 1 and the flow treated at j;
5. upper and lower bounds on degree of treatment.

The flow being piped from location j to j + 1 has been denoted as yp_j and may be either positive or negative, depending on the direction of flow. Using the restriction against bypassing of locations, a flow balance around the j^{th} location gives

$$yp_{j-1} + q_j = y_j + yp_j$$

This merely states that whatever flow comes from (goes to) j - 1 plus the source flow at j must equal the flow treated and disposed to the river at j plus the flow sent to (from) location j + 1. Writing this equation for $j = 1,\ldots,N$, noting that $yp_o = yp_N = 0$, and then solving for yp_j gives

$$yp_j = \sum_{i=1}^{j} q_i - \sum_{i=1}^{j} y_i \tag{2.4}$$

Another inventory constraint requires that all flow in the region pass through a treatment plant into the river. Thus

$$\sum_{j=1}^{N} y_j = \sum_{j=1}^{N} q_j \tag{2.5}$$

Regionalization restriction (2) requires that treatment plants be at least as large as the flow piped in from other locations. This flow is determined by yp_{j-1} for location j. If yp_{j-1} is positive then flow is being piped into location j from its upstream side. Thus

$$y_j \geq yp_{j-1} \qquad \text{when} \qquad y_j > 0$$

and since $yp_{j-1} = y_j + yp_j - q_j$ we have

$$y_j(q_j - yp_j) \geq 0 \tag{2.6}$$

If yp_{j-1} is negative or zero it follows that if any flow is coming into location j it must be coming from its downstream side. This quantity has been called yp_j and would be negative. Therefore

$$y_j \geq - yp_j \qquad \text{when} \qquad y_j > 0$$

Transferring yp_j to the left side and multiplying by y_j gives

$$y_j(y_j + yp_j) \geq 0 \tag{2.7}$$

If in fact no flow is coming into location j from either side then both (2.6) and (2.7) always hold. Similarly (2.6) will always hold when $yp_{j-1} \leq 0$ as will (2.7) when $yp_{j-1} > 0$.

The influent BOD to a given plant, w_j, is dependent on the assignment of sources to treatment plants in a given regional configuration. Suppose that such a configuration has been established by choosing a set of y_j, $j = 1,\ldots,N$ such that (2.5) is satisfied. From the regionalization restrictions (1) and (2), the influent BOD to all plants is determined. This can be demonstrated by first plotting the quantities Σs_j versus Σq_j as j runs from 1 to N. Such a plot for the problem of Fig. 2.2 is shown in Fig. 2.4. Note that the slope of any portion of this curve represents the BOD concentration of the waste at the associated source in lb/mil gal. Now whatever the size of plant 1 is, say y_1, the associated influent BOD can be found by noting what ordinate corresponds to an abscissa of y_1 on Σs vs. Σq curve. Similarly, the influent BOD to plant 2, w_2, is given as the difference between the ordinate at the abscissa value of $y_1 + y_2$ and the quantity w_1. Once having constructed the plot of Σs vs. Σq from the given initial data, if the abscissa is designated as $\sum_{i=1}^{j} y_i$, where j can be between 1 and N, and the ordinate as $W(\sum_{i=1}^{j} y_i)$, then the influent BOD at any location is given by

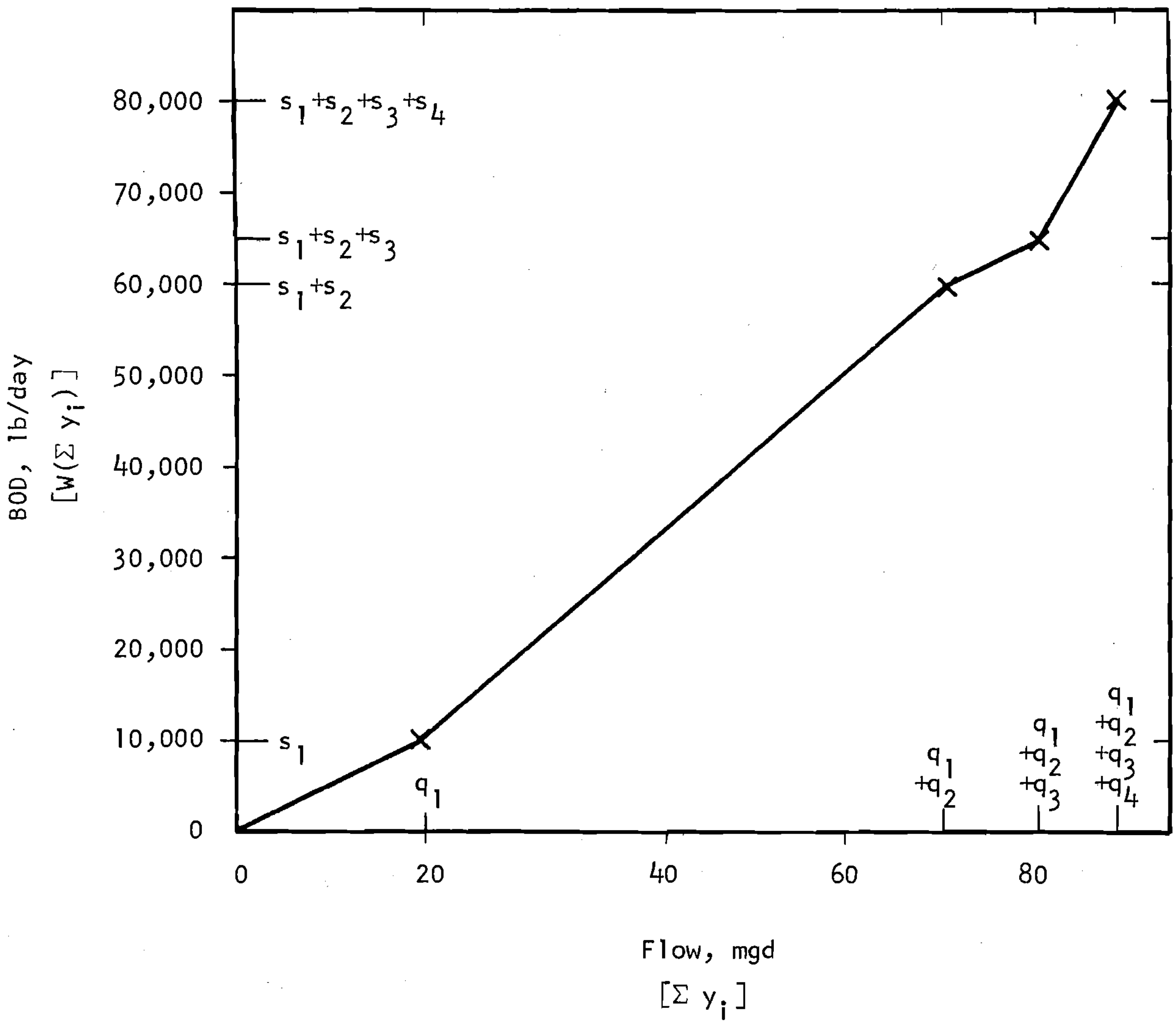


Figure 2.4 The Function W(·) for the Example of Figure 2.2

$$w_j = W(\sum_{i=1}^{j} y_i) - W(\sum_{i=1}^{j-1} y_i) . \quad (2.8)$$

Thus influent BOD at location j can be expressed as a function of the sum of the plant sizes upstream of and including location j. Had regionalization restriction (2) not been made then w_j would depend on more than simply y_j and $\sum_{i=1}^{j-1} y_i$.

An additional constraint which should be placed on the problem is to require that BOD removal be within specified limits. This can be written as

$$L_j \leq 1 - z_j/w_j \leq U_j \quad (2.9)$$

where L_j = lower limit on percent BOD removed at location j

U_j = upper limit on percent BOD removed at location j.

L_j and U_j may assume different values depending on whether a regional plant is built or only source j's own waste is to be treated. The lower bound may represent a policy such as a required minimum of primary treatment so that the river is kept free of floating debris. The upper bound can represent the technologically or economically feasible level of BOD removal available. L_j must be set at no less than zero while U_j at no more than one.

2.5 Structure of Complete Model

Using the expressions derived above the regionalization decision model can be written as the following mathematical program:

$$\text{Minimize Cost} = \sum_{j=1}^{N} P_j(yp_j) + T_j(y_j, z_j, w_j) \quad (2.10)$$

Subject to

$$\sum_{j=1}^{N} a_{ij}(s_j - z_j) \geq \Delta c_i \qquad i = 1,\ldots,M \qquad (2.11)$$

$$yp_j = \sum_{i=1}^{j} q_i - \sum_{i=1}^{j} y_i \qquad j = 1,\ldots,N \qquad (2.12)$$

$$\sum_{j=1}^{N} y_j = \sum_{j=1}^{N} q_j \qquad (2.13)$$

$$\left.\begin{aligned} y_j(q_j - yp_j) &\geq 0 \\ y_j(y_j + yp_j) &\geq 0 \end{aligned}\right\} \qquad j = 1,\ldots,N \qquad (2.14)$$

$$w_j = W(\sum_{i=1}^{j} y_i) - W(\sum_{i=1}^{j-1} y_i) \qquad j = 1,\ldots,N \qquad (2.15)$$

$$L_j \leq 1 - z_j/w_j \leq U_j \qquad j = 1,\ldots,N \qquad (2.16)$$

$$y_j\ ,\ z_j \geq 0 \qquad j = 1,\ldots,N. \qquad (2.17)$$

This program has a nonlinear (and possibly mixed integer) objective function with 4N variables (y, z, w, yp). There are 2N degrees of freedom since yp and w can be substituted for by Eqs. (2.12) and (2.15). Thus, by specifying $y_1, y_2,\ldots,y_N$ and $z_1, z_2,\ldots,z_N$ a solution is determined.

Due to the economies of scale of piping and regional treatment costs, the possibility of fixed charges (and therefore 0-1 variables), and the piecewise linear nonconvex constraints (2.15), the above is a nonconvex program. If the cost functions were free of integer variables one could conceivably apply one of a number of nonlinear programming techniques to solve it (e.g., Gradient Projection (Rosen, 1960), Method of Feasible Directions (Zoutendijk, 1960), SUMT (Fiacco and McCormick, 1964)).

However, the large number of variables and constraints plus the nonconvex piecewise linear constraints would present problems for these methods. In any event, because of the nonconvexities, a globally optimal solution cannot be guaranteed by these techniques. Note that when a regional configuration of treatment plants is predetermined by fixing $y_1, y_2, \ldots, y_N$ and the treatment cost function is convex with respect to the allowable range of percent BOD removal (a reasonable restriction for removals greater than 30-50 percent), then we obtain a convex problem with respect to the remaining variables $z_1, z_2, \ldots, z_N$. Problems of this form have been efficiently solved by Loucks et al. (1967), Hass (1970), and Haimes et al. (1972).

The structure of the model is essentially that of a serial system. The stages are the treatment plant locations. The decisions are the quantity of flow to treat and the degree of treatment at each location. The states are the total flow allocated for treatment upstream of any location and the DO improvement for each reach contributed by reduced BOD discharges of all plants upstream of any location. This structure suggests the use of dynamic programming to solve for the optimal decisions. However, the large number of state variables due to the DO improvement constraints makes a direct solution impractical.

Of the methods available for reducing state dimensionality, Discrete Differential Dynamic Programming (Heidari et al., 1971), and Successive Approximations (Bellman and Dreyfus, 1962) cannot be used because of the nonconvexities and, in the latter method, the coupling of the flow and BOD reduction variables in Constraint (2.16). Instead Lagrange multipliers will be introduced. This, along with separability of the DO constraints and objective function, allows decoupling of the BOD-DO system. The multipliers can be interpreted as prices imposed to ensure just meeting the water quality criteria. Then dynamic programming can be used to decide how much flow is to be treated at each location while at the same time the prices are used to decide how much treatment should be given. An iterative procedure is necessary to choose the optimizing set of multipliers.

This approach for a two state variable problem was first suggested by Bellman and Dreyfus (1962). In a more general context, the use of Lagrange multipliers in nonlinear programming appears in the computational methods known as column generation (Gomory, 1963), generalized linear programming (Dantzig, 1963), Generalized Lagrange Multipliers (Everett, 1963), the dual cutting plane method (Zangwill, 1969), and dual decomposition (Lasdon, 1970). All of these methods can be derived from the notion of duality in nonlinear programming. For nonconvex problems such a strategy will not always succeed and additional measures must often be taken. These ideas will be discussed more fully in the next chapter where an algorithm for solving the regionalization model will be developed.

CHAPTER 3. SOLUTION OF THE MODEL

3.1 Lagrangian Duality in Nonlinear Programming

The method used to solve the regional wastewater treatment problem formulated in Chapter 2 is based on a particular dual approach to nonlinear programming which makes use of Generalized Lagrange Multipliers (GLM) (Everett, 1963). It can be viewed as a special case of the generalized penalty - function/surrogate model recently described by Greenberg (1973). Although specific duality relations have been studied by a number of authors the concepts contained in Lasdon (1970), Geoffrion (1971), and Luenberger (1969) appear best suited for applications to nonconvex problems. What follows is a brief description of these concepts.

Consider the following mathematical program called the primal:

Minimize $f(x)$

Subject to $g(x) \leq 0$

$x \in S$

where $S \subset R^n$, $f:R^n \rightarrow R^1$ and $g:R^n \rightarrow R^m$.

The functions f and g can be any real valued functions while the set S may contain any additional constraints on the decision vector x. In applications it is best to include the "complicating" constraints in g while lumping the simpler constraints into S. The idea behind duality is to make the "complicating" constraints a part of the objective function and then solve a series of less constrained and hopefully easier problems until a certain optimization criterion is met. To do this the Lagrangian function is introduced as

$$L(x,u) = f(x) + ug(x)$$

where $u \in R^m$ is a Lagrange multiplier or dual vector.

For a given u the following function can be evaluated

$$h(u) = \underset{x \in S}{\text{minimum}}\ L(x,u)$$

Evaluating h(u) also determines an x which may or may not feasible in the primal. This function will be called the dual function. Its domain is

$$D(u) = \{\ u : u \geq 0 \text{ and } \min_{x \in S} L(x,u) \text{ exists }\}$$

The reason u is nonnegative is that only infeasible constraint values are to be penalized by adding a positive quantity to the objective function. In what follows it is assumed that the minimum of the Lagrangian exists for any $u \geq 0$ and thus $D(u) = R^{m+}$. From the Weierstrass theorem, sufficient conditions to ensure this are that f and g be continuous and S be compact.

From these simple definitions the following weak duality principle is established. Given a primal feasible x and any $u \geq 0$ then $h(u) \leq f(x)$. This follows immediately from

$$h(u) = \min_{x \in S} f(x) + ug(x)$$

$$\leq f(x) + ug(x) \leq f(x)$$

since $ug(x) \leq 0$ for feasible x and $u \geq 0$. Thus if h(u) is maximized at u* and the resulting x* is primal feasible with $h(u^*) = f(x^*)$ then x* must solve the primal (and is the global minimum of the primal). This naturally leads to the following dual program:

$$\underset{u \geq 0}{\text{Maximize}}\ h(u)$$

or

$$\underset{u \geq 0}{\text{Maximize}} \left\{ \underset{x \in S}{\text{Minimum}} [f(x) + ug(x)] \right\}$$

In an equivalent fashion the primal program can be expressed as a search for values (x^*, u^*) such that

(1) x^* minimizes $f(x) + u^*g(x)$ over $x \in S$

(2) $g(x^*) \leq 0$ and $u^* \geq 0$

(3) $u^* g(x^*) = 0$

(1) is simply the definition of $h(u^*)$; (2) ensures that x^* is primal feasible and $u^* \geq 0$, and (3) ensures that $h(u^*) = f(x^*)$ and is called complementary slackness.

The conditions (1)-(3) are sufficient, but not necessary, for optimality; there is no guarantee that they can be met for any arbitrary primal program. They are equivalent to describing a saddle point of the Lagrangian, i.e., a point (x^*, u^*) such that $L(x^*, u) \leq L(x^*, u^*) \leq L(x,u^*)$ for all $x \in S$, $u \geq 0$ (Lasdon, 1970). It follows then that if the Lagrangian of a problem has a saddle point, conditions (1) - (3) can be met and there exist dual variables such that the maximum of the dual will equal the minimum of the primal. If the problem does not have a Lagrangian saddle point then the maximum of the dual will not equal the minimum of the primal and conditions (2) and (3) cannot be met. However, the value of the dual always serves as a lower bound to the primal. Recall that no provisions have been placed on the form of f, g, and S and that when the dual method succeeds a global minimum is obtained.

When the method fails, i.e., no Lagrangian saddle point exists and conditions (1) - (3) cannot be met, the solution to a closely related problem is easily at hand. The primal can be restated as

Minimize $f(x)$

Subject to $g(x) \leq b$

$x \in S$

where b is an m-vector of right hand sides (r.h.s.). The Lagrangian is

$$L(x,u) = f(x) + u\,g(x) - ub.$$

For any $u^o \geq 0$ an (x^o, u^o) can be found which satisfies conditions (1) - (3) by simply choosing a value of b so that the x^o which solves (1) ($x_o = \{x{:}x$ minimizes $L(x, u^o)$ over $S\}$) also satisfies (2) and (3). By construction then, $b = g(x^o)$. Thus for any $u^o \geq 0$, finding the x^o which minimizes the Lagrangian solves the primal with r.h.s. of $g(x^o)$ (Everett, 1963). In fact, when any component u_i^o of u^o is zero, the corresponding constraint can have a r.h.s. $\geq g_i(x^o)$. (Note that only tight constraints at optimality will have multipliers (u_i) which are not zero. A zero multiplier at optimality implies that the corresponding constraint is superfluous and could have been deleted.) When applying the dual method to the original problem with r.h.s. of zero, if the dual is maximized at u^* and $h(u^*) \neq f(x^*)$ (or x^* is primal infeasible) then x^* still solves the primal only with r.h.s. of $g(x^*)$.

If certain restrictions are placed on f, g, and S then the existence of Lagrangian saddle points can always be guaranteed. The Kuhn-Tucker Saddle Point Theorem (Karlin, 1959; Uzawa, 1958) states that when f and g are convex, the set S is convex, and a constraint qualification is met (such as Slater's condition that there exists an $x^o \in S$ such that $g(x^o) < 0$) then if x^* solves the primal there exists a u^* such that (x^*, u^*) is a Lagrangian saddle point. In fact, under the above conditions, with $S = R^n$ and f and g differentiable, the Kuhn-Tucker necessary conditions for optimality,

$$\nabla_x f(x^*) + u^* \nabla_x g(x^*) = 0$$

$$g(x^*) \leq 0 \ , \ u^* \geq 0$$

$$u^* g(x^*) = 0$$

become equivalent to the sufficient (Lagrangian saddle point) conditions (1) - (3). Other classes of generally nonconvex problems which always have Lagrangian saddle points are geometric programs with posynomials (Duffin et al., 1967) and programs with objective functions as the ratio of a convex function to a positive linear function and subject to linear constraints (Rani and Kaul, 1973).

Up to this point the dual of a nonlinear program has been presented as a search for a Lagrangian saddle point. Another way of viewing duality is in terms of the graph of the optimal value of the primal as a function of the r.h.s. of the constraint set. Consider the optimality function given by

$$w(b) = \min\{f(x) : g(x) \leq b, \ x \in S\}$$

which is defined over the set

$$B = \{b : g(x) \leq b \text{ for some } x \in S\} .$$

For a given r.h.s. vector $b \in B$, the value of $w(b)$ is the optimal value of the primal program

Minimize $f(x)$

Subject to $g(x) \leq b$

$x \in S.$

Geometrically, $w(b)$ represents the lower envelope of the set of points

$$P = \{ [f(x), g(x)] : x \in S \}$$

which are mapped from S into R^{m+1}. Figure 3.1 pictures P and $w(b)$. The original primal problem has r.h.s. of zero. Thus, we are interested in finding $w(0)$.

Evaluating the dual function $h(u)$ for any $u^o \geq 0$ corresponds to finding a supporting hyperplane for the set P with slope $= -u^o$. This follows from

$$h(u^o) = f(x^o) + u^o g(x^o) \leq f(x) + u^o g(x)$$

so that

$$f(x) \geq -u^o g(x) + [f(x^o) + u^o g(x^o)]$$

or $$f(x) \geq -u^o g(x) + h(u^o) .$$

This describes the half-space of a hyperplane with slope $-u^o$, intercept $h(u^o)$, which lies below the set P and contacts it at the point $[f(x^o), g(x^o)]$. Since this point coincides with $w(b)$ (the lower boundary of P) it is evident that x^o solves the primal with r.h.s. of $b^o = g(x^o)$. This demonstrates Everett's Theorem which states that any $u \geq 0$ will yield a solution to a primal with modified r.h.s. Figure 3.2 shows the supporting hyperplane determined by u^o.

Since the function w is nonincreasing, it is evident as in Fig. 3.2 that $h(u)$ is always a lower bound for the optimal primal value, $w(o)$ (i.e., demonstrating the weak duality condition). In maximizing the dual

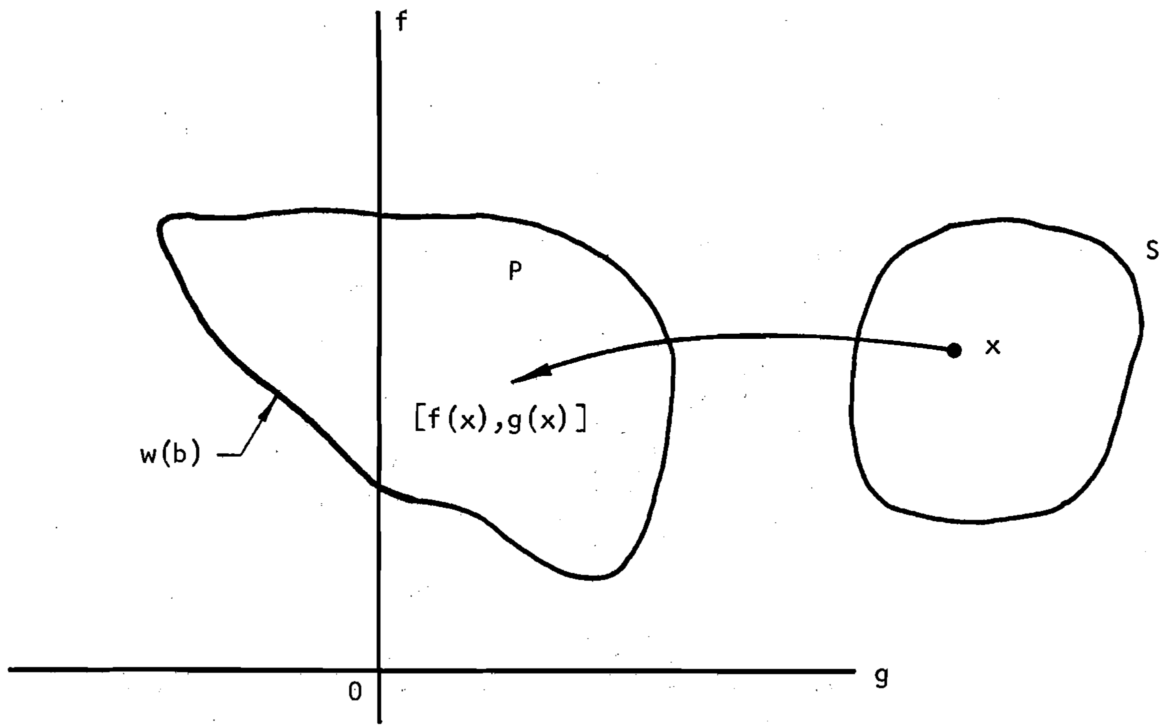


Figure 3.1 The Primal Problem in f-g Space

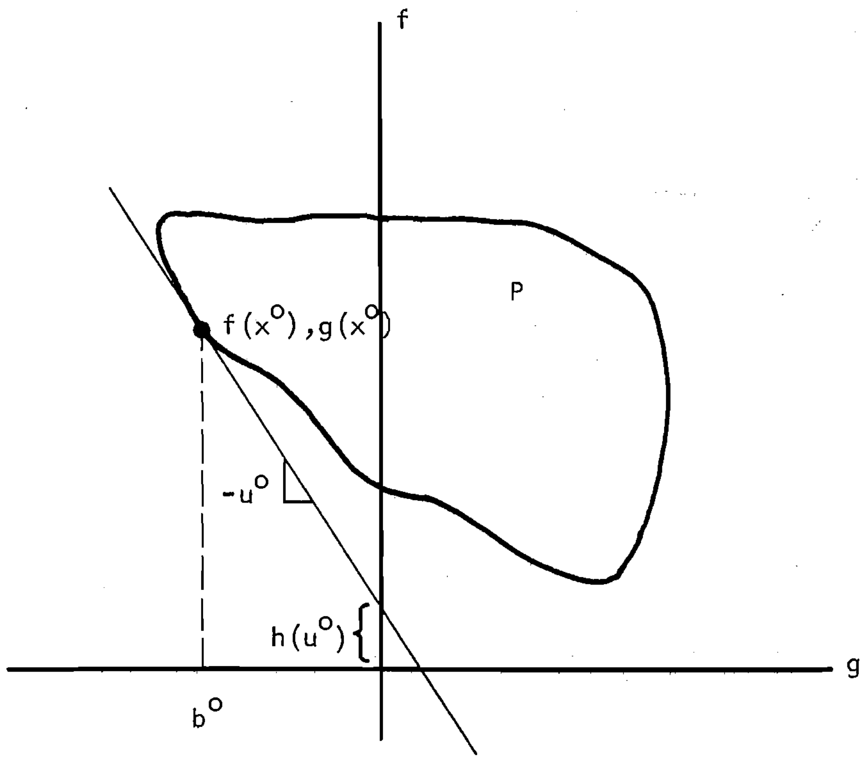


Figure 3.2 The Dual Function in f-g Space

a hyperplane of slope -u is sought which yields the highest intercept. When this intercept coincides with w(0), the optimal value of the primal, then the dual is able to solve the primal. Another way to say this is that the set P be supportable at [w(0), 0]. In Fig. 3.2 it is obvious that such a support exists. For the problem shown in Fig. 3.3 no support exists at [w(0), 0]. When the dual is maximized at u*, h(u*) < w(0). Note that evaluating the dual at u* will yield two alternate x solutions, say $\bar{x}$ and $\bar{\bar{x}}$. These represent the optimal solutions of the primal when the r.h.s. are $\bar{b} = g(\bar{x})$ and $\bar{\bar{b}} = g(\bar{\bar{x}})$, respectively. In fact, as seen in the figure, no support exists between $\bar{b}$ and $\bar{\bar{b}}$ and for all primals with r.h.s. in this interval the dual method fails. Problems such as this demonstrate <u>duality gaps</u>, i.e., gaps in the range of allowable r.h.s. such that the maximum of dual is not equal to the minimum of the primal.

The existence of duality gaps is directly related to the shape of w(b). When the functions f and g are convex and the set S is convex then the set P is convex as is the function w(b) (Luenberger, 1969). Thus P is supportable at all points on its boundary. In particular it will have nonvertical supports at all points along w(b) where b is in the interior of B. This is equivalent to the conclusion of the Kuhn-Tucker Saddle Point Theorem since both imply the success of the dual method for convex programs. The constraint qualification keeps b in the interior of B. (Vertical supports are not allowed since equality of primal and dual need not exist). For nonconvex problems the function w(b) need not be convex. Then supports will exist, and the dual method will succeed only for r.h.s. b where w(b) coincides with the convex hull of P.

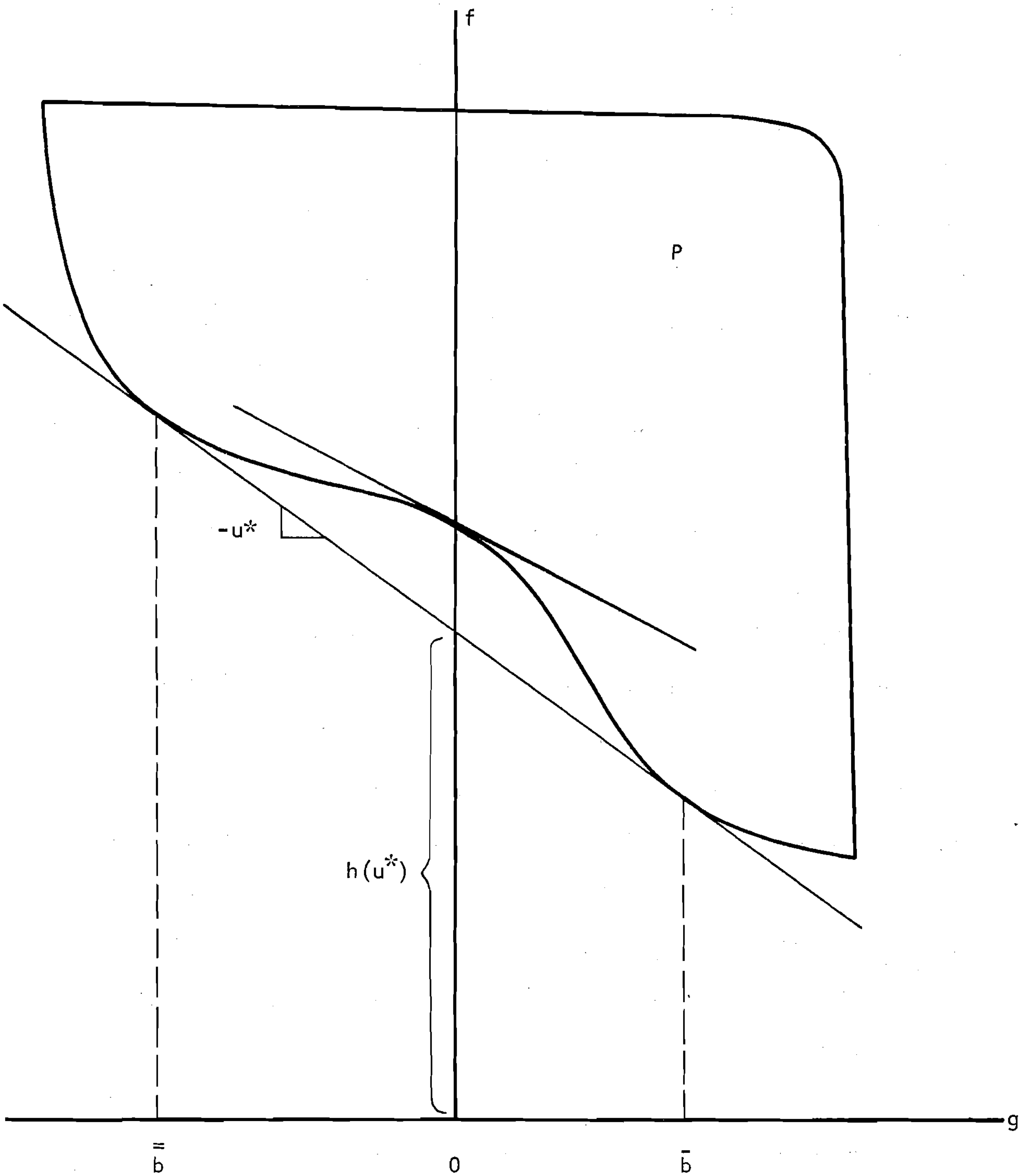


Figure 3.3 An Example of a Duality Gap

Before completing this review of duality two additional points should be noted. First, when a support is found for a particular u^o, if w(b) is differentiable at that point then $-u^o$ is the derivative of w(b), e.g., $-u^o$ measures the decrease in optimal objective function obtainable with an incremental increase in the value of the r.h.s. (Luenberger, 1969). Second, in the nonconvex case, failure of the dual method to find multipliers to support P at [w(0), 0] does not necessarily mean that optimal Lagrange multipliers in the Kuhn-Tucker sense do not exist. As Whittle (1971) shows, when $S = R^n$, f and g are differentiable and g satisfies a constraint qualification then there exist nonnegative multipliers u which define a tangent hyperplane to w(b) at [w(0), 0]. This is shown in Fig. 3.3. The above conditions imply the classical Kuhn-Tucker necessary condition of Lagrangian stationarity (Mangasarian, 1969).

To summarize, the primal is attacked by solving a dual problem

$$\underset{u \geq 0}{\text{Max}} \left\{ h(u) = \underset{x \in S}{\min} \; f(x) + ug(x) \right\} .$$

If (u*, x*) maximizes the dual and if h(u*) = f(x*) with x* primal feasible then x* is the global minimum of the primal. If these conditions do not hold at (u*, x*) then x* solves a primal with right hand side values g(x*). If these are not far from the original values, the solution may still be useful.

In the following sections the regional wastewater treatment problem will be cast in the form of a dual program, computational methods of solving the dual will be discussed, and strategies for solving the original problem when the dual method fails will be developed.

3.2 Formulation and Evaluation of the Dual

The regional wastewater treatment problem has been formulated into the following mathematical program in Chapter 2:

$$\text{Minimize Cost} = \sum_{j=1}^{N} P_j(yp_j) + T_j(y_j, z_j, w_j) \tag{3.1}$$

Subject to:

$$\sum_{j=1}^{N} a_{ij}(s_j - z_j) \geq \Delta c_i \qquad i=1,\ldots,M \tag{3.2}$$

$$yp_j = \sum_{i=1}^{j} q_i - \sum_{i=1}^{j} y_i \qquad j=1,\ldots,N \tag{3.3}$$

$$\sum_{j=1}^{N} y_j = \sum_{j=1}^{N} q_j \tag{3.4}$$

$$\left.\begin{array}{l} y_j(q_j - yp_j) \geq 0 \\ y_j(y_j + yp_j) \geq 0 \end{array}\right\} \qquad j=1,\ldots,N \tag{3.5}$$

$$w_j = W(\sum_{i=1}^{j} y_i) - W(\sum_{i=1}^{j-1} y_i) \qquad j=1,\ldots,N \tag{3.6}$$

$$L_j \leq 1 - z_j/w_j \leq U_j \qquad j=1,\ldots,N \tag{3.7}$$

$$y_j \, , \, z_j \geq 0 \qquad j=1,\ldots,N \tag{3.8}$$

where y_j = flow to be treated at location j

z_j = BOD released after treatment at location j

yp_j = flow piped between location j and j + 1

q_j = source waste flow generated at location j

s_j = source BOD generated at location j

w_j = influent BOD to be treated at location j

$W(\cdot)$ = the piecewise linear graph of Σs vs. Σq

L_j = lower bound on permissible BOD removal efficiency at location j

U_j = upper bound on permissible BOD removal efficiency at location j

P_j = cost of piping as function of yp_j

T_j = cost of treatment as function of y_j, z_j, and w_j

a_{ij} = change in DO in reach i for a unit change in BOD discharged at location j

Δc_i = DO improvement required in reach i

N = number of treatment plant locations

M = number of reaches in river.

Introducing the variable

$$\hat{y}_j = \sum_{i=1}^{j} y_i$$

and the constant

$$b_i = \sum_{j=1}^{N} a_{ij}\, s_j - \Delta c_i$$

results in the following

$$\text{Minimize Cost} = \sum_{j=1}^{N} P_j(yp_j) + T_j(y_j, z_j, w_j) \qquad (3.9)$$

Subject to:

$$\sum_{j=1}^{N} a_{ij}\, z_j \leq b_i \qquad i=1,\ldots,M \qquad (3.10)$$

$$\hat{y}_j = \hat{y}_{j-1} + y_j \qquad j=1,\ldots,N \qquad (3.11)$$

$$\hat{y}_o = 0\ ,\ \hat{y}_N = \sum_{j=1}^{N} q_j$$

$$yp_j = \sum_{i=1}^{j} q_i - \hat{y}_j \qquad j=1,\ldots,N \qquad (3.12)$$

$$\left.\begin{array}{l} y_j(q_j - yp_j) \geq 0 \\ y_j(y_j + yp_j) \geq 0 \end{array}\right\} \qquad j=1,\ldots,N \qquad (3.13)$$

$$w_j = W(\hat{y}_j) - W(\hat{y}_j - y_j) \qquad j=1,\ldots,N \qquad (3.14)$$

$$L_j \leq 1 - z_j/w_j \leq U_j \qquad j=1,\ldots,N \qquad (3.15)$$

$$y_j \ , \ z_j \geq 0 \qquad j=1,\ldots,N. \qquad (3.16)$$

Were this program to be solved directly by dynamic programming there would be one state variable corresponding to (3.11) and M state variables for (3.10). By dualizing with respect to constraints (3.10) the resulting constrained Lagrangian can be minimized by single state dynamic programming. Denoting the Lagrange multipliers or dual variables by $u_1, u_2 \ldots, u_m$, the dual function becomes

$$h(u) = \underset{y,z}{\text{minimum}} \left\{ \sum_{j=1}^{N} P_j(yp_j) + T_j(y_j, z_j, w_j) + \sum_{i=1}^{M} u_i\left(\sum_{j=1}^{N} a_{ij} z_j - b_i\right) \right\}$$

subject to (3.11) - (3.16).

The minimand can be rearranged in completely separable form to yield

$$h(u) = \underset{y,z}{\text{minimum}} \left\{ \sum_{j=1}^{N} [P_j(yp_j) + T_j(y_j, z_j, w_j) + (\sum_{i=1}^{M} u_i a_{ij}) z_j] - \sum_{i=1}^{M} u_i b_i \right\}$$

subject to (3.11) - (3.16).

For a given u, h(u) is evaluated by solving a dynamic program. The stages are $j=1,\ldots,N$ with decision variables y_j and z_j. The state variable is $\hat{y}_j$, i.e., the total flow allocated for treatment at locations 1 through j. Its transition function is Eq. (3.11),

$$\hat{y}_j = \hat{y}_{j-1} + y_j$$

with

$$\hat{y}_o = 0 \ , \quad \hat{y}_N = \sum_{j=1}^{N} q_j \ .$$

The return function for each stage can be expressed as a function of $\hat{y}_j$ and y_j with a minimization with respect to z_j being carried out as follows:

$$R_j(\hat{y}_j, y_j) = \min_{z_j} \{ P_j(yp_j) + T_j(y_j, z_j, w_j) + (\sum_{i=1}^{M} u_i a_{ij}) z_j \}$$

s.t.

$$yp_j = \sum_{i=1}^{j} q_j - \hat{y}_j \tag{3.17}$$

$$w_j = W(\hat{y}_j) - W(\hat{y}_j - y_j) \tag{3.18}$$

$$L_j \leq 1 - z_j/w_j \leq U_j \tag{3.19}$$

$$z_j \geq 0 \ . \tag{3.20}$$

For a given value of $\hat{y}_j$ and y_j the variables yp_j and w_j can be found from (3.17) and (3.18). When the treatment cost function is continuous in z_j the minimizing z_j can be found from the calculus by solving

$$\frac{\partial T_j}{\partial z_j} + \sum_{i=1}^{M} u_i a_{ij} = 0$$

$$w_j(1-U_j) \leq z_j \leq w_j(1-L_j) \ .$$

Finally the recursion relation is

$$F_j(\hat{y}_j) = \min_{y_j} \{ r_j(\hat{y}_j, y_j) + F_{j-1}(\hat{y}_j - y_j) \}$$

$$\text{s.t.} \quad yp_j = \sum_{i=1}^{j} q_i - \hat{y}_j \tag{3.21}$$

$$\left.\begin{aligned} y_j(q_j - yp_j) &\geq 0 \\ y_j(y_j + yp_j) &\geq 0 \end{aligned}\right\} \tag{3.22}$$

$$0 \leq y_j \leq \hat{y}_j \tag{3.23}$$

where $F_j(\hat{y}_j)$ is the optimal return associated with being in state $\hat{y}_j$ after j stages. The initial conditions are

$$\hat{y}_o = 0 \ , \quad F_o(0) = - \sum_{i=1}^{M} u_i \, b_i$$

and the final condition is

$$\hat{y}_N = \sum_{j=1}^{N} q_j \ .$$

Solving the recursion for $j = 1,\ldots,N$ gives the value of the dual function, i.e.,

$$h(u) = F_N(\hat{y}_N) .$$

Due to the complexity of the return function and the constraints, an analytic solution to the recursion equations is not possible. Instead the state variable $\hat{y}$ must be discretized over its allowable range $[0, \sum_{j=1}^{N} q_j]$. We denote the grid of established state values by

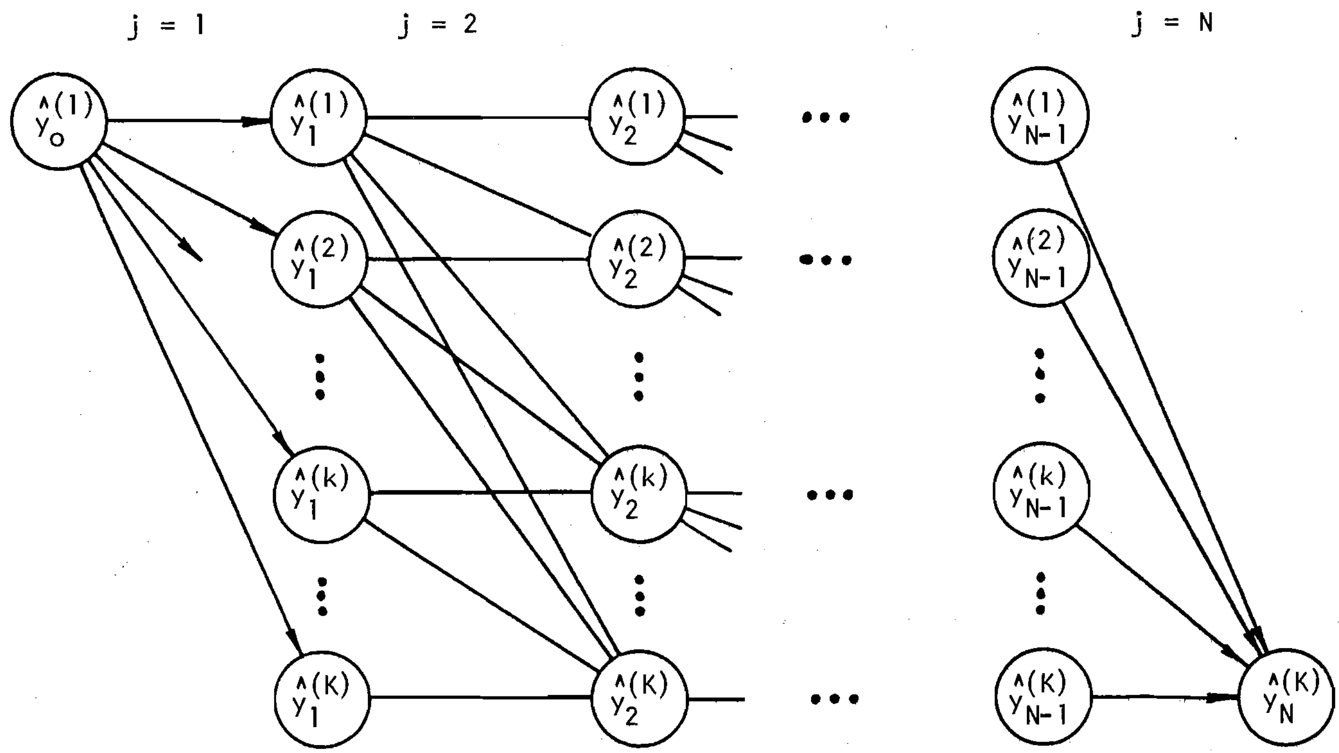
$\{\hat{y}^{(1)}, \hat{y}^{(2)}, \ldots \hat{y}^{(K)}\}$. From the constraints (3.21) - (3.23) of the recursion relation the allowable values of the decision variable y_j are

$$y_j = \begin{cases} \hat{y}^{(k)}, & \text{for } j = 1,\ k = 1,\ldots,K \\ \hat{y}^{(k)} - \hat{y}^{(\ell)}, & \text{for } 2 \le j \le N-1,\ k \ge \ell \\ & \quad k = 1,\ldots,K\ ,\ \ell = 1,\ldots,K \\ & \text{and } \hat{y}^{(k)} \ge \sum_{i=1}^{j-1} q_i\ ,\ \hat{y}^{(\ell)} \le \sum_{i=1}^{j} q_i \text{ for } k \ne \ell \\ \hat{y}^{(K)} - \hat{y}^{(\ell)}, & \text{for } j = N,\ \ell = 1,\ldots,K\,. \end{cases} \tag{3.24}$$

The other decision variable, z_j, can remain continuous over its allowable range.

The state space can be represented as in Fig. 3.4. Each node indicates a level of the state variable, $\hat{y}^{(k)}$. For any arc the difference between its state levels (its end nodes) represents an allowable level of y_j. Associated with each arc is a length which is the value of the return function $R(\hat{y}_j, y_j)$. A path connecting $\hat{y}_o^{(1)}$ with $\hat{y}_N^{(K)}$ represents a feasible solution of the dynamic program. Finding the shortest such path solves the dynamic program optimally and thus evaluates the dual function, h(u). The details of solving a discrete dynamic programming problem are given in Nemhauser (1966).

The introduction of discrete state levels and hence discrete levels of the variable y (size of treatment plant) has changed the original problem since y is no longer continuous. Although theoretically the grid size could be made small enough so that y remained essentially continuous this would make computations impractical. In view of this an additional condition must be added to the list of assumptions made in formulating



$r_1(\hat{y}_1^{(K)}, \hat{y}_1^{(K)} - \hat{y}_o^{(1)})$

$\hat{y}^{(1)} = 0 \; , \quad \hat{y}^{(N)} = \sum_{i=1}^{N} q_i$

Figure 3.4 State Space for the Evaluation of the Dual the regionalization problem in Section 2.1. The allowable levels of treatment plant sizes at all locations are restricted to a finite set of values as specified by the analyst. (Actually the analyst specifies the allowable values for $\hat{y}$ between $[0, \sum_{i=1}^{N} q_i]$. From Eq. (3.24) the allowable values of y result).

When the allowable levels of $\hat{y}$ are $\{0, q_1, q_1 + q_2, \ldots \sum_{i=1}^{N} q_i\}$ then locations either treat or pipe all of their source waste flow (plus any flow piped in from other locations), i.e., no splitting of source flow between treatment and piping is allowed. To illustrate this consider the example of Fig. 2.2 with four locations which generate 20, 50, 10 and 10 mgd respectively. The allowable set of state values would be $\{0, 20, 70, 80, 90\}$. From Eq. (3.24) the allowable values of y_j are

$$y_1 \in \{0, 20, 70, 80, 90\}$$
$$y_2 \in \{0, 20, 70, 80, 90, 50, 60, 10\}$$
$$y_3 \in \{0, 70, 80, 90, 50, 60, 10\}$$
$$y_4 \in \{0, 10, 20, 70, 90\}$$

The state space is shown in Fig. 3.5. For computational efficiency it is recommended that the state space be discretized in this manner. In fact, as shown in Section 4.4, if degree of treatment were not considered as a decision variable, then a grid of this size would always contain the optimum regional configuration.

For treatment costs which are continuous with respect to BOD removed, h(u) will exist for any $u \geq 0$. This follows since by discretizing the state space there is only a finite number of regional configurations, or values of y. Recall that h(u) is evaluated by solving

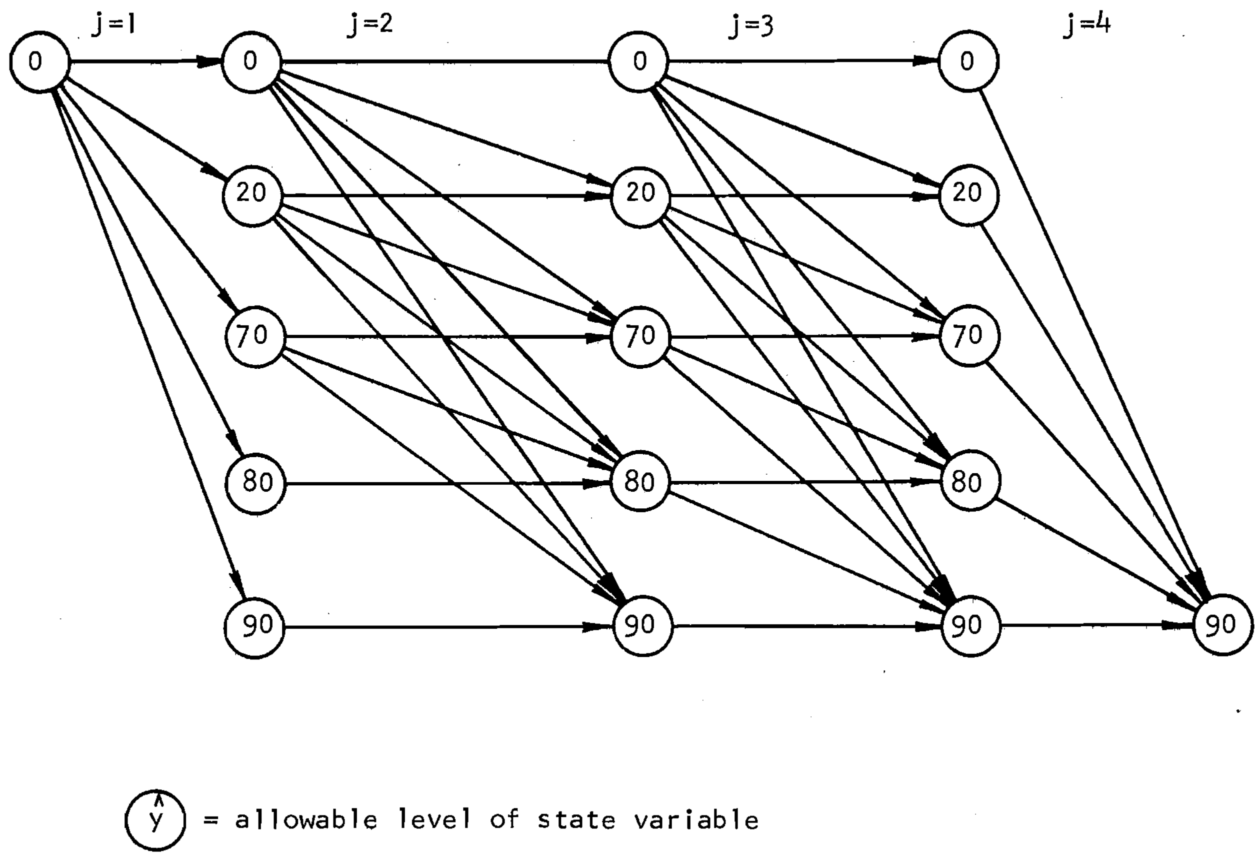


Figure 3.5 State Space for the Example of Figure 2.2

$$\min_{y,z} \left\{ \sum_{j=1}^{N} P_j + T_j + \sum_{i=1}^{M} u_i \left(\sum_{j=1}^{N} a_{ij} z_j - b_i \right) \right\}$$

$$\text{s.t.} \quad (3.11) - (3.16).$$

For fixed y the resulting problem is the minimization of a continuous function in z over a closed and bounded set, and, by the Weierstrass Theorem, this minimum always exists. Thus h(u) exists for any $u \geq 0$.

So far only a single linear segment of waste sources along a river has been considered. The model can include any number of distinct and independent segments. For example, the river may be very wide and piping across is not allowed. Then two linear segments, one for the sources on each side, can be used. This and another example are pictured in Fig. 3.6. Letting the number of such segments be K and subscripting all variables which belong to the k^{th} segment with k, the model can be written as:

$$\text{Minimize} \quad \sum_{k=1}^{K} \sum_{j=1}^{N_k} P_{jk}(yp_{jk}) + T_{jk}(y_{jk}, z_{jk}, w_{jk})$$

Subject to

$$\sum_{k=1}^{K} \sum_{j=1}^{N_k} a_{ijk} z_{jk} - b_i \leq 0 \qquad i = 1,\ldots,M$$

$$\left. \begin{array}{l} \hat{y}_{jk} = \hat{y}_{j-1,k} + y_{jk} \\ \hat{y}_{ok} = 0 \ , \ \hat{y}_{N_k k} = \sum_{j=1}^{N_k} q_{jk} \end{array} \right\} \quad \begin{array}{l} j = 1,\ldots,N_k \\ k = 1,\ldots,K \end{array} \qquad (3.25)$$

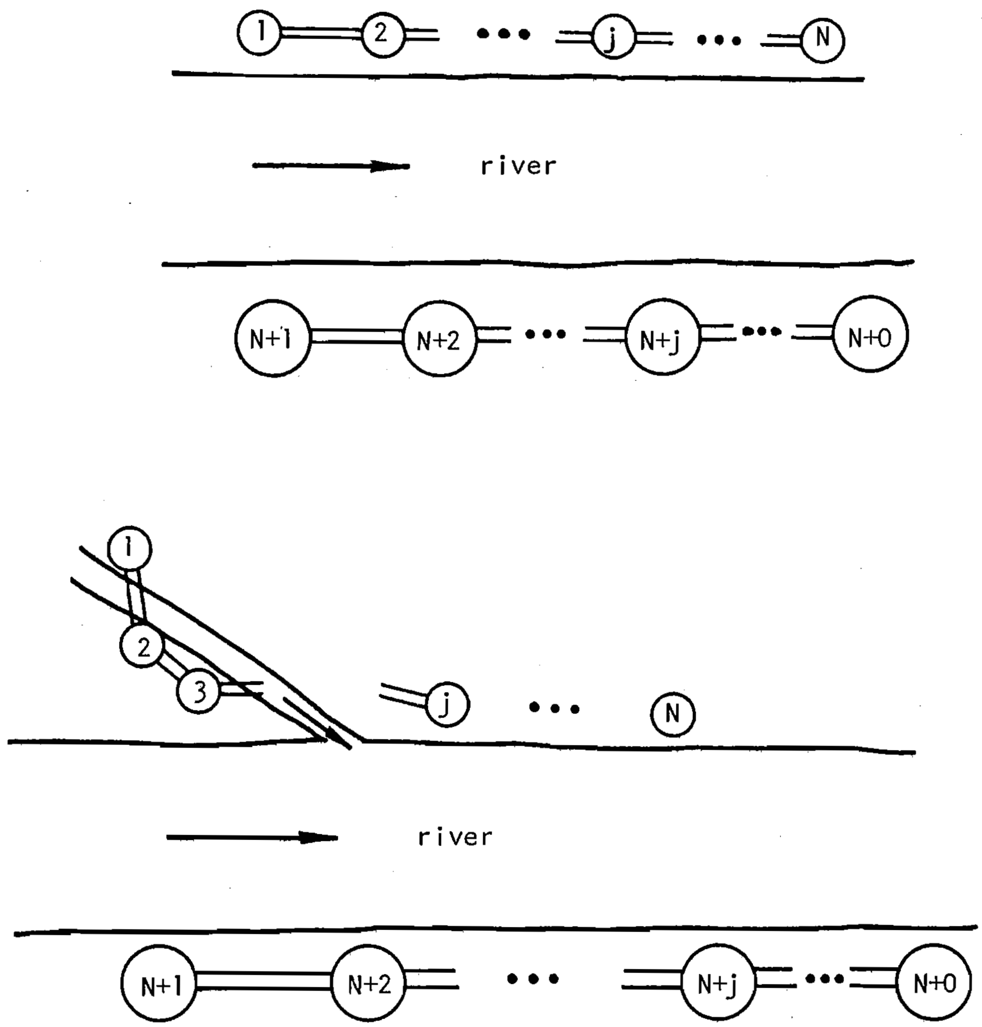


Figure 3.6 Examples of Rivers with Multiple Linear Segments of Waste Sources

$$\left.\begin{aligned}
&yp_{jk} = \sum_{i=1}^{j} q_{ik} - \hat{y}_{jk} \\
&y_{jk}(q_{jk} - yp_{jk}) \geq 0 \\
&y_{jk}(y_{jk} + yp_{jk}) \geq 0 \\
&w_{jk} = W_k(\hat{y}_{jk}) - W_k(\hat{y}_{jk} - y_{jk}) \\
&L_{jk} \leq 1 - z_{jk}/w_{jk} \leq U_{jk} \\
&y_{jk},\ z_{jk} \geq 0
\end{aligned}\right\} \quad \begin{aligned} &j=1,\ldots,N_k \\ &k=1,\ldots,K \end{aligned} \tag{3.26}$$

The dual function is now

$$h(u) = \sum_{k=1}^{K} \left\{ \min_{y_k, z_k} \left[\sum_{j=1}^{N_k} P_{jk} + T_{jk} + \left(\sum_{i=1}^{M} u_i a_{ijk} \right) z_{jk} \right] \right\} - \sum_{i=1}^{M} u_i b_i$$

$$\text{s.t. } (3.25) - (3.26)$$

and can be evaluated by solving K dynamic programs as described above.

3.3 Solution of the Dual

Recall that for the general primal problem

Minimize $f(x)$ subject to $g(x) \leq 0$, $x \in S$

$f: R^n \to R^1$, $g: R^n \to R^m$, $S \subset R^n$

the dual is defined as

$$\underset{u \geq 0}{\text{Maximize}} \left\{ h(u) = \min_{x \in S} f(x) + ug(x) \right\} , \quad u \in R^m .$$

In the previous section it was shown how the regional wastewater treatment problem could be cast in this form and how its dual could be evaluated for any $u \geq 0$. The problem remains of finding the u which maximizes h(u).

Notice that h(u) is a concave function in u since it is the pointwise minimum of a collection of linear functions of u, one for each x in S. Hence the dual is well behaved in the sense that it has no local maxima distinct from the global maximum.

If h(u) is evaluated for any u^o the result is

$$h(u^o) = f(x^o) + u^o \, g(x^o).$$

For any other u,

$$h(u) \leq f(x^o) + ug(x^o).$$

Therefore,

$$h(u) - h(u^o) \leq g(x^o)(u - u^o).$$

This describes a supporting hyperplane to the graph of h(u) which lies above h for all $u \geq 0$ and contacts h(u) at u^o. It follows that $g(x^o)$ is a subgradient of h(u) at u^o and if h were differentiable at u^o its gradient would be $g(x^o)$. Thus, when evaluating the dual function at a given point its gradient, if it exists, is readily at hand. Hence one approach to maximizing the dual is to use a gradient search technique suitably modified to handle nonnegativity conditions. Examples of specific methods are given in Uzawa (1958) and Lasdon (1970).

Falk (1967) shows that when f is strictly convex and g and S are convex then h has continuous first partials for all $u \geq 0$ and hence its gradient always exists. For nonconvex problems the gradient need not exist at all points and hence convergence of a gradient based search cannot be assured.

Another approach makes use of the supporting hyperplanes which are obtained after evaluating h(u) for any $u \geq 0$ (Zangwill, 1969 and Geoffrion, 1970). The idea is to use these hyperplanes as cutting planes on the function h and, as in Kelley's algorithm (1960), solve a series of linear programs, each time generating a new set of u's, until h is maximized. Figure 3.7 shows h(u) for $u \in R^1$ and the hyperplanes (straight lines) obtained after two evaluations of h. In general after K iterations there would be K support hyperplanes to h of the form $h = f(x^k) + u\, g(x^k)$. Now h can be approximated as the minimum point on these supports for any $u \geq 0$ or

$$h(u) \doteq \min_{1 \leq k \leq K} f(x^k) + u\, g(x^k).$$

Notice that such an approximation always exceeds the true value of h as shown in Fig. 3.7. Maximizing over the approximation to h gives

$$\text{Max } h(u) \doteq \max_{u \geq 0} \{ \min_{1 \leq k \leq K} f(x^k) + u\, g(x^k) \}$$

which is equivalent to the following cutting plane linear program,

$$\begin{aligned} \text{CPLP:} \quad & \max v \\ & \text{s.t.} \quad v \leq f(x^k) + u\, g(x^k) \qquad k=1,\ldots,K \qquad (3.27) \\ & \qquad u \geq 0 \text{ , } \quad v \text{ unrestricted.} \end{aligned}$$

Solving for v gives an upper bound on the maximum of the dual while u provides the values of the dual variables for iteration K + 1. The dual is once again evaluated and another constraint such as (3.27) is added

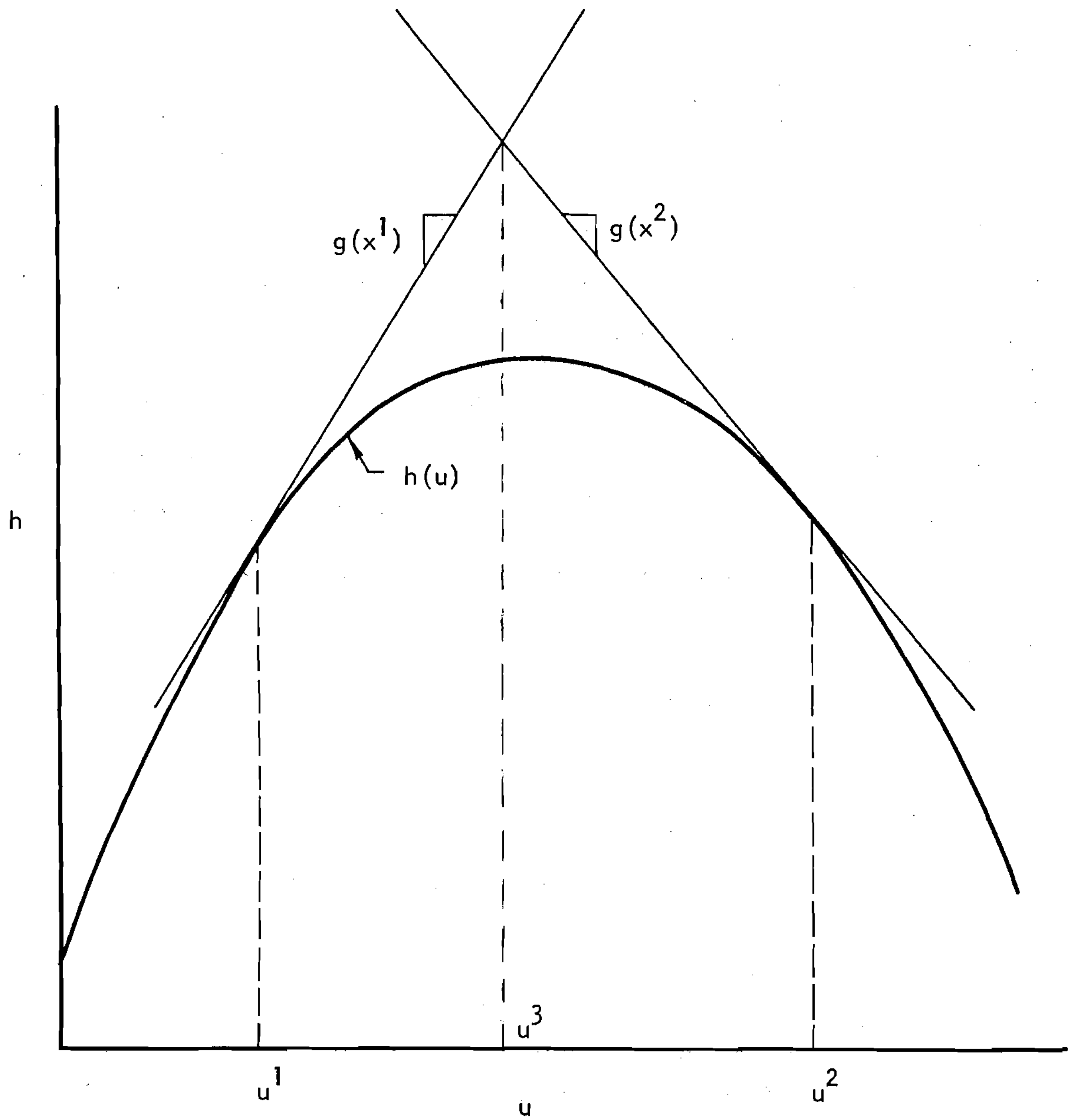


Figure 3.7 Supporting Hyperplanes to the Graph of h(u)

to the linear program which is re-solved. The algorithm can be stopped when v is sufficiently close to the best value of h found.

Taking the dual of the cutting plane LP gives the following column generation LP,

$$\text{CGLP:} \quad \min \sum_{k=1}^{K} \alpha_k f(x^k)$$

$$\text{s.t.} \sum_{k=1}^{K} \alpha_k g_i(x^k) \leq 0 \qquad i=1,\ldots,M$$

$$\sum_{k=1}^{K} \alpha_k = 1$$

$$\alpha_k \geq 0 \qquad k=1,\ldots,K.$$

Now at every iteration a new column is added instead of a constraint so that the size of the basis remains constant. This is the program arrived at by Brooks and Geoffrion (1966) by approximating the original primal over a series of grid points obtained from the first K iterations. It is the equivalent formulation behind the Dantzig-Wolfe Generalized Linear Programming method (Dantzig, 1963) for solving convex nonlinear programs. Remember for convex programs satisfying a constraint qualification, the maximum of the dual always equals the minimum of the primal. Dantzig (1963) and Zangwill (1969) show that for convex programs, solutions of CPLP (CGLP) will yield an infinite sequence of dual (primal) values which contain a limit point which solves the dual (primal) when all constraints (columns) are kept. Greenberg and Robbins (1972) show that this convergence property for the dual still holds even when the primal is nonconvex, convexity being required only to insure that no duality gap occurs. Thus the cutting plane - LP method produces a maximum for h(u) with convergence in the limit guaranteed. Note that the method yields a sequence of nonincreasing upper bounds for the maximum of the dual but that strict improvement of the dual with each iteration is not assured.

The cutting plane (column generation) - LP algorithm should be started with values of $f(x^1)$ and $g(x^1)$ such that x^1 is primal feasible. To obtain such a starting point the dual function could be evaluated for a very large value of u. It follows from the weak duality condition that if the dual is unbounded then the primal problem is infeasible. For the regional wastewater treatment problem a starting point can easily be obtained by simply having each individual source provide as much BOD removal as possible, providing that this results in all DO constraints being satisfied.

Nemhauser and Widhelm (1971) have noted that such cutting plane or column generation methods tend to show slow convergence. They suggest a modification be made on the cuts (constraints) added so that they are more "centrally located" in multiplier (dual) space. Eaves and Zangwill (1971) have presented criteria which allow cuts (constraints) to be dropped from CPLP. O'Neill (1973) has presented computational results on constraint dropping and recommends that the loose constraints of CPLP (or equivalently, the nonbasic columns of CGLP) should be dropped if the Eaves-Zangwill criteria are met. The advantage of dropping cuts or columns is that less computer storage is required. These refinements on the cutting plane - LP procedure have not been implemented in the current study since the computational results as presented in Chapter 5 show the method to work quite satisfactorily.

For the regional wastewater treatment problem, evaluating $h(u)$ for any u^o will yield a primal solution (y^o, z^o, w^o, yp^o).

Let

$$f^o = \sum_{j=1}^{N} P_j(yp_j) + T_j(y_j^o, z_j^o, w_j^o)$$

$$g_i^o = \sum_{j=1}^{N} a_{ij} z_j^o - b_i .$$

Then the procedure for maximizing the dual becomes:

(1) At the Kth iteration solve the following linear program

$$\text{Min} \quad \sum_{k=1}^{K} \alpha_k f^k$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \alpha_k g_i^k \leq 0 \qquad i=1,\ldots,M \tag{3.28}$$

$$\sum_{k=1}^{K} \alpha_k = 1 \tag{3.29}$$

$$\alpha_k \geq 0 \qquad k=1,\ldots,K.$$

Let the dual variables of constraints (3.28) be u^{K+1} and the dual variable of constraints (3.29) be v. Let the best solution of the dual recorded so far be h*. If

$$v \leq h^* + \epsilon \text{ for some } \epsilon > 0 \text{ then stop.}$$

(2) Evaluate $h(u^{K+1})$, f^{K+1} and g^{K+1}. If $h(u^{K+1}) > h^*$ then set $h^* = h(u^{K+1})$. Replace K by K+1 and return to (1).

After the dual is maximized we must check to see if the original primal has been minimized. The procedure is as follows:

Denote the primal solution associated with h* as f*, g*. If

$$\left| h^* - f^* \right| < \xi$$

$$\text{and } g_i^* < \gamma \qquad \text{for} \qquad i=1,\ldots,M$$

where $\xi, \gamma > 0$ are prescribed tolerances then f* is the minimum of the primal.

If the above conditions cannot be met then the dual method has failed to solve the primal. However, the solution to a modified problem is readily at hand. Recall that the required dissolved oxygen (DO) improvements for the primal were denoted by Δc_i, $i = 1,\ldots,M$. Then f*, g* is the optimal solution to a modified primal, with required DO improvements of

$$\Delta c_i' = \Delta c_i - g_i^* \qquad \text{for} \qquad i:u_i^* > 0$$

$$\Delta c_i' \leq \Delta c_i - g_i^* \qquad \text{for} \qquad i:u_i^* = 0 .$$

If these new standards are not far away from the original then such a solution may be satisfactory. In fact this procedure can be carried out at any iteration of the dual maximization algorithm, that is, for any $h^k(u^k)$, f^k and g^k, $1 \leq k \leq K$. In this manner sensitivity information on how the optimal regionalization plan changes with changing DO standards can be obtained with no extra effort. However, if an exact solution to the original problem is desired the methods described in the following section must be employed.

3.4 Structure of the Dual - Gaps and Their Resolution

In Section 3.1 it was shown how the success of the dual method depended on the shape of the optimal value of the primal as a function of the right hand sides of the constraints. If the graph above this curve was supportable when the right hand sides were zero, then the method would work. Otherwise, these right hand sides were in a duality gap. In this section the special structure of the regional wastewater treatment problem will be exploited so that such gaps may be overcome.

If a particular set of y_i, $i = 1,\ldots,N$ is chosen such that constraints (3.3) - (3.6) are met then this amounts to selecting a feasible regionalization configuration since the size of all treatment facilities and piping assignments are specified. The remaining problem is to find out how much BOD treatment each facility should supply so that the dissolved oxygen goals are met at minimum cost. In mathematical programming terms the problem is

$$\text{Min} \qquad f_y(z) = \sum_{j=1}^{N} T_j(z_j) + PC$$

$$\text{s.t.} \qquad \sum_{j=1}^{N} a_{ij}\, z_j - b_i \leq 0 \qquad\qquad i=1,\ldots,M$$

$$L_{yj} \leq z_j \leq U_{yj} \qquad\qquad j=1,\ldots,N$$

where $f_y(z)$ = treatment costs associated with a fixed regional configuration given by y

PC = a constant cost associated with the resulting piping costs when y is fixed

$$\left.\begin{array}{l} L_{yj} = (1 - U_j)\, w_j \\ U_{yj} = (1 - L_j)\, w_j \end{array}\right\} = \text{a constant since } w_j \text{ is fixed by } y$$

and all other symbols are as previously defined. If $T_j(z_j)$ is a convex function then this is a convex programming problem. Earlier it was noted that in the range above 30 to 50 percent removal (which can be converted to an equivalent range on BOD discharged, z_j, since the influent BOD is known) treatment costs are convex. There is a number of methods which can be used to solve convex programs but the one which obviously comes to mind is based on the dual methods previously outlined. Since all functions and feasible regions are convex (and satisfaction of a constraint qualification is assumed) by the Kuhn Tucker Saddle Point Theorem, the dual method will always succeed. The dual for fixed regional configuration, $h_y(u)$ can be evaluated by solving N univariate minimizations as in

$$h_y(u) = \sum_{j=1}^{N} \min_{L_{yj} \leq z_j \leq U_{yj}} \left\{ T_j(z_j) + \left(\sum_{i=1}^{M} u_i a_{ij}\right) z_j \right\} - \sum_{i=1}^{M} u_i b_i + PC.$$

Then the cutting plane (column generation) - LP algorithm can be used to maximize $h(u)$ and find the corresponding minimum for $f_y(z)$. In this form the dual method is the same as Dantzig-Wolfe Generalized Linear Programming (Dantzig, 1963) or Zangwill's Dual Cutting Plane Method (Zangwill, 1969).

To illustrate the nature of the dual to the overall regional treatment problem, the individual duals for each feasible regional configuration, $h_y(u)$, can be plotted as shown in Fig. 3.8 (assuming $u \in R^1$ and a total of three possible configurations). Since each individual dual is of a convex primal the maximum of each gives the optimal cost associated with that particular regional configuration. These correspond to the ordinate values a, b, and c in Fig. 3.8. The lowest of these values, a, represents the optimum configuration and level of treatment for the overall problem. Now the dual to the overall problem, $h(u)$, is the minimum taken over each

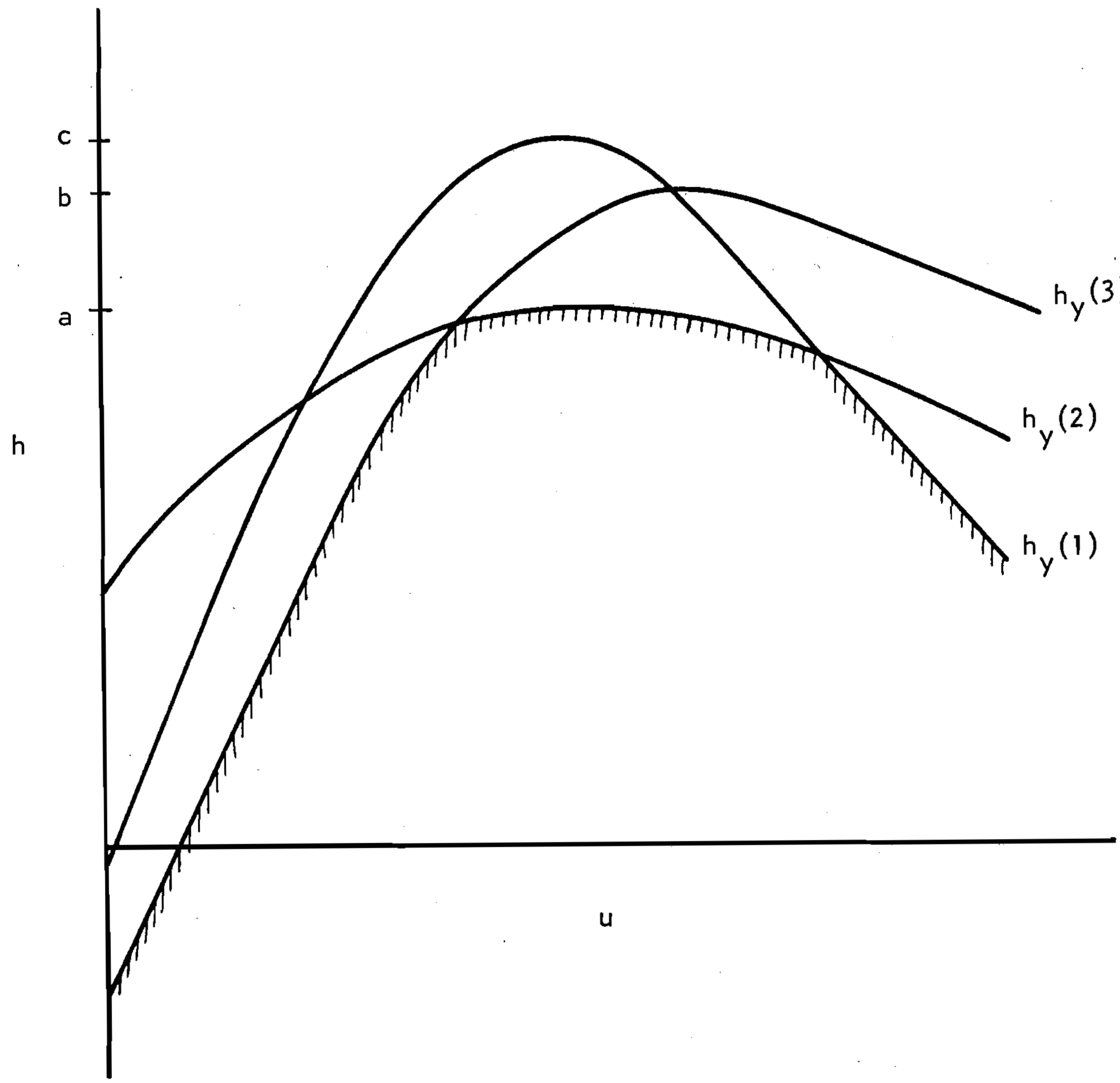


Figure 3.8 The Dual of a Regionalization Problem with No Gap

individual dual curve for any value of u, since when h(u) is evaluated a minimization is performed over y as well as z. Thus h(u) is shown as the hatched curve in Fig. 3.8. Note that in this example the maximum of h(u) has value "a" which, as previously shown, is the optimum value for the overall problem. Therefore, the dual method succeeds and no duality gap results. However, if the individual duals appeared as in Fig. 3.9, the overall minimum is at "a" while the maximum of h(u) has value d < a. Thus the dual method failes to solve the original primal.

This same structure can also be displayed in the M+1 dimensional space of cost versus value of constraint (DO improvement). Let $f_y(z)$ be as before and $g_i(z) = \sum_{j=1}^{N} a_{ij} z_j - b_i$. Consider the optimality function given by

$$w_y(r) = \min \{f_y(z) : g(z) \leq r, L_y \leq z \leq U_y\} .$$

This is the lower envelope of the set of points $P_y = \{[f_y(z), g(z)] : L_y \leq z \leq U_y\}$ which are mapped from R^N to R^{M+1}. As before $w_y(r)$ represents the optimal cost of treatment as a function of dissolved oxygen goals when the regional treatment facility pattern is given by y. Note that r is interpreted as the change in the goals from what they are in the original problem, so w(0) is the optimum solution for the original goals. As was mentioned in the review of duality concepts, since f_y and g are convex, the function w_y is convex. If w_y is plotted for each possible regional configuration y a graph such as Fig. 3.10 may result (assuming a single constraint and three possible configurations). The values a, b, and c are the optimal costs associated with each regional configuration for the original DO goals since they correspond to $w_y(0)$.

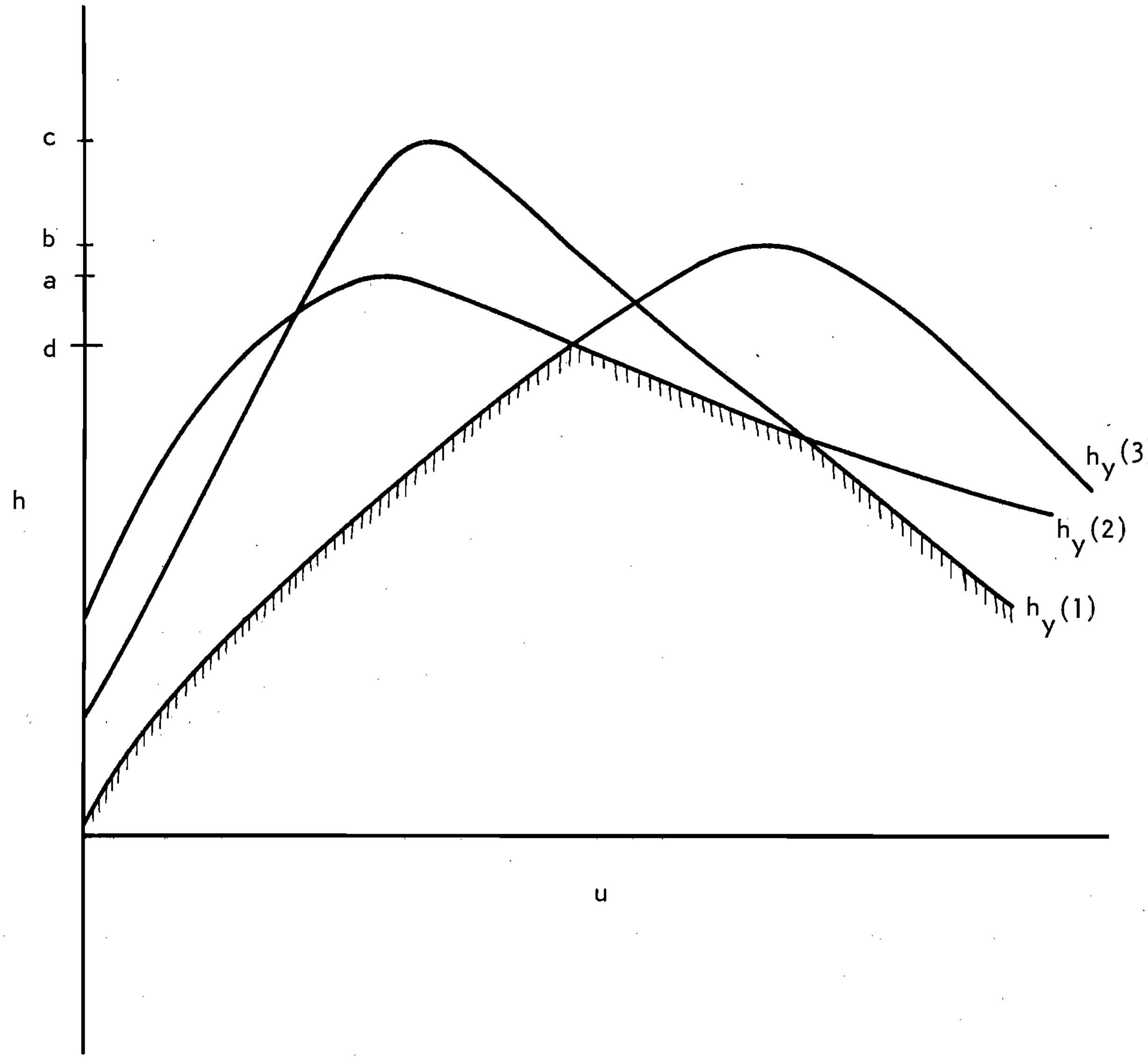


Figure 3.9 The Dual of a Regionalization Problem with a Gap

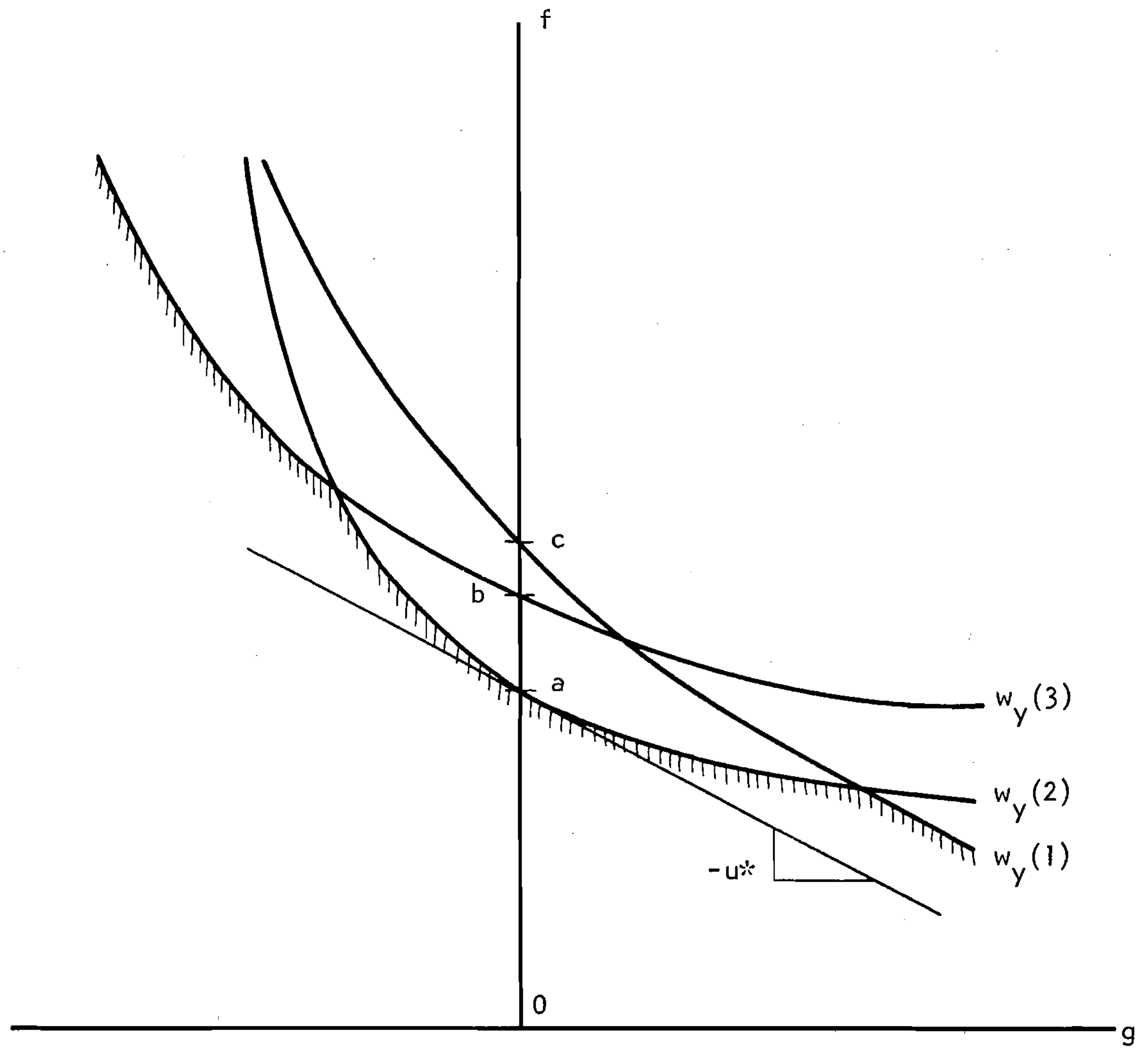


Figure 3.10 The f-g Space Representation of a Regionalization Problem with No Gap

Now the optimality function for the overall problem is given by

$$w(r) = \min_{y} w_y(r)$$

and is shown as the cross hatched curve in Fig. 3.10. Note that w(r) is not convex. Recall that the dual method attempts to find a supporting hyperplane to the graph of w(r) at [w(0), 0]. For the example of Fig. 3.10 such a support exists and thus the dual method is successful. However, in Fig. 3.11 no support exists at [w(0), 0] and the dual method fails, resulting in the lower bound d.

One method suggested for resolving duality gaps involves replacing the linear support of the Lagrangian by a nonlinear support or nonlinear Lagrangian. The Lagrangian can be written as

$$L(x,u) = f(x) + u(g(x))$$

where now u(·) is an m-valued function. Gould (1969) has shown that a saddle point for this more general Lagrangian will solve the original primal just as in the case when u is a vector. Other penalty function concepts have been related to duality by Bellmore et al. (1970), Bazaraa (1973), and Greenberg (1973). Algorithms based on these concepts usually choose some class of penalty function described up to some parameter and then vary this parameter so that a support to [w(0), 0] is achieved. An example is shown in Fig. 3.12 where the nonlinear support is able to dip into the gap region. The problem with these methods is that there is no guarantee that a particular choice of penalty function will resolve a gap. In addition the nonlinearities introduced destroy the separability of the Lagrangian and make its minimization much more difficult.

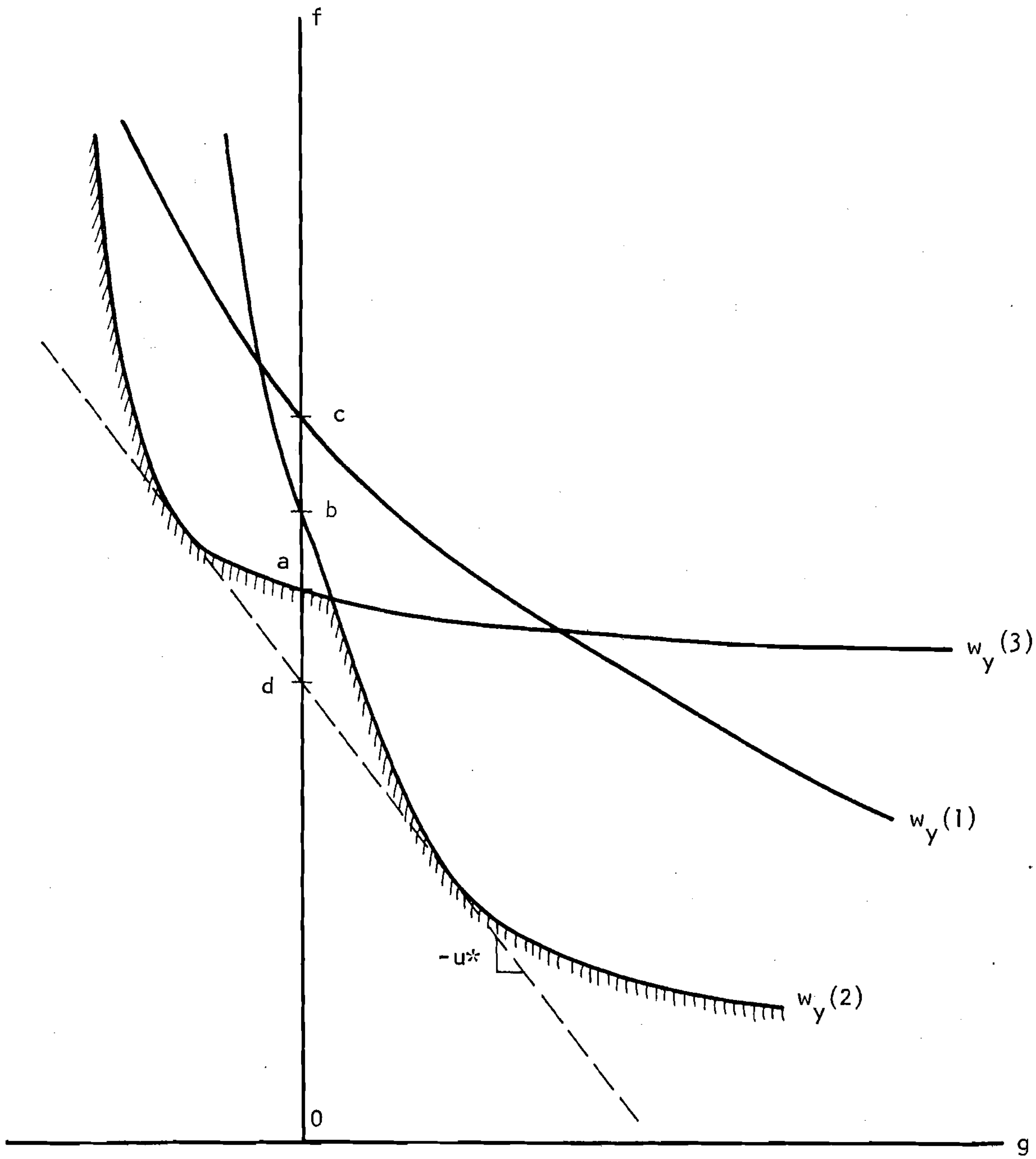


Figure 3.11 The f-g Space Representation of a Regionalization Problem with a Duality Gap

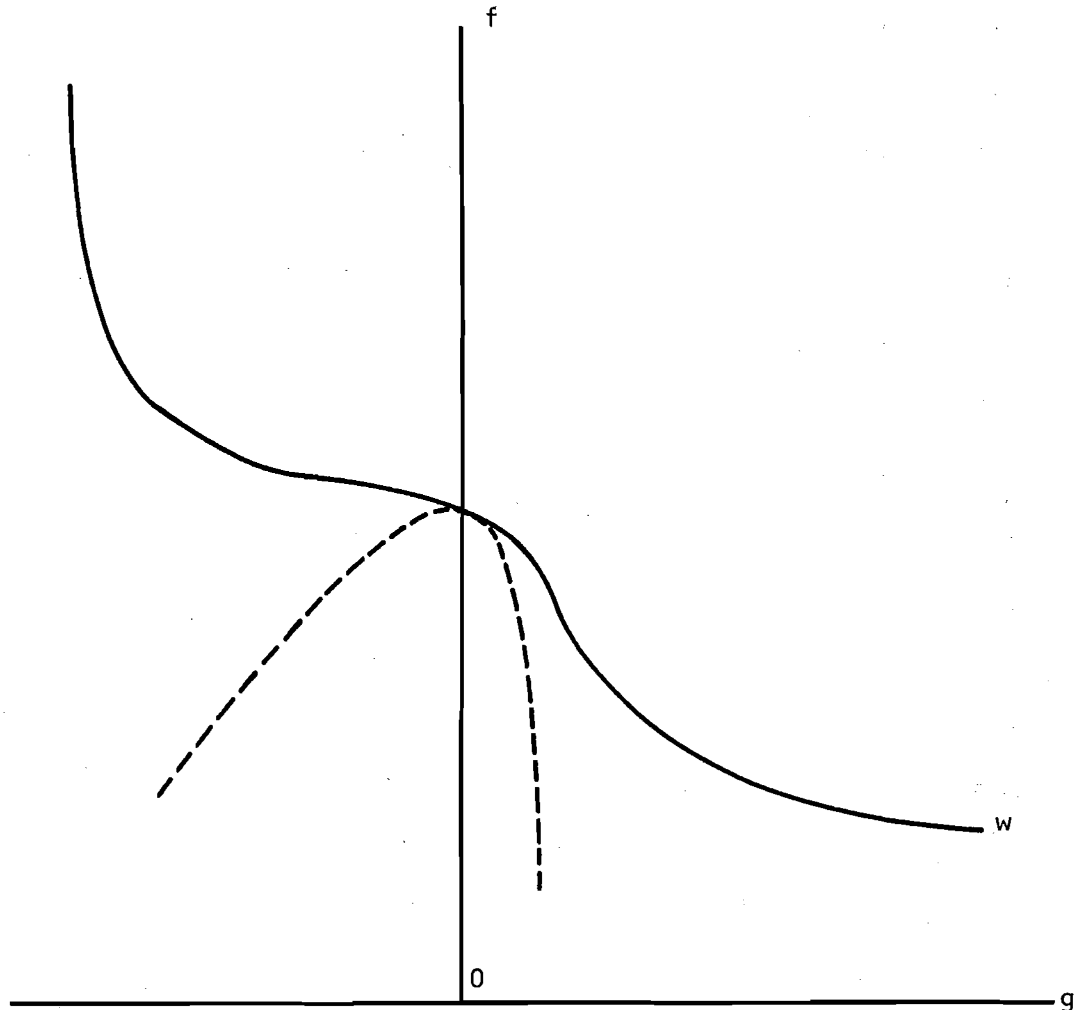


Figure 3.12 Gap Resolution by Nonlinear Supports

Another method for gap resolution is based on a branch and bound approach as suggested by Greenberg (1969). A general discussion of branch and bound can be found in Garfinkel and Nemhauser (1972). It is based on the observation that when in a gap there are at least two regional configurations which result when h(u) is evaluated at the maximizing u. In Figs. 3.9 and 3.11 these configurations would be $y^{(2)}$ and $y^{(3)}$. Such behavior is proved formally in the gap detection theorems of Bellmore et al. (1970) and Greenberg (1969).

To demonstrate how the procedure works we will examine the example of Fig. 3.9 (refer to Fig. 3.13). After we have maximized the dual of the regional wastewater treatment problem we observe that its value is "d" and that a duality gap results (the corresponding primal value is not equal to the maximum of the dual and/or the DO constraints are violated). We note that there are two regional configurations which give this value, $y^{(2)}$ and $y^{(3)}$. For each configuration we can solve the minimum cost degree of treatment problem by the dual method described above and obtain the optimal values "b" for configurations $y^{(2)}$ and "a" for configuration $y^{(3)}$. The lower of these, "a", represents a best feasible solution, or upper bound, for our overall problem. (In fact it is the optimal solution; however, we cannot verify this at this point.) Along with this upper bound, the maximum of the dual, "d", serves as a lower bound. All of this information is displayed in the top graph of Fig. 3.13.

To reduce the difference between these bounds we proceed with a branching. The set of all feasible regional configurations is divided into two subsets, one which contains $y^{(2)}$ but not $y^{(3)}$ and one which contains $y^{(3)}$ but not $y^{(2)}$. Over each subset a new regional wastewater

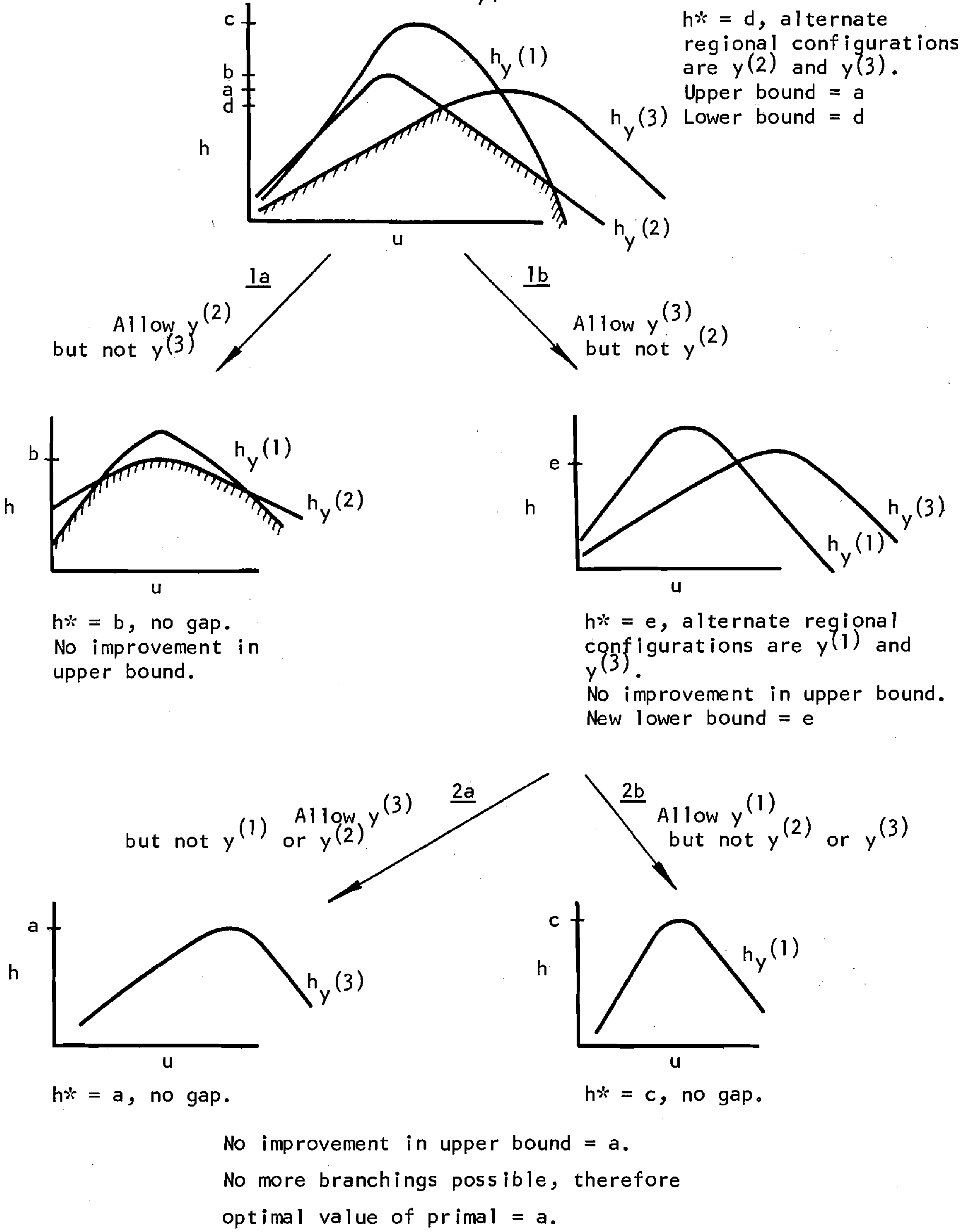


Figure 3.13 Gap Resolution by Branch and Bound as Applied to the Problem of Figure 3.9

treatment problem is created. These are problems 1a and 1b in Fig. 3.13. Solving 1a by the dual method results in no gap. Its optimal primal and dual value is "b" and no improvement in the upper bound is made. For problem 1b another duality gap results. The maximum dual value is "e" and it gives us an improved lower bound for the overall problem. The two regional configurations which result from this gap are $y^{(1)}$ and $y^{(3)}$. Solving the minimum cost degree of treatment problem for each of these gives no improvement in the upper bound.

Next another branching can be made from problem 1b creating problems 2a and 2b as shown in Fig. 3.13. Solving the regional wastewater treatment problem for each results in no duality gap. There is no improvement made in the upper bound ("a") and since no more branchings can be made we conclude that "a" must be the optimal value for the overall regionalization problem.

This example displayed two characteristics which may be misleading. First, since we assumed that there were only three feasible regional configurations the branch and bound procedure amounted to complete enumeration. However, in general there could be a great number of other configurations whose individual dual curves lay well above those of Fig. 3.9 and thus would not enter into the analysis. Second, one may conclude that the optimal solution always corresponded to one of the alternate regional configurations obtained after the initial dual maximization as was the case in the example with $y^{(3)}$. A counter example is shown in Fig. 3.14.

The computational details of the branch and bound gap resolution procedure are as follows. Suppose that the column generation LP has

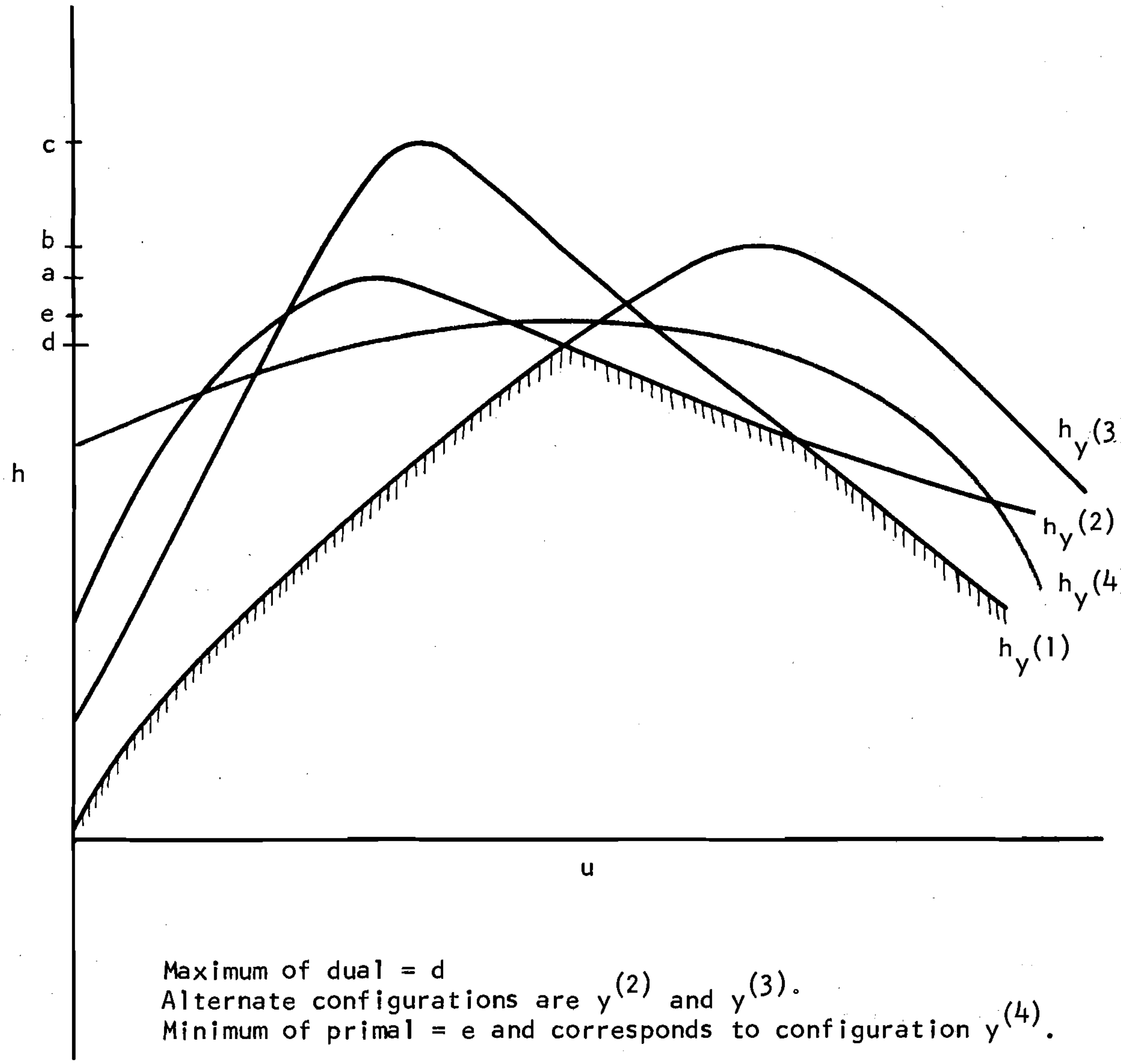


Figure 3.14 A Counter Example to the Duality Gap of Figure 3.9

maximized the dual of the regionalization problem in K iterations and a duality gap results. Denote the maximum of the dual as h* with dual variables u* and regional treatment plant configuration y*. There should be an iteration close to K which has dual value equal (or very close) to h* but a different configuration than y*. If this simple examination fails to produce an alternate configuration (and it has never done so in practice yet) then one can (a) perturb u* slightly and evaluate the dual so that it is close to h* but produces a different regional configuration than y*; (b) for u* find the second best solution to the evaluation of the dual function by using a method to find the kth best solution to a dynamic program (see Elmaghraby, 1970, for example). Now the minimum cost degree of treatment problem is solved for each configuration by the method described at the beginning of this section. We denote these solutions as $f^*_{y^*}1$ and $f^*_{y^*}2$. The solution with lower cost is the best upper bound on the solution to the overall regional treatment problem while the maximum of the dual, h*, is a lower bound.

Although the best branching rule to use is not obvious, it is clear that the allowable set of regional configurations should be split into two mutually exclusive and collectively exhaustive sets, one containing y^{*1} but not y^{*2} and the other y^{*2} but not y^{*1}. Now the dual of the regionalization problem is re-solved over each of these subsets. Its maximum value is a potential improvement on the current lower bound. If no duality gap occurs then this value is also a potential improvement on the upper bound. If a duality gap does occur then there will be another pair of regional configurations identified. Solving the degree of treatment problem for each of these configurations can result in potential improvement in the upper bound. Then this subset can be partitioned into two more subsets and the procedure continued. There is no additional branching from a subset whenever the maximum of its dual is greater than the current upper bound or when no duality gap results. The procedure ends when no more branchings can be made or when the lower and upper bounds are close enough together.

One means for partitioning the set of feasible regional configurations into subsets is as follows. When the two alternate configurations are identified they must differ in at least two components. Select some component of y, say y_ℓ, at which a difference occurs. Suppose $f^*_{y^{*2}} > f^*_{y^{*1}}$ and $0 < y_\ell^{*2} < y_\ell^{*1}$. Now the allowable set of regional configurations can be partitioned into two subsets, one having

$$0 \leq y_\ell \leq y_\ell^{*2}$$

and the other

$$y_\ell^{*2} < y_\ell .$$

The regional treatment problem associated with each subset is solved by the dual method where the above conditions serve as bounds on the allowable size of treatment plant for location ℓ and offer no problem when evaluating the dual function by dynamic programming. Note that since the state space and the allowable plant sizes are discrete the above partitioning assures that the subsets are mutually exclusive and collectively exhaustive. Since the number of regional configurations is finite the procedure will eventually terminate when:

(1) the dual method solves one subset problem with no duality gap and the value of this solution is greater than or

equal to the current lower bound and less than or equal to the current upper bound and

(2) all other subset problems have dual solution values greater than this value.

In terms of f-g space the branching serves to remove some regional configurations from consideration and thus changes the shape of w(r) to permit a support at [w(0), 0]. This branch and bound method for gap resolution produces a sequence of feasible solutions along with lower bounds to the overall regional wastewater treatment problem. Thus, it may be terminated whenever the lowest value of a feasible solution is close enough to the lower bound ensuring that the solution is no worse than a known percentage of the optimum.

3.5 Complete Solution Algorithm

Initialization: Lower bound = $-\infty$

Upper bound = $+\infty$

Dual Solution Phase:

Step 1 - Maximize the dual of the regional wastewater treatment problem by the column generation LP as outlined in Section 3.3. The dual function is evaluated by solving a discrete dynamic programming problem as discussed in Section 3.2. Let the maximum of the dual be h^* and the corresponding primal solution be (y^*, z^*) with objective function value of f^* and DO constraint values of g_i^*, $i = 1,\ldots,M$.

Evaluation Phase:

Step 2 - Determine if the dual method has solved the primal without any duality gap by seeing if f^* is close enough to h^* and all DO constraints are close enough to feasibility. If so then conclude that (y^*, z^*) solves the regionalization problem.

Step 3 - If a duality gap exists then determine what dissolved oxygen goals (y^*, z^*) (or any other primal solution generated) solves for optimally by examining $\Delta c_i - g_i^*$. If such a solution is satisfactory then stop here. Otherwise place this problem and all its information in a list.

Branch and Bound Phase:

Step 4 - Remove from the list the regionalization problem with the lowest h^* value. If the list is empty then the problem is solved. If $h^* >$ lower bound then set lower bound = h^*. Identify two alternative regionalization configurations which led to h^*. Denote these configurations as y^{*1}, y^{*2} and solve the minimum cost degree of treatment problem for each by the convex programming method suggested in Section 3.4. Let these solutions have values $f_y^*{*1}$ and $f_y^*{*2}$. If the smaller of these is less than the current upper bound then replace the upper bound with it and record all solution information. If the upper bound is close enough to the lower bound then stop here.

Step 5 - Construct two new regional wastewater treatment problems by adding bounds on the allowable size of treatment plant at some location so that y^{*1} and y^{*2} cannot be feasible in the same problem (see the method suggested in Section 3.4).

Step 6 - Maximize the dual for each problem as in Step 1. If no gap exists then if $h^* <$ upper bound let upper bound = h^* and record all solution information. If there is a gap and $h^* <$ upper bound then put this problem and its information into the list. Go to Step 4.

3.6 Computational Considerations

When solving the algorithm presented in the previous section there are several procedures which have the potential for reducing computational time. Whenever a branching occurs there are two additional problems created which must have their duals maximized. This branching was necessary because there were two alternate regional configuration solutions when the previous dual problem was maximized. The algorithm requires that each of these configurations be solved for the optimum degree of treatment to meet the water quality goals by convex programming. One can then use the optimal dual variables associated with both of these problems to begin the dual maximization of the problems created after branching. As is evident from Fig. 3.9 these starting dual variable values will bracket the optimal dual solution to the succeeding regionalization problems generated by the branching process. Thus starting with these values rather than arbitrarily large values of the dual variables as would normally be done can possibly save time when maximizing the dual.

Another fact to notice is that when maximizing the dual for a problem created by the branching process it may not be necessary to find the actual maximum value. Once a dual value is found which exceeds the current upper bound for the overall problem, computations can terminate since this subset of solutions can never give a solution less costly than one already at hand.

After solving the dual of the original regionalization problem (or several subsequent problems generated by branching) one may notice that most of the optimal dual variables are zero. This implies that the corresponding constraints in the column generation-LP are strictly satisfied, e.g.

$$\sum_{k=1}^{K} \alpha_k g_i^k < 0 \qquad \text{for} \qquad i \in I \qquad (3.30)$$

where $I = \{i : u_i = 0\}$. One strategy for reducing computations is to restrict the dual maximization to only those u_i not in I, while setting the others equal to zero. This is equivalent to relaxing the column generation LP to include only those constraints with index not in I, and thus reduces the size of the basis. However, upon completing the optimization one must check whether the restriction (relaxation) was valid by seeing if (3.30) actually does hold. If not, then the procedure must be continued and a constraint which was violated must be introduced into CGLP while its corresponding dual variable is released from its value of 0. Further details of restriction and relaxation strategies can be found in Geoffrion (1970).

Most of the algorithm's computation time would probably be spent in evaluating the dual function at each iteration by dynamic programming. A possible means for reducing this effort could be the use of discrete differential dynamic programming (DDDP) (Heidari et al., 1971). DDDP is an iterative process which starts with a trial state path through the state space and performs conventional dynamic programming over those states in the neighborhood of this path. A locally improved solution is obtained which then becomes the trial path on the next iteration. Using this method a local optimum can be found in a relatively short time, providing the initial trial path is close to optimal. Although local minima are of no use in evaluating our dual function, DDDP might still be valuable. Some applications of the algorithm have shown that in maximizing the dual the regional facility patterns produced by dynamic programming at each iteration quickly converge to one or more patterns which have neighboring paths in the state space (see Section 5.2). A possible strategy for utilizing DDDP in maximizing the dual would be to use conventional dynamic programming for the first few iterations until those facility patterns are established which are within the vicinity of the patterns which maximize the dual. Then in subsequent iterations DDDP could be used for evaluating the dual function with its associated reduction in computation time.

In the case where the treatment cost functions are piecewise linear with respect to BOD removal the solution algorithms may have to be augmented. Assume that such functions are described with a single linear segment (what follows will also hold true for more than one segment as long as they form a convex function). Then when evaluating the dual function, at the step where the optimal degree of treatment is computed a solution to the following is required

$$\frac{\partial T_j}{\partial z_j} + \sum_{i=1}^{M} u_i \, a_{ij} = 0$$

$$w_j(1 - U_j) \leq z_j \leq w_j(1 - L_j) .$$

Since T_j is linear in z_j, when $\partial T_j/\partial z_j$ is different from $-\sum_{i=1}^{N} u_i \, a_{ij}$ the optimal z_j will be at one of its bounds. However, when $\partial T_j/\partial z_j = -\sum_{i=1}^{M} u_i \, a_{ij}$ then z_j can be anywhere between its bounds. Whatever value is chosen will not affect the value of the dual function but it will affect the value of the dissolved oxygen constraints

$$\sum_{j=1}^{M} a_{ij} \, z_j - b_i \leq 0 \qquad i=1,\ldots,M.$$

Suppose that the dual has been maximized and such an indeterminate z_j exists. Then the problem is to find the z_j which preserves feasibility at minimum cost.

Recall that for the program

$$\text{Min } f(x) \qquad \text{s.t. } g(x) \leq 0,\ x \in S$$

if the dual method works then a Lagrangian saddle point (x^*, u^*) is found such that

(1) x^* minimizes $f(x) + u^*g(x)$ over S

(2) $u^* \geq 0$ and $g(x^*) \leq 0$

(3) $u^*g(x^*) = 0$.

For the regionalization problem, after the dual has been maximized and assuming no duality gap, if some z_j's are indeterminate in the sense described above then these three optimality conditions can be imposed to find their optimal values. We solve the linear program

$$\text{Minimize } \sum_{j \in J} T_j(z_j)$$

$$\text{Subject to } \sum_{j \in J} a_{ij} z_j + \sum_{j \notin J} a_{ij} z_j - b_i \begin{cases} = 0 & i \in \bar{I} \\ \leq 0 & i \notin \bar{I} \end{cases}$$

$$w_j(1 - U_j) \leq z_j \leq w_j(1 - L_j) \qquad j \in J$$

where $\bar{I} = \{i : u_i > 0\}$ and $J = \{j : z_j \text{ indeterminate}\}$. Note that the z_j, $j \notin J$, which appear in the constraints have their values already known. Likewise the regional facility pattern, its contribution to the objective function and to w_j are also known. If a duality gap existed then this program would not have a feasible solution. In practice, due to computer round-off, it would probably be difficult to identify such indeterminate variables and thus it would be safest to solve the above LP for all z_j, j = 1,...,N. The above considerations are also required when the dual method (or equivalently, the Dantzig-Wolfe Generalized LP method) is used in the duality gap resolution procedure to find the minimum cost degree of treatment for a given regional configuration.

Although the entire dual maximization - branch and bound algorithm could be programmed in closed form for direct implementation on a digital computer, a simpler but more flexible approach can be used. A single program can be written which maximizes the dual of the regionalization problem and solves the minimum cost degree of treatment problem for a fixed regional configuration. Both problems use the same basic input data concerning waste source information, DO transfer coefficients, and required DO improvements. For the regionalization problem these data would be augmented with the bounds on the allowable size of treatment plant at certain locations as required in the branch and bound procedure. For the degree of treatment problem with fixed regional configuration, these data are augmented with the regional configuration being solved for (i.e., the size of treatment plant at each location). The program would work essentially the same for both problems (i.e., performing a dual maximization using a column generation LP) only the regionalization problem would have its dual evaluated by dynamic programming as described in Section 3.2 while the degree of treatment problem would have its dual evaluated by a series of univariate minimizations as described in Section 3.4.

With such a program the analyst would perform the various bookkeeping and branching decisions by hand. By direct examination of the output from a dual maximization he could observe if a duality gap occurred, what the resulting alternate regional configurations were and what modified DO goals had been solved for optimally. From a research point of view this approach permits easy experimentation with different branching and partitioning rules. From an implementation point of view it allows the analyst to make judicious choices for branchings and partitionings based on his insight into the problem. Of course if the number of branchings became very large the process would become unwieldy for the analyst (to say nothing of the large computation time involved). However, as demonstrated in Chapter 5, it appears that in general only a few branchings will be required to either solve a regionalization problem or obtain very tight bounds.

CHAPTER 4. EXTENSIONS OF THE MODEL

4.1 Branched Systems

In the problem formulation described in Chapter 2 it was required that the waste sources lie on a linear configuration (or a discrete number of independent linear configurations) along the river. In this section this restriction is relaxed to allow the sources to have a branched configuration as shown in Fig. 4.1. The following analysis will apply to the case of a single branch only but the general idea can be extended for multiple branches.

Consider the single branch source configuration of Fig. 4.2 The sources along the main stem have been numbered in order starting with the upstream source while the branch sources are denoted with primes. In general, let there be $j = 1,\ldots,L$ sources on the main stem up to the intersection with the branch, $j = L + 1,\ldots,N$ remaining sources on the stem, and $j = 1',\ldots,N'$ sources on the branch. Each source is a potential location for a regional plant and the same regionalization restrictions apply as before - bypassing of sources is not permitted and treatment plants must be large enough to accommodate at least the flow piped in from other locations. In the unbranched problem this was sufficient to allow calculation of the BOD entering plant j based on knowledge of the flow treated at and upstream of j. However, this is no longer so for the branched problem. The question arises as to the order of piping between the branch and the main stem. For example, for the system in Fig. 4.2, if a plant is built at location 1 with capacity greater than $q_1 + q_2$ then it is not clear whether the excess capacity treats flow from source 3, source 2' or some combination of these. Another restriction must be

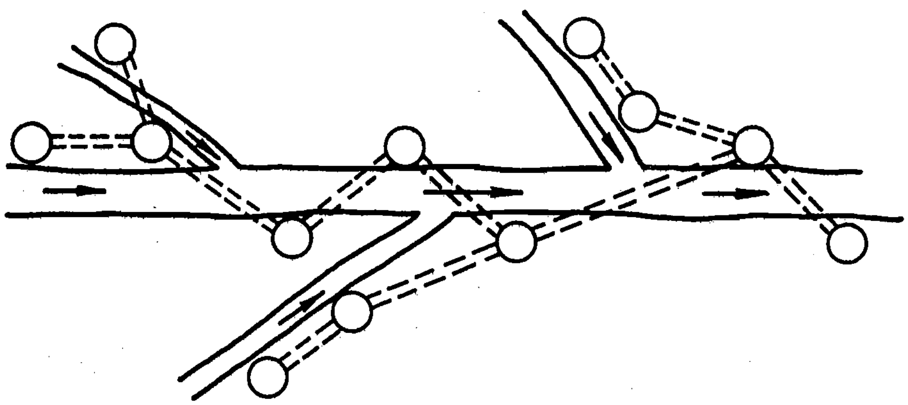


Figure 4.1 A Branched Configuration of Waste Sources

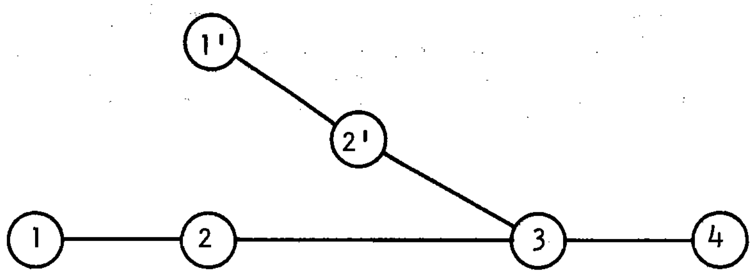


Figure 4.2 A Single Branch Source Configuration

specified to resolve this question.

One approach is to fix the allowable order of piping in advance. For instance one could demand the following two conditions: (i) The order of piping to location 1' is 1',2',...,N', L + 1, L + 2,...,N. (This implies that sources 1,2,...,L cannot send their wastes to any location 1',2',...,N' on the branch. This omission is made here to simplify the notation but in actual practice one could allow these sources to ship to locations on the branch provided the order with respect to the sources L+1,...,N is specified.) (ii) The order of piping to location 1 is 1,2,...,L, any waste generated from the branch sources and not treated on the branch, L + 1,...,N. With these specifications, the regionalization problem can be written as a mathematical program similar to that in Section 2.5 with the variables (y, z, w, and yp) denoting locations on the main stem and (y', z', w', and yp') designating locations on the branch. The dissolved oxygen requirements would appear as

$$\sum_{j=1}^{N} a_{ij}(s_j - z_j) + \sum_{j=1}^{N'} a_{ij}(s_j' - z_j') \geq \Delta c_i \qquad i=1,\ldots,M,M+1,\ldots,M+M'$$

where there are M points on the main stem of the river and M' points on the tributary for which dissolved oxygen goals are specified. Introducing M + M' dual variables, a dual function can be formed as in Section 3.2.

To evaluate the dual function a quasi-two-state variable dynamic program must be solved. We begin with the branch locations 1',...,N' and solve the following recursion

$$F_j(\hat{y}'_j) = \min_{y'_j} [R(\hat{y}'_j, y'_j) + F_{j-1}(\hat{y}'_j - y'_j)]$$

$$\text{s.t.} \quad yp'_j = \sum_{i=1'}^{j} q'_i - \hat{y}'_j \tag{4.1}$$

$$y'_j(q_j - yp'_j) \geq 0 \tag{4.2}$$

$$y'_j(y'_j + yp'_j) \geq 0 \tag{4.3}$$

$$0 \leq y'_j \leq \hat{y}'_j \tag{4.4}$$

where the initial conditions are

$$y'_o = 0 \quad , \quad F_o(0) = - \sum_{i=1}^{M+M'} u_i \, b_i$$

and the final condition is free; that is $F_{N'}(\hat{y}'_{N'})$ is variable in $\hat{y}'_{N'}$. The state space $[0, \sum_{i=1'}^{N'} q_i + \sum_{i=L+1}^{N} q_i]$ is discretized and thus $\hat{y}'_{N'}$ takes on discrete values in this interval. The return function is as before

$$R(\hat{y}'_j, y'_j) = \min_{z'_j} \{ P_j(yp'_j) + T_j(y'_j, z'_j, w'_j) + (\sum_{i=1}^{M+M'} u_i \, a_{ij}) \, z'_j \}$$

$$\text{s.t.} \quad yp'_j = \sum_{i=1'}^{j} q_i - \hat{y}'_j$$

$$w'_j = W'(\hat{y}'_j) - W'(\hat{y}'_j - y'_j)$$

$$L_j \leq 1 - z'_j/w'_j \leq U_j$$

$$z'_j \geq 0$$

where the function $W'(\cdot)$ is obtained by plotting $\sum_{i=1'}^{j} q_i$ against $\sum_{i=1'}^{j} s_i$ as j runs from 1' to N' and then continuing from L + 1 to N.

For each resulting $\hat{y}'_{N'}$ a dynamic programming recursion is solved over locations 1 to N with the source waste quantities at locations L + 1 to N modified as follows. Let

$$Q' = \sum_{i=1'}^{N'} q_i - \hat{y}'_{N'}$$

and

$$S' = W'\left(\sum_{i=1'}^{N'} q_i\right) - W'(\hat{y}'_{N'}).$$

If $Q' \geq 0$ then the source flow at location L + 1 becomes $Q' + q_{L+1}$ and the source BOD becomes $S' + s_{L+1}$. If $Q' < 0$ then beginning at source L + 1 and continuing to source N the source waste quantities are reduced (possibly to zero) by amounts Q' and S'. In other words, with reference to Fig. 4.2, if Q' was -20 mgd and q_3 and q_4 were 10 mgd and 30 mgd respectively then q_3 would become zero and q_4 would become 20 mgd.

The recursion relation is now a function of $y'_{N'}$

$$F_j(\hat{y}_j; \hat{y}'_{N'}) = \min_{y_j} [\, R(\hat{y}_j, y_j; \hat{y}'_{N'}) + F_{j-1}(\hat{y}_j - y_j; \hat{y}'_{N'}) \,]$$

s.t. constraints similar to (4.1 to 4.4) where q_j for j = L + 1,...,N is a function of $y'_{N'}$ as described above. Similarly, when evaluating the return function $R(\hat{y}_j, y_j; \hat{y}'_{N'})$ the function $W(\cdot)$ is also a function of $y'_{N'}$ since the source BOD's for locations L + 1,...,N have been modified as described above. The initial conditions are $\hat{y}_o = 0$, $F_o(0, \hat{y}'_{N'}) = F_{N'}(\hat{y}'_{N'})$ and the final condition is

$$\hat{y}_N(\hat{y}'_{N'}) = \sum_{i=1}^{N} q_i$$

where the $q_i (i = L + 1,\ldots,N)$ are the modified values depending on $\hat{y}'_{N'}$. Finally the value of the dual function is given by

$$h(u) = \min_{\hat{y}'_{N'}} f_N(\hat{y}_N(\hat{y}'_{N'});\ \hat{y}'_{N'}) .$$

The remaining steps for solving the regionalization problem are as described in Chapter 3.

The introduction of a single branch has increased computations in evaluating the dual by a factor equal to the number of possible values of $\hat{y}'_{N'}$. If more than one branch is considered the computations will increase in an exponential manner. A potential method for reducing computations is the use of Fibonacci search to minimize $F_N(\hat{y}_N;\ \hat{y}'_{N'})$ over $\hat{y}'_{N'}$, assuming it is unimodal. In any event, the significant increase in effort required to evaluate the dual function coupled with the fact that the dual must be evaluated a number of times poses serious threats to the computational feasibility of extending the method to branched systems.

We now demonstrate how the methodology of Chapter 3 could be applied to a more general formulation of the regionalization problem. We assume that there are N source locations and each is a potential location for a regional plant. Each source produces q_j mgd of wasteflow with a BOD concentration of s_j lb/mil gal. The sources need not have a linear configuration. Denote the flow piped from location j to location k as yp_{jk} and the associated cost as $P_{jk}(yp_{jk})$. The resulting mathematical program is

$$\text{Minimize Cost} = \sum_{j=1}^{N} \sum_{k=1}^{N} P_{jk}(yp_{jk}) + \sum_{j=1}^{N} T_j(y_j, z_j, w_j)$$

Subject to

$$\sum_{k=1}^{N} yp_{jk} = q_j \qquad j=1,\ldots,N \qquad (4.5)$$

$$y_k = \sum_{j=1}^{N} yp_{jk} \qquad k=1,\ldots,N \qquad (4.6)$$

$$w_k = \sum_{j=1}^{N} yp_{jk}\, s_j \qquad k=1,\ldots,N \qquad (4.7)$$

$$L_j \le 1 - z_j/w_j \le U_j \qquad j=1,\ldots,N \qquad (4.8)$$

$$\left.\begin{array}{ll} yp_{jk} \ge 0 & j=1,\ldots,N \\ z_j \ge 0 & k=1,\ldots,N \end{array}\right\} \qquad (4.9)$$

$$\sum_{j=1}^{N} a_{ij}(q_j\, s_j - z_j) \ge \Delta c_i \qquad i=1,\ldots,M \qquad (4.10)$$

where $T_j(\cdot)$, y_j, z_j, w_j, L_j, U_j, a_{ij}, and Δc_i are previously defined.

The objective function consists of the piping costs plus the treatment plant costs. Constraint (4.5) requires that all source flow be passed through a treatment plant. Eq. (4.6) determines the quantity of flow to be treated at any location while (4.7) determines the associated influent BOD. Constraints (4.8) and (4.9) put limits on allowable BOD removal and require nonnegativity, respectively. Constraint (4.10) is the DO improvement requirement. Note that in this formulation flows which may be piped over common routes are not combined in a single pipe. Specifying each yp_{jk} and z_j determines a solution.

To solve the above program using the dual method of Chapter 3 requires that the following dual function be formed:

$$h(u) = \min_{yp,z} \sum_{j=1}^{N} \sum_{k=1}^{N} P_{jk}(yp_{jk}) + \sum_{j=1}^{N} [T_j(y_j, z_j, w_j) + (\sum_{i=1}^{M} u_i a_{ij}) z_j]$$

$$- \sum_{i=1}^{M} [u_i(\Delta c_i - \sum_{j=1}^{N} a_{ij} q_j s_j)]$$

$$\text{s.t.} \quad (4.5) - (4.9).$$

Evaluating h(u) would certainly be easier than solving the original primal. However, because of the nonconvexities and the large number of constraints this is still a difficult problem. Note that it is essentially the pure facility location problem since the decisions on degree of treatment at each location can be made separately of each other. Providing an efficient means were available for evaluating h(u) one could proceed with the rest of the solution algorithm as described in Chapter 3.

4.2 Partial Regionalization and Bypass Piping

There may be waste dischargers on the river who for one reason or another cannot participate in a regionalization plan. For example some industries may produce wastes which must receive special, separate treatment. Or, as demonstrated in the Delaware Estuary analysis made in the next chapter, cost functions for the treatment of mixed industrial and domestic wastewaters may not be available. In such a case we still desire to find the amount of BOD removal each discharger should provide so that the dissolved oxygen goals are met by all dischargers (including those in the regional treatment plant system) at minimum cost.

Let the waste sources which cannot participate in a regionalization plan be designated $N + 1,\ldots,N+N'$. The mathematical programming formulation of the problem is modified as follows:

(i) the objective function becomes

$$\text{Minimize Cost} = \sum_{j=1}^{N} P_j(yp_j) + T_j(y_j, z_j, w_j) + \sum_{j=N+1}^{N+N'} T_j(z_j)$$

(ii) the dissolved oxygen requirements become

$$\sum_{j=1}^{N+N'} a_{ij}(s_j - z_j) \geq \Delta c_i \qquad i=1,\ldots,M$$

(iii) bounds on BOD removal for sources $j = N+1,\ldots,N+N'$ are

$$L_j \leq 1 - z_j/s_j \leq U_j .$$

Notice that the cost of treatment at the sources $N+1$ to $N+N'$ is a function only of the BOD discharged since the influent waste quantities are the known source quantities, q_j and s_j.

The solution procedure follows the method outlined in Chapter 3 with the dual being evaluated by first solving a dynamic program over locations 1 to N as described in Chapter 3 and then adding to its value the results of N' univariate minimizations of the form

$$\underset{z_j}{\text{Minimize}} \quad T_j(z_j) + \left(\sum_{i=1}^{M} u_i a_{ij}\right) z_j$$

$$\text{Subject to} \quad L_j \leq 1 - z_j/s_j \leq U_j$$

where $j = N+1,\ldots,N+N'$.

The use of bypass piping, that is, treating waste at location j and then piping it for discharge to reach i, offers additional potential savings for meeting a required dissolved oxygen goal. It can be introduced into the model by replacing the variable z_j by z_{ij} (the BOD in the effluent treated at location j and discharged in reach i) and introducing the variable p_{ij} which is 1 if the effluent of treatment at location j is bypassed to reach i and 0 otherwise. The mathematical programming formulation of the problem (including partial regionalization) is modified as follows:

(i) the objective function becomes

$$\text{Minimize Cost} = \sum_{j=1}^{N} \{P_j(yp_j) + \sum_{i=1}^{M} [T_j(y_j, z_{ij}, w_j) + P_{ij}(y_j)]\, p_{ij}\}$$

$$+ \sum_{j=N+1}^{N+N'} \sum_{i=1}^{M} [T_j(z_{ij}) + P_{ij}(q_j)]\, p_{ij}$$

where $P_{ij}(y_j)$ is the cost of piping y_j units of waste flow from location j to reach i in the river.

(ii) the dissolved oxygen requirements become

$$\sum_{j=1}^{N+N'} a_{ij}(s_j - p_{ij}\, z_{ij}) \geq \Delta c_i \qquad i=1,\ldots,M$$

(iii) bounds on BOD removal are

$$\left.\begin{array}{ll} L_j \leq 1 - z_{ij}/w_j \leq U_j & J=1,\ldots,N \\ L_j \leq 1 - z_{ij}/s_j \leq U_j & J=N+1,\ldots,N+N' \end{array}\right\} \quad i=1,\ldots,M$$

(iv) the following constraints are added

$$\sum_{i=1}^{M} p_{ij} = 1 \qquad j=1,\ldots,N+N'$$

$$p_{ij} = 0,\ 1 \qquad \forall\ i,\ j.$$

The modifications necessary in the evaluation of the dual are

(i) in the dynamic programming portion of the calculation the return function is given by

$$R_j(\hat{y}_j,\ y_j) = \min_{z_{ij},p_{ij}} \ P_j(yp_j) + \sum_{i=1}^{M} [T_j(y_j,\ z_{ij},\ w_j) + P_{ij}(y_j)$$

$$+ (\sum_{i=1}^{M} u_i\ a_{ij})\ z_{ij}]\ p_{ij}$$

$$\text{s.t.} \quad L_j \le 1 - z_{ij}/w_j \le U_j \qquad i=1,\ldots,M$$

$$\sum_{i=1}^{M} p_{ij} = 1 \ , \quad p_{ij} = 0,\ 1 \qquad i=1,\ldots,M$$

and yp_j and w_j can be found from knowledge of $\hat{y}_j$ and y_j. The above is evaluated by enumerating over the M possible values of p_{ij} and solving a univariate minimization in z_{ij} at each enumeration.

(ii) the remaining N' separate minimizations (for the sources which cannot regionalize) become

$$\min_{z_{ij},p_{ij}} \sum_{i=1}^{M} [T_j(z_{ij}) + P_{ij}(q_j) + (\sum_{i=1}^{M} u_i\ a_{ij})z_{ij}]\ p_{ij}$$

$$\text{s.t.} \qquad L_j \le 1 - z_{ij}/s_j \le U_j \qquad i=1,\ldots,M$$

$$\sum_{i=1}^{M} p_{ij} = 1 \ , \quad p_{ij} = 0,\ 1 \qquad i=1,\ldots,M$$

and the same enumeration method as described in (i) can be used to solve each of these.

It is evident that this extension to handle bypass piping will also increase the computational effort. Preliminary application to data from the Delaware Estuary indicates that computational feasibility of the above formulation is questionable. Improved performance was obtained by limiting the use of bypass piping to treatment plants which treat their source waste only, the rationale being that regionalization in some sense serves the same purpose of shifting the waste discharge points as bypass piping does. Another strategy would be to limit the number of allowable discharge sections for each polluter to some subset of the entire M reaches.

A second problem concerns the question of duality gap resolution. In theory the same branch and bound procedure as described in Chapter 3 can be applied. When a duality gap occurs there will exist two alternate combined regional treatment plant and bypass piping configurations at the maximum of the dual. Again the optimal level of BOD reduction problem can be solved for each by convex programming and the results used to establish an upper bound. Then two additional problems are created, each with additional constraints on allowable treatment plant sizes and allowable reaches to which effluent can be piped (e.g., p_{ij} set to either 0 or 1 for some j). These are solved by the dual method and the procedure continues in this manner until the lower bound is close enough to the upper bound. Again, some preliminary computational experience has indicated poorer performance of the branch and bound method for gap resolution when the model includes bypass piping. Specifically, the difference between the initially established lower and upper bounds is greater in the case of bypass piping.

4.3 Effluent Charges

In all that has preceded it has been tacitly assumed that some central planning authority exists which, having perfect information concerning the treatment costs of all polluters, is able to solve the regional wastewater treatment model and directly implement the resulting least cost regional plant configuration and BOD reduction plans to meet the desired dissolved oxygen goal. Several authors (Kneese (1964), Hass (1970)) have noted that an appealing vehicle for obtaining desired water quality is the imposition of pollution taxes or effluent charges by a central authority. The response of an individual polluter to an announced effluent charge per unit of BOD discharged would be to reduce BOD discharges to a level where the marginal cost of BOD reduction is equal to the unit charge. The optimal set of charges results in meeting the specified DO goals with minimum treatment costs.

A polluter faced with having to pay a certain charge per lb of BOD discharged would view his treatment cost function as

$$\text{Cost of BOD removal} = T_j(y_j,\ z_j,\ w_j) + (\text{charge})* z_j$$

where T_j is the cost of BOD reduction as a function of wasteflow and BOD removal efficiency. With reference to the dual function developed for the regionalization problem.

$$h(u) = \min_{y,z} \sum_{j=1}^{N} [P_j(yp_j) + T_j(y_j,\ z_j,\ w_j) + (\sum_{i=1}^{M} u_i\ a_{ij})\ z_j] - \sum_{i=1}^{M} u_i\ b_i$$

$$\text{s.t.} \quad (3.11) - (3.16)$$

and in view of the polluter's cost function with an effluent charge as written above, the quantity $(\sum_{i=1}^{M} u_i a_{ij})$ can be interpreted as the effluent charge levied against polluter j. Knowing this charge the polluters can plan their optimal regionalization and BOD reduction strategy by evaluating h(u). The column generation LP used to maximise h(u) attempts to find the least cost convex combination of these regionalization and BOD reduction plans so that DO goals are met. The dual variables of this linear program which correspond to the DO constraints, u, approximate the marginal cost of requiring the desired DO improvement. They serve to define the next set of effluent charges to be levied, $\sum_{i=1}^{M} u_i a_{ij}$. The process continues until the convex combination of polluter responses results in an exact approximation, e.g., max $h(u) = \min \sum_{i=1}^{M} P_j + T_j$, which is simply the optimality condition derived from duality theory.

However, because of the nonconvexities associated with the regionalization problem, this condition may not be realizable. In terms of effluent charges this means that there need not exist a set of effluent charges whose imposition would result in the optimal pollution control plan, given that this plan is found by evaluating h(u) as described above. Such a situation corresponds to the existence of duality gap. Thus it is not always possible to obtain optimal regionalization of wastewater treatment to meet DO goals by means of linear effluent charges.

4.4 Related Regionalization Problems

Two problems are considered which are variations of the original regionalization problem. First is the problem of determining the minimum cost regionalization pattern when treatment levels are set in advance.

As before, it is assumed that the waste sources lie in a linear configuration with each source $j = 1,\ldots,N$, producing q_j units of waste flow. For this problem the cost of constructing and operating a treatment plant at location j is a function of the quantity of flow treated (y_j) only, since the level of waste treatment is set in advance. This cost, $T_j(y_j)$, is assumed concave with respect to y, thus exhibiting the economies of scale which make regionalization worthwhile. Similarly the cost of piping a quantity of flow yp_j between locations j and j+1 is assumed to be concave.

Under these conditions the regionalization restrictions imposed on the more general regionalization problem with water quality goals are automatically satisfied at optimality. Consider first the restriction that a plant must be at least as large as the piped-in flow from other locations. Figure 4.3a shows a situation which violates this condition. The total cost involved is

$$\text{Cost} = T_j(y_j) + P_j(yp_j) + T_{j+1}(y_{j+1})$$

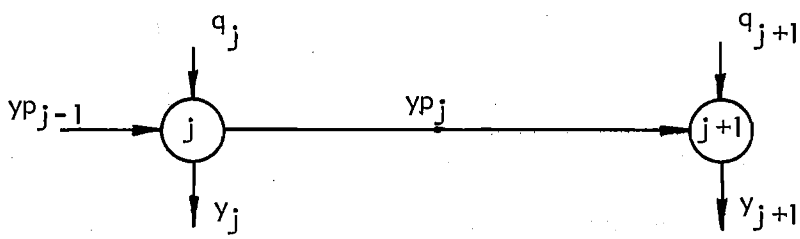
where $yp_j = yp_{j-1} + q_j - y_j$

$$y_{j+1} = q_{j+1} + yp_j \,.$$

Now if one more unit of flow was sent from j to j+1 the change in cost would be

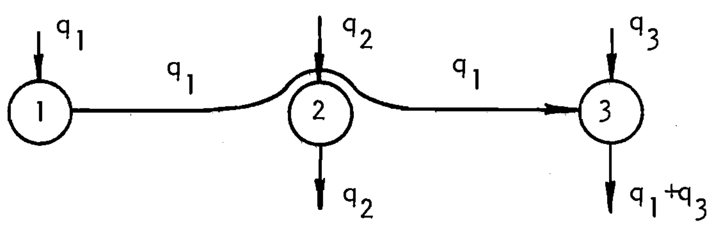
$$\Delta\,\text{cost} = -\dot{T}_j(y_j) + \dot{P}_j(yp_j) + \dot{T}_{j+1}(y_{j+1})$$

where $\dot{T}(y_j) = \left.\frac{dT}{dy}\right|_{y_j}$ and $\dot{P}(yp_j) = \left.\frac{dP}{dyp}\right|_{yp_j}$.



$y_j < yp_{j-1}$

(a)



(b)

Figure 4.3 Regional Facility Patterns Which Violate the Regionalization Restrictions

If $\dot{T}_j(y_j) \geq \dot{P}_j(yp_j) + \dot{T}_{j+1}(y_{j+1})$

then it pays to ship the additional unit to j+1. However, since all functions are concave (decreasing marginal costs with increasing quantity of flow) costs are minimized when all of y_j is shipped. On the other hand, if $\dot{T}_j(y_j) \leq \dot{P}_j(yp_j) + \dot{T}_{j+1}(y_{j+1})$ then it pays to reduce the amount shipped to j+1 to zero.

Notice that the above argument implies an even stronger optimality condition, namely that sources never send only a portion of their flow to other locations. Also, it always pays for a source to ship to another location when

$$\dot{T}_j(q_j) > \dot{P}_j(0^+) + \dot{T}_{j+1}(q_{j+1}) .$$

The other regionalization restriction prohibited bypassing of sources. A violation of this is pictured in Fig. 4.3b. Using the above results, for this situation to hold it follows that

$$\dot{P}_1(0^+) + \dot{P}_2(0^+) + \dot{T}_3(q_3) - \dot{T}_1(q_1) < 0$$

(i.e., it pays to ship from source 1 to source 3) and

$$\dot{P}_1(0^+) + \dot{T}_2(q_2) - \dot{T}_1(q_1) > \dot{P}_1(0^+) + \dot{P}_2(0^+) + \dot{T}_3(q_3) - \dot{T}_1(q_1)$$

(i.e., it is more costly to ship from source 1 to source 2 and treat there than ship from source 1 for treatment at source 3)

and

$$\dot{P}_2(q_1) + \dot{T}_3(q_3 + q_1) - \dot{T}_2(q_2) > 0$$

(i.e., it does not pay to ship 2 to 3 when 1 already ships to 3).

From the second and third inequalities we get

$$\dot{P}_2(0^+) + \dot{T}_3(q_3) < \dot{T}_2(q_2) < \dot{P}_2(q_1) + \dot{T}_3(q_3 + q_1).$$

But because of the concave cost functions we must have

$$\dot{P}_2(0^+) \geq \dot{P}_2(q_1)$$

$$\dot{T}_3(q_3) \geq \dot{T}_3(q_3 + q_1)$$

and thus there is a contradiction. Hence a solution which bypasses any sources such as in Fig. 4.3b can never be optimal. We have shown that the two regionalization restrictions imposed on the more general regionalization problem where degree of waste treatment was considered are in fact optimality conditions for the regionalization problem where degree of waste treatment is not considered.

The mathematical programming statement of this problem is

$$\text{Minimize Cost} = \sum_{j=1}^{N} P_j(yp_j) + T_j(y_j)$$

Subject to

$$yp_j = \sum_{i=1}^{j} q_i - \sum_{i=1}^{j} y_i \qquad j=1,\ldots,N$$

$$\sum_{j=1}^{N} y_j = \sum_{j=1}^{N} q_j$$

$$y_j \geq 0 \qquad j=1,\ldots,N$$

which can be solved by dynamic programming.

Introducting the state variable

$$\hat{y}_j = \sum_{i=1}^{j} y_i$$

the recursion relation is

$$F_j(\hat{y}_j) = \min_{y_j} [P_j(yp_j) + T_j(y_j) + F_{j-1}(\hat{y}_j - y_j)]$$

subject to $\quad yp_j = \sum_{i=1}^{j} q_i - \hat{y}_j$

$$0 \leq y_j \leq \hat{y}_j .$$

The initial conditions are

$$\hat{y}_o = 0 \quad , \quad F_o(0) = 0$$

and the final condition is

$$\hat{y}_N = \sum_{j=1}^{N} q_j .$$

The optimum value is given by $F_N(\hat{y}_N)$. From the results obtained above it is sufficient to discretize the state space to $\{0, q_1, q_1 + q_2, \ldots, \sum_{j=1}^{N} q_j\}$.

The second problem is determining the least cost regionalization pattern and uniform level of BOD reduction provided by all dischargers to meet a given DO goal. Making the same assumptions and regionalization restrictions as before, the mathematical programming model is

$$\text{Minimize cost} = \sum_{j=1}^{N} P_j(yp_j) + T_j(y_j, r)$$

subject to

$$\sum_{j=1}^{N} y_j = \sum_{j=1}^{N} q_j \tag{4.11}$$

$$yp_j = \sum_{i=1}^{j} q_i - \sum_{i=1}^{j} y_i \qquad j=1,\ldots,N \tag{4.12}$$

$$w_j = W\left(\sum_{i=1}^{j} y_i\right) - W\left(\sum_{i=1}^{j-1} y_i\right) \qquad j=1,\ldots,N \tag{4.13}$$

$$\left.\begin{array}{l} y_j(q_j - yp_j) \geq 0 \\ y_j(y_j + yp_j) \geq 0 \end{array}\right\} \qquad j=1,\ldots,N \tag{4.14}$$

$$y_j \geq 0 \qquad j=1,\ldots,N \tag{4.15}$$

$$L \leq r \leq U \tag{4.16}$$

$$z_j = w_j(1 - r) \qquad j=1,\ldots,N \tag{4.17}$$

$$\sum_{j=1}^{N} a_{ij}(s_j - z_j) \geq \Delta c_i \qquad i=1,\ldots,M \tag{4.18}$$

where r = the fraction of BOD removed by treatment and is the same at all locations. Notice that treatment costs are once again a function of the hydraulic size of the plant and the degree of BOD removal provided. Also, reintroduction of water quality goals and a variable treatment level means that the regionalization restrictions are no longer optimality conditions so they are specified through constraints (4.12) - (4.14).

At first glance this appears to be an easier problem than the more general case where treatment levels are allowed to vary among dischargers. One solution approach might be to perform a univariate search on the treatment level r so that the resulting discharges from the minimum cost regional facility pattern just meet the DO goals. The minimum cost facility pattern would be found using the straightforward dynamic programming solution discussed above for the problem where treatment levels are fixed. Such a method would not produce a truly minimum cost solution because the coupling of the level of treatment, the regionalization configuration and the resulting BOD discharges and DO levels (constraints (4.13), (4.17) and (4.18)) is ignored. Instead a more expensive, but feasible, solution would result.

Substituting (4.17) into (4.18) and dualizing with respect to the latter would result in the following dual problem.

$$\max_{u\geq 0} \quad h(u) = \min_{r,y} \sum_{j=1}^{N} [P_j(yp_j) + T_j(y_j, r) + (\sum_{i=1}^{M} u_i a_{ij})(1-r)w_j] - \sum_{i=1}^{M} u_i b_i$$

$$\text{s.t.} \quad (4.11) - (4.16).$$

Evaluation of the dual is now a more difficult task than in the more general regionalization problem. It can be rewritten as

$$h(u) = \min_{L\leq r\leq U} [\min_{y} \sum_{j=1}^{N} P_j + T_j + (\sum_{i=1}^{N} u_i a_{ij})\ (1-r)\ w_j\ ] - \sum_{i=1}^{M} u_i b_i$$

$$\text{s.t.} \quad (4.11) - (4.15).$$

For fixed r, the minimization over y can be obtained by solving a dynamic program.

$$\text{Minimize} \quad \sum P_j(yp_j) + T_j(y_j, r) + (\sum_{i=1}^{M} u_i a_{ij})(1-r)\ w_j$$

$$\text{s.t.} \quad \left.\begin{aligned} \hat{y}_j &= \hat{y}_{j-1} + y_j \qquad\qquad j=1,\ldots,N \\ \hat{y}_o &= 0 \;,\; \hat{y}_N = \sum_{j=1}^{N} q_j \end{aligned}\right\} \tag{4.19}$$

$$yp_j = \sum_{i=1}^{j} q_i - \hat{y}_j \qquad j=1,\ldots,N \tag{4.20}$$

$$w_j = W(\hat{y}_j) - W(\hat{y}_j - y_j) \qquad j=1,\ldots,N \tag{4.21}$$

$$\left.\begin{aligned} y_j(q_j - yp_j) &\geq 0 \\ y_j(y_j + Yp_j) &\geq 0 \end{aligned}\right\} \qquad j=1,\ldots,N \tag{4.22}$$

$$y_j \geq 0 \qquad j=1,\ldots,N \tag{4.23}$$

where as before $\hat{y}_j = \sum_{i=1}^{j} y_i$ and is the state variable. The return function at each stage is

$$R_j(\hat{y}_j,\ y_j;\ r) = P_j(yp_j) + T_j(y_j,\ r) + (\sum_{i=1}^{M} u_i\ a_{ij})(1 - r)\ w_j$$

where $\quad yp_j = \sum_{i=1}^{j} q_i - \hat{y}_j$

$$w_j = W(\hat{y}_j) - W(\hat{y}_j - y_j).$$

The recursion relation is

$$F_j(\hat{y}_j;\ r) = \min_{y_j} [\ R_j(\hat{y}_j,\ y_j;\ r) + F_{j-1}(\hat{y}_j - y_j;\ r)\ ]$$

$$\text{s.t.} \quad (4.20) - (4.23)$$

$$0 \leq y_j \leq \hat{y}_j$$

with initial conditions

$$\hat{y}_o = 0 \quad , \quad F_o(0,r) = 0$$

and final condition

$$\hat{y}_N = \sum_{j=1}^{N} q_j \ .$$

The value of the dual function is now

$$h(u) = \min_{L \leq r \leq U} F_N(\hat{y}_N;\ r).$$

We may find h(u) by using a univariate search on r. Notice that this implies that a series of dynamic programs must be solved at each evaluation of h(u). Solution is made considerably easier when the treatment costs with respect to BOD removal at each location differ only by some constant multiplier. Under this condition the value of r will not affect the solution of the regionalizing dynamic program. Thus the latter is solved only once at any value of r, then the minimizing r is found using the resulting regionalization solution.

The remaining details of maximizing h(u) and resolving duality gaps follow the same procedure as described in Chapter 3. The one exception occurs in the gap resolution procedure when the minimum cost uniform level of treatment to meet the given DO goals must be found for a given regionalization pattern. This is a trivial problem since the optimal r satisfies

$$r = \max_{i=1,\ldots,M} \left[1 - \frac{b_i}{\sum_{j=1}^{N} a_{ij} w_j} \right]$$

where w_j is determined by the known regionalization pattern.

CHAPTER 5. APPLICATION OF THE MODEL

5.1 Delaware Estuary

The basic version of the regionalization model developed in Chapters 2 and 3 will be applied to data from a 72 mile stretch of the Delaware Estuary beginning at Trenton, New Jersey. A number of authors have analyzed the costs of various pollution control policies for the estuary. Those who have examined regionalization include Graves et al. (1970) and Whitlatch (1973). A summary of their approaches was given in Section 1.2. We will use the same data as they did so that a comparison between the methods can be made. The aim here is not to perform a comprehensive analysis of regionalization strategies for the Delaware but rather to show the potential utility of the proposed model in such efforts as compared with currently available methods.

The physical model for relating changes in dissolved oxygen to changes in BOD inflows to the river is based on a finite difference approximation to the differential form of the mass balance equations for DO and BOD. Details of this approach are described in Thomann (1972). The resulting model has the form

$$\sum_{j=1}^{M} a_{ij}(\Delta w_j) = \Delta c_i \qquad i=1,\ldots,M$$

where a_{ij} = a DO transfer coefficient, $\frac{mg/\ell}{lb/day}$

Δw_j = the total change in BOD discharge in reach j (and could be made up of a number of individual discharges which are located in reach j), lb/day

Δc_i = the change in dissolved oxygen in reach i, mg/ℓ

M = number of reaches.

This is similar to the constraints developed for the regionalization model in Section 2.3, except that Δw_j lumps together all $(s_k - z_k)$ for locations k in reach j. Recall that use of a model such as this assumes that the variation in riverflow as a result of regionalization will have negligible effect on the DO transfer coefficients, a_{ij}. For the Delaware Estuary this appears to be a fair assumption (Graves, 1972).

The Delaware has been divided into 30 reaches with lengths between 10,000 and 20,000 feet. Information on the geometry, flows, and dispersion, BOD decay, and reaeration coefficients of each reach is given in Table 7-3, p. 172 of Thomann (1972). Of the 30 reaches only 18 will have waste dischargers in them so there is a total of (18 x 30) or 540 DO transfer coefficients required. Their values can be found in Table 22, p. 97 of Graves et al. (1970) and will not be repeated here.

There are 44 major waste dischargers along the Delaware. Specific information for each is given in Table 5.1. Of the 44, half discharge industrial wastes and half domestic. In keeping with the previous regionalization studies, only the domestic sources will be allowed to regionalize. In the previous studies only nine potential locations, distinct from the waste sources, for regional treatment plants were considered. To make the most of our regionalization model we consider these nine plus all of the 22 domestic waste sources as potential locations for regional plants. A schematic representation of these locations and distances between them as given by Whitlatch (1973) is shown in Fig. 5.1. Note that piping across the river is not allowed. Also we will not allow a source to split part of its flow between treatment at source and

Table 5.1 Source Data for the Delaware Estuary

Source	Reach Location	Domestic or Industrial	Flow, mgd	Raw BOD, lb/day	Present BOD Discharge lb/day
1	1	D	19.7	20400	3060
2	2	D	6.1	11333	1700
3	2	I	250.0	18333	2750
4	3	I	15.2	990	990
5	3	I	4.6	20933	3140
6	4	D	3.0	4444	2000
7	10	I	2.0	11615	7550
8	10	D	4.0	8920	2230
9	10	D	144.7	501000	125250
10	13	D	22.4	85014	59510
11	13	D	7.0	19392	12605
12	14	D	107.6	261708	170110
13	14	I	3.5	24477	15910
14	14	I	10.6	3560	3560
15	14	D	4.9	5867	1760
16	15	I	28.2	62643	21925
17	15	I	51.0	19846	21900
18	15	I	6.8	6085	3955
19	15	D	2.0	3450	2070
20	16	D	118.0	260333	156200
21	16	D	0.7	2671	1870
22	16	I	38.6	39462	25650
23	17	I	0.3	23143	8100
24	17	I	80.4	52554	34160
25	17	D	11.6	21067	3160
26	17	D	4.0	7167	1075
27	17	I	0.7	2700	1890
28	17	D	6.6	11923	7750
29	18	I	14.8	19363	7745
30	18	D	8.6	14550	10185
31	18	I	112.2	5538	3600
32	18	I	5.4	4838	3145
33	19	D	1.4	2800	1820
34	19	I	1.1	40813	32050
35	19	I	106.9	114920	28730
36	19	I	33.3	2890	2890
37	21	D	60.0	107463	85970
38	21	D	1.0	2383	1430
39	21	D	4.8	8480	8480
40	22	I	103.3	169231	110000
41	22	I	10.1	33775	6755
42	23	D	1.2	3117	1870
43	26	I	278.6	3846	2500
44	28	D	1.3	2883	1730
45	3	Potential locations for regional plants			
46	4				
47	5				
48	14				
49	16				
50	17				
51	17				
52	18				
53	23				

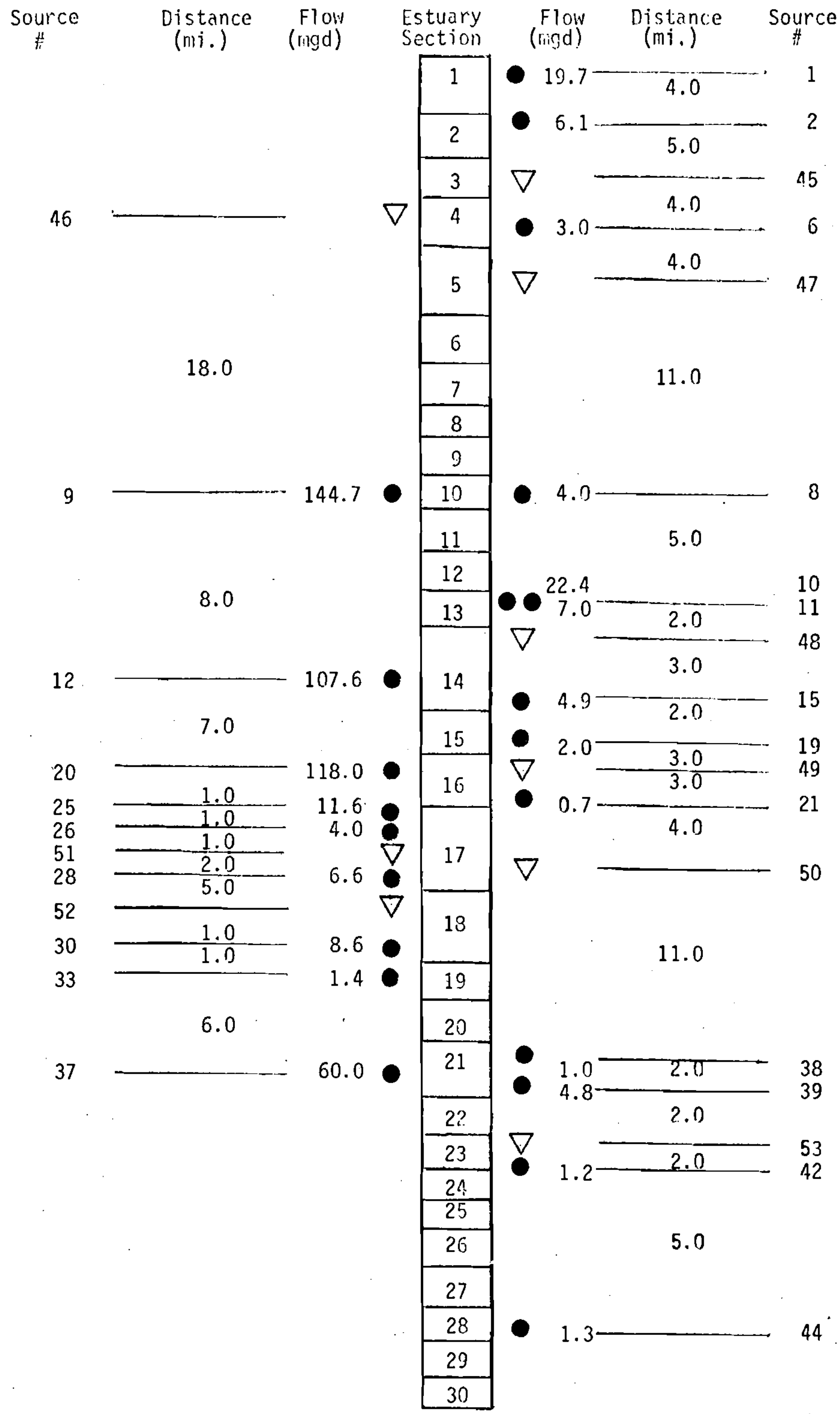


Figure 5.1 Schematic Representation of the Domestic Waste Sources along the Delaware Estuary

shipping to an adjacent location although this can be handled by the model.

The cost of BOD removal for each individual discharger is given as a piecewise linear function of quantity of BOD removed as follows:

$$T_{Lj} = CC(j,1)(s_j - z_j) \qquad 0 \le (s_j - z_j) \le BND(j,1)$$

$$= CC(j,1)\ BND(j,1) + CC(j,2)(s_j - z_j - BND(j,1)) \qquad BND(j,1) \le (s_j - z_j) \le BND(j,2)$$

$$= CC(j,1)\ BND(j,1) + CC(j,2)\ BND(j,2) + CC(j,3)(s_j - z_j - BND(j,2)) \qquad BND(j,2) \le (s_j - z_j) \le BND(j,3)$$

where T_{Lj} = annual cost of providing at-source BOD removal for discharger j, $

$CC(j, 1\text{-}3)$ = cost coefficients, $\$/\frac{\text{lb BOD}}{\text{day}}$

s_j = BOD currently discharged from location j, lb/day

z_j = BOD discharged after additional BOD removal at location j, lb/day.

The values of CC and BND for each of the 44 dischargers are given in Table 9, p. 66 of Graves et al. (1970) and will not be repeated here. Note that T_{Lj} represents the cost of additional BOD reduction over and above what is already being provided. To keep things comparable with the Graves et al. (1970) study no minimum BOD removals (such as primary treatment) are specified for at-source treatment.

The cost of BOD removal at a regional plant is

$$T_{Rj} = (393760)\ y_j^{.75}\ [(.5 - \frac{z_j}{w_j})^3 - (.5 - \frac{w_j}{w_j})^3]$$

where T_{Rj} = annual cost of treatment at a regional plant, \$

w_j = the total of the current BOD discharges (Σs_i) of all sources who ship their wastes to regional plant j, lb/day

$\bar{w}_j$ = the total of the raw source BOD's $(\Sigma \bar{s}_i)$ of all sources who ship their wastes to regional plant j, lb/day

$\bar{s}_j$ = raw source BOD of source j, lb/day

y_j = hydraulic size of regional plant j, mgd.

This cost represents the cost of additional BOD removal provided by a regional plant over and above what the individual sources who ship to the regional plant currently provide. It is derived by Graves et al. (1970) from data in Frankel (1965). A minimum of 50% BOD removal (relative to raw influent BOD) will be required at each regional plant so that treatment costs remain convex with respect to BOD removed. An upper limit of 98% removal is imposed.

Piping costs between adjacent locations are given as in Graves et al. (1970) by

$$P_j(yp_j) = 1865\, d_j \,|\, yp_j \,|^{.6}$$

where P_j = annual cost of piping between location j and j+1, \$

d_j = distance between j and j+1, miles

yp_j = flow piped between j and j+1, mgd.

Observe that all treatment and piping cost functions are continuous. Hence the power of our solution method to handle more general cost functions is not being utilized.

With these data the regionalization model will be used to determine what regionalization pattern and increase in BOD removal is necessary to provide a DO of 3 mg/ℓ in each reach of the Delaware at minimum cost. Based on the "current" (summer of 1964) DO levels, the required improvement for each reach to attain the 3 mg/ℓ goal is shown in Table 5.2. Reaches already at or above 3 mg/ℓ will be required to at least maintain their present levels (i.e., $\Delta c = 0$). One necessary addition to the development of the model is the calculation of $\bar{w}_j$. As for w_j, $\bar{w}_j$ can be expressed as

$$\bar{w}_j = \bar{W}(\hat{y}_j) - \bar{W}(\hat{y}_j - y_j)$$

where $\bar{W}(\circ)$ = the resulting piecewise linear curve when $\Sigma \bar{s}_i$ plotted against Σq_i.

The regionalization model as discussed in Chapters 2 and 3 has been coded in ASA Standard FORTRAN IV for implementation on an IBM 360/75. This same program is also capable of solving the minimum cost degree of treatment problem for a given regionalization pattern as would be required in the duality gap resolution procedure. Because of the piecewise linear costs an additional linear programming calculation as described in Section 3.6 must be made to check on duality gaps and extract optimal values of BOD removals. The stop criterion used in maximizing the dual is that the difference in the upper bound and the best dual value be less than .001% of the best dual value (see Section 3.3).

Solving the model the first time through gave a maximum dual value of \$1.280 x 10^6. Scanning the output revealed that the last two iterations gave dual values approximately equal to this but each corresponding primal solution had a different regionalization pattern. Thus the conditions of

Table 5.2 Required Improvement to Attain 3 mg/ℓ of DO for the Delaware Estuary

Reach	Required DO Improvement, mg/ℓ	Reach	Required DO Improvement, mg/ℓ
1	0.	16	2.0
2	0.	17	2.0
3	0.	18	1.8
4	0.	19	1.6
5	0.	20	0.8
6	0.	21	0.1
7	0.	22	0.
8	0.	23	0.
9	0.	24	0.
10	0.8	25	0.
11	0.7	26	0
12	1.5	27	0.
13	1.8	28	0.
14	1.9	29	0.
15	1.9	30	0.

a duality gap exist. Solving the minimum cost degree of treatment problems for each configuration gave resulting costs of $1.319 x 10^6 and $1.304 x 10^6. Thus a feasible solution (representing our best upper bound) with cost $1.304 x 10^6 is found and from our lower bound of $1.280 x 10^6 it is guaranteed to be within a tolerance of 1.87% of the global minimum. Total time to achieve this result was 77.31 seconds on an IBM 360/75.

To improve on this tolerance and perhaps our upper bound we can proceed with the branch-and-bound method of gap resolution. We will use a branching rule which branches on the first facility location with different size plants in the two alternate regional configurations. Examining these configurations showed that location 8 had a 33.4 mgd plant in the $1.304 x 10^6 solution and a 4.0 mgd plant in the $1.319 x 10^6 solution. Branching on this location created two new problems, problem 1a with the plant size at location 8 constrained to be less than 33.4 mgd and problem 1b with the size constrained to be greater than or equal to 33.4 mgd. It should be remembered that this branching and partitioning rule is, at this stage in our knowledge of the properties of the algorithm, strictly arbitrary. Any other procedure could be used to partition the set of feasible regional configurations so that the two configurations obtained above could not both appear in the same partition.

Proceeding to maximize the dual for problem 1b yielded a value of $1.304 x 10^6, the same cost and regional configuration as our upper bound. Thus no more branching can be done from this solution. Maximizing the dual for problem 1a gave a value of $1.299 x 10^6. This improved the lower bound. Examining this solution indicated a duality gap once again with the least cost degree of treatment for each of the two alternative configurations being \$1.312 x 10^6 and \$1.319 x 10^6. Thus no improvement is offered for our upper bound. At this stage we have a solution (\$1.304 x 10^6) guaranteed to be within a .38% tolerance of the global minimum. Total solution time, including the running of a linear program to extract the degree of treatment variables from the \$1.304 x 10^6 solution was 3.5 minutes. If we were willing to accept this tolerance, a reasonable level considering the uncertainty in the data and the degree of round-off in the computations, then we would be through, claiming that the \$1.304 x 10^6 solution was our global optimum. Alternatively we could continue the branching from the results of problem 1a. Using the same branching rule as before we branch on location 8 once again. Two more problems are created, 2a with allowable size of plant at location 8 less than 26.4 mgd and 2b with allowable size of plant between 26.4 and 33.4 mgd. The results of the dual maximizations for these problems give dual values greater than our existing upper bound. Hence we conclude that no improvement is possible and \$1.304 x 10^6 truly is the global optimum. Total computation time has now reached 5.46 minutes. The steps of the branch and bound procedure are displayed in Fig. 5.2.

An average of 34 iterations was required by the column generation LP to maximize each 30 variable dual problem. It should be noted that none of the potential time-saving methods described in Section 3.6 were used and that each time a regionalization or degree of treatment problem was solved the entire input data was read in.

The resulting solution is given in Fig. 5.3 and Table 5.3. There are 4 regional plants serving a total of 13 sources. Of the sources not served by a regional plant only 4 increase their degree of treatment.

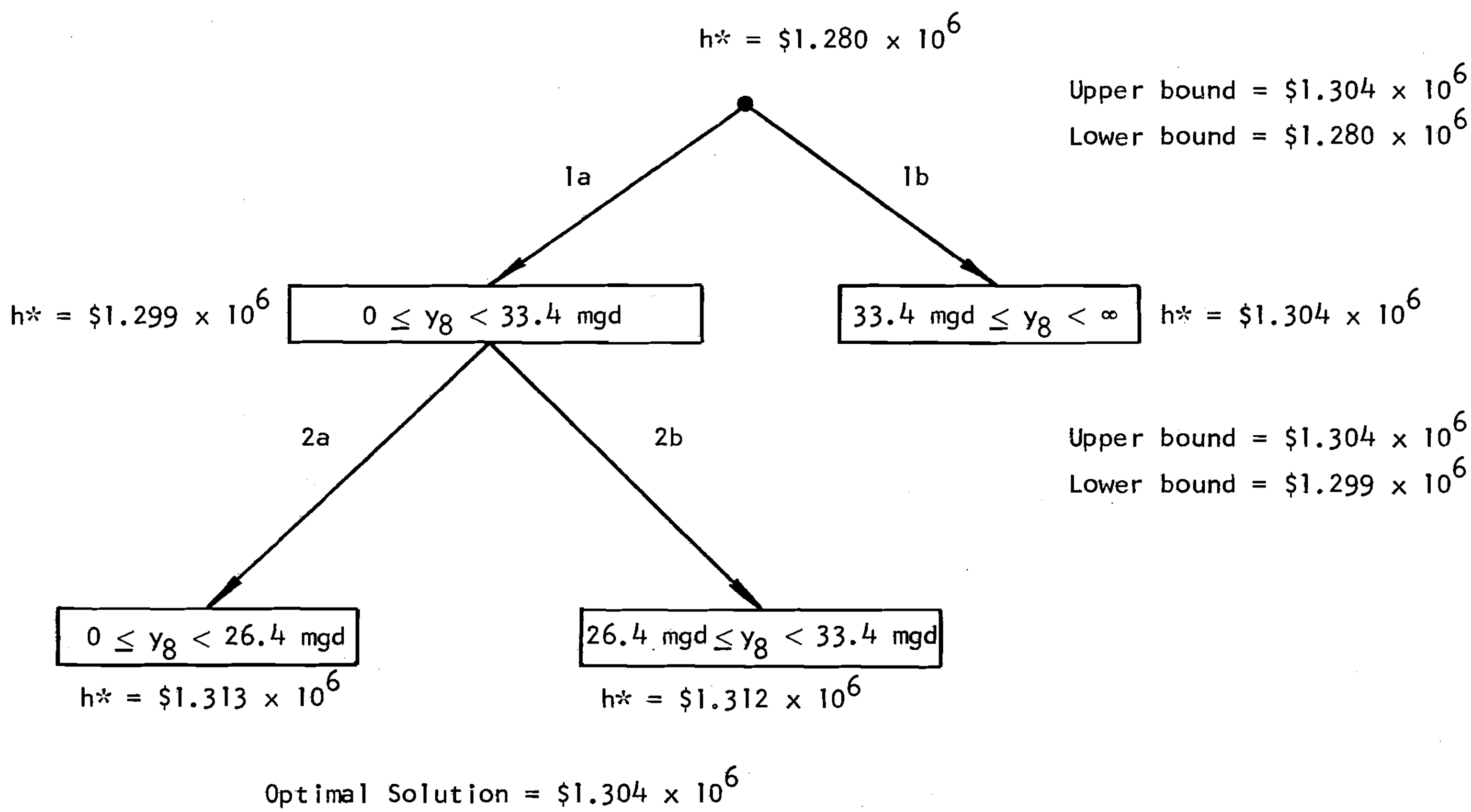


Figure 5.2 The Branch and Bound Gap Resolution Procedure Applied to the Delaware Estuary Problem

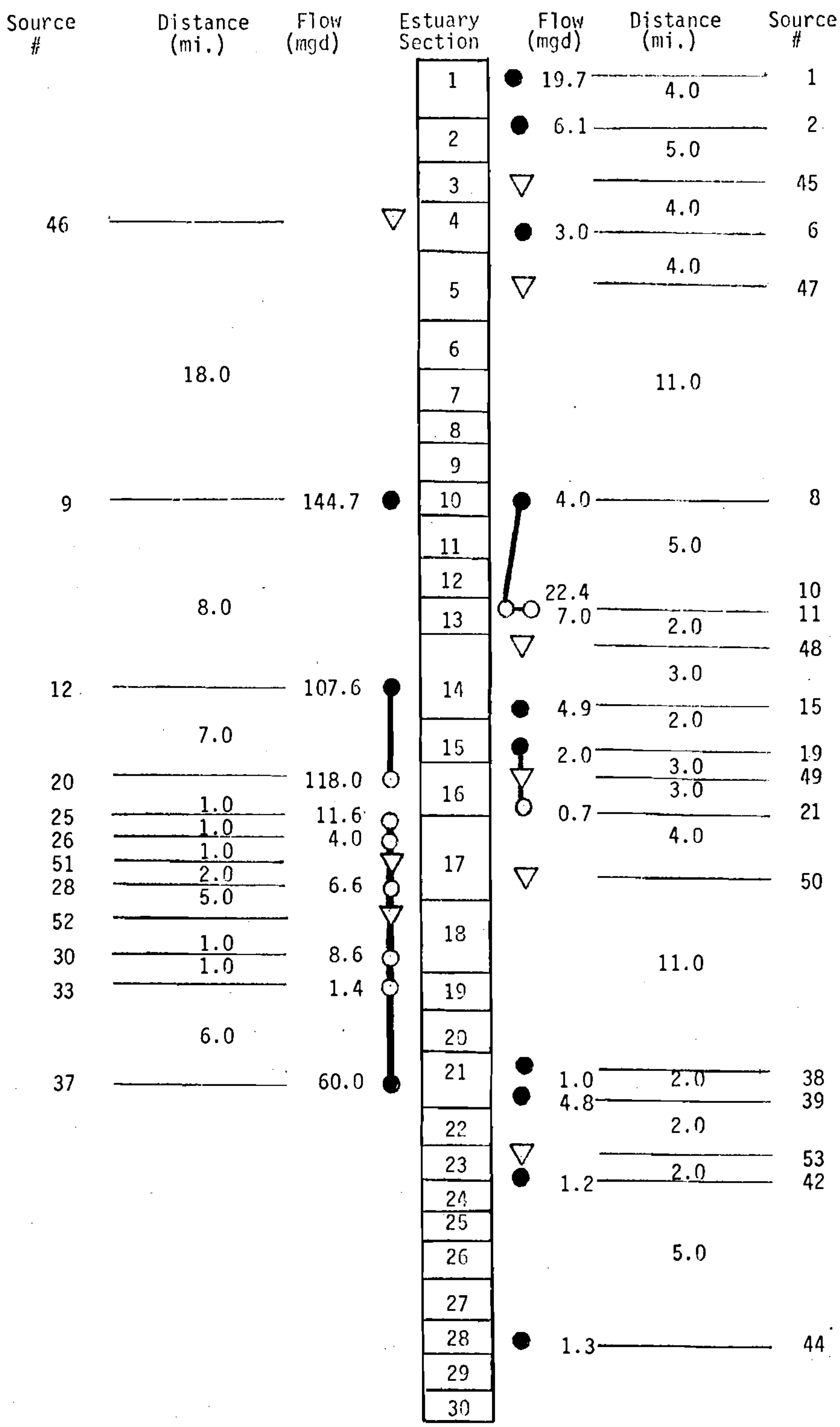


Figure 5.3 Optimal Regional Solution for the Delaware Estuary Problem

Table 5.3 Optimal Regional Solution for the Delaware Estuary

Regional Treatment

Location (Source Number)	Sources Served	% BOD Removal
8	8, 10, 11	75.6
12	12, 20	74.7
19	19, 21	64.6
37	25, 26, 28, 30 33, 37	65.0

Nonregional Treatment

Source	% BOD Removal
7	85.0
9	85.0
13	85.0
27	80.0

The remaining polluters do not increase treatment.

Total cost = $\$1.304 \times 10^6$/yr.

Of the total cost 42.5% is devoted to piping.

Of the previous regionalization studies the best solution was \$2.292 x 10^6 obtained in 10 minutes by Graves et al. (1970) on a machine not specified. Our solution represents a 43% savings (almost a million dollars) from this figure obtained in about 1/3 to 1/2 the time, depending on the stopping criteria used. To make this comparison fair it should be pointed out that the Graves et al. analysis allowed regionalization at only 9 locations while we used these 9 plus the other 22 source locations as well. Had the Graves et al. method been applied to this many potential regional locations it may very well have produced a lower cost solution since it allows for more piping arrangements than our approach. However, the resulting increase in problem size may make the Graves et al. algorithm computationally infeasible. In addition, our cost of \$1.304 x 10^6 for a least cost regionalized solution compares with \$4.1 x 10^6 for the least cost at-source treatment solution and \$10.331 x 10^6 for required secondary treatment at each source.

5.2 Performance under Varied Cost Functions

One of the factors crucial to the utility of the regionalization model is the frequency of occurrence of duality gaps and the difficulty in resolving them. In this section a series of regionalization problems will be solved to study this aspect. Instead of making all input parameters for these problems completely random we have chosen to keep these values close to what one may encounter in an actual application by

(i) using physical data from two real river systems

(ii) using a variety of piping and treatment cost functions as found in the literature.

The assumptions made in solving the problems are that existing treatment plants are ignored (i.e., only completely new treatment facilities can be built), the cost functions for at-source treatment and regional treatment facilities are the same, and no split flows are considered.

One of the river systems considered is actually the subsystem of the Delaware Estuary represented by reaches 11-20. The sources considered are the domestic sources shown in Fig. 5.1 which discharge to reaches 11 to 20. There is a total of 9 locations on one side of the river with 8 on the other side. Information on these sources can be found in Fig. 5.2 and Table 5.1. Two sets of DO goals were considered. The first corresponds to achieving the resulting DO in each reach if each individual source were to provide 85% BOD removal. The second set corresponds to 70% BOD removal by each source. The required DO improvements for each set are shown in Table 5.4. The DO constraints have the same form as those described in Section 5.1.

The second river system considered is the Willamette. Data were abstracted from Liebman (1965). There are 11 sources of waste flow and 4 additional potential locations for regional plants. Information on these is given in Table 5.5. The sources are located on both sides of the river and it is assumed that piping across the river is allowed. The river is divided into 14 reaches. The DO constraint equations are those described in Section 2.3. The coefficients were calculated using the equation in Appendix A. Since the river flow is much greater than the wasteflow discharged the calculations were made with fixed values of riverflow. Values of the river parameters used to obtain the coefficients are given in Table 5.6. Two sets of standards were considered. The first

Table 5.4 DO Goals for the Delaware Subsystem

	Required DO Improvement, mg/ℓ *	
Reach	1	2
11	0.9	0.63
12	1.4	1.0
13	2.1	1.5
14	2.6	1.8
15	2.9	2.0
16	3.1	2.1
17	2.6	1.8
18	2.0	1.3
19	1.4	1.0
20	1.0	0.65

* Set 1 corresponds to the improvement required to attain the resulting DO level if all individual sources were to provide 85 percent BOD removal. Set 2 is for 75 percent BOD removal.

Table 5.5 Source Data for the Willamette River

Source	Time of Flow from Origin, days	Flow, mgd	Raw BOD, lb/day	Present BOD Discharge, lb/day	BOD Decay Coeff., day^{-1}	Distance to Downstream Source, miles
1	0.23	4.83	10000	3300	0.31	6.3
2	0.58	31.3	74000	66600	0.33	15.
3	0.93	0.	0	0	0.35	15.
4	1.27	0.	0	0	0.35	13.
5	1.38	4.16	11000	11000	0.34	4.
6	1.69	12.9	13000	10400	0.34	12.
7	1.72	14.0	11000	9130	0.36	2.
8	2.26	8.4	17000	17000	0.36	16.
9	2.81	0.	0	0	0.35	16.
10	2.89	14.2	8000	8000	0.40	2.
11	3.55	36.8	125000	112500	0.40	17.
12	4.22	0.	0	0	0.35	16.
13	6.29	4.0	48000	48000	0.35	16.
14	7.34	0.33	6000	5400	0.35	8.
15	13.47	40.7	95000	95000	0.30	-

Table 5.6 River Data for the Willamette River

Reach (Control Pt.)	Time of Flow from Origin, days	Reaeration Coeff., $days^{-1}$	Present DO Deficit, mg/ℓ
1	0.25	1.02	0.78
2	1.00	0.72	0.71
3	1.50	0.72	0.77
4	2.00	0.60	0.89
5	2.50	0.60	1.00
6	3.00	0.60	1.08
7	4.00	0.60	1.62
8	5.00	0.10	2.75
9	7.00	0.10	4.09
10	8.00	0.10	4.51
11	9.00	0.10	4.66
12	13.00	0.10	3.95
13	14.00	0.10	4.10
14	15.00	0.10	4.41

Base river flow = 3600 mgd

had an allowable maximum DO deficit of 1 mg/ℓ at the ends of reaches 1 to 7 and 2 mg/ℓ for reaches 8 to 14. The second had the maximum allowable deficits for reaches 8 to 14 reduced to 1.5 mg/ℓ.

Two pairs of both piping and treatment plant cost functions were utilized. They are shown in Table 5.7. Note that the piping function PB introduces a pumping cost if flow is being piped against the gradient (upstream). The two treatment plant cost functions were derived by different authors from the same data presented by Frankel (1965). Note that TB includes fixed costs while TA does not. It is felt that this collection of cost functions represents a good sampling of functional types including smooth, continuous functions, fixed costs, and conditional costs.

For each river system and DO goal, 4 regionalization problems were solved representing all combinations of piping and treatment cost functions shown in Table 5.7. This gave a total of 16 problems. In maximizing the dual function a stop criterion of .0001% was used. At the maximum of the dual, the corresponding primal solution was said to be optimal if

$$\left| \frac{g_i}{\Delta c_i} \right| < .001 \quad \forall \quad i \in \{i : u_i > 0 \ , \quad \Delta c_i > 0\}$$

$$\frac{g_i}{\Delta c_i} < .001 \quad \forall \quad i \in \{i : u_i = 0 \ , \quad \Delta c_i > 0\}$$

$$g_i < .001 \quad \forall \quad i \in \{i : \Delta c_i = 0\}$$

and
$$\left| \frac{f^* - h^*}{h^*} \right| < .001$$

where g_i = required - actual DO improvement in each i

Δc_i = required DO improvement in reach i

u_i = dual variable associated with reach i

Table 5.7 Cost Functions

Piping Costs

PA: Cost \$/yr-mile = $1865\ Q^{.6}$

from Graves et al. (1970)

PB: For piping downstream (no pumping),

*Cost, \$/mile	$= 149653\ Q^{.53088}$	$Q \leq .5$
	$= 154697\ Q^{.5787}$	$.5 < Q \leq 2.5$
	$= 165346\ Q^{.50604}$	$Q > 2.5$

For piping upstream,

*Cost, \$/mile (pipe)	$= 92609\ Q^{.49544}$	$Q \leq .3$
	$= 98228\ Q^{.54427}$	$.3 < Q \leq 1.0$
	$= 98228\ Q^{.58505}$	$1.0 < Q \leq 5.0$
	$= 94100\ Q^{.61173}$	$Q > 5.0$
*Cost, \$ (pumping)	$=414387\ Q^{.75699}$	

from Smith (1971)

* Converted to \$/yr by dividing by a present value factor of 13.

Treatment Plant Costs

TA: Cost, \$/yr = $49.22\ Q^{.75}[8.0(r-.5)^3 + 1]$

from Graves et al. (1970)

TB: $\$/yr = 160.8 + 26.7\ Q + 640.7(r - .45)^2 + 255.7\ Q(r - .45)^2$

from Hass (1970)

where Q = flow handled, mgd

r = fractional removal of BOD

h^* = maximum value of dual function

f^* = corresponding value of primal.

If these conditions were not satisfied then a duality gap was assumed to exist and the branch and bound procedure of Section 3.5 was used to resolve the gap completely. In contrast to the previous section, the branching rule used here was to branch on that location with the largest size difference in the two alternate regional patterns which result after maximizing the dual.

The results of the 16 runs are summarized in Tables 5.8, 5.9, 5.10, 5.11, and 5.12. In the codes used to distinguish the runs the first letter refers to the river system (D for Delaware and W for Willamette), the second digit refers to the DO goal and the next four characters describe which of the cost functions of Table 5.7 was used. For instance, D2-PA-TB means that the Delaware River problem with the second set of DO goals was solved with piping cost function PA and treatment cost function TB.

Table 5.8 shows the results of the initial dual maximization of the 16 problems. Six were able to be solved at this point; the other 10 had duality gaps according to our criteria. However, by Everett's Theorem we can still obtain the optimal solution to a problem with modified DO goals from the results of any iteration in this dual maximization. Using that iteration which gave goals closest to the original, the number of such modified goals and the largest modification for each problem are given in Table 5.8. Note that in most cases the modified problem had DO goals only a few percent away from the original ones. Considering the imprecision involved in measuring DO such results may be entirely satisfactory.

Table 5.8 Results from a Single Dual Maximization

Run	Time sec	# Modified DO Goals	Maximum % Difference in Goal	Cost, $ x 10^6	Dual Value, $ x 10^6
D1-PA-TA	10.20	2	11.14	5.746	5.740
D1-PA-TB	6.52	1	2.067	7.94	7.911
D1-PB-TA	6.87	2	0.545	7.168	7.133
D1-PB-TB	7.45	0	0.	8.361	8.359
D2-PA-TA	11.12	3	130.3	4.920	4.587
D2-PA-TB	10.96	4	1.709	5.072	5.083
D2-PB-TA	10.02	0	0.	5.618	5.618
D2-PB-TB	8.29	0	0.	5.435	5.436
W1-PA-TA	13.94	2	5.522	3.607	3.601
W1-PA-TB	13.54	3	15.99	2.262	2.270
W1-PB-TA	18.55	0	0.	3.859	3.859
W1-PB-TB	11.70	1	0.092	2.510	2.519
W2-PA-TA	15.29	0	0.	3.688	3.688
W2-PA-TB	14.34	1	2.027	2.737	2.701
W2-PB-TA	26.40	0	0.	3.996	3.996
W2-PB-TB	13.82	1	0.311	2.884	2.897

Table 5.9 Feasible Solutions Found from a Single Dual Maximization

Run	Total Time (sec)	Upper Bound $ x 10^6	(Upper-Lower Bound) x 100 / Lower Bound	Optimum Found?
D1-PA-TA	15.89	5.897	2.74	Yes
D1-PA-TB	12.54	7.949	0.49	Yes
D1-PB-TA	12.09	7.134	0.000+	Yes
D1-PB-TB	7.45	8.361	0.0	Yes
D2-PA-TA	22.08	4.699	2.46	Yes
D2-PA-TB	21.56	5.097	0.27	Yes
D2-PB-TA	10.02	5.618	0.0	Yes
D2-PB-TB	8.29	5.435	0.0	Yes
W1-PA-TA	19.36	3.604	0.083	Yes
W1-PA-TB	21.79	2.290	0.85	Yes
W1-PB-TA	18.55	3.859	0.0	Yes
W1-PB-TB	19.57	2.523	0.16	Yes
W2-PA-TA	15.29	3.688	0.0	Yes
W2-PA-TB	22.99	2.714	0.49	No*
W2-PB-TA	26.40	3.996	0.0	Yes
W2-PB-TB	23.16	2.898	0.041	Yes

* 0.32% away from optimum

Table 5.10 Results of Complete Gap Resolution

Run	# Branchings	Optimum Cost, $ x 10^6	Avg. Time Dual (sec)	Avg. # iterations	# Degree of Treat. Probs.	Avg. Time (sec)	Total Time (sec)
D1-PA-TA	2	5.897	7.49	19	3	2.83	45.94
D1-PA-TB	2	7.949	6.52	23	4	2.50	42.60
D1-PB-TA	1	7.134	6.87	17	2	2.61	25.85
D1-PB-TB	0	8.361	7.45	19	0	0	7.45
D2-PA-TA	11	4.699	9.07	22	20	2.5	258.61
D2-PA-TB	1	5.097	9.38	29	3	3.54	38.76
D2-PB-TA	0	5.618	10.02	24	0	0	10.02
D2-PB-TB	0	5.435	8.29	23	0	0	8.29
W1-PA-TA	1	3.604	11.40	15	2	2.70	39.60
W1-PA-TB	3	2.290	14.83	24	4	3.52	117.89
W1-PB-TA	0	3.859	18.55	21	0	0	18.55
W1-PB-TB	1	2.523	13.89	18	2	3.94	49.55
W2-PA-TA	0	3.688	15.29	19	0	0	15.29
W2-PA-TB	2	2.705	14.01	22	3	4.49	83.52
W2-PB-TA	0	3.996	26.40	29	0	0	26.40
W2-PB-TB	1	2.898	19.64	26	2	4.67	68.26

Table 5.11 Solutions to the Delaware Subsystem Problems

Run	Treatment Plant Locations*	Sources Served*	% BOD Removal
D1-PA-TA	20	12,20,25, 26,28,20,33	89
	11	10, 11,15,19,21	78
D1-PA-TB	12	12	86
	20	20,25,26,28,30,33	84
	11	10,11,15,19,21	92
D1-PB-TA	12	12	85
	20	20,25	89
	28	26,28	80
	30	30,33	82
	11	10,11	85
	15	15	67
	19	19	69
	21	21	77
D1-PB-TB	12	12	82
	20	20	89
	26	25,26	84
	28	28	78
	33	30,33	96
	11	10,11	91
	19	15,19,21	66
D2-PA-TA	20	12,20,25 26,28,30,33	75
	11	10,11,15,19,21	68
D2-PA-TB	12	12	70
	20	20,25,26,28	71
	33	30,33	68
	11	10,11,15,19,21	76
D2-PB-TA	12	12	71
	20	20	73
	26	25,26	68
	28	28	66
	33	30,33	68
	11	10,11	71
	15	15	60
	19	19	50
	21	21	66
D2-PB-TB	12	12	68
	20	20	74
	26	25,26	71
	28	28	67
	33	30,33	72
	11	10,11	75
	19	15,19	56
	21	21	56

* See Figure 5.1 for locations of sources.

Table 5.12 Solutions to the Willamette Problems

Run	Treatment Plant Locations*	Sources Served*	% BOD Removal
W1-PA-TA	2	1,2	53
	7	5,6,7,8	53
	11	10,11	55
	15	13,14,15	55
W1-PA-TB	2	1,2	50
	6	5,6	50
	7	7,8	49
	11	10,11	51
	15	13,14,15	56
W1-PB-TA	1	1	67
	2	2	51
	5	5	57
	6	6	55
	8	7,8	57
	11	10,11	62
	13	13	68
	14	14	64
	15	15	50
W1-PB-TB	1	1	66
	2	2	51
	5	5	51
	6	6	48
	7	7	47
	8	8	52
	10	10	47
	11	11	62
	13	13	79
	15	14,15	45
W2-PA-TA	2	1,2	61
	7	5,6,7,8	60
	10	10,11	66
	15	13,14,15	66
W2-PA-TB	2	1,2	58
	6	5,6	55
	7	7,8	54
	10	10,11	70
	13	13	98
	15	14,15	50

Table 5.12 continued

Run	Treatment Plant Locations *	Sources Served*	% BOD Removal
W2-PB-TA	1	1	68
	2	2	65
	5	5	63
	6	6	60
	8	7,8	62
	11	10,11	72
	13	13	83
	14	14	77
	15	15	56
W2-PB-TB	1	1	67
	2	2	56
	5	5	55
	6	6	50
	7	7	49
	8	8	56
	10	10	49
	11	11	74
	13	13	98
	15	14,15	52

*Sources 3, 4, 9, and 12 are dummy sources, i.e. additional potential locations for regional plants

For obtaining the optimal solutions to the original problems, Table 5.9 shows the results of carrying out the initial step in the gap resolution procedure. That is, obtaining a best feasible solution by solving the minimum cost degree of treatment problems associated with the alternate regional configurations of the first dual maximization. Using the maximum value of the dual as a lower bound, these solutions were, in most cases, guaranteed to be within 1% of the optimum, the worst being 2.47% away. In fact, as it turned out after completing the entire gap resolution procedure, all but one of these solutions was indeed the true optimum. The results of the entire branch and bound procedure are summarized in Table 5.10. They show that of the 10 problems with gaps 8 were solved with 2 branchings or less (recall that for each branching two new dual maximization problems are created). For both the Delaware problems, which had 10 dual variables, and the Willamette problems with 14 dual variables, the average number of iterations needed by the column generation LP to maximize each dual problem was 22. All computations were made on an IBM 360/75.

From these results we see that duality gaps can occur frequently. However, when they do occur a solution to a modified problem not very different from the original is immediately at hand and a feasible, upper bound solution can be easily found which is quite close to and has a high probability of being the global minimum. It is this fact which probably makes the branch-and-bound procedure successful in resolving duality gaps. It was observed that in most cases the alternate regionalization patterns associated with duality gaps were very similar, the differences in plant sizes occurring at a pair of locations not very far apart. This might give support to the suggestion of Section 3.6 to speed solution of the dual problem by using discrete differential dynamic programming. A final obvious observation is that as shown in the solutions displayed in Tables 5.11 and 5.12 the use of different cost functions can result in widely varying regional configurations.

CHAPTER 6. SUMMARY AND CONCLUSIONS

The goal of this research has been the development of a practical water quality management decision model for the minimum cost design of a regionalized system of wastewater treatment facilities along a river, subject to water quality criteria. To allow for efficient solutions while maintaining a high degree of accuracy the following restrictions were made in the formulation of the model:

1. Only steady state dissolved oxygen and its interaction with carbonaceous BOD discharged from continuous point sources was considered.

2. All waste sources which could participate in a regional system had to be arranged in a linear configuration along a single river.

3. A regional treatment plant could only serve those sources located sequentially upstream and/or downstream of itself (i.e., no bypassing of sources).

4. A regional plant at a given location had to treat at least the flow piped into that location from other sources.

5. The changes in various river parameters due to changes in river flow as a result of regionalization could be ignored.

Based on these restrictions the formulated mathematical programming model displayed a serial structure which suggested the use of dynamic programming to solve for the optimal decisions. However, the large number of state variables made this approach impractical. Instead, Lagrange multipliers were introduced and a Lagrangian function formed. Using concepts from duality theory for nonlinear programming it was shown that a dual function could be defined as the minimum of the Lagrangian for given values of the multipliers (or dual variables). This minimum could be found by single state discrete dynamic programming. The dual problem then became one of finding those multipliers which maximized the dual function. If the maximum of the dual equaled its corresponding primal value and all water quality constraints were satisfied, then the primal was solved. If not, then a duality gap was said to exist. A branch and bound procedure was presented which could resolve such gaps to obtain the optimal solution.

Some advantages of formulating the model in this manner are:

1. Every source is considered as a potential regional plant and adding additional plant sites only increases the size of the problem linearly; thus a large number of regional plant locations can be considered and can increase chances of reducing costs. For example, in analyzing the Delaware Estuary the model required 155 variables for 31 potential regional facility locations while a previous analysis (Graves et al. (1970)) required over 2,000 variables for only 9 regional facility locations (however, bypass piping of effluents to different river sections was also considered).

2. Flows shipped over common piping routes are automatically combined in a single pipe.

3. Cost functions can be of any form with respect to waste flow handled and a global minimum is still found.

The model was applied to data from a previous regionalization study of the Delaware Estuary and its performance under a variety of cost functions with two smaller river systems was investigated. The performance of the model can be summarized as follows:

1. A 43% less costly solution to the Delaware problem was found in one third the time as compared with Graves et al. (1970).

2. Computer storage requirements for the coded version of the model are moderate. The Delaware problem was solved using 110 kilobytes of storage.

3. The model is capable of generating feasible solutions which are either optimal or within a few percent of optimality very quickly from the results of a single dual maximization. For instance, a solution to the Delaware problem guaranteed to be within 2% of the optimum (and which actually was the optimum) was obtained in little over a minute.

4. Duality gaps can occur frequently (in 11 out of 17 problems solved). However, optimal solutions are immediately available to problems with modified DO goals. In most of the problems examined these goals were within a few percent of the originals. To resolve such gaps and obtain optimal solutions for the original goals, the branch and bound procedure is an effective method, although no conclusions can be made concerning the best branching rule.

This basic version of the model was extended to include branched river systems and bypassing of wastes to other discharge points. However, the computational feasibility of these extensions remains unknown. Keeping the assumption of linear source configuration it was shown how two other regional treatment facility problems could be solved. The first of these eliminates degree of waste treatment as a decision variable while the second requires degree of treatment to be uniform at all facilities. In Appendix A several approaches are given for extending the model to consider the effect of variable river flow due to regionalization on the river parameters which establish the BOD-DO relations. Finally it should be noted that the basic approach to formulating and solving the model under the linear source configuration and other regionalization restrictions could be applied to more general source and allowable regionalization arrangements provided that efficient methods are available for solving the resulting facility location portion of the problem.

The model developed in this research can be effectively utilized in water quality management studies providing its role in such studies is properly understood. To illustrate this point we can take as an example the preparation of a basin plan as specified by Environmental Protection Agency Guidelines (1971). One of the purposes of such a plan is "to maximize the cost effectiveness of investments in pollution abatement and prevention actions required to achieve national water quality objectives". The steps involved in producing the basin plan are shown in Fig. 6.1.

Economic optimization models, such as the regionalization model developed in this work, can aid the planner in synthesizing the most cost effective plan from the multitude of alternatives available to him. However, the amount and quality of the information he obtains will depend on his understanding of the model's assumptions and limitations and the quality of the input data which is used. With reference to the regionalization model it should be noted that there are many questions which must be answered in the basin plan not considered in the model. These include

1. control measures for other pollutants besides carbonaceous BOD

2. control measures for distributed sources such as land runoff and nonsteady sources such as storm water overflows

3. utilization of other control measures such as flow regulation, temporary waste storage, and in-stream aeration

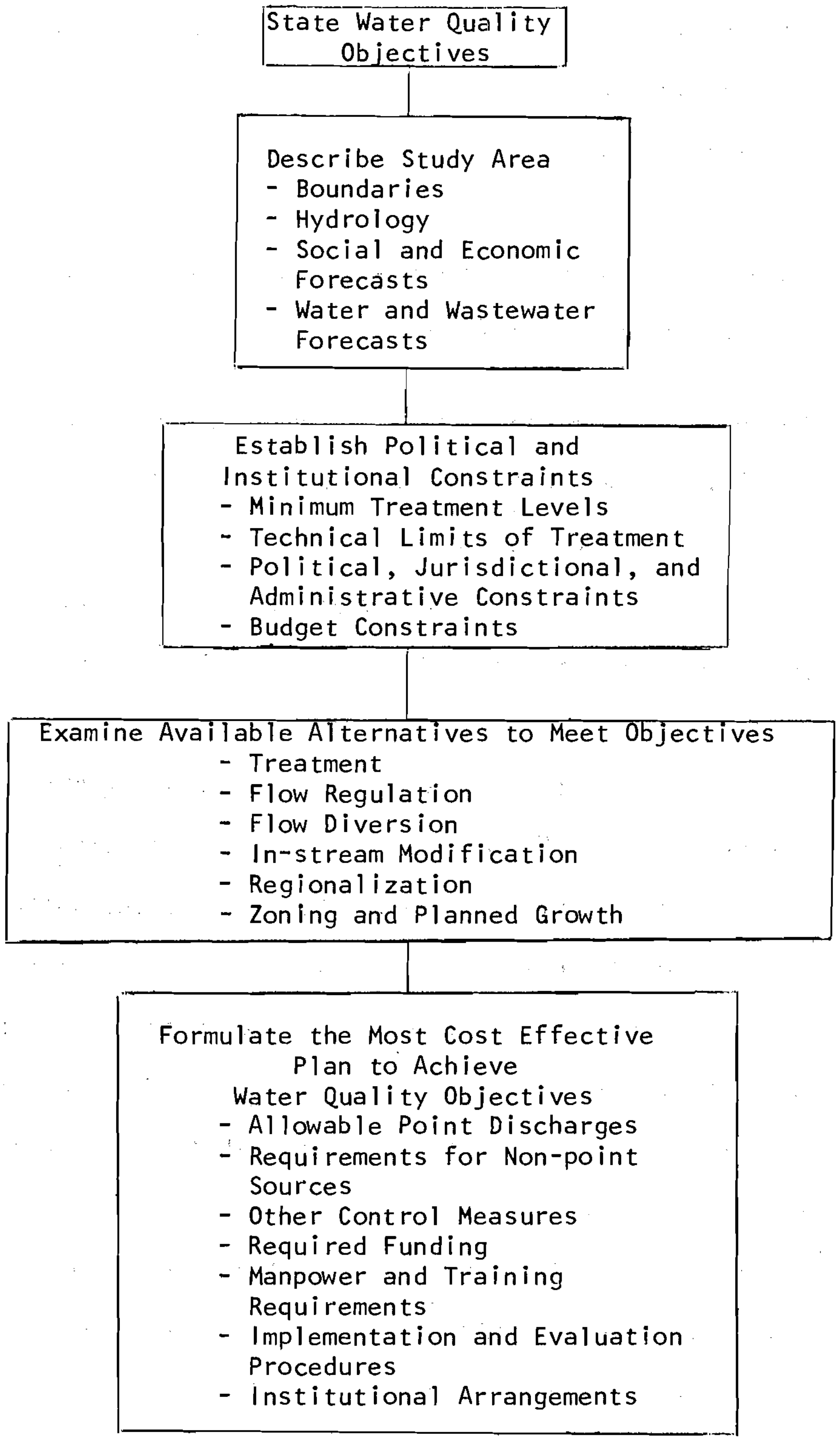


Figure 6.1 Steps in Preparing a Water Quality Management Plan

4. effects of variability in treatment plant performance on water quality for regionalized systems

5. timing the capacity expansion of facilities to accommodate the growth in waste loads

6. resolution of equity problems which arise when one discharger must provide more treatment than another whom he considers his equal.

These omissions obviously preclude the use of the model as the sole basis for making policy decisions. Instead the model, and others like it, should be viewed as a screening device which leads the planner in the direction of the "best" solution. The benefits in terms of the information generated by using the model can far outweigh the costs in man hours of running the model.

A conceptual framework for utilizing optimization models in the basin planning study is shown in Fig. 6.2. A first step is the selection of the appropriate models to be used. This is done on the basis of the political, social and institutional constraints imposed and the kinds of alternatives available. Next the necessary input data for these models are acquired and all model parameters are calculated. On first running of the model this may be a rather crude effort to be strengthened later in the areas which the model shows to be most sensitive.

The models are then run under a variety of input conditions and assumptions. For instance, with respect to the regionalization model, the following types of analyses could be made

1. several runs with increasing waste loadings, to observe the change in regionalization patterns over time as waste production grows in the basin

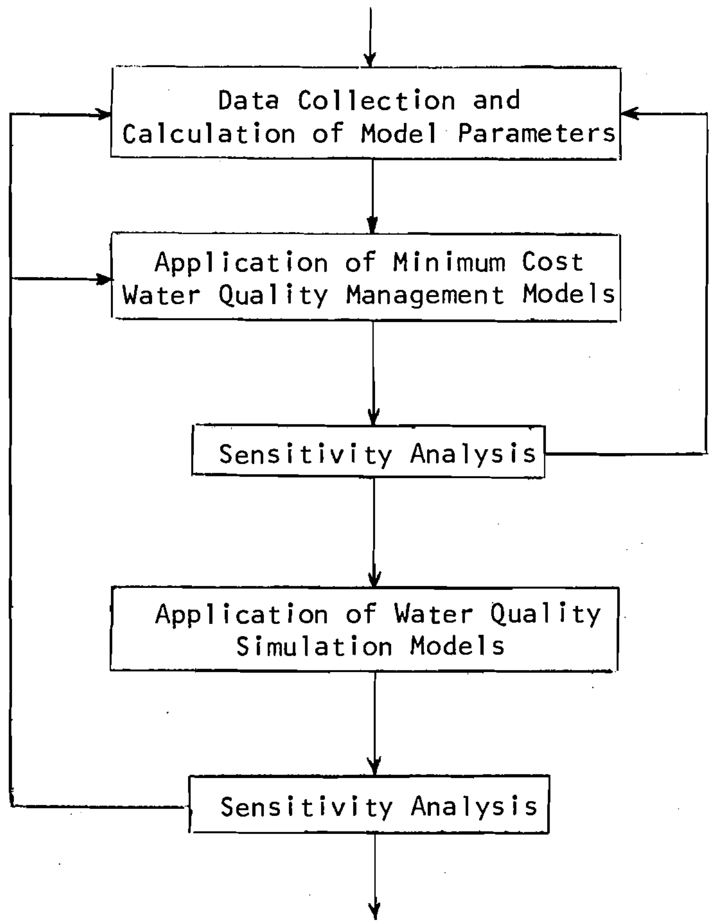


Figure 6.2 Utilization of Optimization Models in Basin Planning

2. selected dischargers already providing a high level of treatment can be excluded from those allowed to regionalize since additional treatment would be of little value for such sources

3. required treatment levels for other pollutants besides BOD can be set in advance thus affecting the magnitudes of the cost functions

4. flow regulation strategies can be considered by fixing the level of flow augmentation in advance and then solving the regionalization model.

In addition the sensitivity of the results to changes in various model parameters can be established. This may necessitate a return to the data collection and parameter estimation phase of the study to obtain more accurate values of these parameters.

The results of the optimization model runs (regional facility arrangements and degree of BOD removal) can then form the input for a more detailed analysis. This analysis would include decisions on those items not explicitly considered in the optimization model, some of which are listed above. The use of simulation models would be helpful at this stage. These models would predict the complete environmental impact of a detailed pollution control strategy. Using the information obtained from the optimization models as a starting point such a strategy could be synthesized and its environmental performance monitored by the simulation model. Certain areas of the strategy may require alteration so that environmental objectives are met. Experimenting with different combinations of alternatives which lead to lower costs would also be done. The output of the simulation studies may generate information on how the input to the optimization models should have been specified. Thus a return to the optimization model could be made to generate another set of inputs to the simulation model. For example, if the regionalization model suggested a highly regionalized system which the simulation model showed would continuously violate water quality standards because of the variability in treatment plant performance then the regionalization model could be run again with the sizes of treatment plants restricted at certain locations.

Another such feedback loop could be created between the simulation studies and the data collection and parameter estimation activities as it became clear that better accuracy would be required. Guiding the synthesis of the basin plan at every stage are the evaluation methods which convert all decisions into cost figures and consider other performance indices such as ease of implementation and equity. The end result of this process is the formulation of a set of water quality management decisions which will obtain the desired water quality objectives in the basin in a highly cost effective manner.

In our example we have assumed that the water quality objectives were established separately from the basin plan. Yet to specify a level of water quality which in some sense provides the greatest measure of social utility requires that the benefits derived from a particular level be compared with the costs of achieving it. Thus we could add to our planning process displayed in Fig. 6.1 the following items:

1. an additional block which evaluates the benefits received from the stated water quality objectives and the formulated basin plan,

2. a feedback loop to the statement of objectives on which the objectives would be reevaluated and suitably modified in relation to benefits, costs, and other intangible social and political factors.

Within this expanded framework our water quality goals can be established on a rational basis.

APPENDIX A. EXTENSIONS OF THE WATER QUALITY CONSTRAINTS

In what follows we show how constraints can be written which relate dissolved oxygen goals at points in a river to wastewater flow and BOD discharges. The river is divided into N reaches. For convenience we place each waste discharger or tributary flow at the beginning of each reach and we are given a known dissolved oxygen (DO) level which must be attained at the end of each reach. Actually the dischargers and downstream DO control points could be placed anywhere in the reach. River parameters, with the exception of BOD, DO, and flow, are assumed constant within each reach. The following notation is used.

$K11_j$ = BOD deoxygenation rate coefficient in reach j, day^{-1}

$K12_j$ = BOD removal rate coefficient in reach j, day^{-1}

$K2_j$ = reaeration rate coefficient in reach j, day^{-1}

t_j = time of flow through reach j, days

b_j = stream BOD at end of reach j, lb/day

c_j = stream DO concentration at end of reach j, lb/mil gal

$(bb)_j$ = stream BOD at beginning of reach j, just downstream of a waste discharger, lb/day

$(cb)_j$ = stream DO concentration at beginning of reach j, just downstream of waste discharger, lb/mil gal

f_j = flow in reach j, mgd

y_j = flow discharged to stream by polluter (or tributary) j at beginning of reach j, mgd

z_j = BOD discharged to stream by polluter (or tributary) at beginning of reach j, lb/day

$(cw)_j$ = DO concentration in flow of jth discharger, lb/mil gal

$(cs)_j$ = saturation DO concentration in reach j, lb/ mil gal.

We assume that temperature effects are negligible so that values of $K11_j$, $K12_j$, $K2_j$ and $(cs)_j$ are known constants for each reach. As shown later, $K2_j$ and t_j are really functions of $(y_1, y_2, \ldots, y_j)$ but for the moment we assume they are constants. Also known in advance are values of y_j and z_j for tributary inflows, $(cw)_j$ for each waste discharger, and initial conditions b_o, c_o, and f_o for a point just above the first discharger. These quantities can be computed based on the existing flow and BOD discharges (the vectors q and s, respectively).

Assuming complete mixing across the cross section and no longitudinal dispersion the Camp-Dobbins form of the Streeter-Phelps equations are

$$b_j = \alpha_j (bb)_j \tag{A.1}$$

$$c_j = (cs)_j (1 - \beta_j) + \beta_j (cb)_j - \frac{\gamma_j b_j}{f_j} \tag{A.2}$$

where

$$\alpha_j = \exp(-K11_j t_j)$$

$$\beta_j = \exp(-K2_j t_j)$$

$$\gamma_j = \frac{K11_j}{K2_j - K12_j} (\alpha_j - \beta_j)$$

From mass balances at the beginning of each reach we obtain

$$(bb)_j = b_{j-1} + z_j \tag{A.3}$$

$$f_j (cb)_j = f_{j-1} c_{j-1} + y_j (cw)_j \tag{A.4}$$

$$f_j = f_{j-1} + y_j \tag{A.5}$$

Substituting (A.3) into (A.1) gives

$$b_j = \alpha_j b_{j-1} + \alpha_j z_j$$

Denoting the N x N matrices A_{11} and A_{12} by

$$A_{11} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -\alpha_2 & 1 & 0 & \cdots & 0 & 0 \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \\ 0 & 0 & 0 & \cdots & -\alpha_N & 1 \end{bmatrix}$$

$$A_{12} = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ 0 & 0 & \cdots & \alpha_N \end{bmatrix}$$

and the N x 1 vector a_1 by

$$a_1 = [\alpha_1 b_o \ , 0, \ldots, 0]^T$$

results in the following matrix equation

$$A_{11} b = A_{12} z + a_1 \qquad (A.6)$$

where $\qquad b = [b_1, b_2, \ldots, b_N]^T$

and $\qquad z = [z_1, z_2, \ldots, z_N]^T.$

Solving for b gives

$$b = A_{11}^{-1} A_{12} z + A_{11}^{-1} a_1 \qquad (A.7)$$

Substituting (A.4) into (A.2) gives

$$f_j c_j = f_j (cs)_j - \beta_j f_j (cs)_j + \beta_j f_{j-1}\, c_{j-1} + \beta_j y_j (cw)_j - \gamma_j b_j$$

Forming the following N x N matrices

$$A_{21} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -\beta_2 & 1 & 0 & \cdots & 0 & 0 \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \\ 0 & 0 & 0 & \cdots & -\beta_N & 1 \end{bmatrix}$$

$$A_{22} = \begin{bmatrix} \beta_1 (cw)_1 & 0 & \cdots & 0 \\ 0 & \beta_2 (cw)_2 & \cdots & 0 \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ 0 & 0 & \cdots & \beta_N (cw)_N \end{bmatrix}$$

$$A_{23} = \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ 0 & 0 & \cdots & \gamma_N \end{bmatrix}$$

$$A_{24} = \begin{bmatrix} (cs)_1(1-\beta_1) & 0 & \cdots & 0 \\ 0 & (cs)_2(1-\beta_2) & \cdots & 0 \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ 0 & 0 & \cdots & (cs)_N(1-\beta_N) \end{bmatrix}$$

and the N x 1 vector a_2,

$$a_2 = [\beta_1 f_o c_o, 0, \ldots, 0]^T$$

results in the following matrix equation

$$A_{21} C f = A_{22} y - A_{23} b + A_{24} f + a_2 \tag{A.8}$$

where

$$f = [f_1, f_2, \ldots, f_N]^T$$

$$y = [y_1, y_2, \ldots, y_N]^T$$

$$C = \begin{bmatrix} c_1 & 0 & \ldots & 0 \\ 0 & c_2 & \ldots & 0 \\ \vdots & & & \\ 0 & 0 & \ldots & c_N \end{bmatrix}$$

Substituting (A.7) for b gives

$$A_{21} C f = A_{22} y - A_{23} A_{11}^{-1} A_{12} z$$

$$- A_{23} A_{11}^{-1} a_1 + A_{24} f + a_2 \tag{A.9}$$

From (A.5), f can be written as

$$f = A_{31} y + a_3 \tag{A.10}$$

where

$$A_{31} = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & & & & & \\ 1 & 1 & 1 & \ldots & 1 & 1 \end{bmatrix}_{N \times N}$$

$$a_3 = [f_o, f_o, \ldots, f_o]^T_{1 \times N}$$

Substituting (A.10) into (A.9) gives

$$\begin{aligned} A_{21}\, C\, A_{31}\, y + A_{21}\, C\, a_3 = {} & A_{22}\, y - A_{23}\, y - A_{23}\, A_{11}^{-1}\, A_{12}\, z \\ & - A_{23}\, A_{11}^{-1}\, a_1 + a_2 \\ & + A_{24}\, A_{31}\, y + A_{24}\, a_3 \end{aligned}$$

It follows that if the c_j's are set at some standard which must be attained then the corresponding inequality constraint set becomes

$$\begin{aligned} A_{21}\, C\, A_{31}\, y + A_{21}\, C\, a_3 \geq {} & A_{22}\, y - A_{23}\, A_{11}^{-1}\, A_{12}\, z \\ & - A_{23}\, A_{11}^{-1}\, a_1 + a_2 \\ & + A_{24}\, A_{31}\, y + A_{24}\, a_3 \end{aligned}$$

which can be rewritten as

$$A_1\, y + A_2\, z + a \leq 0 \qquad \text{(A.11)}$$

where

$$A_1 = A_{22} + A_{24}\, A_{31} - A_{21}\, C\, A_{31}$$

$$A_2 = -A_{23}\, A_{11}^{-1}$$

$$a = A_{23}\, A_{11}^{-1}\, a_1 - a_2 - A_{24}\, a_3 \; .$$

Thus we have expressed DO constraints in terms of flow and BOD discharges. Note that the constraints of (A.11) are linear and separable in the y_j's and z_j's. If we replace the original DO constraints of the regionalization problem (Eqs. (3.9)) with (A.11) the resulting dual function is

$$h(u) = \min_{y,z} \sum_{j=1}^{N} [P_j + T_j + (\sum_{i=1}^{N} (a_{ij})_1 u_i) y_j + (\sum_{i=1}^{N} (a_{ij})_2 u_i) z_j] + \sum_{i=1}^{N} a_i u_i$$

s.t. (3.10) - (3.15).

For a given u this can be evaluated by the dynamic programming recursion discussed in Section 3.2 where now the term $(\sum_{i=1}^{N} (a_{ij})_1 u_i) y_j$ is added to the return function for each stage. Thus this extension of the DO constraints poses no problem for the solution algorithm.

In the above analysis we assumed that $K2_j$ and t_j were constant. However, we know that they are really functions of streamflow, f_j, which can vary due to regionalization, if streamflow is small and wasteflow is large. These parameters can be related to streamflow as follows.

River depth and velocity can be empirically related to flow by (Leopold and Maddock (1953))

$$H = \theta f^{\phi}$$

$$V = \theta' f^{\phi'}$$

where H = depth

V = velocity

$\theta, \theta', \phi, \phi'$ = constants.

For constant velocity throughout a reach, the time of flow in reach j is

$$t_j = (\ell_j/\theta'_j) f_j^{-\phi'_j} \qquad (A.12)$$

where ℓ_j = length of reach j.

Tsivoglou and Wallace (1972), in their review of prediction equations for K2, note that most take the form

$$K2 = \eta \frac{V^{\phi''}}{H^{\phi'''}}$$

where η, ϕ'', and ϕ''' are constants.

Substituting for V and H gives

$$K2_j = \frac{\eta_j \, \theta_j' \, f_j^{(\phi' \phi'' - \phi \phi''')}}{\theta_j} \qquad (A.13)$$

Now the quantities t_j and $K2_j$ have been expressed as nonlinear functions of the streamflow, f_j, which is itself a function of the waste flow discharges,

$$f_j = f_o + \sum_{i=1}^{j} y_i .$$

As a result, the entries in the matrices A_1, A_2, and a of the DO constraints (A.11) are nonlinear functions of the vector of waste flow discharges, y. Since these constraints are now no longer separable in y the dual function of the regionalization problem cannot be evaluated by dynamic programming. In fact its evaluation by any other technique would be such an arduous task as to make the dual solution approach impractical.

One way to save the use of the dual method would be to linearize (A.11) around the value of the existing discharges, (q, s). Using a first order Taylor series expansion of (A.11) and recognizing that A_{11}^{-1} can be expressed as

$$A_{11}^{-1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ \alpha_2 & 1 & 0 & \cdots & 0 & 0 \\ \alpha_2\,\alpha_3 & \alpha_3 & 1 & \cdots & 0 & 0 \\ \vdots & & & & & \\ \alpha_2\,\alpha_3\cdots\alpha_N & \alpha_3\,\alpha_4\cdots\alpha_N & \alpha_4\cdots\alpha_N & \cdots & \alpha_N & 1 \end{bmatrix}$$

gives a linearized approximation of the DO constraints as

$$B_o + B_1(q - y) + B_2(s - z) \leq 0 \qquad \text{(A.14)}$$

where

$$B_o = [A_1\ y + A_2\ z + a\ ]_{\substack{y=q\\z=s}}$$

$$B_1 = [J_y(A_1\ y) + J_y(a)]_{\substack{y=q\\z=s}}$$

$$B_2 = A_2\ \Big|_{\substack{y=q\\z=s}}$$

$J_y(\cdot)$ = N x N Jacobian with respect to y.

The resulting linearized DO constraints are once again separable in y and z and so there is no problem in using the dual method. However, it is not clear how much error is introduced through the linearization. If the nonlinearities in A_1,A_2, and a are large and the optimal y produces a streamflow very different from the original base flow then this approach might be meaningless.

A second approach is to work directly with the stagewise equations (A.1) and (A.2) instead of combining all previous stages. The resulting DO constraints would be

$$c_i \geq c_{(STD)i} \tag{A.15}$$

$$\frac{f_i\ c_i}{\beta_i} = \frac{f_i (cs)_i (1-\beta_i)}{\beta_i} + y_i (cw)_i + f_{i-1}\ c_{i-1} - \frac{\gamma_i\ b_i}{\beta_i} \tag{A.16}$$

$$\frac{b_i}{\alpha_i} = b_{i-1} + z_i \tag{A.17}$$

$$f_i = f_o + \sum_{j=1}^{i} y_j = f_o + \hat{y}_i \tag{A.18}$$

where $c_{(STD)i}$ = DO standard for reach i; α_i, β_i, and γ_i are nonlinear functions of f_i; and f_o, b_o, and c_o are known. Constraints (A.16) and (A.17) introduce 2N coupled equality constraints into the regionalization problem as opposed to just N coupled inequality constraints when the above equations are written in their combined form. If the model is dualized with respect to these constraints then 2N dual variables are required. Let $u_1,\ldots,u_N$ be such variables corresponding to (A.16), and $v_1,\ldots,v_N$ correspond to (A.17).

The dual function for this new formulation becomes

$$\begin{aligned} h(u,v) = \min_{x,y,b,c} \sum_{i=1}^{N} [P_i + T_i - v_i\ z_i + (\frac{u_i\ \gamma_i}{\beta_i} + \frac{v_i}{\alpha_i} - v_{i+1})\ b_i \\ + (\frac{u_i\ f_i}{\beta_i} - u_{i+1}\ f_i)\ c_i - \frac{u_i\ f_i (cs)_i\ (1-\beta_i)}{\beta_i} \\ + u_i\ y_i (cw)_i] - V_1\ b_o - u_i\ f_o\ c_o \end{aligned}$$

$$\begin{aligned} \text{s.t.}\quad & c_{(STD)i} \leq c_i \leq (cs)_i \\ & b_{(min)i} \leq b_i \leq b_{(max)i} \\ & f_i = f_o + \hat{y}_i \\ & (3.10) - (3.15) \end{aligned}$$

where $\left.\begin{array}{l} b_{(min)i} \\ b_{(max)i} \end{array}\right\}$ = reasonable bounds on stream BOD in reach i to insure that a minimum can be attained

and $\left.\begin{array}{l} u_{N+1} \\ v_{N+1} \end{array}\right\} = 0 .$

For a given (u, v) it can be evaluated by the discrete dynamic programming recursion given in Section 3.2 with the return function given by

$$R(\hat{y}_i, y_i) = \min_{z_i, b_i, c_i} \Big[P_i(yp_i) + T_i(y_i, z_i, w_i) - v_i z_i + \Big(\frac{u_i \gamma_i}{\beta_i} + \frac{v_i}{\alpha_i} - v_{i+1}\Big) b_i + \Big(\frac{u_i f_i}{\beta_i} - u_{i+1} f_i\Big) c_i - u_i f_i (cs)_i (1-\beta_i) + u_i y_i (cw)_i \Big]$$

s.t.

$$c_{(STD)i} \leq c_i \leq (cs)_i$$

$$b_{(min)i} \leq b_i \leq b_{(max)i}$$

$$f_i = f_o + \hat{y}_i$$

$$yp_i = \sum_{j=1}^{i} q_j - \hat{y}_i$$

$$w_i = W(\hat{y}_i) - W(\hat{y}_i - y_i)$$

$$L_i \leq 1 - z_i / w_i \leq u_i$$

$$z_i \geq 0$$

and α_i, β_i, and γ_i are nonlinear functions of f_i.

Because of the separability in z_i, b_i, and c_i the above can be rewritten as

$$R(\hat{y}_i, y_i) = \min_{w_i(1-U_i)\leq z_i \leq w_i(1-L_i)} [T_i(y_i, z_i, w_i) - v_i\, z_i]$$

$$+ \min_{b_{(min)i} \leq b_i \leq b_{(max)i}} \left[\frac{u_i\, \gamma_i}{\beta_i} + \frac{v_i}{\alpha_i} - v_{i+1} \right] b_i$$

$$+ \min_{c_{(STD)i} \leq c_i \leq (cs)_i} \left[\frac{u_i\, f_i}{\beta_i} - u_{i+1}\, f_i \right] c_i$$

where w_i and f_i are known from $\hat{y}_i$ and y_i.

The dual problem is now

$$\max\ h(u,v)$$

$$u,v \text{ unrestricted.}$$

The dual variables (u,v) are unrestricted in sign since all the constraints which are dualized are equalities. The column generation linear program for solving the dual now becomes

$$\text{Min} \quad \sum_{k=1}^{K} \lambda_k\, F^k$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \lambda_k\, g_i^{\,k} = 0 \qquad i = 1,2,\ldots,N,N+1,\ldots,2N$$

$$\sum_{k=1}^{K} \lambda_k = 1$$

$$\lambda_k \geq 0 \qquad k = 1,\ldots,K$$

where $$F^k = \sum_{i=1}^{N} P_i(yp^k) + T_i(y_i^{\,k}, z_i^{\,k}, w_i^{\,k})$$

$$g_i^k = \frac{f_i^k c_i^k}{\beta_i^k} - f_i^k (cs)_i \; (1 - \beta_i^k) + y_i^i (cw)_i$$

$$+ f_{i-1}^k c_{i-1}^k - \frac{\gamma_i^k b_i^k}{\beta_i^k} \qquad \text{for } i = 1,\ldots,N$$

$$g_i^k = \frac{b_j^k}{\alpha_j^k} - b_{j-1}^k - z_j^k \qquad \text{for } j = i-N, \quad i = N+1,\ldots,2N.$$

The dual variables $u_1,\ldots,u_N$ correspond to constraints 1 to N while the variables $v_1,\ldots,v_N$ correspond to constraints N+1 to 2N. Another procedure which appears to work well on equality constrained problems (as reported by Greenberg and Robbins (1972)) is the Noose Method of Furman and Weinstein (1970). The gap test and resolution procedure remain the same. However, using Everett's Theorem to find solutions to problems with modified right hand sides is meaningless since in this formulation, due to mass balances, the right hand sides must always be zero.

The advantage of this formulation is that we are now able to take into account all of the effects of varying streamflow due to regionalization on the stream dissolved oxygen. In addition, the potential exists for introducing the effect of temperature. Thus a truly comprehensive regional water quality model applicable to rivers with all ranges of stream flow can be developed. The disadvantage is the doubling in the size of the dual problem with the resulting increase in computational time. The technique of relaxation, discussed in Section 3.6, used to reduce the size of the dual problem is not applicable here since all constraints are equalities. For large systems it is not clear whether the benefits to be gained by using the more accurate approach outweigh the computational costs.

APPENDIX B. COMPUTER PROGRAM

PORST is a computer program designed to maximize the dual of the regionalization model of Chapters 2 and 3 (PORST1) and solve the minimum cost degree of treatment problem for a fixed regional facility configuration (PORST2). It is written in ASA Standard FORTRAN IV for implementation on an IBM 360/75 computer. With reference to the complete solution algorithm as described in Section 3.5, PORST is used whenever the dual to a regionalization problem is maximized or a least cost degree of treatment problem for fixed regional configuration must be solved. The remaining steps of the algorithm are carried out manually by the analyst.

If no duality gap occurs, PORST (PORST1) need be run only a single time. If a gap does occur, the optimal solution to a regionalization problem with different DO goals is readily at hand. If the gap is to be resolved by branch and bound, the alternate regional configurations occurring at the maximum of the dual should be available by inspection of the output of PORST1. Then PORST2 is used to find the minimum cost degree of treatment solutions for these configurations. Finally, PORST1 is run again as the branching process of the algorithm unfolds.

This version of PORST operates under the following restrictions:

(1) The waste sources are divided into two groups: those which can regionalize and those which can't.

(2) The waste sources which can regionalize must lie in at most two distinct linear segments along the river.

(3) As implied by the regionalization restrictions on which the model was formulated, a regional facility can only serve a consecutive sequence of adjacent sources. Only the sources on the ends of such a

sequence can split their waste flow between at-source and regional treatment (or regional treatment at two different locations). The user must specify the fractions into which the source flow can be split.

(4) The same cost function for regional treatment applies at all locations. The cost function for at-source treatment can vary at each location by three parameters, one of which specifies the maximum amount of BOD which can be removed. One suggested functional form is

$$\left.\begin{array}{l}\text{Cost of at-source BOD}\\ \text{removal for source I}\end{array}\right\} = C1(I) * (S(I) - Z(I)) ** C2(I)$$

where C1(I) and C2(I) are coefficients, S(I) is the present BOD discharge, and Z(I) is the BOD discharge after treatment. Note that if a linear cost function is used then an additional linear programming step may be necessary. Refer to Section 3.6.

(5) Should a facility at location I treat only the source flow of location I-1 or I+1 then the cost is calculated using the at-source treatment cost function of location I-1 or I+1.

Both the regionalization problem and the degree of treatment problem for fixed regional configuration have their duals maximized by a column generation linear programming algorithm. For given values of the dual variables (Lagrange multipliers) the dual function for the regionalization problem is evaluated by dynamic programming. The dual function for the degree of treatment problem is evaluated by a series of univariate minimizations. The corresponding primal solution adds another column to a linear program, the solution of which provides an upper bound on the maximum of the dual and the dual variables for the next iteration. Termination occurs when the upper bound is close enough to the best dual value achieved. The procedure is initiated with large dual variable values so that a feasible primal solution may be obtained.

The input requirements concerning waste sources, DO transfer coefficients, and DO improvement goals are the same for both PORST1 and PORST2. To this basic data set PORST1 adds bound on the allowable sizes of facilities at certain locations as given by the branch and bound procedure. PORST2 adds a specific regional facility arrangement as input. To prepare the problem data for input, the user should number the sources consecutively starting with those in the first segment, then those in the second segment and finally those which can't regionalize. Also the reaches in the river should be numbered from 1 to NREACH. If reaches are assumed completely mixed as in the Thomann (1972) BOD-DO model then the reaches to which waste can be discharged are numbered from 1 to MREACH. Otherwise, all individual discharge points are so numbered.

The input data appear in the following order and format:

Title Card

columns 1-6 PORST1 (to maximize the dual of the regional wastewater or treatment problem)

PORST2 (to solve the minimum cost degree of treatment problem for a given regional facility configuration)

columns 7-80 an optional title.

Stop Criteria Card

TOL1, TOL2 (2F10.7) where

TOL1 = fractional difference of maximum dual value from its upper bound which terminates program.

TOL2 = fractional difference of maximum dual value from its upper bound below which detailed output is printed. Default value is TOL1*1.E4.

Problem Size Card

NL1, NL2, NLNR, NREACH, MREACH (5I2) where

NL1 = number of sources in segment one.

NL2 = number of sources in segment two.

NLNR = number of sources which can't regionalize.

NREACH = number of reaches in river.

MREACH = number of reaches which receive waste discharges or number of discharge points (see above).

Source Cards (one for each source I)

NR(I), NSPLIT, Q(I),SBAR(I), S(I), ZMAX(I), C1(I), C2(I), D(I) (2I2,8X, 7F10.2) where

NR(I) = reach or discharge point to which source discharges.

NSPLIT = number of equally spaced breakpoints for source flow to determine allowable levels of split flows. Default value is 1.

Q(I) = source flow, mgd.

SBAR(I)= untreated influent source BOD, lb/day.

S(I) = current effluent BOD discharged, lb/day.

ZMAX(I)= maximum BOD removal by at-source treatment, lb/day.

C1(I), C2(I) } = coefficients in at-source BOD removal cost function.

D(I) = distance to downstream source, miles (not entered for sources which can't regionalize.

DO Transfer Coefficient Cards

A(I,J) ((I=1,NREACH), J=1,MREACH) (8(E9.3,1X)) where

A(I,J) = change in DO in reach I due to unit change in BOD discharged in reach J or discharge point J, mg/l/lb/day.

DO Standards Cards

STD(I) (I=1,NREACH) (8F10.5) where

STD(I) = DO improvement required in reach I, mg/l.

The following cards would be used with PORST1:

Bounds on Facility Sizes

I, YL, YU (I2,8X,2F10.2) where

I = facility location number

YL = lower bound on allowable facility size, mgd

YU = upper bound on allowable facility size, mgd.

There is one such card for each facility which is bounded in size (initially there are none).

The following cards would be used with PORST2:

Regional Configuration Cards

Y(I) (I=1,NL1+NL2) (8F10.2) where

Y(I) = size of facility at location I, mgd.

The program output gives the values of the primal objective, the dual objective, and the upper bound on the latter at each iteration. After TOL2 is reached the following values are printed at each interation:

J = facility location, J=1, NL1 + NL2 + NLNR

Y(J) = size of facility at location J, mgd

Z(J) = BOD discharged, lb/day

I = river reach, I=1, NREACH

U(I) = dual variable, $/mg/l

STD(I) = DO improvement required, mg/l

G(I) = amount which DO in reach I is below STD(I), mg/l.

From the regionalization restrictions, knowing Y(J) and Z(J) for all locations is sufficient to determine the allocation of sources to regional facilities and the percent BOD removal provided.

There are three types of error messages which can appear. One indicates that a starting feasible solution cannot be found and thus says that the problem is infeasible. Another occurs when the facility pattern inputed for PORST2 is infeasible. The third type indicates that something has gone wrong in the column generation linear program (either infeasibility, unboundedness, or failure to terminate).

PORST is composed of the following subprograms:

MAIN program - initializes data, updates column generation LP and dual variables, checks stop criteria, and prints output.

INPUT - subroutine called by MAIN which reads in input data.

DUAL - subroutine called by MAIN which evaluates dual function.

SIMPLE - subroutine called by MAIN which solves column generation LP. Written by R. J. Clasen, the RAND Corporation, Santa Monica, California, November, 1965, SHARE SDA3384.

PCOST - user supplied function called by DUAL which computes cost of piping between locations I and I+1 given flow piped (YIP) and distance (DIST). Negative YIP means flow is piped from I+1 to I.

TCOST1 - user supplied function called by DUAL which computes

Minimize TCOST1 = T(YI,WI,WIB,ZSTAR) + TAX*ZSTAR

Subject to L $\leq$ ZSTAR $\leq$ U where

T = cost of regional treatment as function of facility size (YI), influent BOD based on current discharges (WI), influent BOD based on raw source BOD's (WIB), and effluent BOD (ZSTAR)

TAX = charge per unit of BOD discharged

L = lower bound on effluent BOD as function of WI and WIB

U = upper bound on effluent BOD as function of WI and WIB.

The values of YI, WI, WIB, and TAX are supplied to the function while TCOST1 and ZSTAR are returned.

TCOST2 - user supplied function called by DUAL which computes

Minimize TCOST2 = T(YI,WI,WIB,C1,C2,ZSTAR) + TAX*ZSTAR
ZSTAR

Subject to WI - ZMAX $\leq$ ZSTAR $\leq$ WI where

T = cost of at-source BOD removal as function of source flow (YI), present BOD discharge (WI), raw source BOD (WIB), effluent BOD discharged (ZSTAR), and two constants (C1 and C2)

TAX = charge per unit of BOD discharged

ZMAX = upper bound on allowable BOD removal.

The values of YI, WI, WIB, C1, C2, and TAX are supplied to the function while TCOST2 and ZSTAR are returned.

To give a numerical demonstration of the program problem WI-PB-TB of Section 5.2 will be run. Recall that it is based on data from the Willamette River. There is a single linear segment of 15 waste sources, 4 of which are dummy sources (additional locations for regional plants). Since the DO transfer coefficients are derived from a BOD-DO model with no longitudinal mixing there are 15 discharge points, one for each source. Hence MREACH = 15. There are 14 reaches or control points at which DO goals must be met. The at-source treatment cost function is taken to be the same as the regional treatment cost function so ZMAX, C1, and C2 are not required as input. All costs are in millions of dollars.

The input data are displayed in Fig. B.1 for the initial run of PORST1 to solve the regionalization problem. The resulting output is shown in Fig. B.2. The maximum dual value is obtained at iteration 14. Since the dual value is not within our established criterion of a 0.1% tolerance of the primal a duality gap exists. However, the primal solution of this iteration would be optimal if the DO goal of reach 11 was increased by only 0.0024 mg/l (ignoring the slight infeasibility of reach 1). The alternate regional configurations which occur for this gap can be identified from iterations 14 and 16 (or 15), the only difference being the size of the facilities built at locations 1 and 2. Using PORST2 to find the minimum cost degree of treatment solution for each configuration (for the original DO goals) would show that the regional facility arrangement of iteration 14 gives the better result, $\$2.523 \times 10^6$. From our lower bound of $\$2.519 \times 10^6$ (the maximum dual value of iteration 14) we see that this solution is guaranteed to be within 0.16% of the true optimum. In fact a single branching would show that it actually is optimum.

```
PCRST1 - PROBLEM W1-PB-TB (ALL COSTS IN MILLION DOLLARS)
.000001   .00020
1500001415
 1        4.83      10000.    3300.                                           6.3
 2        31.3      74000.    66600.                                          15.
 3                                                                            15.
 4                                                                            13.
 5        4.16      11000.    11000.                                          4.
 6        12.9      13000.    10400.                                          12.
 7        14.       11000.    9130.                                           2.
 8        8.4       17000.    17000.                                          16.
 9                                                                            16.
10        14.2      8000.     8000.                                           2.
11        36.8      125000.   112500.                                         17.
12                                                                            16.
13        4.        48000.    48000.                                          16.
14        .33       6000.     5400.                                           8.
15        40.7      95000.    95000.
.1541E-06 .5344E-05 .6864E-05 .7734E-05 .8070E-05 .7977E-05 .7086E-05 .8462E-05
.9956E-05 .9999E-05 .9773E-05 .7728E-05 .7144E-05 .6574E-05
0.        .3687E-05 .6269E-05 .7756E-05 .8443E-05 .8531E-05 .7714E-05 .9227E-05
.1076E-04 .1075E-04 .1045E-04 .8122E-05 .7484E-05 .6868E-05
0.        .7224E-06 .4827E-05 .7219E-05 .8455E-05 .8866E-05 .8292E-05 .1003E-04
.1171E-04 .1167E-04 .1132E-04 .8720E-05 .8022E-05 .7350E-05
0.        0.        .2302E-05 .5849E-05 .7833E-05 .8734E-05 .8652E-05 .1072E-04
.1271E-04 .1272E-04 .1236E-04 .9563E-05 .8801E-05 .8067E-05
0.        0.        .1268E-05 .5168E-05 .7401E-05 .8487E-05 .8637E-05 .1084E-04
.1301E-04 .1307E-04 .1274E-04 .9935E-05 .9155E-05 .8401E-05
0.        0.        0.        .3031E-05 .6212E-05 .7938E-05 .8775E-05 .1135E-04
.1389E-04 .1400E-04 .1369E-04 .1072E-04 .9887E-05 .9075E-05
0.        0.        0.        .2858E-05 .6283E-05 .8131E-05 .9010E-05 .1160E-04
.1405E-04 .1410E-04 .1373E-04 .1062E-04 .9774E-05 .8955E-05
0.        0.        0.        0.        .2453E-05 .6041E-05 .8841E-05 .1233E-04
.1563E-04 .1582E-04 .1550E-04 .1211E-04 .1116E-04 .1023E-04
0.        0.        0.        0.        0.        .1997E-05 .7695E-05 .1236E-04
.1686E-04 .1733E-04 .1714E-04 .1367E-04 .1262E-04 .1160E-04
0.        0.        0.        0.        0.        .1382E-05 .8347E-05 .1345E-04
.1780E-04 .1803E-04 .1759E-04 .1349E-04 .1237E-04 .1131E-04
0.        0.        0.        0.        0.        0.        .4647E-05 .1235E-04
.1898E-04 .1968E-04 .1948E-04 .1528E-04 .1404E-04 .1285E-04
0.        0.        0.        0.        0.        0.        0.        .7389E-05
.1709E-04 .1886E-04 .1947E-04 .1663E-04 .1546E-04 .1428E-04
0.
.6883E-05 .1327E-04 .1695E-04 .1871E-04 .1778E-04 .1668E-04
0.
0.        .6472E-05 .1302E-04 .1937E-04 .1875E-04 .1782E-04
0.
0.        0.        0.        0.        .4607E-05 .1086E-04
0.        0.        0.        0.        0.        .08       .62       .750
2.09      2.51      2.66      1.95      2.10      2.41
```

Figure B.1 Input Data for Problem W1-PB-TB

PCRST - PROGRAM TO CPTIMIZE REGICNAL SEWAGE TREATMENT
****

PCRST1 - PROBLEM W1-PB-TB (ALL COSTS IN MILLICN DOLLARS)

```
ITERATION    1   DUAL =  -0.163296E 03   PRIMAL =   0.757536E 01
UPPER BOUND CN DUAL =  0.757536E 01

ITERATION    2   DUAL =   0.229586E 01   PRIMAL =   0.229586E 01
UPPER BOUND CN DUAL =  0.448735E 01

ITERATION    3   DUAL =  -0.923421E 00   PRIMAL =   0.236382E 01
UPPER BOUND ON DUAL =  0.387226E 01

ITERATION    4   DUAL =   0.134150E 01   PRIMAL =   0.386035E 01
UPPER BOUND CN DUAL =  0.291508E 01

ITERATION    5   DUAL =   0.234468E 01   PRIMAL =   0.277605E 01
UPPER BOUND ON DUAL =  0.265680E 01

ITERATION    6   DUAL =   0.250733E 01   PRIMAL =   0.244532E 01
UPPER BOUND CN DUAL =  0.262725E 01

ITERATION    7   DUAL =   0.223449E 01   PRIMAL =   0.275658E 01
UPPER BOUND ON DUAL =  0.254006E 01

ITERATION    8   DUAL =   0.247074E 01   PRIMAL =   0.268421E 01
UPPER BOUND ON DUAL =  0.253516E 01

ITERATION    9   DUAL =   0.251522E 01   PRIMAL =   0.256308E 01
UPPER BOUND ON DUAL =  0.253183E 01

ITERATION   10   DUAL =   0.249094E 01   PRIMAL =   0.254399E 01
UPPER BOUND ON DUAL =  0.252360E 01

ITERATION   11   DUAL =   0.251808E 01   PRIMAL =   0.249251E 01
UPPER BOUND ON DUAL =  0.252057E 01

ITERATION   12   DUAL =   0.251875E 01   PRIMAL =   0.258210E 01
UPPER BOUND ON DUAL =  0.252020E 01

ITERATION   13   DUAL =   0.251879E 01   PRIMAL =   0.252729E 01
UPPER BOUND ON DUAL =  0.251981E 01

ITERATION   14   DUAL =   0.251943E 01   PRIMAL =   0.250962E 01
```

J	Y(J)	Z(J)	I	U(I)	STD(I)	G(I)
1	4.83	4032.72	1	0.97952E 02	0.0	0.00011
2	31.30	36063.37	2	0.0	0.0	-0.10867
3	0.0	0.0	3	0.0	0.0	-0.19348
4	0.0	0.0	4	0.0	0.0	-0.28072
5	4.16	5416.53	5	0.0	0.0	-0.35858
6	12.90	6738.64	6	0.0	0.08000	-0.33664
7	14.00	5774.30	7	0.0	0.62000	-0.13200
8	8.40	8224.90	8	0.0	0.75000	-0.62794
9	0.0	0.0	9	0.0	2.09000	-0.10125
10	14.20	4215.41	10	0.0	2.51000	-0.00776
11	36.80	47538.78	11	0.51129E 00	2.66000	-0.00244
12	0.0	0.0	12	0.0	1.95000	-0.38902
13	4.00	9957.35	13	0.0	2.10000	-0.25961
14	0.0	0.0	14	0.0	2.41000	-0.03162

Figure B.2 Output for Problem W1-PB-TA

15	41.03	55549.99

UPPER BOUND ON DUAL = 0.251951E 01

ITERATION 15 DUAL = 0.251906E 01 PRIMAL = 0.256263E 01

J	Y(J)	Z(J)	I	U(I)	STD(I)	G(I)
1	0.0	0.0	1	0.98652E 02	0.0	-0.00051
2	36.13	40995.75	2	0.0	0.0	-0.11204
3	0.0	0.0	3	0.0	0.0	-0.19024
4	0.0	0.0	4	0.0	0.0	-0.27363
5	4.16	5419.36	5	0.0	0.0	-0.34943
6	12.90	6740.47	6	0.0	0.08000	-0.32665
7	14.00	5775.53	7	0.0	0.62000	-0.12199
8	8.40	8229.92	8	0.0	0.75000	-0.61524
9	0.0	0.0	9	0.0	2.09000	-0.08586
10	14.20	4216.24	10	0.0	2.51000	0.00795
11	36.80	47633.36	11	0.50901E 00	2.66000	0.01295
12	0.0	0.0	12	0.0	1.95000	-0.37717
13	4.00	10030.66	13	0.0	2.10000	-0.24875
14	0.0	0.0	14	0.0	2.41000	-0.02171
15	41.03	55549.99				

UPPER BOUND ON DUAL = 0.251944E 01

ITERATION 16 DUAL = 0.251932E 01 PRIMAL = 0.257234E 01

J	Y(J)	Z(J)	I	U(I)	STD(I)	G(I)
1	0.0	0.0	1	0.98218E 02	0.0	-0.00051
2	36.13	40870.18	2	0.0	0.0	-0.11250
3	0.0	0.0	3	0.0	0.0	-0.19105
4	0.0	0.0	4	0.0	0.0	-0.27473
5	4.16	5404.14	5	0.0	0.0	-0.35077
6	12.90	6730.59	6	0.0	0.08000	-0.32815
7	14.00	5768.91	7	0.0	0.62000	-0.12588
8	8.40	8202.89	8	0.0	0.75000	-0.62344
9	0.0	0.0	9	0.0	2.09000	-0.10053
10	14.20	4211.80	10	0.0	2.51000	-0.00961
11	36.80	47123.82	11	0.52129E 00	2.66000	-0.00590
12	0.0	0.0	12	0.0	1.95000	-0.39408
13	4.00	9635.67	13	0.0	2.10000	-0.26453
14	0.0	0.0	14	0.0	2.41000	-0.03631
15	41.03	55549.99				

UPPER BOUND ON DUAL = 0.251943E 01

Figure B.2 Continued

The following is a list of the major variables used in the program. The dimensions of all arrays are shown in parentheses.

NL1 - number of sources in segment one.

NL2 - number of sources in segment two.

NLNR - number of sources which cannot regionalize.

NLTOT - total number of sources.

NREACH - number of reaches in river.

MREACH - number of reaches which receive waste discharges or number of discharge points.

NSPLIT - number of equally spaced breakpoints which source flow can be divided into to determine allowable levels of split flows.

NSTAT1 - number of state levels for segment one.

NSTAT2 - number of state levels for segment two.

NSTATE - total number of state levels, NSTAT1 + NSTAT2. Equals the sum of the NSPLIT values for all sources which regionalize plus 1.

NR(I) - reach or discharge point for source I, (MREACH).

Q(I) - source flow of source I, (NLTOT).

SBAR(I) - untreated influent source BOD for source I, (NLTOT).

S(I) - current effluent BOD of source I, (NLTOT).

ZMAX(I) - maximum BOD removal by at-source treatment at I, (NLTOT).

C1(I) - coefficient in at-source BOD removal cost function at I, (NLTOT).

C2(I) - same as above.

D(I),DIST - distance to downstream source (NLTOT).

A(I,J) - change in DO in reach I due to unit change in BOD discharged into reach J or discharge point J, (NREACH,MREACH).

STD(I) - DO improvement required for reach I, (NREACH).

YLOWER(I), YL - lower bound on allowable facility size at I, (NL1+NL2).

YUPPER(I), YU - upper bound on allowable facility size at I, (NL1+NL2).

Y(I) - hydraulic size of facility at I, (NLTOT).

Z(I) - BOD discharged at I, (NLTOT)

U(I) - dual variable for reach I, (NREACH).

G(I) - amount which DO is below STD(I) in reach I, value of DO constraint, (NREACH).

B(I) - right hand side of DO constraint, (NREACH).

UA(I), TAX - dot product of U with Ith column of A, (NREACH).

SUMQ(I) - sum of Q(J), J=1, I (NL1+NL2+1).

STATE(1,J) - state levels of waste flow

STATE(2,J) - state levels of present BOD discharges (3,NSTATE).

STATE(3,J) - state levels of raw BOD's

KYPATH(I) - state path of regional configuration of PORST2, (NL1+NL2+1).

R(I,J) - optimal return at location I being in STATE(1,J), (NL1+NL2,NSTATE).

NSOLY(I,J) - optimal state level of flow for location I in STATE(1,J), (NL1+NL2,NSTATE).

SOLZ(I,J) - optimal BOD discharge at location I in STATE(1,J), (NL1+NL2,NSTATE).

H – value of dual objective function.

F - value of primal objective function.

HMAX - maximum value of dual.

HUB - upper bound on dual.

ICODE - is 1 for PORST1, 2 for PORST2.

TOL1 - value of (HUB - HMAX)/HMAX which terminates program.

TOL2 - value of (HUB - HMAX)/HMAX below which detailed output is printed.

ITER - iteration number.

RR - cost of regional solution as determined by dynamic programming.

YIP - flow piped between source locations.

YI - hydraulic size of treatment facility.

WI - influent BOD based on current BOD discharges.

WIB - influent BOD based on raw source BOD's.

PIPE - cost of piping between adjacent locations.

TREAT - cost of treatment plus effluent charge.

ZSTAR - optimal BOD discharge.

UB - dot product of U with B.

M - number of rows in column generation LP.

N - number of columns in column generation LP.

AA(I,J)(A(I,J) in subroutine SIMPLE) - coefficient matrix of column generation LP, (NREACH + 1, NREACH + max. number of iterations).

BB(I) - right hand side of column generation LP, (NREACH + 1).

CC(I) - objective function coefficients of column generation LP, (NREACH + max. number of iterations).

ZZ(I) - solution of column generation LP, (NREACH + max. number of interations).

P(I) - dual variables to column generation LP, (NREACH + 1).

JH,XX,X,PE - temporary storage arrays used in subroutine SIMPLE, all dimensioned at (NREACH + 1).

E - temporary storage array used in SIMPLE, dimensioned at $(NREACH + 1)^2$.

KO(I) - contains solution information for column generation LP, (6).

The maximum number of iterations used in this version of PORST is 50.

A listing of the entire program follows. The functions PCOST, TCOST1, and TCOST2 correspond to the cost functions PB and TB given in Table 5.7.

```
                C***********************************************************
                C PCRST - PROGRAM TO OPTIMIZE REGIONAL SEWAGE           *
                C TREATMENT (BY USING OUTER LINEARIZATION TO            *
                C MAXIMIZE THE DUAL)                                    *
                C***********************************************************
                C PCRST1 - LEAST COST REGIONAL FACILITY PATTERN AND DEGREE
                C           OF TREATMENT.
                C PCRST2 - LEAST COST DEGREE OF TREATMENT FOR GIVEN REGIONAL
                C          FACILITY PATTERN.
                C
0001                  COMMON A(30,30),B(30),STD(30),U(30),G(30)
                    2 ,UA(30),STATE(40,3),Q(55),SUMQ(55),S(55)
                    3 ,SBAR(55),ZMAX(55),C1(55),C2(55),D(55)
                    4 ,YLOWER(55),YUPPER(55),Y(55),Z(55),NR(55)
                    5 ,KYPATH(55),NREACH,MREACH,NL1,NL2,NLNR,NLTOT
                    6 ,NSTAT1,NSTAT2,ICODE,H,F
0002                  DIMENSION E(961),P(31),JH(31),XX(31),X(31),PE(31),
                    2           AA(31,80),BB(31),CC(80),KO(6),ZZ(80)
0003                  DATA HMAX/-1.E20/,HUB/1.E20/
0004                  IN=5
0005                  IOUT=6
                C
                C***READ IN INPUT. INITIALIZE ARRAY OF L.P.
0006                  CALL INPUT
0007                  IF (ICODE.EQ.0) GO TO 1060
0008                  TOL1=H
0009                  TOL2=F
0010                  JJ=NLTOT
0011                  IF (JJ.LT.NREACH) JJ=NREACH
0012                  KB=0
0013                  M=NREACH+1
0014                  DO 10 I=1,NREACH
0015                  CC(I)=0.
0016                  BB(I)=0.
0017               10 AA(M,I)=0.
0018                  BB(M)=1.
0019                  DO 30 I=1,NREACH
0020                  DO 20 J=1,NREACH
0021                  AA(I,J)=0.
0022                  IF (I.EQ.J) AA(I,J)=1.
0023               20 CONTINUE
0024               30 CONTINUE
                C
                C***PERFORM 1ST ITERATION WITH MULTIPLIERS SUFFICIENTLY
                C***LARGE TO PRODUCE A FEASIBLE SOLUTION.
0025                  DO 40 I=1,NREACH
0026               40 U(I)=10.
0027                  ITER=1
```

```
               N=NREACH
      60       CALL DUAL
C
C***PRINT OUT SOLUTION.
               WRITE(IOUT,1010) ITER,H,F
               IF (H.GT.HMAX) HMAX=H
               IF (ABS(HUB-HMAX).GT.ABS(HMAX*TOL2)) GO TO 65
               WRITE(IOUT,1020)
               DO 120 I=1,JJ
               IF (I.GT.NLTOT) GO TO 100
               WRITE(IOUT,1030) I,Y(I),Z(I)
               GO TO 110
     100       WRITE(IOUT,1030)
     110       IF (I.GT.NREACH) GO TO 120
               WRITE(IOUT,1040) I,U(I),STD(I),G(I)
     120       CONTINUE
C
C***UPDATE L.P. ARRAY. SOLVE L.P. TO OBTAIN NEW
C***MULTIPLIERS.
      65       N=N+1
               CC(N)=F
               DO 70 I=1,NREACH
      70       AA(I,N)=G(I)
               AA(M,N)=1.
               ZZ(N)=0.
               IF (ITER.GT.1) KB=1
      75       CALL SIMPLE(KB,M,N,AA,BB,CC,KO,ZZ,P,JH,XX,X,PE,E)
               IF (KO(1).NE.0) GO TO 1000
               HUB=-P(M)
               WRITE(IOUT,2020) HUB
    2020       FORMAT(' UPPER BOUND ON DUAL =',E14.6)
               IF (HUB.LE.HMAX) GO TO 1060
               IF (ABS(HUB-HMAX).LE.ABS(HMAX*TOL1)) GO TO 1060
               DO 80 I=1,NREACH
      80       U(I)=P(I)
               ITER=ITER+1
               IF (N.GE.80) GO TO 1060
               GO TO 60
C
C***ERROR MESSAGES.
    1000       IF (KO(1)-2) 1011,1006,1008
    1011       IF (ITER.NE.1) GO TO 1004
               IF (U(1).GT.1.E6) GO TO 1002
               DO 1001 I=1,NREACH
    1001       U(I)=U(I)**2
               GO TO 60
    1002       WRITE(IOUT,1003)
    1003       FORMAT('0 PRIMAL IS INFEASIBLE')
```

```
                  GO TO 1060
          1004    WRITE(IOUT,1005)
          1005    FORMAT('0 L.P. INFEASIBLE')
                  GO TO 1060
          1006    WRITE(IOUT,1007)
          1007    FORMAT('0 L.P. UNBOUNDED')
                  GO TO 1060
          1008    WRITE(IOUT,1009)
          1009    FORMAT('0 L.P. COULD NOT TERMINATE')
          1010    FORMAT ('0ITERATION ',I4,'    DUAL = ',E14.6,
             2          '    PRIMAL = ',E14.6)
          1020    FORMAT(2X,'J',5X,'Y(J)',6X,'Z(J)',11X,
             2          'I',2X,'U(I)',9X,'STD(I)',2X,'G(I)')
          1030    FORMAT(1X,I2,2F10.2)
          1040    FORMAT('+',30X,I3,1X,E12.5,2(1X,F8.5))
          1060    STOP
                  END
```

```
          SUBROUTINE INPUT
C***READS IN PROBLEM DATA AND DOES PRELIMINARY CALCULATIONS.
C
          COMMON A(30,30),B(30),STD(30),U(30),G(30)
     2    ,UA(30),STATE(40,3),Q(55),SUMQ(55),S(55)
     3    ,SBAR(55),ZMAX(55),C1(55),C2(55),D(55)
     4    ,YLOWER(55),YUPPER(55),Y(55),Z(55),NR(55)
     5    ,KYPATH(55),NREACH,MREACH,NL1,NL2,NLNR,NLTOT
     6    ,NSTAT1,NSTAT2,ICODE,H,F
          DIMENSION TITLE(20)
          IN=5
          IOUT=6
C
C***READ IN PROBLEM TITLE.
          READ(IN,8)ICODE,(TITLE(I),I=1,19)
    8     FORMAT (5X,I1,18A4,A2)
C
C***READ IN STOP AND PRINTOUT CRITERIA.
          READ(IN,2000)H,F
 2000     FORMAT(2F10.7)
          IF (F.EQ.0.) F=H*1.E4
C
C***READ IN NO. LOCATIONS IN EACH SEGMENT OF
C***REGIONAL SYSTEM, NO. LOCATIONS WHICH CAN'T
C***REGIONALIZE, AND NO. REACHES IN RIVER.
          READ(IN,1)NL1,NL2,NLNR,NREACH,MREACH
    1     FORMAT (5I2)
C
C***READ IN FLOW, BOD,COST AND DISTANCE DATA FOR EACH LOCATION.
C***AT SAME TIME CONSTRUCT STATE VECTORS.
          NLTOT=NL1+NL2+NLNR
          STATE(1,1)=0.
          STATE(1,2)=0.
          STATE(1,3)=0.
          QOLD=0.
          NSTATE=1
          DO 1020 I=1,NLTOT
          READ(IN,2)NR(I),NSPLIT,Q(I),SBAR(I),S(I),ZMAX(I), C1(I),
     2                C2(I),D(I)
    2     FORMAT(2I2,6X,7F10.2)
          IF (I.GT.NL1+NL2) GO TO 1020
          IF (NSPLIT.EQ.0) NSPLIT=1
          YLOWER(I)=0.
          YUPPER(I)=1.E20
          SUMQ(I)=QOLD+Q(I)
          QOLD=SUMQ(I)
          IF (Q(I).EQ.0.) GO TO 1005
          DO 1010 J=1,NSPLIT
```

```
              NS=NSTATE+J
              R=FLOAT(J)/NSPLIT
              STATE(NS,1)=STATE(NSTATE,1)+Q(I)*R
              STATE(NS,3)=STATE(NSTATE,3)+SBAR(I)*R
        1010  STATE(NS,2)=STATE(NSTATE,2)+S(I)*R
              NSTATE=NS
        1005  IF (I.EQ.NL1) NSTAT1=NSTATE
        1020  CONTINUE
              NSTAT2=NSTATE-NSTAT1
C
C***READ IN RIVER TRANSFER COEFFICIENTS AND D.O. STDS.
              DO 10 J=1,MREACH
              READ(IN,5)(A(I,J),I=1,NREACH)
           5  FORMAT(8(E9.3,1X))
          10  CONTINUE
              READ(IN,6)(STD(I),I=1,NREACH)
           6  FORMAT(8F10.5)
C
C***PRINT OUTPUT HEADING.
              WRITE(IOUT,7)
           7  FORMAT ('1',10X,'PORST - PROGRAM TO OPTIMIZE REGIONAL ',
            2 'SEWAGE TREATMENT'/11X,'*********************************',
            3 '*************************'//)
           .  WRITE(IOUT,9) ICODE,(TITLE(I),I=1,19)
           9  FORMAT(11X,'PORST',I1,18A4,A2)
             IF (ICODE.EQ.2) GO TO 205
C
C***READ IN BOUNDS ON PLANT SIZES.
         270  READ(IN,300,END=400) I,YL,YU
         300  FORMAT(I2,8X,2F10.2)
              YLOWER(I)=YL
              YUPPER(I)=YU
              GO TO 270
C
C***REGIONAL CONFIGURATION IS GIVEN. READ IN PLANT SIZES
C***AND COMPUTE CORRESPONDING STATE PATH.
         205  NPR=NL1+NL2
              READ(IN,200)(Y(I),I=1,NPR)
         200  FORMAT(8F10.2)
              YSUM=0.
              KYPATH(1)=1
              NS=1
              DO 250 I=1,NPR
              YSUM=YSUM+Y(I)
              DO 210 J=NS,NSTATE
              IF (ABS(YSUM-STATE(J,1)).LT..001) GO TO 240
         210  CONTINUE
              WRITE(IOUT,230)
```

```
  230  FORMAT('0 REGIONAL FACILITY PATTERN INFEASIBLE')
       ICODE=0
       GO TO 1060
  240  KYPATH(I+1)=J
       YSUM=STATE(J,1)
       NS=J
  250  CONTINUE
C
C***FIND R.H.S. OF QUALITY CONSTRAINTS.
  400  DO 70 I=1,NREACH
       B(I)=STD(I)
       DO 60 J=1,NLTOT
   60  B(I)=B(I)-(A(I,NR(J))*S(J))
   70  CONTINUE
 1060  RETURN
       END
```

```
0001          SUBROUTINE DUAL
        C***EVALUATES DUAL FUNCTION BY SOLVING A DYNAMIC PROG.
        C
        C***NOTE - IF A LOCATION DISCHARGES ONLY THE SOURCE FLOW FROM
        C          AN ADJACENT SOURCE, THEN TREATMENT (AND ITS ASSOC-
        C          IATED COST) TAKES PLACE AT THE ADJACENT SOURCE.
        C
0002          COMMON A(30,30),B(30),STD(30),U(30),G(30)
             2 ,UA(30),STATE(40,3),Q(55),SUMQ(55),S(55)
             3 ,SBAR(55),ZMAX(55),C1(55),C2(55),D(55)
             4 ,YLOWER(55),YUPPER(55),Y(55),Z(55),NR(55)
             5 ,KYPATH(55),NREACH,MREACH,NL1,NL2,NLNR,NLTOT
             6 ,NSTAT1,NSTAT2,ICODE,H,F
0003          DIMENSION R(55,55),NSOLY(55,55),SOLZ(55,55)
        C
        C***FORM U*B AND U*COLUMNS OF A.
0004          UB=0.
0005          DO 1010 I=1,NREACH
0006     1010 UB=UB+(U(I)*B(I))
0007          DO 1030 J=1,MREACH
0008          UA(J)=0.
0009          DO 1020 I=1,NREACH
0010     1020 UA(J)=UA(J)+(U(I)*A(I,J))
0011     1030 CONTINUE
        C
        C***RUN THROUGH D.P. RECURSION FOR LOCATIONS IN REGIONAL SYSTEM.
        C***PLACE FLOW SOLUTION IN NSOLY,TREATMENT SOLUTION IN SOLX.
0012          RR=0.
0013          IF (NL1.EQ.0) GO TO 75
0014          IF (ICODE.EQ.2) GO TO 200
        C***SOLVE FOR LOCATIONS IN SEGMENT 1 FIRST.
0015          JLOWER=1
0016          JUPPER=NSTAT1
0017          ILOWER=2
0018          IUPPER=NL1
        C***SOLVE RECURSION FOR LOCATION 1 FOR ALL STATE VALUES.
0019        5 I=ILOWER-1
0020          DO 10 J2=JLOWER,JUPPER
        C***COMPUTE PIPING COSTS.
0021          YIP=SUMQ(I)-STATE(J2,1)
0022          PIPE=PCOST(YIP,D(I))
        C***COMPUTE WASTE QUANTITIES.
0023          YI=STATE(J2,1)-STATE(JLOWER,1)
0024          IF (YI.LT.YLOWER(I).OR.YI.GT.YUPPER(I)) GO TO 110
0025          IF (YI*(Q(I)-YIP).LT.0..OR.YI*(YI+YIP).LT.0.) GO TO 110
0026          WI=STATE(J2,2)-STATE(JLOWER,2)
0027          WIB=STATE(J2,3)-STATE(JLOWER,3)
        C***FIND OPTIMUM TREATMENT LEVEL AND COST.
```

```
C***CHECK IF SOURCE FLOW ONLY IS TREATED.
      IF (ABS(YI-Q(I)).LT..001) GO TO 105
C***CHECK IF ADJACENT SOURCE FLOW ONLY IS TREATED.
      IF (ABS(YI-Q(I+1)).LT..001.AND.Q(I).NE.0.) GO TO 102
C***REGIONAL TREATMENT PROVIDED.
      TREAT=TCOST1(YI,WI,WIB,UA(NR(I)),ZSTAR)
      GO TO 115
C***AT SOURCE TREATMENT PROVIDED AT ADJACENT SOURCE.
  102 TREAT=TCOST2(YI,WI,WIB,UA(NR(I)),ZSTAR,
     2            ZMAX(I+1),C1(I+1),C2(I+1))
      GO TO 115
C***AT SOURCE TREATMENT PROVIDED.
  105 TREAT=TCOST2(YI,WI,WIB,UA(NR(I)),ZSTAR,
     2            ZMAX(I),C1(I),C2(I))
      GO TO 115
  110 TREAT=1.E20
  115 R(I,J2)=RR+TREAT+PIPE
      NSOLY(I,J2)=JLOWER
   10 SOLZ(I,J2)=ZSTAR
C***SOLVE RECURSION FOR REMAINING LOCATIONS.
      JJ=JLOWER
      DO 50 I=ILOWER,IUPPER
      IF (I.LT.IUPPER) GO TO 20
      JJ=JUPPER
   20 DO 40 J2=JJ,JUPPER
      R(I,J2)=1.E20
      DO 30 J1=JLOWER,J2
C***COMPUTE PIPING COSTS.
      YIP=SUMQ(I)-STATE(J2,1)
      PIPE=PCOST(YIP,D(I))
C***COMPUTE WASTE QUANTITIES.
      YI=STATE(J2,1)-STATE(J1,1)
      IF (YI.LT.YLOWER(I).OR.YI.GT.YUPPER(I)) GO TO 125
      IF (YI*(Q(I)-YIP).LT.0..OR.YI*(YI+YIP).LT.0.) GO TO 125
      WI=STATE(J2,2)-STATE(J1,2)
      WIB=STATE(J2,3)-STATE(J1,3)
C***FIND OPTIMUM TREATMENT LEVEL AND COST.
C***CHECK IF SOURCE FLOW ONLY IS TREATED.
      IF (ABS(YI-Q(I)).LT..001) GO TO 120
      IF (I.EQ.IUPPER.OR.Q(I).EQ.0.) GO TO 116
C***CHECK IF ADJACENT SOURCE FLOW ONLY IS TREATED.
      IF (ABS(YI-Q(I-1)).LT..001) GO TO 117
      IF (ABS(YI-Q(I+1)).LT..001) GO TO 118
C***REGIONAL TREATMENT PROVIDED.
  116 TREAT=TCOST1(YI,WI,WIB,UA(NR(I)),ZSTAR)
      GO TO 130
C***AT SOURCE TREATMENT PROVIDED.
  120 TREAT=TCOST2(YI,WI,WIB,UA(NR(I)),ZSTAR,
```

```
     2                 ZMAX(I),C1(I),C2(I))
        GO TO 130
C***AT SOURCE TREATMENT PROVIDED AT ADJACENT SOURCE.
  117   TREAT=TCOST2(YI,WI,WIB,UA(NR(I)),ZSTAR,
     2                 ZMAX(I-1),C1(I-1),C2(I-1))
        GO TO 130
  118   TREAT=TCOST2(YI,WI,WIB,UA(NR(I)),ZSTAR,
     2                 ZMAX(I+1),C1(I+1),C2(I+1))
        GO TO 130
  125   TREAT=1.E20
  130   RTEMP=R(I-1,J1)+TREAT+PIPE
        IF (RTEMP.GE.R(I,J2)) GO TO 30
        R(I,J2)=RTEMP
        NSOLY(I,J2)=J1
        SOLZ(I,J2)=ZSTAR
   30   CONTINUE
   40   CONTINUE
   50   CONTINUE
        RR=R(I,J2)
        IF (I.EQ.NL1+NL2) GO TO 60
C***NOW SOLVE FOR LOCATIONS IN SEGMENT 2.
        JLOWER=NSTAT1
        JUPPER=NSTAT1+NSTAT2
        ILOWER=NL1+2
        IUPPER=NL1+NL2
        GO TO 5
C
C***BACKTRACK THROUGH D.P. TO FIND OPTIMAL FLOW AND TREATMENT
C***VALUES AT ALL LOCATIONS.
   60   NS=NSTAT1+NSTAT2
        DO 70 J=1,IUPPER
        I=IUPPER-J+1
        Y(I)=STATE(NS,1)-STATE(NSOLY(I,NS),1)
        Z(I)=SOLZ(I,NS)
   70   NS=NSOLY(I,NS)
        GO TO 75
C
C***REGIONAL CONFIGURATION IS GIVEN. COMPUTE COSTS.
  200   NPR=NL1+NL2
        DO 210 I=1,NPR
        J1=KYPATH(I)
        J2=KYPATH(I+1)
C***COMPUTE PIPING COSTS.
        YIP=SUMQ(I)-STATE(J2,1)
        PIPE=PCOST(YIP,D(I))
C***COMPUTE WASTE QUANTITIES.
        YI=STATE(J2,1)-STATE(J1,1)
        WI=STATE(J2,2)-STATE(J1,2)
```

```
0097           WIB=STATE(J2,3)-STATE(J1,3)
         C***FIND OPTIMAL TREATMENT LEVEL AND COST.
         C***CHECK IF SOURCE FLOW ONLY IS TREATED.
0098           IF (ABS(YI-Q(I)).LT..001) GO TO 215
0099           IF (I.EQ.1.OR.I.EQ.NL1+1) GO TO 201
0100           IF (Q(I).EQ.0.) GO TO 202
         C***CHECK IF ADJACENT SOURCE FLOW ONLY IS TREATED.
0101           IF (ABS(YI-Q(I-1)).LT..001) GO TO 203
0102       201 IF (I.EQ.NL1.OR.I.EQ.NPR) GO TO 202
0103           IF (Q(I).EQ.0.) GO TO 202
0104           IF (ABS(YI-Q(I+1)).LT..001) GO TO 204
         C***REGIONAL TREATMENT PROVIDED.
0105       202 TREAT=TCOST1(YI,WI,WIB,UA(NR(I)),ZSTAR)
0106           GO TO 220
         C***AT SOURCE TREATMENT PROVIDED AT ADJACENT SOURCE.
0107       203 TREAT=TCOST2(YI,WI,WIB,UA(NR(I)),ZSTAR,
             2             ZMAX(I-1),C1(I-1),C2(I-1))
0108           GO TO 220
0109       204 TREAT=TCOST2(YI,WI,WIB,UA(NR(I)),ZSTAR,
             2             ZMAX(I+1),C1(I+1),C2(I+1))
0110           GO TO 220
         C***AT SOURCE TREATMENT PROVIDED.
0111       215 TREAT=TCOST2(YI,WI,WIB,UA(NR(I)),ZSTAR,
             2             ZMAX(I),C1(I),C2(I))
0112       220 RR=RR+TREAT+PIPE
0113       210 Z(I)=ZSTAR
         C
         C***COMPUTE PORTION OF COST CONTRIBUTED BY NONREGIONAL PLANTS.
0114        75 RNR=0.
0115           IF (NLNR.EQ.0.) GO TO 85
0116           NPR=NL1+NL2+1
0117        74 DO 76 I=NPR,NLTOT
0118           RNR=RNR+TCOST2(Q(I),S(I),SBAR(I),UA(NR(I)),ZSTAR,
             2             ZMAX(I),C1(I),C2(I))
0119           Y(I)=Q(I)
0120        76 Z(I)=ZSTAR
         C
         C***DUAL=D.P. SOLN. + NONREG. SOLN. + U*B
0121        85 H=RR+RNR+UB
         C
         C***FIND VALUE OF EACH CONSTRAINT AND OF PRIMAL.
0122           GU=0.
0123           DO 100 I=1,NREACH
0124           G(I)=B(I)
0125           DO 90 J=1,NLTOT
0126        90 G(I)=G(I)+(A(I,NR(J))*Z(J))
0127       100 GU=GU+(U(I)*G(I))
0128           F=H-GU
```

```
      RETURN
      END
```

```
      FUNCTION PCOST(YIP,DIST)
C***PIPING COST FUNCTION FROM SMITH (1971).
C***ANNUAL COST IN MILLION DOLLARS (PRESENT VALUE
C   FACTOR OF 13.)
      IF (YIP) 40,100,10
   10 IF (YIP.GT..5) GO TO 20
      PCOST=(.149653*.0548*YIP**.53088)*DIST
      GO TO 90
   20 IF (YIP.GT.2.5) GO TO 30
      PCOST=(.154697*.0548*YIP**.5787)*DIST
      GO TO 90
   30 PCOST=(.165346*.0548*YIP**.50604)*DIST
      GO TO 90
   40 YP=-YIP
      IF (YP.GT..3) GO TO 50
      PCOST=(.092609*.0548*YP**.49544)*DIST
      GO TO 80
   50 IF (YP.GT.1.) GO TO 60
      PCOST=(.098228*.0548*YP**.54427)*DIST
      GO TO 80
   60 IF (YP.GT.5.) GO TO 70
      PCOST=(.098228*.0548*YP**.58505)*DIST
      GO TO 80
   70 PCOST=(.0941*.0548*YP**.61173)*DIST
   80 PCOST=PCOST+(.414387*.071*YP**.75699)
   90 RETURN
  100 PCOST=0.
      RETURN
      END
```

```
      FUNCTION TCOST1(YI,WI,WIB,TAX,ZSTAR)
C***TREATMENT COST FUNCTION FROM HASS (1970).
C***ANNUAL COST IN MILLION DOLLARS.
      IF (YI.EQ.0.) GO TO 10
      T1=365.E-6*(160.8+26.7*YI)
      T2=365.E-6*(640.7+255.7*YI)
      X1=TAX*WIB/(2.*T2)
      X2=.55-X1
      ZSTAR=WIB*X2
      IF (ZSTAR.GT..55*WIB) ZSTAR=.55*WIB
      IF (ZSTAR.LT..02*WIB) ZSTAR=.02*WIB
      T3=(.55-(ZSTAR/WIB))**2
      TCOST1=T1+(T2*T3)+(TAX*ZSTAR)
      RETURN
   10 TCOST1=0.
      ZSTAR=0.
      RETURN
      END
```

```
      FUNCTION TCOST2(YI,WI,WIB,TAX,ZSTAR,I)
      TCOST2=TCOST1(YI,WI,WIB,TAX,ZST)
      ZSTAR=ZST
       RETURN
      END
```

```
C  AUTOMATIC SIMPLEX        REDUNDANT EQUATIONS CAUSE INFEASIBILITY
      SUBROUTINE SIMPLE(INFLAG,MX,NN,A,B,C,KO,KB,P,JH,X,Y,PE,E)
      REAL B(1),C(1),P(1),X(1),Y(1),PE(1),E(1)
      INTEGER INFLAG,MX,NN,KO(6),KB(1),JH(1)
      EQUIVALENCE (XX,LL)
C THE FOLLOWING DIMENSION SHOULD BE THE SAME HERE AS IT IS IN CALLER.
      REAL A(31,80)
      REAL AA,AIJT,BB,COST,DT,RCOST,TEXP,TPIV,TY,XOLD,XX,XY,YI,YMAX
      LOGICAL FEAS,VER,NEG,TRIG,KQ,ABSC
      INTEGER I,IA,INVC,IR,ITER,J,JT,K,KBJ,L,LL,M,M2,MM,N
      INTEGER NCUT,NPIV,NUMVR,NVER
C
C          SET INITIAL VALUES, SET CONSTANT VALUES
      ITER = 0
      NUMVR = 0
      NUMPV = 0
      M = MX
      N = NN
      TEXP  =  .5**16
      NCUT  =  4*M + 10
      NVER  =  M/2  +  5
      M2 = M**2
      FEAS = .FALSE.
      IF (INFLAG.NE.0)  GO TO 1400
C* 'NEW'    START  PHASE ONE WITH SINGLETON BASIS
      DO 1402  J = 1,N
        KB(J) = 0
        KQ = .FALSE.
        DO 1403  I = 1,M
          IF (A(I,J).EQ.0.0)  GO TO 1403
          IF (KQ.OR.A(I,J).LT.0.0) GO TO 1402
          KQ = .TRUE.
 1403   CONTINUE
        KB(J) = 1
 1402 CONTINUE
 1400 DO 1401  I = 1,M
        JH (I) = -1
 1401 CONTINUE
C* 'VER'    CREATE INVERSE FROM 'KB' AND 'JH'      (STEP 7)
 1320 VER = .TRUE.
      INVC = 0
      NUMVR   =  NUMVR  +1
      TRIG = .FALSE.
      DO 1101  I = 1,M2
        E(I) = 0.0
 1101 CONTINUE
      MM=1
      DO 1113  I = 1,M
```

```
                E(MM) = 1.0
                PE(I)  =  0.0
                X(I) =  B(I)
                IF (JH(I) .NE.0) JH(I) = -1
                MM = MM + M + 1
      1113 CONTINUE
C                      FORM INVERSE
           JT=0
      1103 JT=JT+1
             IF (KB(JT).EQ.0) GO TO 1102
               GO TO 600
C 600          CALL JMY
C                                    CHOOSE PIVOT
      1114     TY = 0.0
               KQ = .FALSE.
               DO 1104 I = 1,M
                 IF (JH(I).NE.-1.OR.ABS(Y(I)).LE.TPIV)  GO TO 1104
                 IF (KQ) GO TO 1116
                 IF (X(I).EQ.0.)  GO TO 1115
                 IF (ABS(Y(I)/X(I)).LE.TY)  GO TO 1104
                 TY = ABS(Y(I)/X(I))
                 GO TO 1118
      1115       KQ = .TRUE.
                 GO TO 1117
      1116       IF (X(I).NE.0..OR.ABS(Y(I)).LE.TY)  GO TO 1104
      1117       TY = ABS(Y(I))
      1118       IR = I
      1104     CONTINUE
               KB(JT) = 0
C                                    TEST PIVOT
               IF (TY.LE.0.)   GO TO 1102
C                                    PIVOT
               GO TO 900
C 900          CALL PIV
      1102 IF (JT.LT.N) GO TO 1103
C                         RESET ARTIFICIALS
           DO 1109  I = 1,M
             IF (JH(I).EQ.-1)  JH(I) = 0
             IF (JH(I).EQ.0)  FEAS = .FALSE.
      1109 CONTINUE
      1200 VER = .FALSE.
C                 ***           PERFORM ONE ITERATION          ***
C* 'XCK'     DETERMINE FEASIBILITY                          (STEP 1)
           NEG = .FALSE.
           IF (FEAS)  GO TO 500
           FEAS= .TRUE.
           DO 1201  I = 1,M
             IF (X(I).LT.0.0)  GO TO 1250
```

```
            IF (JH(I).EQ.0)  FEAS = .FALSE.
 1201 CONTINUE
C* 'GET'    GET APPLICABLE PRICES                            (STEP 2)
        IF (.NOT.FEAS)  GO TO 501
  500 DO 503 I = 1,M
          P(I) = PE(I)
          IF (X(I).LT.0.)  X(I) = 0.
  503 CONTINUE
        ABSC = .FALSE.
        GO TO 599
 1250 FEAS = .FALSE.
      NEG  = .TRUE.
  501 DO 504  J = 1, M
          P(J) = 0.
  504 CONTINUE
      ABSC = .TRUE.
      DO 505  I = 1,M
        MM = I
        IF (X(I).GE.0.0)  GO TO 507
        ABSC = .FALSE.
        DO 508 J = 1,M
          P(J) = P(J) + E(MM)
          MM = MM + M
  508   CONTINUE
        GO TO 505
  507   IF (JH(I).NE.0) GO TO 505
        IF (X(I).NE.0.) ABSC = .FALSE.
        DO 510  J = 1,M
          P(J) = P(J) - E(MM)
          MM = MM + M
  510   CONTINUE
  505 CONTINUE
C* 'MIN'    FIND MINIMUM REDUCED COST                        (STEP 3)
  599 JT = 0
      BB  =  0.0
      DO 701  J =1,N
        IF (KB(J).NE.0)   GO TO 701
        DT = 0.0
        DO 303 I = 1,M
          DT = DT + P(I) * A(I,J)
  303   CONTINUE
        IF (FEAS)  DT = DT + C(J)
        IF  (ABSC)  DT = - ABS(DT)
        IF (DT.GE.BB)  GO TO 701
        BB = DT
        JT = J
  701 CONTINUE
C  TEST FOR NO PIVOT COLUMN
```

```
            IF (JT.LE.0)  GO TO 203
      C  TEST FOR ITERATION LIMIT EXCEEDED
            IF (ITER.GE.NCUT)  GO TO 160
            ITER = ITER +1
      C* 'JMY'     MULTIPLY INVERSE TIMES A(.,JT)          (STEP 4)
        600 DO 610  I= 1,M
              Y(I) = 0.0
        610 CONTINUE
            LL = 0
            COST = C(JT)
            DO 605   I= 1,M
              AIJT = A(I,JT)
              IF (AIJT.EQ.0.) GO TO 602
              COST = COST + AIJT * PE(I)
              DO  606  J = 1,M
                LL = LL + 1
                Y(J) = Y(J) + AIJT * E(LL)
        606   CONTINUE
              GO TO 605
        602   LL = LL + M
        605 CONTINUE
      C         COMPUTE PIVOT TOLERANCE
            YMAX = 0.0
            DO 620  I   =   1,M
              YMAX = AMAX1( ABS(Y(I)),YMAX )
        620 CONTINUE
            TPIV  =   YMAX * TEXP
      C              RETURN TO INVERSION ROUTINE, IF INVERTING
            IF (VER)  GO TO 1114
      C        COST TOLERANCE CONTROL
            RCOST = YMAX/BB
            IF (TRIG.AND.BB.GE.-TPIV)  GO TO 203
            TRIG = .FALSE.
            IF (BB.GE.-TPIV)  TRIG = .TRUE.
      C* 'ROW'     SELECT PIVOT ROW                          (STEP 5)
      C AMONG EQS. WITH X=0, FIND MAXIMUM Y  AMONG ARTIFICIALS, OR, IF NONE,
      C  GET MAX POSITIVE Y(I) AMONG REALS.
            IR = 0
            AA = 0.0
            KQ = .FALSE.
            DO  1050  I =1,M
              IF (X(I).NE.0.0.OR.Y(I).LE.TPIV)  GO TO 1050
              IF (JH(I).EQ.0)  GO TO 1044
              IF (KQ) GO TO 1050
       1045   IF (Y(I).LE.AA)  GO TO 1050
              GO TO 1047
       1044   IF (KQ) GO TO 1045
              KQ = .TRUE.
```

```
      1047   AA = Y(I)
             IR = I
      1050 CONTINUE
           IF (IR.NE.0)  GO TO 1099
           AA = 1.0E+20
C                        FIND MIN. PIVOT AMONG POSITIVE EQUATIONS
           DO 1010  I  = 1,M
             IF (Y(I).LE.TPIV.OR.X(I).LE.0.0.OR.Y(I)*AA.LE.X(I) ) GO TO 10
             AA = X(I)/Y(I)
             IR = I
      1010 CONTINUE
           IF (.NOT.NEG)  GO TO 1099
C  FIND PIVOT AMONG NEGATIVE EQUATIONS, IN WHICH X/Y IS LESS THAN THE
C MINIMUM X/Y IN THE POSITIVE EQUATIONS, THAT HAS THE LARGEST ABSF(Y)
           BB = - TPIV
           DO 1030  I = 1,M
             IF (X(I).GE.0..OR.Y(I).GE.BB.OR.Y(I)*AA.GT.X(I) )  GO TO 1030
             BB = Y(I)
             IR = I
      1030 CONTINUE
C  TEST FOR NO PIVOT ROW
      1099 IF  (IR.LE.0)  GO TO 207
C* 'PIV'     PIVOT ON (IR,JT)                              (STEP 6)
           IA = JH(IR)
           IF (IA.GT.0)  KB(IA) = 0
       900 NUMPV  =  NUMPV  +  1
           JH(IR) = JT
           KB(JT) = IR
           YI = -Y(IR)
           Y(IR) = -1.0
           LL = 0
C                                   TRANSFORM INVERSE
           DO  904  J = 1,M
             L = LL + IR
             IF (E(L).NE.0.0)  GO TO 905
             LL = LL + M
             GO TO 904
       905   XY = E(L) / YI
             PE(J) = PE(J) + COST * XY
             E(L) = 0.0
             DO 906  I = 1,M
               LL  =  LL + 1
               E(LL) = E(LL)  + XY * Y(I)
       906   CONTINUE
       904 CONTINUE
C                                   TRANSFORM X
           XY  =  X(IR) / YI
           DO 908  I  = 1, M
```

```
            XOLD = X(I)
            X(I) = XOLD + XY * Y(I)
            IF (.NOT.VER.AND.X(I).LT.0..AND.XOLD.GE.0.)  X(I) = 0.
  908 CONTINUE
      Y(IR) = -YI
      X(IR) = -XY
      IF (VER) GO TO 1102
      IF (NUMPV.LE.M) GO TO 1200
C  TEST FOR INVERSION ON THIS ITERATION
      INVC    = INVC    +1
      IF  (INVC.EQ.NVER)   GO TO 1320
      GO  TO  1200
C*  END OF ALGORITHM, SET EXIT VALUES                  ***
  207 IF (.NOT.FEAS.OR.RCOST.LE.-1000.)  GO TO 203
C              INFINITE SOLUTION
      K   =  2
      GO TO 250
C           PROBLEM IS CYCLING
  160 K   =  4
      GO TO 250
C           FEASIBLE OR INFEASIBLE SOLUTION
  203 K   =  0
  250 IF (.NOT.FEAS)  K = K + 1
      DO 1399  J = 1,N
        XX  =  0.0
        KBJ  =  KB(J)
        IF (KBJ.NE.0)  XX = X(KBJ)
        KB(J) = LL
 1399 CONTINUE
      KO(1) = K
      KO(2) = ITER
      KO(3) = INVC
      KO(4) = NUMVR
      KO(5) = NUMPV
      KO(6) = JT
      RETURN
       END
```

LIST OF REFERENCES

Adams, B. J. and R. S. Gemmell, Water Quality Evaluation of Regionalized Wastewater Systems, Research Report No. 70, Water Resources Center, Univ. of Illinois at Urbana-Champaign, August 1973.

Bazaraa, M. S., "Geometry and Resolution of Duality Gaps", Naval Res. Log. Quart., Vol. 20, No. 2, June 1973.

Bellman, R. and S. E. Dreyfus, Applied Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1962.

Bellmore, M., H. J. Greenberg, and J. J. Jarvis, "Generalized Penalty-Function Concepts in Mathematical Optimization", Opns. Res., Vol. 18, No. 2, 1970.

Bhalla, H. S. and R. F. Rikkers, Multi-Time Period, Facilities Location Problems, Publication No. 21, Water Resources Research Center, University of Massachusetts, Amherst, Mass., 1971.

Brooks, R. and A. M. Geoffrion, "Finding Everett's Lagrange Multipliers by Linear Programming", Opns. Res., Vol. 14, No. 6, 1966.

Converse, A. O., "Optimum Number and Location of Treatment Plants", Jour. Water Poll. Control Fed., Vol. 44, No. 8, August 1972.

Dantzig, G. B., Linear Programming and Extensions, Princeton University Press, Princeton, New Jersey, 1963.

Deininger, R. A. and S. Y. Su, "Regional Waste Water Treatment Systems", Paper presented at the 18th International Meeting of the Institute of Management Sciences, Washington, D.C., March 21-24, 1971.

Dobbins, W. E., "BOD and Oxygen Relationships in Streams", Jour. San. Eng. Div., Proc. ASCE, Vol. 90, No. SA3, June 1964.

Duffin, R. J., E. L. Peterson, and C. Zener, Geometric Programming, Wiley, New York, N.Y., 1967.

Eaves, B. C. and W. I. Zangwill, "Generalized Cutting Plane Algorithms", SIAM J. on Control, Vol. 9, No. 4, 1971.

Elmaghraby, S. E., Some Network Models in Management Science, Springer-Verlag, Berlin, 1970.

Enviro Control, Inc., Systems Analysis for Water Quality Management - Survey and Abstracts, Office of Water Programs, U.S. Environmental Protection Agency, September 1971.

Environmental Protection Agency, Guidelines - Water Quality Management Planning, Water Quality Office, Washington, D.C., January 1971.

Everett, H., "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources", Opns. Res., Vol. 11, No. 3, 1963.

Falk, J. E., "Lagrange Multipliers and Nonlinear Programming", J. Math. Anal. Appl., Vol. 19, No. 1, 1967.

Fiacco, A. V. and G. P. McCormick, "The Sequential Unconstrained Minimization Technique for Nonlinear Programming", Man. Sci., Vol. 10, No. 2, 1964.

Frankel, R. J., Economic Evaluation of Water Quality, An Engineering-Economic Model for Water Quality Management, First Annual Report, College of Engineering, Univ. of California, Berkeley, SERL Report No. 65-3, January 1965.

Garfinkel, R. S. and G. L. Nemhauser, Integer Programming, Wiley, New York, N.Y., 1972.

Geoffrion, A. M., "Elements of Large-Scale Mathematical Programming", Man. Sci., Vol. 16, No. 11, 1970.

Geoffrion, A. M., "Duality in Nonlinear Programming: A Simplified Applications-Oriented Development", SIAM Review, Vol. 13, No. 1, 1971.

Graves, G., Extensions of Mathematical Programming for Regional Water Quality Management, U.S. Environmental Protection Agency, April 1972.

Graves, G. W., A. B. Whinston, and G. B. Hatfield, Mathematical Programming for Regional Water Quality Management, FWQA, Department of Interior, 1970.

Graves, G., D. Pingry, and A. Whinston, "Applications of a Large Scale Nonlinear Programming Problem to Pollution Control", AFIPS Conf. Proc., Vol. 39, 1971.

Greenberg, H. J., Lagrangian Duality Gaps: Their Source and Resolution, Technical Report CP-69005, Southern Methodist Univ., Dallas, Texas, 1969.

Greenberg, H. J., "The Generalized Penalty-Function/Surrogate Model", Opns. Res., Vol. 21, No. 1, 1973.

Greenberg, H. J. and T. C. Robbins, Finding Everett's Lagrange Multipliers by Generalized Linear Programming, Parts I-III, Technical Report CP-700008, Southern Methodist Univ., Dallas, Texas, 1972.

Gomory, R. E., "Large and Nonconvex Problems in Linear Programming", Proc. 15th Symp. in Appl. Math., Amer. Math. Society, 1963.

Gould, F. J., "Extensions of Lagrange Multipliers in Nonlinear Programming", SIAM J. on Appl. Math., Vol. 17, No. 6, November 1969.

Haimes, Y. Y., J. Foley, and W. Yu, "Computational Results for Water Pollution Taxation Using Multilevel Approach", Water Resources Bulletin, Vol. 8, No. 4, 1972.

Haimes, Y. Y., M. A. Kaplan, and M. A. Husar, Jr., "A Multilevel Approach to Determining Optimal Taxation for the Abatement of Water Pollution", Water Resources Research, Vol. 8, No. 4, 1972.

Hass, J. E., "Optimal Taxing for the Abatement of Water Pollution", Water Resources Research, Vol. 6, No. 2, 1970.

Heidari, M., V. T. Chow, and D. D. Meredith, Water Resources Systems Analysis by Discrete Differential Dynamic Programming, Department of Civil Engineering, Univ. of Illinois at Urbana-Champaign, January 1971.

Joeres, E. F., J. Dressler, C. C. Cho, and C. H. Falkner, "Planning Methodology for the Design of Regional Waste-water Treatment Systems", to appear in Water Resources Research, 1974.

Karlin, S., Mathematical Methods and Theory in Games, Programming and Economics, Vol. 1, Addison Wesley, Reading, Mass., 1959.

Kelley, J. E., "The Cutting Plane Method for Solving Convex Programs", SIAM J. on Appl. Math., Vol. 8, No. 4, 1960.

Kneese, A. V., The Economics of Regional Water Quality Management, Johns Hopkins Press, Baltimore, Md., 1964.

Lasdon, L. S., Optimization Theory for Large Systems, Macmillan Co., New York, N.Y., 1970.

Leopold, L. B. and T. Maddock, The Hydraulic Geometry of Stream Channels and Some Physiographic Implications, Geological Survey Professional Paper 252, Washington, D.C., 1953.

Liebman, J. C., Optimal Allocation of Stream Dissolved Oxygen Resources," Water Resources Center, Cornell Univ., Ithaca, New York, 1965.

Loucks, D. P., C. S. ReVelle, and W. R. Lynn, "Linear Programming Models for Water Pollution Control", Man. Sci., Vol. 14, No. 4, 1967.

Luenberger, D. G., Optimization by Vector Space Methods, John Wiley and Sons, Inc., New York, N.Y., 1969.

Mangasarian, O. L., Nonlinear Programming, McGraw-Hill, New York, N.Y., 1969.

McConagha, D. L. and A. O. Converse, "Design and Cost Allocation Algorithm for Waste Treatment Systems", Jour. Water Poll. Control Fed., Vol. 45, No. 12, December 1973.

Meier, P. M., "A Branch-and-Bound Algorithm for Regional Water Quality Management", 14th Congress of the International Association for Hydraulic Research, 1971.

Nemhauser, G. L., Introduction to Dynamic Programming, John Wiley and Sons, Inc., New York, N.Y., 1966.

Nemhauser, G. L. and W. B. Widhelm, "A Modified Linear Program for Columnar Methods in Mathematical Programming", Opns. Res., Vol. 19, No. 4, 1971.

O'Connor, D. J., "Oxygen Balance of an Estuary", Jour. San. Eng. Div., Proc. ASCE, Vol. 86, No. SA3, May 1960.

O'Neill, R. P., "Some Computational Results on Column Dropping in Nonlinear Programming", Paper presented at the 44th National ORSA Meeting, San Diego, Cal., 1973.

Pingry, D. E. and A. B. Whinston, "Multigoal Water Quality Planning Model", Jour. Env. Eng. Div., Proc. ASCE, Vol. 99, No. EE6, December 1973.

Pingry, D. E. and A. B. Whinston, "Application of Multigoal Water Quality Planning Model", Jour. Env. Eng. Div., Proc. ASCE, Vol. 100, No. EE1, February 1974.

Rani, O. and R. N. Kaul, "Duality Theorems for a Class of Nonconvex Programming Problems", J. Opt. Theory Appl., Vol. 11, No. 3, 1973.

Rosen, J. B., "The Gradient Projection Method for Nonlinear Programming", SIAM J. Appl. Math., Vol. 8, No. 1, 1960.

Smith, R., "Cost-Effectiveness Task Force - Economics of Consolidating Sewage Treatment Plants by Means of Interceptor Sewers and Force Mains" (U.S. Government Memorandum, Advanced Waste Treatment Laboratory, Environmental Protection Agency, March 10, 1971).

Streeter, H. W. and E. B. Phelps, A Study of the Pollution and Natural Purification of the Ohio River, Public Health Bull. 146, U.S. Public Health Service, 1925.

Thomann, R. V., Systems Analysis and Water Quality Management, Environmental Research and Applications, Inc., New York, N.Y., 1972.

Tsivoglou, E. C. and J. R. Wallace, Characterization of Stream Reaeration Capacity, U.S. Government Printing Office, Washington, D.C., October 1972.

Uzawa, H. in Arrow, K. J., L. Hurwicz, and H. Uzawa, Studies in Linear and Nonlinear Programming, Stanford Univ. Press, California, 1958.

Wanielista, M. P. and C. S. Bauer, "Centralization of Waste Treatment Facilities", Jour. Water Poll. Control Fed., Vol. 44, No. 12, December 1972.

Whitlatch, Jr., E. E., Optimal Siting of Regional Wastewater Treatment Plants (Unpublished Ph.D. Dissertation), The Johns Hopkins Univ., Baltimore, Md., February 1973.

Whittle, P., Optimization under Constraints, Wiley-Interscience, London, 1971.

Yao, K. M., "Regionalization and Water Quality Management", Jour. Water Poll. Control Fed., Vol. 45, No. 3, March 1973.

Zangwill, W. I., Nonlinear Programming: A Unified Approach, Prentice-Hall, Englewood Cliffs, New Jersey, 1969.

Zoutendijk, G., Methods of Feasible Directions, American Elsevier Publishing Co., Inc., New York, N.Y., 1960.