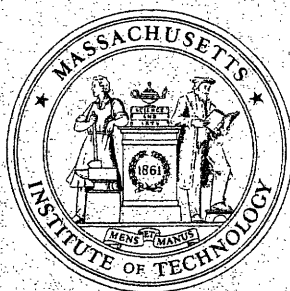


OPERATIONS RESEARCH CENTER

working paper



MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PLANNING INVESTMENTS IN WATER RESOURCES BY
MIXED INTEGER PROGRAMMING:
The Vardar-Axios River Basin

by

Dorothy P. Elliott

OR 072-78

March 1978

Research supported (in part) by the U.S. Army Research Office (Durham)
under Contract DAAG29-76-C-0064.

ABSTRACT

A mixed integer programming model for planning water resources investments is presented. The model is a sequencing model applied to the Vardar-Axios river basin in Yugoslavia and Greece. The structure of the model is outlined, and computational experience is described. The size of the model presented some difficulties, which are discussed along with the results to date. The experience with this model points to areas where further research is needed.

I. Introduction

Many problems in the area of water resources can be modeled as mathematical programming problems. In particular, river basin planning can be facilitated through the use of optimization techniques. Such problems may be solved using tools such as linear programming, network optimization, and mixed integer programming, depending on the nature of the problem. Numerous optimization models have been formulated and tested on specific problems, including single project models, where conflicting uses of the project must be balanced to produce the maximum benefit¹, multiple project models, where flow along the river must be considered², and entire basin models, where the needs of a large region must be taken into account³. Mixed integer programming in particular has been found useful in problems involving scheduling, sequencing, returns to scale, and plant location. One example of such a problem is a sequencing model of the Vardar-Axios river basin (located in Yugoslavia and Greece). The purpose of this paper is to examine the problem and how it was solved, and to see how the techniques used can be generalized and applied to other problems.

II. Statement of Model and Discussion

The model is adapted from the one used by Cohon, Facet, Marks, and Haan⁴ in their work involving project selection and sequencing on the Rio Colorado in Argentina. The project data is provided by a simulation model of the basin. Data required includes the locations, costs, benefits, and possible sizes of the projects, plus flow data for the river in its current state. The task is to determine which projects should be built and in which time period, and, in some cases, which of the proposed sizes. There are budget constraints, water availability limitations, and logical constraints. The objective is to maximize total benefits.

The decision variables of the model are of two types; flow variables and project variables. The river basin is marked with checkpoints where

flow data is available. (Many of these checkpoints are located at proposed sites of projects or immediately downstream.) The flow variables, which are continuous, give the flow at each checkpoint for each time period for a given feasible sequence of projects. (Such a feasible sequence of projects, representing a feasible solution to the integer part of the problem, will be referred to hereafter as a configuration.) The project variables, which are zero-one variables, indicate whether or not a given project is constructed during a given time period. The problem allows for a planning horizon of more than one time period. Each time period is assumed to be long enough to allow for the completion of projects started at the beginning of the time period. The limitation on the number of periods in the planning horizon is that both the number of columns and the number of rows grows as the number of time periods, drastically increasing the computational burden.

The constraints fall into four classes: continuity, conditional, construction, and budget constraints. The continuity constraints balance the flow from one checkpoint (or site) to the next to make sure that the configuration found by the model is consistent with the existing flow conditions in the river basin. These are equality constraints (the only such constraints in the model) and there is one for each site in each time period. The construction constraints simply ensure that a project is built no more than once and in no more than one of the proposed sizes, if more than one possible size is given by the data. There is one of these constraints for each proposed project. The conditional constraints enforce limitations such as precedence relationships (a reservoir must be present for a power plant to be built, for example). The budget constraints guarantee that for each time period, total capital costs for project construction will not exceed a given amount. There is one such constraint for each time period. In the original version of the model there were also population constraints to make sure that the number of people needed to develop the irrigation sites did not exceed the number of people available. (These constraints were later deleted.)

The objective function to be maximized is the sum of the benefits of the construction of a project minus the capital and operation and maintenance costs of building that project. The benefit figures include

benefits from agriculture, power production, flood control, and recreational use, all taken from the data for the simulation model⁵, from which the sizes and locations of the projects were also obtained.

The detailed formulation follows:

Original model:

$$\text{Budget: } \sum_{k=1}^K \sum_{s=1}^S \text{bud}_{si}^k \cdot \text{proj}_{si}^k \leq B_i \quad \text{for all } i$$

$$\text{Population: } \sum_{s=1}^S \text{pop}_{si} \text{IRR}_{si} \leq P_i \quad \text{for all } i$$

$$\text{Construction: } \sum_{i=1}^T \text{proj}_{si}^k \leq 1 \quad \text{for all } s, k$$

$$\text{Conditional: } \text{PPV}_{si} \leq \sum_{j=1}^i \text{RES}_{sj} \quad \text{for all } s, i$$

$$Y_{si} - Z_{si} = 0 \quad \text{for all } s, i$$

Continuity: (general form)

$$D_{si} - D_{s-1,i} - \sum_{j=1}^{i-1} \bar{Y}_s \text{RES}_{sj} + \bar{c}_s \text{RES}_{si} + \sum_{j=1}^i E_s \text{IRR}_{sj} = F_s - F_{s-1} \\ \text{for all } s, i$$

(The form of continuity constraints depends on what projects are proposed at a site. At sites where water returns from an irrigation site there is also a $(1-\mu)E_{s^*}$ term present, s^* indicating location of irrigation site outflow. In addition, each continuity row at a site where a branch flows in has a -1 in the column corresponding to the end site of the inflowing branch.)

Objective function:

$$\max z = \sum_{s=1}^S \sum_{i=1}^T [(\beta_{si}^{\text{PPV}} - \alpha_{si}^{\text{PPV}}) \text{PPV}_{si} + (\beta_{si}^{\text{EXP}} - \alpha_{si}^{\text{EXP}}) \text{EXP}_{si} + \\ (\beta_{si}^{\text{IRR}} - \alpha_{si}^{\text{IRR}}) \text{IRR}_{si} + (\beta_{si}^{\text{Y}} - \alpha_{si}^{\text{Y}}) Y_{si}] - \sum_{s=1}^S \sum_{i=1}^T \alpha_{si} \text{RES}_{si}$$

Decision variables:

D_{si} release from reservoir at site s during period i , or flow at that point if reservoir has not been constructed.

EXP_{si} construction or not of export to another basin from s during i

IRR_{si}	construction or not of irrigation site at s during i
PPV_{si}	construction or not of power plant at s during i
RES_{si}	construction or not of reservoir at s during i
Y_{si}	construction or not of diversion(in) at s during i
Z_{si}	construction or not of diversion(out) at s during i

Data variables

B_i	budget available in time period i
\bar{c}_s	consumptive use of reservoir at s during construction and filling
bud_{si}^k	capital cost of project of type k at s constructed during i
E_s	diversion for irrigation at s
F_s	streamflow measured at s
P_i	total number of persons available to develop irrigation areas in period i
\bar{Y}_s	yield from a reservoir at s
α_{si}^k	value of capital and operation and maintenance costs for a project of type k constructed during i at s
β_{si}^k	value of time stream of benefits resulting from the construction of a project of type k at s during i
μ	irrigation loss coefficient

The model was applied to the Vardar-Axios river basin in two stages. Some revisions were required for the second stage. Population constraints were dropped since it was believed that they would never become binding in the river basin. Diversions, which in the first stage were treated as two projects, one flowing into the basin and the other out, were merged into one. Also, some projects were proposed in several sizes so a new index is added to the project decision variables. Thus the budget and construction constraint left-hand-sides are now also summed over z, as is the objective function. Some conditional constraints were added to make certain irrigation sites dependent on some diversions. Also, a new set of projects was added to represent municipal and industrial water supply projects, which behave in the model somewhat like irrigation sites since a fraction of the water diverted is not returned to the river.

For the first stage a simplified schematic of the basin and proposed projects was used as the data for the model. This first stage included four time periods of five years each, for a planning horizon of 20 years. Seventy-one projects and 113 sites were included, yielding a matrix that was 608 x 736. The breakdown of projects was as follows:

Reservoirs	18
Irrigation sites	24
Power plants	13
Diversions(in)	6
Diversions(out)	7
Exports	<u>3</u>
	71

The second stage was more detailed and resulted in a significantly larger problem. The number of proposed projects became 111. (Some of these had been combined into single projects in the first stage.) In addition, several sizes were proposed for a number of projects, resulting in not 4 but 8 or even 12 integer variables corresponding to some of the projects. This resulted in an increase in the number of integer variables from 284 to 628. The number of sites increased to 194. The size of the matrix became 1002 x 1404. The structure of the planning horizon was not changed.

The water flow data chosen was for July, the driest month in the basin, since if a configuration satisfied water limitations in that month, it definitely would during the rest of the year. The breakdown of projects in the second stage was:

Reservoirs with power plants	10
Reservoirs	20
Power plants	2
Diversions	8
Irrigation sites	47
Municipal & industrial supply	<u>24</u>
	111

III. Computational Experience and Results

The model was set up using a matrix generating language and solved using an interactive branch-and-bound procedure on an IBM 370/168. The initial optimal linear programming solution had an objective function

value of 355.21 (in millions of dollars) as compared to the optimal integer programming solution of 217 (a gap of about 40%). This gap is largely accounted for by the fact that many irrigation sites were partially built in the linear programming relaxation, contributing significantly to the objective function, but had to be eliminated in the integer programming solution because if completely built they would be too large for the amount of water available at the site. This suggested that perhaps alternate sizes should be considered, or that the data might be inaccurate, or both. A more detailed model was called for. Further evidence of this need was the fact that the integer programming solution included no reservoirs and no power plants. This called into question the benefit figures, the sizing of the projects, and some of the precedence relationships. All these considerations suggested a second pass at the river basin, using less aggregate data than in the first stage.

For the second stage more detailed and up-to-date information was obtained. As noted above, changes in the formulation resulted in a larger model, which was consequently much more difficult to handle.

The optimal value for the linear programming relaxation was 29595.123 millions of new dinars (Yugoslavian currency). This corresponds roughly to 1500 million dollars. The large difference from the results of the first stage is due to revisions in benefits estimates as well as an increased number of possible projects. Many projects in this solution were fractional, falling into two classes: some projects were to be completely built, but over all four time periods, and others would be only partially constructed over the entire planning horizon. (This distinction was useful in planning the branch-and-bound strategy.) In this second stage as in the first, the budget was more fully utilized toward the end of the planning horizon. This is not surprising since some expensive projects depended on the prior construction of relatively cheaper ones.

The mixed integer program was attacked using branch-and-bound, and several attempts were made before a productive branching strategy was found. Simply finding a feasible integer solution (other than the trivial one of no project construction) turned out to be an enormous task. The first goal was therefore to find a limited feasible solution: one in which splits in projects over time were allowed, so that the only variables

treated as integer were the slack variables on the construction constraints, which were of the form $x_{11} + x_{12} + x_{13} + x_{14} \leq 1$. The branching strategy used in finding this solution was to force as many projects as possible to be built until the water constraints are violated, and then force the marginal one to zero and proceed. The intermediate solution found in this way had some projects built in stages over the planning horizon but all projects were either completely built or not at all. The objective function value for this solution was 27274.269, about 8% less than the solution to the linear programming relaxation. The budget was at its upper bound for the third and fourth periods.

The next stage in the solution would be to shift the projects around within the time periods, keeping the same solution to the intermediate problem but looking at various feasible solutions stemming from that solution. Of course, there is no guarantee that the optimal solution to the whole problem is a descendant of that intermediate solution. It may well be a descendant of another feasible solution to the intermediate problem. However, finding an integer solution as a descendant of that solution would at least provide a useful lower bound to the problem and indicate a range in which the optimal solution is to be found. Following such a procedure with each of the feasible intermediate solutions would ultimately have to lead to the optimal integer solution. (Unfortunately, this is more easily said than done with a problem of this size. Storage requirements and costs of the iterations needed in branching can become prohibitive long before a solution is reached.)

Once an integer solution is reached, much work remains to be done. Some sensitivity analysis is desirable, especially since budget figures and cost and benefit estimates can be uncertain. This sensitivity analysis is a difficult problem in mixed integer programming. Furthermore, the model (although large) left out some aspects that perhaps should be included, such as power production requirements and agricultural water needs. The objective function also may be oversimplified, since some objectives, like power and recreational use, which were simply translated into dollar values, may be conflicting and involve tradeoffs. This would point to

the need for multiobjective analysis, another difficult area in mixed integer programming.

IV. Future Research

The experience gained in solving this problem indicates first the difficulty of solving such a large-scale problem by branch-and-bound. The process is time-consuming and costly. Even using an interactive branch-and-bound routine, which allowed choice of branches by one thoroughly familiar with the structure of the problem rather than simply random branching, the problem was slow to solve. Furthermore, the storage requirements were prodigious, even for the extremely sparse matrix in this problem. For further work on this problem and others of this type it seems that development of a decomposition strategy would be useful. But, more interestingly, it also seems that thorough treatment of the problem in the areas of sensitivity analysis and multiple objectives will require some ground-breaking work in mixed integer programming techniques. Much research remains to be done here.

References

1. Thomas, Harold A., Jr. and Roger Revelle, "On the Efficient Use of High Aswan Dam for Hydropower and Irrigation," Management Science, Vol. 12, 1966, B296-B311.
2. Roefs, Theodore G., and Lawrence D. Bodin, "Multireservoir Operation Studies," Water Resources Research, Vol. 6, 1970, 410-420.
3. Hufschmidt, Maynard F., "Field Level Planning of Water Resource Systems," Water Resources Research, Vol. 1, 1965, 147-163.
4. Cohon, J.L., T.B. Facet, A.H. Haan, and D.H. Marks, "Mathematical Programming Models and Methodological Approaches for River Basin Planning," Technical Report, Ralph M. Parsons Lab for Water Resources and Hydrodynamics, M.I.T., Cambridge, Mass., 1973.
5. Strzepek, Kenneth M., "Analysis of Multipurpose River Basin Planning: Guidelines for Simulation Modelling," S.M. Thesis, Department of Civil Engineering, M.I.T., Cambridge, Mass., June 1977.