SOME ISSUES AFFECTING OPTIMIZATION MODELS IN WATER RESOURCE PLANNING

Eric F. Wood

August 1974 WP-74-42

Some Issues Affecting Optimization Models

in Water Resource Planning

Eric F. Wood

Abstract

This paper draws heavily upon my experiences with the MIT-Argentina Study. The purpose of the paper is to bring forth some of the issues that affect the role of optimization models within a water resource planning methodology.

Reasonable analysis goes from a problem--a region or part of a river with certain characteristics--to model formulation. It is the problem that puts requirements upon the model and in water resources the stochasticity of the real system has caused difficulties in the modelling step.

Since optimization plays such an important role in the planning methodology, this paper considers current approaches to handling stochastic elements within optimization. The applicability to a real system like the Tisza is discussed and a new alternative approach is offered.

1. Introduction

The analysis of river basin systems is complicated by two considerations. First, the river basin is a complex physical system which is very difficult to model realistically. Second, decision makers must try to choose from among many alternatives. There exists in most river systems an extremely large combination of possible location and sizes--all of which must be considered (at least implicitly) within the analysis. To complicate matters, economic, political and social aspects of river basin development affect the decision making process and pose additional modelling problems.

These two considerations have lead planners to consider two types of models. One type is simulation. Simulation has the advantage of being able to model the physical system extremely well. Unfortunately, the decision variables, such as reservoir sizes, targets, irrigation area etc., must be set a priori to running the model. The simulation model then gives a point inside (or maybe on) the feasible region--there is no indication whether a better combination of decision variables exist. If the simulation model could be run an extremely large number of times then maybe the set of efficient solution could be determined. Usually this is not practical.

The second type of model is an optimization model. These models implicitly consider all combinations of the proposed system and will choose an optimal configuration based on some specified objective function. This optimal configuration will satisfy physical, economic, political, and social feasibility as represented in the constraint set. However, to model the physical system within an optimization model framework requires a large number of assumptions that may lead to an unrealistic representation of the physical system. The degree of realism depends upon the physical system being modelled--in particular upon the stochastic and non-linearity aspects of the system. Stochastic considerations increase the size of the optimization model by requiring more constraints, non-linearity vastly increase computation time or may make solutions impossible. Spofford [1, 2] has considered optimizing models where a non-linear set of constraints--to represent a physical system--have been included within a larger linear programming model. The solution procedures for these models still require additional research, and probably involves nesting a simulation type model (to handle the physical system) within the optimization framework. Such solution techniques still require additional research to make them operational in the real world.

The solution to the correct mix of models, optimization and simulation, depends upon the problem at hand. Planning methodologies vary from each problem and depend upon the time and resource constraints for analysis, computer budget, and the questions that the models should address. Much of the art of successful analysis is contained within the planning methodology--and how the methodology evolves and adapts throughout a study depending upon the models developed and the results obtained. The planning methodology is a dynamic framework for analyzing a planning problem.

Optimization models often play a significant role in planning because many optimization questions arise. Furthermore, optimization models are usually the first step in planning and are often used to screen out non-contenders from the set of feasible combinations of development. The nature of the optimization model should depend upon the question it tries to address. In the Tisza, the stochastic nature of the system is one issue that should be considered in the design of an optimization model. Since most facilities must still be constructed, the investment question (as opposed to the operation of the system) must also be considered. This paper presents some optimization models[1] which have appeared in the literature and discusses the effectiveness of the models in representing an actual system. The Tisza River will always be kept in mind.

Deterministic Models

Consider a model of the form

$$\max \underline{B}^T \underline{x} - \underline{C}^T \underline{x} \quad \text{Maximize Net Benefits}$$

such that

$$\begin{array}{ll} \underline{Ax} \leq \underline{b} & \text{Continuity Constraints} \\ & \text{Technological Constraints} \\ \underline{x} \geq 0 & \text{Policy Constraints} \end{array} \qquad (1)$$

---

[1]Our discussion will be constrained to linear pro- (or L.P. type) models as opposed to considering dynamic programming. In D.P. there exists dimensionality problems when the water resource investment problem is considered; and that problem is the main focus of our discussion.

Such a qualitative representation of the optimization model will fit virtually any model found in the literature. The only differences arise in the degree of detail and complexity within the constraint set, especially with respect to the physical system. For example Poblete [3] investigated the sensitivity of solutions to the number of time periods, and only included continuity constraints. Loucks [4] presented some fairly large formulation with relatively detailed constraint sets-- similarly with the MIT-Argentina Project which had a deterministic mixed-integer L.P. of 629 constraints and 658 continuous variables to represent a 38 site, 3 season system. The degree of detail is an important a priori decision.

Another important decision is whether to use a deterministic model or not. If the river does not exhibit major streamflow variation from year to year then it may be best to investigate the stochasticity in a simulation model. Furthermore, if the yield from the reservoirs is not too large and if demands are fairly constant from year to year then a deterministic model may do a fairly good job at finding some designs for further study.

In the Tisza, there are two problems. Flood control and water supply. Flood control clearly cannot be handled within a deterministic model, as usually formulated. When the events of interest are occurring once every 10, 20, or 100 years it is unrealistic to construct a model that uses the "average" yearly flood. For water supply the issues are not so obvious. Most of the demands are for irrigation but on a supplementary basis. If it is wet, irrigation may not be required at all, but if it is dry the demands may be significant. From talks with the Hungarians, it is not clear if over-year storage reservoirs would be required. That is, there exists enough water on a yearly basis to fulfill demands; the problem is that the water comes at the wrong time of the year. An n season (2 or 3, let say) deterministic model could be formulated which for average demands would determine the sizes of the reservoirs. But it is the short term operation (this year's demand, next year's etc.) that will affect the performance of the system, not the long term averages. There will exist both short term losses and opportunity losses from incorrectly sizing the elements of the system. Since the Tisza's demands are highly stochastic it is questionable whether a deterministic model can provide the necessary planning information. If optimization models are to be applied to the Tisza, then stochastic considerations must be included.

Two-Stage Linear Programming under Uncertainty

Consider a situation where a decision must be made in the first time period which will be affected by an uncertain event occurring in the second period. At the first time period the probability of the events is known but which event takes place is unknown. This is the general framework of the two-stage linear program. In more precise notation:

$$\text{Min } \underline{c}^T \underline{x} + \underline{d}^T \underline{y}$$

such that

$$\underline{Ax} = \underline{b}_1$$

$$\underline{Bx} \pm \underline{Dy} = \underline{b}_2 \quad , \qquad x \leq 0 \quad , \qquad y \geq 0 \quad . \tag{2}$$

The first constraint set contains only the deterministic elements and represents the first stage problem. The second constraint set contains random elements (in both $\underline{B}$ and $\underline{b}_2$) with known probabilities. In water resource problems, the random variables are usually confined to $\underline{b}_2$. This general form was first analyzed by Dantzig and Ferguson [5].

In a water resource context, the reservoir capacity and target level are picked in the first stage and then a random inflow and demand are observed in the second stage.

The objective function has additional costs $\underline{d}^T \underline{y}$ when $\underline{y}$ is non-zero. These costs are short term penalty functions and were introduced by Dantzig [6]. A closer analysis of (2) reveals that if $\underline{x}^*$ is a solution to the first stage and $\underline{y}^*$ is chosen to satisfy the second constraint set, then for $\underline{x}^*$ to be feasible in the program it must be feasible in the second stage regardless of the outcome of the random process in $\underline{b}_2$. While $\underline{x}^*$ gives a feasible solution it may not be a very reasonable solution due to its conservative nature. This can be adjusted by weighing the outcomes by their probability of occurrence. This leads to the use of expectation within the objective function. The objective function can now be written as

$$\text{Min } \underline{c}^T\underline{x} + E_{\underline{b}_2}\left[\underline{d}^T\underline{y}\right]$$

where

$$E_{\underline{b}_2}\left[\underline{d}^T\underline{y}\right] = \sum_k dy^k p_k$$

for the k possible outcomes that can occur, each with probability p_k. Thus the objective function indicates that an $\underline{x}$ is to be found that minimizes the first-stage costs, plus the expected smallest penalty cost which is, itself, a function of the first stage decision variables.

Water resource applications of the two-stage technique have been performed by Dorfman [7], Loucks [8], and Håan [9]. Loucks formulation was quite complex. He considered inflows as a lag-one Markov process and segmented each reservoir into three stages. He was concerned with setting targets for storage at the beginning of the next season given the occurrence during the present season. Unfortunately, his procedure cannot be applied to a second reservoir in series because its inflow depends upon the yet to be determined storage of upstream reservoirs. One would have to go to a n-stage stochastic model, with some sort of transfer constraints between the stages. Such a formulation would be extremely large, since McBean [10] estimated that a two-stage formulation of a water resource problem would require four to five times more constraints than would the deterministic formulation. Yet the two-stage formulation has the advantages that it explicitly considers the economic affects of failing to meet demands. It is upon these affects that decisions are based.

In the Tisza study it is not clear whether such a formulation would prove useful.

To be applicable to a river system such as the Tisza the following considerations would have to be overcome

1. The size of the constraint set may make it computationally difficult or impossible to solve.
2. The problems with multi-reservoirs, especially those in series, would have to be solved. This problem may be handled initially by 'lumping' such reservoirs together. If this is not feasible then a multi-stage formulation would be required.
3. Two-stage models address most effectively the short term operation--i.e. yearly basis. Given some information about last seasons inflows and the state of the reservoir it will set a new target. But long-range development depends upon long-range targets. Such long-range targets must be met with a high level of reliability if a development is to be economically, politically and socially successful.

Chance-Constrained Programming

An alternative model which considers the stochasticity of the system is chance-constraint programming. In the two-stage formulation the first-stage decision variables were permitted to violate the second-stage random events by incurring a penalty. In chance-constrained programming, the decision variables may violate certain 'random' constraints but only by a pre-determined probability. The formulation is:

$$\text{Min } \underline{C}^T\underline{x}$$

such that

$$P[\underline{Qx} \geq \underline{b}] \geq \underline{\alpha} \quad , \qquad \underline{x} \geq 0 \; , \quad 0 \leq \underline{\alpha} \leq 1 \; . \tag{4}$$

The constraint set $\underline{Qx} \geq \underline{b}$ can only be violated with a probability of $1 - \alpha$; thus $\underline{\alpha}$ represents the risk level that is allowed in the system. Chance constrained programming was first introduced by Charnes and Cooper |11|. The formulation in (4) must be converted into a solveable form which leads to a deterministic equivalent formulation. This can be done in the following manner. If the i^{th} constraint is

$$P[\sum_{j=1}^{n} q_{ij} \, x_i \geq b_i] \geq \alpha_i \tag{5}$$

and for b_i there exists a probability density function f_{b_i}

then a b_i^* can be found such that:

$$b_i^* = F_{b_i}(1 - \alpha) \tag{6}$$

where $F_{b_i}(\cdot)$ is the cumulative density function for b_i. Equation (5) can be replaced by:

$$\sum_{j=1}^{n} q_{ij}x \geq b_i^* \tag{7}$$

The chance-constrained programming has received more attention in the water resources literature than has two-stage formulation. This is probably because the formulation is of the same size as a deterministic formulation and insights into a problem by its application can be quite significant. Chance constrained programming has been used by Smith [12] for irrigation designs, Revelle et al [13] for reservoir designs and Hermann and Perkins [14] for pumping facility design. Within the chance-constrained formulation it is not really possible to evaluate the economic consequences of not meeting demands. Planning must deal with the economic implications of alternative designs, so the absence of short-run loss functions may limit the effectiveness of the formulation. Furthermore, the level of risk, α, is set a priori within the model, but the risk level itself is a decision variable. A high value of α implies a very reliable system, low water availability and low downstream development. It leads to conservative development without explicitly evaluating the economic consequences--thus the chance constrained formulation may not lead to the best alternative. A strategy of increasing reservoir capacity (to increase system reliability) while maintaining downstream development level is not available from the chance-constrained formulation.

In the Tisza study, chance-constrained programming may be able to provide meaningful insights into the stochastic nature of the demands and supply. The shortcomings in the above discussion provide a warning to how the results should be interpreted.

An Alternative Approach

During the MIT-Argentina Study the planning methodology that was developed centered around a large deterministic mathematical program and a simulation model, which handled the stochastic elements.

As envisioned, the optimization model would generate an initial configuration which would be passed on to the simulation model where a 'redesign' would take place and a more reliable system would emerge.

This methodology was not completely successful. The optimization model and the simulation model are constructed differently. Some things can be handled well within one framework and not in the other. The deterministic aspects of the optimization model put a relatively large burden on the simulation model which was hampered by the lack of a systematic search technique as it tried to redesign on the response surface. McBean |10| developed a search technique after the conclusion of the analysis phase of the project. McBean indicated that a stochastic screening model can significantly reduce the computational burden of the simulation model since many of the stochastic elements are taken into consideration. Since for large systems stochastic screening models are too burdensome, McBean suggested to use a deterministic screening model then a stochastic optimization model on part of the river basin in conjunction with the use of the simulation model.

An alternative procedure could be the following:

1. Solve the large scale deterministic optimization model and obtain initial values of the capacity and targets.

2. Formulate a network flow model[1] for the river basin using the values from the first step as the given capacities and targets. Stochastic inflows and demand can be generated using a synthetic generator. This generation could consider correlation structures in

---

[1] The network flow model does not address the investment problem very effectively because it is a 'scenario' formulation. That is, it only considers the operation of the pre-set configuration. The formulation has the advantage that it can be solved very quickly and relatively cheaply. Both of these 'properties' can be utilized in my proposed two-step procedure.

time and space and between demands and inflows. The output from the generation process would then be 'fed' into the network flow model and run for many periods. In essence, it would be run for 1 period and the results stored and used for initial values for running the next period etc. For a reasonably large system, it would not be computationally burdensome to run the model for 100 periods. The interesting output from this exercise are the shadow prices on the step one decision variables. For some periods there will be excess water and for some periods there will be insufficient water, giving rise to a distribution of shadow prices for each of the step one decision variables. Therefore, a 'redesign' can take place on those elements of system that have the highest expected shadow prices (or on some other appropriate criteria) and step two can be re-run. This iterative fashion results in a good initial configuration. Whether simulation would be required is not clear, it would depend upon the particular problem and the questions that it is suppose to answer.

This two-step procedure is similar to the two stage formulation in that it is a deterministic formulation followed by a stochastic analysis. It has the advantage that large systems with many time periods can be handled with reasonable computer resources. The step-two formulation is run in a manner similar to a simulation model except that there are implicit search techniques built in. The economic consequences from deficits are explicitly utilized, so that the two-step formulation should be more insightful than a chance-constrained formulation.

How this would work in a real application to a large system is not known. It has never been applied. I do feel though that the Tisza River may be an appropriate application--at least the water supply analysis. The large variances in supply and demand would suggest that stochasticity should be considered in the optimization step.

Conclusions

Drawing from my experiences with the MIT-Argentina Study, I have tried to bring forth many of the issues that affect optimization in river basin planning. Planning methodologies are not unique, but optimization often takes a significant role--usually as the first step. If the stochastic considerations will affect the system evaluation, then it should be included, if possible, in the optimization step. This paper presents some of the current approaches of stochastic optimization that may be relevant to water resource investment planning and suggests a new approach that may not have the difficulties of the current procedures.

References

[1] Spofford, W. "Total Environmental Quality Management Models in R.A. Deininger, ed. Models for Environmental Pollution Control, Ann Arbor Science Publ., Inc., Ann Arbor, Mich., 1973, Ch.19.

[2] Spofford, W. "An Approach to the Flood Protection Investment Problem," unpublished notes, IIASA, August 20, 1974.

[3] Poblete, J.A. and McLaughlin, R. "Time Periods and Parameters in Mathematical Programming for River Basins," T.R. No. 128, Ralph M. Parsons Laboratory, Department of Civil Engineering, MIT Cambridge, Mass., September 1970.

[4] Loucks, D.P. "Stochastic Methods for Analyzing River Basin Systems," T.R. No. 16, Cornell University Water Resources and Marine Science Center, Ithaca, N.Y., August 1969.

[5] Dantzig, G. and Ferguson, A.R. "The Allocation of Aircraft to Routes: An Example of Linear Programming Under Uncertain Demand," Management Science 3, 1956.

[6] Dantzig G. "Linear Programming Under Uncertainty," Management Science, 1, 1955.

[7] Dorfman, R. "Mathematical Models: The Multi-Structure Approach" in Maass, A et al. Design of Water Resource Systems, Howard University Press, 1962.

[8] Loucks, D.P. "Stochastic Methods for Analyzing River Basin Systems," T.R. 6, Cornell University Water Resources and Marine Sciences Center, Ithaca, N.Y., August 1969.

[9] Håan, A.H. A Screening Model for Water Resource Development: Stochastic Considerations unpublished S.M. thesis, Department of Civil Engineering, MIT Cambridge, Mass., February 1972.

[10] McBean, E. "A Marginal Analysis-System Simulation Technique for Multi-Objective Plan Formulation in Water Resources," Phd. Thesis, Dept. of Civil Engineering, MIT, Cambridge, Mass., January 1973.

[11] Charnes, A. and Cooper W. "Chance-constrained Programming," Management Science, 1, 1955.

[12] Smith, D.V. Stochastic Irrigation Planning Models Center for Population Studies, Harvard University, Cambridge, Mass., January 1970.

[13] Ravelle, C., Jones, E. and Kirby, W. "The Linear Decision Rule in Reservoir Management and Design. 1. Development of the Stochastic Model," Water Resources Research, Vol. 5, No. 4, August 1969, pp. 767-777

[14] Hermann, R.M. and Perkins F.E. "Stochastic Linear Models for Multi-Reservoir Systems," T.R. No. 130, Ralph M. Parsons Laboratory, Department of Civil Engineering, MIT, Cambridge, Mass., 1970.