Comp. & Maths. with Appls. Vol. 8, No. 4, pp. 313–318, 1982
Printed in Great Britain.

0097-4943/82/040313-06$03.00/0
Pergamon Press Ltd.

# THE SEQUENTIAL PROBLEM UNDER UNCERTAINTY

## THE DEVELOPMENT OF WATER SYSTEMS

WITOLD F. KRAJEWSKI†

Systems Research Institute, Polish Academy of Sciences, ul. Newelska 6, Warsaw, Poland

Communicated by E. Y. Rodin

(*Received November* 1981)

**Abstract**—The problem of setting up the optimal sequence for construction of water facilities in water systems is presented. The problem is formulated as a deterministic model where all economic parameters of the water plants and the entire system are assumed to be known. Then, uncertainty of those parameters has been taken into account and the appropriate model is described.

### INTRODUCTION

The development of water resource systems is very complex and must deal with a variety of scientific, technological, economical and social issues. This forces the necessity to analyze a significant number of variants and creates new problems, i.e. investments. One such problem which has been almost completely neglected until this time, is that of sequencing the construction of water facilities (including reservoirs, water supply treatment plants, canals, power plants, etc.) which have been determined to be feasible based on technical-economical criteria. This problem was first formulated by Butcher *et al.*[1] and was then developed by others such as Becker and Yeh[2], Erlenkotter[3, 4], Morin and Esobgue[5], Rinaldi *et al.*[6] and others.

In general, the problem consists of minimization (or maximization) of a certain function, $f(\xi)$, where $\xi = \{i, x_i\}$ is a set of pairs of numbers; the first is an object number, the second is the related date of the beginning of the object's realization:

$$\text{minimize (maximize)} \quad f(\xi)$$

$$\text{with constraints:} \quad g_j(\xi) \leqq b_j \qquad \text{for } j = \{1, 2, \ldots, n\}$$

where $g_j(\xi)$ are real functions and $b_j$ are constant coefficients.

Usually, $f(\xi)$ is assumed to be the sum of different types of costs, but this function is also known in other forms. Rinaldi *et al.*[6] minimize a water pollution indicator in the water system to determine the optimal sequence of water treatment plant installations. As constraints, the relations concerning demands are taken. Those demands can be expressed either in natural units (i.e. $m^3/s$ or mg $O_2/1$) or monetary terms. A very important group of constraints deals with the interdependency between particular objects in the system. Their existence affects significantly the sequence of plant realization and their distribution in time.

It should be taken into consideration that formulation and an attempt for solution of the sequential problem can give other advantages than just the optimal solution. Namely, it is possible that the feasible solution does not exist. In this situation, decision making concerning the initiation of system realization can lead to serious perturbances to the realization process.

Of course, in every system that has been realized heretofore, certain sequences have been set up. However, the choices were made on the basis of intuition or professional experience. It is easy to see that for $m$ objects to be ordered $m!$ solutions are possible. Therefore, when $m = 20$ we have $m! = 2.4 \times 10^{18}$. Choosing the best available solution forces the necessity of developing the objective tools with which to do it. It should be expected that setting up the optimal or quassi-optimal solution would save expenses due to the fact that water plants are usually very capital consuming.

†Present affiliation: Hydrologic Research Laboratory, OA/W23, National Weather Service, NOAA, 8060 13th Street, Room 530, Silver Spring, MD 20910, U.S.A.

Most of the investigators have formulated this problem as a dynamic programming problem eventhough it is very useful to describe it as a zero-one integer programming task. We will use this form to help us describe the sequential problem under uncertainty.

## THE SEQUENTIAL PROBLEM FORMULATION

Let's define $n$ as the number of years within the planning horizon and $m$ as the number of objects to be introduced. For each object, investments, maintenance costs, material means, etc. as well as different kinds of effects, are known for each year of its realization and activity.

The following optimization task can be formulated: fix date $x_i$ (for $i = \{1, 2, \ldots, m\}$), the beginning of realization of each of $m$ objects within the planning horizon, in such a way as to optimize the goal function and to get the planned effects without violating existing limits.

Let's assign the following function as a goal function:

$$Z = \sum_{i \in M} \sum_{t=0}^{n-x_i+1} a_{it}(I_{it} + K_{it}) \tag{1}$$

which is to be minimized.

Let the constraints be:

$$\sum_{i \in M} E_{ijp}(x_i) \geqq D_{jp} \qquad \begin{array}{l} \forall j \in N \\ \forall p \in P \end{array} \tag{2}$$

$$\sum_{i \in M} L_{ijq}(x_i) \leqq R_{jq} \qquad \begin{array}{l} \forall j \in N \\ \forall q \in Q \end{array} \tag{3}$$

$$x_k \leqq x_l \qquad k, l \in N \tag{4}$$

$$d_i \leqq x_i \leqq n \qquad i \in M \tag{5}$$

$$x_i \in \mathbf{N} \text{ (the natural numbers set)}$$

where: $M = \{1, 2, \ldots, m\}$; $N = \{1, 2, \ldots, n\}$; $P = \{1, 2, \ldots, v\}$; $Q = \{1, 2, \ldots, u\}$; $a_{it}$ is a discount factor for $i$ object and $t$ year and is computed from the equation

$$a_{it} = (1 + r)^{-t} \tag{6}$$

with discount rate $r$; $I_{it}$ is the investment cost for $i$ object and $t$ year, for $t = 0$ it can be the design cost; $K_{it}$ is the maintenance cost; $E_{itp}$ is the $p$ kind of effect for $i$ object and $t$ year and must be expressed in units which will enable summarization of the effects of $p$ kind over the whole set of objects under consideration; $x_i$ is the starting date of the $i$ object realization; $E_{ijp}(x_i)$ is determined as

$$E_{ijp}(x_i) = \begin{cases} E_{i,j-x_i+1,p} & \text{for } \; j - x_i \geqq 0 \\ 0 & \phantom{\text{for }}\; j - x_i < 0^{\cdot} \end{cases} \tag{7}$$

$L_{ijq}(x_i)$ is material means of $q$ kind for $i$ object and $j$ year. Partially, for $q = 1$ it means the financial funds defined as

$$L_{ijl}(x_i) = \begin{cases} I_{i,j-x_i+1} + K_{i,j-x_i+1} & \text{for } \; x_i - j \leqq 1 \\ 0 & \phantom{\text{for }}\; x_i - j > 1^{\cdot} \end{cases} \tag{8}$$

$D_{jp}$ is a demanded effect of $p$ kind for $j$ year; $R_{jq}$ is a limit of material means of $q$ kind for $j$ year.

The effects stand for such effects as: the amount of water supplied for customers, the electrical energy production, the water quality, the BOD reduction level, etc. The material means are finances, equipment, constructive elements, the efficiency of the building trade, etc.

To simplify the notation in the following discussion, we will limit the range of our interest to the case where $P = \{1\}$ and $Q = \{1\}$.

The above described optimization problem has goal function as well as constraints which are non-continuous and non-linear. As a matter of fact, it is a non-linear integer programming problem. However, it is quite easy to rearrange it to a linear zero-one integer programming task. Let $\alpha = n - j - k + 2$ and $\beta = n - k + 1$. Also, if we have

$$C_{ij} = \sum_{t=0}^{n-j+1} a_{it}(I_{it} + K_{it}) \qquad \forall j \in N \tag{9}$$

then each $C_{ij}$ value stands for the total discounted cost of existance and operation of $i$ object, assuming that its realization started in $j$ year. Thus, the new optimization problem is:

$$\text{minimize} \sum_{i \in M} \sum_{j \in N} c_{ij} x_{ij} \tag{10}$$

with constraints

$$\sum_{j \in N} x_{ij} \leqq 1 \qquad \forall i \in M \tag{11}$$

$$\sum_{i \in M} \sum_{j \in N} e_{ijk} x_{ij} \geqq D_\beta \qquad \forall k \in N \tag{12}$$

where

$$e_{ijk} = \begin{cases} E_{ia} & \text{for} \quad \alpha \geqq 1 \\ 0 & \phantom{\text{for}} \quad \alpha < 1 \end{cases} \tag{13}$$

$$\sum_{i \in M} \sum_{j \in N} L_{ijk} x_{ij} \leqq R_\beta \qquad \forall k \in N \tag{14}$$

where

$$L_{ijk} = \begin{cases} I_{i\alpha} + K_{i\alpha} & \text{for} \quad \alpha \geqq 1 \\ 0 & \phantom{\text{for}} \quad \alpha < 1 \end{cases} \tag{15}$$

$$\sum_{j \in N} 2^j x_{kj} \leqq \sum_{j \in N} 2^j x_{lj} \qquad \text{for} \quad k \in M, \ l \in M \tag{16}$$

$$\sum_{j=l}^{d_i} x_{ij} = 0 \qquad \text{for} \quad i \in M, \ d_i \in \mathbf{N} \cup \{0\} \tag{17}$$

$$x_{ij} = l \text{ or } x_{ij} = 0 \qquad \forall i \in M, \ \forall j \in N. \tag{18}$$

The zero-one integer programming task we obtained has $n \cdot m$ variables and $2 \cdot n + m$ constraints (for the case $P = \{1\}$, $Q = \{1\}$ and without regard to (16) and (17) types).

The constraints of the (11) type say that each object can be realized only once. Each of $n$ inequalities (12) guarantees that demands are satisfied, type (13) ensures that limits are not broken. Constraints (16) and (17) concern subsequence of objects ($k$ before $j$), $x_{ij} = 1$ means, of course, that the $i$ object realization starts in year $j$, $x_{ij} = 0$ indicates the case when it is going or has not started yet.

The above formulated problem can be solved by the well-known algorithm as a Balas method[7]. For bigger tasks, the statistical optimization as well as the heuristic method have been proposed in order to obtain the sub-optimal solution (Krajewski[8]).

## THE SEQUENTIAL PROBLEM UNDER UNCERTAINTY

Our previous consideration dealt with the problem of seeking an optimal sequence of water plant realization assuming that all costs, effects, as well as demand distribution are known over the planning horizon. However, that assumption is a simplification of reality because actual costs are related to the technologies currently used. Also, effects are disturbed by exceptionally bad meteorological and hydrological conditions. Finally, the demand distribution during the planning period might completely differ from those assumed.

Now let's formulate the sequential problem considering uncertain characteristics of demand distribution. In addition, it should be pointed out that the notion "uncertain" has been used on purpose instead of using a "random" term. The reason is that the measure of this uncertainty should not be treated as a probability in a statistical sense.

Let's assume that demands are known for each year within the planning horizon as well as some numbers from the range $\langle 0, 1\rangle$ saying the chance of such demand exists. Let's say that for $j$ year the $k$ value of demand, $D_j^k$, is related to chance $P_j^k$. Also, all other parameters, characterizing the planned plants such as investments, maintenance costs and effects, are known. At the time of making a decision, concerning the sequence of introducing the objects, the real values of demand are not known. Thus, as a result of an arbitrarily chosen sequence, a surplus or an inadequate production ability will occur and will cause economical loses.

If, for instance, the total effect of an entire system constructed with sequence $\pi$ is $E_j^\pi$ for the $j$ year, and demands are $D_j^k$, the loss $s_1$ will occur if $E_j^\pi < D_j^k$. This loss may be estimated as a function of unmet demand. If the relation is $E_j^\pi > D_j^k$, the loss $s_2$ is caused by overinvestment. Thus, the optimal sequence may be defined as a sequence which minimizes the total sum of discounted investments and maintenance costs and a function of $s_1$ and $s_2$ losses.

The measure of loss $s_1$ can be described as follows:

$$M(s_1) = f_1\left[\sum_{q\in N}\sum_{k\in K}\left(D_{n-q+1}^k - \sum_{i\in M}\sum_{j\in N} e_{ijq}x_{ij}\right)_+ \cdot P_{n-q+1}^k\right] \tag{19}$$

and for $s_2$:

$$M(s_2) = f_2\left[\sum_{q\in N}\sum_{k\in K}\left(\sum_{i\in M}\sum_{j\in N} e_{ijq}x_{ij} - D_{n-q+1}^k\right)_+ \cdot P_{n-q+1}^k\right] \tag{20}$$

where $e_{ijq}$ are effects defined as in equation (12) using index $q$ in place of $k$. Also, the symbol $(q)_+$ is defined as:

$$(q)_+ = \max(0, q) \tag{21}$$

and $K = \{1, 2, \ldots, k_1\}$ where $k_1$ is a number of possible demand levels.

Therefore, the search for an optimal sequence of water plants with uncertain demands is resolved by the following optimization problem:

$$\text{minimize} \quad \sum_{i\in M}\sum_{j\in N} c_{ij}x_{ij} + M(s_1) + M(s_2) \tag{22}$$

with restraints:

$$\sum_{i\in M}\sum_{j\in N} L_{ijq}x_{ij} \leqq R_{n-q+1} \qquad \forall q \in N \tag{23}$$

$$x_{ij} = 0 \text{ or } x_{ij} = 1 \qquad \forall i \in M \quad \text{and} \quad \forall j \in N$$

$$\sum_{j\in N} x_{ij} \leqq 1 \qquad \forall i \in M.$$

The limits $L_{ijq}$ in equation (23) concern financial funds and are defined as in equation (13).

The expressions contained in equations (19) and (20) can be simplified by introducing new variables $\vartheta$ and $\phi$. Let's assign:

$$\vartheta_{qk} = \Big(D^k_{n-q+1} - \sum_{i\in M}\sum_{j\in N} e_{ijq}x_{ij}\Big)_+$$

and

$$\phi_{qk} = \Big(\sum_{i\in M}\sum_{j\in N} e_{ijq}x_{ij} - D^k_{n-q+1}\Big)_+ \qquad \forall q \in N. \tag{24}$$

Let's assume then, that the functions $f_1$ and $f_2$ appearing in equations (19) and (20) are linear. Now, substituting (24) into (22), the optimization task is:

$$\text{minimize} \qquad \sum_{i\in M}\sum_{j\in N} c_{ij}x_{ij} + \lambda \sum_{q\in N}\sum_{k\in K} \varphi_{qk}P^k_{n-q+1} + \delta \sum_{q\in N}\sum_{k\in K} \phi_{qk}P^k_{n-q+1}$$

with restraints:

$$\sum_{i\in M}\sum_{j\in N} L_{ijq}x_{ij} \leqq R_{n-q+1} \qquad \forall q \in N \tag{26}$$

$$\sum_{j\in N} x_{ij} \leqq 1 \qquad \forall i \in M$$

$$x_{ij} = 0 \quad \text{or} \quad x_{ij} = 1 \qquad \forall i \in M, \quad \forall j \in N$$

$$\vartheta_{qk} \geqq 0, \quad \phi_{qk} \geqq 0 \qquad \forall q \in N \quad \text{and} \quad \forall k \in K. \tag{27}$$

The $\lambda$ and $\delta$ in the objective function can be interpreted as unit loss coefficients expressed in monetary terms, resulting in surplus or the inadequacy of productive abilities, respectively.

The task described above is an integer programming problem, assuming that all its coefficients are integers. This assumption does not change the generality of the problem.

It is worthwhile to note that the new formulation of the sequential problem did not cause the enlargement of the task. The number of all possible solutions remains unchanged. However, it does not mean that the new problem is easier to solve; on the contrary, it is even more difficult because of the greater number of new variables $\vartheta$ and $\phi$.

The problem can be stated as a zero-one integer programming task:

$$\text{minimize} \qquad \sum_{i\in M}\sum_{j\in N} c_{ij}x_{ij} \tag{28}$$

with restraints:

$$\sum_{i\in M}\sum_{j\in N} E_{ijq}x_{ij} = \sum_{k\in K} D^k_{n-q+1} \cdot P^k_{n-q+1} \qquad \forall q \in N$$

$$\sum_{j\in N} x_{ij} \leqq 1 \qquad \forall i \in M \tag{29}$$

$$\sum_{i\in M}\sum_{j\in N} L_{ijq}x_{ij} \leqq R_{n-q+1} \qquad \forall q \in N \tag{30}$$

$$x_{ij} = 0 \text{ or } x_{ij} = 1 \qquad \forall i \in M \quad \text{and} \quad \forall j \in N.$$

The above approach is an example only and does not deplete other existing possibilities of formulating the sequential problem under uncertainty.

Formally, the described approach can be extended with respect to the uncertain character of other components such as: limits of funds for each year, realization costs and effects of plants' existance. It seems however, that such extention would be premature for various reasons.

First, the estimation of the uncertainty levels would be very difficult for those components, particularly for realization costs. Secondly, the estimation of uncertainty of effects desires very detailed statistical analysis of hydro-meteorological processes affecting those effects. Finally, every attempt to extend the sequential problem seriously enlarges numerical difficulties. Because of the lack of fast and efficient methods of solving the large size of the sequential problems in its determinate form, it is hard to expect the task to be successfully solved considering the uncertainty of the mentioned components. One can judge, that after intensive efforts to obtain effective numerical algorithms, those factors will be able to be considered.

## SUMMARY AND CONCLUSIONS

The sequential problem in water resource systems has been presented. The formulation of its determinate form as well as the consideration of uncertainty has been described. Now, let's mention some other aspects of this problem, related directly or indirectly to the above text.

First, it is worthwhile to point out that the sequential problem for complex systems is a multicriterial problem by its own nature; such approach has not been applied here, but, it would be a natural extention after having the efficient algorithms for solving the one-criterial problem.

Also, the influence of object location has not been considered in the optimal solution. In dealing with certain kinds of objects (i.e. water treatment plants), location is a very important factor affecting the solution and can be taken into consideration by including additional constraints or, for larger systems, by simulation techniques.

## REFERENCES


1. W. S. Butcher, Y. Y. Haimes and W. A. Hall, Dynamic programming for the optimal sequencing of water supply projects. *Water Resources Res.* **5**(6), 1196–1204 (1969).
2. L. Becker and W-G. Yeh W., Optimal timing, sequencing and sizing of multiple reservoir surface water supply facilities. *Water Resources Res.* **10**(1), 57–62 (1974).
3. D. Erlenkotter, Sequencing of interdependent hydroelectric projects. *Water Resources Res.* **9**(1), 21–27 (1973).
4. D. Erlenkotter, Coordinating scale and sequencing decisions for water resources projects. In *Economic modeling for Water Policy Evaluation*, Vol. 3, PP. 97–112. North-Holland/TIMS Studies in the Management Sciences, Amsterdam (1976).
5. T. L. Morin, Solution of some combinatorial optimization problems encountered in water resources development. *Engng Optimization* **1**, 155–167 (1975).
6. S. Rinaldi, R. Soncini-Sessa and H. Stehfest, Optimal sequencing in installing wastewater treatment plants. IIASA RM-75-77.
7. E. Balas, An additive algorithm for solving linear programs with zero-one variables. *Ops Res.* July–August, 517–546 (1965).
8. W. F. Krajewski, Optimal sequencing of water plants in water resource systems. Ph.D. thesis (in Polish), Warsaw Technical University (1980).