SEARCH AND MONTECARLO TECHNIQUES FOR DETERMINING RESERVOIR OPERATING POLICIES.

A. Colorni (°), G. Fronza (°)

(°) Istituto di Elettrotecnica ed Elettronica
Centro Teoria dei Sistemi - Politecnico di Milano

1. INTRODUCTION

The problem of determining the optimal design and operation of a reservoir has been considered under many different viewpoints. Standard engineering procedures such as mass curves techniques [1] or classical hydrologic methods of analysis (see [2] - [5] for instance), have gradually been replaced by the use of mathematical programs, mainly separable [6] and dynamic [7], [8] ones, both in the deterministic and in the stochastic environment. In control theory terminology, these approaches correspond to open loop optimization schemes. It is an apparent drawback in presence of a stochastic input into the system such as the inflow. Of course, more reliable solutions can be obtained by introducing control laws (operating rules), that is when applying to feedback schemes. In this case, the problem of determining the optimal regulation, i.e. the optimal operating rule, is usually turned into a finite dimensional one by assuming a specific class of rules. In the most common case, the release in any period is made to depend upon the total available water in the period (initial storage plus inflow). Because of the constraints on the reservoir storage, the choice of the rule class cannot a priori be quite comfortable, for instance it is not possible to apply to linear functions. As a matter of fact, the great majority of existing applications considers Z-shaped rules, such as the normal ones. This prevents the optimization problem from being formulated as a mathematical program of a standard type. Three main different approaches have been proposed in the literature.

a) Pure simulation, based on the superimposition of a grid in the decision variables space [9].

b) A procedure consisting of the two following steps [10] .

   b1) Solution of the open loop control problem via dynamic programming in the deterministic environment supplied by a long synthetic record (see [11] - [14] for instance and the next section).

   b2) Least squares optimization for choosing the operating rule that yields a sequence of releases best fitting the optimal open loop sequence.

c) Bypass of the question (see the formulations in the stochastic environment in [15] , [16]) by setting the constraints that the "tails" of the Z-shaped rule are never effective, i.e. in no period the reservoir remains either empty or full.

In this case, the release turns out to be a linear function of the total available water. It must remarked, however, that on a monthly or a yearly basis, the constraint that the reservoir must never stay full is not justified by economic reasons and, in design problems, yields unnecesserarily large dam sizes.

In this paper, the simulation approach is followed, while taking a structural property of the optimization problem into account. Such a property enables to apply an efficient search-simulation scheme for determining the solution and hence to obtain a considerable decreas of the computational effort with respect to the "brute force" approach.

2. PROBLEM STATEMENT

Consider the problem of designing and operating a reservoir on a yearly basis, while assuming the following characteristics.

a) The reservoir is regulated by means of the normal operating rule $f(\cdot, x_1, x_2)$ shown in Fig. 1, where

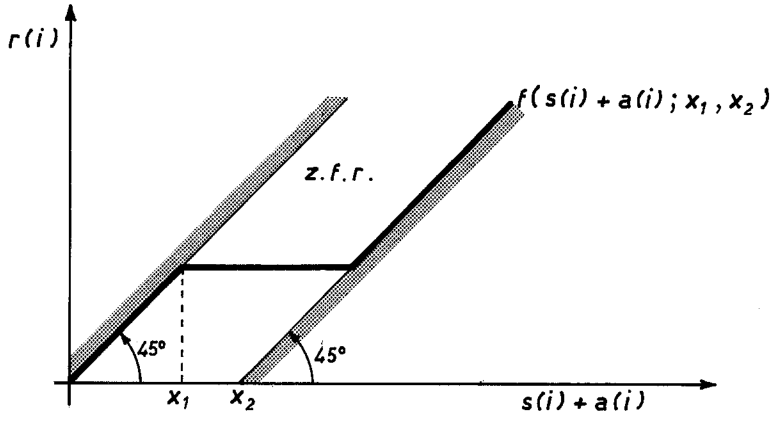

Fig. 1 Normal operating rule (z.f.r. = zone of feasible releases)

$s(i)$ = storage at the beginning of the i-th year, $i=0,1,\dots,n-1$;

$a(i)$ = total inflow during the i-th year;

$r(i)$ = total release during the i-th year;

x_1 = lower bound of the regulation range (decision variable)

x_2 = reservoir capacity (decision variable).

b) The release in any year is required to be not less than x_1, the guaranteed minimum (contract). This means the downstream users are assumed to plan their activities only on the basis of the contract, so that the profit due to the reservoir operation turns out to depend only on x_1, the extra water possibly supplied having no economic value. Let $g(x_1)$ represent such profit (Fig. 2a) and $c(.)$ the total cost function of the reservoir (Fig. 2b).

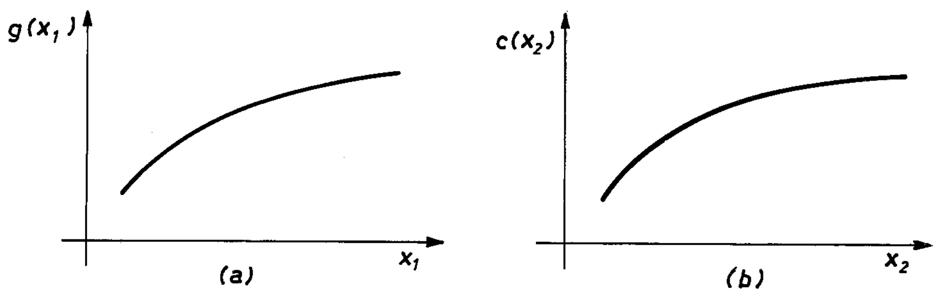

Fig.2 Profit (a) and cost (b) functions

Therefore, if the initial storage is assumed to be a given s_o and the continuity equation

$s(i+1) = s(i) + a(i) - r(i)$

is taken into account, then the reservoir design and regulation problem is turned into the follosing mathematical program

$(x = |x_1 \ x_2|')$:

$$\max_x \ g(x_1) - c(x_2) \tag{2.1}$$

$$s(i+1) = s(i) + a(i) - f\ (s(i) + a(i);\ x_1, x_2) \quad (s(o)=s_o)\ i = 0,1,\ldots,n-1 \tag{2.2}$$

$$s(i) + a(i) \geq x_1 \tag{2.3}$$

$$x_1, x_2 \geq 0 \quad . \tag{2.4}$$

This should be regarded to as a program in the stochastic environment (see [16]), in view of the nature of a(.). In particular, if $x^{(o)}(n)$ denotes the optimal solution of (2.1)-(2.4), that is in correspondence with a planning horizon of n years, the sequence $\left\{x^{(o)}(n)\right\}_{n=1}$ should be considered as a stochastic process.

(i) Such a process is a priori non-stationary because of the arbitrary choice of s_o. However it is reasonable to assume that, for n large, the process tends to become a stationary one.

(ii) The distribution of such "asymptotic" stationary process has small variance, i.e. the long run solution is not much affected by the introduction of a particular realizazion of the inflow process into (2.1) - (2.4).

This may be verified a posteriori by solving (2.1) - (2.4) in correspondence with different synthetic long records of inflows (see below).

In conclusion, it is reasonable to replace n in (2.1)-(2.4) by N = kn (e.g. $n \simeq 20$, $k = 50$) and to introduce a given record $\left\{\hat{a}(i)\right\}_{i=o}^{N-1}$ into (2.2)-(2.3), so that the program is formulated in the deterministic environment :

$$\max_{x} \; g(x_1) - c(x_2) \tag{2.5}$$

$$s(i+1)=s(i)+\hat{a}(i)-f(s(i)+\hat{a}(i);x_1,x_2)\,(s(0)=s_o), \quad i=0,1,\dots,N-1 \tag{2.6}$$

$$s(i)+\hat{a}(i) \geq x_1 \qquad i=0,1,\dots,N-1 \tag{2.7}$$

$$x_1, x_2 \geq 0 \tag{2.8}$$

The sequence $\left\{\hat{a}(i)\right\}_{i=0}^{N-1}$ is usually obtained by the much shorter historical datum through synthetic hydrology methods.

Specifically, the following procedure is adopted

I) The historical data $\left\{\hat{a}(i)\right\}_{i=-m}^{-1}$ are considered as a (partial) realization of an ergodic process. Of course, the distribution of the random variables of the process turns out to be skewed since the inflow is a non-negative variable (usually a lognormal or a Pearsom type III distribution can be assumed). Then the data are normalized or quasi-normalized [13]

$$\hat{b}(i) = h\,(\hat{a}(i))$$

via a proper transformation h(.)

and subsequently standardized

$$\hat{c}(i) = \frac{\hat{a}(i) - \mu_b}{\sigma_b}$$

where μ_b and σ_b represent the mean and variance of the normalized process respectively.

II) If $\left\{c(i)\right\}_i$ is the process that admits $\left\{\hat{c}(i)\right\}_{i=-m}^{-1}$ as a realization, then a model of $\left\{c\,(i)\right\}_i$ is built.

Usually such a model is selected among the ARMA (p,q) stationary models [17]

$$c(i+1)= \phi_1 c(i)+\phi_2 c(i-1)+\dots+\phi_p c(i-p+1)- \varepsilon(i)-\theta_1 \varepsilon(i-1)-\dots- \theta_q \varepsilon(i-q) \tag{2.9}$$

where

$\left\{\varepsilon(i)\right\}_i$ = purely random stationary gaussian process with zero mean and variance σ_ε ;

and

$\left\{\phi_j\right\}_{j=1}^{p}$, $\left\{\theta_k\right\}_{k=1}^{q}$ and σ_ε = model parameters to be estimated on the basis of the data $\left\{\hat{c}(i)\right\}_{i=-m}^{-1}$.

In most cases q = 0 or 1, while p = 0 or 1, or 2.

III) A realization $\left\{\hat{c}(i)\right\}_{i=0}^{N-1}$ is obtained by generating a realization of $\left\{\xi(i)\right\}_i$ by means of a Montecarlo technique (see [18] for example) and subsequently introducing it into (2.9).

IV) A realization $\left\{\hat{a}(i)\right\}_{i=0}^{N-1}$ (synthetic record) is supplied by anti-standardizing and subsequently anti-normalizing the sequence $\left\{\hat{c}(i)\right\}_{i=0}^{N-1}$.

3. PROBLEM SOLUTION

Turning to program (2.5)- (2.8) and letting $s(\cdot, x_1, x_2)$ denote the result of a "system simulation", i.e. the solution of (2.6), it is possible to choose an efficient algorithm by means of the following property (the proof is given in the Appendix).

Proposition

The function $s(i;\cdot,x_2)$ is non-increasing, the fucntion $s(i;x_1,\cdot)$ is non-decreasing.

Furthermore let

$$p(i;x_1,x_2) = s(i;\ x_1,x_2) + a(i) - x_1 \qquad (3.1)$$

Program (2.5) - (2.8) is then turned into the following :

$$\max_x\ g(x_1)-c(x_2) \qquad (3.2)$$

$$p^*(x_1,x_2) \geq 0 \qquad (3.3)$$

$$x_1,x_2 \geq 0 \qquad (3.4)$$

where

$$p^*(x_1,x_2) = \min_{0 \le i \le N-1} p(i;x_1,x_2) .$$

In view of the proposition and (3.1), the functions $p^*(\cdot,x_2)$ and $p^*(x_1,\cdot)$ turn out to be non-increasing and non-decreasing respectively. Hence, the optimal solution of (3.2)-(3.4) apparently lies on the curve implicitly defined by

$$p^*(x_1,x_2) = 0 . \qquad (3.5)$$

Let $x_2(x_1)$ denote the explicit form of (3.5). Since

$$\frac{dx_2(x_1)}{dx_1} = - \frac{\dfrac{\partial p^*(x_1,x_2)}{\partial x_1}}{\dfrac{\partial p^*(x_1,x_2)}{\partial x_2}}$$

the function $x_2(\cdot)$ is a non-decreasing one. No convexity property, however, can be established in correspondence with any inflow datum $\hat{a}(\cdot)$, so that optimality situations of different kinds might occur (Fig. 3). Recall that the feasible region F, which has been defi-

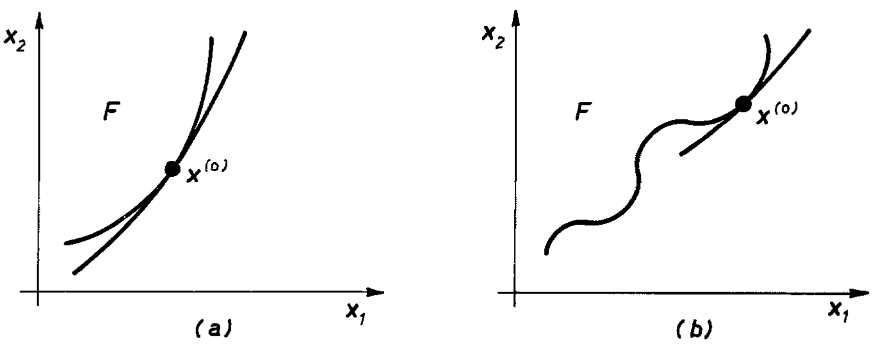

Fig. 3 Different optimality condition (F = feasible region)

ned via simulation is not explicitly known. Then assume the case is the one described in Fig. 3a and consider the solution algorithm, whose k-th step is the following.

i) The data at the beginning of the step is an interval $\mathcal{J}_k=(x_{11}^{(k)},x_{12}^{(k)})$

such that $x_{11}^{(k)} \leq x_1^{(o)} \leq x_{12}^{(k)}$, as well as the pairs $(x_{21}^{(k)}, x_{22}^{(k)})$, $(z_1^{(k)}, z_2^{(k)})$ where $x_{2j}^{(k)} = x_2(x_{1j}^{(k)})$, $z_j^{(k)} = g(x_{1j}^{(k)}) - c(x_{2j}^{(k)})$, $j = 1,2$.

In $\mathfrak{J}_k$ select a couple $(\tilde{x}_{11}^{(k)}, \tilde{x}_{12}^{(k)})$ in accordance with a Fibonacci search scheme [19].

ii) Determine $\tilde{x}_{2j} = x_2(\tilde{x}_{1j}^{(k)})$, $j = 1,2$, that is values of x_2 such that $p^*(\tilde{x}_{11}^{(k)}, \tilde{x}_{21}^{(k)}) = 0$ and $p^*(\tilde{x}_{12}^{(k)}, \tilde{x}_{22}^{(k)}) = 0$ respectively. These zeroes can be evaluated by means of simulation of (2.6), in accordance with a bisection search scheme, since $p^*(x_1,.)$ is a monotonic function.

iii) Compute $\tilde{z}_j^{(k)} = g(\tilde{x}_{1j}^{(k)}) - c(\tilde{x}_{2j}^{(k)})$, $j = 1,2$ and select $\mathfrak{J}_{k+1}$, on the basis of $z_j^{(k)}$ and $\tilde{z}_j^{(k)}$ $j = 1,2$ and through the criterion usually followed in the Fibonacci search.

The algorithm stops when $k = \bar{k}$ such that $\mathfrak{J}_{\bar{k}}$ is smaller than a preassigned interval. The use of search methods allows to obtain the solution with remarkable precision, a characteristic of some interest when dealing with problems where couspicuous profits and costs are involved.

To obtain the same degree of precision by means of grid sampling techniques would imply a much greater amount of computation. Of course if the situation is not the one described in Fig. 3.a, the algorithm may yield a local optimum instead of a global one. Some attempts with different starting intervals $\mathfrak{J}_o$ are usually enough to have a screening of the local solutions.

5. AN EXAMPLE

The algorithm has been used in the case of a reservoir, whose yearly inflow may be described by an AR(1) lognormally distributed stationary process characterized by :

mean = $75 \cdot 10^6$ m^3/sec

st.dev.= $29 \cdot 10^6$ m^3/sec

corr.coeff.= 0.105 .

The value of N has been set equal to 500. Profit and cost functions of the respective forms.

$$g(x_1) = 10^9 \left(1 - e^{-10^{-7}\lambda_1}\right)$$

$$c(x_2) = 10^9 \left(1 - e^{-25 \cdot 10^{-9}\lambda_2}\right)$$

has been used for different values of the parameters λ_1 and λ_2. The results are summarized in Table 1,

λ_1 \ λ_2	0.2	0.6	1.0	1.4	1.8
0.6	35	30	24	18	13
	34	28	22	17	12
	72	49	34	23	15
1.0	31	22	18	15	13
	29	20	16	13	11
	82	63	50	41	33
1.4	25	18	15	13	11
	23	16	13	11	9
	86	70	60	52	45
1.8	21	15	12	11	10
	19	13	10	9	8
	88	75	66	59	53

Table 1 Optimal solutions for different λ_1, λ_2.

where in each box $x_1^{(o)}$ (10^6 m^3), $x_2^{(o)}$(10^6 m^3) and $g(x_1^{(o)})-c(x_2^{(o)})$ (10^7 £) are reported from top to bottom. The computer time required to solve all cases described in Table 1 has been 2.22 min, on a UNIVAC 1108 computer.

5. CONCLUDING REMARKS

The problem of determining the optimal design and contract release of a reservoir has been described in the paper and a solution algorithm, based on the alternative use of well-known search procedure has been proposed.

A general remark concerns the use of policies, such as the normal operating rule, that make the decision on the release in any period also depend on the inflow in the period. Hence the choice of the outflow volume should be based on an information, not available at the beginning of the period, when the decision is actually taken.

A more reasonable viewpoint would be the one of using operating rules based on the storage as well as on inflow forecast, supplied by the stochastic model of the inflow process. This approach, anyway, has not yet been widely investigated in water resources literature.

REFERENCES

[1] W. Rippl, "The Capacity of Storage Reservoir for Water Supply", Proceedings of the I.C.E., Vol. 71, 1883.

[2] H. Hurst, "Methods of Using Long Term Storage in Reservoirs", Proceedings of the I.C.E., paper 6059, 1956.

[3] W. Feller, "The Asymptotic Distribution of the Range of Sums of Indipendent Random Variables", Annual of Mathematical Statistics, Vol. 22, 1951.

[4] A. Anis, H. Lloyd, "On the Range of Partial Sums of a Finite

Number of Independent Normal Variates", Biometrika, Vol. 42, 1955.

[5] P. Moran: "The Theory of Storage", Methuen, London, 1959.

[6] R. Dorfman, "Mathematical Models: the Multistructure Approach" in A. Maass et al.: "Design of Water-Resource Systems", Harvard University Press, Cambridge (USA), 1962.

[7] Hall W., Butcher W., Esogbue A., "Optimization of the Operation of a Multiple Purpose Reservoir by Dynamic Programming", Water Resources Research, June 1968.

[8] Buras N., "Scientific Allocation of Water Resources", American Elsevier, New York, 1972.

[9] Hufschmidt, M., and M. Fiering, "Simulation Techniques for Design of Water-Resource Systems", Harvard University Press, 212 pp., Cambridge, Massachusetts, 1966.

[10] G. Young, "Finding Reservoir Operating Rules", Proceedings of the A.S.C.E., Vol. 93, n. HY6, 1967.

[11] Thomas A., Fiering M., "The Nature of the Storage Yield Function", in "Operations Research in Water Quality Management" Harvard University Water Program, Cambridge, U.S.A., 1963.

[12] N. Matalas, J. Wallis, "Statistical Properties of Multivariate Fractional Noise Processes", Water Resources Research, v.7, n. 6, December 1971.

[13] P.E. O'Connell, "Stochastic Modelling of Long-term Persistence in Steamflow Sequences" Imperial College - London, Hydrology Section, Internal Report 1973-2.

[14] R. Clarke, "Mathematical Models in Hydrology", Irrigation and Drainage Paper 19 - F.A.O. - Roma 1973.

[15] C. Revelle, E. Joeres, W. Kirby, "The Linear Decision Rule in Reservoir Management and Design Development of the Stochastic Model", Water Resources Research, Vol. 5, n. 4, 1969.

[16] Eisel L.M., "Chance Constrained Reservoir Model", Water Resources Research, v. 8, n. 2, April 1972.

[17] G. Box, G. Jenkins, "Time Series Analysis: Forecasting and Control", San Francisco, Holden-day Inc.

[18] J. Hammersley, D. Handscomb, "Monte Carlo Methods", Metheun, London, 1965.

[19] Wilde D., Beightler C. "Foundations of Optimization", Prentice Hall Inc. 1967.

APPENDIX

Proof of the Proposition

Only the property of $s(i;.,x_2)$ is proved, since quite similar arguments hold for $s(i;x_1,.)$.

The proof is based on induction. Consider x_1', x_1'' with $x_1' \geq x_1''$: since, in view of the continuity equation,

$$s(1;\ x_1',x_2)-s(1;x_1'',x_2) = f(s_o+\hat{a}(0);x_1'',x_2) - f(s_o+\hat{a}(0);\ x_1',\ x_2)$$

it turns out that

$$s(1;x_1',x_2) \leq s(1;\ x_1'',\ x_2).$$

Moreover assume that $s(i;x_1',\ x_2) \leq s(i;\ x_1'',\ x_2)$. Then

$$s(i+1);\ x_1',x_2) - s(i+1;x_1'',\ x_2) = s' - s'' + f'' - f'$$

where, for simplicity of notation,

$$s' = s(i;\ x_1',\ x_2)$$
$$s'' = s(i;\ x_1'',\ x_2)$$
$$f' = f(s' + \hat{a}(i);\ x_1',\ x_2)$$
$$f'' = f(s'' + \hat{a}(i);\ x_1'',\ x_2)$$

i) If $s' + \hat{a}(i) \leq x_1'$, then $s(i+1;\ x_1',\ x_2) = 0$, so that

$$s(i+1;\ x_1',\ x_2) \leq s(i+1;\ x_1'',\ x_2).$$

ii) If $s''+a(i) \geq x_1'' + x_2$ then $s(i+1;\ x_1'',\ x_2) = x_2 \geq s(i+1;\ x_1',\ x_2)$.

iii) Finally if

$$x_1' \leq s'+ \hat{a}(i) \leq s'' + \hat{a}(i) \leq x_1'' + x_2 \quad \text{it turns out that}$$

$f' = x_1'$, $f'' = x_1''$ and hence

$$s(i+1;\ x_1',\ x_2) - s(i+1;\ x_1'',\ x_2) = (s'-s'') + (x_1''- x_1') \leq 0$$