
Application of stochastic dynamic programming in optimizing the regulation of hydropower reservoirs

TAN WEIYAN, LIU JIANMIN
Nanjing Hydrological Research Institute, Nanjing, China

HUANG SHOUXIN & FANG SHUXIU
Water Conservancy & Hydroelectric Power Scientific Research Institute, Beijing, China

ABSTRACT This paper discusses a problem in the theory of dynamic programming and Markovian decision process concerning a hydroelectric plant with a long term storage reservoir operating in a power system together with several run-of-river generating plants. The objective is to establish its regulating chart and various operating characteristics. In a working example, this chart will increase both the total guaranteed output in the dry season and the annual generation by 1-2% compared with a conventional one on the basis of a historic runoff record.

INTRODUCTION

Before the formulation of Howard's theory of dynamic programming and Markovian decision process (Howard, 1960), dynamic programming was used only for obtaining an optimal reservoir regulating plan within one year or less (Little, 1955). However, no fixed storage volume at the end of the calculation can be specified due to the randomness of runoff. At the same time, because the theoretical duration of operation is infinite, it is difficult to optimize the expected benefit and meet the demand for the probability of normal power supply. In 1963, we started to use Howard's theory to formulate a mathematical model for the optimal regulation of a long term storage reservoir of a hydroelectric plant (Tan et al., 1963); however, it became obvious that the problem involved a periodic Markovian process which was more general than Howard's. A mathematical proof was soon completed and the hydroelectric plants in series on the River Longxi were taken as a working example (Tan & Xu, 1966). In this paper the method is utilized further in the situation where the output in the dry season from several other run-of-river plants in Sichuan Province will be compensated by the above-mentioned ones. Various expected operating characteristics have been calculated by a probabilistic procedure, and economic loss due to violations of the operation policy was analysed so as to meet the need of practical work.

STATISTICAL TREATMENT OF RUNOFF PROCESS

Neglecting the long memory aspects of runoff fluctuations, runoff is
a general continuous random process with a period of one year. Generally in practice, due to inadequacy of records and for reasons of simplicity, runoff is treated approximately as a random series. Two forms have been adopted. The first is an independent random series. One month is taken as the time interval in the dry season, while one-third of a month is used in the wet season. The runoff of each interval follows a Pearson type III distribution. The second form is a simple Markovian series with correlation between successive time intervals. Lacking an appropriate form of a bivariate Pearson-III distribution, an idea originally proposed by Kartvelischvili (1956) is adopted. First of all, the runoff of each interval is transformed respectively into a normalized normal variate and each pair of new variates so obtained from successive intervals is then assumed to belong to bivariate normal distribution. As a working procedure, the functional transformation is carried out with the aid of a diagram, i.e. a curve relating pairs of variates from Pearson-III and normalized normal distribution, both corresponding to the same probability. In regard to the normality of the bivariate distribution, a statistical testing method given in Hald (1952) may be useful. Denoting the values of the new variates of the ith and (i + 1)th intervals as \( n_{i} \) and \( n_{i+1} \), under the assumption of normality, \( \chi^2 \) calculated from

\[
\chi^2 = \frac{1}{1 - r^2} (n_i^2 - 2rn_in_{i+1} + n_{i+1}^2)
\]

in which \( r \) is the correlation coefficient, will follow a \( \chi^2 \) distribution with the number of degrees of freedom equal to 2. Therefore, when pairs of values \( \chi^2 \) and corresponding logarithmic probability \( \log P(\chi^2) \) scatter closely around a straight line with slope equal to -0.217 and passing through the point (0,1) on a semi-logarithmic paper, the assumption can be accepted. In practical application, the conditional distribution of runoff may be obtained from the conditional normal distribution by an inverse transformation.

The above statistical treatment is only an approximate description of the runoff process. Whether greater benefit can be obtained from the model should be verified through reservoir regulating calculations based on runoff records.

**REGULATING PROCESS OF RESERVOIR**

The regulation process of a reservoir is defined by transition probabilities between successive reservoir states under a given regulating chart.

We take for analysis the treatment of runoff as a simple Markovian process. A year is divided into \( N \) intervals. For the \( n \)th interval \( (t_{n-1} \text{ to } t_n) \), the state variable of the reservoir is a combination of the runoff \( (Q_{n-1}) \) in the preceding interval, and the reservoir storage \( (V_{n-1}) \) at the beginning of the \( n \)th interval. For simplicity, the state variable is taken as discrete, the value field of which is a finite set. The number of elements of storage is \( M \), and that of runoff in each interval is \( L \). A typical element is denoted as \( V_{n-1}^m \) or \( Q_{n-1}^m \) (\( m = 1, 2, \ldots, M \); \( \ell = 1, 2, \ldots, L \)).

The discharge \( q \) for power generation in each interval is taken as
the decision variable. For the nth interval, the dependence of the
decision variable upon the state variable at the beginning of that
interval can be expressed as:

\[ q_n = q_n(V_{n-1}, Q_{n-1}) \quad n = 1, 2, ..., N \tag{2} \]

where the subscript to V denotes an instantaneous value at the
beginning of the nth interval, while those of q and Q denote flows
during intervals. The collection of the above relations for all
intervals within one year is called a policy, and its graphical
representation is the well-known reservoir regulating chart.

The transition between reservoir states within the nth interval
will be random, depending on the conditional probability distribution
function \( F(Q_n/Q_{n-1}) \) and the given regulating chart. In other words,
from that information the state transition probability \( p_{ij} \) that the
reservoir is in state i at the beginning of an interval and j at its
end \((i, j = 1, 2, ..., LM)\) can be decided, and the totality of such
probabilities constitutes a so-called state transition matrix of
order \( LM \times LM \), denoted as \( P[n] \), \( n = 1, 2, ..., N \). Obviously, \( P[n] \)
consists of non-negative elements not greater than 1, and the sum of
each row is equal to 1. The regulating process of the reservoir is
a periodic Markovian process with period \( N \) defined by a set of the
above matrices. Taking one year as an interval, its transition
matrix \( P \) fulfills the following relation:

\[ P = P[1] P[2] ... P[N] \tag{3} \]

Thus, in this case, our problem transforms into the same mathematical
model as Howard's. Starting with some arbitrary water level at the
beginning of a year, a stable probability distribution of reservoir
states, independent of the initial state, will be obtained after
operation over many years. If \( \pi_i \) denotes the probability that the
reservoir is in state i at the beginning of a year (instant \( t_0 \)),
and \( \pi[O] \) is an \( LM \)-dimensional row vector with \( \pi_i \) as its elements,
then we have

\[ \pi[O] = \pi[O] P \tag{4} \]

By adding the obvious relation

\[ \Sigma_i \pi_i = 1 \tag{5} \]

we can get a unique nonzero solution using the theory of systems of
linear equations. On account of the large number of state elements
and the limitation of computer memory, the system usually will not
be solved directly, and iteration is preferable using an arbitrary
\( \pi[O] \) as the initial value substituted into the right side of (4)
until convergence is reached.

With the solution \( \pi[O] \), row vectors of the stable probability
distribution of storage volume at any instant \( t_n \) can be calculated
recurrently:

\[ \pi[n] = \pi[n-1] P[n] \quad n = 1, 2, ..., N-1 \tag{6} \]
CALCULATION OF OPERATING CHARACTERISTICS

After $\mathbf{r}_o$ has been obtained, expected values or the probability distribution of various operating characteristics may further be calculated; the most important of them are stated below:

(a) The stable probability distribution of reservoir storage.
The elements of $\mathbf{r}_n$ correspond, respectively, to the probabilities $P(V^m_n, Q^\lambda_n), m = 1, 2, \ldots, M; \lambda = 1, 2, \ldots, L$, that the reservoir is in state $V^m_n$ and $Q^\lambda_n$. Therefore, the stable probability of storage at $t_n$ is

$$P(V^m_n) = \sum_\lambda P(V^m_n, Q^\lambda_n) \quad m = 1, 2, \ldots, M \quad (7)$$

(b) The expected annual power generation. In general, the discharge from the reservoir is determined by the regulating chart or equation (2). However, because the inflow $Q_n$ follows the conditional probability distribution $P(Q_n/Q_{n-1})$, the storage volume at $t_n$ may lie outside the allowed upper and lower bounds specified individually for each interval. When either bound is reached, the discharge must be changed so as to keep the reservoir storage within the allowed limits. The real mean discharge in that interval $q^*_n$ can then be calculated, and so can the power output $N^*_n$:

$$N^*_n = N^*_n (V^m_{n-1}, Q^\lambda_{n-1}, Q_n) \quad (8)$$

The expected value of $N^*_n$, $E_N$, is then:

$$E_N = \sum_{\Omega} N^*_n (V^m_{n-1}, Q^\lambda_{n-1}, Q_n) P(V^m_{n-1}, Q^\lambda_{n-1}) P(Q_n/Q_{n-1}) \quad (9)$$

in which $\Omega$ denotes a set formed by all possible discrete values of $V^m_{n-1}, Q^\lambda_{n-1}, Q_n$. The expected value of power generation in the nth interval and the whole year can then be found.

The power system will benefit from the output furnished by hydroelectric plants. In China, it is difficult to estimate the economic loss caused by power deficiency, so the benefit will be calculated as follows. When the total real output of the system is greater than or equal to the guaranteed value, the benefit will be considered equal to the output. In the opposite case, according to the concept of "penalty", the deficient power times a penalty coefficient ($c > 0$) will be subtracted from the real output so as to provide an expected benefit including the influence of economic loss. The penalty coefficient is proportional to the probability of normal power supply, so it can be determined through trial-and-error procedure in accordance with a specified probability.

(c) The real probability of normal power supply is defined as the fraction of the time within a year during which real output of the power system is greater than or equal to the guaranteed output. It can be calculated by averaging all the probabilities in all intervals weighted by the interval length.

RECURRENT CALCULATION FOR AN OPTIMAL REGULATING CHART

In view of the randomness of the runoff process, the optimization of
the regulating chart is a multistep random decision process. For the nth interval, we should make a decision \( q_n(i) \) based on the current reservoir state \( i \), so that the state transition probability and the corresponding "reward" are also determined. Let \( P^n_{ij}[q_n(i)] \) denote a probability of the event that starting from state \( i \) at the beginning of the nth interval and making a decision \( q_n(i) \), the state will change to \( j \), which will have an influence on the benefit to be obtained afterwards. Let \( r^n_{ij}[q_n(i)] \) denote an expected benefit in that interval under the same condition. According to the optimality principle in dynamic programming, for any starting state of any interval, the choice of optimal decision \( q^*(i) \) must maximize the total expected benefit to be obtained both in the next interval and in a future long period. Let \( g^n(i) \) denote the expected benefit obtained in the period from \( t_n \) to the end of reservoir operation, under a condition that we start from state \( i \) and always make optimal decisions in succeeding intervals. In that case, the optimality principle can be written as the following recurrence relation

\[
g^n(i) = \max_{q_n(i)} \left\{ \sum_j P^n_{ij}[q_n(i)] \left[ r^n_{ij}[q_n(i)] + g^n(j) \right] \right\} \quad i=1, 2, \ldots, LM
\]

Being obtained simultaneously with \( q^*(i) \), \( g_{n-1}(i) \) can be used in the solution for \( q^*_n(i) \) and \( g_{n-2}(i) \) by (10). Therefore, \( g_n(i) \) is often called the recurrence curve, and it can be moved parallel to a coordinate axis so as to pass through the origin of the coordinates without influence on succeeding computation. The whole procedure starts from some instant far enough in the future and proceeds backwards in time. Of course, for the calculation of the first interval, we must make an initial assumption of \( g_1(i) \). The regulating charts thus obtained approach a limit which is the optimal chart.

For a Markovian decision process with period equal to 1, a mathematical proof of this fact can be found in Howard (1960). For the general case with period \( N \), a proof of the convergence of the iteration process, the optimality of the solution and of its independence of the initial assumption of \( g_n(i) \) was given by Tan & Xu (1966). It is based on a treatment of a periodic chain in the theory of a Markovian process. The chain is simplified to become an equivalent simple chain through augmentation of states. In other words, there are \( LM \) reservoir states at each instant in the original model, while after augmentation the number of possible states reaches \( LMN \). All the states are arranged sequentially in the order of time. In order to preserve the equivalence of the two problems, we define a block matrix \( P' \) of order \( LMN \times LMN \) as the state transition matrix for all intervals:

\[
P' = \begin{bmatrix}
0 & 0 & \ldots & 0 \\
0 & P[1] & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & P[N-1] \\
P[N] & 0 & \ldots & 0
\end{bmatrix}
\]
The original model has thus been transformed into one with period equal to 1. Su & Deininger (1972) also discussed the same subject. Based on the optimal regulating chart, various operating characteristics can be calculated by a probabilistic method as described above, and a chart corresponding to a given probability of normal power supply can then be selected if desired.

The stable recurrence curve at some instant reflects the influence of current reservoir state on the future expected benefit. If a probability distribution of runoff forecast in one or more future intervals is given, it can be used instead of one without a forecast so as to optimize the discharges in those intervals.

OPTIMAL FUZZY DECISION

The foregoing statement is based on a condition that the future operation follows the regulating chart exactly, but there will always be some deviations in practice. In that sense, the so-called optimal chart is not really optimal.

The decision adopted in practice varies around the optimal discharge indicated by the chart within a range from a minimum allowed discharge to the maximum discharge capacity. All possible decisions constitute a fuzzy set which will be referred to as "fuzzy decision". The nearer to the indicated discharge the adopted decision is, the more probable it is. A distribution form readily conceivable is the normal distribution. Its main statistical parameter is the relative standard error $s$, with $s = 0$ for the unfuzzy decision case.

Let $q$ denote the discharge given by the chart, while $q'$ is the real discharge. Let $F[q'(i)]$ denote the normal probability distribution function of $q'$ with mean $q$ and standard deviation $qs$. Let $P_{ij}[q(i)]$ denote the probability of the event that starting from $i$ at the beginning of an interval, with $q(i)$ as its desired discharge and $q'(i)$ its real discharge, the reservoir is in state $j$ at the end of that interval, resulting in an expected benefit $r_{ij}[q(i)]$. Obviously, $P_{ij}[q(i)]$ and $r_{ij}[q(i)]$ can be obtained by the use of the total probability formula, i.e.

$$r_{ij}[q(i)] = \int_{0}^{q'} r_{ij}[q'(i)]dF[q'(i)] \quad (12)$$

The procedure for calculating an optimal chart and its operating characteristics is the same as above.

A CASE STUDY

The total installed power capacity of all hydroelectric plants calculated is 1134.5 MW, about half of that of the whole power system. The plants in series on River Longxi, where there is a plant upstream with a long term storage reservoir and three run-of-the river plants downstream with a negligible unregulated runoff, can be combined into one, the installed capacity of which amounts to 104.5 MW. Other run-of-the river plants in Sichuan Province can also be combined into a single one. Statistics show that the total
output of the latter is virtually independent of the runoff of River Longxi, so some benefit may be caused by the hydrological asynchronism of those rivers. Moreover, due to the effect of the reservoir storage capacity, the nonuniform output within a year may be compensated so as to increase further the guaranteed power output of all hydroelectric plants during the dry season to 360 MW.

Therefore, the criteria for optimal reservoir regulation in this case are as follows. Under the condition that the total guaranteed hydroelectric power will be supplied with a probability of 94-95%, the average annual power generation from the plants on the River Longxi will be maximized. Two variants of the streamflow series are considered, in particular an independent and a Markovian series. The results indicate that, when the penalty coefficient $c = 1.2$, the requirement of guaranteed output and corresponding probability of normal power supply can be satisfied. Also adopted is the alternative $c = 0$, for which the condition in the above criteria has been removed. All optimal charts of these alternatives have been tested by both random runoff model and observed records by finding their corresponding average annual power generation and other operating characteristics listed below. Also available are the results based on an optimal fuzzy decision and using a traditional reservoir regulating chart.

The results, summarized in Table 1, suggest:

(a) The difference between the benefits corresponding to the two different streamflow series models adopted in this case study is rather small, so it is reasonable to use the simpler assumption of independence.

(b) When the reservoir operation cannot be guided exactly by the chart, the benefit will be decreased. Although a distribution of discharge departure is assumed and the optimal fuzzy decision is adopted, the loss in average annual power generation can still reach 4.4% for large departures ($s = 0.5$). Hence the practical operation should be guided by the chart as closely as possible.

<table>
<thead>
<tr>
<th>Regulating chart</th>
<th>Streamflow series model</th>
<th>$c$</th>
<th>$s$</th>
<th>Probability of normal power supply (%)</th>
<th>Average annual power generation (GW h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal</td>
<td>Markovian series</td>
<td>1.2</td>
<td>0</td>
<td>94.03</td>
<td>94.2 463.48 445.65</td>
</tr>
<tr>
<td>Optimal</td>
<td>Independent random series</td>
<td>1.2</td>
<td>0.1</td>
<td>94.04</td>
<td>94.0 463.23 445.36</td>
</tr>
<tr>
<td>Optimal</td>
<td>Markovian series</td>
<td>1.2</td>
<td>0.5</td>
<td>92.94</td>
<td>92.9 442.93</td>
</tr>
<tr>
<td>Optimal</td>
<td>Observed records</td>
<td>0</td>
<td>0</td>
<td>91.15</td>
<td>89.8 466.54 448.12</td>
</tr>
<tr>
<td>Traditional</td>
<td>Observed records</td>
<td>0</td>
<td>0</td>
<td>88.9</td>
<td>439.73</td>
</tr>
</tbody>
</table>

* $A$ - using random streamflow models.
† $B$ - using observed records.
(c) Besides the increase of guaranteed hydroelectric power, the average annual energy production will increase by 1.3% as compared with the traditional chart. When the compensation of power output is not required by the system, the increase will reach 1.9%.

REFERENCES

Hald, A. (1952) *Statistical Theory with Engineering Applications.* John Wiley, New York, USA.


