J. of Spatial Hydrology
Vol.2 No.2

COMBINING PROBABILITY OF EMPTINESS AND MEAN FIRST OVERFLOW TIME OF A DAM TO DETERMINE ITS CAPACITY

Enayetur Raheem*
Lecturer

Sekander Hayat Khan Ph.D.
Professor

Institute of Statistical Research and Training (ISRT)
University of Dhaka, Dhaka- 1000, Bangladesh

ABSTRACT

Probabilistic considerations have been practiced in determining the capacity of a dam after the introduction of probability theory of dams by P. A. P. Moran (1954). Various researchers determined the capacity by using stationary distribution of the dam content, mean of the first emptiness time, and by specifying the probability of overflow of a dam. In this study, after highlighting the methods used by the design engineers using probabilistic consideration at various stages of the process, capacity has been determined by using the probability of emptiness and overflow simultaneously. As an example, riverflow data of Mitta Mitta River of Australia has been considered. The data was available only for a short period of time. So long inflow sequences have been generated by keeping intact the statistical properties of the historical data, and then determined the capacity.

Keywords and Phrases: Stochastic simulation, dam process, behavior analysis

*** Corresponding author (enayet@udhaka.net)**

Introduction

A dam is a barrier constructed across a natural waterway (river, stream etc.) to create a storage, which is generally called a lake or a reservoir. Capacity of a dam is defined as the volume capable of being impounded at the top of the dam. The capacity of a dam should not be too small to meet the demand or it should not be extravagantly large so that its capacity is never or rarely utilized. So, a proper capacity dam is necessary for maximum benefit. Systematic investigation for determining the capacity of a dam dates back from the work of Rippl (1883). Rippl determined the capacity of a dam by the mass curve method. This method is based solely on the historical inflow record. Moran (1954) formulated the probability theory of storage systems, which has now developed into an active branch of applied probability. Gould (1961) suggested Moran-type model to account for both seasonality and serial correlation of inflows to determine the capacity. McMahon (1976) took 156 Australian rivers and used Gould's modified procedure to estimate the theoretical storage capacities for four draft conditions (90%, 70%, 50%, and 30% of mean annual flow) and three probabilities of failure values 2.5%, 5% and 10%). Langbein (1958) gave probability routing method to determine the capacity with correlated annual flow. Hardison (1965) generalized Langbein's probability routing procedure using theoretical distributions of annual flows and assuming serial correlation to be zero. He determined capacity graphically for a given chance of deficiency and variability. The annual storage estimates were shown graphically for Log-normal, Normal and Weibull distributions of annual flows. Melentijevich (1966) obtained expressions for both time dependent and steady state distributions of reservoir content assuming an infinite storage and independent normal inflows. In considering finite reservoirs, Melentijevich used a random walk model and a behavior analysis of 100,000 normally distributed random numbers. From the analysis, he obtained an expression for the density function of the stationary distribution of storage contents.

Phatarfod (1976) suggested a method in determining the capacity which is based on random walk theory and is concerned with finding the probability of the contents of a finite reservoir being equal to or less than some value ℓC where $\ell C > 0$ and C is the reservoir capacity. The physical process of dam fluctuations can be linked to a random walk with impenetrable barriers at full supply and empty conditions. Phatarfod used Wald's identity, which is an approximate technique to solve the problem with absorbing barriers and a relation connecting the two kinds of random walks (McMahon & Mein, 1978). Phatarfod considered annual flows to follow the Gamma distribution and is based on a fixed draft. Among other important techniques, Sequent Peak Algorithm, Alexander's method, Dincer's method etc. (McMahon & Mein, 1978) are important.

Khan (1979) determined the capacity using stationary level of the dam content. Assuming that the successive inputs X_t are either mutually independent and identically distributed or serially correlated, he determined the capacity of a dam such that

$$\Pr(content\ of\ \ the\ \ dam < e\ell K) = P \tag{1}$$

where, *el* is a specified fraction of the capacity and *P* is also given and *K* is the capacity. He also suggested determining the capacity by considering mean of the first emptiness time [Khan (1992)]. He studied for Geometric, Exponential inflows and for arbitrary discrete inputs. He also determined the capacity by specifying the probability of overflow.

In the following, we have considered two important criteria-- probability of emptiness and mean first overflow time together to determine the capacity.

Suggested Approach

The approach uses behavior analysis in determining the capacity of a dam. Behavior analysis is a simulation technique in which various storage capacities are assumed and the inflow sequences are routed through the assumed capacities and how the dam `behaves', is observed. Important behaviors considered here include-- probability of emptiness (PE), probability of overflow (PO), mean first emptiness time (MFE) and mean first overflow time (MFO). MFE is the time at which the dam becomes empty for the first time during its life and MFO is the average time at which the dam overflows for the first time during its life. After routing the generated inflow to various assumed storages, capacity is determined by considering PE and MFO together.

In behavior analysis, the changes in storage content of a finite reservoir are calculated using a mass storage equation, which is given as

$$Z_{t=1} = Z_t + Q_t - D_t - L_t \tag{2}$$

Subject to $0 \le Z_{t+1} \le C$ where,

Z_{t+1} = Storage at the end of t^{th} time period

Z_t = Storage at the beginning of t^{th} time period

Q_t = Inflow during t^{th} time period

D_t = Release during t^{th} time period

L_t = Gross loss during t^{th} time period

C = Storage capacity

Equation (2) indicates that the content at time $t+1$ is obtained by subtracting the demand D_t and loss L_t during time *t* from the content at time *t*, (Z_t) along with the inflow during the time t, (Q_t).

In behavior analysis, it is assumed that the reservoir is initially full ($Z_0 = C$) and inflow and release of water are considered as discrete events. Initially, a sufficiently large capacity is chosen. Then the demanded water is released from the dam at the end of each period. The demand is usually considered as certain fraction of the mean inflow. In our study, we have considered the demand as 75% of mean flow. After releasing the demanded amount of water, there will be no release of water from the dam until the end of that period. During that period, inflow will occur and it will be stored in the dam until released. As part of the stored water will be lost due to many reasons including evaporation, leakage, seepage, etc., we have considered the accumulated loss as 10% of the stored content during a given period.

For determining the capacity of a dam at a certain place, studying the behavior of inflow pattern is essential. Unfortunately, in most cases either the flow records are not available or available for a shorter period. Thus, augmentation of the input series by generating longer series using simulation technique is essential. Keeping the statistical properties of the historical inflow sequences fixed, we first generate inflow data for an expected economic life of the dam. We have generated 1000 inflow sequences and simulation run has been performed to observe the said behaviors. In each run, we have routed 1000 independent inflow sequences through an assumed capacity. Thus, for each capacity we have calculated PE, PO, MFE and MFO, which are based on 1000 independent inflow series. After the first run for an assumed capacity, the behaviors are recorded and another capacity (50 units larger than the previous one) was chosen and similar action was performed. In this way, MFE, MFO, PE and PO, were observed for different capacities of the dam. The optimum capacity is chosen for which the mean first overflow (MFO) time is the largest with a given probability of emptiness (PE). For example, capacity of a dam with 5% chance of emptiness was calculated using the following condition

Capacity = min [Capacity with largest MFO] (3)
Subject to probability of emptiness ≤ 0.05

While determining the capacity in this way, we may encounter three possible situations:

(a) There are several capacities with largest MFO for all of which probability of emptiness is less than or equal to 0.05. In this situation, the smallest capacity will be chosen.
(b) There are several capacities with largest MFO but some of which have probability of emptiness greater than 0.05. In this situation, capacity for which the probability of emptiness is less than or equal to 0.05, will be chosen.
(c) In some situations, there might be no unique large MFO. That is, all the MFO might be the same for a range of capacities. In that case, we will select the capacity based on probability of emptiness only. That is, capacity, for which probability of emptiness is just less than or equal to 0.05, will be chosen.

Model Selection and Data Generation

In our study, 34 years historical inflow record of Mitta Mitta River of Australia has been used. Analyzing the record, it is found that first order autoregressive log-normal distribution was the best fit for the annual inflow records and hence, annual Markov model was appropriate.

Brittan (1961) proposed the following Markov model to represent actual stream flow when the annual streamflow, Q_i, are normally distributed and follow a first-order auto regressive model:

$$Q_{i+1} = \bar{x} + r_1(Q_i - \bar{x}) + t_i S\sqrt{(1-r_1^2)} \quad (4)$$

where

Q_i = Annual flow for i^{th} year

Q_{i+1} = Annual flow for $(i+1)^{th}$ year

$\bar{x}$ = Mean annual historical flow
S = Standard deviation of annual flows
r_1 = Annual lag one serial correlation coefficient
t_i = Normal random variate with a mean of zero and variance of unity.

Using the Markov model proposed by Brittan (1961), Troutman (1978) suggested the following model when the annual inflows are described as first-order autoregressive log normal.

$$X_{i+1} = m_x + r_1(x)(X_i - m_x) + t_i S_x \sqrt{(1-r_1^2(x))} \quad (5)$$

Where, X_i is logarithm of annual flows, the t_i is independent normal distributions with mean 0, variance 1 and $m_x, S_x^2,$ and $r_1(x)$ are the mean, variance and lag-one serial correlation coefficient of the log transformed inflows.

The meaning of such a Markovian flow model is that-- a given flow depends on the preceding flow and a random component alone. One explanation of using this model in generating inflows might be that a high flow in one time period will build up ground water level and thus provide a tendency towards another high flow in the next period. Similarly, ground water will be depleted during the period of low flow and so a low flow is expected to be followed by another low flow.

To generate annual inflows that will possess the same mean, standard deviation and other moments as those of the historical flow distribution, 'we must remember that the procedure reproduces the mean, variance, serial correlation coefficient, and skewness coefficient of the logs of the flows. The serial correlation and the skewness coefficient of the flows themselves will not necessarily be preserved" (Fiering and Jackson, 1971). This distortion may sometimes be important and so Matalas (1967) has suggested procedures for ensuring that the moments of the flows are maintained. Matalas assumed that the number ***a*** is a lower bound on the possible flow values and that if *x* denote a flow, then $y = \log(x - a)$ is normally distributed. He showed that the parameters of *x*'s are related to the parameters of the *y*'s as follows:

$$m_x = a = \exp\left[\frac{s_y^2}{2} + m(y)\right] \quad (6)$$

$$s^2(x) = \exp\{2[s^2(y)] + m(y)\} - \exp[s^2(y) + 2m(y)] \quad (7)$$

$$g(x) = \frac{\exp[2s^2(y)] - 2\exp[s^2(y)] + 2}{[\exp\{s^2(y)\} - 1]^{\frac{3}{2}}} \quad (8)$$

$$r_1(x) = \frac{\exp[s^2(y) r_1(y)] - 1}{\exp[s^2(x) - 1]} \quad (9)$$

Where,

a = Lower bound of the possible inflow values

$m(x)$ = Mean value of the historical inflow sequence

$s^2(x)$ = Variance of the historical inflows

$g(x)$ = Skewness coefficient of the historical inflows

$r_1(x)$ = Lag one serial correlation coefficient of the historical sequence.

To preserve the historical statistics of the flows rather than of their logarithms, the sample statistics $\bar{x}, S^2(x)\, and\, r_1(x)$ are calculated and substituting these values into the equations 6, 7, 8, 9, the estimates $m(y), s^2(y), r_2(y)$ and the lower bound (*a*) were obtained. These estimates were then used in the flow generation model to generate a series $x_1, x_2, ...,$ of logarithms of flows. Finally, a series of inflow was generated using the relation:

$$q_i = \exp(x_i) + a \quad (10)$$

The flows obtained from this procedure have expected parameters $\bar{x}, S^2(x)\, and\, r_1(x)$ as desired. Data have been generated considering 20, 30, 34, 40, 45, 50, 60, 70, 75, 80, 90 and 100 years of expected economic life of the proposed dam. The statistical properties of the generated inflows are given in Table 1.

Table 1: Comparison of parameters of the Historical and Generated Annual Inflows for Various Expected Life-Length

Life of the dam	Mean	St. Deviation	Coeff. Skew	Serial Corr.
Historical	1275 (7.003)	730 (0.5574)	1.50 (-0.0790)	0.061 (0.011)
20	1279 (6.994)	729 (0.5839)	1.51 (-0.5324)	0.058 (0.055)
30	1274 (6.992)	729 (0.5840)	1.51 (-0.5322)	0.058 (0.055)
34	1279 (6.995)	729 (0.5871)	1.51 (-0.5213)	0.058 (0.067)
40	1274 (6.988)	725 (0.5943)	1.48 (-0.5714)	0.047 (0.054)
45	1274 (6.991)	726 (0.5887)	1.47 (-0.5713)	0.050 (0.055)
50	1273 (6.989)	731 (0.5911)	1.55 (-0.5792)	0.059 (0.068)
60	1272 (6.989)	727 (0.5920)	1.51 (-0.6830)	0.058 (0.066)
70	1275 (6.989)	736 (0.5929)	1.56 (-0.5655)	0.053 (0.059)
75	1270 (6.984)	726 (0.5970)	1.44 (-0.6448)	0.065 (0.063)
80	1276 (6.991)	733 (0.5926)	1.55 (-0.5626)	0.058 (0.058)
90	1275 (6.991)	733 (0.5915)	1.56(-0.5685)	0.058 (0.061)
100	1274 (6.985)	734 (0.5952)	1.55 (-0.6021)	0.059 (0.054)

* Values in parenthesis indicate parameters of logarithm of flows

Capacity Determination: An Example

In this study, 0.05 was considered as the annual probability of emptiness. After the data have been generated and the demand has been fixed, we are ready to simulate the dam system using the continuity equation given by Equation (2). As an example, for a dam with expected life of 34 years, the result of the behavior analysis is given in Table 2.

Table 2: Capacity Determination Considering Probability of Emptiness and Mean First Overflow Time Together

Capacity	MFE	MFO	PE	PO
1750	9	3	0.190	0.342
2000	11	4	0.150	0.302
2250	13	4	0.124	0.269
2500	15	4	0.103	0.242
2600	15	5	0.097	0.232
3000	17	5	0.077	0.195
3200	18	5	0.070	0.178
3250	18	5	0.068	0.174
3300	18	6	0.067	0.170
3500	18	6	0.056	0.155
3650	18	6	0.054	0.141
3700	18	6	0.054	0.138
3750	19	6	0.051	0.134
3800	19	6	0.050	0.131
3850	19	6	0.049	0.128
4000	19	6	0.047	0.119
4150	19	6	0.044	0.111
4200	20	7	0.044	0.107
4250	20	7	0.043	0.105
4350	20	7	0.041	0.099
4450	20	7	0.040	0.094
4600	21	7	0.038	0.086
4800	21	7	0.036	0.078
5000	21	7	0.034	0.711
5100	22	7	0.033	0.065
5500	22	7	0.030	0.054
5650	22	7	0.029	0.050
5700	23	6	0.029	0.050
5750	23	6	0.028	0.048
5800	23	6	0.028	0.047
6000	24	6	0.024	0.044

* Selected capacity is underlined

Column (1) of Table 2 shows the capacities at which the generated inflows were routed. Column (2) the mean first emptiness time and column (3) the mean first overflow time. Probability of emptiness (PE) and probability of overflow (PO) are given in columns (4) and (5), respectively.

The rows of Table 2 show various capacities with corresponding MFE, MFO, PE, PO. We see that the largest value of the average first overflow time is 7 years and dam with capacity ranging from 4200-5650 units have average first overflow time of 7 years. This means, a dam with annual capacity between 4200-5650 units will overflow on an average at the 7^{th} year of its life for the first time.

Now applying the proposed technique given by Equation (3) the required annual capacity with probability of emptiness less than or equal to 0.05 is estimated as

Capacity = min [capacity with largest MFO]
= min [4200, 4250, ..., 5650 with MFO = 7 years]
= 4200

In this way, assuming same statistical properties as of the historical inflows, the dam system has been simulated for life-length of 20, 30, 40, 45, 50, 60, 70, 75, 80, 90, 100 years and capacity is determined. The result is shown in Table 3. It shows that higher capacity dams are essential for larger expected economic life, which seems to be a logical assertion.

Table 3: Estimated Capacity of a Dam for Various Life-length Obtained Using the Suggested Approach

Life-length (Years)	**Capacity**	**MFE**	**MFO**	**PE**	**PO**
20	3900	18	7	0.041	0.121
30	4100	20	7	0.042	0.110
34	4200	20	7	0.044	0.107
40	4400	23	8	0.046	0.089
45	4420	25	8	0.047	0.087
50	4750	27	10	0.049	0.070
60	4850	31	11	0.048	0.610
70	5100	33	12	0.049	0.055
75	5250	36	13	0.050	0.045
80	5400	40	15	0.051	0.042
90	5650	42	17	0.051	0.035
100	5800	43	19	0.050	0.030

Conclusion

For 34 years of economic life, the estimated capacity of the dam is 4200 with probability of emptiness 0.044. The estimated capacity ensures that the dam will overflow, for the first time, on an average at the 7^{th} year of its life. With this capacity, the dam is expected to become empty ($34\times 0.044 \approx 2$) twice and overflow four times ($34 \times 0.107 \approx 4$) during its life. The simulated result in Table 3 shows that if the inflow characteristics are same, higher capacity is required for longer expected economic life of the dam. By this technique, capacity of a dam with its expected first overflow time can be determined and hence, the proposed technique might be helpful in designing a dam where early knowledge of overflow is an important consideration.

We have simulated capacities in a 50-unit interval taking upto 100 years of life length. More detailed analysis can be done in the same manner. The study can be extended to monthly capacity determination provided that long sequence of monthly inflow records is available.

References

Brittan, M. (1961) Probability Analysis to the Development of a Synthetic Hydrology for the Colorado River. *In part IV of the ``Past and Probable Future Variations in Streamflow in the Upper Colorado River*". University of Colorado.

Fiering, M. B. and Jackson, B. (1971) Synthetic Streamflows. Springer-Verlag: New York

Gould, B. (1961) Statistical methods for estimating the design capacity of dams. *Journal of the Institution of Engineers*, Australia, 33(12), 405-416

Hardison, C. H. (1965) Storage to augment low flows. In Proceedings of Reservoir Yield Symposium. (Water Research Association).

Khan, S. H. (1979) On probability theory of dams. *Ph. D. thesis, University of Otago,* New Zealand

Khan, S. H. (1992) A note on the capacity of dam. *Pakistan Journal of Statistics,* 8(3B), 35-42

Langbein, W. (1958) Queuing theory and water storage. *Proceedings, Journal of Hydraulic Division, ASCE*, 84 (Paper-1811)

Matalas, N. (1967) Mathematical assessment of synthetic hydrology. *Water Resources Research*, 3(4), 937.

McMahon, T. A., (1976) Preliminary estimation of reservoir storage for Australian streams. *Civil Engineering, Transactions, The Institution of Engineers, Australia, CE18 (2): 55-59.*

McMahon, T. A., and Mein, R. G. (1978) Reservoir capacity and yield. Elsevere Scientific Publications: Nederland.

Moran, P. (1954) A Probability theory of dams and storage system. *Australian Journal of Applied Science*, 5, 116-126.

Melentijevich, M. (1966) Storage equations for linear flow regulations. *Journal of Hydrology,* 4, 201-223

Phatarfod, R. (1976) Some aspects of stochastic reservoir theory. *Journal of Hydrology*, 30, 199-217.

Rippl, W. (1883) The capacity of storage reservoir for water supply. *Inst. Civil Engrs*., 71, 270-278.

Troutman, B. (1978) Reservoir storage with dependent, periodic net inputs. *Water Resources Res.*, 14(3), 395-401.