Reliability Indices for Water Supply Systems

By Richard M. Vogel,[1] A. M. ASCE

Abstract: In the design of hydraulic structures for flood control, it has become standard practice to employ the average return period of a flood as the design event. This study introduces an analogous index for the design of a water supply system: the average return period of a reservoir system failure, defined as the expected number of years until the first reservoir system failure. Here a failure is defined as a year in which the prespecified yield could not be delivered by a water supply system. The mean and variance of the time until the first reservoir system failure are derived, as are the mean and variance of the duration of a reservoir system failure for a simple Markov failure model. Other indices of the reliability of a water supply system are also introduced. The assumption that sequences of reservoir surplus and failures can be represented by a two-state Markov chain is validated for the Pacific Northwest Hydroelectric Power System.

Introduction

The storage-yield relation is the traditional tool used by water resource engineers to determine the required capacity of a storage reservoir to maintain a prespecified reservoir release. Prior to the introduction of stochastic streamflow models by Sudler (28), Barnes (1), Fiering (4,5), and others, the storage-yield relationship was determined by applying the mass curve technique [introduced by Rippl (23)] to the available historical streamflow record. Thus, until the advent of the field of "stochastic hydrology," determination of the reservoir design capacity, S, in the U.S. seemed simple; one just determined the minimum storage that would have been required over the n-year historical period to provide the target yield with absolutely no water shortages. Fiering (5) documents three principal shortcomings of the strict use of the historical record:

1. The analysis is based solely on the historical record, and it is unlikely that the same flow sequence will recur during the active life of the completed structure.
2. The mass diagram (based solely upon the historic flow sequence) does not help the designer to establish or calculate the risk to be taken with regard to water shortages during periods of low flow.
3. The length of the historical record is likely to differ from the economic life of the proposed structure.

Stochastic streamflow models provide the analyst with a flexible tool that can be used to circumvent the shortcomings of the use of the historical required storage alone. Most importantly, synthetic streamflow traces derived from such models may be used to estimate the reliability or probability with which a storage reservoir can deliver scheduled

[1]Asst. Prof., Dept. of Civ. Engrg., Tufts Univ., Medford, MA 02155.

Note.—Discussion open until December 1, 1987. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on July 9, 1986. This paper is part of the *Journal of Water Resources Planning and Management*, Vol. 113, No. 4, July, 1987. ©ASCE, ISSN 0733-9496/87/0004-0563/$01.00. Paper No. 21684.

quantities of water. Such procedures were advocated by Maass, et al. (19). The use of a stochastic streamflow model in conjunction with Rippl's method or its automated equivalent sequent peak algorithm leads to a storage-reliability-yield relationship. Generalized analytic storage-reliability-yield relationships are now available (33,34) that do not require implementation of a stochastic streamflow model. Derived storage-reliability-yield relationships must be used cautiously, as evidenced by the substantial variability associated with design storage, yield, or reliability estimates based upon even relatively long hydrologic records (15,16,33,35).

Although stochastic streamflow models have been available for a few decades, the most common approach in U.S. practice has been to base estimates of the required capacity of a storage reservoir upon application of the sequent peak algorithm (or Rippl's mass curve) to the historical streamflow record. While this approach is still advocated in recent textbooks (22,27,30,31), it ignores the reliability associated with the resulting reservoir design capacity. Other textbooks (17,18) discuss the application of stochastic streamflow models in conjunction with the sequent peak algorithm to generate the cumulative distribution function (cdf) of required reservoir storage capacity S, corresponding to a fixed planning period of length N years. The cdf of S describes the relationship between the required storage capacity to meet a stated yield and the probability of failure-free reservoir operation p over an N-year planning period. Thus p is a measure of the reliability with which a reservoir of size S will provide failure-free operation over an N-year period.

It is the task of the water resource engineer to determine the reservoir storage capacity sufficient to meet the objectives of decision makers and society. In these situations, the engineer is often asked to convert statements of reliability over an N-year planning period to equivalent statements of annual reliability R_a, or vice versa. Relationships between the annual reliability and the reliability over an N-year planning period associated with the design of flood control structures were developed by Thomas (29) and further analyzed by Gumbel (6) and Yen (37). Those nonparametric relationships depend upon the fundamental assumption that the annual peak streamflows are independent and identically distributed random variables. These relations are in widespread use as evidenced by their inclusion in many textbooks on hydrology (2,8,18).

The relationship between annual reservoir system reliability R_a and the probability of failure-free reservoir operations p over an N-year planning period is more complex than for flood events because the sequence of reservoir surplus and failures is characterized by a dependent process. Klemes (12,13) derived reliability indices for the complex structure of sequences of reservoir surplus and failures that arise from reasonable assumptions regarding the character of the inflow and demand processes. Although the work of Klemes may be theoretically correct, the expressions that result are extremely complex and thus difficult to implement. The reliability indices developed by Klemes are difficult to implement because reservoir system states are modeled by an m-state Markov chain; e.g., Klemes (13) used $m = 10$ states of the reservoir system contents in his numerical example.

The primary objective of this study is to develop reliability indices for water supply systems that are easily implemented, yet characterize the likelihood of reservoir system failures, analogous to the relationships developed for flood events now in widespread use. To accomplish this task, the m-state Markov chain model formulation employed by Klemes (13), Moran (20), and others must be simplified considerably at the potential expense of misrepresenting the complex structure of reservoir system failures. Hirsch (9) and Stedinger, et al. (26) have employed a two-state Markov model to characterize sequences of water supply system surplus and failures. Jackson (11) also employed a two-state Markov chain to represent sequences of drought lengths (failure durations). The two-state Markov chain formulation of the sequence of reservoir system surplus and failures advocated by Stedinger, et al. (26) and Hirsch (9) is employed in this study to develop new indices of reservoir system performance.

A Two-State Bernoulli Model of Reservoir System States

In a given year, a reservoir system may be in either one of two states: (1) Failure; or (2) regular operation. Here a failure year is considered one in which the stated yield could not be met, and a regular year is one in which the stated yield is provided or exceeded. The assumption of only two reservoir system states dictates that the reservoir system must be able to pass from one state to the other in any given year. Consider, for example, a reservoir system whose storage capacity is greater or equal to the annual system demand. If such a reservoir system were full at the end of one year, a failure in the following year would be impossible even during an extremely dry year. However, many reservoir systems are subject to failure in every year, and it is these systems that are of interest here.

Now consider a reservoir system with independent annual inflows. Yevjevich (38) found that the no-persistence hypothesis could not be rejected at the 5% significance level for more than 80% of 446 annual streamflow records in western North America. Therefore, independence of the annual inflows for many reservoirs in the western U.S. is a plausible assumption. Now, if such a reservoir system refills every year and inflows are independent, then a failure in a given year occurs only if the demand exceeds the seasonal inflow plus the fixed storage capacity. Therefore, for a fixed demand, each year becomes a Bernoulli trial with a fixed probability of a failure. The trials are independent, and a fixed probability of a failure exists in any given year. The Bernoulli failure model does not acknowledge any persistence in sequences of reservoir system surplus and failures. In practice, large reservoir systems do not always refill after failures, and annual inflows often exhibit serial correlation; thus a more realistic model would exhibit persistence of the failure sequences.

A Two-State Markov Model of Reservoir System States

Acknowledgment of the persistence of sequences of reservoir system failures requires a model that is more complex than the Bernoulli failure model. To keep the analysis simple, a two-state Markov chain represents the next level of complexity, requiring one more parameter. More sophisticated models such as the m-state Markov model employed by Klemes (13), Moran (20), and others may be appropriate for detailed studies, yet the simple two-state Markov model provides insight into the relationships among annual reservoir reliability, reservoir system reliability over an N-year planning period, and the relative persistence of reservoir system failure years.

In this section, storage reservoir behavior is modeled, following Stedinger, et al. (26), concentrating upon both failure and regular (nonfailure) years. Here again, a failure year is one in which the stated yield cannot be met. Relationships are developed to evaluate the distribution of the duration of reservoir system failures and the distribution of the length of time until the first reservoir system failure occurs. While this analysis includes both the year in which the first reservoir system failure occurs as well as the duration of the failure, the actual magnitudes of the failures are ignored. For a more complete discussion of reliability measures associated with the frequency of failure years, failure durations, and magnitude of failures, see Klemes (15), Klemes, et al. (16), and Hashimoto, et al. (10).

Let the row vector $\mathbf{X}_y = (x_{1y}, x_{2y})$ specify the probability that a reservoir system is in either: (1) The failure state; or (2) the regular (nonfailure) state in year y. Also assume, as did Stedinger, et al. (26) and Hirsch (9) that $\mathbf{X}_y$, $y = 1, \ldots, N$ forms a Markov chain with probability transition matrix

$$\mathbf{A} = \begin{bmatrix} 1-r & r \\ f & 1-f \end{bmatrix} \quad \text{(1)}$$

where f = the probability that a failure year follows a regular year; and r = the probability that a regular year follows a failure year. The Markov chain model is given by

$$\mathbf{X}_{y+1} = \mathbf{X}_y \mathbf{A} \quad \text{(2)}$$

Fig. 1 depicts the two-state Markov model. The Markov assumption introduces memory into the process, although the transition of the reservoir system in any given year y is only influenced by the reservoir

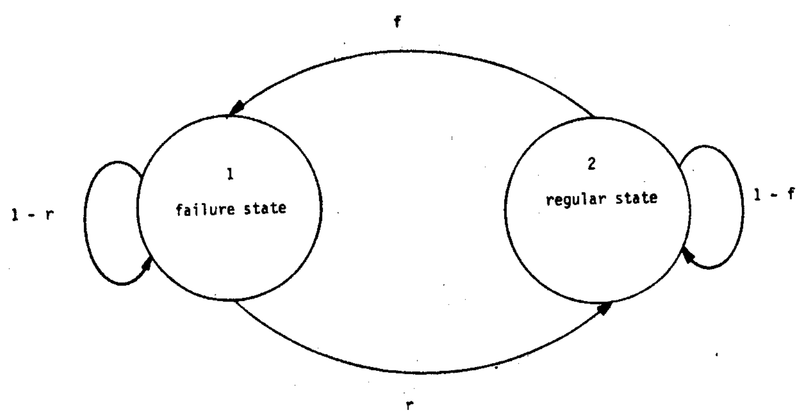

FIG. 1.—Two-State Markov Model of Reservoir System States

system state in the preceding year $y - 1$; more complex models are possible.

As y increases, $\mathbf{X}_y$ reaches a steady-state, and the solution to Eq. 2 becomes

$$\lim_{y\to\infty} \mathbf{X}_y = \left[\frac{f}{r+f}, \frac{r}{r+f}\right] \quad (3)$$

A derivation of this result may be found in Jackson (11). Thus, in the long run, the probabilities that the reservoir system will be in the failure or regular state are $f/(r + f)$ and $r/(r + f)$, respectively, regardless of the initial state of the reservoir system. Thus $r/(r + f)$ is the steady-state probability of regular operation or the annual reliability R_a.

In general, reservoir system failures are persistent phenomena; therefore the probability that the system remains in the regular state from one year to the next, $1 - f$, must be greater than the annual reliability, $R_a = r/(r + f)$. Similarly, persistence dictates that $(1 - r) > f/(r + f)$. This can only occur if $(r + f) < 1$. In fact, if $(r + f) = 1$, then $f = (1 - r)$, $r = (1 - f)$, and the two rows of the probability transition matrix in Eq. 1 are identical. In this case f and r are, respectively, the probabilities of failure and regular operation in each and every year.

When the sequent peak algorithm is employed to determine the smallest reservoir system design capacity required to assure regular or failure-free operation over an N-year planning period with probability p, then p becomes a steady-state probability. This is due to the fact that the sequent peak algorithm wraps the streamflow record around which generates the steady-state solution to the problem posed. If this double-cycling of the streamflow record is not performed, the resulting solution would not necessarily be a steady-state solution.

Now suppose we choose a reservoir capacity equal to S_p, using the sequent peak algorithm. Then the steady-state probability of regular (failure-free) operation over an N-year planning period, p, equals the probability of normal operation the first year $r/(r + f)$ times the probability that subsequent years remain free of failures:

$$p = \frac{r}{r+f}(1 - f)^{(N-1)} \quad (4)$$

The critical parameters of this model that must be estimated are r and f since p and N are usually taken as fixed values. Stedinger, et al. (26) describes the relationship between R_a, r, and f for the cases $p = 0.5$ and $N = 20, 50$. They conclude that knowledge of p and N are not sufficient to determine R_a unless one also knows the value of r or f. They note that in practice, p is not really known either, further complicating the problem.

For the two-state Markov model in Eq. 2, the average length of a reservoir system failure is simply $1/r$. This result is discussed in the next section. Given knowledge of the average length of a reservoir system failure, $1/r$, f may be obtained from Eq. 4, since p and N are usually specified. The resulting values of r and f may be expressed as the system annual reliability $R_a = r/(r + f)$. Stedinger, et al. (26) show that the range of estimates of R_a corresponding to a reasonable range of the mean length of a reservoir system failure is large. Thus Eq. 4 is only useful for estimating R_a from p and N when accurate estimates of r or f can be obtained. This is unlikely to be the case in practice.

Without knowledge of the average length of a reservoir system failure, $1/r$, Eq. 4, derived by Stedinger, et al. (26), is difficult to implement. In this study, Eq. 4 is simplified by conditioning the entire analysis upon the occurrence of regular or nonfailure reservoir operations during the first year. The two-state Markov chain model in Eqs. 1 and 2 and the steady-state solution in Eq. 3 remain unchanged. Now, the steady-state probability of regular operation over an N-year planning period, p, conditioned upon regular reservoir operations in the first year, equals the probability that all years, subsequent to the first year, remain free of reservoir system failures:

$$p = (1 - f)^{(N-1)} \quad (5)$$

Here p no longer depends upon r, and Eq. 5 may be solved directly for f, without resorting to a numerical algorithm as is required in the solution of Eq. 4 as follows:

$$f = 1 - p^{[1/(N-1)]} \quad (6)$$

Furthermore, f is completely specified by our knowledge of p and N and does not require assumptions regarding the value of r.

In the following sections, the two-state Markov model in Eq. 2 and the relationship among p, f, and N in Eqs. 5 and 6 are employed to derive expressions for the probability mass functions for the time to the first reservoir system failure, the duration of a reservoir system failure, and the number of failures in an N-year planning period.

Probability Distribution of Duration of Reservoir System Failure

From the two-state Markov chain model in Eqs. 1 and 2, one may derive the probability mass function for the duration of a reservoir system failure. Once the reservoir system is in the failure state, the probability that it will remain in the failure state for exactly one year is just the probability of one transition from the failure to the regular state r. In general, the probability that a reservoir system failure lasts L years is just the probability of $(L - 1)$ transitions from the failure state to the failure state, $(1 - r)^{(L-1)}$, times the probability of a transition from the failure to the regular state r in the last year. The probability mass function for L becomes

$$P[L = l] = r(1 - r)^{l-1}; \quad \text{for} \quad l \geq 1 \quad (7)$$

Thus L is a geometric random variable with mean

$$\mu_L = \frac{1}{r} \quad (8)$$

and variance

$$\sigma_L^2 = \frac{1 - r}{r^2} \quad (9)$$

TABLE 1.—The Mean, Standard Deviation and Coefficient of Variation of Duration of Reservoir System Failure, L, as Function of Probability that Year of Regular Operation Follows Failure Year, r

r (1)	μ_L (2)	σ_L (3)	C_L (4)
1.0	1.00	0.00	0.00
0.8	1.25	0.56	0.45
0.50	2.00	1.41	0.71
0.33	3.00	2.45	0.82
0.25	4.00	3.46	0.87
0.20	5.00	4.47	0.89
0.10	10.00	9.49	0.95
0.04	25.00	24.49	0.98

A derivation of Eqs. 8 and 9 may be found in texts on probability theory, e.g., see Ross (24).

Thus the coefficient of variation of the duration of a reservoir system failure is

$$C_L = \frac{\sigma_L}{\mu_L} = \sqrt{1-r} \qquad (10)$$

Table 1 reports μ_L, σ_L, and C_L for a range of r values. Knowledge of the transition probability r completely specifies the distribution of L. Table 1 shows that the variability of L, as summarized by C_L, increases as r decreases. The variability of L is substantial.

Average Return Period of Reservoir System Failure

In association with flood studies dealing with sequences of peak annual streamflows, Gumbel (6) and Thomas (29) defined the return period as the interval between flood events, where a flood event is defined as an annual peak flow above some threshold. Alternatively, the return period may be thought of as the number of years until the occurrence of the first flood event. Since Thomas (29), the meaning of the return period has changed. For example, Haan (8) defines the return period as the average elapsed time between occurrences of a flood event. Thus the return period was initially defined as the random time to an event, yet its meaning has changed to become the expected value of that random variable. This study distinguishes between these two definitions by using the terms return period and average return period.

Drawing an analogy to the average return period of a flood, the average return period of a reservoir system failure may be defined as the expected value of the return period, where the return period is the number of years before the occurrence of the first reservoir system failure. Let Z be the year in which the first failure occurs. Then the steady-state probability of the first failure occurring in the Zth year is equal to the probability of regular operation in the first year, followed by $Z-2$ years of regular operation and ending with a failure year. The probability mass function (pmf) of Z becomes

$$P[Z = z] = \begin{cases} 1 - \dfrac{r}{r+f} & \text{if} \quad z = 1 \\ \dfrac{r}{r+f} f(1-f)^{(z-2)} & \text{if} \quad z \geq 2 \end{cases} \qquad (11)$$

Then the average return period for a reservoir system failure is

$$T = \mu_Z = \sum_{z=1}^{\infty} zP[Z = z] \qquad (12a)$$

$$= 1 + \frac{r}{f(r+f)} \qquad (12b)$$

$$= 1 + \frac{R_a}{f} \qquad (12c)$$

The variance of Z is

$$\sigma_Z^2 = \sum_{z=1}^{\infty} (z - \mu_Z)^2 P[Z = z] \qquad (13a)$$

$$= \frac{r[r + f(2 - r - f)]}{f^2(r+f)^2} \qquad (13b)$$

$$= \frac{R_a(2 - f - R_a)}{f^2} \qquad (13c)$$

The average return of a reservoir system failure is a function of the transition probabilities r and f. These are difficult to obtain in practice. Relationships between T, p, r, and N are shown in Fig. 2. Fig. 2 was developed by setting the average length of a reservoir system failure, μ_L

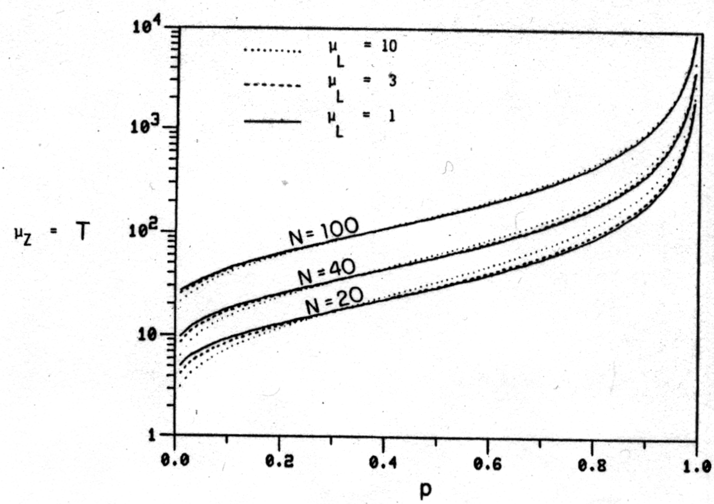

FIG. 2.—Average Return Period of Reservoir System Failure, $\mu_Z = T$, as Function of Probability of N-Year Failure-Free Reservoir Operation, p, and Average Length

$= 1/r$, equal to 1, 3, and 10 years and solving Eq. 4 for the values of f corresponding to combinations of N and p. Since Eq. 4 is a nonlinear function of f, a Newton-Raphson algorithm was employed. The computed values of f and the assumed values of r were then substituted into Eq. 12b to obtain T. Interestingly, the average return period of a reservoir system failure is relatively insensitive to the average length of a reservoir system failure μ_L. In practice the average length of a reservoir system failure is usually less than five years, in which case the dashed lines in Fig. 2 are representative of most realistic situation (i.e., $\mu_L = 1/r = 3$ yrs). In summary, use of Eqs. 4 and 12b with $\mu_L = 3$ provides a reasonable approximation to the relationship between T, p, and N.

Fig. 2 may be used to determine which nonexceedance probability p, associated with N-year failure-free operation, is appropriate for use in design applications. Clearly, reservoir design capacities based upon the upper quantiles of the distribution of required storage (i.e., $p \geq 0.8$) and large planning periods lead to extremely large average return periods. Reasonable designs with average return periods approximately equal to the planning period ($N = T$) correspond to use of S_{50} ($p = 0.5$) as the design capacity. In practice, our estimates of S_p contains a substantial amount of sampling variability leading to a design capacity with an average return period different from that shown in Fig. 2. Vogel (33) examines the impact of sampling variability in estimates of S_p on the resulting variability in the average return period of a reservoir system failure.

Eqs. 11, 12, and 13 can be simplified considerably by conditioning the entire analysis upon the occurrence of regular or nonfailure reservoir operations during the first year, as was done in the derivation of Eqs. 5 and 6. The probability that the first reservoir system failure occurs in the Zth year, conditioned upon regular reservoir operations in the first year, now equals the probability of $Z - 2$ years of regular operation, followed by a failure year. The pmf of the conditional time to the first failure Z^* becomes

$$P[Z^* = z] = \begin{cases} 0 & \text{if } z = 1 \\ f(1 - f)^{z-2} & \text{if } z \geq 2 \end{cases} \qquad (14)$$

where now the conditional average return period for a reservoir system failure is

$$T^* = \mu_{Z^*} = \frac{1 + f}{f} \qquad (15)$$

which may be combined with Eq. 5 to obtain

$$T^* = \frac{2 - p^{[1/(N-1)]}}{1 - p^{[1/(N-1)]}} \qquad (16)$$

Similarly, one obtains from Eq. 14 the variance of Z^* simply as

$$\sigma_{Z^*}^2 = \frac{1 - f}{f^2} \qquad (17)$$

which is identical to the variance of the unconditional return period of a reservoir system failure σ_Z^2 given in Eq. 13 when the annual reliability R_a is equal to one.

TABLE 2.—Comparison of Unconditional Average Return Period, T, with Average Return Period T^*, Conditioned upon Regular Operation in First Year

				T		
p (1)	N (2)	C_{Z^*} (3)	T^* (4)	$\mu_L = 1$ (5)	$\mu_L = 3$ (6)	$\mu_L = 5$ (7)
0.5	20	0.95	29	29	30	31
0.5	40	0.97	58	58	59	60
0.5	100	0.99	144	145	146	147
0.75	20	0.98	68	70	75	80
0.75	40	0.99	137	140	144	149
0.75	100	1.00	346	348	353	358
0.95	20	1.00	372	390	427	464
0.95	40	1.00	762	780	817	854
0.95	100	1.00	1,932	1,950	1,987	2,024

The coefficient of variation of Z^* is

$$C_{Z^*} = \frac{1 - f}{1 + f} \quad \text{(18)}$$

Table 2 documents T^* and C_{Z^*} over a range of values of the probability of N-year failure-free reservoir operation normally encountered in practice. The conditional return period Z^* exhibits the unique property that its mean T^* is in most practical situations approximately equal to its standard deviation σ_{Z^*}.

Table 2 also compares the conditional average return period T^* with the unconditional average return period T given by Eq. 12 for $\mu_L = 1$, 3, and 5 years. The values of T^* are generally equal to, or slightly less than, the values of T. For all practical purposes, the estimator of T^* given by Eq. 16 is recommended here for use in determining the average return period of a reservoir system failure due to its simplicity and general compatibility with the more complex estimator T. The expression for T^* in Eqs. 15 and 16 is a dramatic simplification since it does not require use of a numerical method such as the Newton-Raphson algorithm as is required in the determination of T from Eqs. 12 and 4.

Other Indices of Reservoir System Performance

The average return period of a reservoir system failure T is simply the average number of years prior to the first reservoir system failure. In this section it is shown that $T = \mu_Z$ is not indicative of the distribution of the time until the first reservoir system failure. Perhaps a more reasonable statistic would be to report the "likely recurrence interval," which is defined here to be that interval of time over which reservoir system failures are likely to occur 90% of the time.

The qth percentile of the distribution of the year in which the first reservoir system failure occurs, Z_q, may be obtained by choosing the largest value of Z_q such that

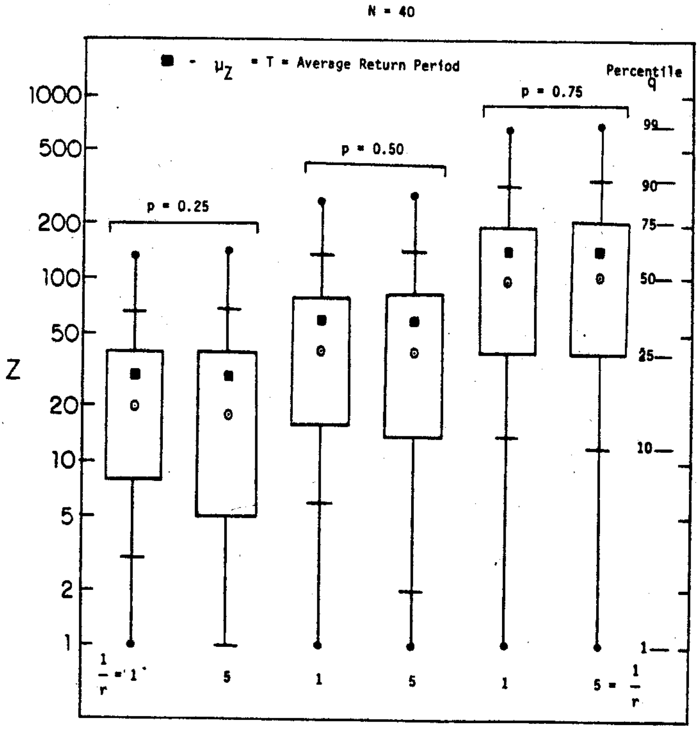

FIG. 3.—Box Plots Showing Distribution of Time to First Reservoir System Failure Z for Probabilities of 40-Year Failure-Free Reservoir Operation (p = 0.25, 0.50, and 0.75) and Average Lengths of Reservoir System Failure, $\mu_L = 1/r$, Equal to One and Five Years

$$\sum_{z=1}^{Z_q} P[Z = z] \leq q \qquad (19)$$

where the pmf of Z is given in Eq. 11. The distribution of Z is shown in Fig. 3, using box plots, for the cases $N = 40$ and $p = 0.25$, 0.50, and 0.75. In each case two distributions are depicted, one with the average failure length μ_L equal to one year and the other with $\mu_L = 5$ years. This range of μ_L should capture the range of practical interest. When $p = 0.75$, the distribution of Z is not very sensitive to the value of μ_L, whereas when $p = 0.25$, the lower quantiles of the distribution of Z ($q < 0.25$) are rather sensitive to the assumed value of μ_L. However, given the extreme variability associated with Z in Fig. 3, the issue of which value of μ_L to use becomes moot.

It is evident from Fig. 3 that the distribution of Z is extremely variable and use of the average return period T, which is simply the expectation of Z, in no manner represents that variability. If one is forced to use a point estimate summarizing the distribution of Z, then it might be more illustrative and informative to choose a lower quantile of Z, e.g., Z_{10} ($q = 0.10$).

Again, estimation of Z_q in Eq. 19 can be dramatically simplified by conditioning the entire analysis upon the occurrence of regular or non-failure reservoir operations during the first year. The conditional qth percentile of the distribution of the year in which the first reservoir system failure occurs, Z_q^*, is obtained by substituting the probability mass function for Z^* given by Eq. 14 into Eq. 19 which yields

$$\sum_{z=2}^{Z_q^*} f(1-f)^{z-2} \leq q \quad \text{(20)}$$

This approach is much simpler than using the unconditional approach because Eq. 20 may be solved for Z_q^*:

$$Z_q^* = \frac{\ln(1-q)}{\ln(1-f)} + 1 \quad \text{(21)}$$

where f is uniquely determined from Eq. 6, given values of p and N. Since the pmf of Z^* and Z are identical when the annual reliability is equal to one, one may expect the pmf of Z^* to approximate the distribution of Z when R_a is close to one, as is often the case in practice.

Application

The reliability indices derived in this study are based upon a two-state Markov chain model of reservoir system surplus and failures. Complex reservoir systems exist for which the Markov assumption and/or the two-state representation may be oversimplifications. Nevertheless, the reliability indices developed here have potential for providing insight into the relationships among annual reservoir reliability, reservoir system reliability over an N-year period, and the average return period of a reservoir system failure analogous to expressions now in widespread use in flood frequency analysis. Given the potential associated with the two-state Markov model formulation described here and in Jackson (11), Hirsch (9), and Stedinger, et al. (26), future research should determine exactly which reservoir systems result in failure sequences that are well-approximated by a two-state Markov model. The following case study examines how well the two-state Markov chain approximates the distribution of reservoir system failures for one particular system, the Pacific Northwest hydroelectric system.

Pacific Northwest Hydroelectric System.—The following example evaluates the ability of the two-state Markov model to represent failure sequences generated from the Pacific Northwest hydroelectric power system. The resources used to supply electric power loads in the Pacific Northwest (Washington, Oregon, Idaho, and western Montana) are predominately hydroelectric. The firm energy load that the Pacific Northwest hydroelectric power system is able to meet is assumed to be the maximum amount of energy that the system would be able to generate, without failure, if the historical streamflows were to recur. Recently Dean and Polos (3) of the Bonneville Power Administration in cooperation with S. J. Burges and D. P. Lettenmaier of the University of Washington employed a stochastic streamflow model in combination with a simulation model of the Pacific Northwest hydroelectric system to examine the likelihood that the system will fail to meet its firm loads. Utilizing the 101-year historical streamflow record available for the Columbia River at The

Dalles, Oregon, S. J. Burges and D. P. Lettenmaier fit a monthly stochastic streamflow model, which was then employed to generate 1,000 sets of 100-year streamflow traces. Each 100-year streamflow trace was routed through a complex simulation model that mimics the regulation of the existing hydroelectric power system. The details of the simulation model of the hydroelectric power system are too involved to report here; the reader is referred to the work of Dean and Polos (3).

One of the failure statistics reported by Dean and Polos was the number of failures that occurred in each of the 1,000 sets of 100-year system simulations. Here a failure is defined as the inability of the system to meet its firm load. In this section the simulated probability mass function of the number of failures over a 100-year period is compared to the theoretical distributions derived from the two-state Markov and the two-state Bernoulli models.

Let X be the number of failures in an N-year period. Again the entire analysis is conditioned upon normal operation in year zero followed by N years of operation in which the system may either fail or perform normally. To obtain an approximate but simple analytic expression for the probability mass function for X, under the two-state Markov model, it is assumed that normal operations also occur in year N. Then

$$P[X = 0] = (1 - f)^{N-1} \quad \text{(22a)}$$

and for $x \geq 1$

$$P[X = x] = (1 - r)^x (1 - f)^{N-x-1} \sum_{j=0}^{x-1} \binom{x-1}{j} \binom{N+1-x}{j+1} \gamma^{j+1} \quad \text{(22b)}$$

where $\gamma = fr/(1 - f)(1 - r)$.

For the two-state Bernoulli failure model, each year represents an independent Bernoulli trial with probability of a failure equal to θ. Under this simple model the number of failures X in an N-year period is distributed binomial:

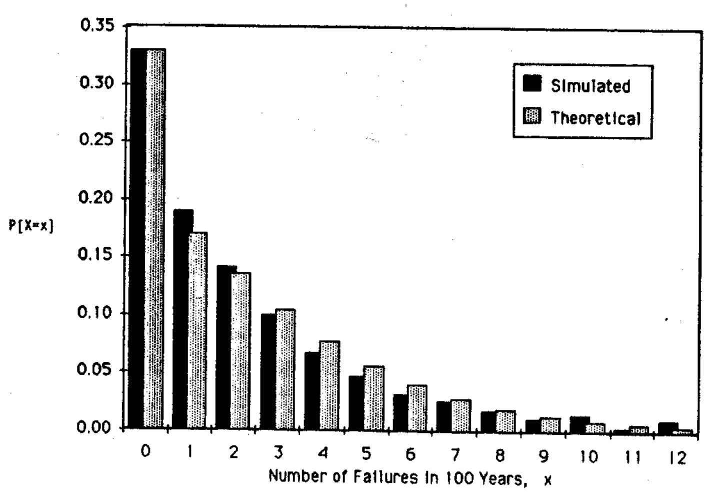

FIG. 4.—Comparison of Simulated and Theoretical Probability Distributions for Number of Failures in 100-Year Period, X, for Pacific Northwest Hydroelectric System

TABLE 3.—Comparison of Cumulative Distribution Functions of Number of Failures in 100-Year Period for Pacific Northwest Hydroelectric System

Number of failures, x (1)	Simulation (2)	$P[X \geq x]$ Two-state Markov model (3)	Two-state Bernoulli model (4)
1	0.670	0.670	0.910
2	0.481	0.499	0.699
3	0.340	0.364	0.425
4	0.240	0.259	0.215
5	0.173	0.182	0.090
6	0.126	0.126	0.032
7	0.095	0.087	0.010
8	0.070	0.059	0.0027
9	0.053	0.040	0.00066
10	0.042	0.027	0.00014
11	0.028	0.018	0.000031
12	0.025	0.012	0.0000080
13	0.015	0.0086	0.0000039
14	0.011	0.0062	0.0000032
15	0.008	0.0046	0.0000031

Note: The cumulative distribution function of X for the simulation experiment is based upon 1,000 replicate 100-year simulation runs.

$$P[X = x] = \binom{N}{x}\theta^x(1 - \theta)^{N-x} \quad (23)$$

Dean and Polos reported 2,400 failures over the 100,000 years of simulated system operation. Therefore the steady-state probability of a failure is equal to 0.024. Thus for the Bernoulli failure model $\theta = 0.024$, while for the Markov failure model $f/(r + f) = 0.024$. Similarly, out of 1,000 sets of 100-year system simulations, Dean and Polos reported 330 of those 1,000 sets had no failures. Therefore from Eq. 5 we obtain $P[X = 0] = (1 - f)^{99} = 0.33$, which yields $f = 0.01114$. Since $f/(r + f) = 0.024$, we obtain $r = 0.4529$. This yields the two parameters of the two-state Markov model of reservoir system failures in the Dean and Polos Pacific Northwest hydroelectric system model. Fig. 4 shows the agreement between the simulated distribution of X and the theoretical distribution of X given in Eq. 22. The agreement is in general excellent. Table 3 compares the simulated cumulative distribution of X with the theoretical cumulative distribution functions derived for the Markov and Bernoulli failure models. Here we observe that the two-state Bernoulli failure model does not adequately represent the likelihood of future reservoir system failures.

The two-state Markov model captures the differential persistence of failures. That is, suppose failures are dependent. Then from Table 3 the probability of no failures over a 100-year period is simply $P[X = 0] = 0.33$. On the other hand, if failures are independent, the probability of 10 or more failures is $P[X \geq 10] = 0.00014$ from Table 3, whereas Dean and Polos obtained $P[X \geq 10] = 0.042$. Thus differential persistence makes it much more likely to go without failures ($P[X = 0]$) or to experience a large number of failures ($P[X \geq 10]$).

Table 3 and Fig. 4 show the adequacy of the two-state Markov model for representing the distribution of the number of failures in a 100-year period for one existing complex water supply system. These results are encouraging. It is hoped that future investigations will examine the adequacy of the two-state Markov model to represent the structure of sequences of reservoir system surplus and failures for a wide class of water supply systems.

Summary

In the design of hydraulic structures for flood control, it has become standard practice to employ the average return period of a flood as the design event. This study developed an analogous index for the design of a water supply system: the average return period of a reservoir system failure. The average return period of a reservoir system failure is derived from the probability distribution of the return period of a reservoir system failure for a simple two-state Markov model. The resulting expressions are simplified dramatically by conditioning the entire analysis upon regular (or nonfailure) reservoir system operations during the first year. The resulting expressions for quantiles of the distribution of return periods of reservoir system failures or the average return period of a reservoir system failure are readily estimated from Eqs. 6, 15, and 21 for a given planning period N and probability of N-year failure-free operation p.

The reliability indices derived in this study are based upon a two-state Markov model of reservoir system states. More sophisticated models may be appropriate for detailed studies; however, the simple two-state Markov model should provide insight into the relationships among annual reservoir system reliability R_a, reliability over an N-year planning period p, and the expected time T^* until a failure, given that one is in a year of regular (nonfailure) operations. The simple reliability indices derived here are not intended to replace reliability indices derived from more complex simulation studies. Rather they were developed to increase our understanding of derived sequences of reservoir system surplus and failures and to assist in the development of explicit statements regarding the likelihood of future reservoir system failures. Application of the two-state Markov model to the Pacific Northwest hydroelectric power system indicated that this simple model can represent the structure of sequences of reservoir system surplus and failures that result from one very complex water supply system. Although further research is required to specify a priori which water supply systems are well approximated by the two-state Markov model formulation, this study describes a general class of systems to which the derived reliability indices should apply. It is hoped that future investigations will verify the adequacy of the two-state Markov model for a wide class of water supply systems.

In a recent national assessment of our nation's water resources, the Water Resources Council (36) concluded that 17 of the nation's 21 water resource regions have or will have a serious problem of inadequate surface-water supply by the year 2000. As increasingly marginal surface-water supply sites are pressed into service, target yields at both existing and proposed sites can only increase. In many instances, increased demands are being met by more efficient management and utilization of existing reservoir systems, rather than by construction of new facilities [for an example of this recent phenomenon see Sheer and Flynn (25) and Palmer, et al. (21)]. Whether new facilities are envisaged or the existing reservoir system is to be operated more efficiently, the storage-reliability-yield relationship is a fundamental ingredient. The use of stochastic streamflow models in conjunction with the sequent peak algorithm may be used to develop the cumulative distribution function of required reservoir system capacities and/or yields. The reliability indices developed here show potential for developing explicit statements regarding the likelihood of future reservoir system failures.

Acknowledgments

The writer is indebted to Jery R. Stedinger, Vit Klemes, Ben C. Yen, and David I. Hellstrom for their constructive suggestions and comprehensive review of early versions of this manuscript. In addition, the writer is indebted to his father, Erwin Vogel, for his assistance in deriving Eq. 22. The writer gratefully acknowledges Lawrence A. Dean, John A. Polos, Stephen J. Burges, and Dennis P. Lettenmaier for making available their simulation results for the Pacific Northwest hydroelectric system.

Appendix.—References

1. Barnes, F. B., "Storage Required for a City Water Supply," *Journal of the Institute of Engineers* (Australia) Vol. 26, No. 198, 1954.
2. Chow, V. T., Ed., *Handbook of Applied Hydrology*, McGraw Hill Book Co. Inc., New York, N.Y., 1964.
3. Dean, L. A., and Polos, J. A., *Frequency of Failures to Meet Firm Loads for the Pacific Northwest Hydroelectric System*, Bonneville Power Administration, Division of Power Supply, Portland, Oreg., 1983, (unpublished report).
4. Fiering, M. B, "Statistical Analysis of the Reservoir Storage-Yield Relations," *Operations Research in Water Quality Management*, H. A. Thomas, Jr., and R. P. Burden, Eds., Harvard Water Resources Group, Cambridge, Mass., 1963.
5. Fiering, M. B, *Streamflow Synthesis*, Harvard University Press, Cambridge, Mass., 1967.
6. Gumbel, E. J., "The Calculated Risk in Flood Control," *Applied Science Research*, Vol. 5A, 1955, The Hague, The Netherlands, pp. 272–280.
7. Gumbel, E. J., "The Return Period of Flood Flows," *Annals of Mathematical Statistics*, Vol. 12, No. 2, 1941, pp. 163–190.
8. Haan, C. T., *Statistical Methods in Hydrology*, The Iowa State University Press, Ames, Iowa, 1977.
9. Hashimoto, T., Stedinger, J. R., and Loucks, D. P., "Robustness of Water Resource Systems," *Water Resources Research*, Vol. 18, No. 1., Feb., 1982, pp. 21–26.
10. Hirsch, R. M., "Synthetic Hydrology and Water Supply Reliability," *Water Resources Research*, Vol. 15, No. 6, Dec., 1979, pp. 1603–1615.
11. Jackson, B. B., "Markov Mixture Models for Drought Lengths," *Water Resources Research*, Vol. 11, No. 1, Feb., 1975, pp. 64–74.
12. Klemes, V., "Applied Storage Theory in Evolution," *Advances in Hydroscience*, Vol. 12, 1979, pp. 79–141.
13. Klemes, V., "Reliability Estimates for a Storage Reservoir with Seasonal Input," *Journal of Hydrology*, Vol. 7, No. 2, 1969, pp. 198–216.

14. Klemes, V., "Reliability of a Water Supply Performed by Means of a Storage Reservoir Within a Limited Period of Time," *Journal of Hydrology*, Vol. 5, 1967, pp. 70–92.
15. Klemes, V., "Some Problems in Pure and Applied Stochastic Hydrology," *Miscellaneous Publication No. 1275*, Agricultural Research Service, U.S. Dept. of Agriculture, Washington, D.C., 1974.
16. Klemes, V., Srikanthan, R., and McMahon, T. A. "Long-Memory Flow Models in Reservoir Analysis: What is their Practical Value?" *Water Resources Research*, Vol. 17, No. 3, Jan., 1981, pp. 737–751.
17. Linsley, R. K., and Franzini, J. B., *Water Resources Engineering*, McGraw-Hill Book Co., New York, N.Y., 1979, pp. 157–158.
18. Linsley, R. K., Kohler, M. A., and Paulhus, J. L. H., *Hydrology for Engineers*, McGraw-Hill Book Co., New York, N.Y., 1982.
19. Maass, A., Hufschmidt, M. M., Dorfman, R., Thomas, H. A., Marglin, A., and Fair, G. M., *Design of Water Resource Systems*, Harvard University Press, Cambridge, Mass., 1962.
20. Moran, P. A. P., "Probability Theory of Dams and Storage Systems," *Australian Journal of Applied Science*, Vol. 5, 1954.
21. Palmer, R. N., Smith, J. A., Cohon, J. L., and ReVelle, C. S., "Reservoir Management in Potomic River Basin," *Journal of the Water Resources Planning and Management Division*, ASCE, Vol. 108, No. 1, Mar., 1982, pp. 47–66.
22. Peavy, H. S., Rowe, D. R., and Tchobanoglous, G., *Environmental Engineering*, McGraw-Hill Book Co., New York, N.Y., 1985, pp. 329–331.
23. Rippl, W., "The Capacity of Storage-Reservoirs for Water-Supply," *Minutes of the Proceedings of the Institute of Civil Engineers*, Vol. LXXI, 1983, pp. 270–278.
24. Ross, S., *A First Course in Probability*, Macmillan Publishing Co., Inc., New York, N.Y., 1976.
25. Sheer, D. P., and Flynn, K., "Water Supply," *Civil Engineering*, ASCE, Vol. 53, No. 6, Jun., 1983, pp. 50–53.
26. Stedinger, J. R., Sule, B. F., and Pei, D., "Multiple Reservoir System Screening Models," *Water Resources Research*, Vol. 19, No. 6, Dec., 1983, pp. 1383–1393.
27. Steel, E. W., and McGhee, T. J., *Water Supply and Sewage*, McGraw-Hill Book Co., New York, N.Y., 1979, pp. 44–47.
28. Sudler, C. E., "Storage Required for the Regulation of Streamflow," *Transactions*, ASCE, Vol. 91, 1927, pp. 622–660.
29. Thomas, H. A., "Frequency of Minor Floods," *Journal of the Boston Society of Civil Engineers*, Vol. 35, No. 1, 1948, pp. 425–442.
30. Vesilind, P. A., and Peirce, J. J., *Environmental Engineering*, Ann Arbor Science, Ann Arbor, Mich., 1982, pp. 81–86.
31. Viessman, W., and Hammer, M. J., *Water Supply and Pollution Control*, 4th ed., Harper and Row, Publishers, Inc., New York, N.Y., 1985, pp. 73–76.
32. Viessman, W., and Welty, C., *Water Management—Technology and Institutions*, Harper and Row, Publishers, Inc., New York, N.Y., 1985, pp. 432–433.
33. Vogel, R. M., "The Variability of Reservoir Storage Estimates," thesis presented to Cornell University, at Ithaca, N.Y., in 1985, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.
34. Vogel, R. M., and Stedinger, J. R., "Generalized Storage-Reliability-Yield Relationships," *Journal of Hydrology*, Vol. 89, Jan., 1987, pp. 303–327.
35. Vogel, R. M., and Stedinger, J. R., "The Value of Stochastic Streamflow Models in Reservoir Design Applications," *Water Resources Research*, (manuscript under review).
36. Water Resources Council, *The Nation's Water Resources, 1975–2000*, Vol. 1 Summary, Water Resources Council, Washington, D.C., 1978.
37. Yen, B. C., "Risks in Hydrologic Design of Engineering Projects," *Journal of the Hydraulic Division*, ASCE, Vol. 96, No. 4, Apr., 1970, pp. 959–966.
38. Yevjevich, V., "Fluctuations of Wet and Dry Years, 2, Analysis by Serial Correlation," *Hydrology Paper No. 4*, Colorado State University, Fort Collins, Colo., 1964.