Journal of Hydrology, 89 (1987) 303–327
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

[1]

GENERALIZED STORAGE–RELIABILITY–YIELD RELATIONSHIPS

RICHARD M. VOGEL and JERY R. STEDINGER

Department of Civil Engineering, Tufts University, Medford, MA 02155 (U.S.A.)
Department of Environmental Engineering, Cornell University, Ithaca, N.Y. 14853 (U.S.A.)

(Received October 21, 1985; accepted after revision August 1, 1986)

ABSTRACT

Vogel, R.M. and Stedinger, J.R., 1987. Generalized storage–reliability–yield relationships. J. Hydrol., 89: 303–327.

Traditionally water resource engineers have employed Rippl's mass curve approach or its automated equivalent sequent peak alogrithm, in conjunction with the historical streamflow sequence to obtain a single estimate of the design capacity of a storage reservoir. More recently stochastic streamflow models have been recommended for use in deriving the probability distribution of the required capacity of a storage reservoir to maintain a prespecified release. The use of stochastic streamflow models in conjunction with the sequent peak algorithm leads to a storage–reliability–yield (S–R–Y) relationship. This study develops approximate but general expressions which describe the over-year S–R–Y relationship when annual streamflows are log normal and follow a first-order autoregressive model. These expressions were developed for three reasons: (1) to provide preliminary design capacity or yield estimates for storage reservoirs governed by over-year storage requirements; (2) to improve our understanding of the S–R–Y relationship; and (3) to facilitate Monte-Carlo experiments which examine the sampling properties of reservoir design capacity and/or yield estimates.

INTRODUCTION

Since the work of Fiering (1963, 1965, 1967) and Svanidze (1964) many investigators have employed stochastic annual streamflow models to examine the probability distribution of over-year reservoir storage capacity (for example, Burges and Linsley, 1971; Wallis and Matalas, 1972; Perrens and Howell, 1972; Lettenmaier and Burges, 1977a, b; Bayazit, 1982). In practice, within-year storage requirements often predominate. Accordingly, a variety of monthly stochastic streamflow models have been developed to investigate the combined within-year over-year storage–reliability–yield (S–R–Y) relationship (for example, Lawrance and Kottegoda, 1977; Hirsch, 1979; Klemes et al., 1981; Stedinger and Taylor, 1982a, b; Stedinger et al., 1985). Although studies which model the within-year storage requirements are more realistic, the S–R–Y relationships which result are difficult to generalize due to the large number of parameters associated with stochastic monthly streamflow models. Thus, general S–R–Y relationships have been developed using stochastic streamflow models, for the simple cases, such as when annual streamflows are normal and arise from a

0022-1694/87/$03.50

first-order autoregressive model (Perrens and Howell, 1972; Bayazit, 1982) and over-year storage requirements predominate.

A primary objective of this study is to develop approximate but general S–R–Y relationships for the realistic situation when annual streamflows have a two-parameter log normal distribution and their logarithms follow a first-order autoregressive process. The derived analytic S–R–Y relationships are used in a subsequent study (Vogel and Stedinger, 1986) to examine the sampling properties of estimates of the design capacity of a storage reservoir.

THE STORAGE-YIELD RELATIONSHIP

Methods available for determining the storage-yield relationship from a streamflow record may be broadly classified into (1) sequential, and (2) nonsequential procedures. Sequential analysis requires routing of the complete streamflow record (or synthetic traces based thereupon) through the reservoir system while accounting for the necessary outflows which may include: water supply, evaporation, seepage losses, minimum downstream releases and other operations. The U.S. Army Corps of Engineers (1967, p. 6; 1975, p. 6.12) emphasized that "sequential analysis is the most accepted method for determining reservoir storage requirements in the United States". Although non-sequential procedures have been advocated for design purposes (McMahon and Mein, 1978, p. 169) these procedures have seen limited use in design applications in the U.S. and are not considered here. The U.S. Army Corps of Engineers (1975, section 6.02) describe the use of non-sequential procedures.

The most commonly used sequential procedure is the mass curve introduced by Rippl (1883) or its automated equivalent sequent peak algorithm. The sequent peak algorithm, developed by Thomas and Burden (1963), is a rather complex algorithm. Loucks (1970) framed the sequent peak algorithm as a linear programming problem (Loucks et al., 1981, pp. 235–236). Louck's (1970) algorithm applied to a sequence of annual streamflows Q_i, $i = 1, \ldots, N$ may be described in symbols by:

$$S = \text{Maximum}\,[S_i] \qquad \text{for } i = 1, \ldots, KN \tag{1}$$

subject to:

$$S_i = \begin{cases} S_{i-1} + \alpha\mu - Q_i & \text{if positive} \\ 0 & \text{otherwise:} \end{cases}$$

$$Q_{i+KN} = Q_i$$

where S_i = storage capacity required at the beginning of period i; Q_i = annual streamflow, year i, $i = 1, \ldots, N$; μ = mean annual streamflow (MAF); α = demand as a fraction of the MAF; K = indicator variable equal to 1 or 2; and N = length of available streamflow record.

The sequent peak algorithm, advanced by Thomas and Burden (1963), uses $K = 2$. Thus the sequence of required storages, S_i, are computed over the period $i = 1, \ldots, 2N$ which is accomplished by repeating the sequence of streamflows. This double-cycling algorithm is used to take care of the situation when the critical low flow sequence occurs at the end of the planning period. The double-cycling sequent peak algorithm generates the steady-state solution to the problem of determining the minimum storage required over an N-year planning period to supply the desired yield of $\alpha\mu$ with no shortages ($\alpha\mu < \bar{Q}$). Although the introduction of double cycling is often attributed to Thomas and Burden, Klemes (1979b, p. 138) points out that it has been used in the past, starting with Stupecky in 1909. A more detailed discussion of Rippl's mass curve technique or its automated equivalent sequent peak algorithm may also be found in Klemes (1978, 1979a).

The sequent peak algorithm in eqn. (1) is presented here to clarify the procedure used to generate S–R–Y relationships in this study. This study employs the double-cycling algorithm ($K = 2$) as opposed to the single-cycling algorithm ($K = 1$) used, for example, by Burges and Linsley (1971), Troutman (1978) and Bayazit (1982). Vogel (1985) provides a comparison of the impact of using the single- versus the double-cycling sequent peak algorithm; he documents situations in which the two procedures yield substantially different results.

A note of caution is appropriate here. When stochastic streamflow models are employed to generate the distribution of S for reservoirs on streams which are highly regulated (i.e., α close to 1) it becomes possible to generate flow sequences with average flow values less than μ leading to infeasible solutions to the problem posed in eqn. (1). Fortunately this situation is only encountered when α approaches unity leading to a storage reservoir of enormous and impractical dimensions. In realistic and practical situations, feasible solutions to eqn. (1) are almost always possible.

REVIEW OF THE LITERATURE

Empirical relations

The first relationships between the over-year design storage capacity, yield, and characteristics of the inflows for a reservoir were developed by Hazen (1914). Hazen's tables provide an approximation to over-year reservoir storage capacity requirements based upon knowledge of the coefficient of variation of the inflows, C_v, and the level of development. Unfortunately his tables were based upon the rather sparse records of annual streamflow available in 1914. Moreover tables developed for one region are not necessarily applicable to another.

Hurst (1951) developed algebraic expressions which relate the required over-year storage S, to the mean μ and variance σ^2 of the inflows as well as the level of development α. Hurst's relationships were of the form:

$$\frac{S}{\sigma} = \exp\left[a + bm\right]\left(\frac{N}{2}\right)^k \tag{2}$$

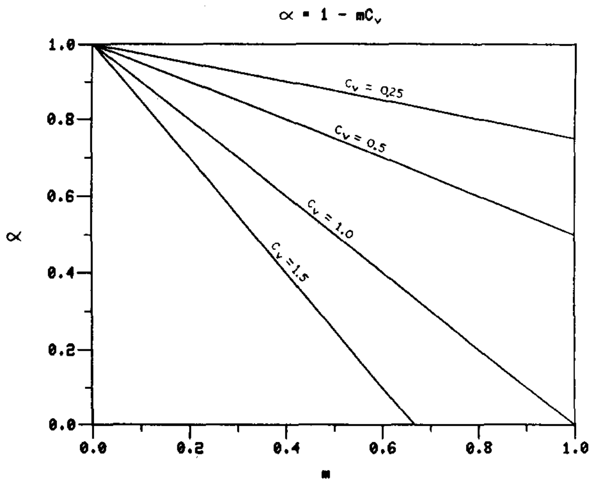


Fig. 1. The level of development α as a function of the standardized inflow m and the coefficient of variation of the inflows C_v.

or:

$$\frac{S}{\sigma} = (a + b\sqrt{m})\left(\frac{N}{2}\right)^k \tag{3}$$

where $m = (1 - \alpha)\ \mu/\sigma$ and a, b and k are constants.

The constant k in eqns. (2) and (3) is the well known Hurst coefficient. Hurst applied the single-cycling sequent peak algorithm to sequences of streamflow, precipitation, temperature, tree ring and varve records to obtain single estimates of S which in turn were used to obtain estimates of the constants a, b and k using graphical curve fitting procedures. These relationships, like Hazen's, are subject to the pitfalls of using a single streamflow trace discussed in Fiering (1967, p. 7) and Vogel and Stedinger (1986) as opposed to using stochastic streamflow models. Most importantly, Hurst (1951) and Hurst et al. (1965) identified the form of these general relationships. However, the expressions in eqns. (2) and (3) are only reasonable approximations over the range considered by Hurst ($0.3 \leqslant m \leqslant 0.8$). Actually m could be any non-negative number since μ and σ are non-negative and the demand as a fraction of the mean annual flow, α, is usually in the interval (0, 1).

Another practical consequence of Hurst's work was his use of the non-dimensional parameter m which has subsequently found use in both analytic investigations in "Water Storage Theory" (Gomide, 1975; Troutman, 1978; Pegram et al., 1980) and Monte-Carlo investigations of the S–R–Y relationship (Perrens and Howell, 1972; Bayazit, 1982). Figure 1 depicts the relationship:

$$\begin{aligned} m &= (1 - \alpha)\frac{\mu}{\sigma} \\ &= \frac{(1 - \alpha)}{C_v} \end{aligned} \tag{4}$$

Thus m may be thought of as the standardized inflow; that is, the mean net inflow ($\mu - \alpha\mu$) standardized by the scale parameter of the inflows. Although m could be any non-negative number, m is normally in the interval (0, 1) for over-year storage problems. For example in the U.S., values of C_v in the range $0.1 \leqslant C_v \leqslant 0.5$ are common, which corresponds to development levels in the range of 50–100% of the mean annual flow (MAF) when m is taken to range from 1.0 to 0.0 respectively. Values of m in excess of unity correspond to situations in which within-year storage requirements become significant relative to over-year storage requirements.

Analytic relations

Hurst's (1951) development of over-year storage-yield relationships using a variety of different natural and simulated records initiated the field of "Water Storage Theory" which has been summarized by Pegram et al. (1980). The subset of "Water Storage Theory" applicable to this study consists of analytic expressions for the distribution of over-year storage based upon a complete specification of the inflow to the reservoir. Recent contributions were made by Gomide (1975) and Troutman (1978), both having derived the probability density function (pdf) of S and its first two moments μ_s, and σ_s^2, which results from application of the single-cycling sequent peak algorithm to realistic models of annual streamflows. For example, when the annual streamflows, Q, are normally distributed and follow a first-order autoregressive model:

$$Q_{i+1} = \mu + \varrho_1 (Q_i - \mu) + \varepsilon_i \sigma \sqrt{1 - \varrho_1^2} \tag{5}$$

where the ε_i are independent normal random disturbances with mean 0 and variance 1, Gomide (1975) derived the pdf of S and its first moment, μ_s. The resulting expressions were so complex that Gomide only presented his results graphically for ϱ_1 equal to 0.0, 0.2, 0.5; m equal to 0.0 (full regulation) and N ranging from 0 to 100 years. For $\varrho_1 = 0$ in eqn. (5) Gomide (1975) presents the pdf of S and its first two moments μ_s and σ_s, when the reservoir is partially regulated (m = 0.0, 0.25, 0.50, 1.0), for planning periods which range from 0 to 50 years. Again the derived expressions were too complex to report; instead Gomide presents the results of selected cases in his figs. 5.7, 5.8, 5.9 and 5.10.

Similarly Troutman (1978) derived the mean μ_s, and variance σ_s of the asymptotic distribution of S when $m = 0$ (full regulation) and inflows are described by an AR(1) log normal model:

$$X_{i+1} = \mu_x + \varrho_1 (x) (X_i - \mu_x) + \varepsilon_i \sigma_x \sqrt{1 - \varrho_1^2(x)} \tag{6}$$

where $X_i = \ln [Q_i]$, the ε_i are independent normal disturbances with mean 0, variance 1 and μ_x, σ_x^2 and $\varrho_1 (x)$ are the mean, variance and serial correlation of the log transformed streamflows. No analytic expressions have been developed for the pdf or moments of the steady-state required storage obtained using the double-cycling sequent peak algorithm.

Experimental relations

With the advent of computer technology and the introduction of stochastic streamflow models by Thomas and Fiering (1962) and Svanidze (1964) it was possible to use Monte-Carlo procedures to generate the pdf of S corresponding to a range of values of μ, σ^2, ϱ_1, N and α. Fiering (1963; 1965; 1967) generated 200 traces of gamma and normally distributed streamflows from a first-order autoregressive model which were subsequently analyzed by the double-cycling sequent peak algorithm. Fiering tabulated the mean μ_s and standard deviation σ_s of the distribution of steady-state storage corresponding to planning periods, N, equal to 10, 25, 50 and 100 years, annual autocorrelations, ϱ_1, of 0.0, 0.1 and 0.2 and levels of development, α, of 0.8, 0.9 and 1.0. Since Fiering held the coefficient of variation of the inflows, C_v, constant at 0.25, his simulations correspond to values of m equal to 0.0, 0.4 and 0.8 ($m = (1 - \alpha)/C_v$).

Burges and Linsley (1971) generated the complete pdf of over-year required storage capacity using eqn. (5) in conjunction with the single-cycling sequent peak algorithm. They suggest that at least 1000 streamflow traces are required to specify the pdf of S.

Perrens and Howell (1972) used a stochastic streamflow model to develop generalized S–R–Y relationships in graphical form. Perrens and Howell used eqn. (5) in conjunction with an algorithm which determined the number of times a reservoir of fixed capacity with a fixed m failed to deliver the target yield over a 14,400-year interval. This algorithm, which allows failures, differs substantially from the sequent peak algorithm. Perrens and Howell plotted S/σ (which they termed standardized capacity) versus $-m$ (which they termed standardized use) to form general graphical relations.

More recently, Bayazit (1982) generated AR(1) normal flows from eqn. (5) and used the single-cycling sequent peak algorithm to obtain general graphical relationships between S/σ and m. Bayazit compared his estimates of the pdf of S, μ_s, and σ_s with Gomide's (1975) analytic expressions for μ_s and σ_s of the *asymptotic* pdf of S; the agreement was poor as is to be expected. Gomide also derived μ_s and σ_s for non-asymptotic cases. In fact, a comparison of Bayazit's results for $\varrho_1 = 0$ to Gomide's non-asymptotic analytic results for normal independent inflows depicted in his figs. 5.7, 5.8, 5.9 and 5.10, show excellent agreement. Thus Bayazit simply compared his results to the wrong analytic results when he concluded that his results differ from Gomide's.

A review of the literature reveals that general graphical S–R–Y relations exist when inflows are AR(1) normal and a single-cycling sequent peak algorithm is used (Bayazit, 1982; Gomide, 1975). However, such relations have not been generalized when the double-cycling sequent peak algorithm is used nor when inflows are AR(1) log normal. No general and convenient S–R–Y relationships exist in analytic form for any of the cases discussed here.

MONTE-CARLO EXPERIMENT

A Monte-Carlo experiment was performed to determine the generalized over-year S–R–Y relationship when inflows are AR(1) log normal (eqn. (6)) and a double-cycling sequent peak algorithm is used. Vogel (1985) summarizes additional experiments for the case when inflows are AR(1) normal and both single-and double cycling sequent peak algorithms are employed. To explore these S–R–Y relationships in detail requires consideration of a wide range of possible inflow parameters, reliabilities and yields.

In the following experiments values of m were chosen to span the region $0.1 \leqslant m \leqslant 1.0$ which includes most over-year storage problems of interest in the U.S. For example, in the U.S., values of C_v in the range $0.2 \leqslant C_v \leqslant 0.5$ are considered quite common which corresponds to demand levels in the range 50–98% of the mean annual flow (MAF) when m is taken to range from 1.0 to 0.1, respectively. Values of the first-order autocorrelation, ϱ_1, of annual streamflows are generally positive (Matalas, 1963). In a study of annual streamflow data from 140 gaging stations around the world, with records of at least 37 years, Yevjevich (1964) found estimates of ϱ_1 to vary from -0.4 to 0.75; however, most values were in the range 0.0–0.4. Similarly, using 106 basins in New England, Vicens et al. (1975) found the mean and standard deviation of estimates of ϱ_1 to be 0.22 and 0.14, respectively. Thus a reasonable range of ϱ_1 for these experiments is $0.0 \leqslant \varrho_1 \leqslant 0.5$. This should capture most cases of practical interest.

Distribution of over-year reservoir storage capacity

Burges and Linsley (1971) suggest that the cumulative distribution function (cdf) of estimates of S derived using the single-cycling sequent peak algorithm with AR(1) normal inflows is well described by the extreme value type I (Gumbel) probability distribution. Burges and Linsley only examined the cases when $\varrho_1 = 0.2$, $N = 40$ and $m = 0.2$, 1.0 and 1.2.

In this section we seek to determine a cdf which approximates the distribution of over-year storage derived with the double-cycling sequent peak algorithm with AR(1) log normal inflows and realistic combinations of m, N and ϱ_1. To accomplish this task, 1000 streamflow traces were generated corresponding to planning periods, N, equal to 20 and 60 years, annual autocorrelations, ϱ_1, equal to 0.0 and 0.3 and values of the standardized inflow, m, equal to 0.1, 0.3, 0.5, 0.7 and 1.0. The double-cycling sequent peak algorithm was applied to each streamflow trace to produce 1000 over-year storage estimates. An extreme value type I (EVI) distribution, a normal (N) distribution and 2- and 3-parameter log normal distributions (LN2, LN3) were fit to each set of 1000 over-year storage estimates as described in the following sections. Filliben's (1975) probability plot correlation coefficient (PPCC) test for normality, extended by Vogel (1986), was used to test the hypothesis that S is distributed N, LN2 and LN3. The PPCC test is quite flexible and may be used to test distributional

TABLE 1

Probability plot correlation coefficient test statistic values for the distribution of over-year storage

ϱ_1	N	m	C_v = 0.25			C_v = 0.50		
			α	$S \sim$ LN3	$S \sim$ EV1	α	$S \sim$ LN3	$S \sim$ EV1
0.0	20	0.1	0.975	0.9954	0.9688	0.95	0.9973	0.9732
0.0	20	0.3	0.925	0.9947	0.9496	0.85	0.9951	0.9545
0.0	20	0.5	0.875	0.9975[a]	0.9778	0.75	0.9976[a]	0.9849
0.0	20	0.7	0.825	0.9989[a]	0.9988[a]	0.65	0.9989[a]	0.9988[a]
0.0	20	1.0	0.750	0.9876[a]	0.9973[a]	0.50	0.9694[a]	0.9918
0.0	60	0.1		0.9984[a]	0.9649		0.9989[a]	0.9666
0.0	60	0.3		0.9989[a]	0.9825		0.9988[a]	0.9832
0.0	60	0.5		0.9989[a]	0.9974[a]		0.9988[a]	0.9973[a]
0.0	60	0.7		0.9994[a]	0.9982[a]		0.9993[a]	0.9983[a]
0.0	60	1.0		0.9982[a]	0.9987[a]		0.9978[a]	0.9959[a]
0.3	20	0.1		0.9967	0.9803		0.9967	0.9829
0.3	20	0.3		0.9976[a]	0.9650		0.9976[a]	0.9652
0.3	20	0.5		0.9993[a]	0.9828		0.9994[a]	0.9834
0.3	20	0.7		0.9997[a]	0.9971[a]		0.9984[a]	0.9980[a]
0.3	20	1.0		0.9801	0.9958		0.9647	0.9917
0.3	60	0.1		0.9979[a]	0.9792		0.9978[a]	0.9833
0.3	60	0.3		0.9991[a]	0.9850		0.9994[a]	0.9777
0.3	60	0.5		0.9993[a]	0.9970[a]		0.9994[a]	0.9871
0.3	60	0.7		0.9996[a]	0.9974[a]		0.9997[a]	0.9972[a]
0.3	60	1.0		0.9992[a]	0.9971[a]		0.9988[a]	0.9981[a]

Note: This table is based upon 1000 replicate experiments.
[a] Cases for which one could not reject the hypothesis that $S \sim$ LN3 or $S \sim$ EV1 using the probability plot correlation coefficient test with a type I error probability of 0.01.

hypotheses for any one or two-parameter distribution which exhibits a fixed shape. The PPCC test has the additional and attractive feature that it is invariant to the parameter estimation procedure. Vogel (1986) summarizes the use of the PPCC test for the normal, log normal and Gumbel distributional hypotheses. In addition, Vogel (1986) provides tables of the PPCC test statistic for sample sizes of interest in Monte-Carlo experiments.

Results

Table 1 summarizes the computed values of $\hat{r}$ when LN3 and EV1 distributions were fit to the values of S and annual flows are AR(1) log normal and a double-cycling sequent peak algorithm is employed. Table 2 displays the results for the normal (N) and 2-parameter lognormal (LN2) distributions which did not fit nearly as well. If we allow type I errors 1% of the time, when the null hypothesis is true, then, from Vogel (1986), the test statistic $\hat{r}$ must exceed 0.9975 and 0.9933 for the normal (lognormal) and Gumbel PPCC tests, respec-

TABLE 2

Probability plot correlation coefficient test statistic values for the distribution of over-year storage

ϱ_1	N	m	$C_v = 0.25$			$C_v = 0.50$		
			α	$S \sim$ N	$S \sim$ LN2	α	$S \sim$ N	$S \sim$ LN2
0.0	20	0.1	0.975	0.8863	0.9952	0.95	0.8954	0.9952
0.0	20	0.5	0.875	0.9157	0.8505	0.75	0.9298	0.9891
0.0	20	1.0	0.750	0.9773	0.6740	0.50	0.9774	0.7329
0.0	60	0.1		0.8801	0.9926		0.8825	0.9921
0.0	60	0.5		0.9573	0.9988[a]		0.9574	0.9985[a]
0.0	60	1.0		0.9718	0.9935		0.9820	0.7807
0.3	20	0.1		0.9115	0.9977[a]		0.9178	0.9978[a]
0.3	20	0.5		0.9164	0.9010		0.9177	0.9852
0.3	20	1.0		0.9531	0.7256		0.9477	0.7763
0.3	60	0.1		0.9075	0.9957		0.9046	0.9963
0.3	60	0.5		0.9531	0.9995[a]		0.9547	0.9995[a]
0.3	60	1.0		0.9541	0.9980[a]		0.9581	0.8540

Note: This table is based upon 1000 replicate experiments.

[a] Cases for which one could not reject the hypothesis that $S \sim$ LN2 using the probability plot correlation coefficient test with a type I error probability of 0.01.

tively. On the basis of Table 1 we must reject the null hypothesis that the distribution of S is EV1 in 24 of the 40 cases represented (60% of the trials). Similarly we must reject the null hypothesis that the distribution of S is LN3 for only 8 of the 40 cases (20% of the cases). It should be noted that the PPCC test was developed for a two parameter normal (or log normal) distribution, yet it is being used here for a three parameter log normal distribution. Use of the percentage points of $\hat{r}$ given in Vogel (1986) for testing the null hypothesis that S is distributed LN3 will lead to fewer rejections of the null hypothesis than one would otherwise anticipate.

Even though we must reject the null hypothesis that S is distributed EV1 or LN3 in many instances, an examination of Table 1 reveals that the estimated probability plot correlation coefficients are always in excess of 0.94, and in most cases are in excess of 0.99. Such large values of $\hat{r}$ indicate that the distributions of S generally are well "approximated" by both the EV1 and LN3 distributions, even if the small discrepancies are statistically significant. In the following analysis the LN3 distribution is used to describe the distribution of S. The values of $\hat{r}$ for the LN3 fits were always in excess of 0.994.

Furthermore, the values of $\hat{r}$ for the LN3 PPCC test are always greater than the values of $\hat{r}$ for the EV1 PPCC test when the standardized inflows m are 0.7 or lower. The only situation in which an LN3 distribution did not provide a good approximation to the distribution of S is when the planning period N is equal to 20 and the standardized inflow m = 1.0. However in these instances the EV1 distribution provides a reasonable and adequate alternative.

Probability plot correlation coefficients were also computed when values of S were fit to normal (N) and 2-parameter log normal (LN2) distributions; the results are displayed in Table 2. In general, the normal distribution provides a poor approximation to the distribution of S for all the cases considered. Although in some instances the LN2 distribution provides a reasonable approximation to the distribution of S, Table 2 documents many instances in which $\hat{r}$ is less than 0.9. These situations occur primarily at the lower demand levels (i.e., $m = 1$). The N, LN2 and EV1 distributions do not appear to fit the distribution of S nearly as well as the LN3 distribution.

Development of storage–reliability–yield relationships

Unfortunately the general analytic form of the S–R–Y relationship is at present unknown when inflows are AR(1) log normal and the double-cycling sequent peak algorithm is employed. Here a Monte-Carlo procedure is used to generate the storage, reliability and yield data. Multivariate regression equations were then fit to that data to obtain approximate but general S–R–Y relationships.

Perrens and Howell (1972), Gomide (1975) and Bayazit (1982) found that when inflows are AR(1) normal, a general relationship exists between the standardized storage capacity S/σ, m, N, and ϱ_1. Thus when the inflows are AR(1) normal, S's distribution is only a function of m, N, ϱ_1 and σ. For AR(1) normal inflows, when the demand level is given, m and σ completely characterize the location, scale and shape of the distribution of inflows since the shape of a normal distribution is fixed. When the inflows are AR(1) log normal an additional parameter, C_v, is required to characterize the shape of the distribution of inflows since an LN2 distribution exhibits non-zero skew.

Since the distribution of S may be well approximated by an LN3 distribution, the pth quantile of the distribution of standardized storage capacity is:

$$\frac{S_p}{\sigma} = \tau_s + \exp(\mu_l + z_p \sigma_l) \tag{7}$$

where:

$$\mu_l = \ln\left[\frac{\mu_s - \tau_s}{\sqrt{1 + \dfrac{\sigma_s^2}{(\mu_s - \tau_s)^2}}}\right]$$

$$\sigma_l^2 = \ln\left[1 + \frac{\sigma_s^2}{(\mu_s - \tau_s)^2}\right]$$

and μ_s, σ_s^2, and τ_s are the mean, variance and lower bound of the distribution of S/σ and z_p is the pth quantile of a standard normal distribution. This study develops a general relationship between S_p/σ and the planning period length

N, the development level m, the autocorrelation of annual streamflows, ϱ_1 and the coefficient of variation of the inflows, C_v. This is accomplished by developing regression equations between μ_s, σ_s, and τ_s and the values of N, m, ϱ_1, C_v and α. A Monte-Carlo experiment was performed to generate values of the dependent variables corresponding to reasonable and practical combinations of the independent variables.

The coefficient of variation, C_v, of most mean annual flow series in the U.S. is in the range $0.1 \leqslant C_v \leqslant 0.5$. As discussed earlier, practical combinations of the other explanatory (independent) variables will likely fall within the intervals: $0.1 \leqslant m \leqslant 1.0$; $20 \leqslant N \leqslant 100$; and $0.00 \leqslant \varrho_1 \leqslant 0.5$.

In the following experiments, values of S/σ were estimated for: $m = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0$; $N = 20, 40, 60, 80, 100$; $\varrho_1 = 0.0, 0.1, 0.2, 0.3, 0.4, 0.5$; and $C_v = 0.1, 0.2, 0.3, 0.4, 0.5$.

The chosen sample space includes 1500 different combinations of the planning parameters. Of those 1500 different combinations, there are only 30 combinations of the streamflow model parameters ϱ_1 and C_v. For each of the 30 combinations of ϱ_1 and C_v, fifty-thousand 100-year streamflow traces were generated from an AR(1) log normal model. Each streamflow trace was subjected to ten double-cycling sequent peak algorithms for each value of N and m. Therefore a total of 1500 different combinations of the dependent and independent variables were generated. The fifty-thousand streamflow traces were required to insure that estimates of μ_s, σ_s and τ_s were of sufficient precision; with fifty-thousand streamflow traces, 95% confidence intervals for estimates of μ_s and σ_s were always within $\pm 1\%$ and $\pm 3.5\%$, respectively, of the estimated values. Ninety-five percent confidence intervals for S_p obtained from eqn. (7) were always within $\pm 1\%$ of the estimated values as long as $0.10 \leqslant p \leqslant 0.95$. Thus the sampling error in estimates of S_p/σ, μ_s, σ_s, and τ_s may be neglected in this study. However the cost of the Monte-Carlo analysis was high because effectively $30 \times 50000 \times 100 = 150$ million generated annual streamflows were required to produce 300 estimates of the dependent variable S/σ.

Regression model selection

The best practical regression equations for μ_s, σ_s and τ_s in terms of m, N, ϱ_1 and C_v had to be identified. On the basis of empirical work by Hurst (1951), theoretical work by Gomide (1975) and Troutman (1978), and experimental work by Svanidze (1964), Perrens and Howell (1972) and Bayazit (1982), one would expect these relationships to be highly nonlinear and fairly complex. Multivariate linear regression procedures were employed to fit several families of non-linear models. Three linear regression equations for ln (μ_s), ln (σ_s) and ln (τ_s) were thus obtained in terms of functions of the explanatory variables m, N, ϱ_1 and C_v.

Suppose that $W_1, W_2, \ldots, W_r$, are functions of m, N, ϱ_1 and C_v, and represent the complete set of variables from which the equations are to be chosen. They may include functions such as cross-products, logarithms, inverses, squares and products of logarithms. In this classical problem one must strike a balance between: (a) including as *many* W_i's as possible to insure that the resulting models produce sufficiently reliable predictions; and (b) including as few W_i's as possible to insure that the resultant models are not too complex for their intended purpose(s).

A compromise is usually required. These models may be used for two purposes: (1) to provide planning level estimates of S_p for use in preliminary feasibility studies; and (2) for research applications which seek to understand the general S–R–Y relationship and to determine the sampling properties of S_p estimates.

Each regression equation must contain at least four explanatory variables (m, N, ϱ_1, and C_v): the regression model simple enough to satisfy criterion (b) if (1) is the intended purpose would have at least five parameters. However, it was found that more than five parameters were generally required to satisfy criterion (a) if either (1) *or* (2) are the intended purposes. We selected a model with as few parameters as is necessary to provide sufficiently reliable predictions for both purposes (1) and (2).

A number of procedures are available for choosing the "best regression equation" from a set of possible regressions. For each set of explanatory variables $W_1, \ldots, W_r$ the stepwise regression procedure was used to generate regression equations with 5–10 parameters. Ten was thought to be an upper limit to assure compliance with criterion (b). Each fitted regression equation results in a set of residuals which are uniquely determined by the set of independent variables used to develop the equations. Therefore the residuals are not random errors since they could be reproduced. Nevertheless, the residuals were evaluated to see if they behaved as random errors with:

(1) $\mathrm{Var}(\varepsilon_i) = \sigma^2$ independent of i

(2) $\mathrm{Cov}(\varepsilon_i, \varepsilon_k) = 0 \quad i \neq k$

(3) $\varepsilon_i \sim$ Normal

These characteristics were examined using standard plots discussed in chapter 3 of Draper and Smith (1981). In addition, Filliben's (1975) probability plot correlation test statistic, $\hat{r}$, was computed to determine if the residuals were essentially normally distributed.

Ordinarily regression equations which manifest the above characteristics can be compared on the basis of standard indices such as the coefficient of determination R^2, the residual mean square s^2, or Mallow's C_p statistic (see Draper and Smith (1981), pp. 299–302). In this study a different approach is taken; each reasonable set of regression equations for μ_s, σ_s, and τ_s was used to estimate S_p/σ in eqn. (7) for a range of m, N, ϱ_1 and C_v. The results were compared with the estimates of S_p/σ derived from the Monte-Carlo experiments. The best set of regression equations was chosen as that set which provided the best estimate of S_p/σ on the basis of graphical comparisons.

TABLE 3

Summary of coefficients and t-ratios of regression equations for μ_s, σ_s^2, and τ_s

θ	μ_s		σ_s		τ_s	
	$\hat{\theta}$	$\hat{t}$	$\hat{\theta}$	$\hat{t}$	$\hat{\theta}$	$\hat{t}$
a	0.237	38	−5.92	−283	0.467	14
b	−1.33	−161	4.89	180	−0.0398	−55
c	1.81	447	−0.000958	−31	0.189	51
d	−1.03	−47	10.0	101	−0.0332	−42
e	0.00621	60	−0.0342	−34	−0.00407	−40
f	0.369	205	−0.520	−196	0.00803	38
g	−0.0562	−81	0.421	219	−0.00403	−20
h	0.100	145	0.0	–	0.0	–

Note: Each regression equation is based upon a sample of size 1500.

Model estimation

The final equations are:

$$\mu_s' = \exp(a + bm)\, \alpha^c m^{m(d\varrho_1 + eN)} N^{(f + g \ln[m])} \left(\frac{1+\varrho_1}{1-\varrho_1}\right)^{h \ln[N]} \tag{8}$$

$$\sigma_s'^2 = \exp\left[a + b\alpha + \frac{cN}{m} + \left(\frac{d}{N} + \frac{e}{m}\right)\left(\frac{1+\varrho_1}{1-\varrho_1}\right)\right] N^{f \ln[m]} \left(\frac{1+\varrho_1}{1-\varrho_1}\right)^{g \ln[N]} \tag{9}$$

$$\tau_s' = a\varrho_1 + \left(bN + \frac{c(1+\varrho_1)}{1-\varrho_1}\right)\ln[m] + N\left[d + \frac{e}{m} + fm \ln[N] + g \ln\left(\frac{1+\varrho_1}{1-\varrho_1}\right)\right] \tag{10}$$

TABLE 4

Summary statistics for regression equations for μ_s, σ_s^2 and τ_s

Summary statistic	μ_s	σ_s^2	τ_s
R^2	99.95	99.72	93.89
s_r	0.01847	0.09377	0.1044
p	8	7	7
C_p	7.47	7.15	6.30
r	0.989	0.985	0.991

Note: The reported statistics are based upon 1500 samples.

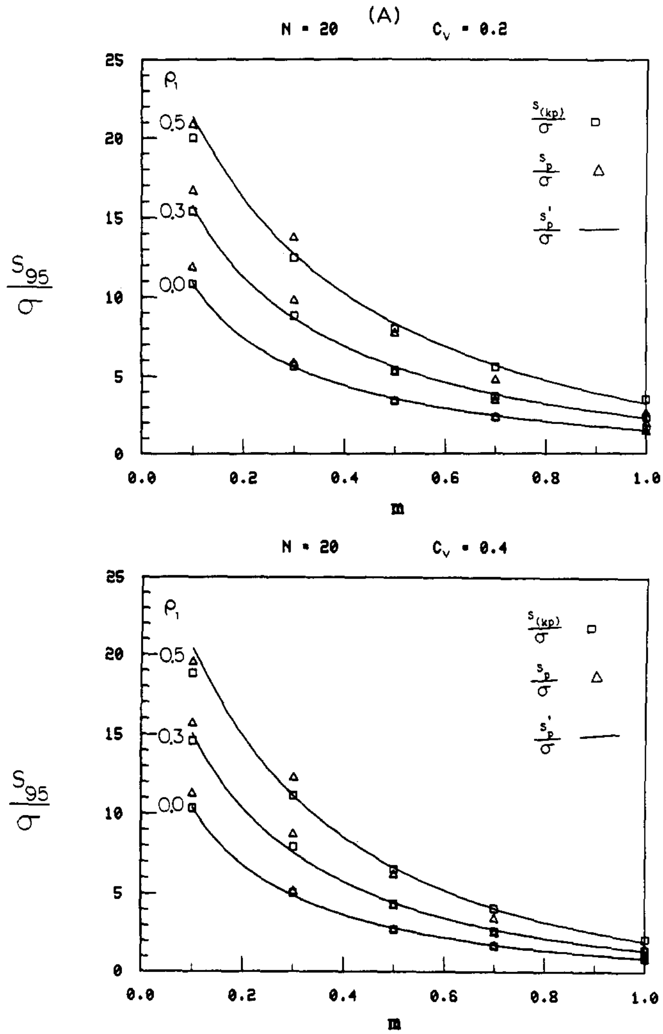
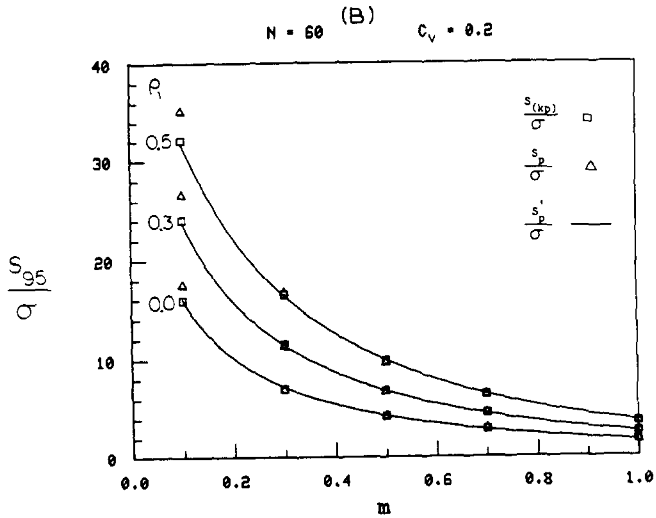
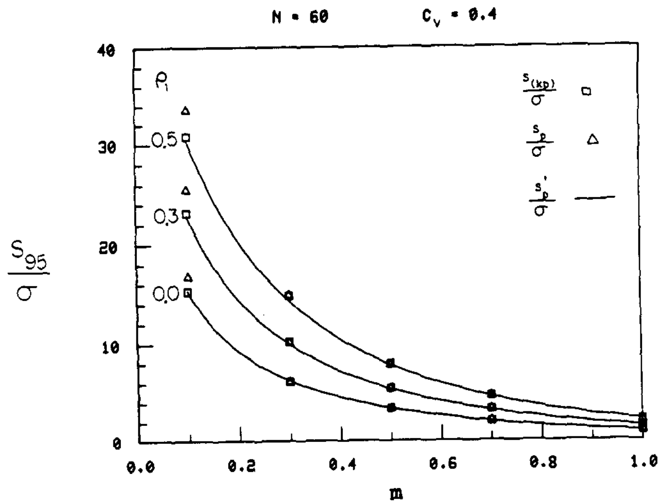


Fig. 2. Comparison of regression estimate S'_p/σ, nonparametric estimate $S_{(kp)}/\sigma$ and parametric estimate S_p/σ of the standardized storage capacity as a function of N, m, ϱ_1 and C_v.

where the parameter estimates are given in Table 3. It was found that use of the demand level, α, as an explanatory variable instead of C_v led to more reasonable relationships. For a given value of m, C_v and α are directly related by eqn. (4), thus eqns. (8) and (9) are strictly functions of m, ϱ_1 and C_v, even though C_v does not appear in these expressions. Values of the t-ratio:

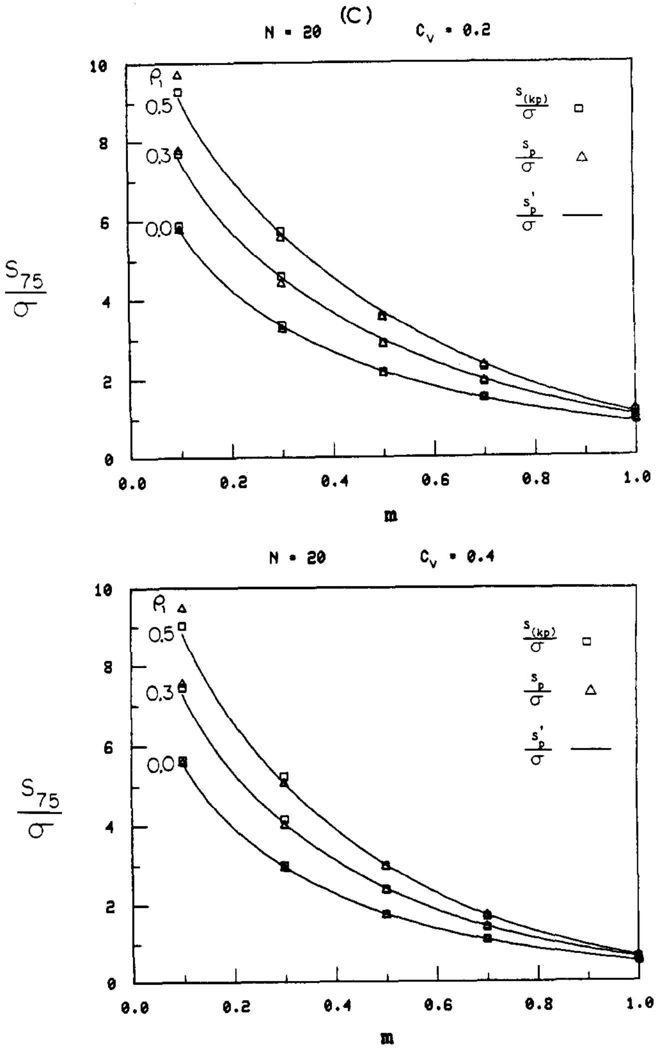
$$\hat{t} = \frac{\hat{\theta}}{\sqrt{\mathrm{Var}\,(\hat{\theta})}} \tag{11}$$





where $\hat{\theta}$ is an estimate of the regression coefficient θ and Var $(\hat{\theta})$, the standard estimate of the estimator's variance if the residuals were independent and homoscedastic random errors, are also displayed in Table 3. Because the t-ratios in Table 3 are always in excess of 14, all the estimated regression coefficients would be statistically significant.

The coefficient of determination, R^2, the square root of the residual mean square, s_r, the number of model parameters, p, Mallow's C_p statistic and Filliben's PPCC for the residuals $\hat{r}$, are summarized in Table 4. In all cases, estimates of C_p are relatively close to p hence the models would be termed "adequate" by Draper and Smith (1981). Models with fewer parameters than shown in Table 4 led to dramatic increases in the C_p statistic indicating "biased" equations. Comparing the computed values of Filliben's $\hat{r}$ in Table 4 with the percentage



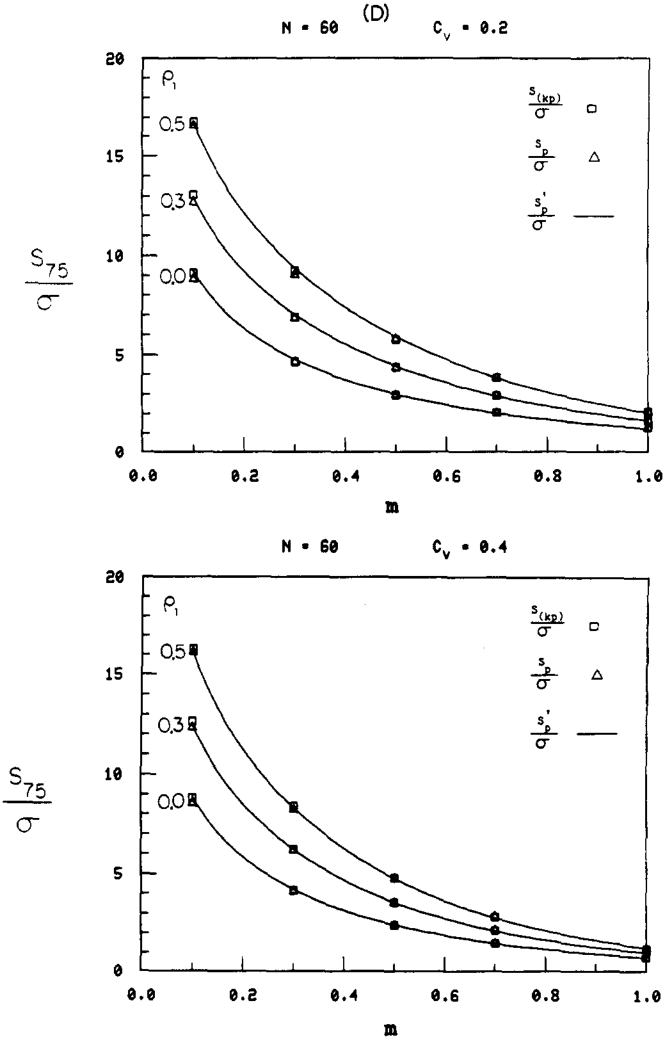
points of $\hat{r}$ in Vogel (1986) we must reject the hypothesis that the residuals are normal at the 1% level; still the values of $\hat{r}$ always exceeded 0.985, so one can conclude that the residuals are approximately normal.

Model prediction errors

The regression eqns. (8)–(10) are complex non-linear functions and the statistics in Tables 3 and 4 provide little insight into the behavior of these models. For this purpose, graphical comparisons are made among the standardized storage capacity S_p/σ computed using three different techniques:

(1) In the Monte-Carlo experiments, k = 50,000 values of S/σ were

(D)



generated for each combination of N, m, ϱ_1 and C_v. A reasonable nonparametric estimate of the pth quantile of S/σ is given by $S_{(kp)}/\sigma$ which is the kpth largest observation.

(2) Efficient estimates of μ_1, σ_1^2 and τ_s in eqn. (7) can be obtained by employing the log-space/quantile-lower bound estimators suggested by Stedinger (1980). Those estimators may be substituted into eqn. (7) to obtain S_p/σ as a quantile of an LN3 distribution. The only difference between S_p/σ and $S_{(kp)}/\sigma$ is the distributional assumption.

(3) Of particular interest are estimates of S_p'/σ derived from the use of the regression equations. The estimator S_p'/σ is derived from substitution of μ_s', σ_s', τ_s' in eqns. (8)–(10) into eqn. (7).

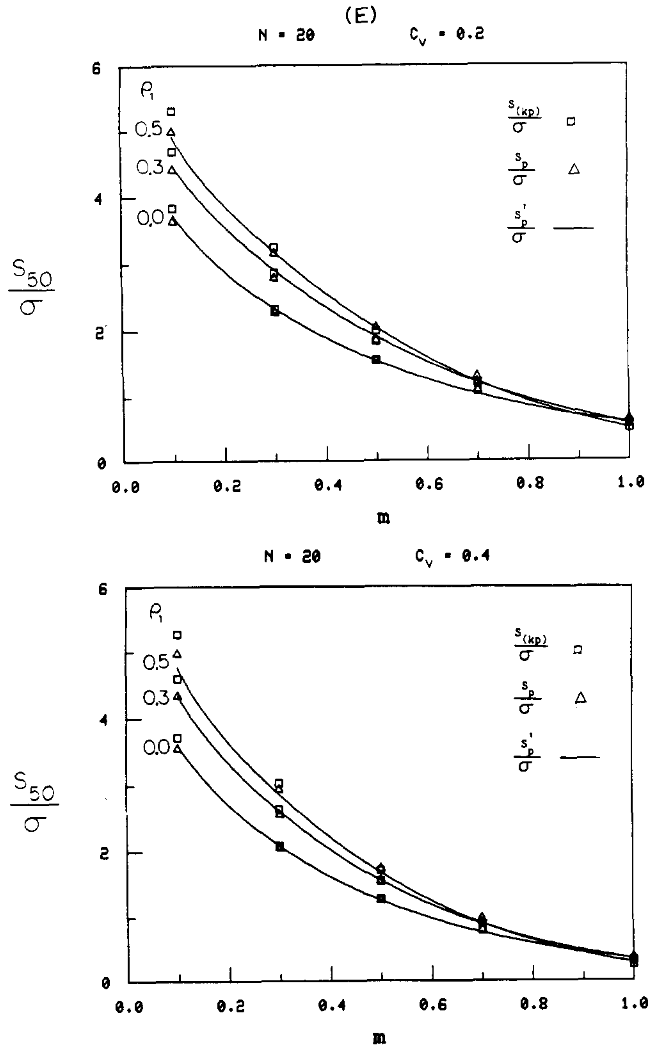
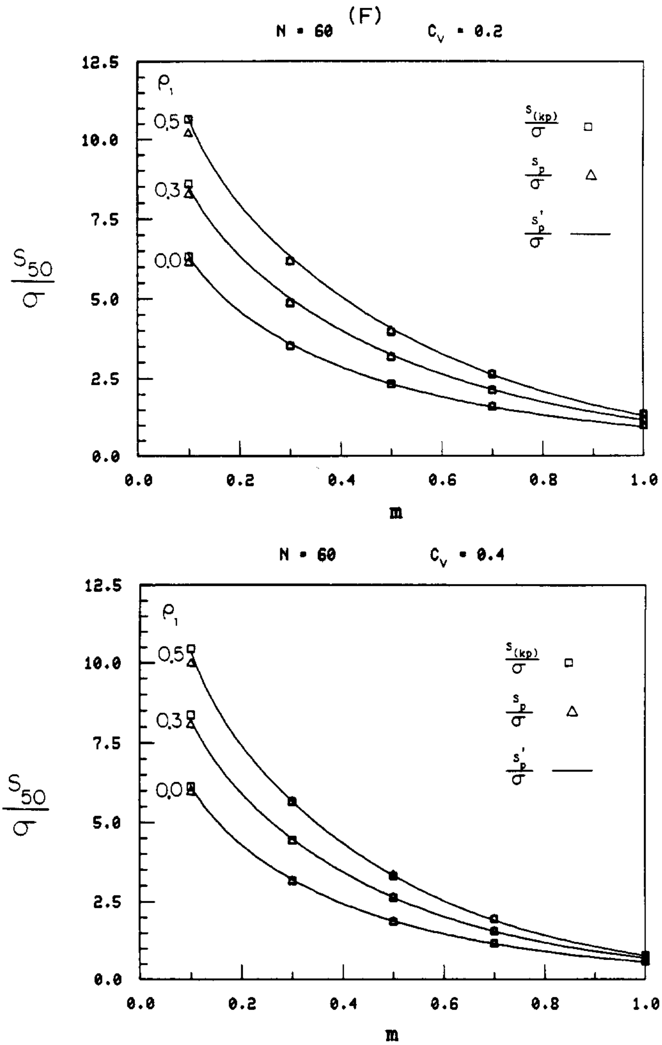


Figure 2 displays $S_{(kp)}/\sigma$, S'_p/σ and S_p/σ as a function of m for $N = 20, 60$; $\varrho_1 = 0.0, 0.3, 0.5$; $p = 0.50, 0.75, 0.95$; and $C_v = 0.2, 0.4$. Unfortunately space only permits a few selected combinations of the independent variables N, m, ϱ_1 and C_v. These combinations are chosen to depict regions in the middle and on the edges of the sample space.

As expected from Table 1 the differences between $S_{(kp)}/\sigma$ and S_p/σ are generally very small; the LN3 distribution provides a reasonable approximation to the distribution of S. As to the regression equations, S'_p/σ provides an excellent approximation to S_p/σ for all values of N, m, ϱ_1 and C_v in Fig. 2. except when $p = 0.5$, $N = 20$, $\varrho_1 \geqslant 0.3$ and $m \leqslant 0.3$ or $m \geqslant 0.7$. These instances raise questions as to the precision of S'_p/σ when one is estimating the median of S with



planning periods of length 20. To evaluate the errors in the estimation of S'_p/σ in a more comprehensive fashion, Tables 5 and 6 display the bias in estimates of S_p defined as $(S'_p - S_p)/S_p$. In general, the regression errors tend to increase as p decreases because S'_p/σ depends heavily upon τ'_s for small p and prediction errors associated with τ_s are large, as evidenced by the relatively low R^2 value listed in Table 4.

It is evident from Tables 5 and 6 that S'_p provides a reasonable approximation to S_p in the region $0.25 \leqslant p \leqslant 0.95$, $0.0 \leqslant m \leqslant 1.0$ and $0.2 \leqslant C_v \leqslant 0.4$. In fact within this region, errors in estimation of S_p using S'_p are less than $\pm 8\%$ of the value of S_p and usually less than $\pm 3\%$. If an estimate of S_p were required for a planning period $N = 20$, then S'_p provides a reasonable estimate of S_p as long

TABLE 5

Bias associated with regression estimates of S_p when $C_v = 0.2$

ϱ_1	N	m	$(S'_p - S_p)/S_p$				
			$p = 0.05$	$p = 0.25$	$p = 0.50$	$p = 0.75$	$p = 0.95$
0.0	20	0.10	0.038	0.031	0.023	0.013	−0.004
0.0	20	0.50	−0.015	0.004	0.001	−0.010	−0.035
0.0	20	1.00	−0.195	−0.046	−0.004	0.027	0.064
0.0	40	0.10	0.023	0.004	−0.002	−0.002	0.003
0.0	40	0.50	−0.005	−0.009	−0.011	−0.012	−0.012
0.0	40	1.00	−0.038	0.038	0.052	0.053	0.039
0.0	60	0.10	−0.004	−0.011	−0.013	−0.012	−0.008
0.0	60	0.50	0.001	−0.008	−0.011	−0.011	−0.009
0.0	60	1.00	−0.080	−0.006	0.019	0.033	0.041
0.0	80	0.10	−0.036	−0.030	−0.025	−0.020	−0.011
0.0	80	0.50	−0.004	−0.005	−0.005	−0.003	−0.001
0.0	80	1.00	−0.097	−0.039	−0.009	0.016	0.045
0.0	100	0.10	−0.062	−0.048	−0.037	−0.026	−0.009
0.0	100	0.50	0.015	0.012	0.011	0.010	0.009
0.0	100	1.00	−0.113	−0.057	−0.025	0.004	0.039
0.3	20	0.10	0.072	0.059	0.040	0.017	−0.023
0.3	20	0.50	−0.106	−0.051	−0.038	−0.033	−0.035
0.3	20	1.00	0.108	−0.040	−0.038	−0.020	0.017
0.3	40	0.10	0.023	0.018	0.017	0.016	0.017
0.3	40	0.50	−0.091	−0.064	−0.044	−0.023	0.008
0.3	40	1.00	0.040	0.038	0.029	0.018	−0.002
0.3	60	0.10	0.005	0.002	0.002	0.004	0.009
0.3	60	0.50	−0.069	−0.045	−0.031	−0.018	−0.001
0.3	60	1.00	−0.040	0.001	0.012	0.016	0.015
0.3	80	0.10	−0.011	−0.011	−0.008	−0.003	0.007
0.3	80	0.50	−0.039	−0.022	−0.015	−0.010	−0.006
0.3	80	1.00	−0.088	−0.036	−0.013	0.005	0.024
0.3	100	0.10	−0.031	−0.025	−0.018	−0.009	0.005
0.3	100	0.50	0.001	0.005	0.004	0.001	−0.006
0.3	100	1.00	−0.120	−0.063	−0.033	−0.008	0.023
0.5	20	0.10	0.310	0.141	0.072	0.014	−0.067
0.5	20	0.50	0.092	−0.001	−0.023	−0.034	−0.040
0.5	20	1.00	1.270	0.029	−0.014	0.009	0.077
0.5	40	0.10	−0.007	0.006	0.009	0.009	0.007
0.5	40	0.50	−0.104	−0.066	−0.047	−0.031	−0.009
0.5	40	1.00	0.207	0.074	0.036	0.011	−0.015
0.5	60	0.10	−0.036	−0.019	−0.009	0.001	0.013
0.5	60	0.50	−0.067	−0.049	−0.037	−0.024	−0.006
0.5	60	1.00	0.068	0.054	0.038	0.019	−0.013

TABLE 5 (Continued)

ϱ_1	N	m	$(S_p' - S_p)/S_p$				
			$p = 0.05$	$p = 0.25$	$p = 0.50$	$p = 0.75$	$p = 0.95$
0.5	80	0.10	−0.049	−0.033	−0.020	−0.007	0.013
0.5	80	0.50	−0.038	−0.022	−0.017	−0.015	−0.016
0.5	80	1.00	−0.011	0.016	0.018	0.015	0.003
0.5	100	0.10	−0.068	−0.046	−0.029	−0.011	0.016
0.5	100	0.50	0.028	0.018	0.007	−0.006	−0.027
0.5	100	1.00	−0.070	−0.022	−0.005	0.004	0.010

as $0.3 \leqslant m \leqslant 0.7$ and $0.25 \leqslant p \leqslant 0.95$, as evidenced by Fig. 2 and Tables 5 and 6; otherwise substantial estimation errors could occur.

SUMMARY

This study presents general over-year Storage–Reliability–Yield (S–R–Y) relationships in an analytic form. Approximate but general expressions are provided for evaluating the quantiles of the distribution of over-year storage capacity as a function of the inflow parameters μ, σ^2, ϱ_1, the demand level α, and the planning period N for log normal inflows. Vogel (1985) provides similar approximate but general S–R–Y relationships for the case when inflows are normal and both single- and double-cycling sequent peak algorithm's are employed ($K = 1$ and 2 in eqn. (1)).

The S–R–Y relations developed here are designed to mimic the task of a stochastic hydrologist. Although these expressions appear complex, their use is much simpler, quicker and cheaper than the alternative of generating synthetic streamflow traces, routing those sequences through the sequent peak algorithm and finally fitting the distribution of required reservoir storage capacity.

Fiering (1963), Svanidze (1964), Burges and Linsley (1971), Wallis and Matalas (1972), Perrens and Howell (1972), Lettenmaier and Burges (1977a, b) Bayazit (1982) and Vogel (1985) have all examined properties of the S–R–Y relationship as a function of combinations of the parameters N, α, C_v and ϱ_1. Those studies in addition to Fig. 2 document the relative importance of the parameters N, m, ϱ_1 and C_v in the determination of S_p. Vogel (1985) and Vogel and Stedinger (1986) document the sampling properties of estimates of S_p treating μ, σ^2, ϱ_1 and α as random variables. In particular, Vogel (1985) and Vogel and Stedinger (1986) use the S–R–Y relationships developed here to show that estimates of the design capacity of a reservoir using stochastic streamflow models are more precise than those obtained by critical period planning which results in a single estimate of the design capacity based upon the single historic streamflow trace.

TABLE 6

Bias asssociated with regression estimates of S_p when C_v = 0.4

ϱ_1	N	m	$(S_p' - S_p)/S_p$				
			p = 0.05	p = 0.25	p = 0.50	p = 0.75	p = 0.95
0.0	20	0.10	0.062	0.040	0.025	0.011	−0.009
0.0	20	0.50	0.025	0.019	0.009	−0.005	−0.028
0.0	20	1.00	−0.154	−0.126	−0.073	−0.014	0.073
0.0	40	0.10	0.031	0.008	0.001	−0.001	0.003
0.0	40	0.50	0.009	−0.005	−0.008	−0.008	−0.002
0.0	40	1.00	−0.038	0.012	0.023	0.027	0.025
0.0	60	0.10	−0.002	−0.008	−0.009	−0.009	−0.005
0.0	60	0.50	0.003	−0.009	−0.011	−0.008	0.000
0.0	60	1.00	−0.102	−0.026	−0.004	0.005	0.006
0.0	80	0.10	−0.038	−0.030	−0.024	−0.017	−0.007
0.0	80	0.50	−0.003	−0.005	−0.004	−0.001	0.006
0.0	80	1.00	−0.117	−0.057	−0.031	−0.012	0.007
0.0	100	0.10	−0.062	−0.048	−0.036	−0.023	−0.003
0.0	100	0.50	0.007	0.010	0.011	0.012	0.011
0.0	100	1.00	−0.132	−0.071	−0.043	−0.022	−0.002
0.3	20	0.10	0.115	0.081	0.052	0.020	−0.031
0.3	20	0.50	−0.034	−0.013	−0.009	−0.011	−0.017
0.3	20	1.00	1.239	−0.113	−0.143	−0.095	0.016
0.3	40	0.10	0.043	0.033	0.027	0.022	0.016
0.3	40	0.50	−0.060	−0.040	−0.022	−0.001	0.030
0.3	40	1.00	0.120	0.001	−0.020	−0.023	−0.013
0.3	60	0.10	0.020	0.014	0.012	0.011	0.012
0.3	60	0.50	−0.048	−0.025	−0.011	0.002	0.019
0.3	60	1.00	0.005	−0.012	−0.016	−0.017	−0.017
0.3	80	0.10	0.000	0.000	0.002	0.005	0.011
0.3	80	0.50	−0.023	−0.006	0.002	0.008	0.013
0.3	80	1.00	−0.072	−0.048	−0.036	−0.025	−0.012
0.3	100	0.10	−0.022	−0.016	−0.009	−0.001	0.011
0.3	100	0.50	0.010	0.017	0.019	0.017	0.012
0.3	100	1.00	−0.131	−0.080	−0.056	−0.036	−0.013
0.5	20	0.10	0.354	0.173	0.093	0.020	−0.085
0.5	20	0.50	0.199	0.051	0.013	−0.006	−0.019
0.5	20	1.00	4.563	0.037	−0.137	−0.095	0.054
0.5	40	0.10	0.027	0.031	0.026	0.018	0.003
0.5	40	0.50	−0.050	−0.032	−0.018	−0.003	0.020
0.5	40	1.00	0.487	0.027	−0.039	−0.048	−0.021
0.5	60	0.10	−0.008	0.001	0.006	0.010	0.015
0.5	60	0.50	−0.034	−0.020	−0.009	0.003	0.020
0.5	60	1.00	0.175	0.024	−0.013	−0.029	−0.036

TABLE 6 (Continued)

ϱ_1	N	m	$(S'_p - S_p)/S_p$				
			$p = 0.05$	$p = 0.25$	$p = 0.50$	$p = 0.75$	$p = 0.95$
0.5	80	0.10	−0.025	−0.015	−0.007	0.003	0.018
0.5	80	0.50	−0.012	0.004	0.009	0.010	0.008
0.5	80	1.00	0.033	−0.004	−0.016	−0.025	−0.033
0.5	100	0.10	−0.050	−0.031	−0.016	0.000	0.023
0.5	100	0.50	0.050	0.041	0.031	0.018	−0.003
0.5	100	1.00	−0.064	−0.043	−0.036	−0.031	−0.026

Although the S–R–Y relations developed here may be useful tools, users should beware of their limitations. Use of the S–R–Y relationship is only reasonable when estimates of m ϱ_1, C_v and N remain within the specified bounds. Table 5 illustrates the potential errors which may arise. As long as one's estimate of m is less than unity, one may expect that over-year storage is a relevant problem.

REFERENCES

Bayazit, M., 1982. Ideal reservoir capacity as a function of yield and risk. J. Hydrol., 58: 1–9.

Burges, S.J. and Linsley, R.K., 1971. Some factors influencing required reservoir storage, J. Hydraul. Div., ASCE, 97(HY7): 977–991.

Draper N.R. and Smith, H., 1981. Applied Regression Analysis. Wiley, New York, N.Y., 2nd ed.

Fiering, M.B., 1963. Statistical analysis of the reservoir storage-yield relations. In: H.A. Thomas Jr. and R.P. Burden, (Editors), Operations Research in Water Quality Management. Harv. Water Resour. Group, Cambridge, Mass.

Fiering, M.B., 1965. The nature of the storage-yield relationship. Proc. Symp. on Streamflow Regulation for Quality Control, Public Health Serv. Publ. 999-WP-30.

Fiering, M.B., 1967. Streamflow Synthesis. Harvard University Press, Cambridge, Mass.

Filliben, J.J., 1975. The probability plot correlation coefficient test for normality. Technometrics, 17(1): 111–117.

Gomide, F.L.S., 1975. Range and Deficit Analysis Using Markov Chains. Hydrol. Pap. 79, Colorado State Univ., Fort Collins, Colo.

Hazen, A., 1914. Storage to be provided in impounding resevoirs for municipal water supply. Trans. Am. Soc. Civ. Eng., 77: 1539.

Hirsch, R.M., 1979. Synthetic hydrology and water supply reliability. Water Resour. Res., 15(6): 1603–1615.

Hurst, H.E., 1951. Long-term storage capacities of reservoirs, Trans. Am. Soc. Civ. Eng., 116: 776.

Hurst, H.E., Black, R.P. and Simaika, Y.M., 1965. Long-Term Storage. Constable & Co., London.

Kartvelishvili, N.A., 1969. Theory of Stochastic Processes in Hydrology and River Runoff Regulation. Isr. Program Sci. Transl., TT68-50478. Distributed through National Technical Information Service, Springfield, Va.

Klemes, V., 1978. Discussion on "Sequent peak procedure: Minimum reservoir capacity subject to constraint on final storage" by K.W. Potter. Water Res. Bull., 14(4): 991–993.

Klemes, V., 1979a. Storage mass-curve analysis in a systems-analytic perspective. Water Resour. Res., 15(2): 359–370.

Klemes, V., 1979b. Applied storage theory in evolution. Adv. Hydrosci., 12: 79–141.

Klemes, V., Srikanthan, R. and McMahon, T.A., 1981. Long-memory flow models in reservoir analysis: What is their practical value? Water Resour. Res., 17(3): 737–751.

Lawrance, A.J. and Kottegoda, N.T., 1977. Stochastic modeling of riverflow time series. J.R. Stat. Soc., Ser. A., 140(1): 1–47.

Lettenmaier, D.P. and Burges, S.J., 1977a. Operational assessment of hydrologic models of long-term persistence. Water Resour. Res., 13(1): 113–124.

Lettenmaier, D.P. and Burges, S.J., 1977b. An operational approach to preserving skew in hydrologic models of long-term persistence. Water Resour. Res., 13(2): 281–290.

Loucks, D.P., 1970. Some comments on linear decision rules and chance constraints. Water Resour. Res., 6(2): 668–671.

Loucks, D.P., Stedinger, J.R. and Haith, D.A., 1981. Water Resource Systems Planning and Analysis. Prentice-Hall, Englewood Cliffs, N.J.

Matalas, N.C., 1963. Autocorrelation of rainfall and streamflow minimums. U.S. Geol. Surv. Pap. 434-B.

McMahon, T.A. and Mein, R.G., 1978. Reservoir Capacity and Yield. Elsevier, New York, N.Y.

Pegram, G.G.S., Salas, J.D., Boes, D.C. and Yevjevich, V., 1980. Stochastic properties of water storage. Hydrol. Pap. 100, Colorado State Univ., Fort Collins, Colo.

Perrens, S.J. and Howell, D.T. 1972. Effects of serially correlated inflows on reservoir behavior. Water Resour. Res., 8(3): 642–647.

Rippl, W., 1883. The capacity of storage-reservoirs for water-supply. Minutes Proc. Inst. Civ. Eng., 71: 270–278.

Stedinger, J.R., 1980. Fitting log normal distributions to hydrologic data. Water Resour. Res., 16(3): 481–490.

Stedinger, J.R. and Taylor, M.R., 1982a. Synthetic streamflow generation, 1: Model verification and validation. Water Resour. Res., 18(4): 909–918.

Stedinger, J.R. and Taylor, M.R., 1982b. Synthetic streamflow generation, 2: Effect of parameter uncertainty. Water Resour. Res., 18(4): 919–924.

Stedinger, J.R., Pei, D. and Cohn, T.A., 1985. A condensed disaggregation model for incorporating parameter uncertainty into monthly reservoir simulations. Water Resour. Res., 21(5): 665–675.

Stupecky, F., 1909. Construction of storage reservoirs. Ceska Matice Technicka, 1 (in Czech).

Svanidze, G.G.., 1964. Principles of Estimating River Flow Regulation by the Monte Carlo Method. Metsniereba, Tbilisi, 271 pp. (in Russian; see Kartvelishvili (1969) for English translation).

Thomas Jr., H.A. and Burden, R.P., 1963. Operations Research in Water Quality Management. Harv. Water Resour. Group, Cambridge, Mass., pp. 1–17.

Thomas, H.A. and Fiering, M.B., 1962. Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation. In: A. Maass, M.M. Hufschmidt, R. Dorfman, H.A. Thomas Jr., S.A. Marglin and G.M. Fair (Editors), Design of Water Resource Systems. Harvard University Press, Cambridge, Mass.

Troutman, B.M., 1978. Reservoir storage with dependent, periodic net inputs. Water Resour. Res., 14(3): 395–401.

U.S. Army Corps of Engineers, 1967. Reservoir Storage-Yield Procedures. Hydrol. Eng. Center, Davis, Calif.

U.S. Army Corps of Engineers, 1975. Reservoir Yield, 8. Hydrol. Eng. Methods Water Resour. Devel.

Vicens, G.J., Rodriguez-Iturbe, I. and Schaake Jr., J.C., 1975. Bayesian generation of synthetic streamflows. Water Resour. Res., 11(6): 827–838.

Vogel, R.M., 1985. The variability of reservoir storage estimates. Ph.D. dissertation, Cornell Univ., Ithaca, N.Y.

Vogel, R.M., 1986. The probability plot correlation coefficient test for the normal, log normal and gumbel distributional hypotheses. Water Resour. Res., 22(4): 587–590.

Vogel, R.M. and Stedinger, J.R., 1986. The value of stochastic streamflow models in reservoir design applications. Water Resour. Res., in prep.

Wallis, J.R. and Matalas, N.C., 1972. Sensitivity of reservoir design to the generating mechanism of inflows. Water Resour. Res., 8(3): 634–641.

Yevjevich, V., 1964. Fluctuations of wet and dry years, 2: Analysis by serial correlation. Hydrol. Pap. 4, Colorado State Univ., Fort Collins, Colo.