Accuracy of Low-Flow Characteristics Estimated by Correlation of Base-Flow Measurements

Manual of Hydrology: Part 2. Low-Flow Techniques

Accuracy of Low-Flow Characteristics Estimated by Correlation of Base-Flow Measurements

By CLAYTON H. HARDISON *and* MARSHALL E. MOSS .

With a section on OUTLINE OF DERIVATIONS

By EDWARD J. GILROY

Manual of Hydrology: Part 2. Low-Flow Techniques

CONTENTS

ILLUSTRATIONS

TABLES

SYMBOLS

[Used in sections by Hardison and Moss]

b	Slope of regression line defined by sample statistics.
β	Slope of the true regression line.
B_G	Logarithm of a base flow at the ungaged site; $\bar{B}_G$ is mean B_G of a sample.
B_U	Logarithm of a base flow at the ungaged site; $\bar{B}_U$ is mean B_U of a sample.
F	Factor used in computing N_U.
I_v	Standard deviation of logarithms of annual low flows.

I_{vG} I_v for gaged site.

I_{vU} I_v for ungaged site.

k Factor for adjusting $s_{B,U}/s_{B,G}$ to I_{vU}/I_{vG}.

K Factor used in computing F.

M Number of independent base-flow measurements.

N Number of annual low flows.

N_G Number of annual low flows observed at gaged site.

N_U Accuracy of estimated T-year low flow at ungaged site, in equivalent years of record.

r Coefficient of correlation computed from sample statistics; $\hat{r}$ is r adjusted for consistency with other graphically computed statistics.

ρ Coefficient of correlation of the population values of B_G and B_U.

R Factor relating standard error of T-year low to I_v and N.

$s_{B,G}$ Standard deviation of the logarithms of the sample of base flows at the gaged site that were used in the regression; $\hat{s}_{B,G}$ is graphically determined $s_{B,G}$ adjusted for consistency with other graphically determined statistics.

$s_{B,U}$ Standard deviation of the logarithms of the sample base flows at the ungaged site that were used in the regression; $s'_{B,U}$ is $s_{B,U}$ based on $s_{B,G}$; $\hat{s}_{B,U}$ is mean of $s'_{B,U}$ and graphically determined $s_{B,U}$.

SE_R Standard error of estimate of a regression, equals $\sqrt{(1-r)^2}s_{B,U}$.

SE_T Standard error of a T-year low flow computed by fitting a log-Pearson Type III distribution to a sample of annual low flows.

$SE_{T,G}$ SE_T at gaged site.

$SE_{T,U}$ SE_T at ungaged site.

$\sigma_{B,G}$ Standard deviation of the population of B_G's of which the base-flow measurements are a sample.

$\sigma_{B,U}$ Standard deviation of the population of B_U's of which the base-flow measurements are a sample.

$\sigma_{T,G}$ Same as $SE_{T,G}$.

T Average interval in years between recurrences of annual lows of equal or greater severity.

V_G Equals $(SE_{T,G})^2$, the variance of an estimated T-year low at the gaged site.

V_R Equals $(SE_R)^2$, the variance about the line of regression.

V_U Equals $(SE_{T,U})^2$, the variance of an estimated T-year low at the ungaged site.

$X_{T,G}$ Logarithm of a T-year low flow at the gaged site.

$\hat{X}_{T,G}$ $X_{T,G}$ estimated from a sample of annual low flows at the gaged site.

$X_{T,U}$ Logarithm of T-year low flow at the ungaged site.

$\tilde{X}_{T,U}$ Estimator of $X_{T,U}$.

z Distance from the mean in standard deviation units; equals $(\hat{X}_{T,G}-\bar{B}_G)/s_{B,G}$.

ACCURACY OF LOW-FLOW CHARACTERISTICS ESTIMATED BY CORRELATION OF BASE-FLOW MEASUREMENTS

By CLAYTON H. HARDISON and MARSHALL E. MOSS

ABSTRACT

The number of discharge measurements required to define an acceptable relation between the base flow at a stream gaging station and that at an ungaged site depends on the statistics of the regression, the accuracy goal, and the length of record at the gaging station used in the regression. Equations and graphs are presented for evaluating the accuracy of a low-flow characteristic estimated from such a relation in terms of the number of years of record that would be required at the ungaged site to give an estimate of comparable accuracy. An outline of the derivations of basic equations is presented in a separate section of the report.

INTRODUCTION

In regression analysis of streamflow characteristics and basin parameters, Thomas and Benson (1970) found that the characteristics of low flow cannot be adequately estimated from basin parameters obtainable from present maps. As an alternative, Riggs (1965) recommends that low-flow frequency characteristics be estimated by relating the logarithms of the base flow at an ungaged site to the logarithms of the base flow at a gaged site through the use of discharge measurements made during periods of base flow. This paper presents a procedure for evaluating the accuracy of low-flow characteristics estimated from such a relation. The low-flow characteristics considered are T-year low flows such as the 2-year 7-day (median annual) low flow and the 10-year 7-day low flow. (T is the average time in years between recurrences of annual low flows smaller than that given as the T-year low flow.) The length of record at the gaged site, number of base-flow measurements, and the standard error of estimate of the regression are all used as parameters. All standard errors, SE, and variance, V, are expressed in log units. Subscripts G and U are used for gaged and ungaged sites, respectively, and

subscripts B, L, and T are used for base flow, annual low flow, and T-year low flow, respectively.

The first part of the report makes practical application of the mathematical equations derived by Edward J. Gilroy in the section "Outline of Derivations." The notation required for the rigorous mathematical derivation in this latter section is considerably more extensive than that shown in the "List of Symbols," and in some instances the notation is different. One difference in notation is that in the first part of the report the logarithmic standard deviation of annual low flows is designated I_v (for index of variability) both for the population and for a sample, whereas in the latter part the logarithmic standard deviation of the population is designated σ_L and that of the sample s_L. Other differences are that in the first part, V_G and V_U are used in place of $\sigma^2_{T,\,G}$ and $\mathrm{Var}(\bar{X}_{T,\,U})$ respectively, and that for the variance about the regression line (square of the standard error of estimate of the regression) V_R is used in place of $(1-\rho^2)\sigma^2_{B,\,U}$. The implicit assumptions discussed in the first part of the report represent the authors' interpretation of the assumptions made by Gilroy in the derivation of the basic equations on which the report is based.

DERIVATION OF ACCURACY EQUATIONS

The accuracy of a regression line at a given value of x is given by

$$V_U = \frac{V_R}{M}[1+z^2]$$

in which V_R, the variance about the regression line, is the square of the standard error of estimate, M is the number of observations used to define the regression, z is the number of standard deviation units between the mean of the x's and the x at which the regression is being used, and V_U is the estimated variance of the corresponding estimate of y about the y that would be given by the true regression line based on population parameters. This equation is valid only if the independent variable is fixed; that is, only if the same values of the independent variable would be used in every repetition of an experiment. In the development of a relation between base-flow measurement at a gaged and an ungaged site, however, the independent variable is not fixed. Furthermore, when the relation is entered with a low-flow characteristic at the gaged site (independent variable) to obtain the corresponding characteristic at the ungaged site, the time-sampling error at the gaged site contributes to the error in the low-flow characteristic estimated for the ungaged site. For these conditions and under the five assumptions given in the next paragraph, Gilroy (see equation

30 in the section "Outline of Derivations") gives the variance of low flows estimated at the ungaged site as

$$\text{Var}(\bar{X}_{T, U}) = \frac{V_R}{M}\left[1 + \frac{1}{M-3} + \frac{z^2 M}{M-3} + \frac{V_G}{\sigma^2_{B, G}}\left(\frac{M}{M-3}\right)\right] + \beta^2 V_G, \quad (1)$$

in which $\text{Var}(\bar{X}_{T, U})$ is the variance of the estimate of the T-year low flow at the ungaged site, V_R is the variance of the departures from the line of regression, V_G is the variance of the T-year low flow at the gaged site, M is the number of independent base-flow measurements, $\sigma^2_{B, G}$ is the variance of the population of base flows of which the base-flow measurements are a sample, β is the slope of the true regression, and z is as previously defined.

The five assumptions implicit in the derivation of equation 1 are:
1. The lower end of the true relation between base-flow measurements is the same as the true relation between annual low flows.
2. The relation between annual low flows is the same as the relation between annual 7-day low flows.
3. The time-sampling error in the low-flow characteristic that is used to enter the regression equation is independent of the variation in the base-flow measurements used in defining the regression.
4. The logarithms of the base flows at the gaged and ungaged sites follow a bivariate normal distribution.
5. The M observations used to define the regression equation are statistically dependent estimates of the base-flow relation.

The first four of these assumptions appear to be reasonably applicable to the conditions under which equation 1 is used in this paper. The fifth assumption can be met by establishing criteria for counting only those measurements that can reasonably be assumed to be independent.

By using $s_{B, G}$ and b as estimates of $\sigma_{B, G}$ and β, respectively, equation 1 becomes

$$V_U = \frac{V_R}{M}\left[1 + \frac{1}{M-3} + \frac{z^2 M}{M-3} + \left(\frac{SE_{T, G}}{s_{B, G}}\right)^2\left(\frac{M}{M-3}\right)\right] + b^2 V_G, \quad (2)$$

in which V_U is the estimator of $\text{Var}(\bar{X}_{T, U})$ and in which $SE_{T, G} = \sqrt{V_G}$ is the standard error of the T-year low flow at the gaged site.

If $SE_{T, G}$, $s_{B, G}$, and V_G are expressed in log units, the square root of V_U as obtained from the solution of equation 2 gives $SE_{T, U}$, the standard error of the estimated low-flow characteristic at the ungaged site, in log units. As accuracy goals recently set by the U.S. Geological Survey for the collection of streamflow data (Carter and Benson, 1970) are expressed in equivalent years of record (number

of years of record required to give a low-flow characteristic of equivalent accuracy) it is desirable to present equation 2 in such a way that the relation between number of base-flow measurements, statistics of the regression, years of record at the gaged site, and the resulting accuracy at the ungaged site in equivalent years of record becomes more obvious. The remainder of this section develops equations and a graph that accomplish this objective. Readers who are interested only in using the graph and equations should skip to the example in the next section.

Letting

$$K = \frac{1+z^2}{1+\dfrac{1}{M-3}+\dfrac{z^2 M}{M-3}+\left(\dfrac{SE_{T,\,G}}{s_{B,\,G}}\right)^2\left(\dfrac{M}{M-3}\right)} \tag{3}$$

equation 2 becomes

$$V_U = \frac{V_R}{M}\left(\frac{1+z^2}{K}\right)+b^2 V_G. \tag{4}$$

If the annual low flows are from a population that can be defined by a log-Pearson Type III distribution, the standard error of a T-year low is given by

$$SE_T = R\frac{I_v}{\sqrt{N}}$$

in which R is as given in table 1, I_v is the standard deviation of the logarithms of the annual lows, and N is the number of observed annual low flows (usually the number of years of record). Thus the variance of the T-year low flow at the gaged site is given by

$$V_G = (SE_{T,\,G})^2 = R^2{}_G(I_{vG})^2/N_G \tag{5}$$

in which N_G is the actual number of years of record at the gaged site. At the ungaged site

$$V_U = R^2{}_U(I_{vU})^2/N_U \tag{6}$$

in which N_U is the number of years of record at the ungaged site that would be required to give a T-year low flow as accurate as that indicated by V_U, the variance of the estimate at the ungaged site. Substitution of equations 5 and 6 in 4 gives

$$\frac{R^2{}_U I_{vU}{}^2}{N_U} = \frac{V_R}{M}\left(\frac{1+z^2}{K}\right)+\frac{b^2 R^2{}_G I_{vG}{}^2}{N_G}. \tag{7}$$

According to Gilroy (see equation 39 in the section "Outline of Derivations") an unbiased estimate of $(I_{vU})^2$ gives

$$\left(\frac{I_{vU}}{I_{vG}}\right)^2 = \left(\frac{s_{B,\,U}}{s_{B,\,G}}\right)^2\left[r^2+\left(\frac{M-4}{M-2}\right)(1-r^2)\right]. \tag{8}$$

TABLE 1.—*Values of R for use in appraising accuracy of a T-year low estimated from observed annual lows from a log-Pearson Type III population of known skew*

[R equals $(SE_T \sqrt{N})/I$, in which SE_T is the standard error of the T-year low, in log units, I, is standard deviation of the logarithms of the annual lows, and N is number of annual lows]

T	R for indicated logarithmic coefficient of skew										
	+2.0	+1.5	+1.0	+0.5	+0.2	0	-0.2	-0.5	-1.0	-1.5	-2.0
2	0.744	0.845	0.933	0.983	0.997	1.000	0.997	0.983	0.933	0.845	0.744
5	.777	.821	.916	1.020	1.102	1.164	1.229	1.328	1.486	1.638	1.734
10	.868	.926	1.029	1.148	1.258	1.350	1.454	1.629	1.956	2.325	2.661
20	.917	1.006	1.134	1.276	1.414	1.534	1.674	1.921	2.416	3.010	3.617
25	.927	1.026	1.163	1.316	1.500	1.591	1.747	2.013	2.560	3.228	3.928
50	.947	1.075	1.246	1.433	1.608	1.763	1.950	2.288	3.006	3.903	4.896
100	.957	1.107	1.313	1.538	1.742	1.925	2.146	2.554	3.438	4.574	5.868

Assuming that the square root of $(I_{vU})^2$ gives an unbiased estimate of I_{vU}, and considering that by definition b equals $r(s_{B,\ U}/s_{B,\ G})$,

$$I_{vU} = \frac{kb}{r} I_{vG},$$

in which

$$k = \sqrt{r^2 + \left(\frac{M-4}{M-2}\right)(1-r^2)}. \tag{9}$$

Then by assuming that the coefficient of skew of the logarithms of the annual low flows is the same at both sites, $R_U = R_G = R$, and by substituting $(SE_R)^2$ for V_R, equation 7 gives

$$\frac{R^2 k^2 b^2 I^2_{vG}}{r^2 N_U} = \frac{SE_R^2}{M}\left(\frac{1+z^2}{K}\right) + \frac{b^2 R^2 I^2_{vG}}{N_G}. \tag{10}$$

For an approximate solution of equation 10, the k^2 and K terms can be replaced by 1.0 and $(M-3)/M$, respectively. (Evaluation of equations 3 and 9 shows that k^2 tends to be several percent less than unity, and the KM tends to be several percent greater than M minus 3, so that the effects of the replacements tend to cancel.) Thus

$$\frac{1}{r^2 N_U} = \frac{SE^2_R(1+z^2)}{(M-3)R^2 b^2 I^2_{vG}} + \frac{1}{N_G}.$$

Letting

$$F = \frac{(M-3)R^2}{1+z^2}\left(\frac{bI_{vG}}{SE_R}\right)^2 \tag{11}$$

and solving for N_U

$$N_U = \frac{F}{r^2\left(1+\dfrac{F}{N_G}\right)}, \tag{12}$$

in which r is the coefficient of correlation of the base-flow measurements.

For values of r less than 0.8, the answers given by equations 11 and 12 become increasingly less exact. In such cases, or if the solution is to be made by electronic computer, V_U should be computed by equation 2, and N_U should be computed by a combination of equations 6 and 8 as follows

$$N_U = \frac{R^2_U(I_{vU})^2}{V_U} = \frac{R^2_G}{V_U}\left(\frac{kbI_{vG}}{r}\right)^2 \tag{13}$$

in which R_U has been assumed to be equal to R_G and in which k is as given by equation 9.

In equations 11 and 13, R, which can be obtained from table 1, depends on the recurrence interval, T, of the low-flow characteristics

and on the skew coefficient of the logarithms of the annual low flows. The values of R given in table 1 were computed from an equation for SE_T given by Hardison (1969).

The solution of equation 12 for F values as much as 200 is given in table 2. Interpolation between the F values shown can be made from curves such as those shown in figure 14, which are plotted from the

TABLE 2.—*Accuracy of low-flow characteristics estimated by correlation of base-flow measurements*

[N_G is the number of years of record at the gaged site, N_U is the accuracy at the ungaged site in equivalent years of record, r is the coefficient of correlation, and F depends on the statistics of the correlation]

N_G	$r^2 N_U$, in years, for indicated F								
	5	10	15	20	30	40	60	100	200
10	3.33	5.00	6.00	6.67	7.50	8.00	8.57	9.09	9.52
12	3.53	5.45	6.67	7.50	8.57	9.23	10.00	10.71	11.32
15	3.75	6.00	7.50	8.57	10.00	10.91	12.00	13.04	13.95
20	4.00	6.67	8.57	10.00	12.00	13.33	15.00	16.67	18.18
25	4.17	7.14	9.37	11.11	13.64	15.38	17.65	20.00	22.22
30	4.28	7.50	10.00	12.00	15.00	17.14	20.00	23.08	26.09
40	4.44	8.00	10.91	13.33	17.14	20.00	24.00	28.57	33.33

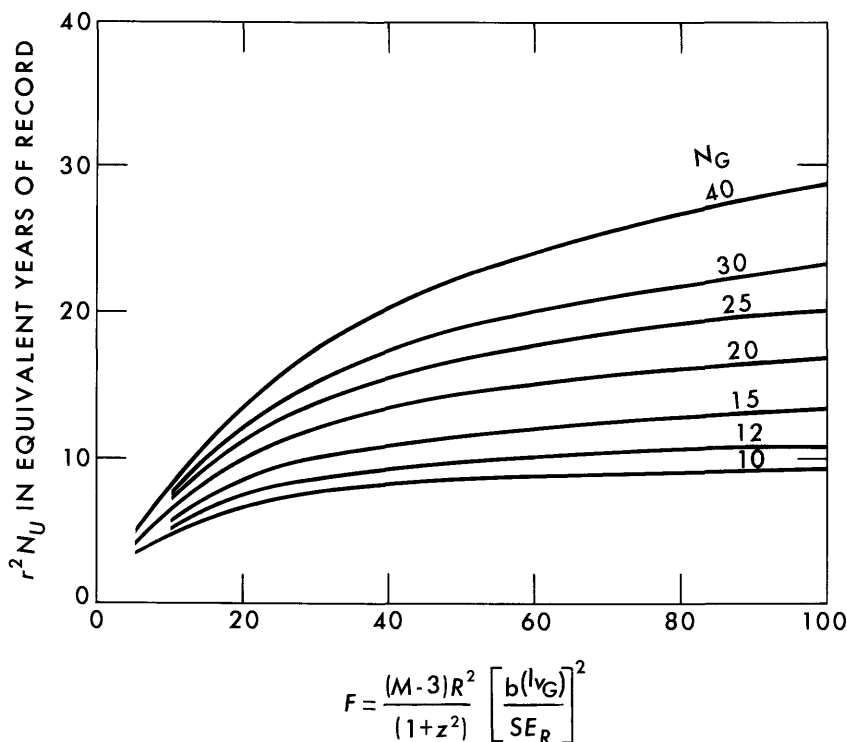$$F = \frac{(M-3)R^2}{(1+z^2)} \left[\frac{b\left(\frac{I_{V_G}}\right)}{SE_R} \right]^2$$

FIGURE 14.—Accuracy of low-flow characteristics estimated by correlation of base-flow measurements.

results given in table 2. The accuracy, in equivalent years of record, can be estimated by using F, computed from the statistics of a plot of base-flow measurements on logarithmic paper, to enter figure 14 with the length of record at the gaged site as a parameter, and thus obtaining a value of $r^2 N_U$ that can be multiplied by $1/r^2$ to give N_U.

A graphical procedure for relating base-flow measurements and for using figure 14 to estimate the accuracy of an estimate obtained from the relation are illustrated by means of an example in the following section.

EXAMPLE

In figure 15, 28 base-flow measurements made on Old Mill Stream branch at Centerville, Md., (drainage area 11.2 sq mi) have been plotted against the corresponding daily discharge at the stream-gaging station on Unicorn Branch near Millington, Md., (drainage area 22.3 sq mi). These daily discharges are used as being equivalent to the base flows that would have been measured at the gaging station on the days the base-flow measurements were made at Centerville and are called base-flow measurements in the remainder of this report. The relation line based on the points in figure 15 is a graphical
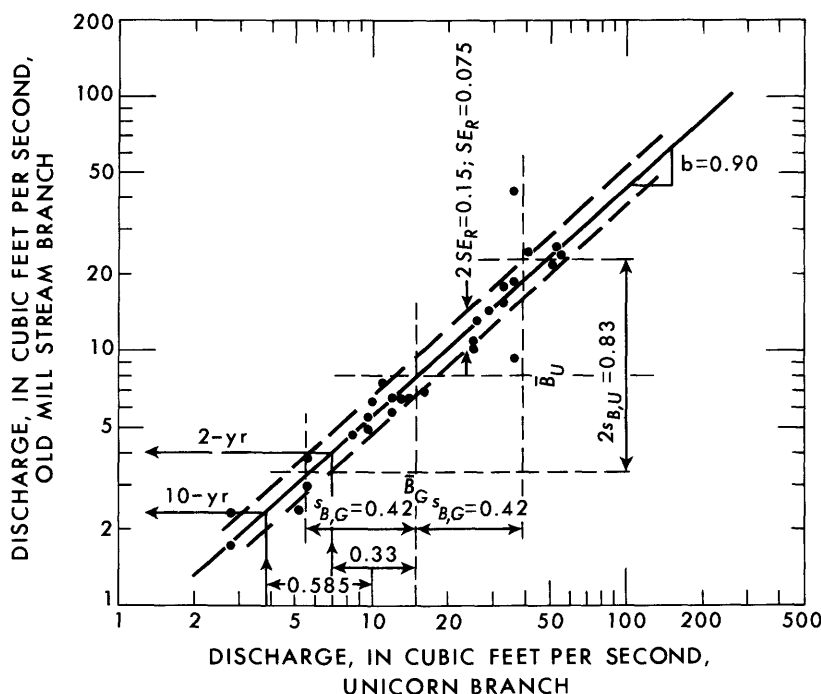

FIGURE 15.—Relation of base-flow measurements, 1952–66, Unicorn Branch near Millington, Md., and Old Mill Stream Branch at Centerville, Md.

estimate of the line of regression for estimating base flow at the un-gaged site from that at the gaged site. As such, the slope should be such as to minimize the sum of the squares of the vertical departures from the line. Horizontal departures should not be considered. Under the assumption that the relation line is applicable to 7-day low flows also, the median annual 7-day low flow of 7.0 cfs (cubic feet per second) at the gaging station gives 4.0 cfs as the estimated median annual low flow on Old Mill Stream Branch. The problem is to appraise the accuracy of this estimate by the use of figure 14.

By assuming the vertical departures from the line of regression, in log units, to be normally distributed, and by drawing parallel lines to exclude the one-sixth of the points that have the greatest positive departure and the one-sixth that have the greatest negative departure, an estimate of the standard error of regression can be ob-tained as half the distance, in log units, between the upper and lower lines. In figure 15, the parallel lines are drawn between the fourth and fifth greatest positive and greatest negative departures, and the re-sulting standard error of estimate, SE_R, is one-half the log of 14.1/10.0, or 0.075 log units. The slope of the regression line, b, is 0.90 as shown in figure 15.

The value of z is computed as $(\hat{X}_{T,\,G} - \bar{B}_G)/s_{B,\,G}$ in which $s_{B,\,G}$ is the standard deviation of the logarithms of the base flow points at Unicorn Branch, and r is computed as $1 - \sqrt{(SE_R/s_{B,\,U})^2}$ in which $s_{B,\,U}$ is the standard deviation of the logarithms of the base flow measurements at Old Mill Stream Branch. Values of $s_{B,\,G}$ and $s_{B,\,U}$ for use in making these computations can be estimated graphically by excluding one-sixth of the largest and one-sixth of the smallest base-flow measurements in the horizontal direction and in the vertical direction respectively and by taking one-half the resulting range in log units as shown in figure 15 ($s_{B,\,G}$ and $s_{B,\,U}$ equal 0.42 and 0.415 respectively). In order to obtain theoretical consistency between the graphically determined statistics, however, values of $s_{B,\,U}$ and $s_{B,\,G}$ should be adjusted and r should be computed as follows:

1. Compute $s'_{B,U} = \sqrt{b^2 s^2_{B,\,G} + SE^2_R} = \sqrt{(0.9)^2(0.42)^2 + (0.075)^2} = 0.385$

2. Compute $\hat{s}_{B,\,U} = (s'_{B,\,U} + \hat{s}_{B,\,U})/2 = (0.385 + 0.415)/2 = 0.400$

3. Compute $\hat{r} = \sqrt{1 - (SE_R/\hat{s}_{B,\,U})^2} = \sqrt{1 - (0.075/0.400)^2} = 0.982$

4. Compute $\hat{s}_{B,\,G} = \hat{s}_{B,\,U}(\hat{r}/b) = 0.400(0.982/0.9) = 0.436$.

Although in this instance r changes only slightly and values of $s_{B,\,U}$ and $s_{B,\,G}$ are changed only a few percent, adjustments will usually be larger because of fewer base flow measurements. Therefore the adjustments should be made before computing values of r and z to use in equations 11 and 12.

The value of $(\hat{X}_{T,\,G} - \bar{B}_G)$ for use in computing z is the distance in log units between the median (logarithmic mean) of the base-flow measurements at Unicorn Branch, which is 15 cfs, and the estimated median annual 7-day low flow at that site, which is 7.0 cfs. It equals the log of $(7/15)$ or -0.33 log units. The resulting z is $-0.33/0.436$ or -0.757 and z^2 is 0.57. The value of $\bar{B}_G$ to use in computing $(\hat{X}_{T,\,G} - \bar{B}_G)$ is taken as being midway of the range of the standard deviation lines so long as the resulting $\bar{B}_G$ is consistent with a similarly computed $\bar{B}_U$. If the computed $\bar{B}_G$ and $\bar{B}_U$ do not plot on the regression line, one or both should be adjusted. It should be noted that both $\bar{B}_G$ and $s_{B,\,G}$ depend on the discharge of the base-flow measurements at the gaged site, and thus depend on when the measurements happen to be made. As such measurements are usually made at various times during the year, the resulting $\bar{B}_G$ and $s_{B,\,G}$ will usually be larger than the corresponding statistics of the annual low flows.

The value for I_{vG}, the standard deviation of the logarithms of 18 annual 7-day low flows at the gaged site (Unicorn Branch), was computed to be 0.17, and the logarithmic skew coefficient, -1.0. (One of the observed annual 7-day low flows in the period 1948–66 was omitted in the computation of these statistics, because it was abnormally low as a result of regulation at a fish hatchery upstream.) For this skew coefficient, table 1 gives an R value of 0.933 for the median annual 7-day low flow.

The only remaining term to be evaluated in equation 11 is M, the number of base-flow measurements, which in the derivation of the equation were assumed to be independent, that is, to be independent estimates of the base-flow relation. As measurements made during the same low-flow period tend not to be independent, M has been used as 16 instead of 28. The figure of 16 was obtained by counting only those measurements made at least 30 days after a previous base-flow measurement. Using this criteria, all five measurements made in 1964–66 were counted, but only 11 of the 23 semimonthly measurements made between October 1952 and October 1953 were counted. Other equally valid rules regarding independent base-flow measurements could no doubt be formulated by studying the causes of such scatter to see how much of it is due to year-to-year variation in the base-flow relation and how much is due to the unequal effect of direct runoff from recent rain.

Using equation 11 and obtaining R from table 1, F for the median annual low flow is

$$F = \frac{(16-3)(0.933)^2}{1+0.57}\left[\frac{(0.9)(0.17)}{0.075}\right]^2 = (7.20)(4.16) = 30.0,$$

and as the median annual 7-day low flow at the gaging station on Unicorn Branch is based on 18 annual 7-day low flows, N_G is 18. For these values of F and N_G, interpolation between the 15- and 20-year curves in figure 14, or solution of equation 12, gives 11.2 years for $r^2 N_U$. As r is 0.982, N_U is $11.2/(0.982)^2$, or 11.6 years. This means that 12 annual 7-day low flows would have to be observed on Old Mill Stream Branch at Centerville to obtain a median annual 7-day low flow as accurate as the 4.0 cfs given by the regression of base-flow measurements.

For the 10-year 7-day low flow, the estimate based on the observed record on Unicorn Branch is 3.9 cfs, and by the relation line in figure 15 the corresponding estimated discharge on Old Mill Stream Branch is 2.3 cfs. In appraising the accuracy of this estimate by use of equation 11, $(\hat{X}_{T, G} - \bar{B}_G)$ equals the logarithm of 3.9/15 or -0.585, z^2 equals $(-0.585/0.436)^2$ or 1.80, and R is 1.956, as given in table 1 for T of 10 and skew coefficient of -1.0. The other terms in the equation remain as they were for the median annual low flow and F becomes 73.9. From figure 14 with an N_G of 18 years, $r^2 N_U$ is 14.5 years and N_U is then equal to 15.0 years. Thus the accuracy of the estimated 10-year 7-day low flow, in equivalent years of record, is 15.0 years as compared to an accuracy of 11.6 years for the median annual 7-day low flow.

DISCUSSION

The fact that the 10-year low flow as estimated in the preceding example has a higher accuracy in equivalent years of record than does the estimated median (2-year) low flow deserves explanation. The fundamental reason behind the change of N_U with change in T is that values of z, $SE_{T, G}$, and V_G in equation 2 all increase with T so that, measured in log units, the standard error of the 10-year low is always greater than that of the 2-year low. The conversion from standard error in log units to equivalent length of record in years, however, depends on the theoretical standard error of a T-year peak based on a given number of years of record at the ungaged site, and thus depends on the R values in table 1. When equation 13 is used to make this conversion, the increase in R^2_G with increase in T may well overbalance the increase in V_U, particularly when R_G increases rapidly with T as it does for large negative skew coefficients. The alternative to estimating the T-year low from a regression of base-flow measurements is to operate a stream-gaging station at the ungaged site for N_U years. At sites where the annual low flows have a negative skew coefficient, more years are required to obtain the same standard error in log units for a 10-year low than at sites with positive skew coefficients. Had the logarithmic skew coefficient for the annual

low flows in the example been $+1.0$ instead of -1.0, N_U for the 10-year low would have been less than for the 2-year low.

In the example shown in figure 15, the base-flow measurements cover a considerably larger range in discharge than do the annual low flows, and the resulting z values tend to make the estimates less accurate than would be the case if the mean of the base-flow measurements were nearer the mean of the annual low flows. Although fewer measurements would be required to give comparable accuracy if the measurements were made only when the base flow was at or below the median annual low flow, other considerations make it more feasible to use base-flow measurements over a wider range of base flow.

The main reason for obtaining base-flow measurements over a wide range of base flow is that the measurements needed to define the relation and to appraise its accuracy can be obtained in fewer years as there is more opportunity each year. An additional reason is that a wider range increases the coefficient of correlation and thus tends to remove the need for applying a "noise" adjustment to the regressed estimates (Matalas and Jacobs, 1964, equation 26). The magnitude of the bias introduced by neglecting such an adjustment can be appraised by comparing a T-year discharge picked from the regression line with one picked from a line that crosses the regression line at the median annual low flow and that has a slope equal to I_{vG}/I_{vU}, which can be estimated from equation 8.

The relation curve defined by concurrent base-flow measurements or by concurrent daily base-flow discharges will in most cases be applicable also to 7-day annual low flows, as the difference in recession rates will be reflected both in the slope of the relation and in the ratios of 7-day low flow to 1-day low flow at the two sites. The applicability of the relation to longer low-flow periods such as those for 30-day and 60-day annual low flows depends on hydrologic factors other than base-flow recession and should be investigated by studies of gaging-station records in the area being studied before the regression line is used to estimate such characteristics at ungaged sites.

The curves shown in figure 14 can be used in planning a data-collection program by observing how N_U is affected by changes in N_G, the length of record at the gaged site, and by M, the number of independent base-flow measurements. If, for the stations used in the preceding example, the objective had been to obtain 10-year accuracy at the ungaged site, $r^2 N_U$ would be 9.64, and the corresponding F for an N_G of 18 would be 20.8. Since $M-3$ is proportional to F, the number of base-flow measurements required to define the median annual 7-day low flow could be decreased enough to make $M-3$ equal to $(16-3)$ times $20.8/30.0$, or 9.0, from which the number of measurements would be 12. For the 10-year low, 13 times $20.8/73.9$ gives 2.8, from which the

required number of measurements would be 6. If the number of measurements were to remain at 16, however, F for the median annual 7-day low flow would remain at 30.0, and N_G (for $r^2 N_U$ of 9.64) could be decreased to 15 years of record. For the 10-year low, F would remain at 70.3 and only 11 years of record would be required at the gaged site. These examples assume that the measurements added or subtracted would be randomly selected and that the distribution of the revised number would be approximately the same as before. A greater improvement could possibly be obtained by planning to have the additional measurements made where they would do the most good in defining the relation.

If the base-flow relations at many low-flow partial record sites are to be evaluated, perhaps the most efficient approach would be to use the accuracy of some of the relations as evaluated by the procedure described in this paper to prepare a summary table of M, N_G, SE_R, and N_U that could then be used as a guide to estimate whether the accuracy objective had been reached at the other sites.

If a procedure could be devised to give reduced weight to "outliers" in the base-flow relation that result from having an abnormally greater amount of direct runoff at one site than at the other, it would be feasible to develop a computer program that would not only establish the regression between the logarithms of the base-flow measurements, but would provide for entering the regression with various T-year low flows to obtain the estimated low-flow characteristic at the ungaged site, for making adjustment for the bias in the regressed line, and for computing the accuracy of the estimate in equivalent years of record. If such a program were to be developed, the accuracy in equivalent years of record at the ungaged site could be computed by relating V_U given by equation 2 to N_U by use of equation 13. Even though a computer were to be used to evaluate the accuracy of an estimated low-flow characteristic at specific sites, the curves presented in figure 14 would still be useful in program planning.

In planning a network of low-flow partial-record stations and associated stream gaging stations, the objective should be to obtain an optimum balance between the cost of additional base-flow measurements and the cost of increasing the length of record at the gaging station. To help in developing such a plan, z values for use in computing F could be estimated by relating z values from existing relations of base-flow measurements to the average number of base-flow measurements per year. Similarly, the standard error of estimate, SE_R, and the coefficient of correlation from existing relations could be related to distance between geologically similar drainage basins. Then these results could be used with the equations and graphs given in this paper to determine how the required number of base-flow

measurements vary with distance from a gaging station and with length of record at the gaging station.

OUTLINE OF DERIVATIONS

By EDWARD J. GILROY

BACKGROUND

Base flow measurements made at a site on a stream where no observations of annual low flows have been obtained are used to estimate a simple linear regression between the logarithms of these base flows and the logarithms of concurrent base flows at a nearby stream gaging station. The regression equation is then evaluated by using the logarithm of an estimated T-year low flow at the gaged site to obtain an estimate of the logarithm of the T-year low flow at the ungaged site. This section of the report presents the mathematical model underlying this procedure along with (1) a derivation of the variance of the estimator of the logarithm of the T-year low flow at the ungaged site, and (2) an unbiased estimator of the variance of the logarithms of annual low flows at the ungaged site by using the technique suggested by Matalas and Jacobs of adding noise to each regression estimate. The variance of the linear regression estimator is obtained by using bivariate normal regression theory with certain assumptions regarding the nature of the data.

STATEMENT OF THE PROBLEM

Let B_U and B_G denote the logarithms of base flows at the ungaged site and the gaged site, respectively. There are M base flow measurements available at both the gaged and ungaged sites. The corresponding logarithms of these $2M$ base flow measurements may be arranged as ordered pairs in the array

$$B_U(1), B_U(2), \ldots, B_U(M)$$
$$B_G(1), B_G(2), \ldots, B_G(M)$$

(14)

where for a fixed j, $1 \leq j \leq M$, $B_U(j)$ and $B_G(j)$ are logarithms of concurrent measurements at the two sites.

Let L_U and L_G denote the logarithms of annual low flows at the ungaged site and the gaged site, respectively. There are N_G annual low-flow observations at the gaged site and no annual low-flow observations at the ungaged site. The corresponding logarithms of there N_G annual low-flow observations may be written as

$$L_G(1), L_G(2), \ldots, L_G(N_G).$$

(15)

The sample statistics pertinent to the derivations in this section are:

$$\bar{B}_U = \frac{1}{M} \sum_{j=1}^{M} B_U(j)$$

$$\bar{B}_G = \frac{1}{M} \sum_{j=1}^{M} B_G(j)$$

$$\bar{L}_G = \frac{1}{N_G} \sum_{j=1}^{N_G} L_G(j)$$

$$S^2_{B,\,U} = \frac{1}{M-1} \sum_{j=1}^{M} [B_U(j) - \bar{B}_U]^2$$

$$S^2_{B,\,G} = \frac{1}{M-1} \sum_{j=1}^{M} [B_G(j) - \bar{B}_G]^2$$

$$S^2_{L,\,G} = \frac{1}{N_G-1} \sum_{j=1}^{N_G} [L_G(j) - \bar{L}_G]^2$$

$$C = \frac{1}{M-1} \sum_{j=1}^{M} [B_U(j) - \bar{B}_U][B_G(j) - \bar{B}_G]$$

$$b = C/S^2_{B,\,G}$$

$$r = C/(S_{B,\,G} S_{B,\,U}).$$

Applying the classical linear least squares technique to the data represented by the array (14), a linear estimate, $\hat{B}_U(j)$, of the logarithm of the j-th base flow at the ungaged site is obtained. $B_U(j)$ is written as

$$\hat{B}_U(j) = \bar{B}_U + b[B_G(j) - \bar{B}_G]. \qquad (16)$$

The main interest in this section lies in using equation 16 to estimate statistical properties of L_U, the logarithm of annual low flows at the ungaged site, for which no direct measurements are available. The statistical properties considered are the mean, variance, and the $\frac{1}{T}$-th percentile, known as the T-year annual low flow. Two methods of estimation are used. Method I, which is used for estimating the T-year annual low flow, consists of first using the observations of L_G in array (15) to estimate the T-year annual low flow at the gaged site and then using this estimate in place of $B_G(j)$ in equation 16 to obtain an estimate of the T-year annual low flow at the ungaged site, Method II consists in first using equation 16—modified by the addition of a noise term—and the observations of L_G in array (15) to estimate the corresponding annual low flows at the ungaged site. These estimates are then operated on as though they were actual observations to compute estimates of the mean and the variance of the annual low flows at the ungaged site.

The rationale for these two methods of estimation is that, based on past experience, it is reasonable to assume that the logarithms of annual low flows and the logarithms of the T-year annual low flows at the two sites follow the same linear relationship as do the logarithms of base flows at the two sites.

For the method I estimator, interest centers on the variance of the estimator. For the method II estimator, interest centers on obtaining an unbiased estimator of the variance of the annual low flows at the ungaged site.

STATISTICAL MODEL

The assumptions made in order to derive the statistical properties of the estimators given in methods I and II above are not the same for the two methods. The assumptions which are common to both cases are three in number and are the following:

Assumption 1.—B_U and B_G are jointly distributed random variables following a bivariate normal probability distribution with parameters given by their means $\mu_{B,\,U}$ and $\mu_{B,\,G}$, their variances $\sigma^2_{B,\,U}$ and $\sigma^2_{B,\,G}$ and their correlation coefficient ρ_B.

Assumption 2.—For the M observations listed in array (14), it is true that nonconcurrent observations are statistically independent observations of the base flow population described in assumption 1.

Assumption 3.—L_G is a random variable and the N_G observations given in array (15) are statistically independent of one another and also statistically independent of the observations in array (14).

For the derivation of the sampling variance of a method I estimator, two additional assumptions are made. Let $X_{T,\,U}$ and $X_{T,\,G}$ denote the logarithm of the T-year annual low flow at the ungaged and gaged sites, respectively. Let $\hat{X}_{T,\,G}$ denote the estimate of $X_{T,\,G}$ obtained from the data array (15).

Assumption 4.—$X_{T,\,U}$ and $X_{T,\,G}$ are linearly related as

$$X_{T,\,U} = \mu_{B,\,U} + \beta[X_{T,\,G} - \mu_{B,\,G}] \qquad (17)$$

where

$$\beta = \rho_B \frac{\sigma_{B,\,U}}{\sigma_{B,\,G}}.$$

Assumption 5.—The estimate, $\hat{X}_{T,\,G}$, of $X_{T,\,G}$ may be written as

$$\hat{X}_{T,\,G} = X_{T,\,G} + d \qquad (18)$$

where d is a random variable with mean zero and variance, $\sigma^2_{T,\,G}$. The random variable d is independent of the random variables B_U and B_G.

For the derivation of the sampling statistics of a method II estimator the following alternate assumptions are made in place of assumptions 4 and 5 above:

Assumption 4'.—L_U and L_G are jointly distributed random variables following a bivariate normal distribution with parameters given by their means, $\mu_{L, U}$ and $\mu_{L, G}$, their variances, $\sigma^2_{L, U}$ and $\sigma^2_{L, G}$, and their correlation coefficient ρ_L.

Assumption 5'.—The following relations are assumed to hold between the parameters of the distributions of B_U, B_G, and L_U, L_G:

$$\sigma^2_{L, U}\big|\sigma^2_{L, G} = \sigma^2_{B, U}\big|\sigma^2_{B, G} \tag{19}$$

$$\beta_L = \beta_B = \beta \tag{20}$$

Since by definition

$$\beta_B = \rho_B \sigma_{B, U}\big|\sigma_{B, G}$$

and

$$\beta_L = \rho_L \sigma_{L, U}\big|\sigma_{L, G}$$

equations 19 and 20 imply

$$\rho_L = \rho_B = \rho. \tag{21}$$

Obviously any two of the three equations 19, 20, and 21 could be assumed and the third would follow directly.

VARIANCE OF THE ESTIMATOR OF THE T-YEAR LOW FLOW (METHOD I)

The method I estimator $\check{X}_{T, U}$ of $X_{T, U}$ is defined by

$$\check{X}_{T, U} = \bar{B}_U + b(\hat{X}_{T, G} - \bar{B}_G). \tag{22}$$

From the assumptions made above

$$E(\check{X}_{T, U}) = \mu_{B, U} + \beta[X_{T, G} - \mu_{B, G}].$$

By reason of equation 17, it follows that

$$E(\check{X}_{T, U}) = X_{T, U}$$

and so $\check{X}_{T, U}$ is an unbiased estimater of $X_{T, U}$.

The variance of $\check{X}_{T, U}$ is given by

$$\mathrm{Var}(\check{X}_{T, U}) = E(\check{X}_{T, U} - X_{T, U})^2$$

which, by reason of equations 17, 18, and 22, may be rewritten as

$$\mathrm{Var}(\check{X}_{T, U}) = E\{[\bar{B}_U - \mu_{B, U}] + (b - \beta)X_{T, G} + bd - b\bar{B}_G + \beta\mu_{B, G}\}^2. \tag{23}$$

As a result of assumptions 1 and 2, the following well known property of bivariate normal random variables may be applied. Each pair, $B_U(j)$, $B_G(j)$ of observations may be written as

$$B_U(j) = \mu_{B,\ U} + \beta(B_G(j) - \mu_{B,\ G}) + e(j) \tag{24}$$

where the $e(j)$ are independently normally distributed with mean zero and variance $(1-\rho^2)\sigma^2_{B,\ U}$ (Kenney and Keeping, 1951, Chapter VIII).

From equation 24 it follows that

$$\bar{B}_U - \mu_{B,\ U} = \beta(\bar{B}_G - \mu_{B,\ G}) + \bar{e} \tag{25}$$

and

$$b = \beta + \frac{C_{e,\ G}}{S^2_{B,\ G}} \tag{26}$$

where

$$\bar{e} = \frac{1}{M} \sum_{j=1}^{M} e(j)$$

$$C_{e,\ G} = \frac{1}{M-1} \sum_{j=1}^{M} (e(j) - \bar{e})(B_G(j) - \bar{B}_G).$$

It follows from equations 25 and 26 that equation 23 may be rewritten as

$$\mathrm{Var}(\tilde{X}_{T,\ U}) = E[\bar{e} + (b - \beta)(X_{T,\ G} - \bar{B}_G) + bd]^2.$$

The technique used to determine the expectation indicated on the right side of equation 27 is to first take a conditional expectation for fixed values of B_G and then find the unconditional expectation of the resulting expression. Thus.

$$E\{[\bar{e} + (b - \beta)(X_{T,\ G} - \bar{B}_G) + d]^2 | B_G\}$$

$$= E(\bar{e}^2 | B_G) + (X_{T,\ G} - \bar{B}_G)^2 E[(b - \beta)^2 | B_G]$$

$$+ E(b^2 | B_G)E(d^2 | B_G) + 2(X_{T,\ G} - \bar{B}_G)E[\bar{e}(b - \beta) | B_G]$$

$$+ E(\bar{e}b | B_G)E(d | B_G)$$

$$+ 2(X_{T,\ G} - \bar{B}_G)E(b(b - \beta) | B_G]E[d | B_G]. \tag{28}$$

Equation 28 holds since d is independent of B_U, B_G and the B_G's are fixed with respect to this conditional expectation.

From equation 26 and the fact that the random variables $e(i)$ and $e(j)$ are independent for $i \neq j$, equation 28 reduces to

$$E[(\bar{e} + (b - \beta)X_{T,\ G} + bd)^2 | B_G)$$

$$= \frac{(1-\rho^2)\sigma^2_{B,\ U}}{M}\left\{1 + \frac{(X_{T,\ G} - \bar{B}_G)^2 + \sigma^2_{T,\ G}}{\frac{M-1}{M}\rho^2_{B,G}}\right\}$$

$$+ \beta^2\sigma^2_{T,\ G}. \tag{29}$$

Now taking the unconditional expectation of the right side of equation 29 gives

$$\text{Var}(\hat{X}_{T,\,U}) = \frac{(1-\rho^2)\sigma^2_{B,\,U}}{M}\Big\{1+\frac{1}{M-3}$$

$$+\frac{M}{M-3}\frac{(X_{T,\,G}-\mu_{B,\,G})^2}{\sigma^2_{B,\,G}}+\frac{M}{M-3}\frac{\sigma^2_{T,\,G}}{\sigma^2_{B,\,G}}\Big\}+\beta^2\sigma^2_{T,\,G}. \qquad (30)$$

Equation 30 follows from equation 29 since $\bar{B}_G$ and $S^2_{B,\,G}$ are independent by reason of the normality assumption and since

$$E[(S^2_{B,\,G})^{-1}] = \frac{M-1}{(M-3)\ \sigma^2_{B,\,G}}.$$

Thus equation 30 gives the variance of $\hat{X}_{T,\,U}$, the estimator of the T-year low flow.

ESTIMATOR OF VARIANCE OF ANNUAL LOW FLOWS (METHOD II)

Estimates, $\hat{L}_U(j)$, $j=1, 2, \ldots, N_G$, of the missing L_U observations which correspond to the L_G observations given in array (15) are given by

$$\hat{L}_U(j) = \bar{B}_U + b(L_G(j)-\bar{B}_G)+\alpha\theta(1-r^2)^{1/2}\frac{S_{B,\,U}S_{L,\,G}}{S_{B,\,G}}f(j)$$

where $f(j)$ is a normally distributed random component with mean zero and variance one, θ is a parameter which equals 1 if "noise" is to be added and 0 if "noise" is not to be added, and α is a parameter to be determined so as to yield an unbiased estimate of $\sigma^2_{L,\,U}$. The introduction of the noise term is due to Matalas and Jacobs (1964). Because no observations at all have been made on the random variable L_U, the term $S_{B,U}\,S_{L,G}|S_{B,G}$ is used as a temporary estimate of $\sigma_{L,\,U}$ by reason of equation 19.

Let $\tilde{\mu}_{L,\,U}$ and $\tilde{\sigma}^2_{L,\,U}$ denote method II estimators of the mean, $\mu_{L,\,U}$, and the variance $\sigma^2_{L,\,U}$ of L_U. Then

$$\tilde{\mu}_{L,\,U}=\frac{1}{N_G}\sum_{j=1}^{N_G}L_U(j)$$

$$\sigma^2_{L,\,U}=\frac{1}{N_g-1}\sum_{j=1}^{N_G}(\hat{L}_U(j)-\tilde{\mu}_{L,\,U})^2.$$

Let

$$\bar{f}=\frac{1}{N_G}\sum_{j=1}^{N_G}f(j)$$

$$S^2_f=\frac{1}{N_G-1}\sum_{j=1}^{N_G}(\bar{f}(j)-f)^2$$

$$S_{f,\,L}=\frac{1}{N_G-1}\sum_{j=1}^{N_G}(\bar{f}(j)-f)(L_G(t)-\bar{L}_G).$$

Then $\tilde{\mu}_{L,\,U}$ and $\tilde{\sigma}^2_{L,\,U}$ may be written as

$$\tilde{\mu}_{L,\,U} = \bar{B}_U + b(\bar{L}_G - \bar{B}_G) + \alpha\theta(1-r^2)\frac{S_{B,\,U}S_{L,\,G}}{S_{B,\,G}}\bar{f}$$

$$\tilde{\sigma}^2_{L,\,U} = b^2 S^2_{L,\,G} + \alpha^2\theta^2(1-r^2)\frac{S^2_{B,\,U}S^2_{L,\,G}}{S^2_{B,\,G}}S^2_f$$

$$+2\alpha\theta(1-r^2)^{1/2}\frac{S_{B,\,U}S_{L,\,G}}{S_{B,\,G}}bS_{f,\,L}.$$

Using the Matalas-Jacobs argument against the use of different sets of "noise" terms by different investigators, the terms $\bar{f}$, S^2_f, and $S_{f,\,L}$ are replaced by their best estimates 0, 1, 0, respectively, leading to the estimates

$$\tilde{\mu}_{L,\,U} = \bar{B}_U + b(\bar{L}_G - \bar{B}_G)$$

$$\tilde{\sigma}^2_{L,\,U} = b^2 S^2_{L,\,G} + \alpha^2\theta^2(1-r^2)\frac{S^2_{B,\,U}S^2_{L,\,G}}{S^2_{B,\,G}}. \tag{31}$$

The addition of the noise term is seen not to affect the estimator $\tilde{\mu}_{L,\,U}$. The estimator $\tilde{\mu}_{L,\,U}$ is unbiased since

$$E(\tilde{\mu}_{L,\,U}) = \mu_{B,\,U} + \beta(\mu_{L,\,G} - \mu_{B,\,G}).$$

But

$$\mu_{L,\,U} = \mu_{B,\,U} + \beta(\mu_{L,\,G} - \mu_{B,\,G})$$

by reason of assumptions 1, 4', and 5'. Therefore

$$E(\tilde{\mu}_{L,\,J}) = \mu_{L,\,U}.$$

Because the N_G annual low-flow measurements are independent of the base flows, the expectation of $\tilde{\sigma}^2_{L,\,U}$ may be written as

$$E(\tilde{\sigma}^2_{L,\,U}) = E(b^2)E(S^2_{L,\,G}) + \alpha^2\theta^2 E\frac{(1-r^2)S^2_{B,\,U}}{S^2_{B,\,G}}E(S^2_{L,\,G}). \tag{32}$$

Using the technique described in the derivations of method I estimators it can be shown that

$$E(S^2_{L,\,G}) = \sigma^2_{L,\,G} \tag{33}$$

$$E(b^2) = \frac{\sigma^2_{B,\,U}}{\sigma^2_{B,\,G}}\frac{1-4\rho^2+M\rho^2}{M-3} \tag{34}$$

$$E(1-r^2)\frac{S^2_{B,\,U}}{S^2_{B,\,G}} = \frac{\sigma^2_{B,\,U}}{\sigma^2_{B,\,G}}(1-\rho^2)\frac{(M-2)}{M-3}. \tag{35}$$

Combining equations 32 through 35 $E(\tilde{\sigma}^2_{L,\,U})$ may be written as

$$E(\tilde{\sigma}^2_{L,\,U}) = \sigma^2_{L,\,U}\left\{1 + \frac{(1-\rho^2)}{M-3}[\alpha^2\theta^2(M-2) - (M-4)]\right\}. \tag{36}$$

In order that $\tilde{\sigma}^2_{L,\ U}$ be an unbiased estimator of $\sigma^2_{L,\ U}$, θ and α must take the values $\theta = 1$ and

$$\alpha^2 = \frac{M-4}{M-2}.$$ (37)

when substituting α^2 from equation 37 into equation 31 gives

$$\tilde{\sigma}^2_{L,\ U} = S^2_{L,\ G}\left[b^2 + \frac{M-4}{M-2}(1-r^2)\frac{S^2_{B,\ U}}{S^2_{B,\ G}}\right]$$ (38)

as the unbiased estimator of $\sigma^2_{L,\ U}$. Since

$$b^2 = r^2\frac{S^2_{B,\ U}}{S^2_{B,\ G}}$$

b can be eliminated from equation 38 to give

$$\tilde{\sigma}^2_{L,\ U} = S^2_{L,\ G}\left(\frac{S_{B,\ U}}{S_{B,\ G}}\right)^2\left[r^2 + \left(\frac{M-4}{M-2}\right)(1-r^2)\right].$$ (39)

REFERENCES

Carter, R. W., and Benson, M. A., 1970, Concepts for the design of streamflow data programs: U.S. Geol. Survey open-file report, 33 p.

Hardison, C. H., 1969, Accuracy of streamflow characteristics, *in* Geological Survey research 1969; U.S. Geol. Survey Prof. Paper 650–D, p. D210–D214.

Kenney, J. F., and Keeping, E. S., 1951, Mathematics of Statistics, pt . 2, 2d ed., Princeton, N.J., D. Van Nostrand Co., 429 p.

Matalas, N. C., and Jacobs, B., 1964, A correlation procedure for augmenting hydrologic data: U.S. Geol. Survey Prof. Paper 434–E, p. E1–E7.

Riggs, H. C., 1965, Estimating probability distributions of drought flows: Water and Sewage Works, v. 112, no. 5, p. 153–157.

Thomas, D. M., and Benson, M. A., 1970, Generalization of streamflow characteristics from drainage basin characteristics: U.S. Geol. Survey Water-Supply Paper 1975, 55 p., 16 figs.

.

Manual of Hydrology: Part 2. Low-Flow Techniques

*This volume was published
as separate chapters A and B*

UNITED STATES DEPARTMENT OF THE INTERIOR

ROGERS C. B. MORTON, *Secretary*

GEOLOGICAL SURVEY

V. E. McKelvey, *Director*

CONTENTS

[Letters designate the separately published chapters]