Flood Frequency Analysis With a Generalized Skew Coefficient

GARY D. TASKER

U.S. Geological Survey, Reston, Virginia 22092

The Hydrology Committee of the Water Resources Council (1976) has recently issued guidelines for flood frequency analyses. One aspect of these guidelines concerns the estimate of skewness. Monte Carlo experiments are used to gain some insights into sensitivity of estimates of peak flows to errors in mapped skew coefficients. The optimum factor by which to weight a sample skew coefficient and a generalized skew coefficient is a function of sample size, population skew coefficient, and map error.

INTRODUCTION

Hydrologists, engineers, planners, and designers often are required to estimate magnitude and frequency of floods at a specific site from records of past flows. A frequently used approach to this problem is to estimate the parameters of an assumed distribution by the method of moments by using annual peak flows from past records. The *Hydrology Committee of the Water Resources Council* [1976] recommends computation of the first three moments as follows.

If the assumed distribution is characterized by three or fewer parameters, as is usually the case, then the parameters may be estimated in terms of the relations of the parameters to the mean $\bar{\chi}$, variance σ^2, and skew coefficient γ. These parameters can be estimated, respectively, as

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{N} \chi_i \qquad (1)$$

$$S^2 = \frac{1}{N-1} \sum_{i=1}^{N} (\chi_i - \bar{\chi})^2 \qquad (2)$$

$$G = \frac{N}{(N-1)(N-2)S^3} \sum_{i=1}^{N} (\chi_i - \bar{\chi})^3 \qquad (3)$$

where χ_i denotes the logarithmic value of an annual peak flow for a record length N.

Because estimates of γ are biased and subject to large sampling errors, the *Hydrology Committee of the Water Resources Council* [1976] has suggested the use of regional skew maps to obtain a generalized skew estimate $\bar{G}$ for a specific site. Moreover, the committee has suggested using a weighted average of G and $\bar{G}$,

$$G' = WG + (1 - W)\bar{G} \qquad (4)$$

where the weighting factor W may assume a value on the interval [0, 1].

In this brief report, results of Monte Carlo experiments are used to gain some insights into sensitivity of estimates of peak flows to errors in mapped skew coefficients. In particular, the influence of our alternative skew weighting factors in estimating peak flows is examined.

EXPERIMENTAL DESIGN

A series of N random numbers was generated to simulate the logarithmic values of a series of annual peak flows. The numbers were generated from a Pearson type 3 distribution with mean μ, standard deviation σ, and skew coefficient γ. Five hundred samples each of size N = 10, 20, 30, 40, 50, 70, and 90 were generated with $\mu = 3.0$, $\sigma = 0.25$, and a randomly generated value of γ. For each of the 500 samples, γ was randomly generated from a normal distribution with mean equal to an assumed standard deviation S_m of skew coefficients about isolines on a skew map.

The experiment was repeated for each combination of $\bar{G}$ = 0.4(0.2)0.8, and S_m = 0.1, 0.4, 0.55, 0.75, and 1. The standard deviation of station values of the skew coefficient about the isolines of the skew map provided by the *Hydrology Committee of the Water Resources Council* [1976] is 0.55.

By using the simulated data, 10-, 50-, 100-, and 500-year peak flows were estimated by the following methods:

Method 1. Sample statistics $\bar{X}$, S, and G calculated from (1), (2), and (3) were used to estimate peak flows. This method is equivalent to using a skew weighting factor of $W = 1.0$.

Method 2. Sample statistics $\bar{X}$ and S along with $\bar{G}$ were used to estimate peak flows. This method is equivalent to using a skew weighting factor of $W = 0$.

Method 3. Sample statistics $\bar{X}$ and S along with a weighted average skew coefficient defined by (4) were used to estimate peak flows, where the weighting factor $W = W_1$ recommended by the *Hydrology Committee of the Water Resources Council* [1976] is defined as

$$\begin{aligned} W_1 &= 0 \qquad N < 25 \\ W_1 &= (N - 25)/75 \qquad 25 \le N \le 100 \\ W_1 &= 1 \qquad >100 \end{aligned} \qquad (5)$$

Method 4. Sample statistics M and S along with a weighted average skew coefficient were used to estimate peak flows, where $W = W_2$ is defined as

$$W_2 = N/(N + 20) \qquad (6)$$

The derivation of W_2 is explained below.

Development of weighting factor W_2. Two independent estimates of a statistic may be combined to form a better (smaller variance) estimate of the statistic than either of the two independent estimates does. The best linear combination of independent estimates is

$$G' = \frac{V_{\bar{G}}G + V_G\bar{G}}{V_G + V_{\bar{G}}} \qquad (7)$$

in which $V_{\bar{G}}$ and V_G are the variances of $\bar{G}$ and G, respectively. Assuming that G and $\bar{G}$ are independent estimates of γ, G' is a better estimate of γ than either G or $\bar{G}$. Let $W = V_{\bar{G}}/(V_G + V_{\bar{G}})$; then (4) may be written $G' = WG + (1 - W)\bar{G}$. The value of W is determined if the values of $V_{\bar{G}}$ and V_G are known. The value of $V_{\bar{G}}$ can be estimated from the map of skew coefficients as the squared value of the standard deviation of station values of skew coefficient about the isolines of the skew map. The value of $V_{\bar{G}}$ estimated for the skew map provided by the

Paper number 8W0013.

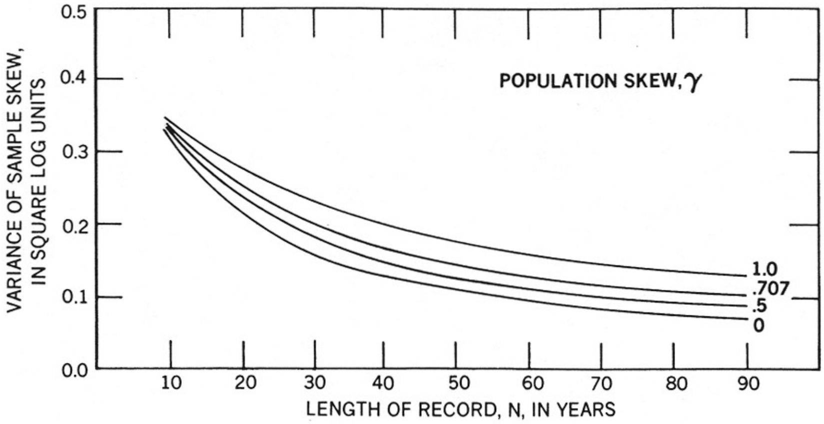

Fig. 1. Variance of sample skew as a function of N and γ. The data are from *Wallis et al.* [1974].

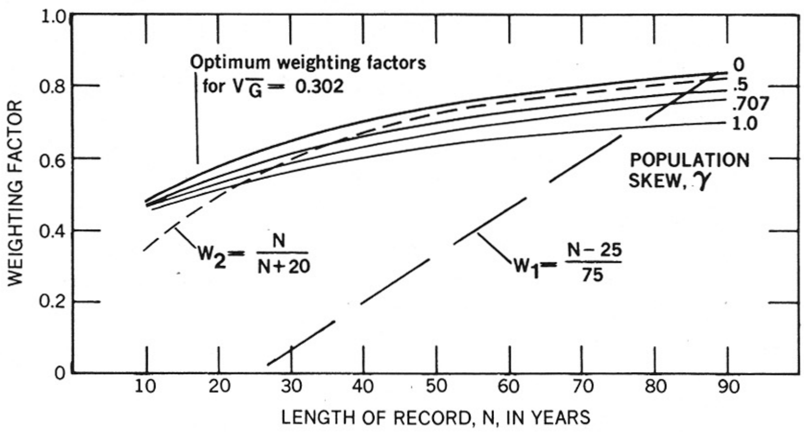

Fig. 2. Weighting factors for computing weighted averages of G and $\bar{G}$ as a function of N and γ. Optimum weighting factors assume a skew map error of 0.55.

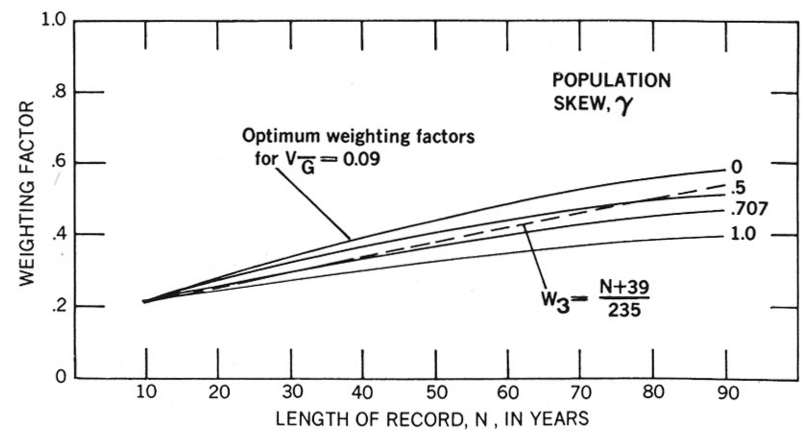

Fig. 3. Weighting factors for computing weighted averages of G and $\bar{G}$ as a function of N and γ. Optimum weighting factors assume a skew map error of 0.3.

Hydrology Committee is 0.55^2, or 0.302. The values of V_G have been defined empirically by *Wallis et al.* [1974] for specific distributions and selected values of sample size N and population skew coefficient γ. For the Pearson type 3 distribution these values are shown graphically in Figure 1. Assuming $V_{\bar{G}} = 0.302$, W as a function of N and γ is shown in Figure 2. In practice, only the value of N is known without error. Therefore an approximation of W given by $W_2 = N/(N + 20)$ is proposed. The graph of W_2 as a function of N is shown on Figure 2 along with the graph of W_1.

The functional relationship between W_2 and N is based on a value of S_m of 0.55. If another value of S_m is assumed, a new approximation of W could be made. For example, *Hardison* [1974] estimates the standard error of map skew east of the Mississippi River as 0.3. The optimal weighting factors would be as shown in Figure 3. An appoximation to the optimal

TABLE 1. Root-Mean-Square Error of 50-Year Peak Discharge for Indicated Sample Size Using Indicated Method as a Function of Map Skew and Map Skew Error

Map Skew Error	10 Samples				30 Samples				50 Samples				90 Samples			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
								Map Skew −0.4								
0.1	0.132	0.116	0.116	0.114	0.075	0.065	0.064	0.065	0.061	0.054	0.051	0.055	0.045	0.040	0.043	0.042
0.4	0.138	0.134	0.134	0.127	0.081	0.085	0.083	0.074	0.065	0.077	0.065	0.061	0.048	0.067	0.047	0.046
0.55	0.146	0.137	0.137	0.130	0.094	0.102	0.099	0.088	0.065	0.093	0.075	0.064	0.054	0.088	0.053	0.053
0.75	0.143	0.162	0.162	0.145	0.093	0.126	0.122	0.096	0.075	0.113	0.089	0.074	0.055	0.112	0.056	0.057
1.0	0.165	0.184	0.184	0.166	0.098	0.146	0.139	0.104	0.084	0.140	0.107	0.086	0.062	0.139	0.065	0.067
								Map Skew −0.2								
0.1	0.150	0.129	0.129	0.130	0.088	0.078	0.078	0.080	0.068	0.059	0.058	0.062	0.048	0.044	0.046	0.045
0.4	0.160	0.146	0.146	0.144	0.086	0.087	0.085	0.080	0.074	0.082	0.072	0.070	0.057	0.066	0.055	0.054
0.55	0.163	0.166	0.166	0.159	0.093	0.101	0.098	0.086	0.077	0.096	0.080	0.073	0.057	0.081	0.056	0.056
0.75	0.162	0.165	0.165	0.156	0.102	0.132	0.127	0.103	0.079	0.114	0.093	0.080	0.062	0.108	0.062	0.063
1.0	0.168	0.191	0.191	0.174	0.112	0.154	0.148	0.117	0.088	0.139	0.110	0.091	0.067	0.132	0.068	0.069
								Map Skew 0.0								
0.1	0.157	0.137	0.137	0.139	0.104	0.089	0.089	0.094	0.075	0.066	0.066	0.070	0.057	0.049	0.055	0.055
0.4	0.176	0.155	0.155	0.156	0.102	0.100	0.099	0.096	0.076	0.085	0.076	0.073	0.055	0.069	0.053	0.053
0.55	0.177	0.162	0.162	0.161	0.111	0.116	0.115	0.106	0.085	0.096	0.084	0.081	0.061	0.090	0.061	0.061
0.75	0.187	0.181	0.181	0.175	0.108	0.125	0.123	0.106	0.085	0.114	0.095	0.084	0.065	0.111	0.065	0.066
1.0	0.185	0.204	0.204	0.189	0.130	0.167	0.163	0.132	0.095	0.142	0.115	0.097	0.078	0.132	0.079	0.080
								Map Skew 0.2								
0.1	0.182	0.157	0.157	0.161	0.104	0.086	0.086	0.094	0.082	0.069	0.071	0.076	0.060	0.051	0.058	0.057
0.4	0.190	0.162	0.162	0.166	0.118	0.106	0.106	0.108	0.089	0.089	0.085	0.085	0.065	0.072	0.063	0.063
0.55	0.190	0.182	0.182	0.179	0.107	0.110	0.108	0.102	0.090	0.100	0.090	0.086	0.066	0.090	0.065	0.065
0.75	0.203	0.193	0.193	0.190	0.125	0.137	0.134	0.120	0.099	0.122	0.106	0.098	0.072	0.108	0.071	0.071
1.0	0.216	0.200	0.200	0.198	0.120	0.159	0.154	0.125	0.107	0.159	0.129	0.109	0.082	0.136	0.082	0.083
								Map Skew 0.4								
0.1	0.202	0.174	0.174	0.179	0.114	0.094	0.095	0.104	0.091	0.077	0.079	0.085	0.072	0.059	0.070	0.069
0.4	0.197	0.177	0.177	0.178	0.121	0.109	0.108	0.111	0.100	0.094	0.092	0.095	0.075	0.080	0.073	0.072
0.55	0.205	0.190	0.190	0.188	0.121	0.122	0.120	0.115	0.097	0.104	0.096	0.094	0.079	0.093	0.078	0.078
0.75	0.212	0.203	0.203	0.200	0.129	0.142	0.140	0.128	0.103	0.122	0.107	0.100	0.086	0.113	0.085	0.085
1.0	0.224	0.226	0.226	0.218	0.143	0.165	0.161	0.143	0.105	0.156	0.127	0.108	0.085	0.132	0.085	0.086
								Map Skew 0.6								
0.1	0.219	0.187	0.187	0.195	0.139	0.112	0.113	0.126	0.103	0.084	0.088	0.096	0.079	0.065	0.076	0.075
0.4	0.228	0.199	0.199	0.205	0.138	0.125	0.125	0.129	0.108	0.095	0.095	0.101	0.081	0.082	0.079	0.078
0.55	0.234	0.213	0.213	0.215	0.141	0.136	0.135	0.134	0.106	0.111	0.104	0.102	0.088	0.100	0.086	0.086
0.75	0.233	0.216	0.216	0.217	0.146	0.137	0.136	0.137	0.106	0.121	0.109	0.104	0.089	0.115	0.088	0.088
1.0	0.243	0.240	0.240	0.236	0.159	0.180	0.177	0.160	0.119	0.146	0.129	0.120	0.095	0.138	0.095	0.095

Numerals 1–4 denote methods.

weighting factors within the ranges of $0 \leq N \leq 90$ and $0 < \gamma < 1.0$ could be $W_3 = (39 + N)/235$. If users develop their own generalized skew coefficient, it would be necessary to estimate the accuracy of the generalized relationship in order to develop optimal weighting factors.

Experimental Results

The performance of each of the four methods described above is judged by how well each estimates peak flows. The criterion for judging performance is root-mean-square errors of the 10-, 50-, 100-, and 500-year peak flows. The root-mean-square error (rmse) for the T-year event is given by

$$\text{rmse}_T = \left[\frac{\Sigma (\hat{X}_T - X_T)^2}{500} \right]^{1/2} \qquad (8)$$

in which $\hat{X}_T$ is the estimated logarithmic value of the T-year peak and X_T is the true logarithmic value of the T-year peak of the underlying Pearson type 3 distribution.

Partial results for the 50-year peak discharge are given in Table 1. Results for the entire experiment indicate that method 4 ($W = N/(N + 20)$) yields the smallest rmse when errors in map skew are between 0.4 and 0.75. For very accurate skew maps ($S_m = 0.1$), method 2 generally yields the lowest value of rmse. For relatively inaccurate skew maps ($S_m = 1.0$), method 1 generally yields the lowest value of rmse. In addition, the results in Table 1 show that using weighting factor W_1 yields generally higher values of rmse than using sample skew without weighting when the map skew error is 0.55 or more and the skew coefficient is less than zero.

Conclusions

The sensitivity of estimated peak flows to errors in generalized skew coefficients is examined by Monte Carlo experiments. Experimental results, which were consistent over all recurrence intervals tested, are sufficient to draw the following general conclusions.

1. The optimum factor to weight G and $\bar{G}$ is a function of N, γ, and S_m.

2. Use of a generalized skew coefficient can improve the accuracy of estimated peak flows provided its accuracy is evaluated and taken into account in the weighting procedure.

3. Determining a weighted average skew coefficient using

the weighting factor recommended by the *Hydrology Committee of the Water Resources Council* [1976] often results in a poorer estimate of the population skew coefficient than using the sample skew coefficient alone.

REFERENCES

Hardison, C. H., Generalized skew coefficients of annual floods in the United States and their application, *Water Resour. Res.*, *10*(4), 745–752, 1974.

Hydrology Committee of the Water Resources Council, Guidelines for determining flood-flow frequency, *Bull. 17*, U.S. Water Resour. Counc., Washington, D. C., 1976.

Wallis, J. R., N. C. Matalas, and J. R. Slack, Just a moment!, *Water Resour. Res.*, *10*(2), 211–219, 1974. (Appendix available through National Technical Information Service, U.S. Department of Commerce, Springfield, Virginia.)

(Received July 25, 1977;
accepted January 3, 1978.)