WATER RESOURCES RESEARCH, VOL. 23, NO. 4, PAGES 715–727, APRIL 1987

Plotting Positions for Historical Floods and Their Precision

ROBERT M. HIRSCH

U.S. Geological Survey, Reston, Virginia

JERY R. STEDINGER

Department of Environmental Engineering, Cornell University, Ithaca, New York

Plotting positions are needed for situations where, in addition to a systematically recorded annual flood series, one would have a record of any large floods which occurred during an extended historical period, if they occurred. Many of the published estimators are based on uncensored sampling theory which is not appropriate for such data sets. Here such historical and systematic flood records are viewed as resulting from a partially censored sampling experiment. Plotting positions are derived for such experiments using both classical and Bayesian viewpoints. In general, it is impossible to construct highly accurate estimates of the exceedance probabilities of the largest floods using only their rank, the number of observed historical floods, and the lengths of the historical period and the systematic record. For the largest flood, the coefficient of variation of exceedance-probability estimators is of the order of 1, as it is for complete systematic records. Examples illustrate the bias and precision of a variety of plotting position formulas. The differences among the different plotting positions are generally small in comparison to the sampling variability. However, plotting positions which are unbiased with uncensored samples are often the most biased when used with a combination of historical and systematic data. Three appendices consider the effect of misspecification of the length of the historical period, the effect of misspecification of the threshold of perception or observation level, and plotting positions for situations with several perception thresholds.

INTRODUCTION

Probability plotting positions are used for the graphical display of flood peaks and serve as estimates of the probability of exceedance of those values. Probability plots allow a visual examination of the adequacy of the fit provided by alternative parametric flood frequency models. They also provide a nonparametric means of forming an estimate of the data's probability distribution by drawing a line by hand or automated means through the plotted points. While most statistically oriented hydrologists prefer efficient parameter estimation procedures (such as probability weighted moments, maximum likelihood, and Bayesian methods), the value of a visual display of the empirical frequency curve is not in dispute.

Cunnane [1978] provides a review and discussion of the statistical and hydrologic literature on this subject for complete samples. Here our focus is on the special case when, in addition to the flows observed during the period of systematic gaging, one also knows the time and approximate magnitude of the large floods which occurred during a preceding (or even subsequent) period (see, for example, *Gerard and Karpuk* [1979]). This knowledge may be derived from written records, high water marks, or geomorphological and botanical evidence. For simplicity, such floods will all be referred to here as "historical floods."

Historical flood peaks reflect the frequency of large floods and thus should be incorporated into flood frequency analyses [*Interagency Advisory Committee on Water Data*, 1982]. They can also be used to judge the adequacy of estimated flood frequency relationships. For this latter purpose, appropriate plotting positions or estimates of the average exceedance probabilities associated with the historical peaks and the remainder of the data are desired. Efficient maximum likelihood procedures for estimating the parameters of a given frequency distribution with such data are discussed elsewhere [*Cohen*, 1959, 1976; *Leese*, 1973; *Condie and Lee*, 1982; *Stedinger and Cohn*, 1986].

Although bias is examined in this paper, we will not explicitly consider the question of whether plotting positions should be mean (or median) unbiased in terms of exceedance probabilities or mean unbiased in terms of flood magnitudes for a given exceedance probability. These questions are best considered in the context of the frequency estimation procedure to be employed, the design decisions arising from it, and the losses associated with overdesign and underdesign. Such considerations are beyond the scope of this paper.

This paper emphasizes the development of a probabilistic model of flood records which include historical information, the use of that model to develop reasonable plotting position formulas, and the evaluation of those estimators as well as others. All estimators are shown to be relatively imprecise, and seldom is one estimator much better than others. Our major concerns are with the correct interpretation of the information conveyed by historical flood data and with the recognition of the limited precision of estimates of the exceedance probabilities of historical floods.

REVIEW OF PLOTTING POSITION FORMULAS FOR HISTORICAL FLOODS

A variety of formulas have been proposed for estimating the exceedance probability of flood discharges in a record of annual floods which includes systematic flood records as well as records of some historical floods. These exceedance probability estimates are generally referred to as plotting positions. The major citations on this subject include *Benson* [1950], *Dalrymple* [1960], *Beard* [1962, 1975], *Qian* [1964], *Chen et al.* [1974], *Cong et al.* [1979], *Gerard and Karpuk* [1979], *Interagency Advisory Committee on Water Data (IACWD)* [1982], and *Zhang* [1982].

Paper number 6W4477.
0043-1397/87/006W-4477$05.00

While these formulas differ, all have in common the idea that it may be possible to determine the number and magnitudes of the largest floods which occurred within a period which is longer than the systematic record. This is accomplished by careful examination of physical evidence and historical flood documentation. The floods which can be unequivocally ranked within some long historical period can be assigned plotting positions based on their rank and the extended record length. The particular formula used varies (typically, Hazen, Gringorten, or Weibull [see *Cunnane*, 1978]). The treatment of floods which cannot be ranked within this longer period, but only within the shorter systematic record, is also a matter of concern. In this paper we develop a new approach based on the probabilistic mechanism which gives rise to the records themselves.

To discuss the various formulas we introduce some standard notation. Let n be the length of the historical period (in years) over which a set of flood events can be unequivocally ranked. This period need not be continuous. Generally, it contains a systematic record period (of s years in length; $s \leq n$) during which all annual flood discharges were recorded.

The complete flood record consists of g observed floods where $n > g \geq s$. Let the floods be ranked in descending order so that $X_{(1)} \geq X_{(2)} \geq \cdots \geq X_{(g)}$. Among these floods there is a subset, $X_{(1)}$ through $X_{(k)}$, which are known to have ranks 1 through k over the investigated (historical) period of length n. These k floods may be referred to as "extraordinary floods." Some or even all may have occurred during the systematic record period. Let e be the number of extraordinary floods from the systematic record, where $e \leq k$. Note that $g = s + k - e$.

More complex cases with multiple historical record lengths are possible: sometimes it may be possible to identify a few floods as being the k largest in some long period n, while it may also be possible to identify k' floods which are the largest in a shorter period of length n'. Typically, this latter set of floods will include some in the former set as well as others of a lower magnitude. This situation arises when the information for a shorter historical period is more complete than for a longer period. Flood events of intermediate size which occurred in the distant past may have left no evidence available today, while such evidence would be available had they occurred more recently. Appendix C deals with this more complex case.

Some authors who deal with historical floods (such as *Benson* [1950], *Dalrymple* [1960], *Chen et al.* [1974], and particularly *Gerard and Karpuk* [1979]) mention the idea of a threshold. They describe a situation in which all floods greater than the threshold over the n years are known. In fact, whenever one uses historical flood data, one must be certain that the record $\{X_{(1)} \cdots X_{(g)}\}$ contains all floods during the n-year period which exceeded some particular level. If this were not so, then the accurate ranking of the largest k floods over the entire n-year period would not be possible. Thus there is always a threshold, X_0, less than $X_{(k)}$, such that the number and magnitude of all floods greater than X_0 in the n years are known.

Benson [1950] reports a case where X_0 is known precisely: it is based on a particular stage, 18 feet. All floods exceeding this stage were recorded. *Chen et al.* [1974] report the case where the threshold is known with less precision. They report that the city of Ankang, China, was destroyed by a flood in 1583. From written and available records it is known that the city existed in the same location since the Sung Dynasty (A.D. 960–1279). Since at least that time the 1583 flood is the only flood to severely damage the city. Thus we can conclude that there is a threshold, X_0 (smaller than the flood of 1583), which has only been exceeded once ($k = 1$). We do not know the precise value of X_0, although with some effort one could develop a good approximation. However, it is not necessary to determine X_0 accurately for the analysis here.

It will be assumed initially that X_0 is known and that one can unequivocally classify all floods into two subsets: the k floods greater than X_0 and the $g - k$ floods less than X_0. Appendix B considers uncertainty in X_0 estimates. It will also be assumed that n is known. In practice, it may be difficult to determine n because the completeness of documentary records may decline gradually as the investigated period extends into the remote past. The earliest documented historical flood places a lower bound on n; Appendix A considers the consequences of using that lower bound rather than the true value of n.

EXISTING METHODS

Many plotting positions for historical or extraordinary floods are based on a Weibull or a Hazen type formula. For floods below X_0, there are several such formulas with relatively minor differences. *Benson* [1950], *Cong et al.* [1979], *IACWD* [1982], and *Zhang* [1982] all propose a Weibull (denoted W herein) plotting position formula for floods above the threshold:

$$\hat{p}_i = \frac{i}{n+1} \qquad i = 1, \cdots, k \tag{1}$$

Beard [1962, 1975] proposes a Hazen (denoted H herein) plotting position formula

$$\hat{p}_i = \frac{i - 0.5}{n} \qquad i = 1, \cdots, k \tag{2}$$

Here $\hat{p}_i$ is the estimated exceedance probability of the ith largest flood. Others employ Gringorten's formula [*Natural Environment Research Council (NERC)*, 1975, pp. 81–82 and 177; *Beable and McKerchar*, 1982], which yields results slightly less extreme than Hazen's (see *Gringorten* [1963*a*] or *Cunnane* [1978]).

Benson [1950], *Cong et al.* [1979], and *IACWD* [1982] present general rules for assigning plotting positions to floods below X_0. Several other papers present examples or give rules for certain cases but do not present a universally applicable formula. Instead of a general formula, NERC presents several examples [*NERC*, 1975, Table 2.45, p. 178]. The table includes only cases where $k > e$ and the largest flood in n years occurs in the historical period. No guidance is given for cases with $k = e$ (the case where it is known that the eth largest systematic flood was not exceeded at any time in the entire $n - s$ year historical period) or for cases where $k > e$ but the highest flood occurs in the systematic record.

We have tried to infer a general rule from their examples and the description NERC provides. In general, NERC uses a *Gringorten* [1963*a*] plotting position formula. Above the threshold their rule is

$$\hat{p}_i = \frac{i - 0.44}{n + 0.12} \qquad i = 1, \cdots, k$$

Below the threshold they appear to use

$$\hat{p}_i = \frac{i - k + e - 0.44}{s + 0.12} \qquad i = k + 1, \cdots, g$$

There are at least two problems with this rule. One is nonmonotonicity: a large flood may be assigned a higher exceedance probability than a smaller flood. An example of this occurs with $s = 20$, $n = 36$, $k = 2$, and $e = 0$ (this is identical to case 35/1 in Table 2.45 of *NERC* [1975], except that s was 8 in that example). The plotting positions would be 0.0155 for the largest flood, 0.0432 for the second-largest flood, and 0.0278 for the third-largest flood. We do not know how they would deal with such a situation or other situations not covered in NERC's Table 2.45. Another problem is that large gaps can occur between the probabilities assigned to the largest floods and the $g - k$ floods less than X_0; we return to this below.

Beable and McKerchar [1982, p. 36] propose the NERC approach but only describe three cases: records with $e = 0$ (their case i), records with $k = e$ (their case ii), and records with $k = 2$, $e = 1$ (their "special case"). More guidance is needed in order to evaluate their scheme.

Three very similar Weibull-based plotting position formulas are evaluated in this paper for plotting floods less than X_0.

Benson's rule (denoted W-B) is

$$\hat{p}_i = \frac{k}{n+1} + \frac{n-k}{n+1} \cdot \frac{i-k}{s-e} \qquad i = k+1, \cdots, g \tag{3}$$

Cong et al.'s rule (denoted W-C) is

$$\hat{p}_i = \frac{k}{n+1} + \frac{n-k+1}{n+1} \cdot \frac{i-k}{s-e+1} \qquad i = k+1, \cdots, g \tag{4}$$

and the IACWD rule (denoted W-I) is

$$\hat{p}_i = \frac{k+1/2}{n+1} + \frac{n-k}{n+1} \cdot \frac{i-k-1/2}{s-e} \qquad i = k+1, \cdots, g \tag{5}$$

Only Benson discusses the particular reasons for the proposed formula. He recognized the possibility of a substantial discontinuity in the probabilities assigned to floods above the threshold and those below, if one naively applied the Weibull plotting formulas to each data set separately. He observed:

> In order to arrive at consistent results, it is necessary to obtain an array of peaks properly representative of the single long period. The known peaks, historical and recent, must be combined in the proper proportions in order to obtain such an array.

The plotting positions for systematic-record floods below the threshold must be adjusted to reflect the additional information provided by the historical flood record if the historical flood data, and the systematic record, are to be analyzed jointly in a consistent and statistically efficient manner. The next section develops a model of historical and systematic flood data which serves as the basis of statistically reasonable and consistent plotting positions.

A PROBABILISTIC MODEL OF HISTORICAL FLOOD RECORDS

Both the W and H formulas were developed for and are appropriate to complete (uncensored) samples. Our concern here is with partially censored records. That is, records which are uncensored for s years and censored (from below) for $n - s$ years. To build an appropriate probabilistic model of such flood records, we ask: Why is it that we may know today about some floods which occurred long ago? In general, the answer is that some floods were noteworthy. They may have caused severe damage. We know about them because they were all large enough to exceed some threshold of perception. If an annual flood was not large enough to be noticed or to leave marked physical evidence of its occurrence, then no record of its magnitude would persist. Thus if k extraordinary floods were observed over an n-year period, their plotting positions should not be assigned as if their only distinguishing feature is that they are the k largest floods to have occurred. They must also have been large enough to exceed some perception threshold, or else we would not be segregating them out for special attention [see *Gerard and Karpuk*, 1979].

Consider the example presented by *Zhang* [1982]. His 1851 flood was the largest flood in at least the period 1851–1980 to occur on the Mississippi River at Keokuk, Iowa; the second-largest flood in this period occurred in 1973. In neither of these years was his streamgage in operation, but there was a 24-year period between those years when his gage operated. Did the 1851 newspaper say "Largest flood to occur between now and 1980 (at least) breaks levees and floods Keokuk?" Certainly not. Newspaper editors and residents of the flood-plain took note of the 1851 flood not because it would not be exceeded for 130 years but because it was extraordinarily large. It exceeded some threshold of perception by flooding their homes, their businesses, and other significant markers.

A newspaper may have reported in 1973, "Second largest flood since 1851 inundates city." They could say that either because the 1973 flood exceeded some other flood known to be second largest since 1851, or because it was a damaging flood and no other damaging flood has occurred since 1851. In both the 1851 and 1973 flood cases, these floods were observed because they were extraordinary events. Smaller floods did not cause records to be left. The number of extraordinary floods observed during the historical period is a random variable reflecting the number of floods which happened to exceed the perception threshold.

The implicit assumption of equations (1) and (2) is that the k events are known solely because they are the k largest. If this were the case, then the data may properly be treated by rules appropriate to complete samples. Unfortunately, such estimators ignore the fact that for the k floods to have been recorded they must have been large enough ($> X_0$) to be noteworthy. We have records of these k floods because they were large, and not because they were the k largest.

Given this understanding, we suggest an alternative plotting position formula that recognizes that there is some probability of exceedance (p_e) of the threshold X_0. Thus the actual number of exceedances of X_0 in n years, k, is a binomial random variable (assuming year-to-year independence) with parameter p_e. Given that model, the true exceedance probability of the ith largest flood (p_i), $i \le k$, falls in the range $(0, p_e)$; the true exceedance probability of floods less than X_0, $i > k$, falls in the range $(p_e, 1)$. In fact, the expectation of p_i given p_e and particular values of k and e is

$$\begin{aligned} E[p_i \mid p_e, k, e] &= \frac{i}{k+1} \cdot p_e & i&=1, \cdots, k \\ E[p_i \mid p_e, k, e] &= p_e + (1-p_e) \cdot \frac{i-k}{s-e+1} & i&=k+1, \cdots, g \end{aligned} \tag{6}$$

A classical estimate of p_i can be obtained by substituting the method of moments estimator $\hat{p}_e$ of p_e into equation (6), where $\hat{p}_e = k/n$. The resulting estimator is

$$\hat{p}_i = \frac{i}{k+1} \cdot \frac{k}{n} \qquad i - 1, \cdots, k \tag{7}$$

$$\hat{p}_i = \frac{k}{n} + \frac{n-k}{n} \cdot \frac{i-k}{s-e+1} \qquad i = k+1, \cdots, g \tag{8}$$

We call this the E formula (for exceedance).

An alternative to the E formula would be a Bayesian estimator which incorporates the uncertainty as to the value of p_e. Suppose that, based on a general assessment of the physical location of the perception threshold, one describes the relative likelihood that p_e assumes different values by a beta prior probability density function:

$$\xi(p_e) \propto p_e^{\alpha'-1}(1-p_e)^{\beta'-1}$$

The omitted constant is $\Gamma(\alpha'+\beta')/[\Gamma(\alpha')\Gamma(\beta')]$. Then the posterior probability density function $f(p_e|k, n)$ for p_e, given that one observes k exceedances of X_0 in n years, is proportional to the product of $\xi(p_e)$ and the probability of k exceedances in n years:

$$f(p_e|k, n) \propto \binom{n}{k} p_e^{k}(1-p_e)^{n-k}\xi(p_e)$$

$$\propto p_e^{\alpha'+k-1}(1-p_e)^{n+\beta'-k-1}$$

[*DeGroot*, 1970]; this is just an application of Bayes theorem. Thus the posterior distribution for p_e (given k) is beta with parameters:

$$\alpha'' = \alpha' + k \tag{9}$$

$$\beta'' = \beta' + n - k \tag{10}$$

Let p_i be the exceedance probability of the ith largest of these k floods $(1 \le i \le k)$. Then the ratio p_i/p_e has a beta distribution with $\alpha = i$ and $\beta = k - i + 1$ [*Gumbel*, 1958]. Hence, conditional on given p_e and k,

$$E[p_i|p_e, k] = \frac{ip_e}{k+1} \tag{11}$$

$$\text{Var}\,[p_i|p_e, k] = \frac{i(k+1-i)p_e^2}{(k+1)^2(k+2)} \tag{12}$$

The posterior moments for p_i are

$$E[p_i] = \frac{i}{k+1}E[p_e] = \frac{i}{k+1}\frac{\alpha'+k}{\alpha'+\beta'+n} \tag{13}$$

$$\begin{aligned}\text{Var}\,[p_i] &= E\{\text{Var}\,[p_i|p_e]\} + \text{Var}\,\{E[p_i|p_e]\}\\ &= \frac{i(k+1-i)}{(k+1)^2(k+2)}\frac{\alpha''(\alpha''+1)}{(\alpha''+\beta'')(\alpha''+\beta''+1)}\\ &\quad + \frac{i^2}{(k+1)^2}\frac{\alpha''\beta''}{(\alpha''+\beta'')(\alpha''+\beta''+1)}\\ &= \frac{i(k+1-i)\alpha''(\alpha''+1)(\alpha''+\beta'') + i^2\alpha''\beta''(k+2)}{(k+1)^2(k+2)(\alpha''+\beta'')(\alpha''+\beta''+1)}\end{aligned} \tag{14}$$

The posterior variance for p_i reflects both the uncertainty as to the exact value of p_i within $[0, p_e]$ as well as the residual uncertainty as to the value of p_e given only k, n, and the prior information specified by α' and β'.

One should use a proper data-based prior for p_e when possible. However, in many cases the effort required to derive such a prior is not justified given the modest influence of the prior distribution on the posterior distribution of the plotting positions. For the problem at hand, one would expect that $p_e \cong 0$ and certainly that p_e is not near 1. Reasonable values of α' and β' would be $\frac{1}{2}$ and 9.5, respectively. Note that $\alpha' = \beta' = \frac{1}{2}$ corresponds to a reasonable "noninformative" prior [*Box and Tiao*, 1973], but here we do have some prior prejudices. With $\alpha' = \frac{1}{2}$ and $\beta' = 9.5$, the prior mean for p_e is $\frac{1}{20}$, corresponding to the 20-year flood, and p_e near 1 is very unlikely. Note that the posterior distribution of p_i is reasonably insensitive to the actual values α' and β' provided $\alpha' \ll k$ and $\beta' \ll n$. Here our interest is in $k \ge 1$ and n large (30 or more).

The same approach can be used for estimating p_i for $i > k$, resulting in the Bayesian (denoted B) plotting position formula:

$$\hat{p}_i = \frac{i}{k+1}\left[\frac{\alpha'+k}{\alpha'+\beta'+n}\right] \qquad i = 1, \cdots, k \tag{15}$$

$$\hat{p}_i = \frac{\alpha'+k}{\alpha'+\beta'+n} + \left(1 - \frac{\alpha'+k}{\alpha'+\beta'+n}\right)\cdot\frac{i-k}{s-e+1} \qquad i = k+1, \cdots, g \tag{16}$$

The prior parameters α' and β' are taken to be 0.5 and 9.5, respectively.

EXPECTATION OF p_i AND $\hat{p}_i$, $i \le k$

Returning to a classical statistical mode of thinking, one (but not the only) consideration in selecting an appropriate plotting position formula is its bias. The expectation of p_i, where $p_i = \text{Prob}\,(X \ge X_{(i)})$, for $k \ge i$ can be determined from (11) and from the distribution of k:

$$E[p_i|k \ge i] = \sum_{k=i}^{\infty} \frac{ip_e}{k+1}\cdot \Pr\,[k|k \ge i] \tag{17}$$

In (17), $\Pr\,[k|k \ge i]$ is the probability of having exactly k exceedances of X_0 in n years, given that there are i or more exceedances.

The situations of greatest importance are those where floods were observed over a relatively long period of n years and where the expected number of floods above X_0

$$E[k] = np_e \tag{18}$$

is relatively small. For independent observations, k will have a binomial distribution. However, for small np_e the binomial distribution will be indistinguishable from a Poisson distribution with

$$\lambda = E[k] = np_e \tag{19}$$

Thus

$$E[p_i|k \ge i] \simeq \sum_{k=i}^{\infty}\left\{\frac{i\lambda}{n(k+1)}\frac{\lambda^k e^{-\lambda}}{k!}\left(1 - \sum_{h=0}^{i-1}\frac{\lambda^h e^{-\lambda}}{h!}\right)^{-1}\right\}$$

$$= \frac{i\left[1 - e^{-\lambda}\sum_{j=0}^{i}\lambda^j/j!\right]}{n\left[1 - e^{-\lambda}\sum_{j=0}^{i-1}\lambda^j/j!\right]} \tag{20}$$

In the special case when $i = 1$, (20) reduces to

$$E[p_1|k \ge 1] = \frac{1-(1+\lambda)e^{-\lambda}}{n(1-e^{-\lambda})} \tag{21}$$

For $\lambda > 5$, $E[p_1|k \ge 1]$ is about $1/n$, which is approximately the Weibull plotting position. For $\lambda < 0.5$, $E[p_1|k \ge 1]$ is about $\lambda/(2n)$. When $\lambda = 1.25$, $E[p_1|k \ge 1] = 1/2n$, the Hazen plotting position.

Figure 1 shows the relationship between λ and $E[p_1|k \ge 1]$ for $n = 150$. Also shown is $E[\hat{p}_1|k \ge 1]$, where $\hat{p}_1$ is based on the W, H, E, and B formulas (equations (1), (2), (7), and (15), respectively). The expectation of $\hat{p}_1$ is determined in the same

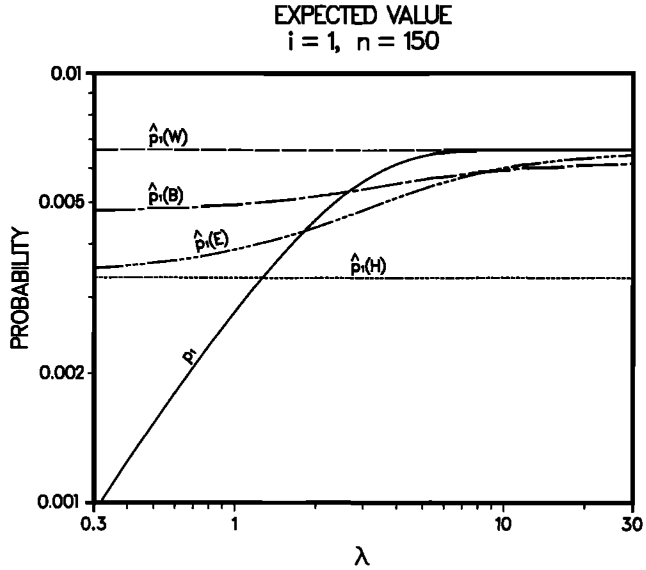

Fig. 1. Expected value of p_1 and of $\hat{p}_1$, given $k \geq 1$, for the plotting position formulas (W, Weibull; H, Hazen; E, exceedance; B, Bayesian), for $n = 150$ as a function of λ (the expected value of k).

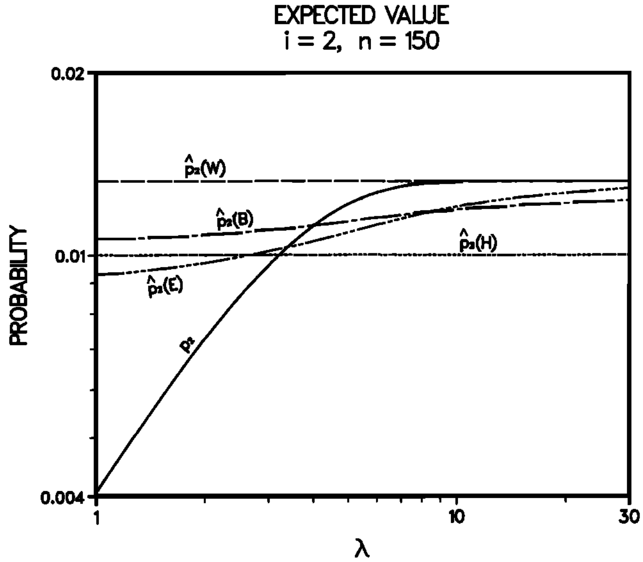

Fig. 2. Expected value of p_2 and of $\hat{p}_2$, given $k \geq 2$, for the plotting position formulas (W, Weibull; H, Hazen; E, exceedance; B, Bayesian), for $n = 150$ as a function of λ (the expected value of k).

fashion as $E[p_1 | k \geq 1]$ except that the particular formula for $\hat{p}_1$ is substituted in place of $i\lambda/n(k + 1)$ in equation (20). Of course, $\hat{p}_1$ is a constant under the W and H formulas. Under the E and B rules, $\hat{p}_1$ is a random variable because it depends on k.

For large λ (above about 7) the expectation of p_1 converges to the W plotting position. For small λ (less than about 1) the expectation of p_1 is lower than both the H and W plotting positions. The W and H plotting positions form upper and lower bounds for the expectation of the E formula, $\hat{p}_1(E)$.

The expectation of the B plotting position is similar in form to the expectation of the E plotting position except that it is less variable and as $\lambda \rightarrow 0$ it converges to $0.75/(n + 10)$ rather than $0.5/n$. Similar plots for smaller and larger values of n reveal the same pattern; they differ only in terms of the value of λ where the various W, E, B, and H curves cross or converge.

Figure 2 is the same as Figure 1 except that it is for $E[p_2 | k \geq 2]$ and $E[\hat{p}_2 | k \geq 2]$. As in Figure 1, the W plotting position is an upper bound on $E[p_2 | k \geq 2]$, and the expectation of the E plotting position approaches it for large λ. However, unlike the case for $i = 1$, the H plotting position is not a lower bound on the E plotting position.

These results do not provide a convincing case for acceptance or rejection of any of the four formulas inasmuch as no compelling argument has been made here, or in any other discussions of this topic, to demonstrate why unbiasedness is necessarily a desirable goal. It should be noted that unbiased estimation of p_i is the primary theoretical motivation for the use of the popular Weibull formula. These results show that, at least for small λ, this unbiasedness does not hold. In fact, for small values of λ (say $\lambda < 3$ for $k = 1$ or 2 and $n = 150$) the W formula is the most biased.

While Figures 1 and 2 and the preceding discussion follow the classical approach to the analysis of bias, that approach may be inappropriate if not misleading in this instance. The bias of the estimators appears to be quite large in Figure 1 when $\lambda < 1$. However, the probability of getting at least one exceedance of X_0 in $n = 150$ years when $\lambda < 1$ decreases rapidly with decreasing λ. Because these estimators are only employed if X_0 is exceeded at least once, these large biases occur in cases which rarely arise.

A second limitation with the bias analysis of Figures 1 and 2 is that it ignores the dependence of the expected value of p_i, in (11), on the value of k ($i \leq k$). Thus if k is larger, one expects for fixed p_e and fixed i that p_i will be smaller. What is really of interest is $E[p_i | k]$. Of course, this is what the Bayesian procedure provides for any given prior distribution on p_e.

VARIABILITY OF p_i FOR $i \geq k$

In addition to bias, one should consider the variability in the quantity one is attempting to estimate, p_i. For uncensored samples, the exceedance probability of the ith largest observation has a beta distribution with

$$E[p_i] = \frac{i}{(n + 1)} \tag{22}$$

$$\text{Var}\ [p_i] = \frac{i(n + 1 - i)}{(n + 1)^2(n + 2)} \tag{23}$$

[*Gumbel*, 1958]. For the most extreme event, $i = 1$, and n large,

$$E[p_i] \cong 1/n \qquad \text{Var}\ [p_i] \cong (1/n)^2$$

Thus the standard deviation of the true exceedance probability p_1 of the largest flood is about $1/n$; this also is the root-mean-square error of the estimator $p_1 = 1/n$. Thus the coefficient of variation of p_1 is nearly 1; in a very real sense, no ploting position formula (such as $(i - a)/(n + 1 - 2a)$ for some a) can accurately estimate the true exceedance probability associated with the largest observation in a random uncensored sample (see also *Gringorten* [1963*b*]). In fact, the quartiles of the distribution of p_1 are approximately $0.29/(n + 1)$ and $1.38/(n + 1)$.

The relative root-mean-square error (RRMSE) of an estimator of p_i for the uncensored case is

$$\text{RRMSE}(\hat{p}_i) = \{[E(\hat{p}_i) - E(p_i)]^2 + \text{Var}\ (p_i)\}^{1/2}/E(p_i) \tag{24}$$

For the Weibull estimator,

$$\mathrm{RRMSE}(\hat{p}_i) = \left[\frac{i(n+1-1)}{n+2}\right]^{1/2} \Big/ i \qquad (25)$$

This is approximately $1/(i)^{1/2}$ for small i and large n. For the Hazen estimator,

$$\mathrm{RRMSE}(\hat{p}_i) = \left\{\left[\frac{i}{n+1} - \frac{2i-1}{2n}\right]^2 + \frac{i(n+1-i)}{(n+1)^2(n+2)}\right\}^{1/2} \Big/ \frac{i}{n+1} \qquad (26)$$

For small i and large n, this is approximately $(4i + 1)^{1/2}/2i$.

In the censored case the errors can first be evaluated conditional upon the observed value of k. The expected value of p_i, given k, is $ip_e/(k + 1)$, and its variance is $i(k + 1 - i)p_e^2/[(k + 1)^2(k + 2)]$. Thus the mean square error (MSE) of any estimator $\hat{p}_i$ of p_i, given k, is

$$\mathrm{MSE}[\hat{p}_i \mid k] = \left[\frac{ip_e}{k+1} - \hat{p}_i\right]^2 + \frac{i(k+1-i)p_e^2}{(k+1)^2(k+2)} \qquad (27)$$

For a given value of λ (equal to np_e, which is the expected value of k), one can calculate numerically the average mean square error for $\hat{p}_i$:

$$\mathrm{MSE}[\hat{p}_i] = \sum_{k=i}^{n} \mathrm{MSE}[\hat{p}_i \mid k] \cdot \Pr[k \mid k \geq i] \qquad (28)$$

From this MSE one obtains the relative root-mean-square error $\mathrm{RRMSE}[\hat{p}_i]$ as $(\mathrm{MSE}[\hat{p}_i])^{1/2}/E[p_i]$. These values are shown in Table 1 for $i = 1$ and in Table 2 for $i = 2$. Both tables consider $n = 50$, 100, 200, and 400, and $\lambda = 1$, 2, 3, 4, 5, 10, and 20. The RRMSE in the uncensored case is in the rightmost column ($\lambda = n$). The B estimator is not considered in the uncensored case because there is no uncertainty about p_e (it equals 1.0). The case in Table 2 with $i = 2$ and $\lambda = 1$ is not shown because it is relatively improbable and gives misleadingly large RRMSE values.

The tables show that RRMSE is very insensitive to n except perhaps for W and B at low values of λ; when $\lambda \geq 10$, the RRMSE values are virtually identical to their uncensored values ($\lambda = n$). For the higher values of λ and for $i = 1$, the H method has higher RRMSE, while the other three methods have about the same. For $i = 2$ (or $i = 3$ or 4, tables not shown) with high λ, the RRMSE values are almost indistinguishable.

For low values of λ, W has distinctly higher RRMSE values than the others. Among the other three methods there is no consistently minimum RRMSE method. The fact that RRMSE values for low λ can be lower than for $\lambda = n$ (the uncensored case) is because p_i has a lower variance in the former case. Even with the substantial bias that occurs with low λ, the low variance results in small RRMSE values despite the high biases.

Tables 1 and 2 suggest that selection of a plotting position formula is a matter of settling for the "best of a bad lot." This is true also for the uncensored case, $\lambda = n$. About all that can be said is that when λ is low, the W formula will give the largest RRMSE and H wil give the lowest RRMSE. For high λ values, H has the highest RRMSE, and there is almost no difference among the other estimators.

TABLE 1. Relative Root-Mean-Square Error of p_1

		λ Values							
n	Method	1	2	3	4	5	10	20	n
50	W	1.88	0.94	0.88	0.91	0.94	0.98	0.98	0.98
50	H	0.88	0.87	0.95	1.02	1.05	1.10	1.10	1.10
50	E	1.10	0.91	0.94	0.97	0.99	0.99	0.98	0.98
50	B	1.15	0.85	0.90	0.95	0.98	1.00	1.00	···
100	W	1.91	0.95	0.88	0.91	0.94	0.99	0.99	0.99
100	H	0.88	0.87	0.96	1.02	1.06	1.11	1.11	1.11
100	E	1.11	0.92	0.95	0.98	1.00	1.00	0.99	0.99
100	B	1.28	0.87	0.90	0.94	0.97	1.00	1.00	···
200	W	1.93	0.95	0.88	0.91	0.94	0.99	1.00	1.00
200	H	0.88	0.87	0.96	1.02	1.06	1.11	1.11	1.11
200	E	1.11	0.92	0.95	0.98	1.00	1.01	1.00	1.00
200	B	1.36	0.88	0.90	0.94	0.97	1.00	1.00	···
400	W	1.94	0.95	0.88	0.91	0.95	1.00	1.00	1.00
400	H	0.88	0.87	0.96	1.02	1.06	1.11	1.12	1.12
400	E	1.11	0.92	0.95	0.98	1.00	1.01	1.00	1.00
400	B	1.40	0.89	0.90	0.94	0.97	1.00	1.00	···

W, Weibull; H, Hazen; E, exceedance; B, Bayesian.

TABLE 2. Relative Root-Mean-Square Error of p_2

		λ Values							
n	Method	1	2	3	4	5	10	20	n
50	W	···	1.19	0.70	0.62	0.63	0.68	0.69	0.69
50	H	···	0.77	0.58	0.61	0.65	0.72	0.73	0.73
50	E	···	0.80	0.64	0.65	0.67	0.70	0.69	0.69
50	B	···	0.75	0.60	0.63	0.66	0.72	0.71	···
100	W	···	1.21	0.71	0.63	0.63	0.69	0.70	0.70
100	H	···	0.77	0.58	0.61	0.65	0.73	0.74	0.74
100	E	···	0.81	0.65	0.65	0.68	0.71	0.70	0.70
100	B	···	0.85	0.62	0.62	0.65	0.71	0.71	···
200	W	···	1.23	0.72	0.63	0.64	0.70	0.70	0.70
200	H	···	0.78	0.58	0.61	0.65	0.74	0.74	0.74
200	E	···	0.81	0.65	0.66	0.68	0.72	0710	0.70
200	B	···	0.92	0.64	0.63	0.65	0.71	0.71	···
400	W	···	1.24	0.72	0.63	0.64	0.70	0.70	0.70
400	H	···	0.78	0.59	0.61	0.65	0.74	0.75	0.75
400	E	···	0.81	0.65	0.66	0.68	0.72	0.71	0.70
400	B	···	0.96	0.65	0.63	0.65	0.71	0.71	···

W, Weibull; H, Hazen; E, exceedance; B, Bayesian.

TABLE 3. Relative Root-Mean-Square Error of p_1 Computed Using Posterior Distribution for p_1 (Prior Alpha = 0.50; Prior Beta = 9.50)

		k Values							
n	Method	1	2	3	4	5	10	20	n
50	W	1.23	1.11	1.07	1.05	1.04	1.01	1.01	0.98
50	H	1.11	1.07	1.06	1.06	1.05	1.05	1.06	1.10
50	E	1.11	1.03	1.01	1.00	1.00	1.00	1.00	0.98
50	B	1.09	1.03	1.01	1.00	1.00	0.99	0.98	···
100	W	1.19	1.08	1.05	1.03	1.02	1.00	1.00	0.99
100	H	1.13	1.09	1.08	1.08	1.08	1.08	1.08	1.11
100	E	1.13	1.05	1.02	1.01	1.00	1.00	0.99	0.99
100	B	1.10	1.04	1.02	1.01	1.00	0.99	0.99	···
200	W	1.17	1.07	1.04	1.03	1.02	1.00	1.00	1.00
200	H	1.14	1.11	1.10	1.10	1.09	1.10	1.10	1.11
200	E	1.14	1.06	1.03	1.02	1.01	1.00	1.00	1.00
200	B	1.10	1.04	1.02	1.01	1.01	1.00	1.00	···
400	W	1.16	1.07	1.04	1.02	1.02	1.00	1.00	1.00
400	H	1.15	1.11	1.11	1.10	1.10	1.10	1.11	1.12
400	E	1.15	1.06	1.03	1.02	1.01	1.00	1.00	1.00
400	B	1.10	1.05	1.03	1.02	1.01	1.00	1.00	···

W, Weibull; H, Hazen; E, exceedance; B, Bayesian.

TABLE 4. Relative Root-Mean-Square Error of p_2 Computed Using Posterior Distribution for p_2 (Prior Alpha = 0.50; Prior Beta = 9.50)

		k Values							
n	Method	1	2	3	4	5	10	20	n
50	W	···	0.85	0.80	0.77	0.76	0.73	0.72	0.69
50	H	···	0.75	0.72	0.71	0.70	0.70	0.70	0.73
50	E	···	0.74	0.72	0.71	0.71	0.71	0.71	0.69
50	B	···	0.74	0.72	0.71	0.70	0.69	0.69	···
100	W	···	0.81	0.77	0.75	0.74	0.72	0.71	0.70
100	H	···	0.75	0.73	0.72	0.72	0.71	0.72	0.74
100	E	···	0.76	0.73	0.72	0.71	0.70	0.70	0.70
100	B	···	0.75	0.73	0.72	0.71	0.70	0.70	···
200	W	···	0.79	0.76	0.74	0.73	0.71	0.71	0.70
200	H	···	0.76	0.74	0.73	0.73	0.73	0.73	0.74
200	E	···	0.77	0.74	0.72	0.72	0.71	0.70	0.70
200	B	···	0.75	0.73	0.72	0.72	0.71	0.70	···
400	W	···	0.79	0.75	0.74	0.73	0.71	0.71	0.70
400	H	···	0.76	0.74	0.74	0.74	0.73	0.74	0.75
400	E	···	0.78	0.74	0.73	0.72	0.71	0.71	0.70
400	B	···	0.76	0.73	0.72	0.72	0.71	0.71	···

W, Weibull; H, Hazen; E, exceedance; B, Bayesian.

An important conclusion is that some description of the uncertainty with which p_i can be estimated may be desirable. By using a Bayesian approach we can estimate the posterior distribution and moments of p_i based on the observed values of n and k. In fact, the Bayesian approach allows for the development of interval estimates for p_i.

BAYESIAN DERIVATION OF PRECISION

The average mean square error in (28) provides a reasonable description of the average precision of exceedance probability estimators. However, as is the case with many such descriptions of average or prior precision (see *Cornfield* [1969], especially pp. 620–622 and 650–651, and *Berger* [1980, pp. 18–21, 103–104]), it fails to describe the conditional or posterior precision with which such exceedance probabilities can be estimated given the value of k actually observed. Rather than specifying the average precision in repeated trials (with different values of k but a fixed value of λ), one can use a Bayesian approach to quantify the precision with which p_1 can be specified for the particular value of k observed (and λ unknown).

Equations (13) and (14) give the posterior expectation and variance of p_i. Tables 3 and 4 present the Bayesian RRMSE of each of the estimators given various k and n. The Bayesian estimator employed $\alpha' = \frac{1}{2}$ and $\beta' = 9.5$, as did the prior distribution of p_e used to compute the RRMSE values. The use of the same α' and β' values in the construction of the B estimator and in the RRMSE calculation insures that the B estimator will have the lowest RRMSE. The sensitivity of these results (for $i = 1$) to the prior parameters α' and β' will be discussed below.

Tables 3 and 4 show a lack of sensitivity of the RRMSE to n and k if k is 3 or larger. The methods give similar RRMSE values except that W has higher RRMSE values for small k, H has higher RRMSE value for large k, and E generally comes very close to achieving the lowest possible RRMSE value for every k.

The sensitivity of these results to the selected prior is illustrated by Tables 5 and 6, which report the RRMSE(p_1) for $\alpha' = 0.3$, $\beta' = 5.7$ (Table 5) and $\alpha' = 1.0$, $\beta' = 19.0$ (Table 6). In both tables the B estimator was based on the original prior distribution, $\alpha' = 0.5$, $\beta' = 9.5$. The probability distribution of p_e under each of the three priors is displayed in Table 7 and Figure 3. Our subjective evaluation is that $\alpha' = 1.0$, $\beta' = 19.0$ is a prior which is excessively informative and that $\alpha' = 0.3$, $\beta' = 5.7$ is excessively noninformative. Together they bound a range of reasonable priors. Even with these very large differences in priors, the B estimator remains the best or within 2% of the best for all cases in Tables 5 and 6. The E estimator also remains very close to best; W remains worst for low k; H remains the worst for high k. About the only effect one can observe from changing the prior is the changes in the RRMSE values for low values of k. This occurs because when k is low, the historical record provides less information and the prior values of α and β have a larger influence on the posterior distribution of p_e.

Tables 3 and 4 (or ones developed with different priors) provide the hydrologist with a statement about the precision of an estimate of p_i based on k and n, two quantities known in

TABLE 5. Relative Root-Mean-Square Error of p_1 Computed Using the Posterior Distribution for p_1 Obtained With Relatively Noninformative Prior (Prior Alpha = 0.30; Prior Beta = 5.70)

		k Values							
n	Method	1	2	3	4	5	10	20	n
50	W	1.34	1.14	1.08	1.05	1.03	1.00	0.99	0.98
50	H	1.16	1.09	1.07	1.07	1.07	1.07	1.07	1.10
50	E	1.16	1.06	1.02	1.01	1.00	0.99	0.99	0.98
50	B	1.15	1.06	1.02	1.01	1.00	0.99	0.99	···
100	W	1.31	1.13	1.07	1.04	1.03	1.00	1.00	0.99
100	H	1.17	1.11	1.09	1.09	1.09	1.09	1.09	1.11
100	E	1.17	1.07	1.03	1.02	1.01	1.00	0.99	0.99
100	B	1.16	1.07	1.03	1.02	1.01	1.00	0.99	···
200	W	1.30	1.12	1.06	1.04	1.03	1.01	1.00	1.00
200	H	1.18	1.12	1.10	1.10	1.10	1.10	1.10	1.11
200	E	1.18	1.07	1.04	1.02	1.01	1.00	1.00	1.00
200	B	1.17	1.07	1.04	1.02	1.01	1.00	1.00	···
400	W	1.29	1.12	1.06	1.04	1.03	1.01	1.00	1.00
400	H	1.18	1.12	1.11	1.10	1.10	1.10	1.11	1.12
400	E	1.18	1.08	1.04	1.03	1.02	1.00	1.00	1.00
400	B	1.17	1.07	1.04	1.02	1.02	1.00	1.00	···

W, Weibull; H, Hazen; E, exceedance; B, Bayesian.

TABLE 6. Relative Root-Mean-Square Error of p_2 Computed Using Posterior Distribution for p_1 Obtained With Relatively Informative Prior (Prior Alpha = 1.00; Prior Beta = 19.00)

		k Values							
n	Method	1	2	3	4	5	10	20	*n*
50	W	1.05	1.05	1.05	1.05	1.05	1.05	1.05	0.98
50	H	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.10
50	E	1.03	0.99	0.99	0.99	1.00	1.02	1.04	0.98
50	B	0.99	0.99	0.99	0.99	0.99	0.99	1.00	...
100	W	1.01	1.01	1.01	1.01	1.01	1.01	1.01	0.99
100	H	1.07	1.07	1.07	1.07	1.07	1.07	1.07	1.11
100	E	1.07	1.01	1.00	0.99	0.99	1.00	1.00	0.99
100	B	1.01	1.00	0.99	0.99	0.99	0.99	0.99	...
200	W	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
200	H	1.09	1.09	1.09	1.09	1.09	1.09	1.09	1.11
200	E	1.09	1.03	1.01	1.00	1.00	1.00	1.00	1.00
200	B	1.02	1.00	1.00	1.00	1.00	1.00	1.00	...
400	W	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
400	H	1.10	1.10	1.10	1.10	1.10	1.10	1.10	1.12
400	E	1.10	1.04	1.02	1.01	1.01	1.00	1.00	1.00
400	B	1.02	1.01	1.00	1.00	1.00	1.00	1.00	...

W, Weibull; H, Hazen; E, exceedance; B, Bayesian.

practice. It does demand the adoption of a particular prior to p_e but is rather insensitive to the selected prior distribution provided it is of a reasonable shape ($\alpha \leq 1$) with most of its probability in the range $0 < p_e < 0.1$. We turn now to other issues and a few examples.

PRECISION OF PLOTTING POSITIONS FOR LARGE FLOODS BELOW X_0

Five estimators of p_i are considered for floods less than X_0 (that is for $i > k$). They are denoted W-B for Weibull-Benson (equation (3)), W-C for Weibull–Cong et al. (equation (4)), W-I for Weibull-IACWD (equation (5)), E for exceedance (equation (7)), and B for Bayesian (equation (16) with $\alpha' = 0.5$, $\beta' = 9.5$). We have not carried out a classical analysis of the bias and precision of the $\hat{p}_i$ for these smaller floods ($i > k$) for two reasons. First, the differences between the plotting position values from each of the formulas are rather small. Second, the development and presentation of results would be highly complex because of the joint probabilities of k and e for given i, n, s, and p_e. A Bayesian error analysis is less complex and more to the point. For given n, k, and $s - e$, and a prior distribution for p_e with parameters α' and β', the posterior mean of p_i is given in equation (16). The posterior variance is

$$\begin{aligned}\mathrm{Var}\,[p_i] = \{&(k+s-e-i+1)\beta''[(i-k)(\beta''+1)(\alpha''+\beta'')\\ &+(k+s-e-i+1)\alpha''(s-e+2)]\}\\ &\cdot[(s-e+1)^2(s-e+2)(\alpha''+\beta'')^2(\alpha''+\beta''+1)]^{-1} \qquad (29)\end{aligned}$$

Table 8 presents the relative root-mean-square errors for each of the five estimators for the largest observed flood less than X_0 ($i = k + 1$, where $k = 1, 2, 3, 4, 5, 10, 20, n = 150$, and $s - e = 10, 25, 50, 100$). The table shows that except where k is very small and $s - e$ large, all formulas except W-I have virtually equal relative root-mean-square errors; W-I has a RRMSE as much as 11% higher than the other formulas. In the low k, high $s - e$ cases the B formula has about a 2% lower RRMSE than the W-B, W-C, or E formulas. When the choice of priors is changed, without changing α' or β' in the B estimator, this apparent advantage vanishes, and the W-B, W-C, E, and B estimators yield the same RRMSE values. Tables for $i = k + 2$ and $k + 3$ reveal similar patterns with reduced RRMSE values and smaller differences between W-I and the other formulas.

When $i = g$ (the smallest flood), W-B will have a RRMSE which is large in comparison to W-C, B, or E, and W-I will have RRMSE which approximates that of W-C, B, or E. However, these cases are of little concern.

TABLE 7. Prior Probabilities That $1/p_e$ Falls in Different Intervals, as a Function of α' and β', the Parameters of the Three Prior Distributions Used in Tables 3–6

	$\alpha' = 1.0$, $\beta' = 19.0$	$\alpha' = 0.5$, $\beta' = 9.5$	$\alpha' = 0.3$, $\beta' = 5.7$
$1/p_e > 1000$	0.019	0.108	0.232
$100 < 1/p_e < 1000$	0.155	0.226	0.226
$10 < 1/p_e < 100$	0.691	0.503	0.375
$1/p_e < 10$	0.135	0.163	0.167

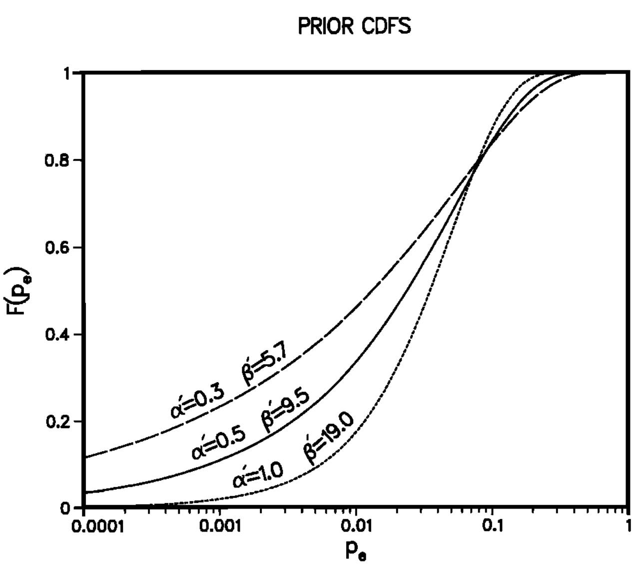

Fig 3. Three prior cumulative distribution functions for p_e, the probability of exceedance of the threshold.

TABLE 8. Relative Root-Mean-Square Error of p_i for the Largest Flood Below X_0, $i = k + 1$ (Prior Alpha = 0.50; Prior Beta = 9.50)

		k Values						
$s - e$	Method	1	2	3	4	5	10	20
10	W–B	0.832	0.784	0.741	0.702	0.667	0.531	0.372
10	W–C	0.830	0.782	0.739	0.700	0.665	0.528	0.368
10	W–I	0.923	0.868	0.818	0.773	0.733	0.576	0.393
10	E	0.830	0.782	0.739	0.700	0.665	0.528	0.368
10	B	0.830	0.782	0.739	0.700	0.665	0.528	0.368
25	W–B	0.788	0.704	0.636	0.582	0.536	0.388	0.256
25	W–C	0.789	0.705	0.637	0.582	0.536	0.387	0.255
25	W–I	0.872	0.774	0.697	0.633	0.580	0.410	0.259
25	E	0.789	0.705	0.637	0.582	0.536	0.388	0.256
25	B	0.787	0.703	0.636	0.582	0.536	0.387	0.254
50	W–B	0.715	0.610	0.536	0.481	0.437	0.313	0.214
50	W–C	0.716	0.611	0.537	0.481	0.438	0.313	0.214
50	W–I	0.776	0.665	0.570	0.507	0.458	0.319	0.212
50	E	0.716	0.611	0.536	0.481	0.438	0.313	0.215
50	B	0.710	0.607	0.534	0.480	0.437	0.313	0.212
100	W–B	0.655	0.546	0.476	0.427	0.390	0.286	0.203
100	W–C	0.656	0.547	0.477	0.427	0.390	0.286	0.203
100	W–I	0.679	0.560	0.485	0.433	0.394	0.286	0.202
100	E	0.655	0.546	0.476	0.427	0.389	0.286	0.204
100	B	0.640	0.539	0.472	0.425	0.389	0.286	0.201

W–B, Weibull–Benson; W–C, Weibull–Cong et al.; W–I, Weibull–IACWD; E, exceedance; B, Bayesian.

SELECTED EXAMPLES OF THE DISTRIBUTION OF p_i AND THEIR RELATIONSHIP TO $\hat{p}_i$

IACWD [1982] presents an example data set to demonstrate the plotting position and frequency estimation techniques proposed therein. The example is from the Big Sandy River at Bruceton, Tennessee. They report the following information (expressed in our notation): $n = 77$, $s = 44$, $k = 3$, $e = 0$. Using this information and a prior distribution for p_e with $\alpha' = 0.5$ and $\beta' = 9.5$, for $i = 1, 2, 3$, and 4, we computed the p_i by all available methods and estimated the 0.05, 0.25, 0.5, 0.75, and 0.95 quantiles of the p_i distribution. For $i = 1, 2, 3$ this was done by a Monte Carlo experiment using the appropriate values of n, $s - e$, k, α', and β'; for $i = 4$ this was done by approximating the posterior distribution of p_4 by a beta distribution with the appropriate mean and variance. These distribution quantiles and $\hat{p}_i$ values are shown in Figure 4. Also shown are the quantiles with different prior distributions. Figure 4 suggests several things. First, $F(p_i)$ is rather insensitive to the prior distribution for p_e, with the greatest sensitivity occurring for $i = 1$. For $i < k$ the W formula plotting positions are very nearly at the seventy-fifth percentile of the posterior distribution. The B formula is by definition the posterior mean of p_i; it is always greater than the posterior median of p_i. The H formula p_i is rather inconsistent with respect to $F(p_i)$: for $i = 1$ it approximates the median, for $i = 2$ it approximates the mean, and for $i = 3$ it exceeds the mean. In all cases the E formula is very close to the posterior mean of p_i. For $i = 4$ the five plotting positions cluster quite closely together, with W-B, W-C, and E below the mean and above the median and W-I below the median.

Figure 5 is a similar description of another case: $n = 150$, $s = 50$, $e = 0$, $k = 1$. This example maximizes the importance

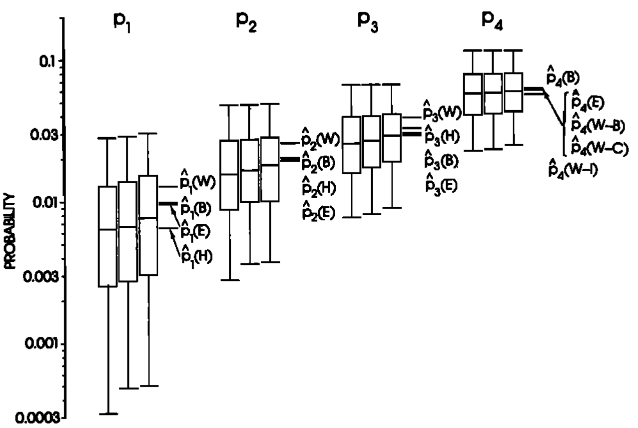

Fig. 4. Box plots (5, 25, 50, 75, and 95 percentiles) of the posterior distribution of p_i for $i = 1, 2, 3$, and 4, $n = 77$, $k = 3$, $s - e = 44$ for three different prior distributions for p_e. The parameters of the prior (α', β') are (left) 0.3, 5.7, (center) 0.5, 9.5, and (right) 1.0, 19.0. The horizontal lines are the point estimates of p_i.

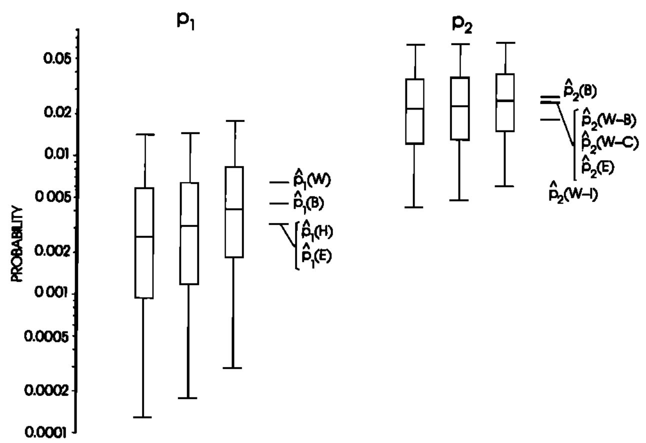

Fig. 5. Box plots (5, 25, 50, 75, and 95 percentiles) of the posterior distribution of p_i for $i = 1$ and 2, $n = 150$, $k = 1$, $s - e = 50$ for three different prior distributions for p_e. The parameters of the prior (α', β') are (left) 0.3, 5.7, (center) 0.5, 9.5, and (right) 1.0, 19.0. The horizontal lines are the point estimates of p_i.

of the choice of a prior for p_e (the influence of the prior is greatest when $k = 1$). For the largest flood, the W $\hat{p}_1$ estimate lies above the upper quartile given the prior with $\alpha' = 0.5$ and $\beta' = 9.5$ as well as the prior with $\alpha' = 0.3$ and $\beta' = 5.7$. For the other prior, the estimate is between the median and upper quartile, but very close to the upper quartile. Both the H and the E $\hat{p}_1$ values happen to be close to the posterior median. The differences in the posterior distributions are small compared to their spread, even in this worst case. For $i = 2$ the influence of the selected prior is very small. W-B, W-C, and E values of $\hat{p}_2$ are very close together (but not equal) and are intermediate between the posterior mean and median of p_2. The W-I value of $\hat{p}_2$ lies below the median. For all situations considered (many more than those shown in these two figures), plotting according to the formula of the *IACWD* [1982] would be internally very inconsistent, with $\hat{p}_i$, $i = 1, \cdots, k$, being approximately at the upper quartile of p_i, and $\hat{p}_i$, $i > k$, being below the median of the p_i.

These examples demonstrate that the W or H formulas, based on uncensored sampling theory, result in inadequate point estimates of p_i for the cases considered here. They occur at a variety of positions with respect to the posterior distribution of p_i, even outside the interquartile region in some cases. For those uncomfortable with a Bayesian estimator because of the reliance on a prior distribution, the E estimator is ideal. It generally falls relatively close to the posterior mean under a wide variety of priors and has its origin in a maximum likelihood estimate of the parameter of a binomial process. For those who are comfortable with a Bayesian approach, the B estimator with the prior of one's own choice is appropriate. If one prefers a median unbiased estimator, one can use the Bayesian approach to find such an estimate.

Even more important than these conclusions is that any estimate of p_i is very imprecise, especially for the important small-i case. Figures 4 and 5 demonstrate that any point estimate of p_i, based on i, n, k, and s, has a potential for being substantially in error. The fifth and ninety-fifth percentiles of the posterior distribution of p_1 are approximately two orders of magnitude apart. Tables 3, 5, and 6 also demonstrate this imprecision in that regardless of which estimator is used, the relative root-mean-square error is approximately 1. A correct understanding of the mechanism giving rise to historical flood records makes it possible to describe this imprecision in the form of intervals containing p_i with specified probabilities.

CONCLUSIONS

This paper has addressed a host of issues related to the assignment of plotting positions to partially censored data sets. An early section of the paper set the foundation for the logical assignment of plotting positions to historical and systematic-record floods in the partially censored flood records available in practice. That analysis led to the development of the exceedance estimator based upon a maximum likelihood estimator of p_e and the Bayesian estimator based upon the posterior distribution of p_e, given the values of k and n and a prior distribution for p_e.

Subsequent sections examined the bias and the precision of these estimators and of the Weibull and Hazen formula which other hydrologists have employed. While our new plotting position formulas provide some increases in precision, the analysis generally showed that it is impossible to make very accurate estimates of the exceedance probability p_i of the largest floods using only i, n, k, and s. Given that point estimates of p_i (based on i, n, and k) are so inaccurate, it becomes important to consider the likely range in which the actual value of p_i may fall. Such an analysis requires an appreciation of the sampling experiment giving rise to the partially censored flood series and is easily performed within a Bayesian statistical context. Separate sections in the appendix address (1) the appropriate choice of n and the consequences of values often employed, (2) the impact of uncertainty as to the value of X_0, and (3) the derivation of plotting positions for situations with several censoring thresholds.

APPENDIX A: THE EFFECT OF MISSPECIFICATION OF n

The parameter n is the length of the period during which any floods greater than X_0 would have left evidence which would be discovered in the course of an investigation of historical floods. The factors which determine the beginning of this period are the status of occupation and use of the floodplain and the development of record keeping in the particular area. The E and B formulas presented in this paper are based on the sampling of a binomial distribution where it is assumed that the beginning and ending of the sampling period are fixed. The W and H formulas are "borrowed" from noncensored sampling procedures but carry with them the implicit assumption that the beginning and ending of the sampling period are fixed and not a function of the magnitude of floods which occurred within those n years.

In many of the literature examples (specifically, *Benson* [1950], *Dalrymple* [1960], *IACWD* [1982], and *Zhang* [1982]), one finds that n is always taken to be the period from the first extraordinary flood to the present. For example, *IACWD* [1982] states in their example that the Big Sandy River at Bruceton, Tennessee, has flood records for 1897, 1919, 1927, and 1930–1973. They then consider the historical period to be 1897–1973 ($n = 77$ years). This conclusion could only be reasonable if those who investigated this flood record found that their information sources (newspapers, letters, diaries, government or business records) reached back only to 1897 and no earlier. While these examples do not indicate what those authors would do in all cases, the NERC study is more specific. If a single extraordinary flood as observed $n - s$ years before the beginning of an s-year systematic record, then their rule is [*NERC*, 1975, p. 177] to plot that flood as "the largest in a sample of size" n. We see no basis for this presumption.

The question that the investigator needs to ask is: "Is it possible that there would have been an extraordinary flood in 1896 and I would not have found out about it?" We would suggest that the answer is very likely "no." If there had been one, then information about the 1897 flood would probably have taken note of these "back-to-back" events. One could, in fact, go back in time year-by-year posing the question: "Would I know about an extraordinary flood in this year if one had occurred?" At some point, as history gets sketchier, the subjective answer will become "It is quite possible that an extraordinary flood in this year would leave no record which my investigation would uncover." It is this answer which should determine n. The common practice of using the earliest known flood year as the beginning of the record will mean that the n value selected is a lower bound on the true n and thus, under any of the plotting position formulas, the $\hat{p}_i$ (for small i) will be higher than it should be with the correct n. Furthermore, n would be a random variable which is dependent on the flood-producing process itself; that would be a

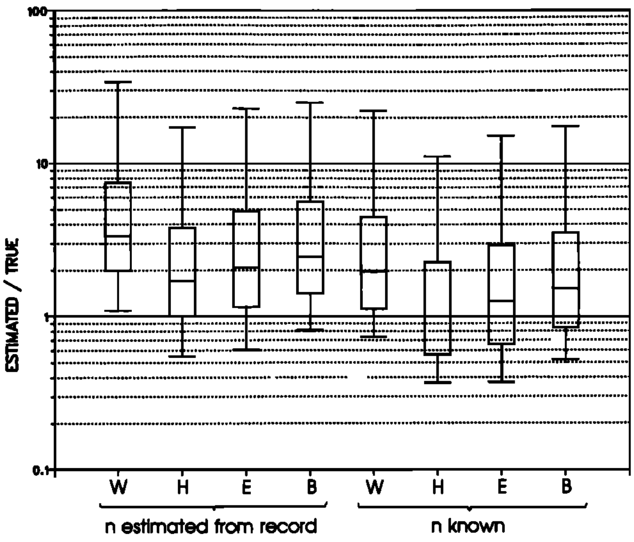

Fig. 6. Box plots (5, 25, 50, 75, and 95 percentiles) of the ratio $\hat{p}_1/p_1$ for each of four plotting position formulas (W, H, E, and B), given either the true record length ($n = 150$) or $\hat{n}$, the estimated record length based on the time of the first extraordinary flood or first systematic flood ($s = 50$).

violation of the assumption of all of the plotting position formulas.

To illustrate the magnitude of this problem, a Monte Carlo experiment was performed. Flood records with $n = 150$ and $s = 50$ years were generated (100 years of historical information followed by 50 years of systematic record). The value of p_e was set to 0.01. Only those records with $k \geq 1$ were considered. For each record, $\hat{n}$ rather than n was employed to calculate the plotting positions, where $\hat{n}$ was set to the number

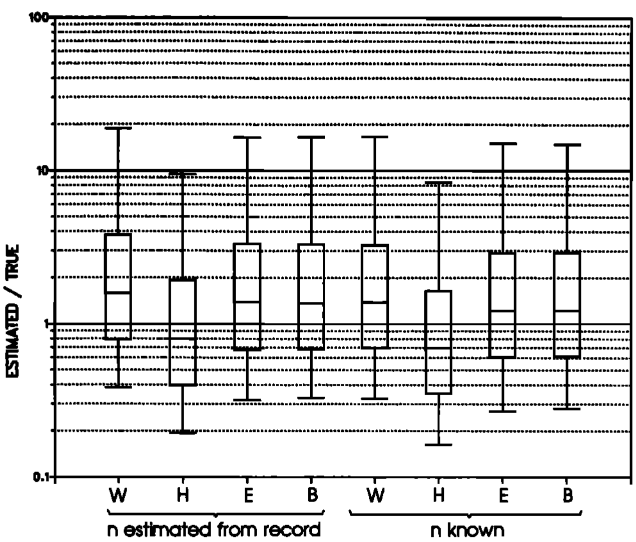

Fig. 7. Box plots (5, 25, 50, 75, and 95 percentiles) of the ratio $\hat{p}_1/p_1$ for each of four plotting position formulas (W, H, E, and B), given either the true record length ($n = 150$) or $\hat{n}$, the estimated record length based on the time of the first extraordinary flood or first systematic flood ($s = 50$).

of years from the first exceedance to the end of the record, inclusive, or to 50 if the first exceedance of X_0 occurred during the systematic record. The value of $\hat{p}_1$ was computed by each of the formulas (W, H, E, and B) using n and also using $\hat{n}$ in place of n. Figure 6 shows box plots (of the 5, 25, 50, 75, and 95 percentiles) of the ratio $\hat{p}_1/p_1$ for each formula. In the case of the W formula, using $\hat{n}$, there is a greater than 0.95 probability that $\hat{p}_1 > p_1$ and a greater than 0.5 probability that $\hat{p}_1 > 3p_1$. What the figure shows is that with this method of estimating n, all plotting positions have severe upward bias with a greater than 0.75 probability that $\hat{p}_1 > p_1$. Figure 7 is based on a similar Monte Carlo experiment with $p_e = 0.05$. In this case the effect of the errors in n are much less severe because $\hat{n}$ is generally only slightly less than n. Clearly, the estimation of n based on the date the first extraordinary flood occurred exacerbates the severe imprecision of any of the plotting position formulas and the severe upward bias that exists in the W formula and less so in the E and B formulas. Every effort should be made to establish n accurately on the basis of the quality of historical evidence and not on the basis of the occurrence of the first extraordinary flood.

APPENDIX B: EFFECT OF THE UNCERTAINTY IN X_0

As stated earlier, the threshold X_0 may not be precisely known, but it may be bounded given the flood records available. X_0 must be smaller than the smallest flood observed in the nonsystematic period of the record, but how much smaller? If there is a systematic-record flood just slightly smaller than the smallest nonsystematic-record flood, should the large systematic record flood be classified as being above the threshold or below the threshold? When X_0 is not known and cannot be determined from the documentary evidence available, then a convenient approximation is to let X_0 be slightly less than the smallest flood in the nonsystematic record. The result of this is that one or more systematic record floods which exceeded the true threshold may be classified as being below the threshold. Thus the presumed value of k may be too low.

In order to assess the effect of such a misclassification of floods in the systematic record, a Monte Carlo experiment was run with $n = 150$, $s = 50$, and $p_e = 0.01$ or 0.02. Using the plotting position rules W, E, and B, the plotting position of the largest flood in the 50-year systematic record was evaluated. Two thousand repetitions were run, first with the true value of X_0 known and then with X_0 estimated as being just less than the smallest nonsystematic-record flood. For all of the rules, the average plotting position was about 8–10% higher when the estimated X_0 was used than when the true X_0 was used. Of course, in most samples this uncertainty in X_0 did not affect the specified plotting positions, while in a relative few, large differences sometimes occurred. Given the uncertainty inherent in these estimates, this additional error is not particularly troublesome. However, it does demonstrate the desirability of attempting to ascertain the true value of X_0 rather than relying on the smallest historical flood.

APPENDIX C: CASES OF MULTIPLE THRESHOLDS

It is possible that for different segments of the historical period there may be different thresholds. There may be a long early period with a very high threshold, followed by a later period with a lower one which is, in turn, followed by a period of systematic record. This scenario is plausible for two reasons. The first is the progressive intensification of floodplain development. As communities become more populated and develop over time, there may be increased settlement of more flood-prone locations as the less flood-prone locations become occupied. The second is more detailed record keeping in later years. At an earlier stage of community development there is little in the way of organized record keeping, and only events of catastrophic magnitude would be recorded permanently. In later years the government, the press, and industry may become more organized and collect and record more events, including floods which cause only modest amounts of damage. *Gerard and Karpuk* [1979] provide a multiple threshold example.

A set of plotting positions can be determined in this multiple-threshold case. The approach here follows that of the E formula, and the E formula is a special case of this more general rule. However, a different notation is more convenient.

Define a series of thresholds X_1, X_2, $\cdots$, X_m such that $X_1 < X_2 < \cdots < X_m$. In the case where some systematic record exists, X_1 will equal zero. That is, systematic sampling is viewed as a special case of censored sampling with the threshold equal to zero. For convenience, define $X_{m+1} = \infty$.

For each threshold j ($j = 1, 2, \cdots, m$), consider two variables. A_j is the number of floods Q, such that $X_j \le Q < X_{j+1}$. B_j is the number of floods Q, where $Q \le X_j$. Note that the total number of floods in the complete record that can be ranked and plotted is

$$G = \sum_{j=1}^{m} A_j$$

Thus if $j = 1$, G is equivalent to g (or $k + s - e$) in the body of the paper. The number of systematic-record floods below the lowest nonzero threshold (X_2) is B_1. This is equivalent to $s - e$ in the body of the paper.

Our aim is to estimate the probability of exceedance of each of the m thresholds, p_{ej}. From these values we can estimate the exceedance probability p_i for each observed flood in a manner which assures that the estimates $\hat{p}_{ej}$ and $\hat{p}_i$ have the properties

$$\hat{p}_{ej} \le \hat{p}_i < \hat{p}_{ej+1}$$

whenever $X_j \le Q_i < X_{j+1}$ and $\hat{p}_1 < \hat{p}_2 < \cdots < \hat{p}_G$. The probability of exceedance of a threshold p_{ej} is defined as

$$p_{ej} = \text{Prob}\,[Q \ge X_j]$$

This can be reexpressed as

$$p_{ej} = \text{Prob}\,[Q \ge X_{j+1}] + \text{Prob}\,[X_j \le Q < X_{j+1} \,|\, Q < X_{j+1}] \cdot \text{Prob}\,[Q < X_{j+1}]$$

$$p_{ej} = p_{e,j+1} + \text{Prob}\,[X_j \le Q < X_{j+1} \,|\, Q < X_{j+1}] \cdot (1 - p_{e,j+1})$$

This formulation suggests the possibility of recursive estimation. We can begin by estimating the probability of exceedance of the highest threshold and work downward. The conditional probability Prob $[X_j \le Q < X_{j+1} | Q < X_{j+1}]$ is readily estimated by the method of moments (identical to the maximum likelihood estimator) as $A_j/(A_j + B_j)$. Thus

$$\hat{p}_{ej} = \hat{p}_{e,j+1} + \left(\frac{A_j}{A_j + B_j}\right) \cdot (1 - \hat{p}_{e,j+1})$$

which may be solved for $j = m, m - 1, \cdots, 2, 1$. Note that $\hat{p}_{e,m+1} = 0$, because $X_{m+1} = \infty$. When there is some systematic record, $\hat{p}_{e1} = 1.0$ because B_1 must equal zero.

The assignment of specific plotting positions to the G known floods can be generalized to the formula

$$\hat{p}(i) = (1 - \hat{p}_{ej}) + (\hat{p}_{ej} - \hat{p}_{e,j+1}) \cdot \frac{r - a}{A_j + 1 - 2a}$$

where the flood Q_i is in the range $X_j \le Q_i < X_{j+1}$ and r is the rank of the ith flood among the A_j floods in that range; note that

$$r = i - \sum_{q<j} A_q$$

The constant a is a number in the range (0, 0.5). If $a = 0$, the spacing between plotting positions is according to Weibull; if $a = 0.44$, the spacing is Gringorten; and if $a = 0.5$, the spacing is Hazen.

Acknowledgment. The authors would like to express their appreciation to Timothy A. Cohn for reformulating the multiple threshold case.

REFERENCES

Beable, M. E., and A. I. McKerchar, Regional flood estimation in New Zealand, *Water & Soil Tech. Publ. 20*, Min. of Works and Dev., Wellington, New Zealand, 1982.
Beard, L. R., Statistical methods in hydrology, *Rep. CW-151*, Civ. Works Invest. Proj., U.S. Army Corps of Eng., Sacramento, Calif., 1962.
Beard, L. R., *Hydrologic Engineering Methods for Water Resources Development*, vol. 3, *Hydrologic Frequency Analysis*, Hydrologic Engineering Center, U.S. Army Corps of Engineers, Davis, Calif., 1975.
Benson, M. A., Use of historical data in flood-frequency analysis, *Eos Trans. AGU*, *31*(3), 419–424, 1950.
Berger, J. O., *Statistical Decision Theory: Foundations, Concepts, and Methods*, Springer-Verlag, New York, 1980.
Box, G. E. P., and G. C. Tiao, *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading, Mass., 1973.
Chen, C., Y. Yeh, and W. Tan, The important role of historical flood data in the estimation of spillway design floods, report, Min. of Water Conservancy and Electr. Power, Peking, China, June 1974.
Cohen, A. C., Jr., Simplified estimators of the normal distribution when samples are singly censored or truncated, *Technometrics*, *1*(3), 217–237, 1959.
Cohen, A. C., Jr., Progressively censored sampling in the log-normal distribution, *Technometrics*, *18*(1), 347–351, 1976.
Condie, R., and K. Lee, Flood frequency analysis with historic information, *J. Hydrol. Amsterdam*, *58*, 47–61, 1982.
Cong, Shuzheng, et al., Statistical testing research on the methods of parameter estimation in hydrological computation, report, 27 pp., East China Coll. of Hydraul. Eng., Nanking, China, 1979.
Cornfield, J., The Bayesian outlook and its application, *Biometrics*, *25*(24), 617–657, 1969.
Cunnane, C., Unbiased plotting position—A review, *J. Hydrol. Amsterdam*, *37*, 205–222, 1978.
Dalrymple, T., Flood-frequency analyses, *U.S. Geol. Surv. Water Supply Pap.*, *1543-A*, 1960.
DeGroot, M. H., *Optimal Statistical Decisions*, McGraw-Hill, New York, 1970.
Gerard, R., and E. W. Karpuk, Probability analysis of historical flood data, *J. Hydraul. Div. Am. Soc. Civ. Eng.*, *105*(HY9), 1153–1165, 1979.
Gringorten, I. I., A plotting rule for extreme probability, *J. Geophys. Res.*, *68*, 813–814, 1963*a*.
Gringorten, I. I., Envelopes for order observations applied to extreme meteorological extremes, *J. Geophys. Res.*, *68*(3), 815–826, 1963*b*.
Gumbel, E. J., *Statistics of Extremes*, Columbia University Press, New York, 1958.
Interagency Advisory Committee on Water Data (IACWD), Guidelines for determining flood flow frequency, *Bull. 17B*, Hydrol. Subcomm., Office of Water Data Coord., U.S. Geol. Surv., Reston, Va., 1982.

Leese, M. N., Use of censored data in the estimation of Gumbel distribution parameter for annual maximum flood, *Water Resour. Res.*, *9*(6), 1534–1542, 1973.

Natural Environment Research Council (NERC), *Flood Studies Report*, vol. 1, London, 1975.

Qian, T., The determination of plotting position of the flood in the presence of historical flood data (in Chinese), *Shui Li Xuebao*, *2*, 50–54, 1964.

Stedinger, J. R., and T. Cohn, Flood frequency analysis with historical and paleoflood information, *Water Resour. Res.*, *22*(5), 785–793, 1986.

Zhang, Y., Plotting positions of annual flood extremes considering extraordinary values, *Water Resour. Res.*, *18*(4), 859–864, 1982.

R. M. Hirsch, U.S. Geological Survey, WGS-Mail Stop 410, Reston, VA 22092.

J. R. Stedinger, Department of Environmental Engineering, Cornell University, Ithaca, NY 14853.

(Received March 14, 1986;
revised November 18, 1986;
accepted November 19, 1986.)