Fallacy of the 500-year Flood: A Cautionary Note on Flood Frequency Analysis

Scott A. Lecce
Department of Geography
East Carolina University

The flood of 1999 on the Tar River in eastern North Carolina was the largest in nearly 100 years of stream flow records, where recurrence interval estimates at several gaging stations exceeded 500 years. Nevertheless, the estimation of recurrence intervals for low frequency, large magnitude floods involves considerable uncertainty. This paper uses annual flood records from four gaging stations in the Tar River basin to demonstrate the level of inaccuracy associated with flood frequency analysis (FFA). The margin of error (90% confidence interval) for recurrence interval estimates of large floods on the Tar River are suggestive of the inaccuracy of flood frequency curves, which show that the 100-year flood may be under or overestimated by as much as 1.5-2 times. Although FFA is necessary for the effective management of floodplains, estimates of discharge for various recurrence intervals should be evaluated in the context of several significant limitations: they are often based on short records, the underlying assumptions are routinely violated, and the margins of error are usually large.

Introduction

Torrential rainfall from hurricanes Dennis and Floyd produced the great flood of 1999 that was the most costly disaster in the history of North Carolina. Many news accounts touted the event as the "flood of the century" and reported the probability of experiencing such an event as one in 400 or 500 years (e.g., Royal 2000). Unquestionably, the magnitude of the flood was exceptionally large, however, the temptation to assign a probabilistic definition to the flood provides an opportunity to reexamine flood frequency analysis and the accuracy of recurrence interval estimates. The accuracy with which low frequency, large magnitude events like the 100-year flood are estimated has important implications for floodplain management because current federal flood insurance programs are linked to the 100-year floodplain. The purpose of this paper is to assess the difficulties inherent in flood frequency analysis using examples drawn from the Tar River basin in North Carolina.

Background: A Primer on Flood Frequency Analysis

Flood frequency analysis (FFA), or extreme value analysis, is based on the notion that the magnitude and frequency of extreme events can be estimated by fitting theoretical probability distributions to flood events (Gumbel 1941, 1958). Estimates are made of the probability that a certain discharge will be equaled or exceeded in any given year. This is usually applied to the *annual flood*, the largest peak discharge of each year of record, and is expressed as the exceedence probability p. The inverse of the exceedence probability ($1/p$) is the *recurrence interval* or *return period* T. For example, if the calculated exceedence probability for a peak discharge of 50,000 ft^3/s is 0.02, there is a 2% chance in any given year that this discharge will be equaled or exceeded. The recurrence interval for this example is 50 years, which means that over a long period of time the 50-year flood (50,000 ft^3/s) will occur an average of once every 50 years. It is important to recognize that the recurrence interval implies nothing about the time sequence of floods. In other words, it does not mean that the 50-year flood will occur exactly every 50 years. In fact, the 50-year flood could be equaled or exceeded in successive years or more than once in the same year.

Although the recurrence interval and the exceedence probability are the most commonly used probability estimates associated with floods, they only provide probabilities for individual years. If we want to know the probability of a flood magnitude occurring once over some longer timperiod, then:

(1) $$P = 1-(1-p)^n$$

where P is the probability that an event will occur *once* during a time period of n years and p is the exceedence probability. The probability of more than one event occurring over a time period greater than one year could also be calculated using the binomial distribution:

(2) $$P = \left(\frac{n!}{y!(n-y)!}\right) p^y(1-p)^{n-y}$$

where P is the probability that an event will occur *more than once* during a time period greater than one year, n is the time period (years), y is the number of occurrences (i.e., floods), and p is the exceedence probability. For example, using equation 1 the probability of the 50-year flood occurring once during the duration of a typical home mortgage (30 years) is 0.45 (Table 1). So while there is only a 2% chance of experiencing a 50-year flood in a single year, the chance of this event occurring once over a 30 year period is much higher (45%). During a human life time of 70 years, there is more than a three-in-four chance that the 50-year flood will be equaled or exceeded (p = 0.76). Even a 500-year flood with a 0.2% chance of occurring in a single year, has a 13% chance of occurring once during a 70 year time span. It seems likely that the public perception of flood risk would be quite different if probabilities were stated using equation 1 because it demonstrates that although large floods may be unlikely in any single year, the odds are fairly high that a large flood will occur over an extended period of time.

The objective of FFA is to relate the magnitude of flood events to their frequency of occurrence through the use of probability distributions (Chow et al. 1988). Although many statistics are based on the normal distribution, flood series are not normally distributed. Instead, flood distributions are negatively

Table 1. Percentage probability of the N-year flood occurring during a particular time span.

Time Span (yrs)	N = Return Period (yrs)							
	5	10	20	50	100	200	500	1000
1	20	10	5	2	1	-	-	-
2	36	19	10	4	2	1	-	-
5	67	41	23	10	5	2	1	-
10	89	65	40	18	10	5	2	1
20	99	88	64	33	18	10	4	2
30	-	96	79	45	26	14	6	3
50	-	99	92	64	39	22	10	5
100	-	-	99	87	63	39	18	10
200	-	-	-	98	87	63	33	18
500	-	-	-	-	99	92	63	39
1000	-	-	-	-	-	99	86	63

Modified from Smith and Ward (1998).

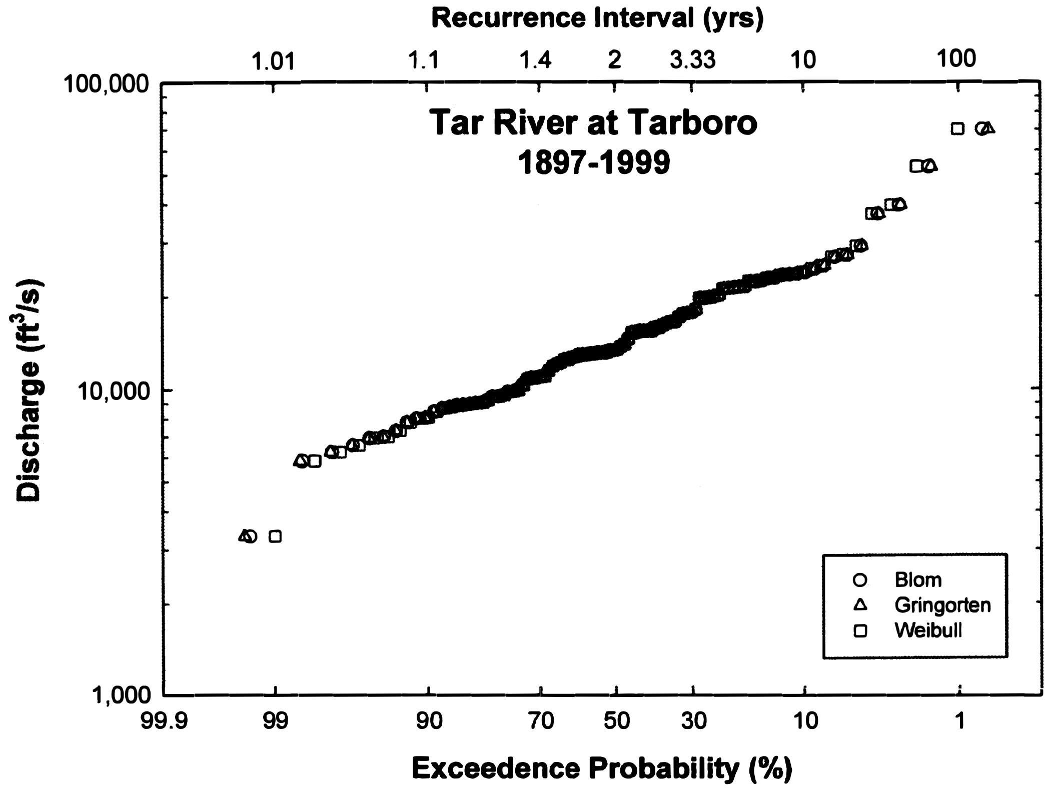

Figure 1. The flood series for the Tar River at Tarboro plotted using different plotting position formulas.

skewed, that is, small floods occur more often than large floods. Nevertheless, there is little theoretical rationale to guide the selection of the most appropriate, negatively skewed, probability distribution. In fact, different distributions are used by the United States (log-Pearson Type III) and the United Kingdom (log-Gumbel or Extreme Value Type II), and the selection of the log-Pearson Type III distribution by the U.S. Water Resources Council (1981) was met with considerable opposition (Benson 1968, 1969; Kisiel 1969). The method of moments or maximum likelihood estimates are used to fit the probability distribution to the empirical data. This produces a flood frequency curve that is used to estimate the discharge of any *n-year* event. Although the method of maximum likelihood is considered superior to the methods of moments for fitting the data to a particular probability distribution, the former is more computationally complex.

In order to assure that the theoretical (fitted) probability distribution fits the flood series, the empirical data are plotted on specially designed probability paper that linearizes the flood frequency curve for a particular distribution. Quantitative measures such as the chi-squared statistic and the Kolmogorov-Smirnov test may be used to assess which distribution best fits the data, however, a graphical comparison is often equally useful. Probability paper cannot be constructed for the log-Pearson Type III distribution because a different probability scale would be needed for each value of the coefficient of skewness, therefore, lognormal probability paper is usually used. The data are plotted by calculating plotting positions, which assign a

Table 2. Gaging record and drainage area for stations in the Tar river basin.

Station	Gage ID	Period of Record	Record Length (yrs)	Drainage Area (mi^2)
Little Fishing Creek near White Oak	02082950	1960-1999	40	177
Tar River at US 401 at Louisburg	02081747	1964-1999	36	427
Tar River at NC 97 at Rocky Mount	02082585	1977-1999	23	925
Tar River at Tarboro	02083500	1897-1900, 1906-1999	98	2,183

probability value to each flood discharge to be plotted. The most common plotting position used for floods is the Weibull formula:

$$(3) \qquad T = \frac{n+1}{m}$$

where n is the number of floods, m is the rank of each flood (ranked from largest to smallest where the largest is $m = 1$), and p is the inverse of equation 3. Although the Weibull formula remains widely used in hydrology (U.S. Water Resources Council 1981), it has been criticized for under-estimating the recurrence interval of large magnitude floods (Cunnane 1978). A variety of alternative plotting position formulas may be used, many of which have the general form:

$$(4) \qquad T = \frac{n+1-2a}{m-a}$$

where the parameter $a = 0$ for Weibull's formula, $a = 0.375$ for Blom's, and $a = 0.44$ for Gringorten's (Chow et al. 1988). Figure 1 shows that the plotting positions calculated by equation 4 can be significantly different for the larger magnitude floods. In order to obtain unbiased plotting positions, Cunnane (1978) found that Blom's plotting position should be used for the normal (or lognormal) distribution and Gringorten's for the Gumbel (Extreme Value Type I). The value of a for the log-Pearson Type III distribution depends on the value of the coefficient of skewness, with $a > 0.375$ for positively skewed data, and $a < 0.375$ for negatively skewed data (Chow et al. 1988).

Data

Historical records of annual floods at four stations in the Tar River basin were obtained from the USGS (Table 2; Figure 2). The USGS has published preliminary discharge estimates for the 1999 flood at each of the stations, which represent a range of drainage

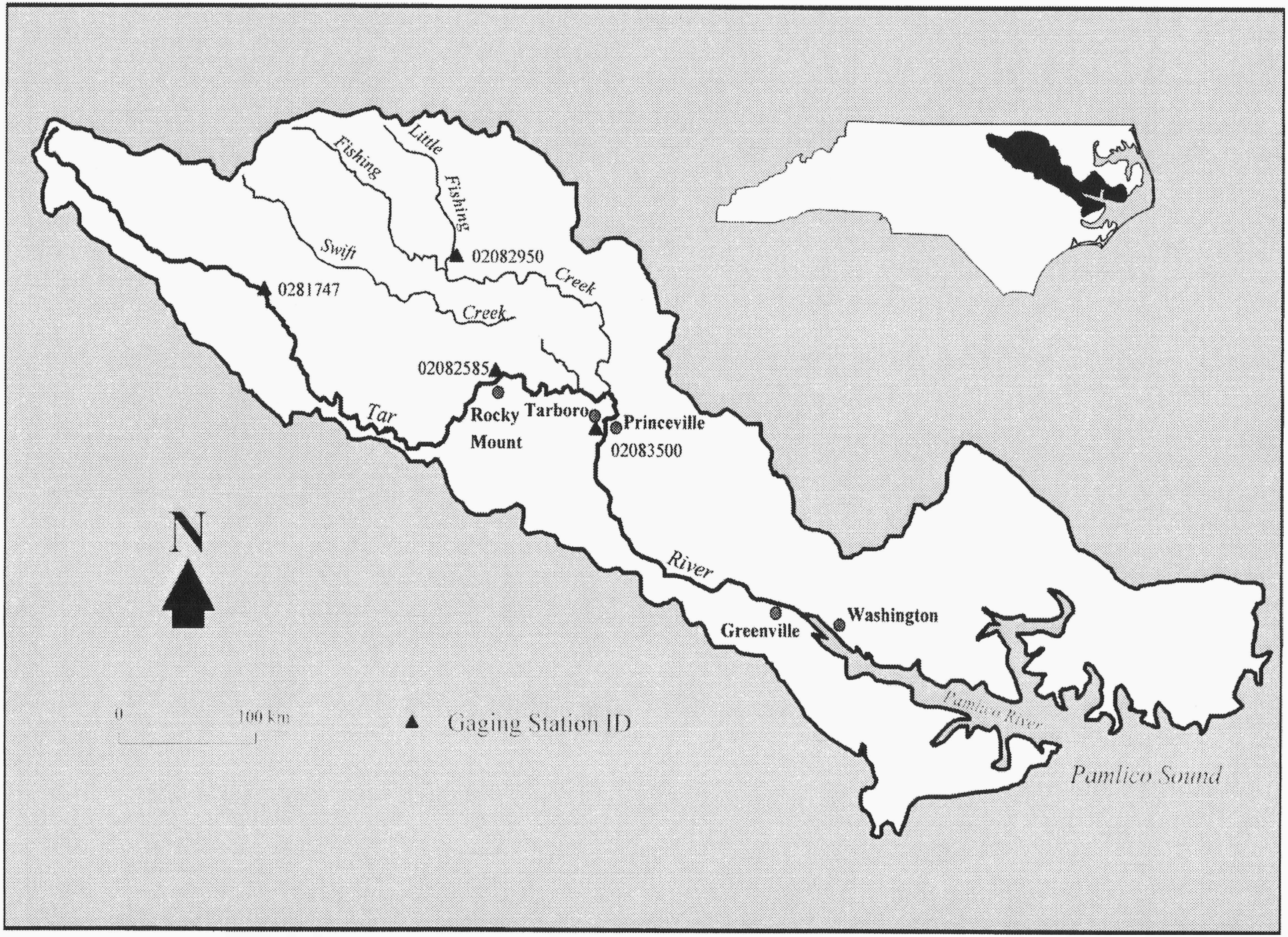

Figure 2. Map of the Tar River basin, North Carolina, showing the locations of the four USGS gaging stations.

areas and record lengths. Flood frequency curves were produced by fitting the log-Pearson Type III distribution to the data using the method of moments. Blom's formula ($a = 0.375$ in equation 4) was used to plot the data because it is closer to being unbiased than Weibull's. Confidence limits were calculated following the procedures outlined in Chow et al. (1988).

Results

FFA assumes that the period of record sampled is representative of the distribution of annual floods that would occur over a very long period of time. This is, of course, unlikely because record lengths are usually less than 50 years and seldom as long as 100 years. Confidence limits around the flood frequency curve define the degree of statistical uncertainty associated with recurrence interval estimates derived from the curve. Figure 3 shows 5% and 95% confidence limits for the flood frequency curves at the four stations in the Tar River basin. That is, there is a 90% chance that the curve at a given recurrence interval should be located between these confidence limits.

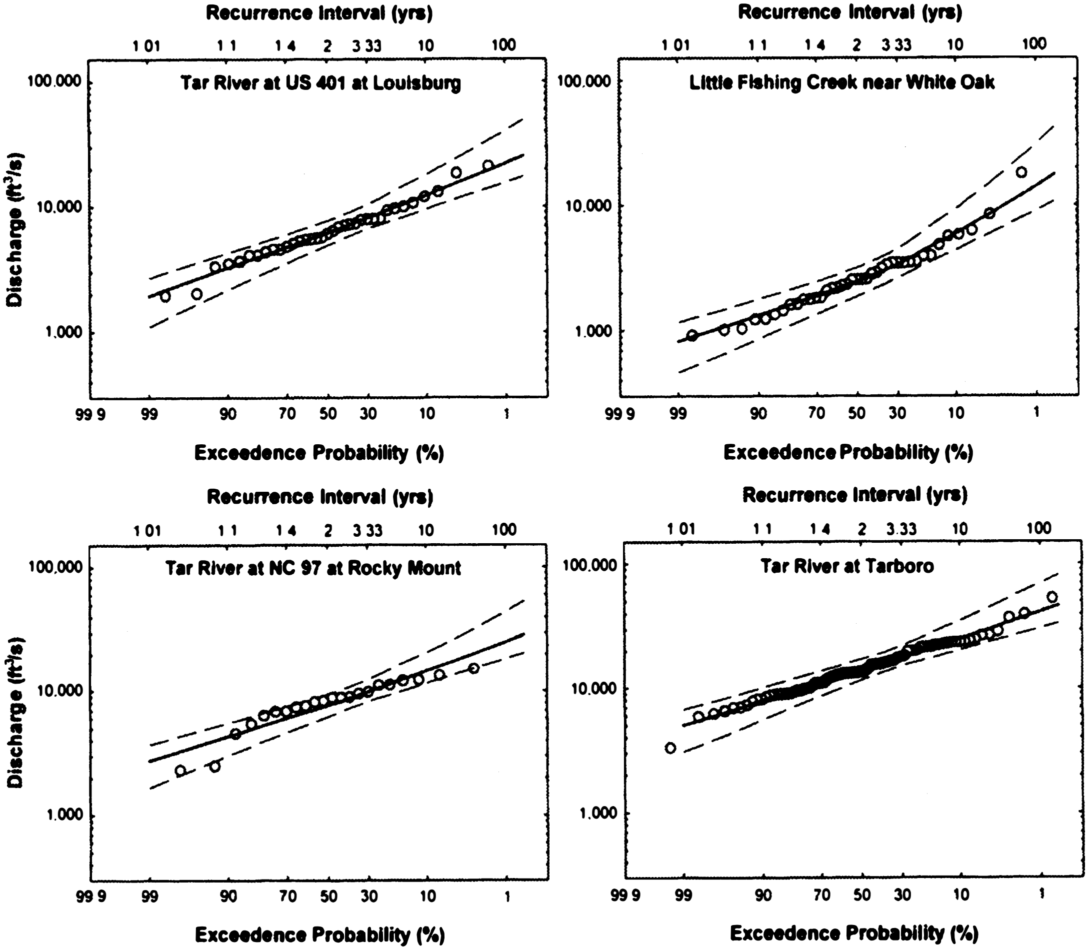

Figure 3. Flood frequency curves (solid lines) and 90% confidence limits (dashed lines) on lognormal probability paper for four gaging stations in the Tar River basin. Data plotted using Blom's formula.

The range of discharge between the confidence limits is indicative of the relative inaccuracy of flood frequency curves. For example, the 100-year flood discharge for the Tar River at Tarboro (which has an unusually long record at 97 years) is 40,868 ft^3/s. The confidence limits indicate, however, that there is a 90% chance that the 100-year flood lies between 30,362 ft^3/s and 67,960 ft^3/s, a range of 37,598 ft^3/s. Increasing the confidence interval to 95% or 99% would increase this range substantially. Furthermore, the logarithmic axes used in Figure 3 provide a misleading visual display of changes in the confidence interval as the recurrence interval increases. When plotted on arithmetic probability paper (Figure 4), the confidence interval increases exponentially for the larger recurrence intervals. Thus, there is a large degree of statistical "uncertainty" associated with recurrence interval estimates of large flood discharges, even with large data sets like the Tar River at Tarboro. The technique produces good results for the small floods, but not for the large events with which we are most interested. Because obtaining improved estimates of large floods would require a much longer period of record, this problem is largely unavoidable.

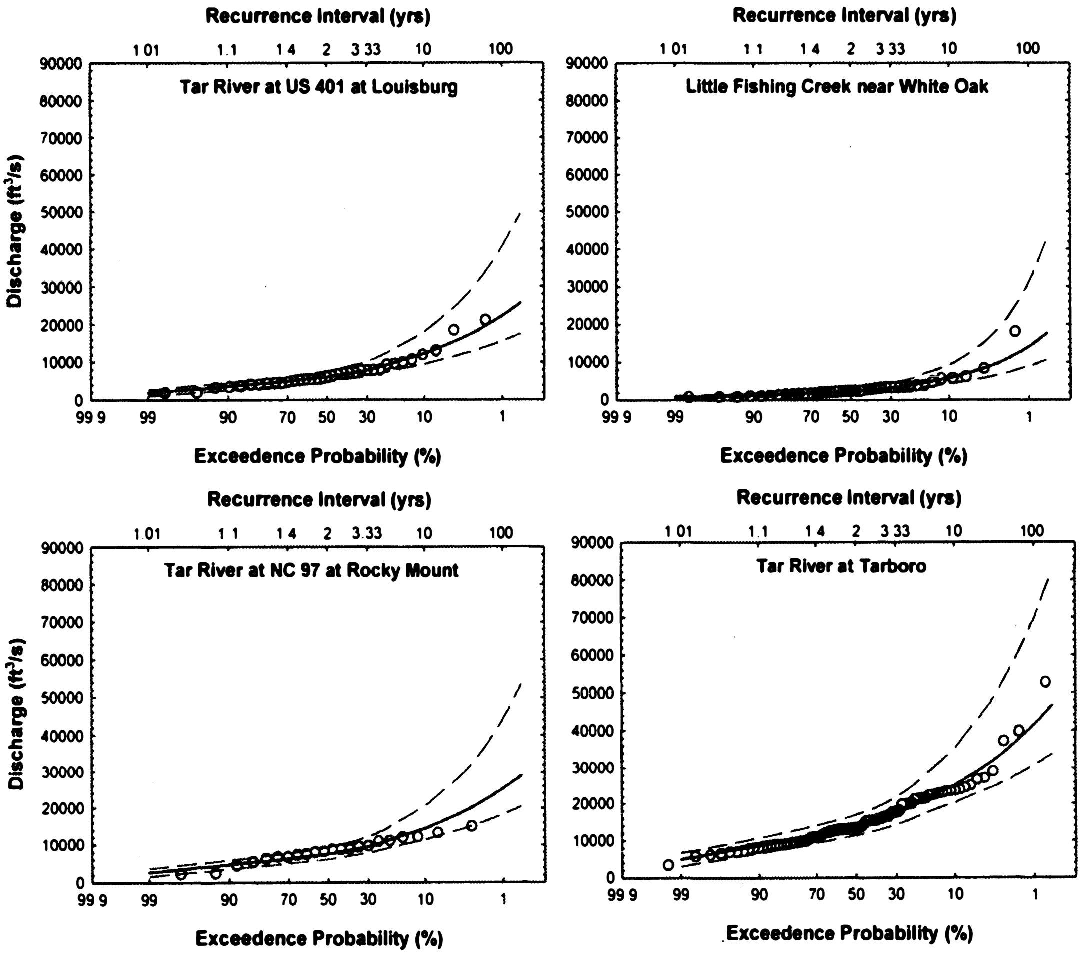

Figure 4. Flood frequency curves (solid lines) and 90% confidence limits (dashed lines) on an arithmetic probability plot for four gaging stations in the Tar River basin. Data plotted using Blom's formula.

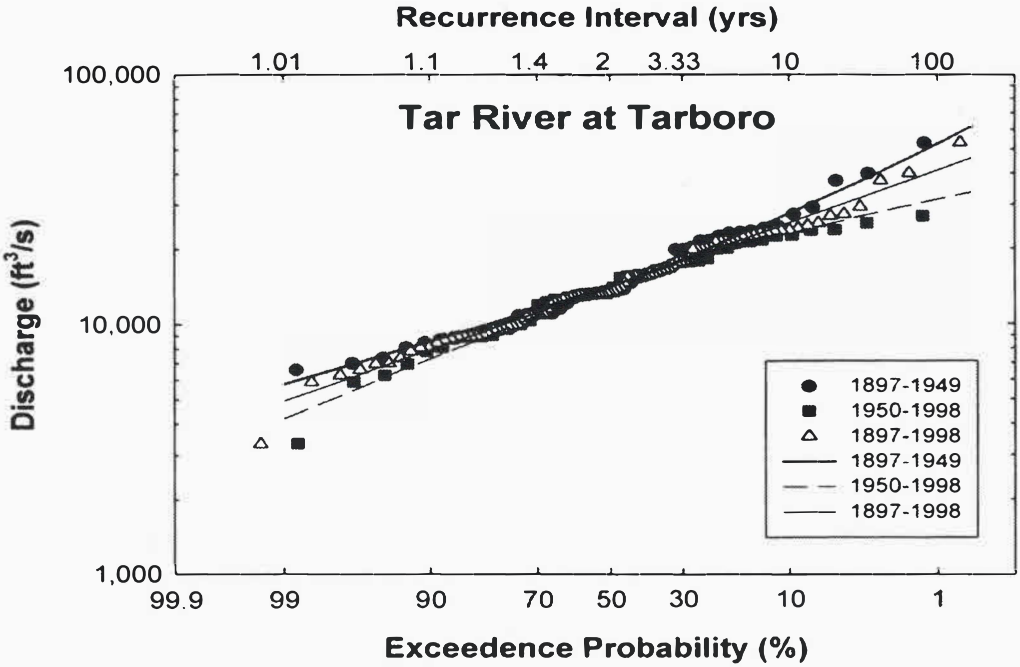

Figure 5. Flood frequency curves for the Tar River at Tarboro.

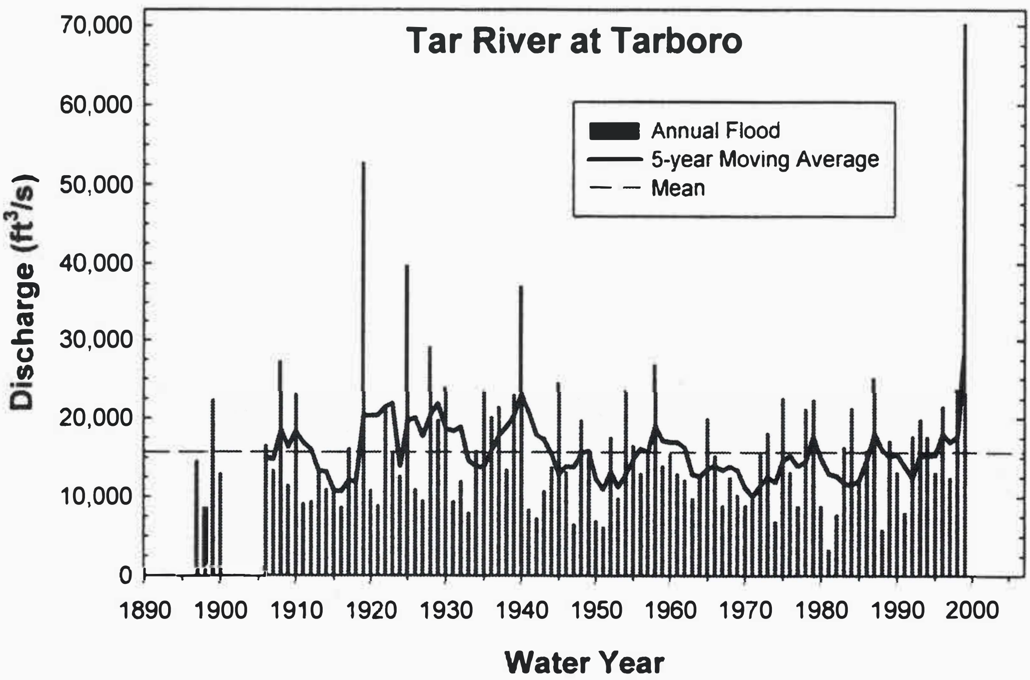

Figure 6. Historical variations in annual flood magnitudes for the Tar River at Tarboro.

Table 3. Flood frequency estimates for the Tar River at Tarboro, 1897-1949, 1950-1998.

	1897-1949			1950-1998			Change		
Recurrence Interval (yrs)	Q (ft^3/s)	5% Conf. Limit (ft^3/s)	95% Conf. Limit (ft^3/s)	Q (ft^3/s)	5% Conf. Limit (ft^3/s)	95% Conf. Limit (ft^3/s)	Q (%)	5% Conf. Limit (%)	95% Conf. Limit (%)
1.01	5,764	3,629	7,607	4,190	2,431	5,729	-27	-33	-25
2	14,123	11,463	17,271	13,989	11,641	16,934	-1	-2	-2
5	21,537	17,587	28,461	19,545	16,215	25,359	-9	-8	-11
10	27,466	21,847	39,103	22,829	18,633	31,050	-17	-15	-21
25	36,211	27,584	56,865	26,551	21,214	38,008	-27	-23	-33
50	43,724	32,190	73,754	29,056	22,882	42,962	-34	-29	-42
100	52,144	37,106	94,231	31,360	24,377	47,695	-40	-34	-49
200	61,598	42,390	118,970	33,479	25,723	52,186	-46	-39	-56

The annual flood series for the Tar River at Tarboro can be used to illustrate another element of statistical uncertainty associated with FFA. If this series was split in half, the 48 year and 49 year sub-series would still be longer than those at most USGS gaging stations. If these sub-series were representative of the long-term distribution of flood discharges, they should plot similarly in Figure 5. Clearly, this is not the case. Because many more large floods were experienced during the first half of the century (Figure 6), the flood frequency curve calculated for the 1897-1949 period would produce a 100-year flood discharge of 52,144 ft^3/s, while the 1950-1998 series would give a 100-year flood discharge of only 31,360 ft^3/s (Table 3). Furthermore, the 100-year flood for the 1897-1949 period is not even within the 90% confidence interval for 1950-1998 series (24,377-47,695 ft^3/s). Thus, if the collection of gaging station data had been initiated in 1950, rather than 1897, the 100-year discharge would be 9,508 ft^3/s lower than that obtained using the full record (1897-1998), and 20,784 ft^3/s lower than that for the 1897-1949 period.

Although FFA assumes that the flood series is stationary (i.e., the mean and variance are constant through time), periods of high rainfall and drought appear to cluster. Such non-stationarity in the historical record might be explained by climatic trends or cycles. The allocation of the flow of the Colorado River, which began in 1922 with the partitioning of water rights between the upper basin states and the lower basin states, provides a useful example of decade-scale variability of stream flows and adverse effects on water management decisions. The apportionment of water rights on the Colorado River was, unfortunately, based on records from an unusually wet period. Average stream flow during 1896-1930 was much higher (17 million acre-feet) than that from 1931-1965 (13 million acre-feet). Nevertheless, flows during this anomalous period played an important role in the over-appropriation of the river's water between competing states (Graf 1985). Non-stationarity in the flood series may also be generated by a variety of human activities such as urbanization, deforestation, agriculture, channelization, levees, damming by road crossings, and human-induced global warming. These all suggest that historical flood records might not be a good guide to future flood risks.

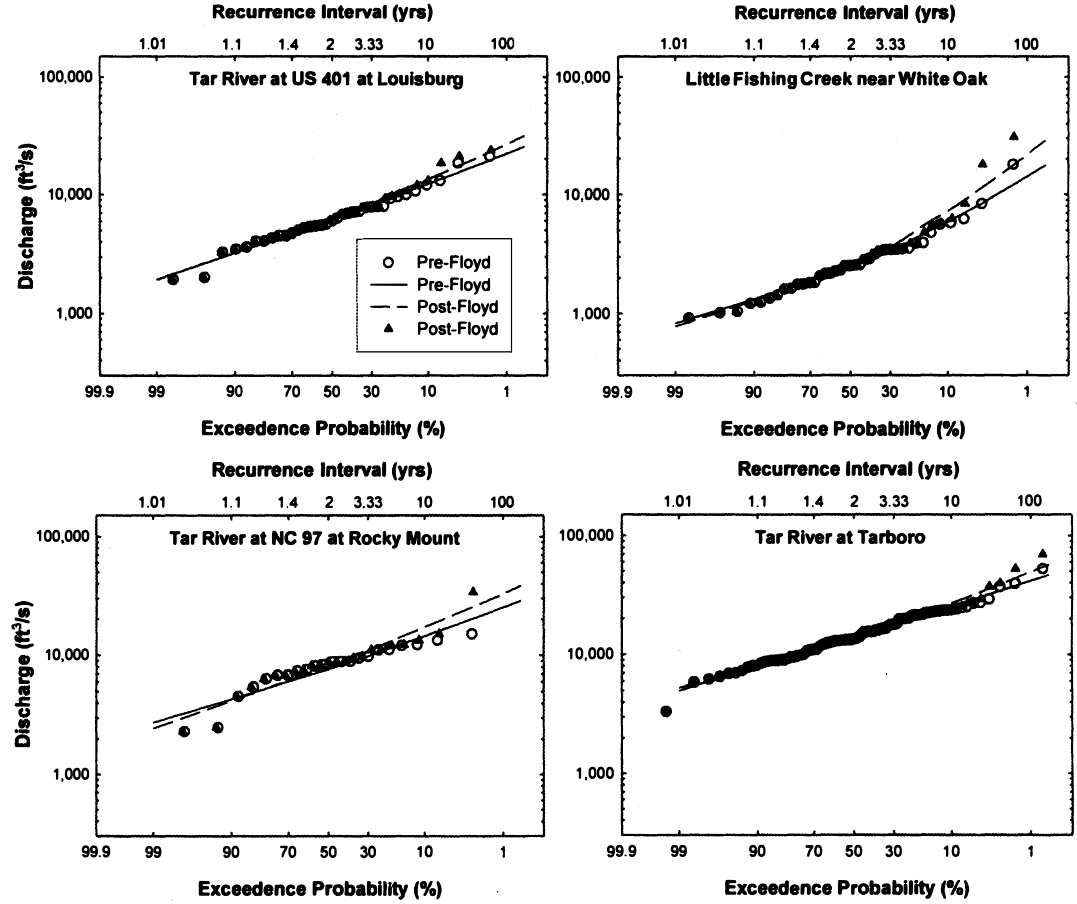

Figure 7. Effect of adding the 1999 flood on flood frequency estimates.

A fundamental problem with FFA is that flood records are too short to estimate low frequency events accurately. Because flood distributions are negatively skewed, few observations exist for the large events that are used to fit the high end of the flood frequency curve. This may be illustrated by adding the 1999 flood to the flood series at the four gaging stations in the Tar River basin (Figure 7). The magnitude of the 100-year flood increases at each station, from a minimum of 18% to a maximum of 54% (Table 4). Although the 1999 flood would be treated as an outlier by the USGS (U.S. Water Resources Council 1981) at three of the four stations (exception: Tar River at US 401 at Louisburg), the fact remains that a single large flood, even one that passes the outlier test, would increase recurrence interval estimates substantially. This effect is even more pronounced where the record length is short.

Another fundamental assumption of FFA is that the flood series is homogeneous, that is, the underlying population of floods is generated by only one type of event. Despite this, most hydrologists recognize that flood series consist of mixed populations, and thus, violate the homogeneity assumption (Hirschboeck 1988). For example, Diehl and Potter (1987) and Knox (1988) have shown that failing to separate the flood series into seasonal sub-populations (i.e., summer thunderstorm floods and spring snowmelt floods) can give unrealistic estimates of the magnitude and frequency of floods. In the Tar

Table 4. Pre-and post-Floyd 100-year discharge estimates.

Station	Pre-Floyd			Post-Floyd			Change		
	Q (ft^3/s)	5% Conf. Limit (ft^3/s)	95% Conf. Limit (ft^3/s)	Q (ft^3/s)	5% Conf. Limit (ft^3/s)	95% Conf. Limit (ft^3/s)	Q (%)	5% Conf. Limit (%)	95% Conf. Limit (%)
Little Fishing Creek near White Oak	14,106	9,000	30,903	21,767	12,641	56,399	54	41	83
Tar River at US 401 at Louisburg	22,198	15,630	40,571	26,753	18,161	52,247	21	16	29
Tar River at NC 97 at Rocky Mount	25,291	18,309	44,124	32,980	22,593	63,278	30	23	43
Tar River at Tarboro	40,868	30,362	67,960	48,302	34,679	85,638	18	14	26

River basin, and elsewhere in the southeastern U.S., mixed distributions can be a problem because floods are generated by a variety of meteorological mechanisms (Lecce 2000a, 2000b). For example, hurricanes are often responsible for generating the largest floods on record in North Carolina (Zembrzuski et al. 1987). Although most would agree that separating flood series into homogeneous sub-populations would improve flood frequency estimates, this is rarely done in practice (Knox 1988).

Conclusion

A review of probabilistic estimates of flood frequency showed that although large floods are unlikely in any single year, the odds are considerably higher that a large flood will occur over an extended period of time. An examination of annual floods at four stations in the Tar River basin, North Carolina, illustrates the difficulties inherent in estimating the recurrence intervals of large floods using traditional flood frequency analysis. Flood frequency estimates are sensitive to large floods, particularly where flood records are short, and 90% confidence limits suggest that the 100-year flood discharge may be under or overestimated by as much as 1.5-2 times. The gaged period of record may also not be representative of the long-term distribution of flood discharges if cyclic variations in climate are significant.

Because flood distributions are negatively skewed, because we are interested in that part of the distribution with which we are most uncertain (i.e., high magnitude, low frequency events), and because flood records are inevitably too short to effectively deal with the infrequent events, flood frequency estimates should be regarded as best guesses based on historical data. Although this paper has focused on the degree of uncertainty associated with estimates of the 100-year flood, attempting to estimate recurrence intervals for larger magnitude events like the flood of 1999 is fraught with even more uncertainty. Perhaps the most meaningful description of the flood of 1999 is the catch-phrase "the flood of the century". With this, there can be no debate.

Acknowlegment

This paper is contribution number 51 to the Tobacco Road Research Team.

References

Benson, M.A. (1968). Uniform flood-frequency estimating methods for federal agencies. *Water Resources Research* 4:891-908.

_______ (1969). Reply to comments on uniform flood-frequency estimating methods for federal agencies. *Water Resources Research* 5:911.

Chow, V.T., Maidment, D.R., and Mays, L.W. (1988). *Applied hydrology.* New York: McGraw-Hill, 572p.

Cunnane, C. (1978). Unbiased plotting positions – a review. *Journal of Hydrology* 37:205-222.

Diehl, T., and **Potter, K.W.** (1987). Mixed flood distributions in Wisconsin,. In *Hydrologic frequency modeling*, ed. V.P. Singh, pp. 213-226. Boston: D. Reidel Publishing.

Graf, W.L. (1985). *The Colorado River: Instability and basin management.* Washington, DC: Association of American Geographers, 86 p.

Gumbel, E.J. (1941). The return period of flood flows. *Annals of Mathematical Statistics* 12:163-190.

_______ (1958). *Statistics of extremes.* Columbia University Press: New York, 375 p.

Hirschboeck, K.K. (1988). Flood hydroclimatology. In *Flood geomorphology*, eds. V.R. Baker, R.G. Kochel, and P.C. Patton, pp. 27-49. New York: John Wiley and Sons.

Kisiel, C.C. (1969). Comments on 'Uniform flood-frequency estimating methods for federal agencies' by Manual A. Benson. *Water Resources Research* 5:910.

Knox, J.C. (1988). Climatic influence on Upper Mississippi Valley floods. In *Flood geomorphology*, eds. V.R. Baker, R.G. Kochel, and P.C. Patton, pp. 279-300. New York: John Wiley and Sons.

Lecce, S.A. (2000a). Spatial variations in the timing of annual floods in the southeastern United States. *Journal of Hydrology*, 235:151-169.

_______ (2000b). Seasonality of flooding in North Carolina. *Southeastern Geographer*, in press.

Royal, J. (2000). Floyd's flood one in 400 or 450 years. *The Daily Reflector.*

Smith, K., and **Ward, R.** (1998). *Floods: Physical processes and human impacts.* John Wiley and Sons: Chichester, 394 p.

U.S. Water Resources Council. (1981). Guidelines for determining flood flow frequency. *Bulletin 17B*, Washington, D.C.

Zembrzuski, T.J., **Hill, C.L.**, **Weaver, J.C.**, **Coble, R.W.**, and **Gunter, H.C.** (1987). North Carolina: floods and droughts. In *National water summary 1988-89, Floods and droughts, U. S. Geological Survey Water-Supply Paper 2375*, eds. R.W. Paulson, E.B. Chase, and D.W. Moody, p. 425-434.