

Local Polynomial Technique for Flood Frequency Analysis

Somkiat Apipattanavis^{1,2}, Balaji Rajagopalan^{1,3} and Upmanu Lall⁴

¹Department of Civil, Environmental and Architectural Engineering, University of Colorado at Boulder, Boulder, USA

²Royal Irrigation Department, Bangkok, Thailand

³Co-operative Institute for Research in Environmental Sciences, University of Colorado, Boulder, USA

⁴Department of Earth and Env. Engg., Columbia University, New York, NY

Abstract

Nonparametric estimators of at-site flood frequency using annual maximum flow data present an alternative to methods that a priori assume a specific probability distribution function. They approximate a wide class of distribution functions. Past work in this direction using kernel density and quantile estimators is extended here using a higher order approximation technique, locally weighted polynomial regression, for estimating the quantile function. An empirical investigation of the performance of this method relative to selected alternatives and for selected target distributions is presented here.

Introduction

Flood frequency analysis entails relating the magnitude of annual maximum flood events to their frequency of occurrence at a gauged site. The typical interest is estimating extreme flood quantiles i.e. - 100-year or 500-year flood, from a small number of observations (~50 to 100 years) for design of hydraulic structures such as dams, culverts and bridges. Traditional parametric methods for this problem assume that the annual maximum floods are independent and identically distributed and drawn from a population with a known Probability Density Function (PDF). An appropriate PDF is selected from a candidate set or mandated by a regulatory agency for at site applications. Typical distributions that are prescribed by agencies such as USBR, USGS and widely used in practice are Log-Pearson Type III, Log-Normal and Extreme Value distributions (see Kite, 1977, Chow et al., 1988). There are statistical tests to discriminate between choices of distributions including L-moments methods (see Kite, 1977; Vogel 1986; Hosking, 1990; Vogel and McMartin, 1991). However, it is often difficult to discriminate between candidate models for a given data set, and the best fit criteria emphasize the bulk of the distribution rather than its tails. Consequently, there is considerable uncertainty as to the best underlying model for the estimation of the upper flood quantiles.

Nonparametric methods, on the other hand, do not assume a distributional form to the data. Rather, the flood magnitude at any quantile is estimated by locally smoothing the empirical quantile function of the data or estimating the PDF using a kernel based estimator. Because the method is “local”, in that estimates of the function at a point are based on data points in its neighborhood – this provides the ability to better capture an arbitrary features exhibited by the data and furthermore, easily portable across sites. For the estimation of tail quantiles, an extrapolation rather

than interpolation of the empirical quantile function is needed. The “local” estimation procedure inherent in nonparametric flood frequency analysis translates into a model for tail probability estimation. Traditional tail probability estimators consider specific models of tail behavior whose parameters are to be estimated. Typically, a threshold beyond which the tail probability model should be applied also needs to be inferred from the data. Moon et al (1993) demonstrated that kernel based methods often performed better in practice than some of the tail probability models that are commonly used. In this paper, we present a higher order nonparametric estimation scheme that improves further on the kernel quantile estimations presented by Moon et al (1994).

Nonparametric flood frequency estimators were developed and studied by Schuster and Yakowitz (1985), Adamowski (1985, 1989), Adamowski and Feluch (1990), Bardley (1988, 1989) and more recently Lall et al. (1993), Moon et al. (1993) and Moon and Lall (1994). Lall et al. (1993) developed a kernel based quantile estimator, where in, a kernel density estimator is used to estimate the probability distribution function and consequently, the quantiles of interest. They also showed that parametric estimates based on the cumulative distribution function are more appropriate than those based on density estimates in the flood frequency context. Kernel density based estimators while easy to implement, suffer from (1) loss of efficiency of estimation with respect to the true distribution, (2) an uncertain and likely negligible ability to extrapolate beyond the data (Lall et al., 1993) and (3) over smooth the distribution function. Adamowski (1989) suggested a variable bandwidth kernel density estimator that addresses the extrapolation problem. Later Moon and Lall (1994) developed a nonparametric kernel based regression estimator for quantiles. Here, the empirical quantile function is smoothed using a kernel regression

estimator. They find that both, the density and regression based estimators are competitive to other estimators. However, both these nonparametric quantile estimators suffer from boundary problems, i.e., the tail quantiles are biased (Lall et al., 1993; Moon and Lall, 1994). Here, we present a local polynomial (Loader, 1999) based estimator that improves upon the kernel regression estimator.

The local polynomial estimator is first described. We then compare the performance of this estimator with traditional parametric estimators on a suite of synthetic data set, followed by their comparison on two streamflow data sets.

Local Polynomial Estimator

Given an n -year historical record of annual maximum floods, we can define the empirical quantile function through the following set of ordered pairs: (X_i, Y_i) , $i = 1, 2, \dots, n$ where $X_i = (i - 0.25)/(n + 0.50)$, $Y_i =$ ranked annual maximum flood data. The X_i is the so-called plotting positions, and one can use any other formula of interest for the purpose. Here Adamowski's (1981) formula is used.

Then, we consider a general model for the quantile function as:

$$Y_i = \mu(X_i) + \varepsilon_i \quad (1)$$

Where $\mu(.)$ is a nonlinear function, ε_i are assumed to be identically distributed errors with mean 0 and finite variance, and $X_i \in [0, 1]$.

In this context if we consider the estimation of the T year flood, then we are interested in an estimate $\mu(X_T)$ such that $X_T = 1 - 1/T$. The specific proposal advanced here is that $\mu(X_T)$ be estimated using locally weighted polynomial regression, where we assume that $\mu(X_T)$ is a general function that is continuous and has $(p-1)$ derivatives. Hence, it is reasonable to approximate $\mu(X_T)$ using a local

polynomial of order p , following Taylor series arguments. “Local”, here refers to an approximation in the neighborhood of X_T . The size of the neighborhood depends on the smoothness of the target regression function and on the nature of the residual process, ε_i .

For details as to the specific local polynomial estimation method (LOCFIT) used here see Loader (1999). The estimation algorithm is summarized below:

1. For any point of estimate, X_T , nearest neighbors (i.e. nearest data points), $k(= \alpha n)$, are identified, where α varies from 0 to 1 (when $\alpha = 1$ then all the data points are neighbors to X_T). The bandwidth $h(X_T)$ of this window of k neighbors around X_T is the distance to the k^{th} neighbor. For tail quantiles, this translates into the number of upper order statistics that are used to fit a polynomial tail quantile model.
2. Each of the k data pairs used is then weighted according to the distance to T via a weight function (e.g. Bisquare, Tricubic etc.). The Bisquare weight is given as $W(u_i) = (15/16) * (1 - u_i^2)^2$, where $u = (X_i - X_T) / h(X_T)$, and $|u| \leq 1$.
3. Within the smoothing window (i.e. with the k neighbors), $\mu(X)$ is approximated by a polynomial order p . For example, a local quadratic model would be

$$\mu(X) = a_0 + a_1(X) + a_2(X)^2 \quad (2)$$

The coefficients of the polynomial a_0 , a_1 and a_2 are obtained by minimizing the weighted least squares function,

$$\sum_{i=1}^k W_i(X_T)(Y_i - \mu(X_i))^2 \quad (3)$$

These steps are repeated for each estimation point.

The key parameters identify are the optimal number of neighbors k and the order of polynomial p . These are obtained via minimization of a Generalized Cross Validation (GCV) function described below. If $h(X)$ is too small, insufficient data fall within the smoothing window, the estimated quantile value will have a very high variance. On the other hand, if $h(X)$ is too large, the quantile estimate may have a large bias. Therefore, the bandwidth must be chosen to compromise this bias-variance trade-off. Similar to the bandwidth, the degree of the local polynomial p , affects the bias-variance trade-off.

It often suffices to choose a low order polynomial and concentrate on choosing the bandwidth to obtain a satisfactory fit (Loader, 1999). Typically, in parametric regression, mean squared error is used to assess the performance of the fit. However, this is a poor indicator of future performance of the model (i.e. predictive error). Craven and Wahba (1979) developed a measure called the GCV, similar to Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC), that approximates predictive risk.

$$GCV(\alpha, p) = n \frac{\sum_{i=1}^n (Y_i - \hat{\mu}(X_i))^2}{\left(1 - \sum_{i=1}^n h_{ii}\right)^2} \quad (4)$$

Where n is the sample size, $Y_i - \hat{\mu}(X_i)$ is the residual and h_{ii} are the diagonal terms of the hat matrix H . The hat matrix can be estimated using standard linear regression procedures. For fairly small datasets Loader (1999) suggests the use of the Cross Validation (CV) function:

$$CV(\alpha, p) = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{\mu}_{-i}(X_i))^2 \quad (5)$$

Where $\hat{\mu}_{-i}(X_i)$ denotes the leave- X_i -out estimate of $\hat{\mu}(X_i)$. That is, each X_i is removed from the dataset in turn, and the local regression estimate computed from the remaining $n-1$ data points.

Loader (1999, page 29) also developed an approximate confidence and prediction intervals for the estimates from the local polynomial, assuming the residuals to be normally distributed locally (within the neighborhood of k data points). If the prediction points are not normally distributed, the prediction intervals from this approach will not be correct, even asymptotically. Therefore, we used the bootstrap approach for obtaining the intervals – in this, a large number (1000) of bootstrap samples were generated and the quantile estimates obtained from the methodology applied to each of the sample, the 5th and 95th percentiles of the estimates from the bootstrap samples 90% intervals. Bootstrap approach is nonparametric in that, no distributional assumption of the data is required and is consistent with the nonparametric LOCFIT estimator proposed here.

Applications

We tested the LOCFIT quantile estimator on a suite of synthetic data sets and two streamflow data sets. We also compared with traditional parametric estimators.

Synthetic experiments

To simulate the “choice” of models that a practitioner may face, we considered a set of probability distribution models as “parents” for the at-site flood generation process, and similarly for the estimation of quantiles. The LOCFIT procedure is considered as an alternative for estimation across the suite of “parents”. The setting is of interest where a public regulatory agency may have mandated the use of a specific distributional model across all enterprises as a “best practice”. This has

been the case in the U.S., since the publication of Bulletin 17B (Interagency Advisory Committee on Water Data, 1982), U.S. Water Resources Council procedures were adopted. Our hypothesis is that the nonparametric procedure will be competitive against parametric alternatives, where a mix of parent populations may be appropriate across the country or region. We test this hypothesis by examining the success or lack thereof versus the proper specification and mis-specification of the parametric model. Consequently, we generated 500 samples of size 75 each from the following parent populations:

$$1) \text{ Log-Normal: } f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right], \quad x > 0 \text{ where } y = \log x,$$

$$\mu_y = \bar{y}, \quad \sigma_y = s_y$$

$$2) \text{ Log-Pearson type III: } f(x) = \frac{\lambda^\beta (x-\varepsilon)^{\beta-1} e^{-\lambda(x-\varepsilon)}}{x\Gamma(\beta)}, \quad \log x \geq \varepsilon \text{ where}$$

$$y = \log x, \quad \lambda = s_y/\sqrt{\beta}, \quad \beta = [2/C_s(y)]^2, \quad \varepsilon = \bar{y} - s_y\sqrt{\beta} \text{ (assuming } C_s(y) \text{ is positive)}$$

$$3) \text{ Extreme Value Type I: } f(x) = \frac{1}{\alpha} \exp\left[-\frac{x-u}{\alpha} - \exp\left(-\frac{x-u}{\alpha}\right)\right],$$

$$-\infty < x < \infty \text{ where } \alpha = \sqrt{6}s_x/\pi, \quad u = \bar{x} - 0.5772\alpha$$

$$4) \text{ Mixture of Normal: } f(x) = c_1N(\mu_1, \sigma_1) + c_2N(\mu_2, \sigma_2), \quad -\infty < x < \infty \text{ where } c_1 \text{ and } c_2 \text{ are weighted constant and } c_1 + c_2 = 1.$$

In each case we compare the performance of LOCFIT for selected tail quantiles relative to properly and improperly specified parametric models. The above distributions were chosen for the synthetic experiments because they are from the exponential family-which is widely considered in practice.

Real Data

We also applied the LOCFIT estimator to two streamflow data sets (i) Annual maximum flow on the Santa Cruz at Tucson, Arizona for the period 1915-2000 and (ii) 3-day annual maximum flow on the American river at Folsom dam, CA for the period 1905-2001. The LOCFIT estimate at these sites are compared to Log-Normal (LN), Log-Pearson type III (LPIII), Extreme Value I distribution (EVI). Confidence intervals for the traditional methods were also computed using the bootstrap method described earlier, for estimates at selected return periods.

Results

Synthetic Data

The LOCFIT estimator and the parametric estimators are applied to each synthetic data and we estimated the 10-, 50-, 100-, 250- and 500-year return period magnitudes. These estimates from the simulations are shown as boxplots along with the true values as a solid line. Box sizes provide the variance of the estimates.

It can be seen from Figures 1-4 that the LOCFIT estimator exhibits good performance for all the tested distributions. The variance of the estimates from LOCFIT increases (bigger boxes) as the return period increases – more so for return periods 250 and 500 years. This is to be expected from standard regression theory, as LOCFIT extrapolates beyond the range of the data at higher return periods and hence, has larger variance. It performs well especially with the mixture distribution (Figure 4) when compared with other estimators. LPIII estimator is also good for all the distributions with relatively smaller variance in comparison to LOCFIT. This is because, LPIII performs very well for exponential family of distributions, as is the

case with the distributions tested here. However, for the mixture distribution (Figure 4) the variances from LPIII are higher than those from LOCFIT. LN estimator works well for simulations from EVI and LN distributions (Figure 1 and 2), but performs badly in all other cases. The EVI estimator exhibits poor performance in all the distributions except of course, the simulations from EVI. The performance of the LOCFIT estimator on the mixture distribution is comparable to kernel-based nonparametric flood frequency estimators (Lall et al., 1993; Moon and Lall, 1994).

Real Data

The LOCFIT estimator and the traditional methods are applied to the Santa Cruz annual maximum flow data. As shown in Figure 5a, the LOCFIT estimates closely follow and smooth the empirical quantiles (shown as circles) in the tails. The parametric methods fit well at lower return periods, but for higher return periods they grossly underestimate relative to the observed quantiles, with EVI performing especially poorly. The cross-validated quantile estimates (Figure 5b) appear to be within the 90% confidence interval (obtained from the bootstrap approach). Residuals from the cross-validated estimates were found to be normally distributed with little autocorrelation (figures not shown) indicating the goodness of the LOCFIT model.

In the case of American River (Figure 6) as well, the LOCFIT estimator follows the empirical quantiles quite well, while parametric methods perform poorly in the tails. Also the Log-Normal estimator appears to overestimate the higher return period flows and EVI seems to underestimate. The confidence intervals from the LOCFIT estimator are generally tighter than those from the parametric models (Table 2). Similar results are found when applying these

methods to the data from the 1905-1945 (Figures 7a and b) and 1946-2001 (Figures 7c and d) sub periods (Tables 3 and 4). These sub periods were chosen because the Folsom dam was built in 1945. The Folsom dam was a case study of a recent National Research Council report (Potter et al., 1999) - where they suggest climate variability as being a reason for increased annual maximum flows in the latter sub period.

Summary

Locally weighted polynomial regression technique, a nonparametric approach, is applied to flood frequency estimation. The estimation is “local” and therefore, has the ability to capture any smooth distribution that generated the data. Unlike its parametric counterparts, no prior assumption of the underlying distribution is required, which makes it portable across sites. This also improves upon kernel-based nonparametric estimators developed in the past in that it is easy to implement. Good performance on a variety of synthetic and real data sets is observed. Thus, the proposed approach provides an attractive robust alternative for single-site flood frequency estimation that is completely data-driven and portable across sites. Multivariate extensions of this approach to forecasting regional flood quantiles conditioned on large-scale ocean-atmospheric information (e.g., Sankarasubramanian and Lall, 2003) are underway.

Acknowledgements

The first author thanks the Government of Thailand for his graduate student fellowship and also to Royal Irrigation Department, Thailand for their support. Partial support for the second and third authors from National Science Foundation through grant EAR 9973125 is thankfully acknowledged.

References

- Adamowski, K., 1981. Plotting formula for flood frequency. *Water Resources Research Bulletin*, 17(2), 197-202.
- Adamowski, K., 1985. Nonparametric kernels estimation of flood frequencies. *Water Resources Research* 21(11), 1585-1590.
- Adamowski, K., 1989. A Monte Carlo comparison of parametric and nonparametric estimation of flood frequencies. *Journal of Hydrology* 108, 295-308.
- Adamowski, K., and W. Feluch, 1990. Nonparametric flood-frequency analysis with historical information. *Journal of Hydraulic Engineering* 116(8), 1035-1047.
- Adamowski, K., and C. Labaliuk, 1987. Estimation of flood frequencies by a nonparametric density procedure, in *Hydrologic Frequency Modeling*, edited by V.P. Singh, p.p. 97-106, D. Reidel, Norwell, Mass.
- Bardley, W. E., 1988. Toward a general procedure for analysis of extreme random events in the Earth sciences. *Math. Geol.* 20(5), 513-528.
- Bardsley, W. E., 1989. Using historical data in nonparametric flood estimation. *Journal of Hydrology* 108, 249-255.
- Chow, V. T., D. R. Maidment and L. W. Mays, 1988. *Applied Hydrology*. 572 p.p., McGraw Hill, New York.
- Craven, P. and G. Wahba, 1979. Smoothing noisy data with spline functions. *Numerische Mathematik* 31, 377-403.
- Hosking, J. R. M., 1990. L-moment: Analysis and estimation of distributions using linear combinations of order statistics. *J. R. Stat. Soc., Ser. B* 52(2), 105-124.
- Interagency Advisory Committee on Water Data, 1982. Guidelines for determining flood flow frequency: Bulletin 17B of the hydrology committee, USGS. Office of Water Data Coordination. Reston Virginia.

- Kite, G. W., 1977. Frequency and Risk Analyses in Hydrology. Water Resources Publications. Fort Collins, Colo.
- Lall, U. and J. Niu, 1989. Variable bandwidth kernel density estimation (abstract). Eos Trans. AGU 70, 324.
- Lall, U., Y.-I. Moon and K. Bosworth, 1993. Kernel Flood Frequency Estimators: Bandwidth Selection and Kernel Choice. Water Resources Research 29(4), 1003-1015.
- Loader, C., 1999. Local Regression and Likelihood, 290 p.p., Springer, New York.
- Moon, Y.-I., U. Lall, and K. Bosworth, 1993. A comparison of tail probability estimators. Journal of Hydrology 151, 343-363.
- Moon, Y.-I., and U. Lall, 1994. Kernel function estimator for flood frequency analysis. Water Resources Research 30(11), 3095-3103.
- Potter, K. et al., 1999. Committee on American River Flood Frequencies, Improving American River Flood Frequency Analysis, Water Science and Technology Board, Commission on Geosciences, Environment, and Resources National Research Council, National Academy Press, 120 p.p., Washington, D.C.
- Sankarasubramanian, A., and U. Lall, 2003. Flood quantiles in a changing climate: Seasonal forecasts and causal relations. Water Resources Research 39(5), 1134.
- Schuster, E., and S. Yakowitz, 1985. Parametric/nonparametric mixture density estimation with application to flood-frequency analysis. Water Resources Research Bulletin 21(5), 797-803.
- Vogel, R. M., 1986. The probability plot correlation coefficient test for the normal, lognormal, and Gumbel distribution hypothesis. Water Resources Research 22(4), 587-590.

Vogel, R. M., and D. E. McMartin, 1991. Probability plot goodness-of-fit and skewness estimation procedures for the Pearson type 3 distribution. *Water Resources Research* 27(12), 3149-3158.

Webb R. H. and J. L. Betancourt, 1992. Climatic variability and flood frequency of the Santa Cruz River, Pima County, Arizona, U.S. Geological Survey, Water Supply Paper; 2379.

Table 1 Comparison of 90% confidence intervals for various estimators at 100- and 500-year flood for the Santa Cruz River (1915-2000)

Estimator	100-year Flood				500-year Flood			
	Estimated Flood (cms.)	Confidence Interval			Estimated Flood (cms.)	Confidence Interval		
		upper limit (cms.)	lower limit (cms.)	Interval width (cms.)		upper limit (cms.)	lower limit (cms.)	Interval width (cms.)
EVI	860	1,144	529	615	1,124	1,506	672	834
LN	896	1,199	657	542	1,374	1,936	949	987
LP III	960	1,508	573	935	1,540	3,007	746	2,261
LOCFIT	1,462	1,552	630	922	1,817	1,952	671	1,281

Table 2 Comparison of 90% confidence intervals for various estimators at 100- and 500-year flood for the American River (1905-2000)

Estimator	100-year Flood				500-year Flood			
	Estimated Flood (cms.)	Confidence Interval			Estimated Flood (cms.)	Confidence Interval		
		upper limit (cms.)	lower limit (cms.)	Interval width (cms.)		upper limit (cms.)	lower limit (cms.)	Interval width (cms.)
EVI	3,134	3,602	2,578	1,024	4,071	4,692	3,336	1,356
LN	4,801	6,090	3,677	2,413	8,099	10,640	5,943	4,697
LP III	4,470	5,700	3,342	2,358	7,203	10,132	4,846	5,286
LOCFIT	3,360	3,515	2,868	647	3,530	3,768	3,001	767

Table 3 Comparison of 90% confidence intervals for various estimators at 100- and 500-year flood for the American River (1905-1945)

Estimator	100-year Flood				500-year Flood			
	Estimated Flood (cms.)	Confidence Interval			Estimated Flood (cms.)	Confidence Interval		
		upper limit (cms.)	lower limit (cms.)	Interval width (cms.)		upper limit (cms.)	lower limit (cms.)	Interval width (cms.)
EVI	2,842	3,418	2,158	1,260	3,648	4,399	2,738	1,661
LN	3,659	4,819	2,619	2,200	5,537	7,599	3,750	3,849
LP III	3,726	5,024	2,543	2,481	5,706	8,667	3,402	5,265
LOCFIT	2,927	3,045	2,417	628	3,046	3,228	2,355	873

Table 4 Comparison of 90% confidence intervals for various estimators at 100- and 500-year flood for the American River (1946-2000)

Estimator	100-year Flood				500-year Flood			
	Estimated Flood (cms.)	Confidence Interval			Estimated Flood (cms.)	Confidence Interval		
		upper limit (cms.)	lower limit (cms.)	Interval width (cms.)		upper limit (cms.)	lower limit (cms.)	Interval width (cms.)
EVI	3,340	4,027	2,514	1,513	4,370	5,267	3,267	2,000
LN	5,383	7,527	3,563	3,964	9,644	14,120	5,993	8,127
LP III	5,645	7,784	3,697	4,087	10,434	16,582	5,872	10,710
LOCFIT	3,502	3,671	2,984	687	3,619	3,961	2,932	1,029

Figure Captions

Figure 1. Boxplots of estimates of 10, 50, 100, 250 and 500-year return period of data from EVI distribution. The estimates are from (a) True distribution (EVI), (b) LOCFIT (c) Log Normal, and (d) Log Pearson III estimators. The solid line in all the figure is the true value.

Figure 2. Same as Figure 1 but for data from Log-Normal distribution.

Figure 3. Same as Figure 1 but for data from Log Pearson III distribution.

Figure 4. Same as Figure 1 but for data from Mixture Normal distribution.

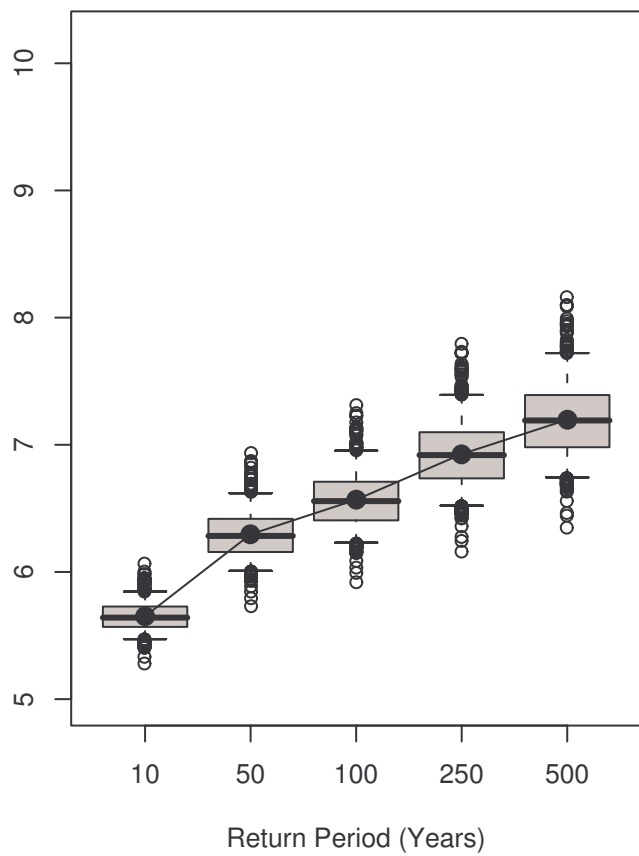
Figure 5. For the Santa Cruz River, AZ annual maximum flows, (a) Quantile estimates from LOCFIT and other parametric alternatives, (b) Cross-validated quantile estimates from LOCFIT along with the 90% confidence intervals.

Figure 6. Same as Figure 5 but for 3-day annual maximum flows at American River, CA.

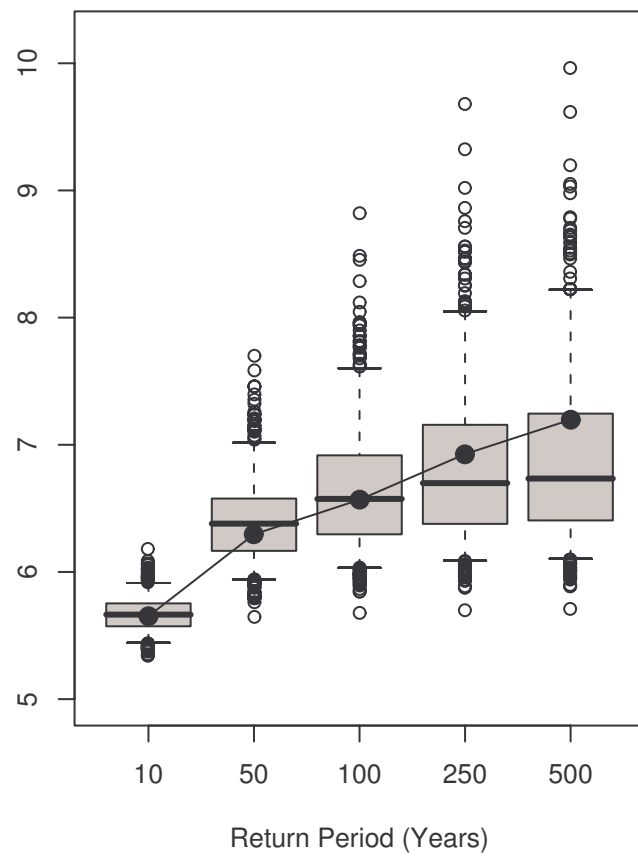
Figure 7. (a) and (b) same as Figure 5 (a) and (b), respectively, but for the 3-day annual maximum flows at American River, CA, for the 1905-1945 period. (c) and (d) are for the 1946-2001 period.

Figure-1

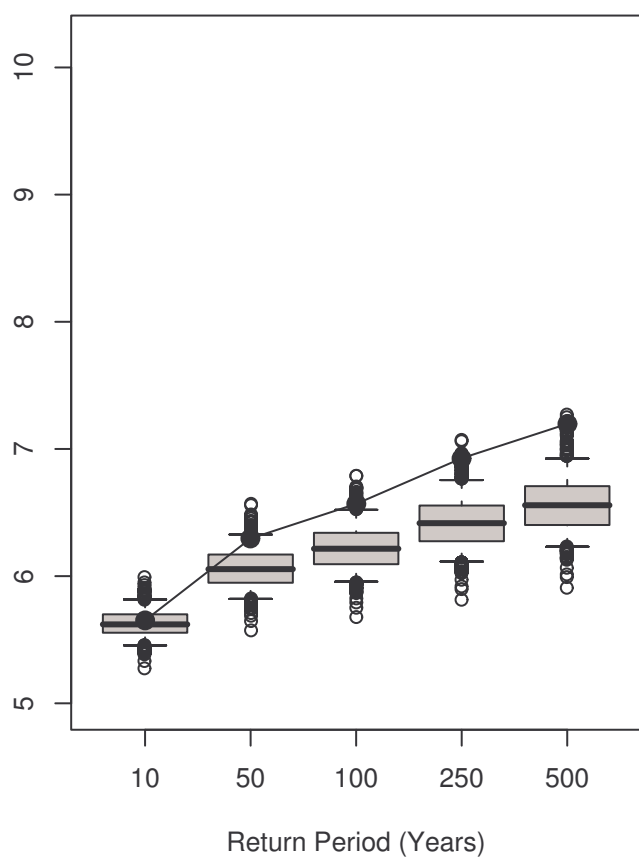
(a) True Distribution (EVI)



(b) Locfit



(c) Log Normal



(d) Log Pearson III

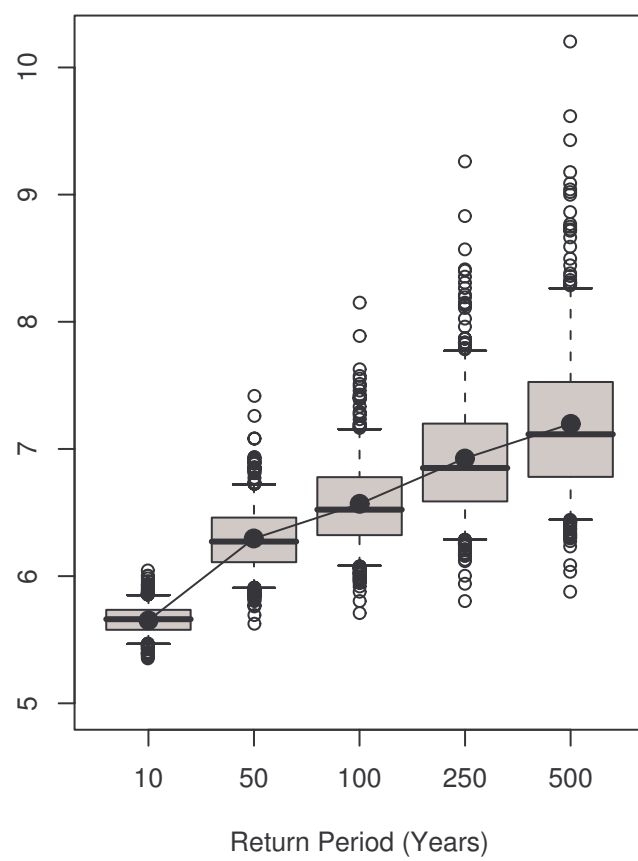
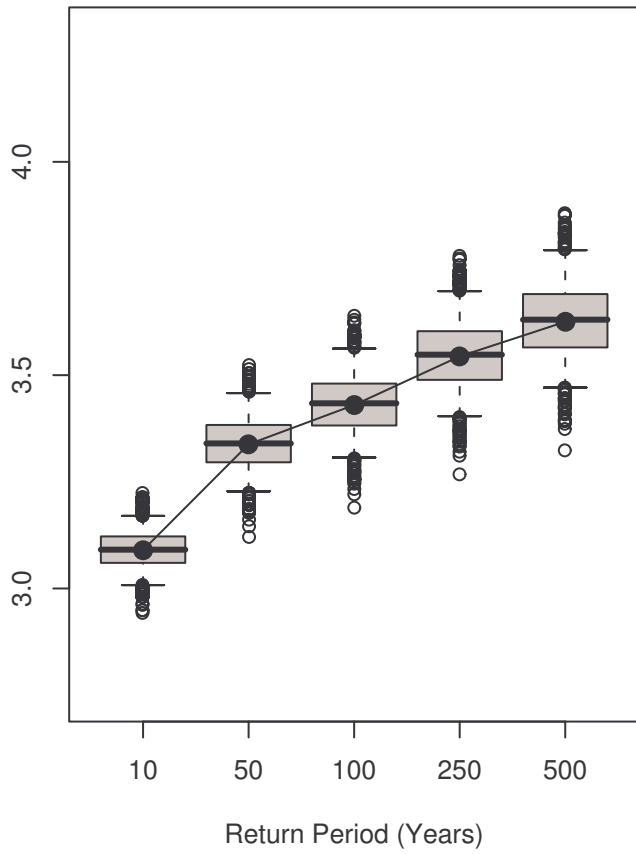
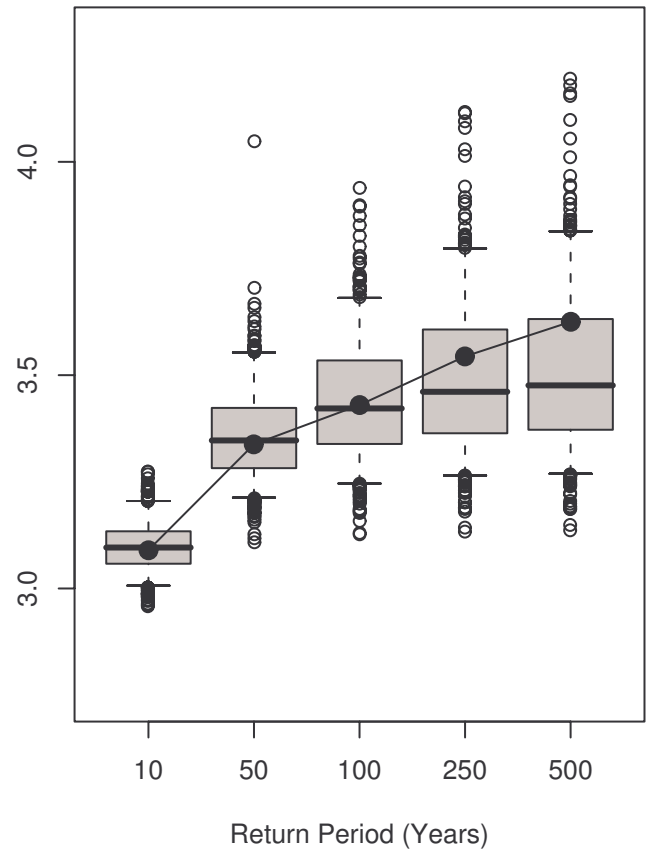


Figure-2

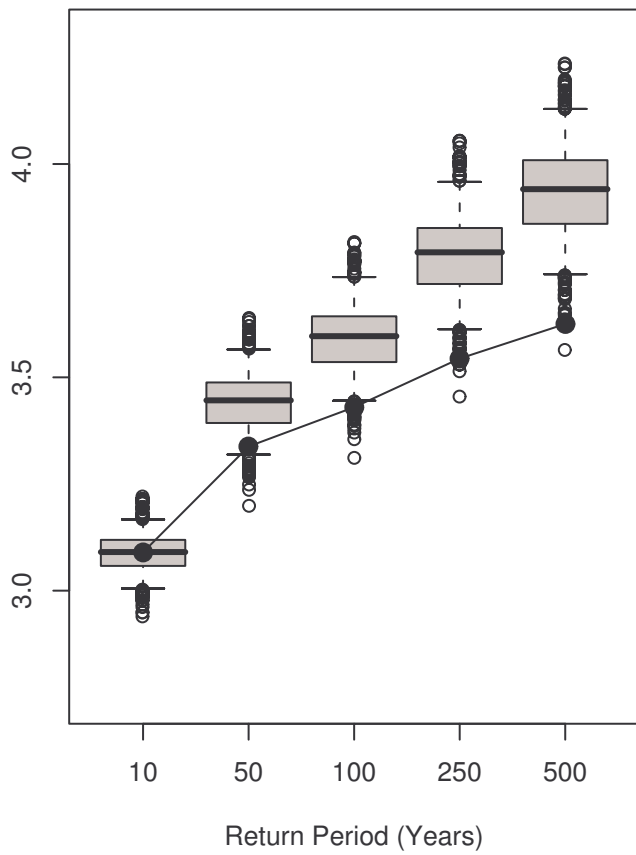
(a) True Distribution (LN)



(b) Locfit



(c) EVI



(d) Log Pearson III

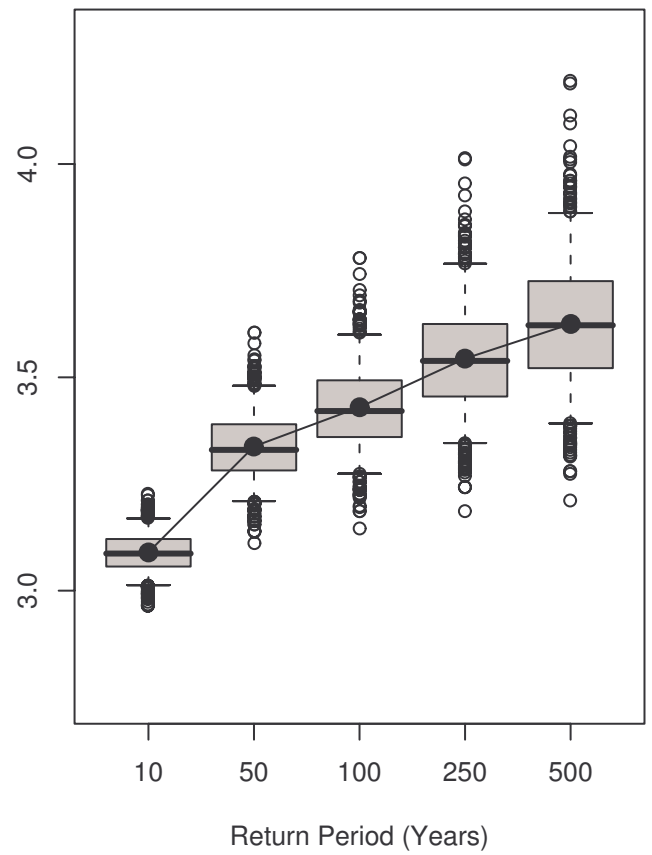
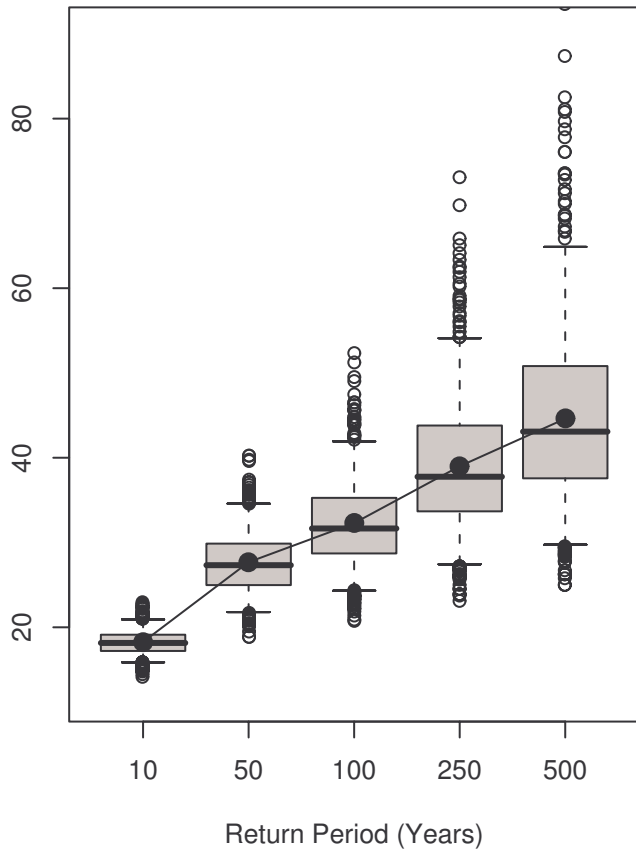
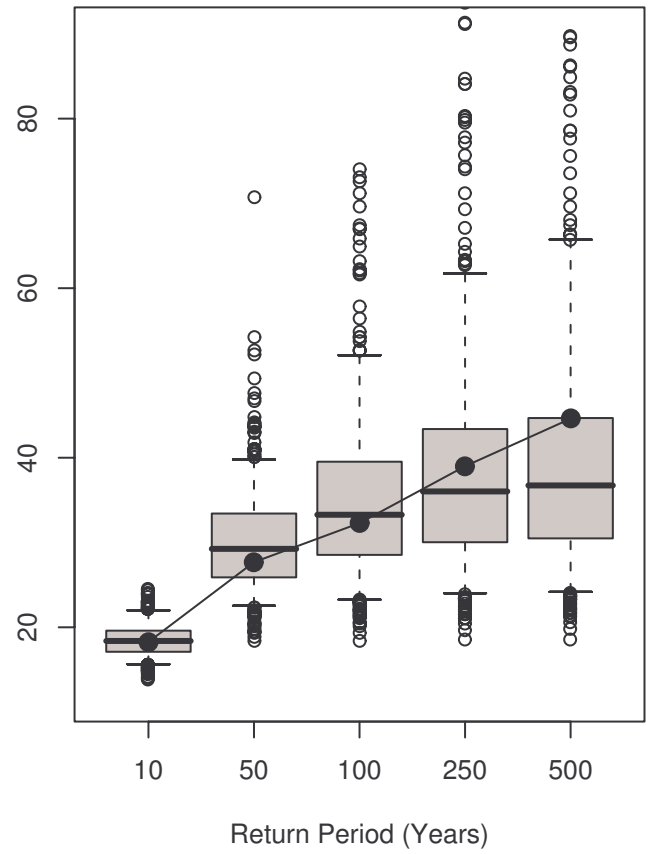


Figure-3

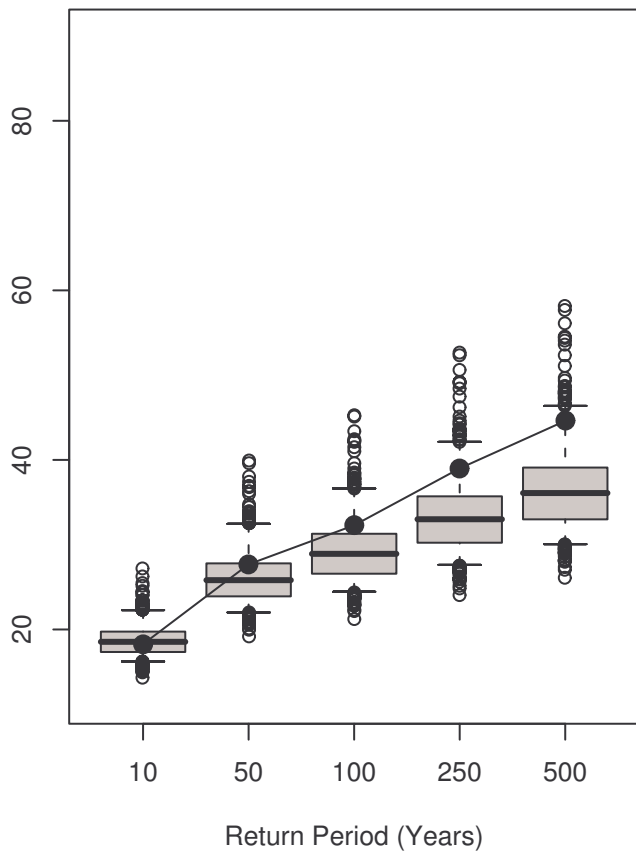
(a) True Distribution (LP3)



(b) Locfit



(c) EVI



(d) Log Normal

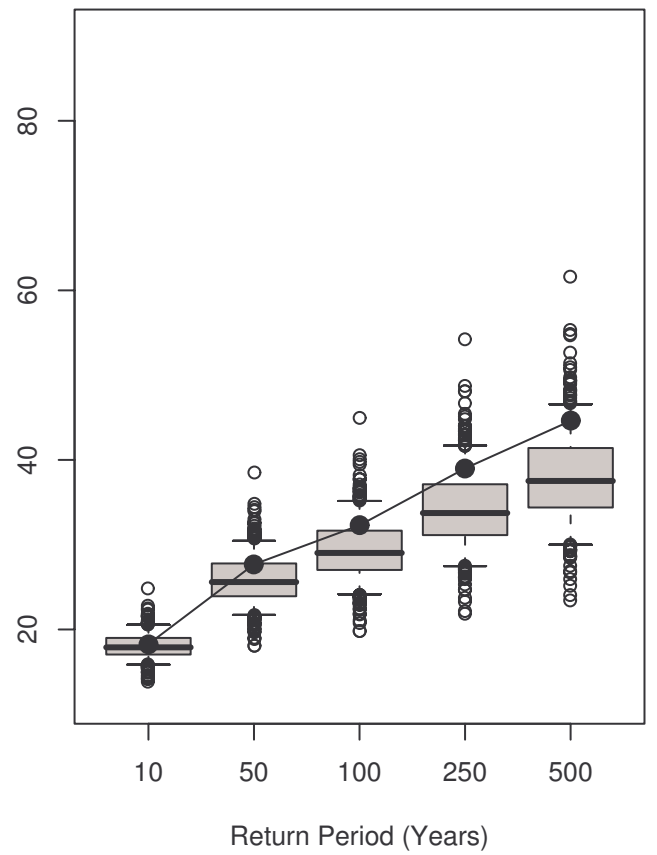
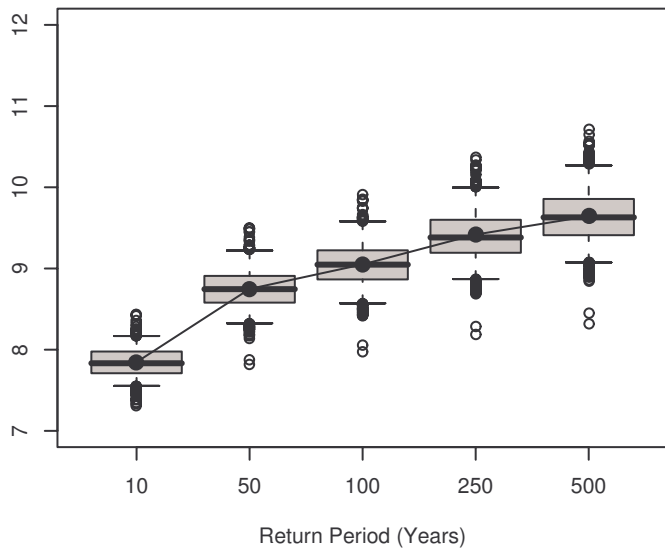
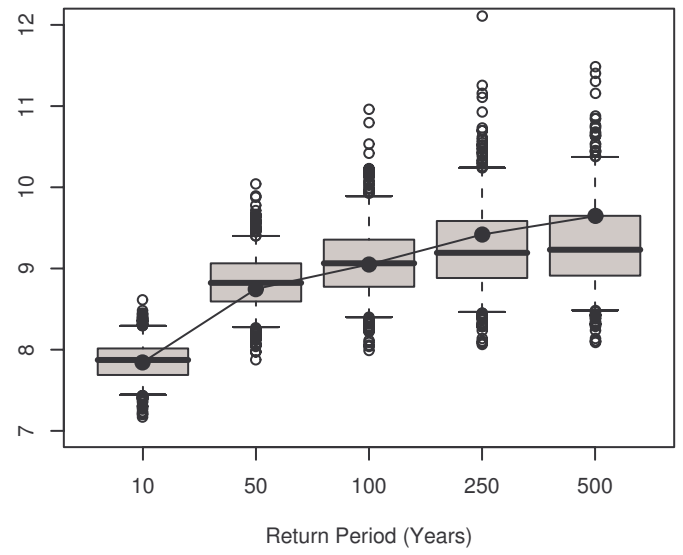


Figure-4

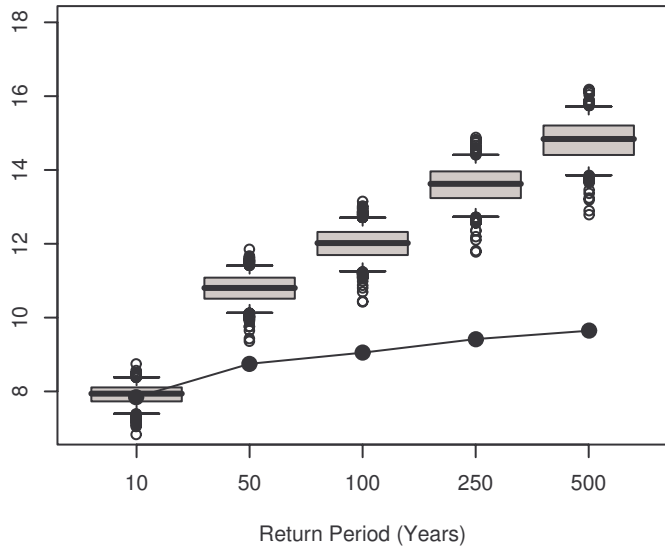
(a) True Distribution (Mix Normal)



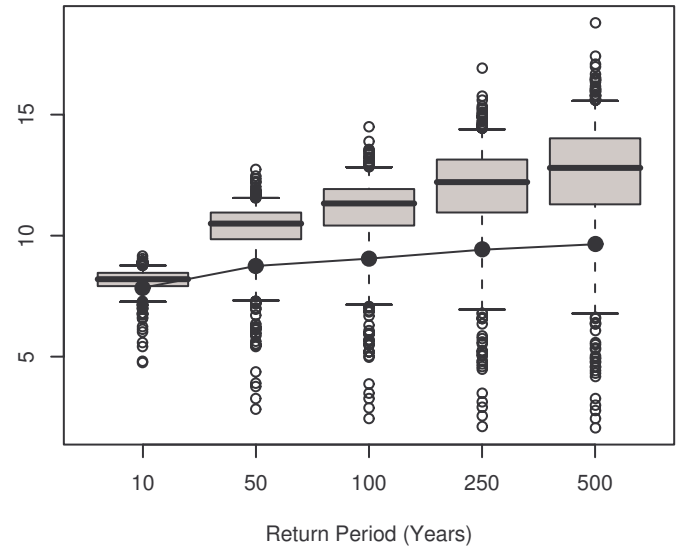
(b) Locfit



(c) EVI



(d) Log Pearson III



(e) Log Normal

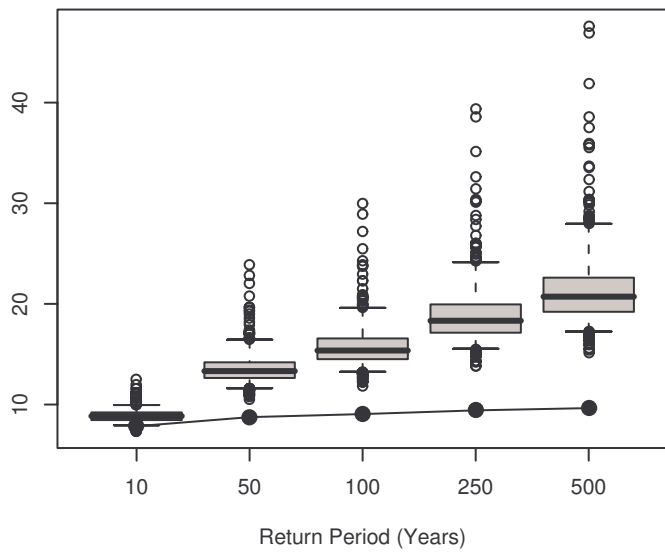
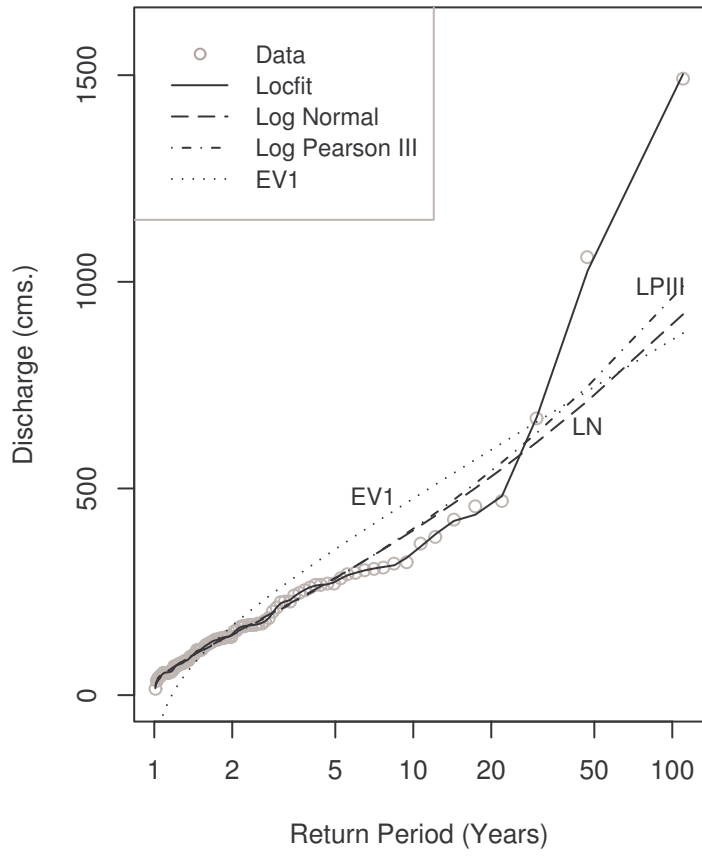


Figure-5

(a)



(b)

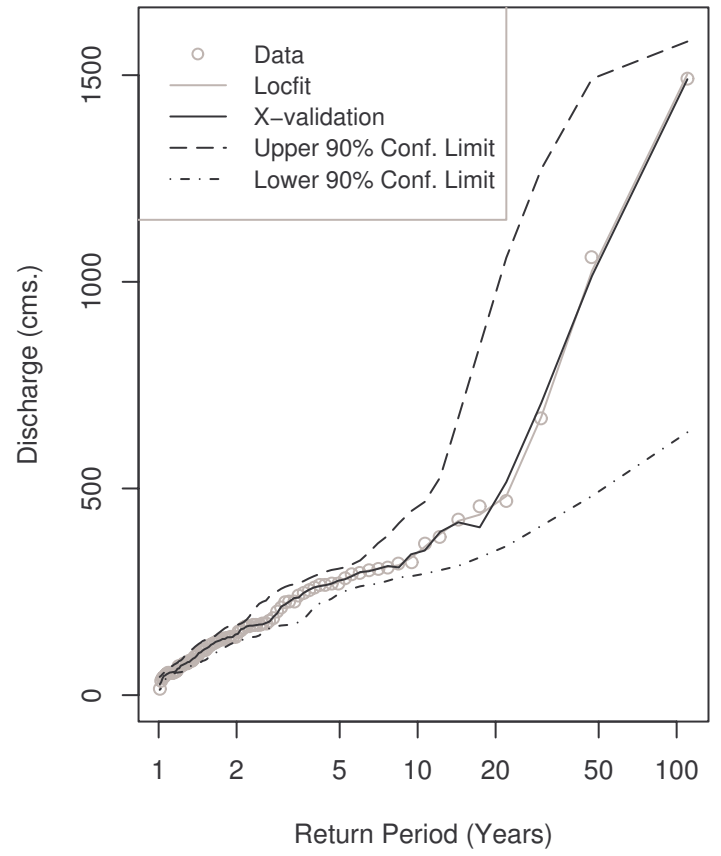
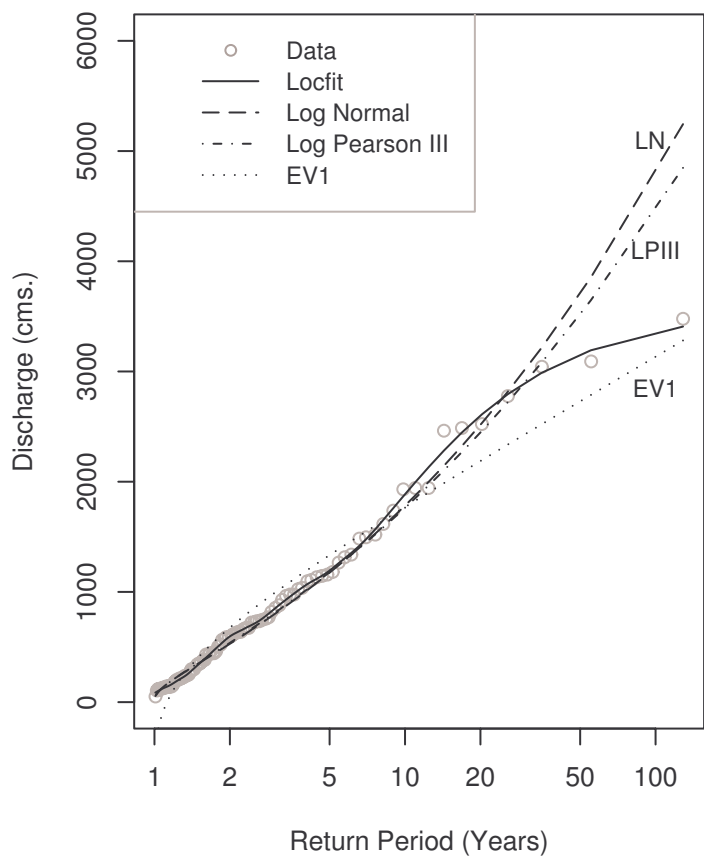


Figure-6

(a)



(b)

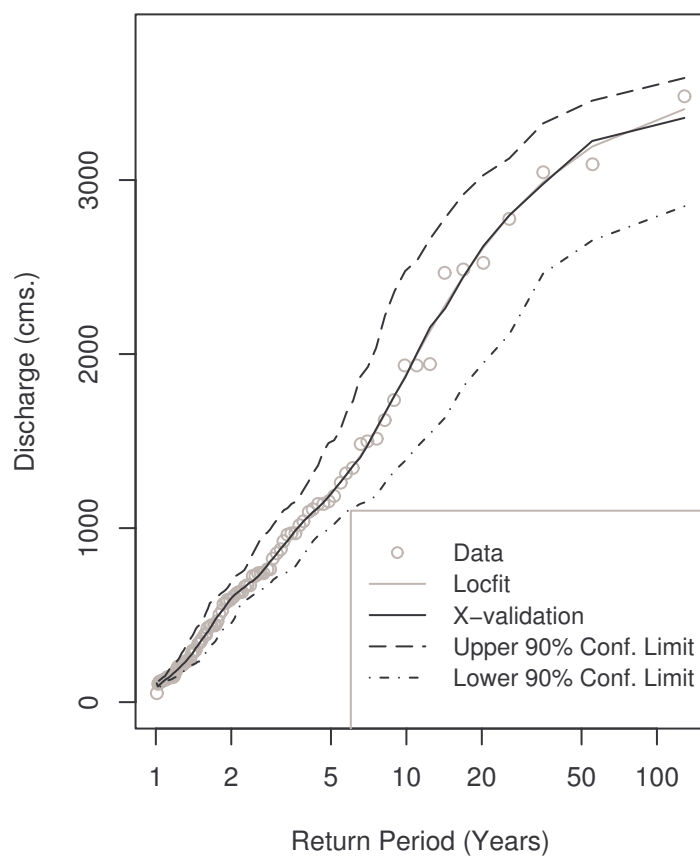
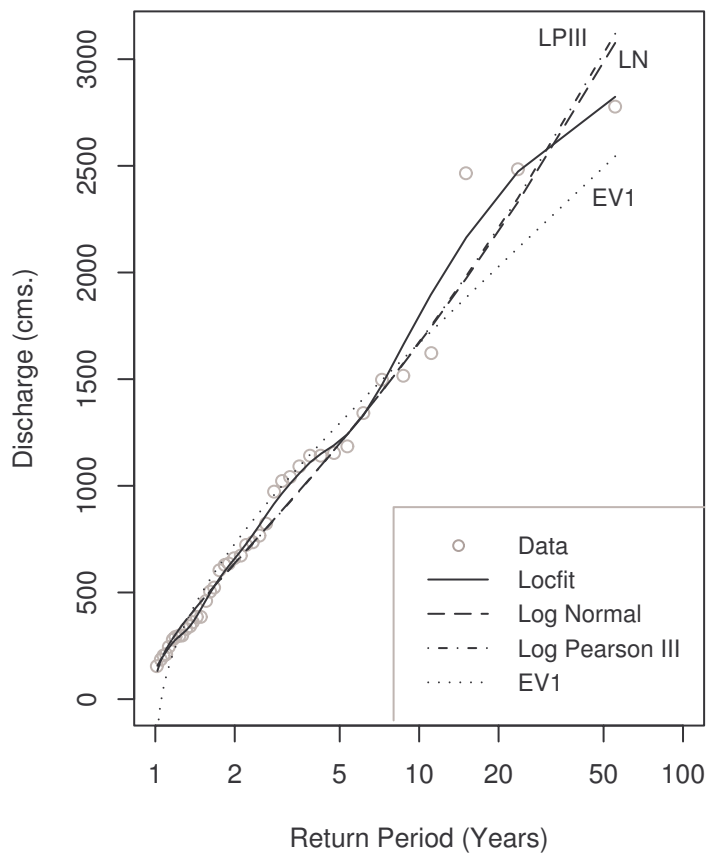
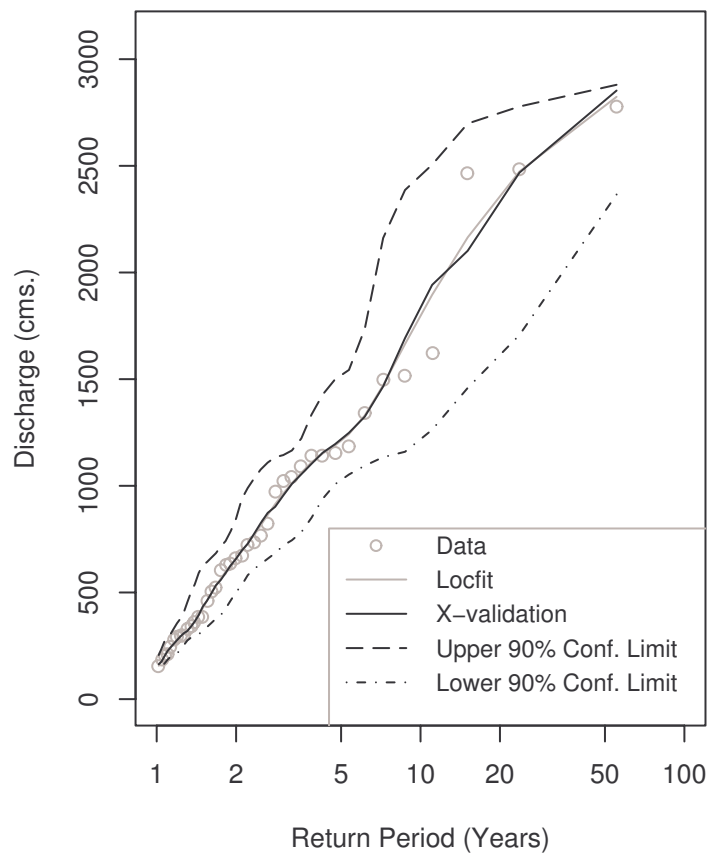


Figure-7

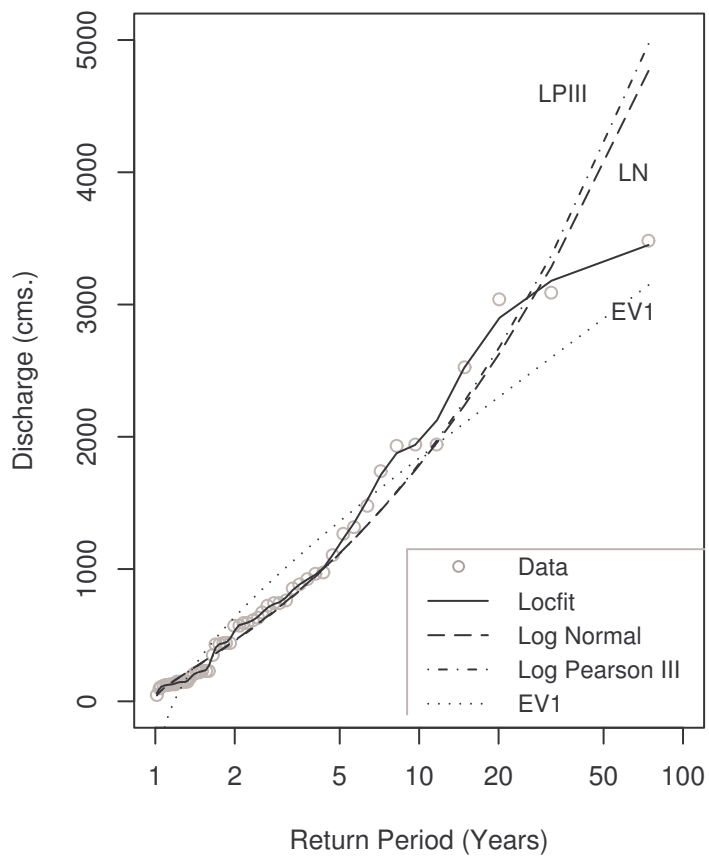
(a)



(b)



(c)



(d)

