A MONTE CARLO COMPARISON OF PARAMETRIC AND NONPARAMETRIC ESTIMATION OF LOW FLOW FREQUENCIES

Berhanu Adugna
Department of Hydraulic Engineering
Arbaminch Water Technology Institute

ABSTRACT

Low streamflow, which is one aspect of drought, constitutes one of the extremes of the hydrological regime. Among the low flow characteristics of rivers, low flow frequency analysis, that is fundamental to a wide range of design and operational problems in area of both water quality and quantity, is the one. This paper deals with the introduction of the nonparametric methods in the low flow frequency analysis and then make comparative evaluation on the magnitude of low flow quantile corresponding to a given return period with the parametric statistics. A currently used approach to low flow frequency analysis is based on the assumption made that the distribution function describing the annual minimum low flow data is known, which is never known exactly. Recently, nonparametric method of estimating probability distribution functions have been developed, which doesn't require a distributional assumption. This involves the use of a suitable smoothing function known as a kernel. The fixed kernel nonparametric method is proposed and developed for estimating low flow quantiles. Based on annual minimum low flow data and Monte Carlo Simulation Experiments, the proposed model is compared with Weibull models both for its descriptive and predictive ability. Computed results showed that the fixed kernel estimator has small bias and root mean square error in low flow quantile estimates. Application of the model to data from the Blue Nile at Eldeim (Sudan) and Komati (South Africa) rivers have shown that the nonparametric approach is viable alternative to the Weibull models. It is, therefore, concluded that the nonparametric method is accurate, uniform, and particularly suitable for the multimodal data.

INTRODUCTION

Although investigation of extremes of hydrological events (maxima and minima) attracts a great deal of research, the methodology of low flow computation is much less reflected in the available hydrological literature than the theory of floods. Some of the techniques that are commonly used to study low flows are:(1) Flow duration curves;(2) Low-Flow spells; and(3) Low flow frequency analysis. This paper only deals with single site analysis of low flow frequency analysis by using two different distinct approaches- *parametric and nonparametric frequency analysis* procedures.

A currently used approach to low flow frequency analysis is based on the concept of parametric statistical inference. In this analysis the assumption is made that the distribution function describing annual minimum flow data is known. Several estimation methods exist which may be used in these circumstances to obtain estimates of parameters and quantiles.

However, in hydrological context, there is no compelling evidence in favor of any one parametric distribution or fitting procedure. Some of the situations that cause problems with parametric methods are: selection of a particular distribution, parameter estimation, most of commonly used distributions are unimodal and sometimes a method can completely fail to produce a solution for no obvious reason. Therefore, it is evident that the parametric method, which depends on prior knowledge of the particular distribution function, has its limitations [4] and, as pointed out by [6], *" no amount of statistical refinement can overcome the disadvantage of not knowing the frequency distribution involved"*.

To overcome some of the limitations of the parametric method, there have been recent developments in the theory of nonparametric statistics that could be another possible approach to the low flow frequency analysis to estimate the probability density function nonparameterically. Such a method allows the annual minimum flow data to "speak for themselves", i.e. it does not require assumption of any functional form of density; it has also the ability of estimating multimodal distributions and therefore, can be very attractive in hydrological applications, and is well worth considering. The application of nonparametric methods for the estimation of flood quantiles have been investigated in the last decade. For flood studies, it has been shown that the nonparametric approach is competitive and presents a viable alternative to the parametric models [1,2,3,8]. This study is an attempt to introduce this important data analytic tool closer to practical application in the low flow frequency analysis. Hence, the main theme of this paper is the development of an alternative, relatively new, nonparametric kernel estimation procedures for the low flow frequency analysis and compares the quantile estimates with parametric models based on real world data and Monte Carlo simulation technique using the criteria of the descriptive ability and predictive ability of a model.

PARAMETRIC MODELS

Determination of design low flow is one of the most commonly performed analyses in hydrology. Very often these are obtained from statistical low flow frequency analysis. Such an analysis assumes a priori knowledge of probability distribution, i.e, a histogram of observed low flow data is assumed to be a sample of a population which has an underlying parent distribution. As a result, the histogram of the observed data is fitted to a single probability density function of some well known statistical distribution. The assumed distribution is fitted to the data based on the values of parameters estimated from that data. The required magnitude return period relationship is derived from the assumed distribution. The term "parametric" is used to describe this procedure. A number of distributions have been found to be applicable low flow hydrology.

In a study of low flows of rivers selected from all parts of USA, Matalas [10] investigated the suitability of four theoretical probability distributions for low flow data. The principal requirements of a suitable distributions to fit an annual minimum flow series is that it should be skewed and must have a finite lower limit greater than or equal to zero [15]. The Weibull distribution satisfies both these criteria. It has often been recommended for use as the statistical distribution for low flow frequency analysis.

Weibull Distribution and Parameter Estimation

The Weibull distribution is a reverse form of the Extreme Value Type III (EV3) distribution used in flood frequency analysis. That is, if a variate x conforms to the EV3 distribution, then $-x$ conforms to the Weibull distribution. The probability density function (pdf), $f(x)$ and the associated cumulative density function (df), $F(x)$ of the Weibull distribution are given by Eqs. (1) and (2), respectively.

$$f(x) = (\frac{\theta_1}{\theta_2})(x-\theta_3)^{(\theta_1-1)}\exp\left[-\frac{(x-\theta_3)^{\theta_1}}{\theta_2}\right] \quad (1)$$

$$F(x) = (\frac{\theta_1}{\theta_2})(x-\theta_3)^{(\theta_1-1)}\exp\left[-\frac{(x-\theta_3)^{\theta_1}}{\theta_2}\right] \quad (2)$$

where θ_1, θ_2 and θ_3 are shape, scale and lower bound displacement parameters respectively, which must be positive. $F(x)$ is also defined as the probability of a drought exceeding x. Weibull distribution can occur in the two parameter or three parameter form depending on third parameter, the lower bound. When the lower bound parameter θ_3 equals zero (i.e. for small streams), the distribution takes the two parameter form.

The Parameters of the Weibull distribution can be estimated by either graphical curve fitting method or objective methods such as the Method of Ordinary Moments (MOM), Method of Maximum Likelihood (ML) or by the Probability Weighted Moments (PWM). These procedures are described by Nathan and McMahon [11] and Gunasekara [7]. Nathan and McMahon [11], using 987 samples of natural low flow data from 134 catchments located in southeastern Australia, have reported a study which analyses the relative performance of the MOM, ML and PWM and found that the three estimation methods provide distinct sets of quantile estimates. An extensive simulation comparison of the six procedures formulated by the combination of the two forms of the Weibull distribution and the three objective methods of fitting is presented by Gunasekara [7]. His investigation was based on Monte Carlo experiments using large amount of synthetically generated data covering a wide range of populations. Both the above investigators have observed that the methods can fail to fit the distribution. The failures can be caused by the lower bound parameter being estimated is either less than zero or greater than the observed minimum discharge. Sometimes a method can completely fail to produce a solution for no obvious reason. All investigations, so far, have not resulted in a universally accepted practice for choosing a particular method of fitting the Weibull distribution for the use in low flow frequency analysis.

NON-PARAMETRIC MODEL

The nonparametric density estimation has been a part of the statistics literature since the 1960's (and with ever increasing prominence there), only recently has its walking its way into stochastic hydrology literature. A recent statistical literature review of nonparametric density estimation can be found in [16, 17].

The nonparametric approach does not make any assumption regarding the parent distribution. In its simplest form the nonparametric approach would comprise of smoothing the histogram of the observed flows and carefully extrapolating the tail of this histogram. The required magnitude return period relationship can be read directly from the smoothed histogram. The difficulty of the method clearly lies in extrapolation beyond the observed data.

Among the several different approaches that can be used in estimating the probability density function of observations by the non-parametric method, the kernel estimator, which is the most popular and well developed, theoretically, nonparametric estimation procedure, is used in this study. In its simplest form, the fixed kernel estimate of the probability density function $\hat{f}(x)$at each point x is defined for univariate data by [1]:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \qquad (3)$$

where h is a positive constant smoothing factor because it controls the degree of smoothness that the resulting function exhibits, $K()$ is a probability density function, referred to as the smoothing kernel function; and a sample of observations $x_1,\ldots,x_n$ form a random sample from the unknown distribution. The kernel estimate is non-negative and integrates to one.

The Choice of Kernel Function for Nonparametric Curve Estimation

In nonparametric frequency , the choice of the kernel function is not critical to the performance of the method as various kernels lead to comparable estimates [13]. However, it must satisfy the following conditions [16]:

$$\int K(x)dx = 1.0 \qquad (4)$$

$$\int xK(x)dx = 0.0 \qquad (5)$$

$$\int x^2K(x)dx = C \neq 0.0 \qquad (6)$$

Where C is the kernel variance.

Because different kernels will lead to different estimates in tails of distribution, a kernel with a variable degree of symmetry is currently being investigated as a solution to this problem [9]. A kernel following a normal distribution, in other words a Gaussian kernel, with standard deviation equal to h, was selected in this study and is given by:

$$K(x) = \frac{1}{\sqrt{2\pi}h}\exp\left\{-\frac{(x-x_i)^2}{2h^2}\right\} \qquad (7)$$

Estimation of Smoothing Factor h

The selection of the smoothing factor h is, however, critical. Many papers that address the problem of determining an optimal kernel use Minimumazation of integrated mean square error IMSE as optimal criterion. Various numerical algorthims for computing h perform similar to each other [2]. They are all close to the optimal value that has predicted theoretically. This value which is only suitable for quadratic kernel is given by Eq. (8)

$$h = \sum_{j=2}^{n}\sum_{i=1}^{j-1}\frac{(x_j - x_i)}{\sqrt{5}n(n - \frac{10}{3})} \qquad (8)$$

One method of computing h is to minimize, by means of a cross-validation technique, IMSE [16]:

$$\begin{aligned} IMSE &= E\int[\hat{f}(x) - f(x)]^2dx \\ &= E\{\int\hat{f}(x)^2dx - 2\int\hat{f}(x)f(x)dx \\ &\quad + \int f(x)^2dx\} \end{aligned} \qquad (9)$$

where $\hat{f}(x)$ is an estimate of the unknown density function $f(x)$.

Because IMSE must be minimized with regard to h, the last term in Eq. (9), which does not involve $f(x)$, can be discarded. Thus, the sum of the first two terms are to be minimized. Scott and Terrell [14] have shown that the cross-validation procedure leads to consistent and asymptotically optimal nonparametric density estimates. In the cross validation, estimates of $f(x)$ are constructed each time using all the data points except one. Thus, $\hat{f}_{-i}(x)$is the nonparametric kernel estimate ignoring a single data point, x_i:

$$\hat{f}_{-i}(x) = \frac{1}{(n-1)h} \sum_{j \neq i} K\left(\frac{(x - x_j)}{h} \right) \qquad (10)$$

It has been shown [16] that:

$$E \frac{1}{n} \sum_{i} \hat{f}_{-i}(x_i) = E \int \hat{f}(x) f(x) dx \qquad (11)$$

Inserting Eq. (11) into Eq. (9) and ignoring its last term, the risk function to be minimized, *R(h)*, which depends on the smoothing factor ***h***, is:

$$R(h) = E \left\{ \int \hat{f}(x)^2 dx - \frac{2}{n} \sum_{i=1}^{2} \hat{f}_{-i}(x_i) \right\} \qquad (12)$$

where $\hat{f}_{-i}(x)$ is the density estimate based on the entire data set except for x_i, beginning with an assumed h. The basic principle of the least squares cross-validation is to construct an estimate of *R(h)* from the data themselves and then to minimize this estimate over h to give the smoothing factor. After some computations, it has been shown [13] that Eq. (12), for normal kernel as defined by Eq. (7), is equivalent to:

$$R(h) = \frac{1}{2\sqrt{\pi n}h} \left\{ 1 + \sum_{i=1}^{j} \sum_{j=1}^{n} \frac{2}{n} \exp(\frac{d_{ij}}{4}) - \sum_{i=1}^{j} \sum_{j=1}^{n} \frac{4\sqrt{2}}{n-1} \exp(\frac{d_{ij}}{2}) \right\} \qquad (13)$$

where $d_{ij} = -((xi-xj)/h)2$.

Setting the derivative of *R(h)* with respect to h equal to zero results in the following equation:

$$\frac{1}{2\sqrt{\pi}nh} \left\{ \sum_{i=1, i \neq j}^{n} \sum_{j=1}^{n} \exp(\frac{d_{ij}}{4}) ((1 - \frac{4\sqrt{2n}}{n-1} \exp(\frac{d_{ij}}{4}) (\frac{x_i - x_j}{h} - 1)(\frac{x_i - x_j}{h} + 1) - 1) \right\} = 0 \qquad (14)$$

Therefore, the value of h can be determined numerically by solving Eq. (14) by using single optimization methods for given sample and kernel. A fast algorithm for calculating h can be done with golden section search method. The initial guess for the optimization may be obtained by solving Eq. (8). Then using an optimal value of h, it is possible to estimate the density function by Eq. (3).

NUMERICAL APPLICATIONS

Annual minimum low flow series for the Blue Nile at Eldeim station (Sudan)and Komati (South Africa) rivers were selected for numerical demonstrations. The statistical characteristics of these rivers are shown in Table 1. As an illustration to the methods only the results of the Blue Nile will be presented. The discharge data for the Blue Nile at Eldeim station for the period (1921-1991), based on 10-day means were obtained from Ministry of Irrigation and Water Resources-Sudan. Discharge data for the 1921-1991 based on monthly basis is also available. No missing data was encountered. The data was subject to quality control procedures. The annual minimum flow series for the Blue Nile at Eldeim station was formed by selecting the lowest flow occurring in each year of record.

Table 1: Data Selected for the Analysis

River	Period of flow records, years	Sample size n	Arithmetic mean ($10^6 m^3/d$)	Standard Deviation ($10^6 m^3/d$)	Coefficient of variation	Coefficient of Skewness
Blue Nile	1921-91	71	8.173282	2.87358	1.117870	1.117870
Komati	1941-90	50	2.905241	1.567646	0.2208143	0.2208143

The nonparametric kernel density estimation procedure requires only one parameter, the smoothing factor h which is optimized for the given sample. It is obtained by solving Eq. (14), and the value is $h = 0.412695$. The nonparametric density estimation for the Blue Nile at Eldeim is calculated by Eq. (3)and the results are presented in Fig. 1(a). It can be observed that the probability density of annual minimum flow series at Blue Nile is distinctly bimodal. Therefore, if nonparametric density analysis had not been performed, it would have been possible to mistakenly conclude that the annual minimum flow of the Blue Nile at Eldeim come from a Weibull distribution (unimodal); but, in fact, the river follows a multimodal distribution . Hence, nonparametric frequency analysis can be considered as an important screening tool in distribution identification. This ability to analyze multimodal density by nonparametric method is particularly useful in hydrology.

The frequency curve is plotted on the EV1 paper on which the Gringorton formula

$$F(i) = \frac{i - 0.44}{n + o.12} \quad i = 1,2,\ldots,n \tag{15}$$

is used to calculate the plotting positions for recorded sample data. Fig. 1(b) shows that the nonpara-metric kernel estimator can fit the real data points closely.

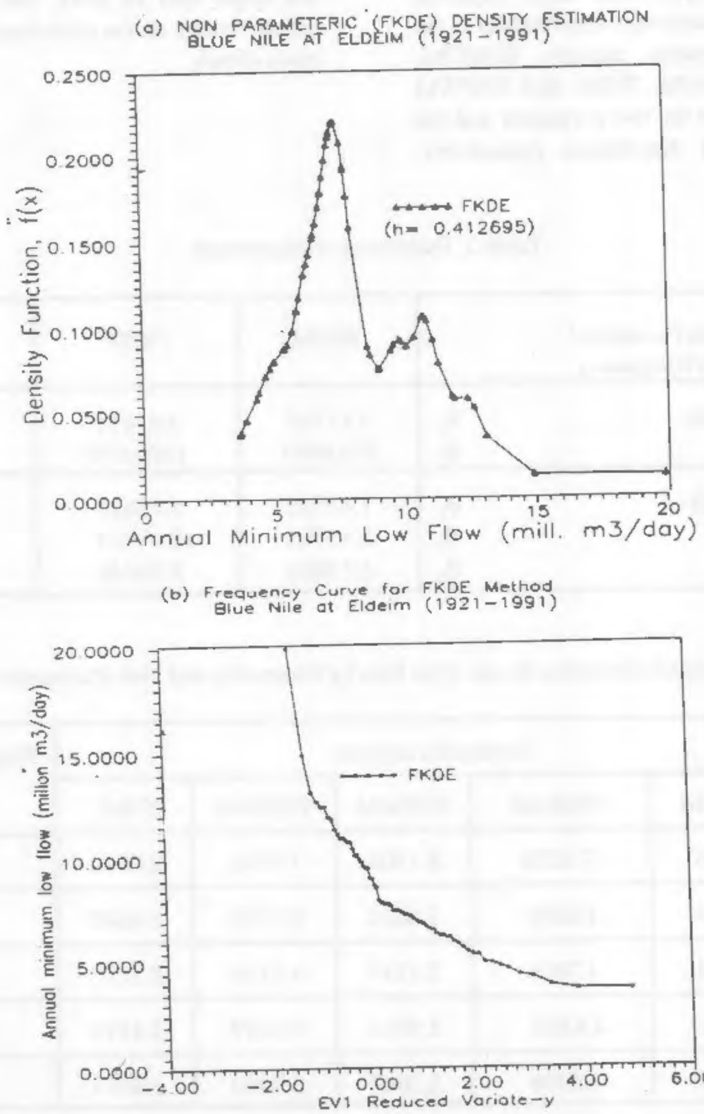

Figure 1 Nonparametric fixed kernel density estimation and frequency curve at Blue Nile at Eldeim (Sudan)

COMPARISON OF NONPARAMETRIC WITH PARAMETRIC MODELS

The criteria of the descriptive ability and the predictive ability of a model are used for comparison of the proposed nonparametric method with the Weibull parametric models. The predictive ability criterion relates to the ability of a chosen model to describe or fit the low flow data while the predictive ability relates to procedure's ability to achieve its assigned task that is to have minimum bias and maximum efficiency for quantile estimates[5].

Test of Descriptive Ability

The annual minimum low flow series at Blue Nile was fitted to the two parameter and the three parameter Weibull distribution coupled with three objective parameter estimation procedures. This resulted in six different parametric models, namely, 2P/MOM, 2P/ML, 2P/PWM, 3P/MOM, 3P/ML and 3P/PWM where 2P and 3P denotes the two parameter and the three parameter Weibull distribution respectively. Table 2 gives estimation of parameters by different methods. Only the 3P/Ml model failed to obtain a solution because the estimated lower bound parameter was calculated as being less than zero.

Table 3 summarizes theoretical low flow quantile and compared with those estimated by the nonparametric model. The results in Table 3 indicates that the nonparametric method gives low flow quantile estimates which are comparable with those obtained by applying different assumed Weibull distributions. The frequency curves of quantile estimates by the Weibull models and nonparametric method are plotted on the EV1 probability paper with recorded data points, see figure 2. It can be seen that the two parameter Weibulll models are underestimated while the three parameter Weibull models are overestimated when return periods are larger than 10 years. The nonparametric kernel estimators can, on the other hand, fit the observed data more closely.

Table 2. Estimation of Parameters

Distribution Parameter/ Method of Estimation		MOM	PWM	ML
2P(2 Parameter) Weibull	θ_1	3.11310	3.41473	2.935927
	θ_2	979.8093	1880.079	662.7628
3P(3 parameter) Weibull	θ_1	1.41702	1.76022	1.4576
	θ_2	8.19769	18.09323	9.9946
	θ_3	4.15804	3.56040	-395.87

Table 3: Quantile Estimates for the Blue Nile by Parametric and Non-Parametric Models

Return	Parametric analysis					Nonparametric analysis
period(yr)	2P/MOM	3P/MOM	2P/PWM	3P/PWM	2P/ML	FKDE
2	8.1226	7.5658	8.1700	7.7676	8.0679	7.7891
5	5.6439	5.6895	5.8624	5.7701	5.4840	6.1162
20	3.5194	4.7006	3.8114	4.5188	3.3237	4.4149
50	2.6091	4.4392	2.9012	4.1249	2.4199	3.8488
100	2.0849	4.3298	2.3647	3.9400	1.9077	3.5872

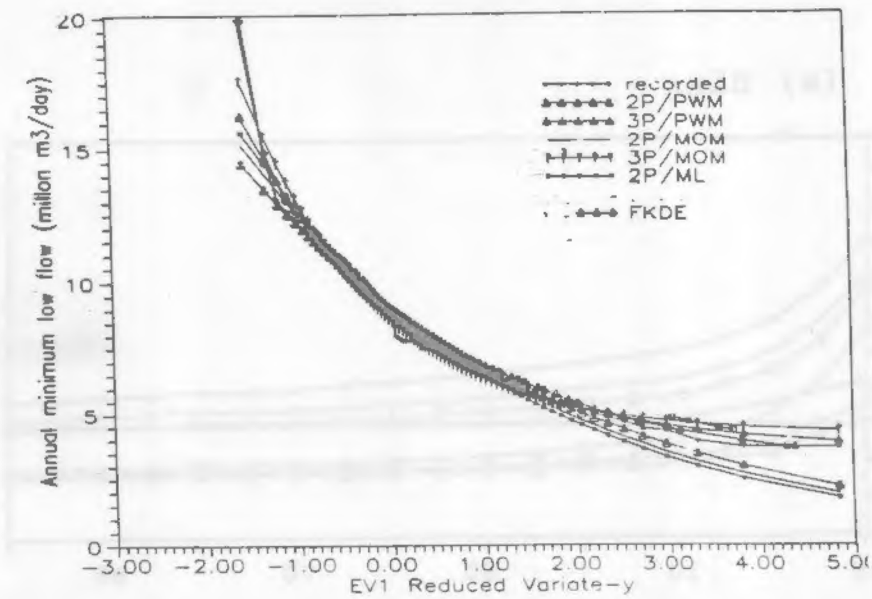

Figure 2 Comparison of parametric and nonparametric low flow frequency curve at Blue Nile at Eldeim (Sudan). Sample size n=71, Mean=8.173*10^6 (m^3/day), Cv=0.3516 and Cs=1.17564

Test of predictive ability

Synthetically generated data were used to compare the performance of the various methods used in this study of low flow frequency analysis. A population conforming to the Weibull distribution was chosen, largely because of its recommendation and it fits a wide range of low flow observations. Such a choice is rather arbitrary and the credibility of simulation rests on the credibility of the Weibull parent to simulate empirical low flow data. It, of course, should be recognized that the Weibull distribution does not encounter all diverse hydrological regions. However, having accepted the Weibull distribution as the low flow parent, it is possible to obtain population estimates of the T-year low flow for the simulated data.

Synthetic samples were generated from the Weibull distribution with given parent parameters and sample size. The range of the sample lengths used was 10 to 100 with an increment of 5 and the return periods used were 100, 50, 20, 5, and 2 years. Altogether 12 synthetic populations were used with θ_1 varying from 1.5 to 4.0, θ_2 varying from 0.3 to 4 and θ_3 varying from 0 to 0.5, covering almost all practical situations. The proposed nonparametric kernel estimator and six Weibull models were fitted to each sample, and the bias and root mean square error (RMSE) of quantile estimates were calculated by Eqs. (16) and (17).

$$BIAS(T) = \frac{1}{M}\sum_{i=1}^{M}\frac{\hat{X}_{i,T}-X_T}{X_T} * 100\% \qquad (16)$$

and

$$RMSE(T) = \left|\frac{1}{M}\sum_{i=1}^{M}\left(\frac{\hat{X}_{i,T}-X_T}{X_T}\right)^2\right|^{0.5} * 100\% \qquad (17)$$

Where T is the return period; M is the number of Monte Carlo repetitions(M = 1000 in this study. 1000 may arguably not be a large enough number of samples to produce the true values of BIAS and RMSE, but will suffice to compare the performances of the methods); $\hat{X}_T$ and X_T represents the estimated and the theoretical quantile values, respectively.

Results of the predictive ability tests were plotted against the sample length (for example, BIAS versus sample length). All the plots were repeated for different populations and also for different return periods. A typical results obtained for population θ_1 = 2.0, θ_2 = 1 and θ_2= 0.1 are shown in Figs. 3 and 4 for illustration.

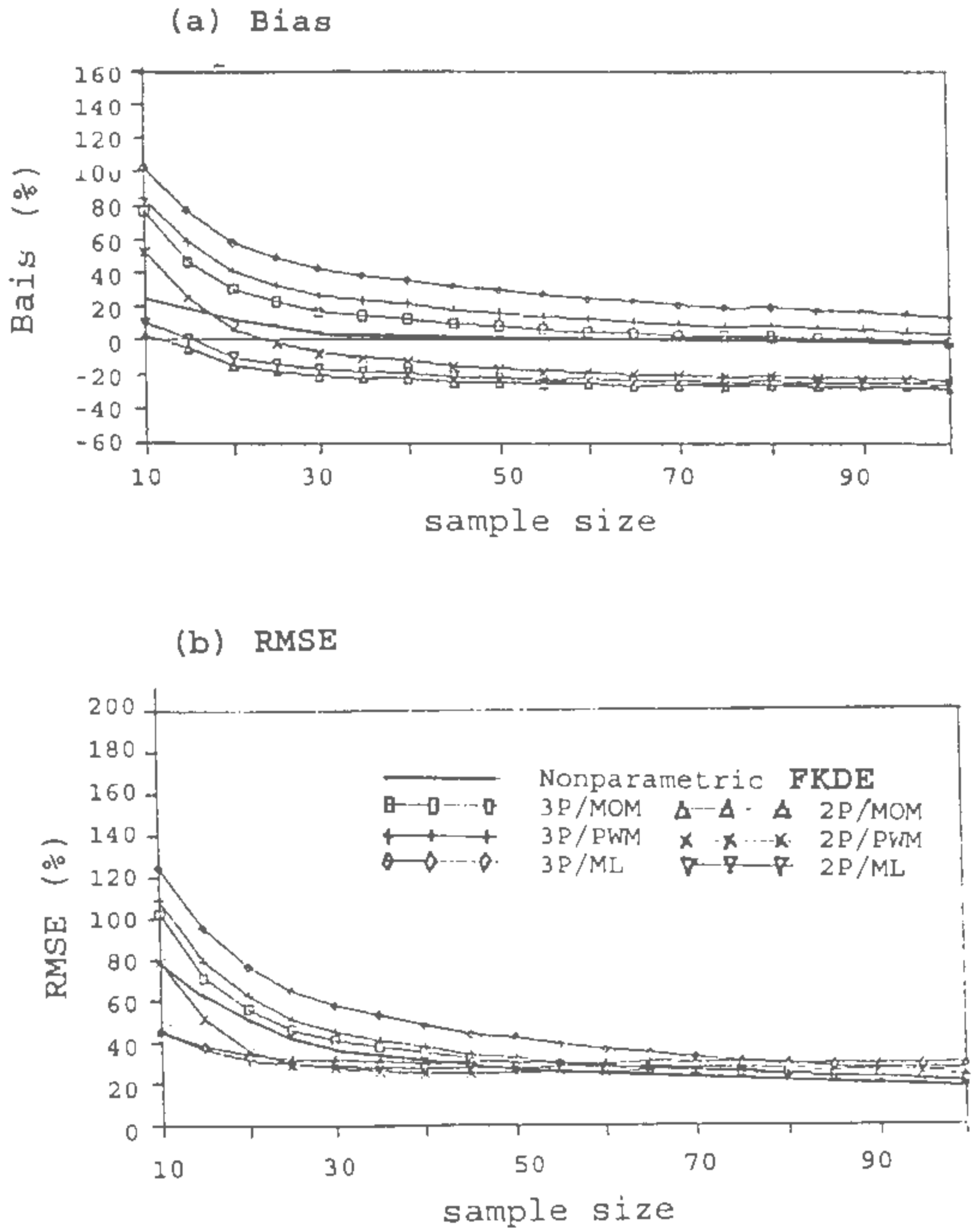

Figure 3 Bias and RMSE of 100 year flow quantile as a function of sample size forpopulation No.1 (θ_1=2.0,θ_2=1,θ_3=0.1 and $X_{T=100}$=0.200), (a) BIAS and (b) RMSE

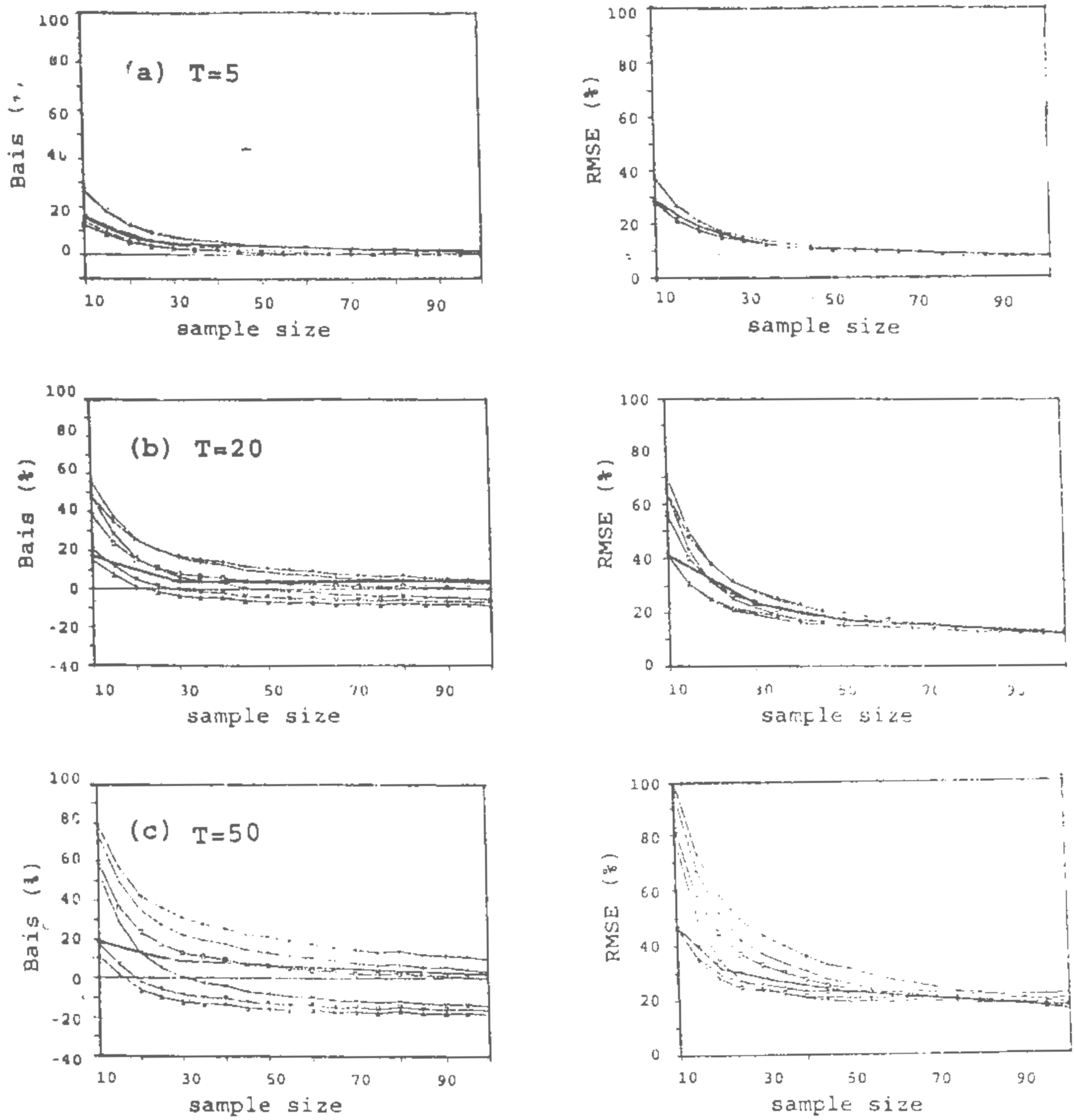

Figure 4 Bias and RMSE of different low flow quantile as a function of sample size for population No. 1($\theta_1 = 2.0$, $\theta_2 = 1$, $\theta_3 = 0.1$ and $X_{T=5} = 0.572$, $X_{T=20} = 0.327$ and $X_{T=50} = 0.242$), (a) return period = 5, (b) return period = 20, (c) return period = 50

The following clear observation can be made based on the simulation results.

(1) In contrast to the flood frequency, there is no significant difference among the three fitting procedures (MOM, ML, and PWM) in the low flow quantile estimation for the Weibull distributions. These figures show that the MOM performs better than the ML and PWM, and ML has the largest bias and RMSE for small size samples.

(2) The two parameter Weibull models perform better than the three parameter Weibull models for small sample size. When the sample size increase, the three parameter Weibull models are better than the two parameter Weibull models, and can obtain small bias and RMSE in quantile estimation. Simulation experiments show the conclusions drawn for 100 year low flow is equally valid for 50 year as well as 20 year return periods. For low flow quantiles of very low return periods such as 5 or 2 years, there is no significant difference between any procedures.

(3) This study reveals the usefulness of nonparametric technique in analysis of low flow data generated by distinctly different hydrological processes. Nonparametric fixed kernel estimator result in small bias in quantile estimation for different sample sizes and parent parameters. Its RMSE is close to the two parameter Weibull models when sample size is small and has the smallest values when sample size is increased.

(4) Since the sample comes from the Weibull family, this is obviously advantageous to the Weibull models. It is expected that the bias and RMSE estimated by the Weibull models will increase if the sample comes from other parent distributions. However, the nonparametric model does not need any parent assumption and its density function is estimated directly from the samples, and there is no any relationship with parents.

CONCLUSIONS

To overcome some of the deficiencies of the conventional parametric approaches, there has been a recent trend to develop new nonparametric approaches. In this study, a recently developed nonparametric method is proposed, investigated and compared with the parametric methods. Results in tables and figures illustrate an important difference in the application philosophy of parametric and nonparametric methods. The nonparametric model has some advantageous, mainly it is distribution free and can fit different data well, and disadvantageous such as the limitation of extrapolation. They do also nothing to avoid the assumption of independent. The nonparametric method usually gives good results for data which is long and can indicate the presence of multimodality in the analyzed data. The nonparametric model can always fit sample data whereas the Weibull distribution fails in some instances particular for three parameter form with small sample size. Simulation results demonstrate that evenif the distribution of low flows was known (i.e, Weibull), then the nonparametric method produces small bias and RMSE. The experiments, therefore, clearly suggest that the nonparametric method be a viable alternative and provides another useful tool in low flow frequency analysis.

REFERENCES

[1] **Adamowski, K.,** "Nonparametric kernel estimation of flood frequency," Water Resources Research, vol 21, No 11,pp 1585-1590, 1985.

[2] **Adamowski, K.,** "A Monte Carlo comparison of parametric and nonparametric estimation of flood frequencies. Journal of Hydrology, vol 108, pp 295-308, 1989.

[3] **Bardsely, W.E.**, "A simple parameter-free flood magnitude estimator," Hydrological Science Journal, vol 34, No 2, pp 129-137, 1985.

[4] **Cunnane, C.**," Factors affecting choice of distribution for flood series," Journal of Hydrological Science, vol 30, No (1,3), pp 25-36, 1985.

[5] **Cunanne, C.**,"Statistical distribution for flood frequency analysis. WMO Operational Hydrology Report, No.33, 72pp.

[6] **Dooge, J C E**.," Looking for hydrologic laws," Water Resources Research, vol 22, No 9, pp 465-585, 1986.

[7] **Gunasekara, T A G.**," Evaluation of flood and low flow frequency analysis procedures," PhD Thesis, National University of Ireland, Galway, 224pp, 1991.

[8] **Guo, S L.**,"Nonparametric variable kernel estimation with historical flood and palaeoflood information," Water Resources Research, vol 27, No 1, pp 91-98, 1991.

[9] **Lall, U , Moon, Y and Bosworth, K.**,"Kernel flood frequency estimators: Bandwidth selection and kernel choice," Water Resources Research, vol 29, No 4, pp 1003-1015, 1993.

[10] **Matalas, N C.**,"Probability distribution of low flows," US Geological Survey Professional Paper, 434-A, 27pp, 1963.

[11] **Nathan, R J and McMahon, T A.**,"Practical aspects of low flow frequency analysis," Water Resources Research, vol 26, No 9, pp 2135-2147, 1990.

[12] **Rao, B L S P.**, "Nonparametric Function Estimation," Academic, San Diego, California, p 522, 1983.

[13] **Rudemo, M.**, "Emperical Choice of Histogram and Kernel Density Estimation," Scandinavian Journal of Statistics, vol 9, pp 65-78, 1982.

[14] **Scoot, D W and Terrel, G R.**," Biased and unbiased cross-validation in density estimation," Journal of the American Statistical Association, vol 82, pp 1131-1146, 1986.

[15] **Shaw, M E.**," Hydrology in practice," McGraw Hill, New York, 1980.

[16] **Silverman, B W.**," Density estimation for statistics and data analysis," New York: Chapman and Hall, 1986.

[17] **Tapia, R A, and Thompson, J R.**," Nonparametric probability density estimation," John Hopkins University Press, Baltimore, MD, 1978.