



Article

State-of-the-Art Statistical Approaches for Estimating Flood Events

Muhammad Fawad ^{1,2}, Felício Cassalho ³, Jingli Ren ¹, Lu Chen ⁴ and Ting Yan ^{5,*}

¹ Zhengzhou Key Laboratory of Big Data Analysis and Application, Henan Academy of Big Data, Zhengzhou University, Zhengzhou 450052, China; fawadkhn42@gmail.com (M.F.); renjl@zzu.edu.cn (J.R.)

² School of Mathematics and Statistics, Zhengzhou University, Zhengzhou 450001, China

³ Department of Civil, Environmental, and Infrastructure Engineering, George Mason University, 4400 University Dr, Fairfax, VA 22030, USA; fcassalh@gmu.edu

⁴ College of Hydropower & Information Engineering, Huazhong University of Science & Technology, Wuhan 430074, China; chen_lu@hust.edu.cn

⁵ Department of Statistics, School of Mathematics and Statistics, Central China Normal University, Wuhan 430079, China

* Correspondence: tingyant@mail.ccnu.edu.cn

Abstract: Reliable quantile estimates of annual peak flow discharges (APFDs) are needed for the design and operation of major hydraulic infrastructures and for more general flood risk management and planning. In the present study, linear higher order-moments (LH-moments) and nonparametric kernel functions were applied to APFDs at 18 stream gauge stations in Punjab, Pakistan. The main purpose of this study was to evaluate the impacts of different quantile estimation methods towards water resources management and engineering applications by means of comparing the state-of-the-art approaches and their quantile estimates calculated from LH-moments and nonparametric kernel functions. The LH-moments ($\eta = 0, 1, 2$) were calculated for the three best-fitted distributions, namely, generalized logistic (GLO), generalized extreme value (GEV), and generalized Pareto (GPA), and the performances of these distributions for each level of LH-moments ($\eta = 0, 1, 2$) were compared in terms of Anderson–Darling, Kolmogorov–Smirnov, and Cramér–Von Mises tests and LH-moment ratio diagrams. The findings indicated that GPA and GEV distributions were best fitted for most stations, followed by GLO distribution. The quantile estimates derived from LH-moments ($\eta = 0, 1, 2$) had a lower relative absolute error, particularly for higher return periods. However, the Gaussian kernel function provided a close estimate among nonparametric kernel functions for small return periods when compared to LH-moments ($\eta = 0, 1, 2$), thus highlighting the importance of using LH-moments ($\eta = 0, 1, 2$) and nonparametric kernel functions in water resources management and engineering projects.

Keywords: water resources management; extreme events; return periods; probability distributions; nonparametric



Citation: Fawad, M.; Cassalho, F.; Ren, J.; Chen, L.; Yan, T. State-of-the-Art Statistical Approaches for Estimating Flood Events. *Entropy* **2022**, *24*, 898. <https://doi.org/10.3390/e24070898>

Academic Editor: Donald J. Jacobs

Received: 4 June 2022

Accepted: 27 June 2022

Published: 29 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Extreme environmental events have always been a vital part of human history. Floods, rainstorms, droughts, and windstorms are some of the manifestations of these events, which cause enormous destruction. Flood is often referred to as one of the most devastating natural disasters in terms of damage to property, infrastructure, and the environment, even threatening human lives.

Quantile estimations of floods, commonly extracted from annual peak flow discharges (APFDs), are of great importance for the description of such events. These estimates allow for the assessment of flood characteristics by associating their magnitude to a corresponding frequency, from which mitigating hydraulic structures and management practices may be designed. The objective of flood frequency analysis (FFA) is to obtain a flood quantile

magnitude estimator for one or more stations on the river system. Depending on the magnitude of the flood, an estimate of its return period may be required. The common interest in estimating quantiles of extreme floods for different return periods, i.e., 50-year, 100-year, or 500-year flood, is the design of hydraulic structures such as dams, culverts, and bridges [1–3].

The parametric estimation method (LH-moments) and nonparametric method (kernel functions) were used in the present study to estimate the flood magnitude at different return periods. In parametric methods, the APFDs are assumed to be independent and identically distributed, as well as being drawn from a population with a known probability distribution function (PDF). In addition, an adequate PDF is selected from a set of candidate PDFs using a robust goodness-of-fit test. As described by [4], extensive literature for FFA is available, not only for the detailed description of PDFs, but also for parameter estimation. The commonly used PDFs for modeling of the APFDs include: generalized extreme value (GEV), generalized logistic (GLO), generalized Pareto (GPA), Pearson type 3 (P3), log Pearson type 3 (LP3), Weibull (WEI), extreme value type 1, extreme value type 2, normal, log normal, gamma, and exponential [4–9].

The widely used parameter estimation techniques for PDFs include maximum likelihood, L-moments, LH-moments, and method of moments. The main drawback of the maximum likelihood and method of moments is that the product moments of the APFDs are similarly affected by the small sample size. Furthermore, the higher moments (e.g., coefficient of variation and skewness) are greatly influenced by the extremes observation in the data series [6]. On the other hand, L-moments are less affected by extreme observations in the data series [10]. Wang [11] introduced LH-moments, which are the generalization of L-moments, using higher order, i.e., “H”, L-moments. Wang found that LH-moments produce consistent quantile estimates for large return periods, since LH-moments provide greater weight to the larger values in the APFD series and hence better fits to the upper tail of PDFs. This characteristic is even more relevant when sample sizes are small, which is a common reality of streamflow monitoring in developing countries, including Pakistan [8,9,12–14]. Various studies have been conducted on LH-moments in different regions of the world [11,15–23]. Most of these researchers have focused on comparing LH-moments with L-moments by using goodness-of-fit-tests for different PDFs.

The aforementioned studies were based on parametric methods, which require a priori PDF selection. Alternative nonparametric methods do not assume the APFD series in a distributional form. Many studies on FFA based on the nonparametric approach have been conducted [24–33], among which kernel function estimators have stood out for producing the most reliable nonparametric methods. Adamowski [24,25] proposed a nonparametric kernel estimation for FFA and conducted a Monte-Carlo simulation experiment in order to compare the nonparametric approach with parametric PDFs, namely, LP3 and WEI distributions. His results show that the nonparametric method produces more accurate estimates than parametric methods, but the probability of extrapolation is lower than the highest data observed in the sample. Lall et al. [34] stressed the choice of a kernel function that reflects the shape and bandwidth of FFA. Kernel function and bandwidth selection techniques are implemented in three situations: Gaussian data, skewed data, and mixed data.

By definition, the extreme event is rare and often occurs in a short time period; therefore, estimating floods for large return periods is a challenging task, often leading to gross errors in the estimation of quantiles. Another problem is the identification of a suitable statistical model. The standard methods of estimation, including maximum likelihood, method of moments, least squares, may not give consistent quantile estimates for large return periods when the sample size is small. Therefore, we need an appropriate method of estimation that gives consistent quantile estimates [35]. The current research attempts to highlight the significance of LH-moments ($\eta = 0, 1, 2$) and nonparametric kernel functions in FFA. On the one hand, this allows for comparison of the two estimation approaches within this framework, while on the other hand, it observes the effect of utilizing RAE to evaluate

quantiles for various return periods. This study does not imply that nonparametric kernel functions are always the best fit and should be used in place of adequately constructed parametric methods. However, the possibility of obtaining the best fit for nonparametric kernel functions may be more significant when using FFA, as these approaches give precise estimations for observed flood variables [36]. Therefore, the current study emphasizes the significance of using LH-moments ($\eta = 0, 1, 2$) and nonparametric kernel functions to develop a comprehensive framework for at-site FFA. The steps in the framework are as follows: (i) selection of the optimized PDFs with a comprehensive level of LH-moments ($\eta = 0, 1, 2$) for the annual peak flow of 18 stations; (ii) estimation of quantiles for various flood return periods through LH-moments ($\eta = 0, 1, 2$) and nonparametric kernel functions; and (iii) comparison of various flood return periods in terms of RAE. This paper is arranged into five sections. Section 2 contains a detailed step-by-step explanation of the methodology. The details of the study area are presented in Section 3. The results and discussions of the study are provided in Section 4. Section 5 summarizes the conclusions of the study.

2. Methods

2.1. Linear Higher Order-Moments (LH-Moments)

LH-moments were proposed by [11] as the expectations of a linear combination of higher-order statistics. Let n be the sample size drawn from the distribution $F(x) = Pr(X \leq x)$; then the four LH-moments are defined as follows:

$$\lambda_1^\eta = E[X_{(\eta+1):(\eta+1)}] \quad (1)$$

$$\lambda_2^\eta = \frac{1}{2}E[X_{(\eta+2):(\eta+2)} - X_{(\eta+1):(\eta+2)}] \quad (2)$$

$$\lambda_3^\eta = \frac{1}{3}E[X_{(\eta+3):(\eta+3)} - 2X_{(\eta+2):(\eta+3)} + X_{(\eta+1):(\eta+3)}] \quad (3)$$

$$\lambda_4^\eta = \frac{1}{4}E[X_{(\eta+4):(\eta+4)} - 3X_{(\eta+3):(\eta+4)} + 3X_{(\eta+2):(\eta+4)} - X_{(\eta+1):(\eta+4)}] \quad (4)$$

where

$$E[X_{(j:n)}] = \frac{m!}{(j-1)!(m-1)!} \int_0^1 x(F) F^{j-1} (1-F)^{m-j} dF \quad (5)$$

When $\eta = 0$, LH-moments become equal to L-moments [10]. As η increases, it reflects more and more characteristics of the upper part of the PDFs and extreme events in data. Here, λ_1^η provides the location of the distribution, λ_2^η is the spreadness of the distribution, λ_3^η represents how the upper part of the distribution is asymmetric, and λ_4^η measures the peakedness of the upper parts of the distributions. For $\eta = 0, 1, 2$, the LH-moments are referred to as L-moments, L1-moments, and L2-moments, respectively. The LH-moment ratios are described below.

$$\tau_2^\eta = \frac{\lambda_2^\eta}{\lambda_1^\eta} \quad (6)$$

$$\tau_3^\eta = \frac{\lambda_3^\eta}{\lambda_2^\eta} \quad (7)$$

$$\tau_4^\eta = \frac{\lambda_4^\eta}{\lambda_2^\eta} \quad (8)$$

Let $x_1 \leq x_2 \leq \dots \leq x_n$ be the order statistic; then the unbiased estimators of LH-moments are given below.

$$\hat{\lambda}_1^\eta = \frac{1}{\binom{n}{\eta+1}} \sum_{i=1}^n \binom{i-1}{\eta} x_i \tag{9}$$

$$\hat{\lambda}_2^\eta = \frac{1}{2 \binom{n}{\eta+1}} \sum_{i=1}^n \left[\binom{i-1}{\eta+1} - \binom{i-1}{\eta} \binom{n-i}{1} \right] x_i \tag{10}$$

$$\hat{\lambda}_3^\eta = \frac{1}{3 \binom{n}{\eta+1}} \sum_{i=1}^n \left[\binom{i-1}{\eta+2} - 2 \binom{i-1}{\eta+1} \binom{n-i}{1} + \binom{i-1}{\eta} \binom{n-i}{2} \right] x_i \tag{11}$$

$$\hat{\lambda}_4^\eta = \frac{1}{4 \binom{n}{\eta+1}} \sum_{i=1}^n \left[\binom{i-1}{\eta+3} - 3 \binom{i-1}{\eta+2} \binom{n-i}{1} + 3 \binom{i-1}{\eta+1} \binom{n-i}{2} - \binom{i-1}{\eta} \binom{n-i}{3} \right] x_i \tag{12}$$

The sample LH-moment ratios are as follows:

$$\hat{\tau}_2^\eta = \frac{\hat{\lambda}_2^\eta}{\hat{\lambda}_1^\eta} \tag{13}$$

$$\hat{\tau}_3^\eta = \frac{\hat{\lambda}_3^\eta}{\hat{\lambda}_2^\eta} \tag{14}$$

$$\hat{\tau}_4^\eta = \frac{\hat{\lambda}_4^\eta}{\hat{\lambda}_2^\eta} \tag{15}$$

2.2. Estimation of the Parameters of the Selected PDFs Based on LH-Moments

Several PDFs for fitting flood series exist in FFA; among them, the GLO, GEV, and GPA distributions have been recommended in Pakistan by various researchers to model extreme flood events [8,9,13,14,37–39]. Based on the comprehensive literature, three PDFs were selected in the current study, i.e., GLO, GEV, and GPA. The parameters of each respective PDF were estimated using the LH-moments and given below.

2.2.1. Generalized Logistics (GLO) Distribution

GLO distribution is a generalized variant of the logistic distribution [40,41] that has been applied in recent years to assess extreme value events. Since GLO was recognized as the acceptable approach for FFA in the UK [42], its use in hydrology has gained popularity [43]. The PDF, distribution function (DF), and quantile function (QF) of the GLO distribution are expressed, respectively, by:

$$f(x) = \frac{1}{\alpha} \left[1 - k \left(\frac{x - \zeta}{\alpha} \right) \right]^{\left(\frac{1}{k}-1\right)} \left[1 + \left\{ 1 - k \left(\frac{x - \zeta}{\alpha} \right) \right\}^{\left(\frac{1}{k}\right)} \right]^{-2} \tag{16}$$

$$F(x) = \left[1 + \left\{ 1 - k \left(\frac{x - \zeta}{\alpha} \right) \right\}^{\left(\frac{1}{k}\right)} \right]^{-1} \tag{17}$$

$$x(F) = \zeta + \frac{\alpha}{k} \left[1 - \left(\frac{1-F}{F} \right)^k \right] \tag{18}$$

where x is the APFDs, k is the shape parameter, ζ is a location, and α is a scale parameter. These parameters are estimated by [17] using LH-moments and given below:

$$\hat{\alpha} = \frac{\Gamma(\eta + 2)(\eta + 2)\hat{\beta}_{n+1} - (\eta + 1)\hat{\beta}_n}{\Gamma(\eta + 1 - \hat{k})\Gamma(1 + \hat{k})} \tag{19}$$

$$\hat{\xi} = (\eta + 1)\hat{\beta}_n - \frac{\hat{\alpha}}{\hat{k}} \left[1 - \frac{\Gamma(\eta + 1 - \hat{k})\Gamma(1 + \hat{k})}{\Gamma(\eta + 1)} \right] \tag{20}$$

$$\hat{k} = - \frac{(\eta + 3)(\eta + 2)\hat{\beta}_{n+2} - [(\eta + 2)^2 + (\eta + 2)(\eta + 1)]\hat{\beta}_{n+1} + (\eta + 1)^2\hat{\beta}_n}{(\eta + 2)\hat{\beta}_{n+1} - (\eta + 1)\hat{\beta}_n} \tag{21}$$

2.2.2. Generalized Extreme Value (GEV) Distribution

The GEV was developed by [44] as a feasible tool for extreme value analysis, and it has gained widespread favor in FFA. The GEV distribution PDF, DF, and QF are written, respectively, as:

$$f(x) = \frac{1}{\alpha} \left[1 - k \left(\frac{x - \xi}{\alpha} \right) \right]^{\left(\frac{1}{k}-1\right)} e^{-[1-k(\frac{x-\xi}{\alpha})]^{\left(\frac{1}{k}\right)}} \tag{22}$$

$$F(x) = \exp \left\{ - \left[1 - k \left(\frac{x - \xi}{\alpha} \right) \right]^{\frac{1}{k}} \right\} \tag{23}$$

$$x(F) = \xi + \frac{\alpha}{k} \left[1 - (-\ln F)^k \right] \tag{24}$$

where k, ξ, α are shape, location, and scale parameters, respectively, which are estimated by [11] and are described below:

$$\hat{\alpha} = \frac{\hat{k}[(\eta + 2)\hat{\beta}_{n+1} - (\eta + 1)\hat{\beta}_n]}{\Gamma(1 + \hat{k}) \left[(\eta + 1)^{-\hat{k}} - (\eta + 2)^{-\hat{k}} \right]} \tag{25}$$

$$\hat{\xi} = (\eta + 1)\hat{\beta}_n - \frac{\hat{\alpha}}{\hat{k}} \left[1 - (\eta + 1)^{-\hat{k}} \Gamma(1 + \hat{k}) \right] \tag{26}$$

$$\hat{k} = a_0 + a_1 \left[\tau_3^{\eta} \right] + a_2 \left[\tau_3^{\eta} \right]^2 + a_3 \left[\tau_3^{\eta} \right]^3 \tag{27}$$

2.2.3. Generalized Pareto (GPA) Distribution

Pickands [45] proposed GPA distribution, and several scholars have widely acknowledged it as the logical alternative for evaluating severe events [46,47]. The PDF, DF, and QF for GPA distribution are given, respectively, by:

$$f(x) = \frac{1}{\alpha} \left[1 - \frac{k}{\alpha} (x - \xi) \right]^{\left(\frac{1}{k}-1\right)} \tag{28}$$

$$F(x) = 1 - \left[1 - \frac{k}{\alpha} (x - \xi) \right]^{\frac{1}{k}} \tag{29}$$

$$x(F) = \xi + \frac{\alpha}{k} \left[1 - (1 - F)^k \right] \tag{30}$$

The location (α), scale (ξ), and shape (k) parameters were estimated by [17] and are stated below:

$$\hat{\alpha} = - \frac{\hat{k}\Gamma(\eta + 3 + \hat{k})\Gamma(\eta + 2 + \hat{k})(\eta + 2)\hat{\beta}_{\eta+1} - (\eta + 1)\hat{\beta}_\eta}{(\eta + 1)!\Gamma(1 + \hat{k}) \left[(\eta + 2)\Gamma(\eta + 2 + \hat{k}) - \Gamma(\eta + 3 + \hat{k}) \right]} \tag{31}$$

$$\hat{\xi} = (\eta + 1)\hat{\beta}_n - \frac{\hat{\alpha}}{\hat{k}} \left[1 - \frac{(\eta + 1)\Gamma(\eta + 1)\Gamma(1 + \hat{k})}{\Gamma(\eta + 2 + \hat{k})} \right] \quad (32)$$

$$\hat{k} = - \frac{-5 - 2\eta + \frac{(\eta + 3)(\eta + 3)\hat{\beta}_{\eta+2} - (\eta + 1)\hat{\beta}_\eta}{(\eta + 2)\hat{\beta}_{\eta+1} - (\eta + 1)\hat{\beta}_\eta}}{-1 + \frac{(\eta + 3)\hat{\beta}_{\eta+2} - (\eta + 1)\hat{\beta}_\eta}{(\eta + 2)\hat{\beta}_{\eta+1} - (\eta + 1)\hat{\beta}_\eta}} \quad (33)$$

2.3. Goodness-of-Fit (GOF) Tests

Four statistical measures, namely, the Anderson–Darling (AD) test, Kolmogorov–Smirnov (KS) test, Cramér–Von Mises (CVM) test, and LH-moment ratios (LH-ratios) diagram, were used in this study to determine the GOF tests for the selection of PDFs using LH-moments ($\eta = 0, 1, 2$). The PDFs for the APFDs that produced the smallest values for all of these GOF measures (AD test, KS test, and CVM test) were determined as a best-fit, and hence they were chosen for further estimation of quantiles. These GOF tests were previously applied to peak flow data and are frequently used to choose the best-fitting PDFs in FFA [8,9,35,48].

The AD test is used to evaluate the fit of an observed distribution function (DF) to its theoretical DF. The AD test gives a higher weight to the PDF's tail, which is a necessary feature in modeling extreme events [35,48]. Heo et al. [49] describe the AD test statistic as follows:

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i - 1) \log F(x_i) - \frac{1}{n} \sum_{i=1}^n (2n - 2i - 1) \log F(x_i) \quad (34)$$

Here, A^2 denotes the test result, n represents the sample size, x is the variable being studied, and $F(x_i)$ denotes the DF.

The KS test is based on the empirical DF and is used to assess whether a sample is drawn from a hypothesized continuous distribution. Assuming we have a random sample $(x_1, x_2, x_3, \dots, x_n)$ from some distribution, then empirical DF is as follows:

$$F_n(x) = \frac{1}{n} [\text{Number of observations} \leq x] \quad (35)$$

The maximum vertical distance between the theoretical PDF and empirical DF determines the KS statistic (D). The KS test statistic (D) is as follows:

$$D = \max_{1 \leq i \leq n} \left[F(x_i) - \frac{i-1}{n}, \frac{i}{n} - F(x_i) \right] \quad (36)$$

where x_i is the i th order statistic, n signifies the size of the random sample, and $F(x_i)$ indicates the theoretical DF.

The CVM test, an alternative to the KS test, is used to compare DF to a given empirical DF. Let $(x_{1:n} \leq x_{i:n} \leq x_{n:n})$ be the order statistics of a sample size n ; then the CVM test is suggested in [50]:

$$\omega^2 = \frac{1}{12n} + \sum_{i=1}^n \left[\frac{2i-1}{2n} - F(x_i) \right] \quad (37)$$

Hosking [10] initially proposed the L-moment ratio diagram as the simplest way to determine the best-fitted distribution for the actual data. The L-moment ratio diagram is extended to each level of LH-moments ($\eta = 0, 1, 2$) [11]. The LH-ratio diagram is based on the relationships between LH-skewness and LH-kurtosis. Therefore, this allows better discrimination between the PDFs, and hence the identification of parent distribution can also be achieved.

2.4. Quantile Estimates for Different Return Periods of Floods Based on LH-Moments

Various scientific fields are interested in estimating quantiles corresponding to different return periods. The return period is also known as a recurrence interval, defined as the average of inter-event times between flood events [4]. Sometimes, the hydrologist wants to know the chances of a flood reaching or exceeding a specific magnitude over a set time period. This is known as the probability of occurrence or the probability of exceedance. The probability that the exceedance for a given flood (q) with a return period (T) may be exceeded once in T years is computed as follows:

$$P(Q_T \leq q) = \frac{1}{T} \quad (38)$$

Equation (39) shows the cumulative probability of non-exceedance as follows:

$$F(Q_T) = P(Q_T \leq q) = 1 - P(Q_T > q) = 1 - \frac{1}{T} \quad (39)$$

Equation (39) is used to calculate the magnitude of a flood for given return periods. We can obtain quantile estimates for different return periods by substituting $[F(Q_T) = 1 - 1/T]$ in the quantile function of the GLO, GEV, and GPA distributions, as described below.

$$\text{GLO} \Rightarrow \hat{x}_T = \hat{\xi} + \frac{\hat{\alpha}}{\hat{k}} \left[1 - (T-1)^{-\hat{k}} \right] \quad (40)$$

$$\text{GEV} \Rightarrow \hat{x}_T = \hat{\xi} + \frac{\hat{\alpha}}{\hat{k}} \left[1 - \left\{ -\log \left(1 - \frac{1}{T} \right) \right\}^{\hat{k}} \right] \quad (41)$$

$$\text{GPA} \Rightarrow \hat{x}_T = \hat{\xi} + \frac{\hat{\alpha}}{\hat{k}} \left[1 - T^{-\hat{k}} \right] \quad (42)$$

Equations (40)–(42) above are used to calculate the quantile associated with the required return periods for the GLO, GEV, and GPA distributions at different levels of LH-moments ($\eta = 0, 1, 2$).

2.5. Quantile Estimates for Different Return Periods of Floods Based on Nonparametric Kernel Function

The nonparametric kernel function is based on kernel smoothing of the empirical QF of the variable under study [32]. Let $(x_1, x_2, x_3, \dots, x_n)$ be the series of observed APFDs arranged in ascending order; then the mathematical form of the kernel estimator of the nonparametric kernel function is expressed as [51]:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (43)$$

where $K(\cdot)$ refers to the kernel function prescribed type (Epanechnikov, Gaussian, Biweight, or Triweight), n denotes the observation's sample size, and h represents the bandwidth or smoothing parameter that controls the variance of the nonparametric kernel function [25,34,52]. The relation between DF and density of the nonparametric kernel function is as follows:

$$\hat{F}_h(x) = \int_{-\infty}^x \hat{f}_h(t) dt = \frac{1}{n} \sum_{i=1}^n H\left(\frac{x - x_i}{h}\right) \quad (44)$$

where

$$H(x) = \int_{-\infty}^x K(t) dt$$

Equation (44) is widely used in hydrology to calculate the quantiles associated with various return periods. The quantiles obtained by using a nonparametric kernel distribution estimator are as follows (for details, see [53,54]):

$$\hat{x}_T = \hat{F}_h^{-1}\left(1 - \frac{1}{T}\right) \quad (45)$$

Estimating the nonparametric kernel density method necessitates the selection of a kernel function, $K(\cdot)$, and the computation of a smoothing parameter, or bandwidth, h (as shown in Equation (44)). The choice of $K(\cdot)$ is less critical, and different types of kernel functions that provide good results can be used. This study applied the Epanechnikov, Gaussian, Biweight, and Triweight kernel functions, which are commonly used in the literature [24,25,53–57]. The expressions of standard kernel functions are given below:

$$\text{Epanechnikov} \Rightarrow k(x) = \begin{cases} \frac{3}{4}(1 - x^2); & \text{if } |x| \leq 1 \\ 0; & \text{otherwise} \end{cases} \quad (46)$$

$$\text{Gaussian} \Rightarrow k(x) = \frac{1}{\sqrt{2\pi}}e^{(-\frac{x^2}{2})}; \quad -\infty \leq x \leq \infty \quad (47)$$

$$\text{Biweight} \Rightarrow k(x) = \begin{cases} \frac{15}{16}(1 - x^2)^2; & \text{if } |x| \leq 1 \\ 0; & \text{otherwise} \end{cases} \quad (48)$$

$$\text{Triweight} \Rightarrow k(x) = \begin{cases} \frac{35}{32}(1 - x^2)^3; & \text{if } |x| \leq 1 \\ 0; & \text{otherwise} \end{cases} \quad (49)$$

In Equation (44), the smoothing parameter h plays a crucial role in the kernel estimator. In practice, selecting an effective technique for computing h for an observed data sample is a more complicated task due to the influence of the bandwidth on the shape of the associated estimate. If the value of h is small, we will obtain an undersmoothed estimator with a large variation. On the other hand, if h is large, the resultant estimator will be extremely smooth and will be farther away from the function that we are attempting to estimate [32,53]. In the context of the nonparametric kernel function, least-squares cross-validation, plug-in, and cross-validation procedures were considered for bandwidth selection. Overall, all the methods performed well both theoretically and practically; however, the least-squares cross-validation method needs a very large sample size to achieve satisfactory findings [58,59]. The two different plug-in bandwidth approaches were investigated by [58,60]. The cross-validation method was developed by [61], which showed promising results. Further, it was discovered that both cross-validation and plug-in bandwidths produced similar results in terms of DF estimation, but cross-validation had a clear disadvantage in terms of computation time [53,54,59]. The plug-in method, which has been employed previously in similar studies, yielded excellent results [54,60,62]. We used the plug-in approach suggested by [60] to determine the bandwidth for nonparametric kernel estimation of the DF of APFDs in this framework. The interested reader can obtain more theoretical information and a comprehensive explanation of the Polansky and Baker plug-in approach for bandwidth selection of the nonparametric DF; for details, see [59,60].

3. Study Area and Data

Climate change has significantly impacted the entire world. Its impacts may vary from increases in the magnitude and frequency of natural disasters, such as floods and droughts, to the extinction of species and the spread of vector-borne diseases. However, the effects of climate change are not equally observed across the globe. As a matter of fact, developing countries are much more vulnerable to climate change-related hazards, mainly due to their lack of proper infrastructure. For instance, Pakistan has suffered significant

economic losses in the past 6 to 7 years as a result of the recent increases on the melting rate of South Asia's glaciers, which leads to more frequent and severe floods.

Pakistan has a population of around 208 million people encompassing an area of approximately 796,000 km². The country is bounded by the Himalayan Mountains to the north, India to the east, Iran to the west, and the Indian Ocean to the south (Figure 1). Its altitudes vary from 8500 m in the northern regions to 0 m in the coastal regions, thus having a strong orographic influence over monsoon winds. Pakistan's climate is usually considered hot and dry, being classified as semi-arid in the south and dry cold in the north by means of Koppen climate classes [63].

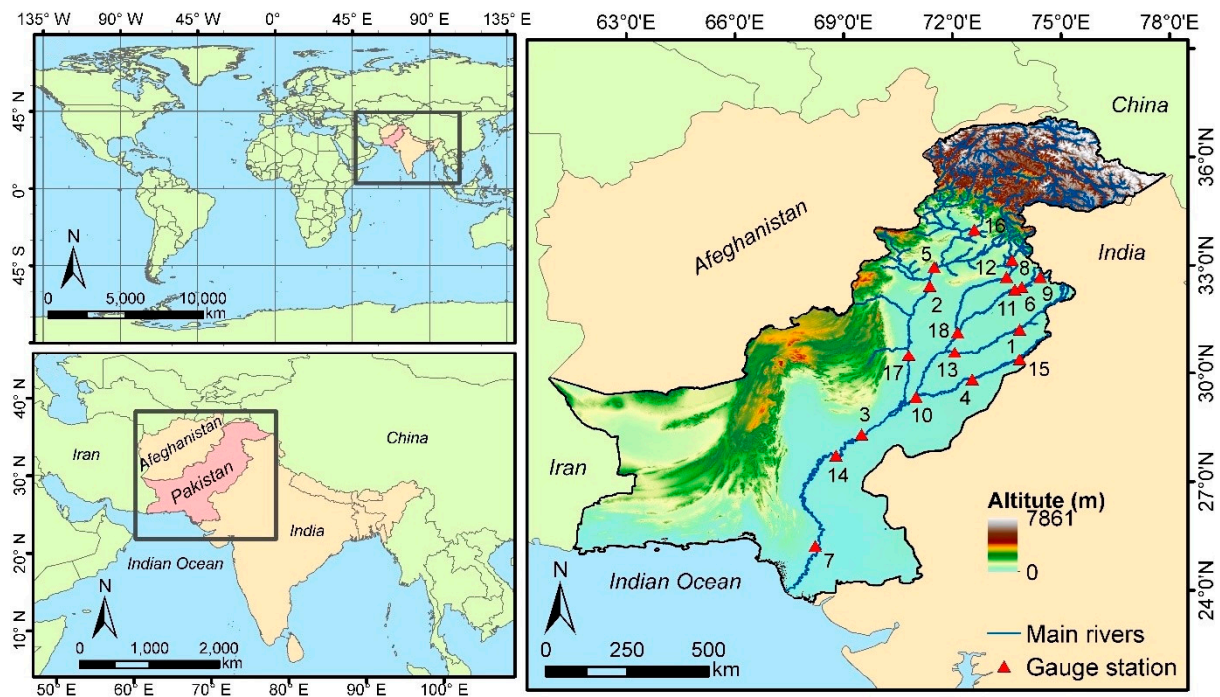


Figure 1. Geographical locations of the eighteen sites of Punjab, Pakistan, used in this study.

The Indus River and its tributaries, i.e., Sutlej, Beas, Ravi, Chenab, Jhelum, Swat, and Kabul, are vital to the economy of Pakistan, as they are the main source of water for irrigation, industry, and urban water supply. However, this river network is also responsible for economic losses through large flood events, most frequent in the Punjab and Sindh regions, affecting not only fertile agricultural lands, but also large urban centers near the river network. This situation is aggravated when considering future climate change scenarios, as higher variabilities in precipitation and glacier melting are projected, as well as rises in sea level and storm surges, leading to stress of current drainage network systems, especially during the monsoon season.

In this study, the annual maximum peak flows data of 18 sites of Pakistan, located on five rivers, namely, Indus, Jhelum, Chenab, Ravi, and Sutlej, were used in this study. The geographical location of these river sites is given in Figure 1. The data for these sites were collected from the hydrology department of the Water and Power Development Authority (WAPDA) and the Federal Flood Commission. Most of the annual maximum peak flows at sites were recorded in the peak of the monsoon season (from July to September). The annual peak flow discharge of 18 stations was measured in cusec. Summary statistics for 18 stations are given in Table 1. The highest mean peak flow discharge was recorded at the Guddu station (609,909.423), and the lowest mean peak flow discharge was observed at Islam station (49,089.45). The highest standard deviations were observed at Kotri, Sukkur, Guddu, Qadirabad, Khanki, and Trimmu stations. The range of skewness varied from 0.552 to 4.240, while Guddu had the smallest skewness, and Mangla had the largest skewness. Similarly, the kurtosis varied from -0.645 to 22.770, with Mangla, Kotri, Rasul, and Tarbela

having the highest Kurtosis, while the lowest kurtosis was observed at Sukkur, Guddu, Qadirabad, and Marala stations.

Table 1. Summary statistics of 18 stations.

Name of Stations	River	Latitude (North)	Longitude (East)	Period (Years)	Mean	Standard Deviation	Skewness	Kurtosis	Minimum Peak Flow	Maximum Peak Flow
Tarbela	Indus	33.99	72.61	1960–2013	386,962.963	87,785.537	2.626	11.806	273,000	835,000
Kalabagh	Indus	32.95	71.50	1968–2013	464,719.956	151,843.363	1.186	2.102	237,297	936,453
Chashma	Indus	32.43	71.38	1971–2013	475,333.046	149,635.274	1.22	3.727	214,045	1,038,873
Taunsa	Indus	30.50	70.80	1958–2013	452,791.554	140,793.102	0.804	2.144	182,372	959,991
Guddu	Indus	28.30	69.50	1962–2013	609,909.423	284,534.413	0.552	−0.557	170,831	1,176,150
Sukkur	Indus	27.72	68.79	1982–2013	546,609.594	309,470.519	0.629	−0.645	126,130	1,172,000
Kotri	Indus	25.22	68.22	1970–2013	395,262.068	379,599.333	3.705	18.290	47,100	2,409,000
Mangla	Jhelum	33.15	73.65	1960–2013	132,481.778	136,385.297	4.240	22.770	20,460	932,700
Rasul	Jhelum	32.68	73.50	1970–2013	134,418.386	161,219.596	3.582	15.787	19,702	952,170
Marala	Chenab	32.68	74.43	1960–2013	308,572.407	196,419.272	1.097	0.227	93,150	792,765
Khanki	Chenab	32.40	73.92	1925–2013	351,963.191	242,710.633	1.494	1.391	97,058	1,086,460
Qadirabad	Chenab	32.33	73.73	1970–2013	356,547.704	247,771.998	1.030	0.106	76,336	948,520
Trimmu	Chenab	31.14	72.15	1968–2013	261,376.217	194,828.961	1.099	0.1693	42,756	706,433
Panjinad	Chenab	29.33	71.00	1960–2013	260,134.722	193,661.339	0.980	0.554	17,833	802,516
Balloki	Ravi	31.22	73.86	1922–2013	87,914.728	64,039.572	2.183	6.180	14,000	399,356
Sidhani	Ravi	30.58	72.07	1925–2013	64,143.427	56,691.878	2.159	4.916	8488	296,086
Sulemanki	Sutlej	30.38	73.86	1975–2013	70,254.923	84,914.177	2.267	5.865	1506	399,453
Islam	Sutlej	29.82	72.55	1974–2013	49,089.45	63,209.754	2.362	6.497	1231	306,425

4. Results and Discussion

The AD test, KS test, and CVM test were applied to each station in order to choose the best-fit PDF among the GEV, GLO, and GPA. The selection of the best fitted PDF for each station was based on the smallest GOF tests among the three PDFs. The best fitted PDF results for each station according to GOF tests at a 5% significance level are reported in Table 2.

Table 2. GOF results for GEV, GLO, and GPA distributions using LH-moments ($\eta = 0, 1, 2$).

Stations	L-Moments ($\eta = 0$)			L1-Moments ($\eta = 1$)			L2-Moments ($\eta = 2$)		
	AD Test	KS Test	CVM Test	AD Test	KS Test	CVM Test	AD Test	KS Test	CVM Test
Tarbela	GEV(0.984)	GEV(0.975)	GEV(0.973)	GEV(0.395)	GEV(0.605)	GEV(0.670)	GEV(0.233)	GEV(0.332)	GLO(0.360)
Kalabagh	GLO(0.954)	GLO(0.938)	GLO(0.972)	GLO(0.855)	GLO(0.785)	GLO(0.876)	GLO(0.642)	GLO(0.682)	GLO(0.696)
Chashma	GLO(0.983)	GLO(0.965)	GLO(0.970)	GLO(0.943)	GLO(0.895)	GLO(0.951)	GLO(0.801)	GLO(0.836)	GLO(0.853)
Taunsa	GLO(0.811)	GEV(0.537)	GLO(0.705)	GLO(0.832)	GEV(0.606)	GLO(0.761)	GLO(0.762)	GLO(0.627)	GLO(0.753)
Guddu	GEV(0.740)	GLO(0.878)	GEV(0.769)	GEV(0.765)	GLO(0.889)	GEV(0.781)	GEV(0.742)	GEV(0.923)	GEV(0.790)
Sukkur	GPA(0.990)	GPA(0.978)	GPA(0.992)	GPA(0.944)	GPA(0.962)	GPA(0.960)	GPA(0.963)	GPA(0.988)	GPA(0.971)
Kotri	GEV(0.974)	GEV(0.837)	GEV(0.924)	GEV(0.947)	GEV(0.821)	GEV(0.916)	GEV(0.859)	GEV(0.831)	GLO(0.887)
Mangla	GLO(0.956)	GLO(0.900)	GLO(0.932)	GEV(0.803)	GEV(0.864)	GEV(0.916)	GEV(0.537)	GLO(0.851)	GLO(0.877)
Rasul	GEV(0.946)	GEV(0.984)	GLO(0.939)	GEV(0.962)	GEV(0.988)	GEV(0.950)	GPA(0.928)	GEV(0.991)	GPA(0.943)
Marala	GPA(0.969)	GPA(0.973)	GPA(0.974)	GPA(0.758)	GPA(0.823)	GPA(0.880)	GPA(0.735)	GPA(0.875)	GPA(0.787)
Khanki	GEV(0.693)	GPA(0.868)	GEV(0.744)	GEV(0.612)	GEV(0.713)	GEV(0.712)	GEV(0.465)	GEV(0.741)	GEV(0.655)
Qadirabad	GPA(0.995)	GPA(0.996)	GPA(0.999)	GEV(0.930)	GPA(0.983)	GPA(0.988)	GPA(0.943)	GPA(0.985)	GPA(0.968)
Trimmu	GPA(0.779)	GPA(0.778)	GPA(0.726)	GEV(0.683)	GPA(0.679)	GPA(0.699)	GEV(0.698)	GPA(0.622)	GPA(0.648)
Panjinad	GPA(0.908)	GEV(0.879)	GEV(0.878)	GPA(0.914)	GPA(0.885)	GPA(0.894)	GPA(0.933)	GPA(0.897)	GPA(0.906)
Balloki	GEV(0.582)	GEV(0.624)	GEV(0.551)	GEV(0.486)	GEV(0.621)	GEV(0.517)	GEV(0.325)	GEV(0.621)	GEV(0.480)
Sidhani	GEV(0.978)	GEV(0.990)	GEV(0.974)	GEV(0.971)	GEV(0.992)	GEV(0.969)	GEV(0.933)	GEV(0.982)	GEV(0.957)
Sulemanki	GPA(0.996)	GPA(0.994)	GPA(0.998)	GPA(0.998)	GPA(0.991)	GPA(0.998)	GPA(0.999)	GPA(0.990)	GPA(0.998)
Islam	GPA(0.900)	GPA(0.753)	GPA(0.877)	GPA(0.931)	GPA(0.693)	GPA(0.885)	GPA(0.936)	GPA(0.715)	GPA(0.886)

In the case of $\eta = 0$, it is clear from Table 2 that GPA and GEV distributions were best fitted for seven stations, while GLO was best fitted for four stations according to the AD test. Similarly, the KS test results well matching the results of the AD test except for Taunsa, Guddu, Khanki, and Panjinad stations. According to the KS test results for $\eta = 0$, the most

appropriate PDFs were GPA and GEV. Investigating the CVM test results for $\eta = 0$, the GEV distribution was suitable for most stations, followed by GPA distribution, which was selected for six stations. Overall, in the case of $\eta = 0$, the results of the AD test, KS test, and CVM test indicated that the GEV and GPA distributions were the most adequate for most stations.

Moving forward to $\eta = 1$ in Table 2, the AD test selected the GEV distribution for the highest number of stations (10 stations out of 18) followed by GPA distribution, which was selected for five stations. However, the KS test and CVM test selected the GEV and GPA distributions for the same number of stations, 8 and 7, respectively. Considering AD test, KS test, and CVM test results for $\eta = 1$, the GEV distribution was selected for eight stations, the GPA distribution for seven stations, and the GLO distribution for only three stations. Finally, for $\eta = 2$, it is observed from Table 2 that the AD test and CVM test selected GPA distribution for eight stations, GEV distribution for seven stations, and GLO distribution for the remaining four stations. On the other hand, GPA and GEV distributions each yielded the best-fit for seven stations based on the KS test.

It is also worthwhile to mention that the GOF test produced different results for Mangla, Rasul, and Panjnad stations as we increased the value $\eta = 1, 2$. However, in general, increasing $\eta = 1, 2$ had no effect on the results for most stations.

The LH-ratio diagram is a useful tool that simplifies analysis, demonstrating the versatility of how various PDFs plot. Furthermore, it can be shown that PDFs can have several different skewness and kurtosis values, rendering them more valuable for analyzing the shape of the distribution. The LH-ratio diagram for $\eta = 0, 1$, and $\eta = 2$ of 18 stations was plotted in Figure 2. It is observed from Figure 2, for $\eta = 0$, that most of the scatter points were between the GEV and GPA distributions curves, whereas a few scattered points were closed to the GLO distribution curve. Therefore, according to the L-ratio diagram ($\eta = 0$), GEV and GPA were the most suitable PDFs for the annual peak flow series of 18 stations. By observing Figure 2 in the case of $\eta = 1$ and $\eta = 2$, most of scattered points were closed to the GPA distribution curve, followed by GEV. We also note from Figure 2 that as we increased $\eta = 1, 2$, peak flow series tended to follow GPA and GEV distributions. Overall, the findings obtained from GOF tests were generally in good agreement with the LH-ratio diagram for most stations.

Further the relation between the return period and APFD was also established for GEV, GLO, and GPA distributions. Figure 3 shows the curves for Balloki, Taunsa, and Islam stations, highlighting how well the APFD series at lower return periods and upper return periods were estimated by LH-moments ($\eta = 0, 1, 2$). It is seen from Figure 3 that the GEV, GPA, and GLO distributions well fitted the observed APFD series at lower and higher return periods. Figure 3 indicates that as the level of LH-moments increased ($\eta = 1, 2$), the GEV, GLO, and GPA distributions performed well in reflecting the extreme tail at higher return periods. In Figure 3, it is noted that most of the observations fell within 2–50 years ($0.02 \leq p \leq 0.5$), implying that hazardous flood events with low probability or large return periods (100 and 500) have rarely occurred at these stations.

In the planning and design of hydrological systems, it is critical to determine the return period since a given flood event. Further, we calculated quantiles for different return periods using LH-moments ($\eta = 0, 1, 2$) and nonparametric kernel functions; results for Tarbela, Kalabagh, Qadirabad, and Trimmu stations are reported in Figure 4. The results of quantile estimates in Figure 4 can be interpreted as follows: for Tarbela station's return period of 500 years, the GEV distribution ($\eta = 0$) produced quantile (847,553.3) is the threshold value of flow that may occur once every 500 years on average. In other words, there is only a 0.2% chance that in a return period of 500 years, one-time discharge (peak flow) will exceed the threshold value (847,553.3) and consequently a flood will occur. At the same time, 99.9% is the chance that the one-time discharge (peak flow) will be less than the threshold value (847,553.3) in a return period of 500 years.

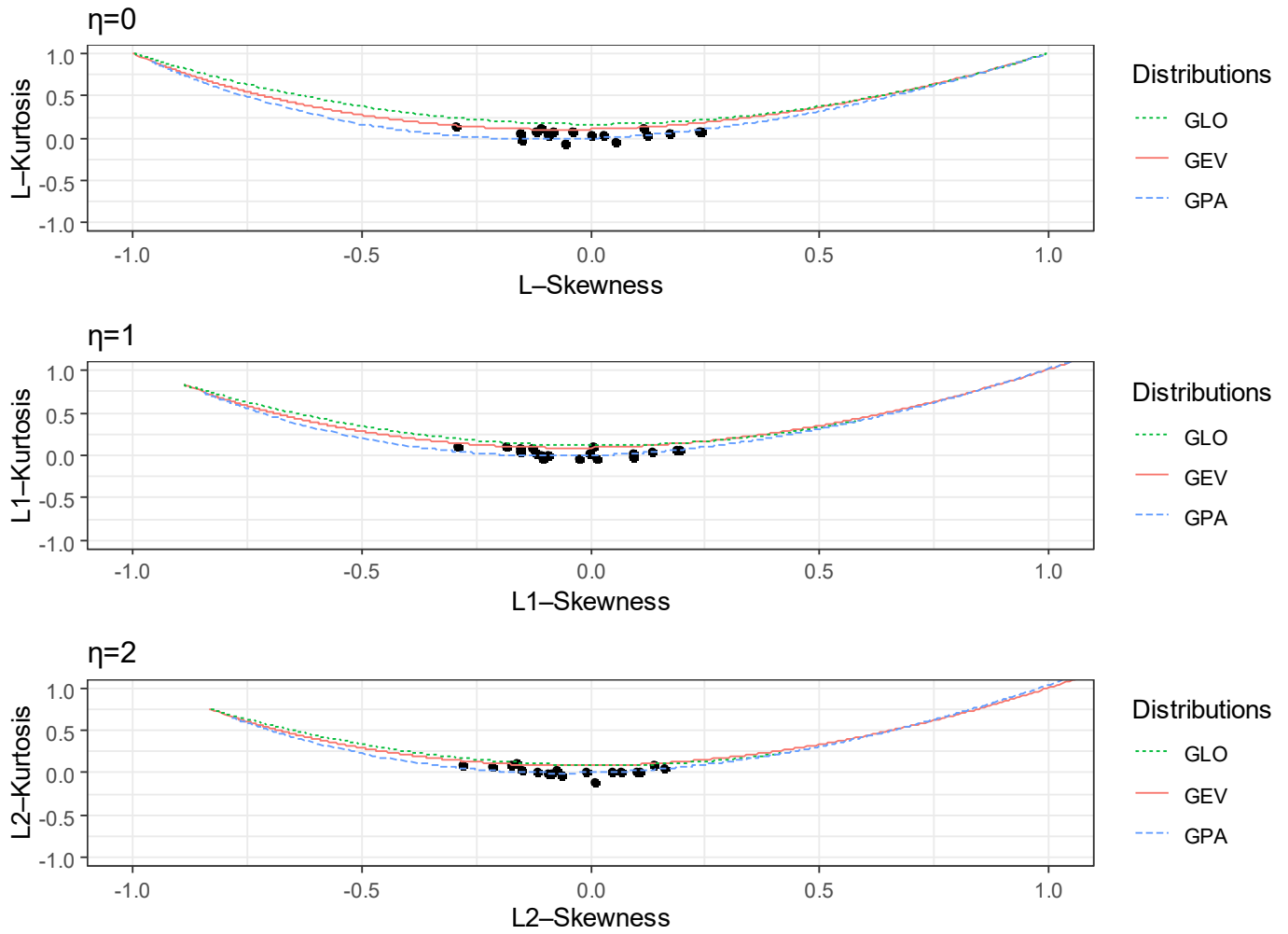


Figure 2. LH-ratio diagram for 18 stations.

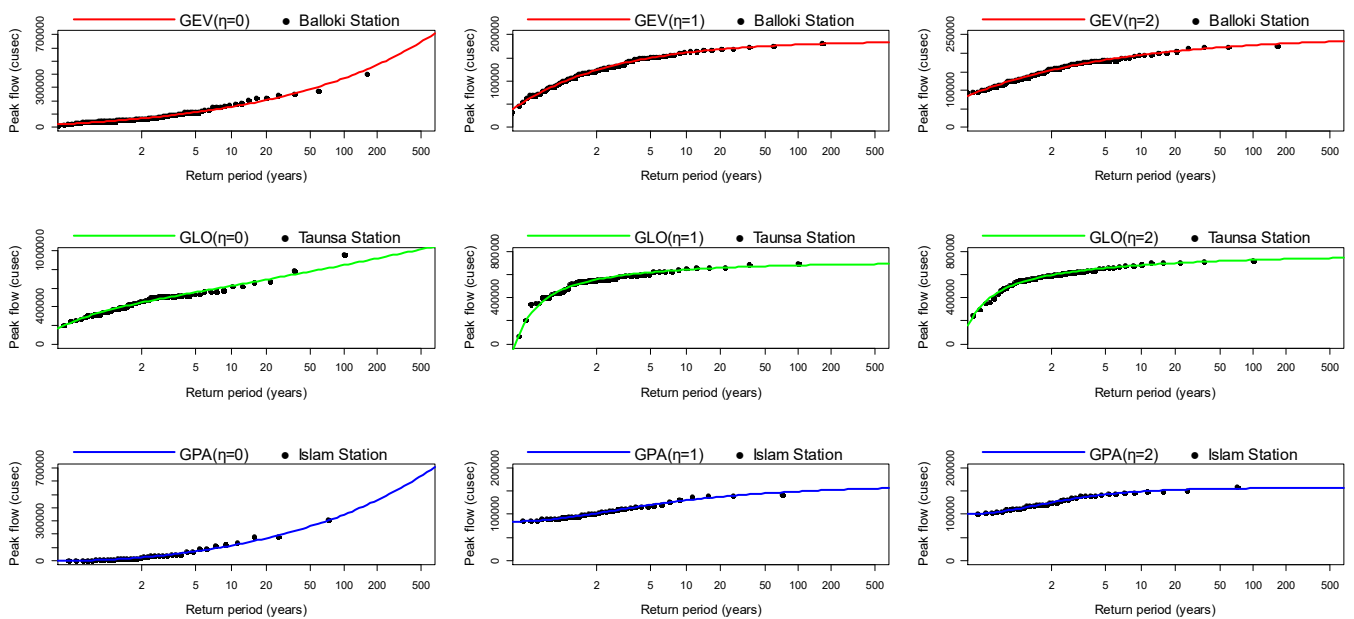


Figure 3. Relation between return period and annual peak flow based on the LH-moments.

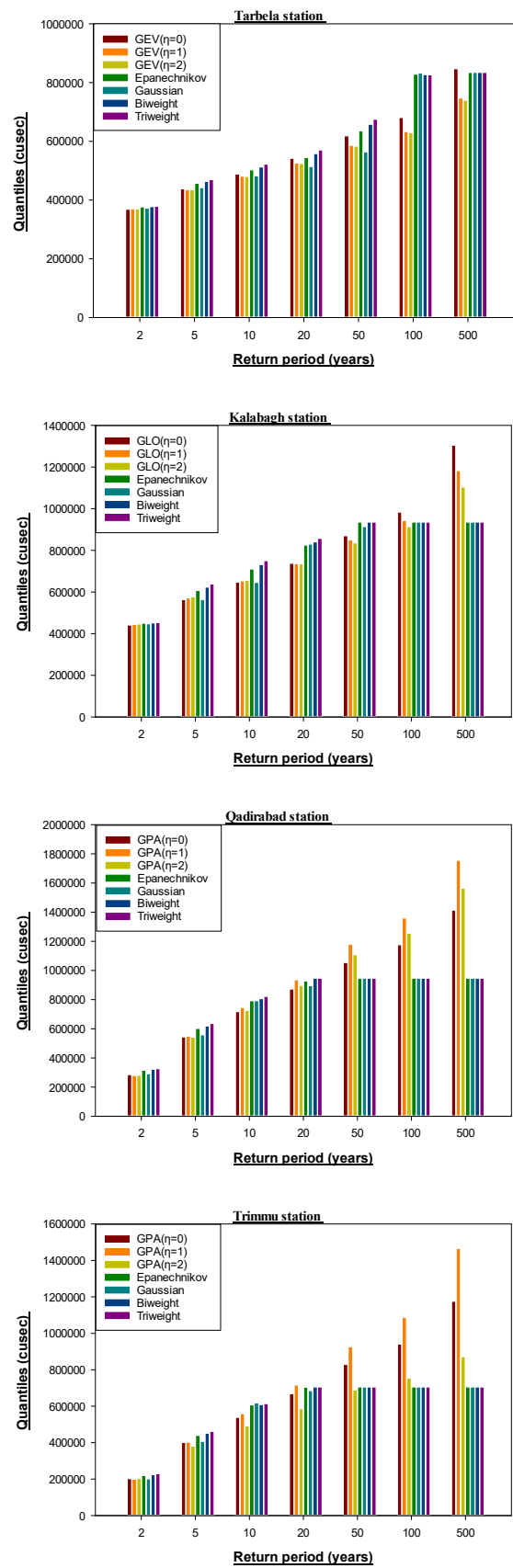


Figure 4. Quantiles for best-fit PDFs and nonparametric kernel functions for four randomly selected stations.

We also investigated the impact of LH-moment choice ($\eta = 0, 1, 2$) and nonparametric kernel functions on estimating quantiles associated with predefined return periods via relative absolute error (RAE). The RAE is an assessment of the difference between the actual flood estimate and the flood estimate by the best-fit PDF. The RAE was calculated using the following equation, described in [35,48,64].

$$RAE = \left| \frac{X - Y}{Y} \right| \tag{50}$$

where X is the actual peak flow, and Y denotes the design quantile estimate obtained through the LH-moments and the nonparametric kernel function.

Tables 3 and 4 summarize the RAE associated with each station using LH-moments and a nonparametric kernel function, emphasizing the significance of evaluating techniques for diverse return periods. It is crucial to investigate the discrepancy between actual flood estimates and quantiles produced via PDF and the nonparametric kernel function. Although all these PDFs passed GOF testing, there were still considerable discrepancies in quantile estimations. These discrepancies are significant to policymakers, planners, and decision-makers.

Table 3. RAE of quantile estimates for GEV, GLO, and GPA distributions using LH-moments ($\eta = 0, 1, 2$).

Station Name	Best Fitted Distribution	2	5	10	20	50	100	500
Tarbela	($\eta = 0$)	0.008	0.009	0.015	0.028	0.053	0.077	0.156
	GEV ($\eta = 1$)	0.004	0.005	0.008	0.011	0.017	0.022	0.035
	($\eta = 2$)	0.003	0.004	0.006	0.008	0.011	0.014	0.02
Kalabagh	($\eta = 0$)	0.011	0.012	0.023	0.041	0.072	0.103	0.202
	GLO ($\eta = 1$)	0.009	0.012	0.016	0.021	0.03	0.039	0.063
	($\eta = 2$)	0.009	0.013	0.015	0.018	0.024	0.029	0.045
Chashma	($\eta = 0$)	0.01	0.012	0.022	0.036	0.061	0.085	0.156
	GLO ($\eta = 1$)	0.01	0.014	0.017	0.022	0.03	0.037	0.059
	($\eta = 2$)	0.011	0.014	0.016	0.019	0.025	0.031	0.048
Taunsa	($\eta = 0$)	0.009	0.012	0.019	0.029	0.046	0.061	0.106
	GLO ($\eta = 1$)	0.01	0.014	0.016	0.02	0.027	0.033	0.05
	($\eta = 2$)	0.01	0.013	0.015	0.019	0.025	0.03	0.046
Guddu	($\eta = 0$)	0.016	0.016	0.025	0.041	0.065	0.087	0.144
	GEV ($\eta = 1$)	0.013	0.015	0.022	0.032	0.047	0.06	0.09
	($\eta = 2$)	0.012	0.015	0.02	0.028	0.039	0.047	0.067
Sukkur	($\eta = 0$)	0.023	0.026	0.032	0.053	0.086	0.112	0.172
	GPA ($\eta = 1$)	0.021	0.024	0.03	0.048	0.076	0.096	0.14
	($\eta = 2$)	0.018	0.021	0.027	0.043	0.066	0.082	0.114
Kotri	($\eta = 0$)	0.03	0.041	0.043	0.081	0.159	0.235	0.483
	GEV ($\eta = 1$)	0.019	0.021	0.03	0.047	0.075	0.099	0.163
	($\eta = 2$)	0.017	0.021	0.028	0.039	0.054	0.067	0.096
Mangla	GLO ($\eta = 0$)	0.042	0.046	0.073	0.079	0.159	0.235	0.47
	GEV ($\eta = 1$)	0.016	0.018	0.026	0.042	0.068	0.09	0.148
	GLO ($\eta = 2$)	0.016	0.018	0.026	0.037	0.056	0.072	0.119
Rasul	GEV ($\eta = 0$)	0.056	0.067	0.089	0.096	0.174	0.256	0.505
	GEV ($\eta = 1$)	0.022	0.034	0.036	0.065	0.113	0.157	0.285
	GPA ($\eta = 2$)	0.018	0.022	0.027	0.044	0.069	0.087	0.124
Marala	($\eta = 0$)	0.022	0.024	0.031	0.052	0.09	0.124	0.214
	GPA ($\eta = 1$)	0.017	0.018	0.025	0.041	0.069	0.092	0.151
	($\eta = 2$)	0.013	0.014	0.019	0.031	0.049	0.063	0.094

Table 3. Cont.

Station Name	Best Fitted Distribution	2	5	10	20	50	100	500
Khanki	($\eta = 0$)	0.021	0.026	0.036	0.056	0.112	0.166	0.337
	GEV ($\eta = 1$)	0.012	0.017	0.021	0.04	0.073	0.103	0.189
	($\eta = 2$)	0.009	0.01	0.016	0.025	0.04	0.053	0.087
Qadirabad	($\eta = 0$)	0.025	0.029	0.034	0.057	0.098	0.133	0.226
	GPA ($\eta = 1$)	0.022	0.023	0.029	0.048	0.079	0.105	0.166
	($\eta = 2$)	0.018	0.019	0.025	0.041	0.065	0.082	0.119
Trimmu	($\eta = 0$)	0.027	0.032	0.035	0.06	0.105	0.145	0.253
	GPA ($\eta = 1$)	0.022	0.024	0.03	0.05	0.083	0.109	0.175
	($\eta = 2$)	0.017	0.02	0.025	0.042	0.064	0.079	0.109
Panjnad	GEV ($\eta = 0$)	0.02	0.031	0.035	0.058	0.099	0.135	0.235
	GPA ($\eta = 1$)	0.018	0.026	0.03	0.048	0.071	0.086	0.112
	GPA ($\eta = 2$)	0.015	0.022	0.026	0.044	0.068	0.079	0.096
Balloki	($\eta = 0$)	0.023	0.026	0.038	0.056	0.114	0.17	0.343
	GEV ($\eta = 1$)	0.011	0.014	0.02	0.035	0.061	0.084	0.149
	($\eta = 2$)	0.008	0.01	0.014	0.021	0.032	0.04	0.06
Sidhani	($\eta = 0$)	0.028	0.028	0.047	0.06	0.123	0.182	0.362
	GEV ($\eta = 1$)	0.013	0.02	0.023	0.042	0.073	0.099	0.173
	($\eta = 2$)	0.012	0.012	0.019	0.03	0.046	0.059	0.093
Sulemanki	($\eta = 0$)	0.053	0.059	0.085	0.09	0.173	0.253	0.512
	GPA ($\eta = 1$)	0.037	0.042	0.054	0.076	0.138	0.193	0.355
	($\eta = 2$)	0.029	0.039	0.044	0.069	0.118	0.158	0.263
Islam	($\eta = 0$)	0.058	0.068	0.091	0.096	0.175	0.256	0.518
	GPA ($\eta = 1$)	0.042	0.044	0.063	0.08	0.149	0.211	0.397
	($\eta = 2$)	0.031	0.039	0.046	0.07	0.121	0.164	0.278

Table 4. RAE of quantile estimates for Epanechnikov, Gaussian, Biweight, and Triweight kernel functions.

Station Name	Kernel Function Type	2	5	10	20	50	100	500
Tarbela	Epanechnikov	0.027	0.033	0.046	0.064	0.182	0.346	0.728
	Gaussian	0.014	0.02	0.027	0.049	0.079	0.12	0.24
	Biweight	0.029	0.047	0.064	0.087	0.211	0.39	0.74
	Triweight	0.03	0.06	0.081	0.107	0.23	0.31	0.5
Kalabagh	Epanechnikov	0.01	0.024	0.101	0.124	0.2	0.23	0.33
	Gaussian	0.004	0.01	0.018	0.023	0.032	0.07	0.125
	Biweight	0.006	0.013	0.125	0.127	0.14	0.19	0.21
	Triweight	0.014	0.016	0.145	0.149	0.17	0.198	0.24
Chashma	Epanechnikov	0.004	0.078	0.081	0.12	0.183	0.263	0.58
	Gaussian	0.005	0.01	0.024	0.068	0.088	0.2	0.534
	Biweight	0.003	0.055	0.127	0.152	0.214	0.434	0.63
	Triweight	0.002	0.029	0.128	0.178	0.239	0.488	0.678
Taunsa	Epanechnikov	0.019	0.089	0.097	0.101	0.103	0.121	0.2
	Gaussian	0.014	0.017	0.02	0.025	0.046	0.067	0.167
	Biweight	0.019	0.115	0.116	0.122	0.129	0.222	0.29
	Triweight	0.019	0.134	0.139	0.14	0.153	0.267	0.32
Guddu	Epanechnikov	0.013	0.014	0.023	0.091	0.182	0.311	0.671
	Gaussian	0.003	0.005	0.017	0.042	0.11	0.224	0.422
	Biweight	0.02	0.023	0.033	0.11	0.196	0.375	0.76
	Triweight	0.025	0.023	0.054	0.127	0.215	0.46	0.845
Sukkur	Epanechnikov	0.032	0.035	0.069	0.139	0.216	0.297	0.532
	Gaussian	0.027	0.03	0.035	0.06	0.1	0.19	0.383
	Biweight	0.035	0.041	0.089	0.171	0.342	0.441	0.72
	Triweight	0.035	0.046	0.089	0.199	0.438	0.564	0.783

Table 4. Cont.

Station Name	Kernel Function Type	2	5	10	20	50	100	500
Kotri	Epanechnikov	0.095	0.131	0.162	0.191	0.257	0.501	0.732
	Gaussian	0.04	0.055	0.083	0.15	0.2	0.295	0.527
	Biweight	0.05	0.11	0.13	0.158	0.222	0.45	0.69
	Triweight	0.073	0.124	0.15	0.181	0.245	0.489	0.705
Mangla	Epanechnikov	0.107	0.127	0.132	0.154	0.231	0.476	0.845
	Gaussian	0.051	0.068	0.099	0.134	0.198	0.345	0.695
	Biweight	0.12	0.153	0.183	0.203	0.282	0.523	0.912
	Triweight	0.135	0.17	0.185	0.212	0.292	0.545	0.989
Rasul	Epanechnikov	0.069	0.092	0.105	0.15	0.315	0.605	1.21
	Gaussian	0.06	0.09	0.099	0.13	0.265	0.55	0.999
	Biweight	0.083	0.101	0.163	0.193	0.386	0.71	1.421
	Triweight	0.085	0.105	0.183	0.213	0.412	0.8	1.89
Marala	Epanechnikov	0.03	0.045	0.055	0.08	0.12	0.223	0.525
	Gaussian	0.028	0.04	0.053	0.069	0.1	0.193	0.412
	Biweight	0.055	0.075	0.097	0.13	0.274	0.498	0.875
	Triweight	0.067	0.091	0.104	0.198	0.32	0.53	0.995
Khanki	Epanechnikov	0.081	0.09	0.124	0.175	0.243	0.475	0.822
	Gaussian	0.043	0.075	0.106	0.135	0.203	0.422	0.79
	Biweight	0.099	0.109	0.141	0.19	0.275	0.49	0.918
	Triweight	0.103	0.116	0.142	0.203	0.303	0.503	1.116
Qadirabad	Epanechnikov	0.046	0.061	0.076	0.122	0.17	0.328	0.631
	Gaussian	0.029	0.052	0.076	0.105	0.152	0.298	0.608
	Biweight	0.058	0.073	0.079	0.139	0.185	0.347	0.675
	Triweight	0.063	0.079	0.096	0.151	0.196	0.365	0.692
Trimmu	Epanechnikov	0.038	0.083	0.108	0.244	0.331	0.644	1.976
	Gaussian	0.025	0.058	0.101	0.175	0.305	0.563	1.107
	Biweight	0.033	0.063	0.106	0.261	0.36	0.682	2.19
	Triweight	0.033	0.073	0.125	0.273	0.36	0.705	2.806
Panjnad	Epanechnikov	0.034	0.073	0.095	0.109	0.136	0.275	0.595
	Gaussian	0.047	0.057	0.083	0.105	0.126	0.234	0.498
	Biweight	0.038	0.073	0.117	0.128	0.143	0.283	0.607
	Triweight	0.064	0.073	0.126	0.156	0.17	0.303	0.67
Balloki	Epanechnikov	0.031	0.047	0.077	0.119	0.177	0.219	0.445
	Gaussian	0.04	0.049	0.064	0.108	0.133	0.212	0.414
	Biweight	0.028	0.043	0.092	0.124	0.192	0.324	0.59
	Triweight	0.029	0.046	0.105	0.135	0.205	0.335	0.67
Sidhani	Epanechnikov	0.068	0.081	0.095	0.155	0.218	0.402	0.851
	Gaussian	0.041	0.051	0.075	0.131	0.197	0.359	0.738
	Biweight	0.085	0.096	0.105	0.174	0.29	0.507	0.907
	Triweight	0.098	0.104	0.118	0.192	0.317	0.541	0.942
Sulemanki	Epanechnikov	0.068	0.112	0.146	0.182	0.268	0.573	1.165
	Gaussian	0.063	0.078	0.112	0.148	0.233	0.438	0.91
	Biweight	0.071	0.087	0.137	0.185	0.271	0.518	1.154
	Triweight	0.071	0.082	0.132	0.155	0.251	0.502	1.123
Islam	Epanechnikov	0.099	0.145	0.191	0.245	0.399	0.745	1.168
	Gaussian	0.067	0.11	0.154	0.21	0.367	0.61	0.929
	Biweight	0.109	0.168	0.217	0.268	0.409	0.778	1.481
	Triweight	0.118	0.19	0.26	0.329	0.418	0.819	1.921

As shown in Table 3, the GLO, GEV, and GPA distributions produced very small errors for all stations. In Table 3, the results of GEV distribution for Tarbela, Guddu, Kotri, Khanki, Balloki, and Sidhani stations based on L2-moment ($\eta = 2$) produced a small error

for all return periods. However, it was noticed that for return periods of 5 years, the GEV distribution provided the same result for Kotri and Guddu stations based on L1-moment ($\eta = 1$) and L2-moment ($\eta = 2$).

According to Table 3, the RAE findings for return periods (2, 5, and 10 years) for GLO distribution based on L-moment ($\eta = 0$), L1-moment ($\eta = 1$) and L2-moment ($\eta = 2$) for Kalabagh, Chashma, and Taunsa stations had a fairly close error, whereas for the remaining return periods, L2-moment ($\eta = 2$) had a clear edge over the L-moment ($\eta = 0$) and L1-moment ($\eta = 1$). It was also observed, as seen in Table 3, that the GPA distribution using L2-moment ($\eta = 2$) produced a minimal error for all return periods for Sukkur, Marala, Qadirabad, Sulemanki, and Islam stations.

Table 3 indicates that for Mangla station, the RAE for GLO and GEV distributions were estimated through L-moment ($\eta = 0$), L2-moment ($\eta = 2$) and L1-moment ($\eta = 1$), respectively; however, the GLO and GEV distributions produced the same amount of error for the return periods (2, 5 and 10 years) when using the L2-moment ($\eta = 2$) and L1-moment ($\eta = 1$), but the L2-moment ($\eta = 2$) had the edge for return the periods of 20, 50, 100, and 500 years. Similarly, the findings for Rasul and Panjnad stations indicated that GPA distribution using L2-moment ($\eta = 2$) had lower values of RAE at all return periods than did GEV distribution when using L-moment ($\eta = 0$) and L1-moment ($\eta = 1$).

Moreover, as can be seen in Table 3, the findings of RAE for all stations using L-moment ($\eta = 0$), L1-moment ($\eta = 1$), and L2-moment ($\eta = 2$) at the return periods of (2, 5, 10, 20 years) were close, but L1-moment ($\eta = 1$) and L2-moment ($\eta = 2$) had a slight advantage over L-moment ($\eta = 0$). However, for high return periods (50, 100, 500 years), the L2-moment ($\eta = 2$) performed better than the L1-moment ($\eta = 1$) and L-moment ($\eta = 0$). We also observed, as seen in Table 3, that with the increasing level of LH-moments ($\eta = 0, 1, 2$), the error became smaller, especially for high return periods. In other words, the L2-moment ($\eta = 2$) yielded the lowest error compared to the L1-moment ($\eta = 1$) and L-moment ($\eta = 0$). Additionally, it was found that there were a few overlaps among these PDFs for certain return periods. For example, for the Mangla station at return periods of 2, 5, and 10 years, the GEV and GLO distributions yielded the same amount of error (0.016, 0.018, and 0.026). This implies that the performance of the PDFs was the same for a certain return period.

Table 4 compares quantile estimates in terms of RAE for Epanechnikov, Gaussian, Biweight, and Triweight kernel functions for all stations. It is evident from Table 4 that the Gaussian kernel function had the lowest RAE throughout all return periods among the Epanechnikov, Biweight, and Triweight kernel functions for Tarbela, Kalabagh, Taunsa, Guddu, Sukkur, Kotri, Mangla, Rasul, Marala, Khanki, Qadirabad, Trimmu, Sidhani, Sulemanki, and Islam stations.

The results in Table 4 for Chashma station indicate that the Triweight kernel function had a low RAE for the return period of 2 years, followed by the Biweight and Epanechnikov kernel functions; however, the Gaussian kernel function performed better for the return periods of 5, 10, 20, 50, 100, and 500 years. Similarly, the findings for the Balloki station revealed that the Biweight kernel function had an edge at the return periods of 2 and 5 years. Additionally, we also observed that the Epanechnikov kernel function had a lower RAE than Gaussian, Biweight, and Triweight kernel functions for the Panjnad station at a return period of 2 years, whereas the Epanechnikov, Biweight, and Triweight kernel functions performed equally well for the return period of 5 years. In accordance with the findings of the previous study [55], the results of our investigation demonstrated that among nonparametric kernel functions, the Gaussian kernel function performed best for the observed flood.

Finally, to evaluate the performance of the LH-moments ($\eta = 0, 1, 2$), the nonparametric kernel functions in terms of the RAE measures of quantile estimates were calculated and given in Tables 3 and 4. The findings reveal that the LH-moments ($\eta = 0, 1, 2$) led to more accurate estimates for most of the stations than did the nonparametric kernel function. On the other hand, it was also noted that the nonparametric kernel function performed better than LH-moments ($\eta = 0, 1, 2$) for Kalabagh, Chashma, and Guddu stations at return periods

of 2, 5, and 10 years. Besides that, among nonparametric kernel functions, the Gaussian kernel function provides very close estimates for smaller return periods as compared to LH-moments ($\eta = 0$). Similar findings were reported by Adamowski et al. [52], who evaluated L-moments and nonparametric methods for the annual maxima and partial duration flood series. However, in the case of L2-moments ($\eta = 2$), we found significant differences, specifically in the higher quantile estimates, with nonparametric kernel functions. This ensures that the L2-moments ($\eta = 2$) accurately estimate the extreme quantiles for the current dataset compared to any other approach considered in this work.

5. Conclusions

Estimating quantiles is a widespread practice in hydrology, and it is often used in the planning, design, and operation of a hydraulic system. In this study, we employed LH-moments ($\eta = 0, 1, 2$) and a nonparametric kernel function to estimate the peak flow series at 18 stations in Punjab, Pakistan. Based on the findings of this study, the following conclusions may be drawn:

Main findings of the paper: The findings of the AD test, KS test, CVM test, and LH-ratio diagram indicate that the best fits PDFs for estimating peak-flow data are GPA, followed by GEV and GLO distributions. It was identified that by raising the value of ($\eta = 1, 2$) in the LH-moments, the GOF test produced different findings for Mangla, Rasul, and Panjnad stations; nevertheless, increasing ($\eta = 1, 2$) did not affect the results for the rest of the stations. The magnitudes of quantile estimates obtained using the nonparametric kernel function are greater than those obtained through LH-moments ($\eta = 0, 1, 2$). Overall, the LH-moments ($\eta = 0, 1, 2$) accurately estimate the quantile in terms of RAE for most of the stations; however, for Kalabagh, Chashma, and Guddu stations at return periods of 2, 5, and 10 years, nonparametric kernel function provide smaller RAE than LH-moments ($\eta = 0, 1, 2$). The L-moment ($\eta = 0$), L1-moment ($\eta = 1$), and L2-moment ($\eta = 2$) provide relatively close estimates of quantile errors for all stations at the return periods of 2, 5, 10, and 20 years; moreover, L2-moments ($\eta = 2$) yielded the lowest error for the higher return period of 50, 100, and 500 years among L-moment ($\eta = 0$), L1-moments ($\eta = 1$), and nonparametric kernel functions. We also found that among nonparametric kernel functions for small return periods, the Gaussian kernel function provides a very close estimate compared to LH-moments ($\eta = 0, 1, 2$).

Limitations of this work and future research: Further research is needed on nonparametric kernel functions, specifically for large return periods, to improve the results in terms of RAE. This is the first application of nonparametric kernel functions in flood frequency analysis of 18 stations in Punjab, Pakistan. This research may be expanded by integrating all of Pakistan's river gauging stations in order to determine the best estimation methods for the whole country.

Broader impacts: The findings of this research will enhance recommendations for future development to preserve current infrastructure and minimize economic damage due to floods. Additionally, the findings will also aid in designing and implementing flood mitigation measures, such as more effective stormwater management.

Author Contributions: Conceptualization, M.F. and J.R.; Data curation, M.F.; Formal analysis, M.F.; Funding acquisition, T.Y.; Investigation, M.F., J.R. and L.C.; Methodology, M.F. and T.Y.; Software, M.F. and F.C.; Supervision, T.Y.; Validation, L.C. and T.Y.; Visualization, M.F. and F.C.; Writing—original draft, M.F. and F.C.; Writing—review & editing, F.C., J.R., L.C. and T.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors are grateful to the Water and Power Development Authority, Pakistan, and the Federal Flood Commission, Pakistan, for providing the required data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Stedinger, J.R.; Griffis, V.W. Flood frequency analysis in the United States: Time to update. *J. Hydrol. Eng.* **2008**, *13*, 199–204. [[CrossRef](#)]
2. De Michele, C.; Salvadori, G.; Canossi, M.; Petaccia, A.; Rosso, R. Bivariate statistical approach to check adequacy of dam spillway. *J. Hydrol. Eng.* **2005**, *10*, 50–57. [[CrossRef](#)]
3. Haddad, K.; Rahman, A. Regional flood frequency analysis in eastern Australia: Bayesian GLS regression-based methods within fixed region and ROI framework—Quantile regression vs. parameter regression technique. *J. Hydrol.* **2012**, *430*, 142–161. [[CrossRef](#)]
4. Cunnane, C. Statistical distributions for flood frequency analysis. In *Operational Hydrology Report*; WMO: Geneva, Switzerland, 1989.
5. Bobée, B.; Cavadias, G.; Ashkar, F.; Bernier, J.; Rasmussen, P. Towards a systematic approach to comparing distributions used in flood frequency analysis. *J. Hydrol.* **1993**, *142*, 121–136. [[CrossRef](#)]
6. Haddad, K.; Rahman, A. Selection of the best fit flood frequency distribution and parameter estimation procedure: A case study for Tasmania in Australia. *Stoch. Hydrol. Hydraul.* **2010**, *25*, 415–428. [[CrossRef](#)]
7. Hamed, K.; Rao, A.R. *Flood Frequency Analysis*; CRC Press: Boca Raton, FL, USA, 2019.
8. Ahmad, I.; Fawad, M.; Mahmood, I. At-site flood frequency analysis of annual maximum stream flows in Pakistan using robust estimation methods. *Pol. J. Environ. Stud.* **2015**, *24*, 2345–2353. [[CrossRef](#)]
9. Ahmad, I.; Fawad, M.; Akbar, M.; Abbass, A.; Zafar, H. Regional frequency analysis of annual peak flows in Pakistan using linear combination of order statistics. *Pol. J. Environ. Stud.* **2016**, *25*, 2255–2264. [[CrossRef](#)]
10. Hosking, J.R.M. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1990**, *52*, 105–124. [[CrossRef](#)]
11. Wang, Q.J. LH moments for statistical analysis of extreme events. *Water Resour. Res.* **1997**, *33*, 2841–2848. [[CrossRef](#)]
12. Hussain, Z.; Pasha, G.R. Regional flood frequency analysis of the seven sites of Punjab, Pakistan, using L-moments. *Water Resour. Manag.* **2008**, *23*, 1917–1933. [[CrossRef](#)]
13. Hussain, Z. Application of the regional flood frequency analysis to the upper and lower basins of the Indus River, Pakistan. *Water Resour. Manag.* **2011**, *25*, 2797–2822. [[CrossRef](#)]
14. Afreen, S.; Muhammad, F. Flood frequency analysis of various dams and barrages in Pakistan. *Irrig. Drain.* **2011**, *61*, 116–128. [[CrossRef](#)]
15. Lee, S.H.; Maeng, S.J. Comparison and analysis of design floods by the change in the order of LH-moment methods. *Irrig. Drain. J. Int. Comm. Irrig. Drain.* **2003**, *52*, 231–245. [[CrossRef](#)]
16. Hewa, G.A.; Wang, Q.J.; McMahan, T.; Nathan, R.J.; Peel, M.C. Generalized extreme value distribution fitted by LH moments for low-flow frequency analysis. *Water Resour. Res.* **2007**, *43*, W06301. [[CrossRef](#)]
17. Meshgi, A.; Khalili, D. Comprehensive evaluation of regional flood frequency analysis by L- and LH-moments. I. A revisit to regional homogeneity. *Stoch. Environ. Res. Risk Assess.* **2009**, *23*, 119–135. [[CrossRef](#)]
18. Meshgi, A.; Khalili, D. Comprehensive evaluation of regional flood frequency analysis by L- and LH-moments. II. Development of LH-moments parameters for the generalized Pareto and generalized logistic distributions. *Stoch. Hydrol. Hydraul.* **2008**, *23*, 137–152. [[CrossRef](#)]
19. Bhuyan, A.; Borah, M.; Kumar, R. Regional flood frequency analysis of North-Bank of the River Brahmaputra by using LH-moments. *Water Resour. Manag.* **2009**, *24*, 1779–1790. [[CrossRef](#)]
20. Gheidari, M.H.N. Comparisons of the L- and LH-moments in the selection of the best distribution for regional flood frequency analysis in Lake Urmia Basin. *Civ. Eng. Environ. Syst.* **2013**, *30*, 72–84. [[CrossRef](#)]
21. Ahmad, I.; Abbass, A.; Saghir, A.; Fawad, M. Finding probability distributions for annual daily maximum rainfall in Pakistan using linear moments and variants. *Pol. J. Environ. Stud.* **2016**, *25*, 925–937. [[CrossRef](#)]
22. Shabri, A. Comparisons of the LH moments and the L moments. *Matematika* **2002**, *18*, 33–43.
23. Deka, S.C.; Borah, M.; Kakaty, S.C. Statistical analysis of annual maximum rainfall in North-East India: An application of LH-moments. *Theor. Appl. Climatol.* **2010**, *104*, 111–122. [[CrossRef](#)]
24. Adamowski, K. Nonparametric kernel estimation of flood frequencies. *Water Resour. Res.* **1985**, *21*, 1585–1590. [[CrossRef](#)]
25. Adamowski, K. A Monte Carlo comparison of parametric and nonparametric estimation of flood frequencies. *J. Hydrol.* **1989**, *108*, 295–308. [[CrossRef](#)]
26. Adamowski, K.; Feluch, W. Nonparametric flood-frequency analysis with historical information. *J. Hydraul. Eng.* **1990**, *116*, 1035–1047. [[CrossRef](#)]
27. Schuster, E.; Yakowitz, S. Parametric/nonparametric mixture density estimation with application to flood-frequency analysis. (*JAWRA*) *J. Am. Water Resour. Assoc.* **1985**, *21*, 797–804. [[CrossRef](#)]
28. Adamowski, K.; Labatiuk, C. Estimation of flood frequencies by a nonparametric density procedure. In *Hydrologic Frequency Modeling*; Springer: Berlin/Heidelberg, Germany, 1987; pp. 97–106. [[CrossRef](#)]
29. Bardsley, W. Using historical data in nonparametric flood estimation. *J. Hydrol.* **1989**, *108*, 249–255. [[CrossRef](#)]
30. Guo, S.L. Nonparametric variable kernel estimation with historical floods and paleoflood information. *Water Resour. Res.* **1991**, *27*, 91–98. [[CrossRef](#)]
31. Moon, Y.-I.; Lall, U.; Bosworth, K. A comparison of tail probability estimators for flood frequency analysis. *J. Hydrol.* **1993**, *151*, 343–363. [[CrossRef](#)]

32. Moon, Y.-I.; Lall, U. Kernel quantite function estimator for flood frequency analysis. *Water Resour. Res.* **1994**, *30*, 3095–3103. [[CrossRef](#)]
33. Silverman, B.W. *Density Estimation for Statistics and Data Analysis*; CRC Press: Boca Raton, FL, USA, 1986.
34. Lall, U.; Moon, Y.-I.; Bosworth, K. Kernel flood frequency estimators: Bandwidth selection and kernel choice. *Water Resour. Res.* **1993**, *29*, 1003–1015. [[CrossRef](#)]
35. Cassalho, F.; Beskow, S.; de Mello, C.; de Moura, M.M.; Kerstner, L.; Ávila, L.F. At-site flood frequency analysis coupled with multiparameter probability distributions. *Water Resour. Manag.* **2017**, *32*, 285–300. [[CrossRef](#)]
36. Lall, U. Recent advances in nonparametric function estimation: Hydrologic applications. *Rev. Geophys.* **1995**, *33*, 1093–1102. [[CrossRef](#)]
37. Ahmad, I.; Almanjahie, I.M.; Hameedullah; Chikr-Elmezouar, Z.; Laksaci, A. Artificial neural network modeling for annual peak flows: A case study. *Appl. Ecol. Environ. Res.* **2019**, *17*, 6917–6935. [[CrossRef](#)]
38. Khan, M.; Hussain, Z.; Ahmad, I. A comparison of quadratic regression and artificial neural networks for the estimation of quantiles at ungauged sites in regional frequency analysis. *Appl. Ecol. Environ. Res.* **2019**, *17*, 6937–6959. [[CrossRef](#)]
39. Khan, M.; Hussain, Z.; Ahmad, I. Regional flood frequency analysis, using L-moments, artificial neural networks and OLS regression, of various sites of Khyber-Pakhtunkhwa, Pakistan. *Appl. Ecol. Environ. Res.* **2021**, *19*, 471–489. [[CrossRef](#)]
40. Zelterman, D. Parameter estimation in the generalized logistic distribution. *Comput. Stat. Data Anal.* **1987**, *5*, 177–184. [[CrossRef](#)]
41. Zelterman, D. Order statistics of the generalized logistic distribution. *Comput. Stat. Data Anal.* **1988**, *7*, 69–77. [[CrossRef](#)]
42. Reed, D. *Flood Estimation Handbook: Overview*; Institute of Hydrology Wallingford: Wallingford, UK, 1999.
43. Ashkar, F.; Mahdi, S. Fitting the log-logistic distribution by generalized moments. *J. Hydrol.* **2006**, *328*, 694–703. [[CrossRef](#)]
44. Jenkinson, A.F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Q. J. R. Meteorol. Soc.* **1955**, *81*, 158–171. [[CrossRef](#)]
45. Pickands, J., III. Statistical inference using extreme order statistics. *Ann. Stat.* **1975**, *3*, 119–131.
46. Hosking, J.; Wallis, J. An index flood procedure for regional rainfall frequency analysis. *EOS Trans. Am. Geophys. Union* **1987**, *68*, 312.
47. Singh, V.P.; Guo, H. Parameter estimation for 3-parameter generalized pareto distribution by the principle of maximum entropy (POME). *Hydrol. Sci. J.* **1995**, *40*, 165–181. [[CrossRef](#)]
48. Beskow, S.; Caldeira, T.L.; de Mello, C.R.; Faria, L.C.; Guedes, H.A.S. Multiparameter probability distributions for heavy rainfall modeling in extreme southern Brazil. *J. Hydrol. Reg. Stud.* **2015**, *4*, 123–133. [[CrossRef](#)]
49. Heo, J.-H.; Shin, H.; Nam, W.; Om, J.; Jeong, C. Approximation of modified Anderson–Darling test statistics for extreme value distributions with unknown shape parameter. *J. Hydrol.* **2013**, *499*, 41–49. [[CrossRef](#)]
50. Csörgő, S.; Faraway, J.J. The exact and asymptotic distributions of Cramér-von Mises statistics. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1996**, *58*, 221–234. [[CrossRef](#)]
51. Parzen, E. On estimation of a probability density function and mode. *Ann. Math. Stat.* **1962**, *33*, 1065–1076. [[CrossRef](#)]
52. Adamowski, K.; Liang, G.-C.; Patry, G.G. Annual maxima and partial duration flood series analysis by parametric and non-parametric methods. *Hydrol. Process.* **1998**, *12*, 1685–1699. [[CrossRef](#)]
53. Quintela-Del-Río, A. On bandwidth selection for nonparametric estimation in flood frequency analysis. *Hydrol. Process.* **2011**, *25*, 671–678. [[CrossRef](#)]
54. Francisco-Fernández, M.; Quintela-Del-Río, A. Comparing simultaneous and pointwise confidence intervals for hydrological processes. *PLoS ONE* **2016**, *11*, e0147505. [[CrossRef](#)]
55. Vittal, H.; Singh, J.; Kumar, P.; Karmakar, S. A framework for multivariate data-based at-site flood frequency analysis: Essentiality of the conjugal application of parametric and nonparametric approaches. *J. Hydrol.* **2015**, *525*, 658–675. [[CrossRef](#)]
56. Rajagopalan, B.; Lall, U.; Tarboton, D.G. Evaluation of kernel density estimation methods for daily precipitation resampling. *Stoch. Hydrol. Hydraul.* **1997**, *11*, 523–547. [[CrossRef](#)]
57. Adamowski, K. Regional analysis of annual maximum and partial duration flood data by nonparametric and L-moment methods. *J. Hydrol.* **2000**, *229*, 219–231. [[CrossRef](#)]
58. Altman, N.; Léger, C. Bandwidth selection for kernel distribution function estimation. *J. Stat. Plan. Inference* **1995**, *46*, 195–214. [[CrossRef](#)]
59. Quintela-del-Río, A.; Estévez-Pérez, G. Nonparametric kernel distribution function estimation with kerdiet: An R package for bandwidth choice and applications. *J. Stat. Softw.* **2012**, *50*, 1–21. [[CrossRef](#)]
60. Polansky, A.M.; Baker, E.R. Multistage plug—in bandwidth selection for kernel distribution function estimates. *J. Stat. Comput. Simul.* **2000**, *65*, 63–80. [[CrossRef](#)]
61. Bowman, A.; Hall, P.; Prvan, T. Bandwidth selection for the smoothing of distribution functions. *Biometrika* **1998**, *85*, 799–808. [[CrossRef](#)]
62. Quintela-del-Ri, A.; Francisco-Fernández, M. Nonparametric functional data estimation applied to ozone data: Prediction and extreme value analysis. *Chemosphere* **2011**, *82*, 800–808. [[CrossRef](#)]
63. Sarfaraz, S.; Arsalan, M.H.; Fatima, H. Regionalizing the climate of Pakistan using Köppen classification system. *Pak. Geogr. Rev.* **2014**, *69*, 111–132.
64. Fawad, M.; Yan, T.; Chen, L.; Huang, K.; Singh, V.P. Multiparameter probability distributions for at-site frequency analysis of annual maximum wind speed with L-moments for parameter estimation. *Energy* **2019**, *181*, 724–737. [[CrossRef](#)]