Communications for Statistical Applications and Methods
2017, Vol. 24, No. 5, 493–505

https://doi.org/10.5351/CSAM.2017.24.5.493
Print ISSN 2287-7843 / Online ISSN 2383-4757

A data-adaptive maximum penalized likelihood estimation for the generalized extreme value distribution

Youngsaeng Lee[a], Yonggwan Shin[a], Jeong-Soo Park[1,a]

[a]Department of Statistics, Chonnam National University, Korea

Abstract

Maximum likelihood estimation (MLE) of the generalized extreme value distribution (GEVD) is known to sometimes over-estimate the positive value of the shape parameter for the small sample size. The maximum penalized likelihood estimation (MPLE) with Beta penalty function was proposed by some researchers to overcome this problem. But the determination of the hyperparameters (HP) in Beta penalty function is still an issue. This paper presents some data adaptive methods to select the HP of Beta penalty function in the MPLE framework. The idea is to let the data tell us what HP to use. For given data, the optimal HP is obtained from the minimum distance between the MLE and MPLE. A bootstrap-based method is also proposed. These methods are compared with existing approaches. The performance evaluation experiments for GEVD by Monte Carlo simulation show that the proposed methods work well for bias and mean squared error. The methods are applied to Blackstone river data and Korean heavy rainfall data to show better performance over MLE, the method of L-moments estimator, and existing MPLEs.

Keywords: annual maximum daily rainfall, Beta distribution, bootstrap-based selection, L-moments estimation, penalty function, quantile estimation

1. Introduction

Generalized extreme value distribution (GEVD) has been used widely as a significant modelling tool to make an inference of extreme events such as heavy rainfall, wind speed, snowfall, earthquake and other related disciplines (Castillo *et al.*, 2005; Coles 2001; Katz *et al.*, 2002; Zhu *et al.*, 2013). Several estimation methods for GEVD parameters were developed in previous studies such as maximum likelihood estimation (MLE) or the method of L-moments estimation (L-ME). It is also found that MLE sometimes over-estimates the positive value of shape parameter ξ (rightly heavy tail case) in a small sample size. Consequently, it causes large bias and variance of extreme upper quantiles. In order to solve this problem, Coles and Dixon (1999) proposed a maximum penalized likelihood estimation (MPLE) that gives a penalty for a large value of ξ by considering an exponential penalty function. Martins and Stedinger (2000) considered a Beta probability density function (pdf) which can be treated as a prior for Bayesian approach.

The two hyper-parameters (HP) should be specified in both of Martins-Stedinger and Coles-Dixon penalty (or prior) functions. They selected a specific value for the HP based on experimentation and experience. Martins and Stedinger (2000) used a Beta(6, 9) pdf. This choice of prior restricts shape parameter to a plausible range ($-0.5 \le \xi \le 0.5$) consistent with rainfall and flood flows observed

[1] Corresponding author: Department of Statistics, Chonnam National University, 77 Yongbong-ro, Buk-gu, Gwangju 61186, Korea. E-mail: jspark@jnu.ac.kr

Published 30 September 2017 / journal homepage: http://csam.or.kr
© 2017 The Korean Statistical Society, and Korean International Statistical Society. All rights reserved.

worldwide (Huard *et al.*, 2010). The actual variance of the generalized extreme value (GEV) random variable is infinite when $\xi > 0.5$; in addition, the MLE of parameters is non-regular (asymptotic optimality is no more guaranteed) when $\xi < -0.5$. However, their selection of the HP in a Beta distribution, which plays a very important role in this framework, depends on subjective experiences and experiments that are not fully explained. Thus, Park (2005) introduced a Monte Carlo simulation-based systematic way to select the HP of the Beta pdf. His recommendation is Beta(2.5, 2.5) pdf.

However, the selection of the HP are irrelevant to the present data. There might be no problem if we treat the Beta pdf as a prior probability which was specified before the current data is given. But, if we treat the Beta pdf as a penalty function, it might be desirable to select the HP using given data because the data shows what HP to use. Thus, in this paper, we propose data-adaptive methods to select the HP of a Beta pdf, based on the MLE of shape parameter ξ. These methods are compared with the existing approaches. The performance evaluations for GEVD are conducted through a simulation study to illustrate that such a Beta penalty with optimally selected HP produces desirable quantile estimates.

Section 2 describes the GEVD, estimations methods and penalty functions. The proposed methods are presented in Section 3. Simulation study to evaluate the performance of the proposed methods is given in Section 4. Real data examples with Blackstone river flood discharge rate and Korean heavy rainfall are provided in Section 5. Discussion and conclusion are given in Section 6.

2. Generalized extreme value distribution and parameter estimation

The cumulative distribution function of the GEVD is as follows (Choi, 2015; Coles, 2001):

$$F(x;\mu,\sigma,\xi) = \exp\left\{-\left[1+\xi\frac{(x-\mu)}{\sigma}\right]^{-\frac{1}{\xi}}\right\}, \quad \text{if } \xi \neq 0, \tag{2.1}$$

for $1+\xi(x-\mu)/\sigma > 0$, where μ, $\sigma > 0$ and ξ are location, scale and shape parameter, respectively. The case for $\xi = 0$ in (2.1) is well known as the Gumbel distribution. By inverting (2.1), quantiles of GEVD are given by

$$x_p = \mu - \frac{\sigma}{\xi}\left[1-\{-\log(p)\}^{-\xi}\right], \quad \text{for } \xi \neq 0. \tag{2.2}$$

Estimates of x_p are obtained after substituting the estimates of (μ,σ,ξ) into (2.2) for various p-values.

Regarding the range of ξ, it is reported that ξ usually lies between -0.5 and 0.5 in hydrological practice. In addition, the GEVD has finite variance when $\xi < 0.5$ and is regular when $\xi > -0.5$. Thus ξ is confined between -0.5 and 0.5 in Martins and Stedinger (2000) and Park (2005) while it is extended to near 1.0 in Coles and Dixon (1999) (Figure 1).

2.1. Maximum likelihood estimation

Under the assumption that the observations $x_1, x_2, \ldots, x_n$ are independent variables having the GEVD, the negative log-likelihood function of (μ,σ,ξ) is

$$-l(\mu,\sigma,\xi) = n\ln\sigma + \left(1+\frac{1}{\xi}\right)\sum_{i=1}^{n}\ln\left[1+\frac{\xi(x_i-\mu)}{\sigma}\right] + \sum_{i=1}^{n}\left[1+\frac{\xi(x_i-\mu)}{\sigma}\right]^{-\frac{1}{\xi}}, \tag{2.3}$$

provided that $1+\xi((x_i-\mu)/\sigma) > 0$ for $i = 1,\ldots,n$. The MLE of μ,σ and ξ can be obtained by minimizing (2.3). We have to use a numerical optimization algorithm such as Newton type optimizer

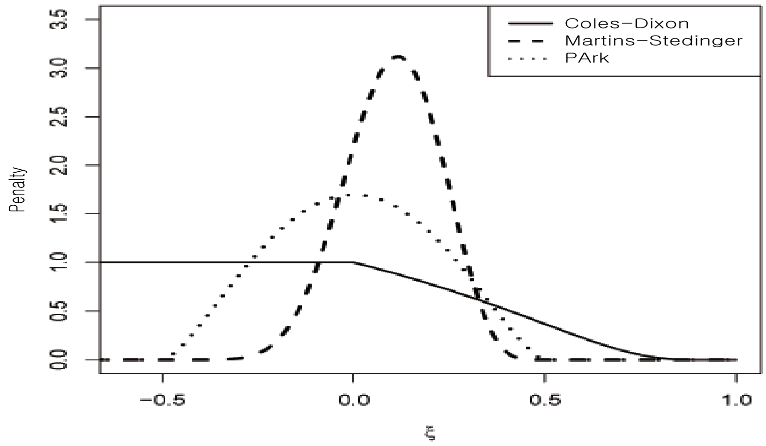

Figure 1: *Penalty functions of Park, Martins-Stedinger, and Coles-Dixon.*

since no explicit minimizer is available in minimizing (2.3). In this study, R package 'ismev' was used to obtain the MLE (Coles, 2001).

2.2. Method of L-moments estimation

The L-moments were introduced by Hosking (1990) as a linear combination of expectations of order statistics. The natural estimator of L-moments based on an observed sample of data is a linear combination of the ordered data values. The r^{th} sample L-moments (l_r) defined by Hosking (1990) are unbiased estimators of the population L-moments.

The method of L-ME obtains parameter estimates by equating the first k sample L-moments to the corresponding population quantities. The L-ME of GEVD are given as (Hosking *et al.*, 1985; Zhu *et al.*, 2013):

$$\hat{\mu} = l_1 - \frac{\hat{\sigma}}{\hat{\xi}} \left\{1 - \Gamma\left(1 + \hat{\xi}\right)\right\}, \tag{2.4}$$

$$\hat{\sigma} = \frac{l_2 \hat{\xi}}{\left(1 - 2^{-\hat{\xi}}\right)\Gamma\left(1 + \hat{\xi}\right)}, \tag{2.5}$$

$$\hat{\xi} = 7.8590c + 2.9554c^2, \tag{2.6}$$

where $c = 2/(3 + \hat{\tau}_3) - \log(2)/\log(3)$, and $\hat{\tau}_3$ is the L-skewness defined as $\hat{\tau}_3 = l_3/l_2$ from the sample. It is known that the L-ME works better than the MLE for small sample size. Moreover it is less sensitive to outlier (Hosking and Wallis, 1997).

These L-moments and L-ME have been used widely in many research fields including meteorology, civil engineering, and hydrology (for example, Busababodhin *et al.*, 2016; Meshgi and Khalili, 2009; Murshed *et al.*, 2014; Zhu *et al.*, 2013). We used R package 'lmom' developed by Hosking (2015) to calculate the sample L-moments and the L-ME of GEVD.

2.3. Maximum penalized likelihood estimation

For the small sample size, the MLE sometimes gives poor performance and over-estimates the large positive value of ξ severely. Consequently it causes large bias and variance of extreme upper quantiles. In order to solve this problem, Coles and Dixon (1999) and Martins and Stedinger (2000) proposed to use penalty functions on the positive value of ξ. The penalized negative log-likelihood to be minimized to obtain the MPLE is

$$l_{pen}(\mu, \sigma, \xi) = -\ln(L(\mu, \sigma, \xi)) + \ln(p(\xi)), \tag{2.7}$$

where $p(\xi)$ is a penalty function on ξ.

Coles and Dixon (1999) proposed the following penalty function;

$$p(\xi) = \begin{cases} 1, & \text{if } \xi \le 0, \\ \exp\left\{-\lambda\left(\dfrac{1}{1-\xi} - 1\right)^{\alpha}\right\}, & \text{if } 0 < \xi < 1, \\ 0, & \text{if } \xi \ge 1, \end{cases} \tag{2.8}$$

for non-negative values of α and λ. For the HP, Coles and Dixon (1999) suggested to use the combination $\alpha = 1$ and $\lambda = 1$ (Figure 1). For this function, we call it Coles-Dixon (CD) penalty as an abbreviation. Martins and Stedinger (2000) proposed the following penalty function, a Beta(α,β) pdf on ξ between -0.5 and 0.5:

$$p(\xi) = \frac{(0.5+\xi)^{\alpha-1}(0.5-\xi)^{\beta-1}}{B(\alpha,\beta)}, \tag{2.9}$$

where $B(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$ is the beta function. They chose $\alpha = 9$ and $\beta = 6$ based on prior hydrological information and experiments. For this function, we call it MS penalty as an abbreviation. Park (2005) recommended to use $\alpha = 2.5$ and $\beta = 2.5$. Yoon *et al.* (2010) considered a full Bayesian approach for the selection of HP. We denote the MPLEs using CD, MS, and Park penalties as the MPLE-CD, MPLE-MS, and MPLE-P, respectively.

3. Proposed approach

3.1. Selection of HP based on distance to ξ estimator

In order to select the HP (α,β) in the Beta pdf on ξ between -0.5 and 0.5 as in (2.9), we considered the distance between the estimator (MLE for the first time) and MPLE of ξ. The proposed method is:

• **Method SHM** (selection of the HP based on the MLE):

Step 1. Compute the MLE of GEV parameters, and denote it $\hat{\xi}_M$.

Step 2. Find the (α,β) which minimize the distance $|\hat{\xi}_M - \hat{\xi}(\alpha,\beta)|$, where $\hat{\xi}(\alpha,\beta)$ is the MPLE for given (α,β). Denote such selected HP as $(\alpha^\star,\beta^\star)$.

The estimator is now obtained by the MPLE with a Beta($\alpha^\star,\beta^\star$) penalty function. In this method, we still respect the MLE from data but restrict $\hat{\xi}$ to be in $[-0.5, 0.5]$ and to follow a Beta distribution. In minimizing the distance between MLE and MPLE of ξ in Step 2, a grid search was applied by

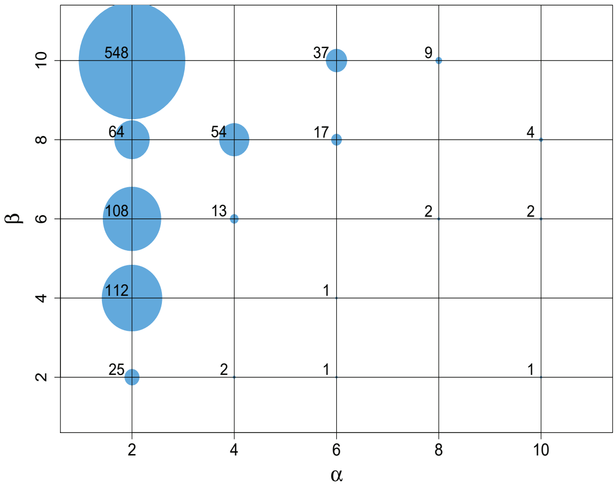

Figure 2: *Distribution of selected values* $(\alpha^\star, \beta^\star)$ *from the selection of HP based on the maximum likelihood estimator (SHM) method for* $\xi = 0.4$ *and for sample sizes* $n = 30$.

changing α and β from 2 to 14 by increment 2. Then, a finer grid search is performed around the coarsely found optimal set of $(\alpha^\star, \beta^\star)$. That is, the distance computation over the grids of $(\alpha^\star - 1,\ \alpha^\star - 0.5,\ \alpha^\star,\ \alpha^\star + 0.5,\ \alpha^\star + 1) \times (\beta^\star - 1,\ \beta^\star - 0.5,\ \beta^\star,\ \beta^\star + 0.5,\ \beta^\star + 1)$ are tried. We refer this method as a selection of the HP based on the MLE, or SHM in abbreviation.

Figure 2 shows distributions of selected values $(\alpha^\star, \beta^\star)$ from the SHM method. It was calculated for 1000 random samples generated from GEVD for $\xi = 0.4$ and for sample sizes $n = 30$. The value $(2, 10)$ was selected the most often.

Figure 3 shows probability density plot of $\hat{\xi}$ obtained by the methods considered in this paper. It was calculated for 1,000 random samples generated from GEVD for $\xi = 0.4$ and for sample sizes $n = 30$. It was drawn using the function "density" in the R program in which a Gaussian kernel density estimation method is implemented. The legends 'pbeta(2.5, 2.5)', 'pbeta(6, 9)', and 'proposed' in this figure stand for the MPLE with Park, Martins-Stedinger, and the SHM method, respectively. Figure 3 ahows that MPLE-P and MPLE-MS under-estimate severely. MLE has big variance. The proposed methods work well and has smaller variance than MLE.

3.2. Bootstrap-based selection of the hyperparameters

We will obtain the sample distribution of $\hat{\xi}$ by using bootstrap samples, and then find the (α, β) which minimize the distance between the distribution of Beta(α, β) and the sample distribution of $\hat{\xi}$. For a given data set, we obtain B MLEs of ξ from B bootstrap samples. The relative frequency is then calculated for each category. Here the number (K) of categories and the width of category are automatically selected by "hist" function in R software. Usually K is chosen between 10 and 15. Only the estimates $\hat{\xi}$ in $[-0.5, 0.5]$ are used when computing relative frequencies. That means that the estimates outside $[-0.5, 0.5]$ are eliminated; subsequently, the total number of the remaining estimates is then used as a denominator in computing relative frequencies. This is done to make a set of relative frequencies form a kind of probability mass function inside the interval $[-0.5, 0.5]$.

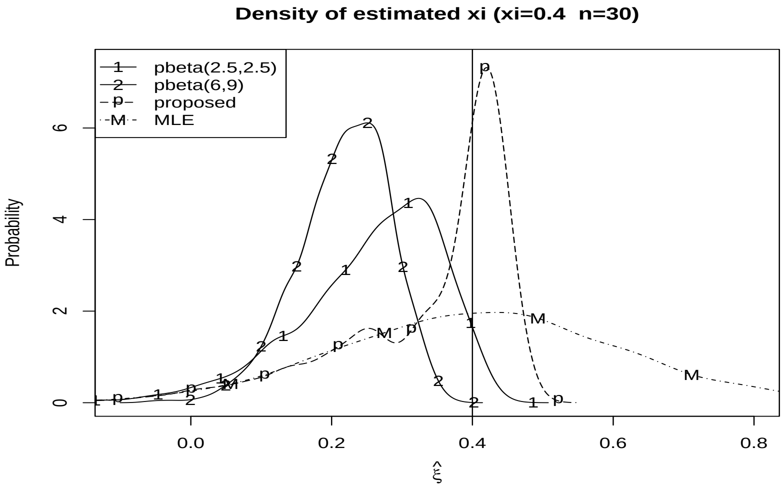

Figure 3: *Probability density plot of* $\hat{\xi}$ *for* $\xi = 0.4$ *and for sample sizes* $n = 30$*. The vertical line represents for the true value of* ξ*. The legends Pbeta*$(2.5, 2.5)$*, Pbeta*$(6, 9)$*, and 'proposed' stand for the maximum penalized likelihood estimation with Park, and Martins-Stedinger, and the SHM (selection of hyperparameters based on the maximum likelihood estimator) method, respectively.*

• **Method BSHM** (bootstrap based selection of the HP using the MLE):

After the above computation, the following measure of discrepancy is minimized with respect to $(\alpha,\ \beta)$;

$$Q(\alpha,\beta) = \sum_{k=1}^{K} \left[p\left(\hat{\xi}_{(k)};\alpha,\beta\right) - \hat{f}\left(\hat{\xi}_{(k)}\right)\right]^2, \tag{3.1}$$

where $p(\hat{\xi}_{(k)};\alpha,\beta)$ is, as in (2.9), the pdf of Beta(α,β) at the center point of k^{th} category, and $\hat{f}(\hat{\xi}_{(k)})$ is the relative frequency of $\hat{\xi}$ at k^{th} category. A grid search that is the same as the algorithm in the above subsection is used. Because of computational costs, we set $B = 100$. We call this method a bootstrap based selection of the HP using the MLE, or BSHM in short.

3.3. Selection of the hyperparameters minimizing prediction squared error

As an analogy to the selection of the smoothing parameter, we consider the following prediction squared error (pse);

$$\text{pse}(\alpha,\beta) = \frac{1}{n}\sum_{i=1}^{n} \left[\hat{x}_{(i)}(\alpha,\beta) - x_{(i)}\right]^2, \tag{3.2}$$

where $x_{(i)}$ is the i^{th} order statistic from the original data, and $\hat{x}_{(i)}(\alpha,\beta)$ is the quantile estimate for the plotting position $p_{i:n} = (i - .35)/n$, as recommended by Hosking *et al.* (1985) for the GEVD. This is obtained by plugging the MPLE with HP (α,β) into the quantile function (2.2). The HP which minimize pse(α,β) are selected. A grid search that is the same as the algorithm in the above subsection is used with $B = 200$. We call this method as a selection of the HP using the pse criterion, or SHPSE in abbreviation. Table 1 provides a description on the abbreviated names of the estimation methods considered in this paper.

Table 1: Description on the abbreviated names of the estimation methods considered in this paper

Methods	Description	Details
SHM	Selection of hyperparameters using the MLE	Subsection 3.1
BSHM	Bootstrap based selection of hyperparameters using the MLE	Subsection 3.2
SHPSE	Selection of hyperparameters minimizing prediction squared error	Equation (3.2)
MPLE-P	Maximum penalized likelihood estimation using Park's penalty	Beta(2.5, 2.5) pdf
MPLE-MS	Maximum penalized likelihood estimation using Martins-Stedinger's penalty	Beta(9, 6) pdf
MPLE-CD	Maximum penalized likelihood estimation using Coles-Dixon's penalty	Equation (2.8)
L-ME	Method of L-moments estimation	Subsection 2.2
MLE	Maximum likelihood estimation	Subsection 2.1

pdf = probability density function.

Table 2: The bias of ξ estimators obtained by proposed and other estimation methods as the true ξ ranges from −0.49 to 0.49 with sample size of 30

Methods	−0.49	−0.4	−0.3	−0.2	−0.1	0	0.1	0.2	0.3	0.4	0.49	Sum
SHM	0.031	−0.014	−0.021	−0.024	−0.021	**−0.020**	−0.013	−0.013	−0.014	**−0.022**	−0.070	**0.265**
BSHM	0.031	−0.019	−0.026	−0.027	−0.021	**−0.017**	**−0.008**	**−0.006**	**−0.007**	**−0.022**	−0.073	**0.258**
SHPSE	0.041	**−0.005**	**−0.009**	**0.005**	0.046	0.080	0.073	0.039	−0.056	−0.148	−0.238	0.740
MPLE-P	0.152	0.097	0.063	0.032	**0.008**	**0.017**	−0.039	−0.070	−0.103	−0.142	−0.197	0.921
MPLE-MS	0.361	0.293	0.235	0.174	0.115	0.058	**0.003**	−0.059	−0.119	−0.182	−0.248	1.846
MPLE-CD	−0.051	−0.051	−0.035	−0.03	−0.028	−0.036	−0.040	−0.054	−0.062	−0.073	−0.106	0.567
L-ME	**0.002**	**−0.011**	**−0.002**	**−0.003**	**−0.007**	−0.022	−0.023	−0.039	−0.046	−0.064	−0.096	0.316
MLE	−0.051	−0.051	−0.034	−0.028	−0.021	**−0.020**	−0.012	**−0.011**	**0.002**	**0.021**	**0.017**	**0.270**

Sum is obtained by the summation of the absolute biases. Table 1 provides descriptions on the abbreviated names of methods.

4. Monte Carlo simulation

In order to evaluate the performance of the proposed method, we compared the proposed estimator to the other estimators by Monte Carlo simulation study. For the comparison of accuracy, we calculated the bias and the root mean squared error (RMSE) of ξ estimators;

$$\mathrm{Bias}(\xi) = \frac{1}{M}\sum_{i=1}^{M}\left(\hat{\xi}_i - \xi\right), \tag{4.1}$$

and

$$\mathrm{RMSE}(\xi) = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\hat{\xi}_i - \xi\right)^2}. \tag{4.2}$$

We have generated $M = 1{,}000$ random samples from GEVD for sample sizes $n = 30, 60$ and for given shape parameters $\xi \in (-0.5, 0.5)$. Other parameters are fixed as location $\mu = 0$ and scale $\sigma = 1$, because these parameters are location and scale equivariant.

Table 2 and Figure 4 show the results of the bias of ξ estimates (Bias(ξ)). There are negative biases both in the MLE for negative ξ and in the L-ME for positive ξ. The MPLE-P and MPLE-MS are good only for near zero ξ, but are worst for ξ far from zero. The sizes of biases for $\xi = 0.4$ are larger than those for $\xi = 0.2$. Based on the summation of the absolute biases (sum), BSHM method works the best and SHM and MLE work well. MPLE-MS is the worst in general. Table 3 and Figure 5 show the results of the RMSE of ξ estimates (RMSE(ξ)). The MLE and L-ME work similarly while those are not good for ξ far from zero. SHM and BSHM and methods work well for general ξ, specially

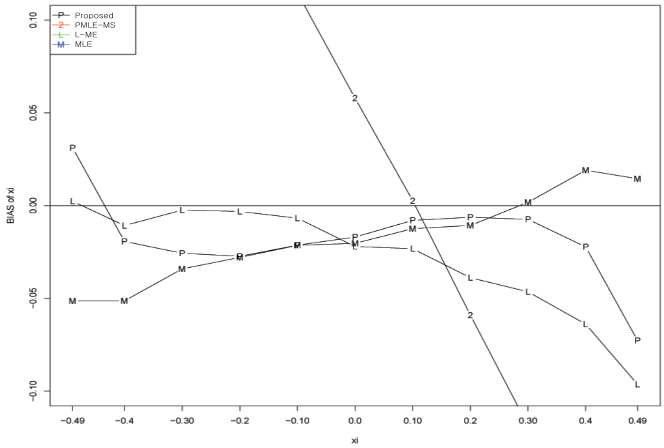

Figure 4: *The bias of ξ estimators obtained by four estimation methods (MLE, L-ME, MPLE-MS, and SHM) with sample size of 30. Description on the abbreviated names of methods is given in Table 1.*

Table 3: The root mean squared error of ξ estimators obtained by proposed and other estimation methods as the true ξ ranges from −0.49 to 0.49 with sample size of 30

Methods	−0.49	−0.4	−0.3	−0.2	−0.1	0	0.1	0.2	0.3	0.4	0.49	sum
SHM	**0.080**	**0.098**	0.135	0.145	0.157	0.175	0.178	0.173	0.174	**0.141**	**0.139**	**1.595**
BSHM	**0.078**	**0.097**	0.133	0.149	0.162	0.179	0.182	0.181	0.173	**0.137**	**0.136**	**1.607**
SHPSE	0.114	0.117	0.169	0.215	0.267	0.295	0.265	0.225	0.200	0.222	0.289	2.377
MPLE-P	0.158	0.122	**0.114**	**0.106**	**0.112**	0.115	0.132	0.138	0.158	0.178	0.227	**1.561**
MPLE-MS	0.366	0.300	0.243	0.185	0.132	**0.090**	**0.068**	**0.089**	**0.137**	0.192	0.256	2.057
MPLE-CD	0.167	0.155	0.162	0.151	0.151	0.162	0.162	0.165	0.172	0.161	0.178	1.786
L-ME	0.159	0.151	0.150	0.138	0.140	0.153	0.162	0.167	0.185	0.183	0.211	1.798
MLE	0.167	0.155	0.164	0.154	0.160	0.177	0.182	0.186	0.200	0.203	0.213	1.960

Table 1 provides descriptions on the abbreviated names of methods.

for ξ far from zero. Note that SHPSE work badly. MPLE-MS works well only for ξ between 0.0 and 0.3, but worst for ξ far from zero. The MPLE-P works well for ξ between −0.3 and 0.0, and is the best in the sense of the summation of the RMSEs (sum). From the results based on both bias and mean squared error criterion, we would conclude that the SHM and BSHM work better than the other estimation methods.

5. Real data examples

5.1. Blackstone River data

We considered the data of the annual flood discharge rates of the Blackstone River at Woonsocket, RI, USA, given in Pericchi and Rodriguez-Iturbe (1985), and Mudholkar and Hutson (1998). This data is for a period of 37 years with unit of ft^3/s.

To judge the overall goodness-of-fit, we use the Kolmogorov-Smirnov (K-S) statistic and the *average scaled absolute error* (ASAE) (Castillo *et al.*, 2005),

$$\text{ASAE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|x_{(i)} - \hat{x}_{(i)}\right|}{x_{(n)} - x_{(1)}}, \tag{5.1}$$

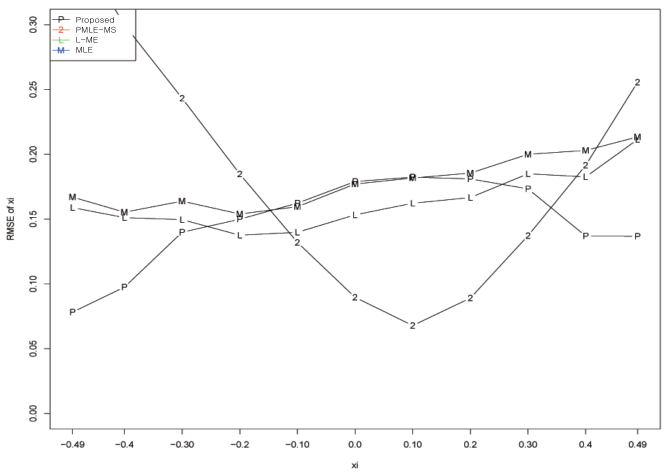

Figure 5: *The root mean squared error (RMSE) of ξ estimators obtained by four estimation methods (MLE, L-ME, MPLE-MS, and SHM) with sample size of 30. Table 1 provides descriptions on the abbreviated names.*

Table 4: Result of the analysis and comparison of the estimation methods for the Blackstone river data

	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\xi}$	ASAE	CM	CIU	CIL	CI range
SHM	4432.9	1866.2	0.268	0.027	0.126	0.472	0.100	0.372
BSHM	4474.1	1845.9	0.203	0.029	**0.114**	0.326	0.086	**0.240**
SHPSE	4469.0	1846.8	0.210	0.029	0.115	0.368	0.074	0.294
MPLE-P	4453.1	1851.4	0.232	0.028	0.118	0.411	0.084	0.327
MPLE-MS	4467.8	1846.3	0.211	0.029	0.115	0.350	0.082	0.268
MPLE-CD	4441.5	1857.2	0.250	0.027	0.122	0.481	0.087	0.394
L-ME	4256.8	1441.3	0.479	**0.017**	**0.109**	0.616	−0.046	0.662
MLE	4430.6	1867.3	0.269	0.027	0.127	0.533	0.097	0.436

Estimates of μ, σ, and ξ, ASAE, CM, CIU, and CIL, and the range of CI. Table 1 provides descriptions on the abbreviated names of methods. ASAE = average scaled absolute error; CM = Cramer-von Mises Statistic; CIU = upper bounds of 95% confidence interval of ξ; CIL = lower bounds of 95% confidence interval of ξ.

where $x_{(i)}$ are the ascendingly ordered observations, and $\hat{x}_{(i)}$ is obtained from the quantile function (2.2) with the estimates plugged in the equation for the plotting position $p_{i:n} = (i - 0.35)/n$, as recommended by Hosking and Wallis (1997) for the GEVD.

Table 4 provides the estimation of μ, σ, and ξ along with ASAE, Cramer–von Mises statistic (CM), upper and lower bounds of 95% confidence interval (CIU and CIL) of ξ), and the range of confidence interval (CI range). To compute the confidence intervals of $\hat{\xi}$, the profile likelihood approach was used for MLE based methods, while the bootstrap ($B = 1{,}000$) method was used for L-ME. The profile likelihood of GEV for ξ_0 is defined as (Coles, 2001)

$$L_p(\xi_0; \mathbf{x}) = \max_{\mu,\sigma|\xi_0} L(\mu, \sigma, \xi_0; \mathbf{x}), \tag{5.2}$$

where $L(\cdot)$ is the likelihood function for given data $\mathbf{x} = x_1, \ldots, x_n$. The L-ME works well for criteria of ASAE and CM, while BSHM works well for CM criterion and CI range. Note that the CI range of L-ME is the worst.

Table 5: Estimates of shape parameter ξ, ASAE, and K-S statistic for the annual maximum of daily rainfall data in twelve Korean sites, for five estimation methods

Location		MPLE-P	MPLE-MS	SHM	(α^*, β^*)	BSHM	(α^*, β^*)	MLE
Daegwallyeong	$\hat{\xi}$	0.221	0.196	0.285		0.165		0.283
n = 42	ASAE	0.017	0.017	**0.016**	(4, 12)	0.017	(10, 14)	**0.016**
	K-S	0.064	**0.061**	0.072		0.062		0.072
Daejeon	$\hat{\xi}$	0.199	0.176	0.311		0.240		0.304
n = 44	ASAE	0.026	**0.025**	0.031	(4, 14)	0.027	(6, 14)	0.031
	K-S	0.096	0.098	**0.083**		0.092		0.084
Pohang	$\hat{\xi}$	0.233	0.212	0.264		0.213		0.267
n = 64	ASAE	**0.017**	**0.017**	**0.017**	(2, 4)	0.018	(8, 14)	**0.017**
	K-S	0.073	**0.072**	0.073		**0.072**		0.073
Gunsan	$\hat{\xi}$	0.228	0.201	0.286		0.286		0.286
n = 45	ASAE	0.034	0.034	**0.033**	(4, 12)	**0.033**	(4, 12)	**0.033**
	K-S	0.108	**0.104**	0.116		0.116		0.116
Wando	$\hat{\xi}$	0.232	0.199	0.312		0.295		0.313
n = 42	ASAE	**0.033**	0.034	0.034	(4, 14)	0.034	(4, 12)	0.034
	K-S	0.103	0.109	**0.091**		0.094		**0.091**
Seogwipo	$\hat{\xi}$	0.206	0.186	0.250		0.234		0.250
n = 52	ASAE	**0.026**	**0.026**	0.027	(2, 4)	0.027	(6, 14)	0.027
	K-S	0.086	0.089	**0.082**		0.084		**0.082**
Buyeo	$\hat{\xi}$	0.248	0.215	0.314		0.200		0.319
n = 41	ASAE	0.020	0.020	**0.018**	(4, 14)	0.021	(8, 12)	0.019
	K-S	0.087	0.082	0.095		**0.080**		0.096
Imsil	$\hat{\xi}$	0.214	0.187	0.287		0.287		0.292
n = 42	ASAE	**0.035**	**0.035**	0.038	(4, 12)	0.038	(4, 12)	0.039
	K-S	0.089	0.092	**0.081**		**0.081**		0.082
Jeongeup	$\hat{\xi}$	0.224	0.202	0.263		0.204		0.267
n = 43	ASAE	0.021	0.022	**0.019**	(2, 4)	0.023	(8, 14)	**0.019**
	K-S	0.083	0.084	**0.082**		0.084		**0.082**
Haenam	$\hat{\xi}$	0.292	0.240	0.401		0.338		0.405
n = 42	ASAE	0.021	0.026	**0.012**	(2, 10)	0.017	(4, 14)	**0.012**
	K-S	0.068	0.077	0.061		0.065		**0.060**
Goheung	$\hat{\xi}$	0.233	0.203	0.309		0.250		0.301
n = 41	ASAE	0.020	0.023	**0.013**	(4, 14)	0.019	(6, 14)	0.014
	K-S	0.081	0.083	**0.079**		0.080		**0.079**
Yeongdeok	$\hat{\xi}$	0.269	0.222	0.377		0.411		0.384
n = 40	ASAE	0.024	0.028	**0.023**	(2, 8)	0.025	(2, 12)	**0.023**
	K-S	0.095	0.098	0.085		**0.082**		0.084
No. of the best	ASAE	4	4	8		1		6
No. of the best	K-S	0	3	6		4		5

The selected values $(\alpha^\star, \beta^\star)$ from SHM and BSHM are provided. Table 1 provides descriptions on the abbreviated names of methods. ASAE = average scaled absolute error; K-S = Kolmogorov-Smirnov.

5.2. Korean heavy rainfall data

In this section, we compared the performance of the proposed method with existing MPLEs using Korean heavy rainfall data. Annual daily maximum precipitation (unit: mm) record are considered for 75 weather stations which has at least 20 years observation (Korea Meteorological Administration, 2016). The stations with relatively large positive MLE value of $\hat{\xi}$ (> 0.25) are selected. Table 5 shows the 12 stations.

The estimates of ξ, ASAE, and K-S statistic for various methods are given in Table 5 for the 12 stations. The selected values $(\alpha^\star, \beta^\star)$ from SHM and BSHM are provided. Based on the number of the best at the bottom of Table 5, the SHM method works best compared to other methods.

6. Discussion

In this paper, we restrict the range of shape parameter to be in $(-0.5, 0.5)$ because the variance of GEV random variable is infinite when $\xi > 0.5$ and the MLE of parameters is non-regular (asymptotic optimality is no more guaranteed) when $\xi < -0.5$. Therefore, our concern was concentrated on Beta pdf only. One can release this restriction to $(-1.0, 1.0)$. For this case, one can use the Coles-Dixon (CD) penalty function (2.8). Based on our experimental experience on the CD penalty function, the HP $\alpha = 1, \lambda = 1$ work well. If a data-adaptive selection of HP (α, λ) in CD penalty function is recommended, one can consider a criterion like a pse(α, λ) similar as in (3.2) to choose the best α and λ from the data.

In Step 2 of the SHM method, one can consider refining the grid search using the increment 0.5 or 0.2. That may be able to find the HP that makes the density function very narrow and concentrate to the MLE, so that the MPLE be near the MLE. However, to our experiments showed that refining the grid did not improve the results obtained using the coarse grid as reported in this paper.

One can consider a criterion with the help of the method of L-ME. For example, the distance between population L-moments (calculated from the MPLE) and sample L-moments, i.e., $|\tau_3 - t_3|$ can be used to select the HP, where τ_3 and t_3 are population L-skewness and sample L-skewness, respectively. The HP minimizing the above distance is then selected. We expect this MPLE, guided by L-ME, to work well because L-ME works better than the MLE for small sample size. It is a topic for future research.

We tried the following bootstrap based criterion, using the pse as in Efron and Tibshirani (1993);

$$\text{pse}^*(\alpha,\beta) = \frac{1}{n}\sum_{i=1}^{n}\left[\hat{x}^*_{(i)}(\alpha,\beta) - x_{(i)}\right]^2, \tag{6.1}$$

where $\hat{x}^*_{(i)}(u, v)$ is the same as the above but is obtained from the MPLE for a bootstrap sample x^*_i, $i = 1, 2, \ldots, n$. Averaging this quantity pse$^*(u, v)$ over B bootstrap samples provides an estimate of the expected prediction squared error. We denote this average by $\widehat{\text{pse}}(u, v)$. The HP minimizing $\widehat{\text{pse}}(u, v)$ can be selected; however, the disadvantage of this bootstrap approach was not reported it in this paper because it a computational cost, and it did not work well in our brief simulation study.

7. Conclusion

A data adaptive method to select the HP of Beta pdf on the shape parameter of GEVD is presented in a MPLE framework that enables the data to tell us what HP to use. For given data, the optimal HP is obtained from the minimum distance between the MLE and MPLE. The performance evaluation experiments for GEVDs by Monte Carlo simulation show that the proposed estimators often work well. Blackstone river data and Korean heavy rainfall data are fitted to illustrate the usefulness of the proposed methods. Our recommendation is to use the SHM method among some estimations considered in this study. The details of the SHM are described in Subsection 3.1. A computer program for the proposed methods developed using R software is available upon request from the corresponding author.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2016R1A2B4014518). Lee's work was supported by Basic Science

Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2017R1A6A3A11032852).

References

Busababodhin P, Seo YA, Park JS, and Kumphon B (2016). LH-moment estimation of Wakeby distribution with hydrological applications, *Stochastic Environmental Research and Risk Assessment*, **30**, 1757–1767.

Castillo E, Hadi AS, Balakrishnan N, and Sarabia JM (2005). *Extreme Value and Related Models with Applications in Engineering and Science*, Wiley-Interscience, New Jersey.

Choi H (2015). A note on the dependence conditions for stationary normal sequences, *Communications for Statistical Applications and Methods*, **22**, 647–653.

Coles S (2001). *An Introduction to Statistical Modeling of Extreme Values*, Springer, New York.

Coles SG and Dixon MJ (1999). Likelihood-based inference for extreme value models, *Extremes*, **2**, 5–23.

Efron B and Tibshirani RJ (1993). *An Introduction to the Bootstrap*, Chapman & Hall/CRC, Boca Raton.

Hosking JRM (1990). L-moments: analysis and estimation of distributions using linear combinations of order statistics, *Journal of the Royal Statistical Society Series B (Methodological)*, **52**, 105–124.

Hosking JRM (2015). Package 'lmom', version 2.5, Retrieved June 9, 2016, from: https://cran.r-project.org/web/packages/lmom/lmom.pdf

Hosking JRM and Wallis JR (1997). *Regional Frequency Analysis: An Approach based on L-Moments*, Cambridge University Press, Cambridge.

Hosking JRM, Wallis JR, and Wood EF (1985). Estimation of the generalized extreme-value distribution by the method of probability weighted moments, *Technometrics*, **27**, 251–261.

Huard D, Mailhot A, and Duchesne S (2010). Bayesian estimation of intensity-duration-frequency curves and of the return period associated to a given rainfall event, *Stochastic Environmental Research and Risk Assessment*, **24**, 337–347.

Katz RW, Parlange MB, and Naveau P (2002). Statistics of extremes in hydrology, *Advances in Water Resources*, **25**, 1287–1304.

Korea Meteorological Administration (2016). Annual daily maximum precipitation record for 75 weather stations, Retrieved May 1, 2017, from: http://www.kma.go.kr

Martins ES and Stedinger JR (2000). Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data, *Water Resources Research*, **36**, 737–744.

Meshgi A and Khalili D (2009). Comprehensive evaluation of regional flood frequency analysis by L- and LH-moments. I. A re-visit to regional homogeneity, *Stochastic Environmental Research and Risk Assessment*, **23**, 119–135.

Mudholkar GS and Hutson AD (1998). LQ-moments: analogs of L-moments, *Journal of Statistical Planning and Inference*, **71**, 191–208.

Murshed MS, Seo YA, and Park JS (2014). LH-moment estimation of a four parameter kappa distribution with hydrologic applications, *Stochastic Environmental Research and Risk Assessment*, **28**, 253–262.

Park JS (2005). A simulation-based hyperparameter selection for quantile estimation of the generalized extreme value distribution, *Mathematics and Computers in Simulation*, **70**, 227–234.

Pericchi LR and Rodriguez-Iturbe I (1985). On the statistical analysis of floods. In Atkinson AC,

Fienberg SE (Eds), *A Celebration of Statistics* (pp. 511–541), Springer, New York.
Yoon S, Cho W, Heo JH, and Kim CE (2010). A full Bayesian approach to generalized maximum likelihood estimation of generalized extreme value distribution, *Stochastic Environmental Research and Risk Assessment*, **24**, 761–770.
Zhu J, Forsee W, Schumer R, and Gautam M (2013). Future projections and uncertainty assessment of extreme rainfall intensity in the United States from an ensemble of climate models, *Climatic Change*, **118**, 469–485.

Received June 12, 2017; Revised July 30, 2017; Accepted August 21, 2017