International Journal of Advanced and Applied Sciences, 9(2) 2022, Pages: 104-108

Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Parameter estimate for three-parameter kappa distribution using LH-moments approach

Zahrahtul Amani Zakaria [1, *], Jarah Moath Suleiman Ali [1], Wan Nur Syahidah Wan Yusoff [2]

[1]*Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Kuala Terengganu, Malaysia*
[2]*Centre for Mathematical Sciences, College of Computing and Applied Sciences, Universiti Malaysia Pahang, Pahang, Malaysia*

A R T I C L E I N F O

Article history:
Received 27 June 2021
Received in revised form
27 September 2021
Accepted 9 December 2021

Keywords:
Higher-order moments
L-moments
Probability distribution
Rainfall forecasting

A B S T R A C T

The method of higher-order L-moments (LH-moment) was proposed as a more robust alternative compared to classical L-moments to characterize extreme events. The new derivation will be done for Mielke-Johnson's Kappa and Three-Parameters Kappa Type-II (K3D-II) distributions based on the LH-moments approach. The data of maximum monthly rainfall for Embong station in Terengganu were used as a case study. The analyses were conducted using the classical L-moments method with $\eta = 0$ and LH-moments methods with $\eta = 1$, $\eta = 2$, $\eta = 3$ and $\eta = 4$ for a complete data series and upper parts of the distributions. The most suitable distributions were determined based on the Mean Absolute Deviation Index (MADI), Mean Square Deviation Index (MSDI), and Correlation (r). Also, L-moment and LH-moment ratio diagrams were used to represent visual proofs of the results. The analysis showed that LH-moments methods at a higher order of K3D-II distribution best fit the data of maximum monthly rainfalls for the Embong station for the upper parts of the distribution compared to L-moments. The results also proved that whenever η increases, LH-moments reflect more and more characteristics of the upper part of the distribution. This seems to suggest that LH-moments estimates for the upper part of the distribution events are superior to L-moments in fitting the data of maximum monthly rainfalls.

1. Introduction

Floods are extreme events that often occur in a short period of time and may kill people and destroy property. Therefore, necessary to design the hydraulic structures to estimate the risks caused by such extreme events. Analysis of frequency is utilized to find suitable probability distributions for the major event (Samantaray and Sahoo, 2021). Statistical analysis of major events is frequently carried for forecasting significant return period occurrences. The presence of extreme data in flood data makes it hard to investigate the qualities of statistical distributions because extreme events are inherently occasional and happen over a short time period (Bhat et al., 2019).

The approximation of how often a certain event will take place is called frequency analysis (Hosking and Wallis, 1997). There are many statistical methods used in frequency analysis and estimation of the parameters of statistical distribution to predict the probability of upcoming events based on installation the previous observations on selected statistical distributions (Harun et al., 2017; Makhtar et al., 2016; Zhou et al., 2017; Sharafi et al., 2021; Murshed et al. 2018). Some methods that rely on moments that have been used over a long period in frequency analysis are not always satisfactory especially when the sample is small. Hosking (1990) introduced the L-moments approach, which has a major role in probability distributions parameter estimation. This approach has become one of the broadest methods used for frequency analysis (Rao and Hamed, 2000).

L-moments estimates have characterized by a range of characteristics that make them superior to other estimates which depend on the moments. The most important of these characteristics, it's more robust when there are extreme values in the data compared to conventional moments and their ability to describe a wide range of distributions. Compared to estimates of maximum likelihood, L-moments exist whenever the distribution arithmetic mean

* Corresponding Author.
Email Address: zahrahtulamani@unisza.edu.my (Z. A. Zakaria)
https://doi.org/10.21833/ijaas.2022.02.011
Corresponding author's ORCID profile:
https://orcid.org/0000-0002-7198-1403

exists and they are more accurate and more favorable in the circumstance of small trials or when the estimates of the maximum likelihood are not available or have undesirable characteristics, or are difficult to calculate. L-moments are similar to conventional moments but can be projected by a linear combination of order statistics (Mackenzie and Winterstein, 2011). Statistical analyses are used in many scientific fields, some of these may be simple and easy to analyze or some may be complex due to some characteristics like a large range, skewness, or variation. Thus, the disparity between the upper values and the lower values is very common in many fields. So, should choose the best methods to describe and analyze the data for getting the best estimate. LH-moments are suggested by Wang (1997) which could have extra robust properties in comparison to L-moments for describing the greater events in numbers and the upper part of distributions. In other words, LH-moments are considered a generalization of L-moment. Therefore, for extreme values, LH-moments are also studied for approximating the parameters of distributions.

Using the LH-moments in the estimation of predictions leads to reduce the undesirable influence of the sample's small quantities (Zakaria et al., 2018). The result of LH-moment estimation of a four-parameter kappa distribution for estimating heavy-tail quantiles suggests that when the L-moment method did not succeed in giving a suitable solution, the LH-moment approach is useful in handling data following a four-parameter kappa distribution (Murshed et al., 2014). In a similar study, LH-moment estimation of Wakeby distribution was used with hydrological applications (Busababodhin et al. 2016).

A comparative case study was done by Bora et al. (2017) for analysis rainfall data of the northeast region of India using L-moment and LH-moment, the result of the comparative explained that the L1-moment method is significantly more effective for analysis of the data than the other methods which used in Park et al. (2009).

Accordingly, this study initiated the derivation parameters of three-parameter kappa Type-II (K3D-II) distribution using LH-moments. The study also focused on assesses the performance of LH-moments compared to the conventional L-moments.

2. Methodology

2.1. Three-parameter kappa distribution

The Three-Parameter Kappa Type-II distribution (K3D) was proposed by Busababodhin et al. (2016) and it is considered as a special type of a four parameters kappa distribution (Hosking, 1994; Park and Kim, 2007). This distribution has three parameters which are μ, α and β, where μ is a location, α is a shape and β is a scale for $\mu \leq min_{1\leq i\leq n}(x_i)$, $\alpha > 0, \beta > 0$ and $x > 0$.

The probability distribution function and quantile function of the K3D is given by the following equations:

$$F(x) = \left(\frac{x-\mu}{\beta}\right)\left(\alpha + \left(\frac{x-\mu}{\beta}\right)^{\alpha}\right)^{-\frac{1}{\alpha}} \tag{1}$$

$$x(F) = \mu + \beta\left(\frac{\alpha F^{\alpha}}{1-F^{\alpha}}\right)^{\frac{1}{\alpha}}, \qquad 0 < F < 1 \tag{2}$$

2.2. L-Moments of K3D distribution

Hosking developed the L- moment theory based on order statistics and he defined the L-moments are the linear combinations of probability-weighted moments (PWMs) (Hosking and Wallis, 1997). Greenwood et al. (1979) defined and summarized the theory of PWMs as:

$$\beta_r = \int_0^1 x(F)F^r dF \tag{3}$$

The first four L-moment are defined as:

$$\begin{aligned} \lambda_1 &= \beta_0, \\ \lambda_2 &= 2\beta_1 - \beta_0, \\ \lambda_3 &= 6\beta_2 - 6\beta_1 + \beta_0, \\ \lambda_4 &= 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \end{aligned} \tag{4}$$

The L-moment ratios defined and calculated by Park et al. (2009) as:

$$\tau_2 = \frac{\lambda_2}{\lambda_1}, \tag{5}$$

$$\tau_3 = \frac{\lambda_3}{\lambda_2}, \tag{6}$$

$$\tau_4 = \frac{\lambda_4}{\lambda_2} \tag{7}$$

With

$$\tau_r = \frac{\lambda_r}{\lambda_2}, r \geq 3$$

Now, use L-moments to estimate the parameters of K3D distribution as:

$$\hat{\beta} = \frac{l_2}{\hat{\alpha}^{\frac{1}{\hat{\alpha}}-1}\left(2B\left(\frac{3}{\hat{\alpha}}, 1-\frac{1}{\hat{\alpha}}\right) - B\left(\frac{2}{\hat{\alpha}}, 1-\frac{1}{\hat{\alpha}}\right)\right)} \tag{8}$$

$$\hat{\mu} = l_1 - \hat{\beta}\hat{\alpha}^{\frac{1}{\hat{\alpha}}-1} B\left(\frac{2}{\hat{\alpha}}, 1 - \frac{1}{\hat{\alpha}}\right) \tag{9}$$

and the values of α are found by solving the equation

$$\tau_3 = \frac{6B\left(\frac{4}{\alpha}, 1-\frac{1}{\alpha}\right) - 6B\left(\frac{3}{\alpha}, 1-\frac{1}{\alpha}\right) + B\left(\frac{2}{\alpha}, 1-\frac{1}{\alpha}\right)}{2B\left(\frac{3}{\alpha}, 1-\frac{1}{\alpha}\right) - B\left(\frac{2}{\alpha}, 1-\frac{1}{\alpha}\right)} \tag{10}$$

because the equation of τ_3 depends on α only, given that the values of τ_3 are known.

2.3. LH-moments of K3D distribution

LH-moments are considered as a generalization of L-moments; based on linear combinations of higher-order statistics for describing the larger events in the data and the upper part of distributions. The relationship between LH-moments and PWMs according to (Murshed et al., 2014) can be written as:

$$\lambda_1^{\eta} = B_{\eta}$$
$$\lambda_2^{\eta} = \frac{1}{2!}(\eta+2)B_{\eta+1} - B_{\eta} \quad (11)$$
$$\lambda_3^{\eta} = \frac{1}{3!}(\eta+3)\left[(\eta+4)B_{\eta+2} - 2(\eta+3)B_{\eta+1} + (\eta+2)B_{\eta}\right]$$
$$\lambda_3^{\eta} = \frac{1}{4!}(\eta+4)\left[(\eta+6)(\eta+5)B_{\eta+3} - 3(\eta+5)(\eta+4)B_{\eta+2} + 3(\eta+4)(\eta+3)B_{\eta+1} - (\eta+3)(\eta+2)B_{\eta}\right]$$

where β_r is called the standard PWMs and can be written as:

$$B_r = \frac{\int_0^1 x(F)F^r dF}{\int_0^1 F^r dF} = (r+1)\int_0^1 x(F)F^r dF = (r+1)\beta_r \quad (12)$$

The LH-moments ratios are written by:

$$\tau_3 = \frac{1}{3!}\frac{(\eta+3)}{(\eta+2)}\left(\frac{(\eta+4)(\eta+3)B\left(\frac{\eta+4}{\alpha},1-\frac{1}{\alpha}\right)-2(\eta+3)(\eta+2)B\left(\frac{\eta+3}{\alpha},1-\frac{1}{\alpha}\right)+(\eta+2)(\eta+1)B\left(\frac{\eta+2}{\alpha},1-\frac{1}{\alpha}\right)}{(\eta+2)B\left(\frac{\eta+3}{\alpha},1-\frac{1}{\alpha}\right)-(\eta+1)B\left(\frac{\eta+2}{\alpha},1-\frac{1}{\alpha}\right)}\right) \quad (13)$$

$$\tau_4 = \frac{2}{4!}\frac{(\eta+4)}{(\eta+2)}\left(\frac{(\eta+6)(\eta+5)(\eta+4)B\left(\frac{\eta+5}{\alpha},1-\frac{1}{\alpha}\right)-3(\eta+5)(\eta+4)(\eta+3)B\left(\frac{\eta+4}{\alpha},1-\frac{1}{\alpha}\right)+3(\eta+4)(\eta+3)(\eta+2)B\left(\frac{\eta+3}{\alpha},1-\frac{1}{\alpha}\right)-(\eta+3)(\eta+2)(\eta+1)B\left(\frac{\eta+2}{\alpha},1-\frac{1}{\alpha}\right)}{(\eta+2)B\left(\frac{\eta+3}{\alpha},1-\frac{1}{\alpha}\right)-(\eta+1)B\left(\frac{\eta+2}{\alpha},1-\frac{1}{\alpha}\right)}\right) \quad (14)$$

Now, the parameters of K3D distributions can be estimated using different LH-moments levels as:

$$\hat{\beta} = \frac{l_2^{\eta}}{\frac{(\eta+2)}{2!}\hat{\alpha}^{\frac{1}{\hat{\alpha}}-1}\left[(\eta+2)B\left(\frac{\eta+3}{\alpha},1-\frac{1}{\alpha}\right)-(\eta+1)B\left(\frac{\eta+2}{\alpha},1-\frac{1}{\alpha}\right)\right]} \quad (15)$$

$$\hat{\mu} = \hat{\lambda}_1^{\eta} - (\eta+1)\hat{\beta}\hat{\alpha}^{\frac{1}{\hat{\alpha}}-1}B\left(\frac{\eta+2}{\hat{\alpha}},1-\frac{1}{\hat{\alpha}}\right) \quad (16)$$

and the values of α are found by solving the equation of τ_3 because the equation of τ_3 depends on α only, given that the values of τ_3 are known.

In this study, parameter estimates for LH-moments with different levels (η=0, 1, 2, 3, and 4) are derived in order to investigate their performance in estimating large quantile data. L1-, L2-, L3- and L4-moments are used to represent LH-moments for the different levels of η=0, 1, 2, 3, and 4 respectively.

3. Case study

This study utilized the annual maximum rainfall series obtained from the Department of Drainage and Irrigation System Terengganu of Embong station located in Terengganu, which is on the east coast of Peninsular Malaysia. The data contains measurements of daily rainfalls in millimeters from the year 2011 until 2016. Terengganu has a strong tropical monsoon climate, with relatively uniform temperatures within the range of 21°C to 32°C. The weather is dry and warm from January till April with humidity in the lowlands consistently high, between 82-86 percent annually. The mean rainfalls of Embong station are 98.894 mm with a standard deviation of 75.355mm. The skewness and kurtosis are 1.493 and 1.807 respectively. The L- and LH-moments ratios of different levels of LH-moments were calculated and summarized in Table 1.

4. Results and discussion

The analysis aims to investigate the performance of L-moments compared to LH-moments in estimating rainfall frequency analysis by fitting the data to the K3D distribution. The effects of the different levels of LH-moments are also observed by utilizing data from the lower and upper parts of the distribution.

The performances of L- and LH-moments are measured based on Mean Absolute Deviation Index (MADI) and Mean Square Deviation Index (MSDI) as follows:

$$MADI = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - x(F_i)}{x_i}\right| \quad (17)$$
$$MSDI = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - x(F_i)}{x_i}\right)^2 \quad (18)$$

The criterion of MADI and MSDI were calculated for K3D distribution using L- and LH-moments for a complete data series of $0 \le F \le 1$, lower part distribution of $0.3 \le F \le 1$ and upper part distribution of $0.6 \le F \le 1$ and $0.9 \le F \le 1$.

Tables 2 and 3 present the values of MADI and MSDI for $0 \le F \le 1$, $0.3 \le F \le 1$, $0.6 \le F \le 1$ and $0.9 \le F \le 1$ of K3D distribution, respectively. The bold values indicate the smallest values for each column. Based on these tables, clearly showed that L-moments methods have the smallest MADI and MSDI values for the complete data whereas LH-moments have the smallest values of MADI and MSDI for the larger values of data. Hence, the K3D distribution using L-moments is the best methods for a complete data, while L4-Moments method is the best at the upper part of the distribution at $0.6 \le F \le 1$ and $0.9 \le F \le 1$ and can be noted that L4-moments at $0.9 \le F \le 1$ is better than L4-Moments at $0.6 \le F \le 1$. In this case, it seems to suggest that the LH-moments method at higher order could improve the estimation of rainfall at larger return periods for the K3D distribution.

The performance of different levels of LH-moments in estimating the real data can be illustrated by fitting the K3D distribution to the annual maximum rainfall series of Embong station as shown in Fig. 1. Based on Fig. 1, results of a high level of LH-moments (L3 and L4) are poor for the lower part of distribution but perform better for the upper part of the distribution. This indicates that the curves fitted by a higher level of LH-moments give a better estimation for the higher quantiles. Therefore, LH-moments gave more satisfying results than L-moment in high quantile estimation.

Table 1: L- and LH-moments

Methods	L-	L1-	L2-	L3-	L4-
t_2	0.397	0.397	0.397	0.397	0.397
t_3	0.352	0.352	0.352	0.352	0.352
t_4	0.182	0.182	0.182	0.182	0.182

Table 2: Values of MADI of K3D distribution

Methods	$0 \le F \le 1$	$0.3 \le F \le 1$	$0.6 \le F \le 1$	$0.9 \le F \le 1$
L-	**0.0850**	**0.0620**	0.0460	0.0160
L1-	0.1290	0.0640	0.0440	0.0140
L2-	0.2390	0.0740	0.0400	0.0110
L3-	0.4090	0.1010	0.0360	0.0090
L4-	0.6240	0.1430	**0.0320**	**0.0070**

Table 3: Values of MSDI of K3D distribution

Methods	$0 \le F \le 1$	$0.3 \le F \le 1$	$0.6 \le F \le 1$	$0.9 \le F \le 1$
L-	**0.0090**	**0.0070**	0.0050	0.0020
L1-	0.0280	0.0070	0.0050	0.0020
L2-	0.1420	0.0100	0.0050	0.0010
L3-	0.4810	0.0250	0.0040	0.0009
L4-	1.1670	0.0630	**0.0040**	**0.0007**

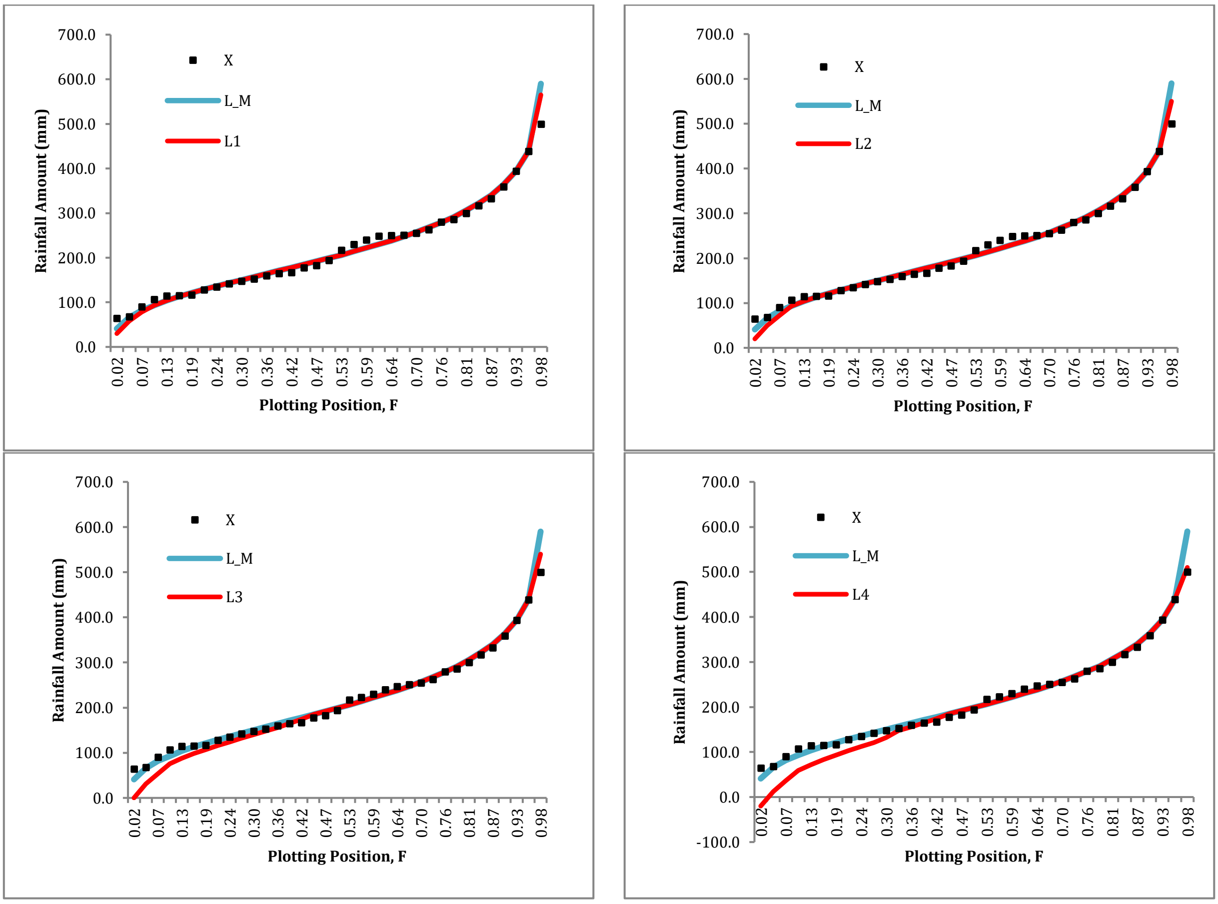

Fig. 1: Fitting the K3D distribution of monthly maximum rainfall data of Embong station for different levels of L-moments and LH-moments. Y-axis is the amount of rainfall (mm) and X-axis is a plotting position

5. Conclusion

The paper describes briefly the study carried out for estimation of K3D distribution by adopting LH-moments in analyzing rainfall data of Embong Station in Terengganu, Malaysia. The following conclusions are drawn from the study:

i. The study presents the selection of suitable levels of LH-moments in estimating large quantile data evaluated by MADI and MSDI tests.

ii. The LH-moments method at higher-order $0.9 \le F \le 1$ could improve the estimation of rainfall at larger return periods for the K3D distribution

iii. High levels of LH-moments (L3 and L4) are poor for the lower part of distribution but perform better for the part of the distribution.

iv. This study suggested that LH-moments gave more satisfying results than L-moment in high quantile

Acknowledgment

The authors would like to thank the Department of Irrigation and Drainage, Ministry of Natural

Resources and Environment, Malaysia for providing the rainfall data and Center for Research Excellence and Incubation Management, Universiti Sultan Zainal Abidin, Malaysia for funding this research.

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Bhat MS, Alam A, Ahmad B, Kotlia BS, Farooq H, Taloor AK, and Ahmad S (2019). Flood frequency analysis of river Jhelum in Kashmir basin. Quaternary International, 507: 288-294. https://doi.org/10.1016/j.quaint.2018.09.039

Bora DJ, Borah M, and Bhuyan A (2017). Regional analysis of maximum rainfall using L-moment and LH-moment: A comparative case study for the northeast India. Mausam, 68(3): 451-462. https://doi.org/10.54302/mausam.v68i3.677

Busababodhin P, Am Seo Y, Park JS, and Kumphon BO (2016). LH-moment estimation of Wakeby distribution with hydrological applications. Stochastic Environmental Research and Risk Assessment, 30(6): 1757-1767. https://doi.org/10.1007/s00477-015-1168-4

Greenwood JA, Landwehr JM, Matalas NC, and Wallis JR (1979). Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. Water Resources Research, 15(5): 1049-1054. https://doi.org/10.1029/WR015i005p01049

Harun NA, Makhtar M, Aziz AA, Mohamad M, and Zakaria ZA (2017). Implementation of apriori algorithm for a new flood area prediction system. Advanced Science Letters, 23(6): 5419-5422. https://doi.org/10.1166/asl.2017.7390

Hosking JR (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. Journal of the Royal Statistical Society: Series B (Methodological), 52(1): 105-124. https://doi.org/10.1111/j.2517-6161.1990.tb01775.x

Hosking JR (1994). The four-parameter kappa distribution. IBM Journal of Research and Development, 38(3): 251-258. https://doi.org/10.1147/rd.383.0251

Hosking JRM and Wallis JR (1997). Regional frequency analysis: An approach based on L-moments. Cambridge University Press, Cambridge, UK. https://doi.org/10.1017/CBO9780511529443

MacKenzie CA and Winterstein SR (2011). Comparing L-moments and conventional moments to model current speeds in the North Sea. In the 61st Annual IIE Conference and Expo Proceedings, Reno, USA.

Makhtar M, Harun NA, Abd Aziz A, Zakaria ZA, Abdullah FS, and Jusoh JA (2016). An association rule mining approach in predicting flood areas. In the International Conference on Soft Computing and Data Mining, Springer, Bandung, Indonesia: 437-446. https://doi.org/10.1007/978-3-319-51281-5_44

Murshed MS, Am Seo Y, and Park JS (2014). LH-moment estimation of a four parameter kappa distribution with hydrologic applications. Stochastic Environmental Research and Risk Assessment, 28(2): 253-262. https://doi.org/10.1007/s00477-013-0746-6

Murshed MS, Seo YA, Park JS, and Lee Y (2018). Use of beta-P distribution for modeling hydrologic events. Communications for Statistical Applications and Methods, 25(1): 15-27. https://doi.org/10.29220/CSAM.2018.25.1.015

Park JS and Kim TY (2007). Fisher information matrix for a four-parameter kappa distribution. Statistics and Probability Letters, 77(13): 1459-1466. https://doi.org/10.1016/j.spl.2007.03.002

Park JS, Seo SC, and Kim TY (2009). A kappa distribution with a hydrological application. Stochastic Environmental Research and Risk Assessment, 23(5): 579-586. https://doi.org/10.1007/s00477-008-0243-5

Rao AR and Hamed KH (2000) Flood frequency analysis. CRC Press, Boca Raton, USA.

Samantaray S and Sahoo A (2021). Estimation of flood frequency using statistical method: Mahanadi River basin, India. H2Open Journal, 3(1): 189-207. https://doi.org/10.2166/h2oj.2020.004

Sharafi M, Rezaei A, and Tavangar F (2021). Comparing estimation methods for the three-parameter kappa distribution with application to precipitation data. Hydrological Sciences Journal, 66(6): 991-1003. https://doi.org/10.1080/02626667.2021.1903010

Wang QJ (1997). LH moments for statistical analysis of extreme events. Water Resources Research, 33(12): 2841-2848. https://doi.org/10.1029/97WR02134

Zakaria ZA, Suleiman JMA, and Mohamad M (2018). Rainfall frequency analysis using LH-moments approach: A case of Kemaman Station, Malaysia. International Journal of Engineering Technology, 7(2): 107-110. https://doi.org/10.14419/ijet.v7i2.15.11363

Zhou CR, Chen YF, Huang Q, and Gu SH (2017). Higher moments method for generalized Pareto distribution in flood frequency analysis. In the IOP Conference Series: Earth and Environmental Science, IOP Publishing, Qingdao, China, 82: 012031. https://doi.org/10.1088/1755-1315/82/1/012031