

LQ-moments: Parameter Estimation for Kappa Distribution (LQ-momen: Anggaran Parameter untuk Taburan Kappa)

ANI SHABRI* & ABDUL AZIZ JEMAIN

ABSTRACT

Identification of the true statistical distributions for various hydrologic data sets is a major problem facing engineers. The four-parameter kappa distribution is a combination of the established distribution including the Generalised Extreme Value (GEV), Generalised Logistic (GL), Generalised Pareto (GP) and the Gumbel distribution were considered in this study. The main objective of this study was to develop the method of LQ-moments for the kappa distribution. The performance of the LQ-moments was compared with L-moments through eight problems using published data sets. The results showed that the performance of both methods, the LQ-moments and L-moments worked equally well.

Keywords: Kappa distribution; kernel quantile; L-moments; LQ-moments

ABSTRAK

Mengenalpasti taburan yang sebenar untuk pelbagai set data hidrologi merupakan masalah utama yang dihadapi oleh jurutera. Taburan kappa empat parameter adalah gabungan taburan yang terkenal termasuk taburan GEV, GL, GP dan Gumbel dipertimbangkan dalam kajian ini. Objektif utama kajian ini adalah untuk membangunkan kaedah LQ-momen untuk taburan kappa. Keupayaan kaedah LQ-momen dibandingkan dengan kaedah L-momen melalui lapan contoh permasalahan menggunakan set data yang telah diterbitkan. Hasil kajian menunjukkan keupayaan kaedah LQ-momen adalah sama dengan kaedah L-momen.

Kata kunci: Kuantil kernel; L-momen; LQ-momen; taburan kappa

INTRODUCTION

The modelling of extreme events such as floods, wind storms or heavy rains is essential in the design of water-related structures, in agriculture, weather modification, monitoring climate changes and flood-plain management. Knowledge related to the distribution of extreme events is of great importance for the design of water-related structure. Furthermore, there is no universal distribution to fit all the extreme observations. One of the extreme value distributions for the maximum values is the four parameter kappa distribution. The four-parameter Kappa distribution introduced by Hosking (1994) is considered because several established distributions including the GEV distribution, generalised logistic (GL) distribution, generalised Pareto (GP) distribution and Gumbel distribution which have been used for the modelling of extreme events, are special cases of the Kappa distribution (Park & Park 2002a).

Many parameter estimation methods have been proposed to fit statistical distribution to hydrological data. The current method of estimation uses linear functions of expected order statistics, namely L-moments estimates. For the estimation of parameter Kappa, the method of L-moments estimation procedure has been used (Hosking 1990). The main advantages of using the method of L-moments are that the parameter estimates are more reliable (i.e. smaller mean-squared error of estimation) than the method of moments estimates, particularly from small

samples, and are usually computationally more tractable than maximum likelihood estimates. However, the method of L-moments estimation of Kappa still requires Newton-Raphson iteration. Moreover, the parameter space is restricted to ensure the existence of the L-moments and the uniqueness of parameters. Therefore the routine provided by Hosking (1999) sometimes fails to get the method of L-moments estimates for some data sets, because no unique solution exists inside of the restricted region. Parida (1999) used the kappa distribution with the method of L-moments estimates in modeling Indian summer monsoon rainfall.

Park and Jung (2002b) attempted to use kappa distribution with the maximum likelihood estimates on the summer extreme daily rainfall over South Korea. Singh et al. (2003) developed an entropy-based method for estimating parameters of the four-parameter kappa distribution. The entropy-based method is evaluated and compared with the method of moments, L-moments, and the maximum likelihood estimation using four sets of data on annual maximum rainfall and on annual peak flow discharge. The results of estimation show that the entropy or maximum likelihood method and L-moments method enable the four-parameter Kappa distribution to fit the data well and are comparable, whereas the method of moment estimator worked poorly for all cases.

Mudolkar and Hutson (1998) extended L-moments to new moment like entities called LQ-moments. They found

LQ-moments always exists, are often easier to compute than L-moments, and in general behave similarly to the L-moments. Ani and Jemain (2006) developed and improved the method of LQ-moments for the Extreme Values Type 1 (EV1) distribution. The performances of the proposed method were compared with the estimators based on method of L-moments, moments and maximum likelihood. The simulations studies show that the method of LQ-moments outperforms the method of maximum likelihood over the entire sample size n considered but always perform better than the L-moments and moments methods. Wan Zin et al. (2009) determined the best fitting distribution of annual maximum rainfall in Peninsular Malaysia based on methods of L-moment and LQ-moment.

The objective of this study was to develop the method of LQ-moments for the four-parameter kappa distribution. The popular quantile estimator namely the weighted kernel quantile (WKQ) estimator was proposed to estimate the quantile function. The performances of the proposed estimators of the kappa distribution were compared with the estimators based on conventional L-moments, using data on annual maximum of daily precipitation, annual maximum peak flow and annual maximum gust wind speed.

FOUR PARAMETERS KAPPA DISTRIBUTION

The cumulative distribution function of the kappa distribution is

$$F(x) = \left\{ 1 - h \left[1 - \frac{h}{\delta} (x - \omega) \right]^{1/k} \right\}^{1/h} \quad (1)$$

and the probability density function is

$$f(x) = \frac{1}{\delta} \left[1 - \frac{k}{\delta} (x - \omega) \right]^{1/k-1} \left\{ 1 - h \left[1 - \frac{k}{\delta} (x - \omega) \right]^{1/k} \right\}^{1/h-1} \quad (2)$$

where ω is a location parameter, δ is a scale parameter, and h and k are shape parameters. The quantile function (inverse cumulative distribution function) for the kappa distribution is expressed as

$$x(F) = \omega + \frac{\delta}{k} \left[1 - \left(\frac{1-F^h}{h} \right)^k \right] \quad (3)$$

DEFINITION OF LQ-MOMENTS ESTIMATORS

The r th LQ-moments ξ_r was defined by Mudolkar and Hutson (1998) as

$$\xi_r = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} \tau_{p,\alpha}(X_{r-k:r}), \quad r = 1, 2, \dots \quad (4)$$

where

$$\begin{aligned} \tau_{p,\alpha}(X_{r-k:r}) &= pQ_{X_{r-k:r}}(\alpha) + (1-2p)Q_{X_{r-k:r}}(1/2) + pQ_{X_{r-k:r}}(1-\alpha) \\ &= pQ[B_{r-k:r}^{-1}(\alpha)] + (1-2p)Q[B_{r-k:r}^{-1}(1/2)] + pQ[B_{r-k:r}^{-1}(1-\alpha)] \end{aligned} \quad (5)$$

$\tau_{p,\alpha}(X_{r-k:r})$ is the quick estimator of location and $B_{r-k:r}^{-1}(\alpha)$ is the quantile of a beta random variable with parameter $r - k$ and $k + 1$, and $Q(\cdot)$ denotes the quantile estimator. The first four LQ-moments of the random variable X are defined as

$$\xi_1 = \tau_{p,\alpha}(X), \quad (6)$$

$$\xi_2 = \frac{1}{2} [\tau_{p,\alpha}(X_{2:2}) - \tau_{p,\alpha}(X_{1:2})], \quad (7)$$

$$\xi_3 = \frac{1}{3} [\tau_{p,\alpha}(X_{3:3}) - 2\tau_{p,\alpha}(X_{2:3}) + \tau_{p,\alpha}(X_{1:3})], \quad (8)$$

$$\xi_4 = \frac{1}{4} [\tau_{p,\alpha}(X_{4:4}) - 3\tau_{p,\alpha}(X_{3:4}) + 3\tau_{p,\alpha}(X_{2:4}) - \tau_{p,\alpha}(X_{1:4})]. \quad (9)$$

The ratios of LQ-moments are defined as

$$\eta_2 = \frac{\xi_2}{\xi_1}, \eta_3 = \frac{\xi_3}{\xi_2}, \eta_4 = \frac{\xi_4}{\xi_2} \quad (10)$$

where η_2 (LQ-CV), η_3 (LQ-Skewness), and η_4 (LQ-Kurtosis) are alternate measures of coefficient of variation, skewness (CS), and kurtosis (Ck), respectively. Given a ranked sample of size n , $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$, the r th sample LQ-moments is given by

$$\hat{\xi}_r = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} \hat{\tau}_{p,\alpha}(X_{r-k:r}), \quad r = 1, 2, \dots \quad (11)$$

where

$$\begin{aligned} \hat{\tau}_{p,\alpha}(X_{r-k:r}) &= p\hat{Q}_{X_{r-k:r}}(\alpha) + (1-2p)\hat{Q}_{X_{r-k:r}}(1/2) + p\hat{Q}_{X_{r-k:r}}(1-\alpha) \\ &= p\hat{Q}[B_{r-k:r}^{-1}(\alpha)] + (1-2p)\hat{Q}[B_{r-k:r}^{-1}(1/2)] + p\hat{Q}[B_{r-k:r}^{-1}(1-\alpha)] \end{aligned} \quad (12)$$

is the quick estimator of location and $B_{r-k:r}^{-1}(\alpha)$ is the quantile of a beta random variable with parameter $r - k$ and $k + 1$, and $\hat{Q}(\cdot)$ denotes the sample quantiles estimators.

In our study, the approximation of the quantile estimator, $\hat{Q}(\cdot)$ called the weighted kernel quantile estimator (WKQ) is proposed. The WKQ is given by

$$\hat{Q}(u) = \sum_{i=1}^n \left[n^{-1} K_h \left(\sum_{j=1}^i w_{j,n} - u \right) \right] X_{i:n} \quad 0 < u < 1 \quad (13)$$

where

$$K_h(\bullet) = (1/h) K(\bullet/h)$$

$$w_{i,n} = \begin{cases} \frac{1}{2} \left(1 - \frac{n-2}{\sqrt{n(n-1)}} \right), & i = 1, n \\ \frac{1}{\sqrt{n(n-1)}}, & i = 2, 3, \dots, n-1. \end{cases} \quad (14)$$

and $K(t) = (2\pi)^{-1/2} \exp(-t^2/2)$ is the Gaussian Kernel, $h = [u(1-u)/n]^{1/2}$ is an optimal bandwidth (Sheather & Marron 1990).

METHODS OF PARAMETER ESTIMATION

METHOD OF LQ-MOMENTS

Parameter estimates are obtained in LQ-moments, as in the case of L-moments, by equating moments of the distributions with the corresponding sample moments. The resulting equations are then solved simultaneously for the unknown parameters. From equations (6)–(12) and equation (3) for quantile function, $Q(u)$ of Kappa, the expressions for the LQ-moments of the Kappa distribution are given as follow

$$\xi_1 = \omega + \frac{\delta}{k} t_{p,\alpha}(X_{1:1}) \quad (15)$$

$$\xi_2 = \frac{\delta}{2k} [t_{p,\alpha}(X_{2:2}) - t_{p,\alpha}(X_{1:2})] \quad (16)$$

$$\eta_2 = \frac{\delta}{2k} \frac{[t_{p,\alpha}(X_{2:2}) - t_{p,\alpha}(X_{1:2})]}{\omega + \frac{\delta}{k} t_{p,\alpha}(X_{1:1})} \quad (17)$$

$$\eta_3 = \frac{\frac{1}{3}[t_{p,\alpha}(X_{3:3}) - 2t_{p,\alpha}(X_{2:3}) + t_{p,\alpha}(X_{1:3})]}{\frac{1}{2}[t_{p,\alpha}(X_{2:2}) - t_{p,\alpha}(X_{1:2})]} \quad (18)$$

$$\eta_4 = \frac{\frac{1}{4}[t_{p,\alpha}(X_{4:4}) - 3t_{p,\alpha}(X_{3:4}) + 3t_{p,\alpha}(X_{2:4}) - t_{p,\alpha}(X_{1:4})]}{\frac{1}{2}[t_{p,\alpha}(X_{2:2}) - t_{p,\alpha}(X_{1:2})]} \quad (19)$$

where

$$t_{p,\alpha}(X_{r-k:r}) = pQ_0[B_{r-k:2}^{-1}(\alpha)] + (1-2p)Q_0[B_{r-k:r}^{-1}(1/2)] + pQ_0[B_{r-k:r}^{-1}(1-\alpha)] \quad (20)$$

and

$$Q_0(F) = 1 - \left(\frac{1-F^h}{h} \right)^k. \quad (21)$$

In both equations (18) and (19), η_3 and η_4 are estimated by sample LQ-moments. Therefore, equations (18) and (19) constitute a system of two simultaneous non-linear equations in terms of k and h , whose solution will yield those particular values of k and h parameters of kappa distribution by the LQ-moments method. The other parameter δ and ω then can be obtained from equations (15) and (17), also constitute a system of two simultaneous non-linear equations. These equations can be solved by Newton-Raphson iteration.

METHOD OF L-MOMENTS

Hosking (1990) defined the r th L-moments λ_r as

$$\lambda_r = \sum_{j=0}^r (-1)^{r-j} \binom{r}{j} \binom{r+1}{j} \beta_j, \quad r = 0, 1, 2, \dots \quad (22)$$

where $\beta_r = \int_0^1 x^r F^r dF$. The probability-weighted moments (PWM) of the kappa distribution for are given by Hosking (1994) as

$$r\beta_{r-1} = \xi + \frac{\alpha}{k} \left[1 - \frac{r\Gamma(1+k)\Gamma(r/h)}{h^{1+k}\Gamma(1+k+r/h)} \right], \quad h > 0; \quad k > -1 \quad (23)$$

$$r\beta_{r-1} = \xi + \frac{\alpha}{k} \left[1 - r^{-k}\Gamma(1+k) \right], \quad h = 0; \quad k > -1 \quad (24)$$

$$r\beta_{r-1} = \xi + \frac{\alpha}{k} \left[1 - \frac{r\Gamma(1+k)\Gamma(-k-r/h)}{(-h^{1+k})\Gamma(1-r/h)} \right], \quad h < 0; -1 < k < -1/h \quad (25)$$

and for

$$r\beta_{r-1} = \xi + \alpha [\gamma + \log h + \Psi(1+r/h)], \quad h > 0 \quad (26)$$

$$r\beta_{r-1} = \xi + \alpha [\gamma + \log r], \quad h = 0 \quad (27)$$

$$r\beta_{r-1} = \xi + \alpha [\gamma + \log(-h) + \Psi(-r/h)], \quad h < 0 \quad (28)$$

where $\gamma = 0.5772$ is Euler's constant, and ψ is the digamma function. Estimation of the parameter kappa distribution using L-moments requires the solution of equations (23)–(28) with β_r estimated by $b_r = \frac{1}{n} \sum_{i=1}^n \frac{(i-1)(i-2)\dots(i-r)}{(n-1)(n-2)\dots(n-r)} X_{r:n}$. No explicit solution is possible, but the equations can be solved by Newton-Raphson iteration (Hosking 1994).

APPLICATIONS

The parameter estimation of a distribution can be evaluated and compared using real-world data or Monte Carlo simulated data and both are accepted techniques and used in hydrology. However the conclusion reached about the performance of a parameter estimation method using Monte Carlo simulated data can be significantly different from those reached using real world data, as noted by Haktanir (1992). The objective here was to show the LQ-moments method to estimate the parameters of kappa distribution and provide a limited comparison with other methods using real-world data.

To illustrate the use of LQ-moments for the fitting a Kappa distribution, four sets of annual maximum peak flow for the Beskonak river in Southern Anatolia (Haktanir 1997), Spey river and Kelvin river in Scotland (Ahmad et al. 1988), Feather River in Oroville (Mudholkar & Hutson 1998) and Huai river in China (Song & Ding 1998), two sets of annual maximum of daily precipitation in Kwangju and Kangnung, Korea (Park & Park 2002) and a set of annual maximum gust wind speed occurring in Hong Kong (Xiao et al. 2006) were employed.

It is acknowledged that these data sets are not sufficiently large to reach definitive conclusions but will be suffice to demonstrate the LQ-moments method, which is the objective of this study. The name and characteristics of the data are illustrated in Table 1.

The relative error (RR) was used as performance index for comparing the methods of LQ-moments and L-moments of parameter estimation. The RR in the computed distribution fit to the observed data was defined as

$$RR = \frac{1}{n} \sum_{i=1}^n \left(\frac{x(F_i) - x_c(F_i)}{x(F_i)} \right)^2$$

in which $x(F_i)$ and $x_c(F_i)$ denote the observed and computed values, respectively, for a given value of F_i . The observed values corresponds to a return period which was computed by using the Gringorton plotting position formula

$$F_i = (i - 0.44)/(n + 0.12)$$

where i is the rank assigned to each data point in the sample with a value of one for the lowest value, two for the second smallest and so on, with n for the highest value.

Initially, parameters of LQ-moments method were estimated using combinations of the quick estimators parameters (α and p) values in the ranges 0 to 0.5. In the computer iterations the values of α and p were chosen in small steps by adding 0.01, and all possible combination of α and p were examined in order to find the best combination in terms of RR.

The optimal combinations of α and p for LQ-moments method, the values of the parameters of Kappa distribution and the values of RR based on the two different estimation methods, obtained from all data sets are summarized in Table 2.

Table 2 shows that the optimal combination of α and p for LQ-moments method for each cases, are different. For data at Kwangju, L-moments of Kappa distribution fails to give valid estimates for this data because no unique solution exists.

Observed and computed frequency curve of the Kappa distributions with the four parameters estimated by L-moments and LQ-moments for eight the eight data sets are plotted in Figures 1(a-h). The observed data values are plotted against the corresponding EV1 reduced variates $-\log(-\log F_i)$ using Gringorton plotting position formula.

These figures show that the L-moments and LQ-moments are in close agreement and it is difficult to distinguish which method is the best. However in terms of RR, LQ-moments are equivalent to the L-moments for

TABLE 1. The Name and characteristics of data used in parameter estimation

Case	Name	n	Data mean	Standard Deviation	Skewness		Kurtosis	
					Cs	τ_3	Ck	τ_4
1	Spey	31	145.303	67.258	1.821	0.402	6.395	0.242
2	Kelvin	35	81.611	16.448	0.732	0.150	3.278	0.162
3	Feather	59	70265.085	52023.524	1.039	0.232	3.628	0.109
4	Beskonak	51	888.6176	465.687	1.343	0.232	5.218	0.180
5	Huai	31	830.452	544.245	1.082	0.282	3.263	0.154
6	Kwangju	61	116.802	49.878	1.718	0.223	7.746	0.252
7	Kangnung	88	135.833	64.758	0.956	0.222	3.235	0.136
8	Hong Kong	45	38.878	10.201	1.127	0.275	3.401	0.174

TABLE 2. Parameters and RR of L-moments and LQ-moments in different cases

Case	Method	LQ-Moments Parameter			Distribution Parameter			RR
		α	p	k	h	δ	ω	
1	LQ-Moments	0.175	0.25	-0.339	0.232	34.757	105.480	0.0020
	L-Moments			-0.210	0.722	44.534	93.895	0.0020
2	LQ-Moments	0.25	0.45	-0.053	-0.761	9.989	77.234	0.0016
	L-Moments			-0.039	-0.266	11.681	75.987	0.0002
3	LQ-Moments	0.20	0.35	0.246	0.977	73990.0	8702.0	0.0047
	L-Moments			0.178	0.842	68755.4	15641.2	0.0086
4	LQ-Moments	0.20	0.35	0.144	0.493	472.351	541.824	0.0031
	L-Moments			-0.104	-0.044	317.604	675.788	0.0033
5	LQ-Moments	0.075	0.15	0.194	1.082	731.765	163.286	0.0080
	L-Moments			0.007	0.675	538.887	360.980	0.0067
6	LQ-Moments	0.10	0.35	-0.137	-0.411	29.737	97.991	0.0041
	L-Moments			-	-	-	-	-
7	LQ-Moments	0.11	0.28	0.010	0.305	60.330	93.716	0.0029
	L-Moments			0.058	0.471	64.489	87.927	0.0030
8	LQ-Moments	0.10	0.25	-0.119	-0.372	6.642	35.285	0.0028
	L-Moments			-0.069	0.378	8.267	32.061	0.0015

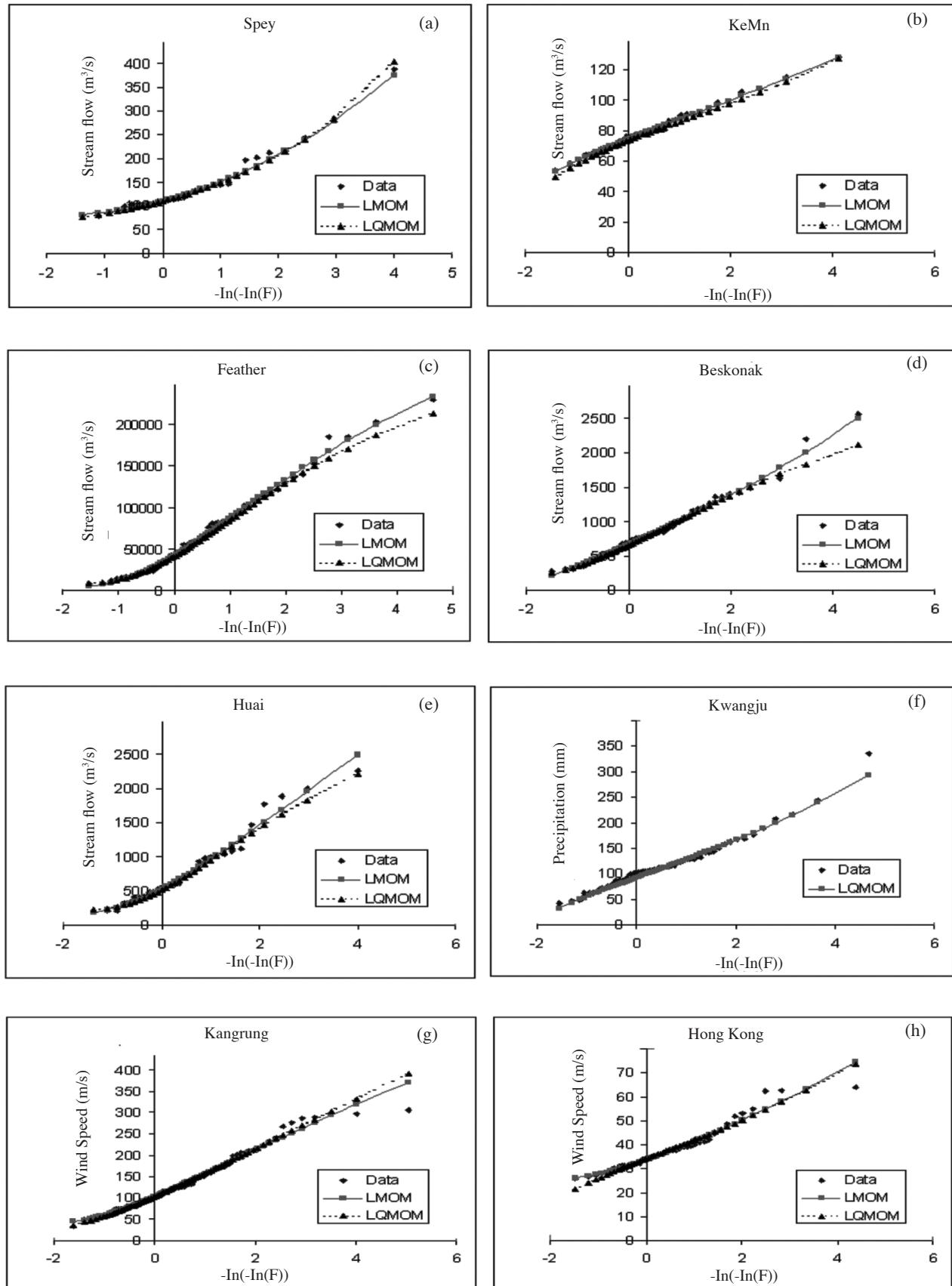


FIGURE 1. Comparison of observed and computed frequency curves of theoretical distribution with data for (a) Spey, (b) KeMn, (c) Feather, (d) Beskonak, (e) Huai, (f) Kwangju, (g) Kangrung and (h) Hong Kong

data in case 1, produced the best estimates for data in cases 3, 4 and 7 while L-moments gave the best estimates for data in cases 2, 5, 8.

CONCLUSIONS

The method of LQ-moments was used to derive estimators of parameters and quantiles of the four-parameter Kappa distribution. The results obtained are compared with those obtained by using the method of L-moments. The Kappa distribution which has received only limited attention from the hydrologic community, can be widely used because several established distributions including the distribution of GEV, GL, and Gumbel are special cases of the Kappa distribution. Because of its useful applications, its parameters need to be evaluated precisely, accurately and efficiently. We investigate the performance of these estimators using two data sets on annual maximum of daily precipitation, five data sets on annual maximum peak flow and a data set on annual maximum gust wind speed.

The results show the methods of LQ-moments and L-moments are comparable and enable the four-parameter kappa distribution to fit the data well. The LQ-moments in general behave similarly to the L-moments and can always be calculated whereas the method of L-moments sometimes fails to give estimates of kappa distribution.

REFERENCES

- Ahmad, M.I., Sinclair, C.D. & Werritty, A. 1988. Log-Logistic flood frequency analysis. *Journal of Hydrology* 98: 205-224.
- Ani, S. & A.A. Jemain. 2006. LQ-moments: application to the Extreme Value Type I distribution. *Journal of Applied Sciences* 6(5): 993-997.
- Haktanir, T. 1997. Self-determined probability-weighted moments method and its application to various distributions. *Journal of Hydrology* 194: 180-200.
- Hosking, J.R.M. 1990. L-moments: Analysis and estimation of distribution using linear combinations of order statistics. *J. Roy. Statist. Soc. B.* 52: 105-124.
- Hosking, J.R.M. 1994. The four-parameter Kappa distribution. *IBM J. Res. Dev.* 38(3): 251-258.
- Hosking, J.R.M. 1999. LMOMENTS: Fortran routines for L-moment computations and regional frequency analysis, Version 3.02.
- Mudholkar, G.S. & Hutson, A.D. 1998. LQ-moments: analogs of L-moments. *Journal of Statistical Planning and Inference* 71: 191-208.
- Parida, B.P. 1999. Modelling of Indian summer monsoon rainfall using a four-parameter Kappa distribution. *Int. Journal Climate* 19: 1389-1398.
- Park, J.S. & Jung, H.S. 2002a. Modelling Korea extreme rainfall using a Kappa distribution and maximum likelihood estimate. *Theor. Appl. Climatol.* 72: 55-64.
- Park, J.S. & Park, B.J. 2002b. Maximum likelihood estimation of the four-parameter Kappa distribution using the penalty method. *Computers & Geosciences* 28: 65-68.
- Sheather, S.J. & Marron, J.S. 1990. Kernel quantile estimators. *Journal of the American Statistical Association* 85: 410-416.
- Singh, V.P., Asce, F. & Deng, Z.Q. 2003. Entropy-based parameter estimation for Kappa distribution. *Journal of Hydrologic Engineering* 8(2): 81-92.
- Song, D. & Ding, J. 1988. The application of probability weighted moments in estimating the parameters of the Pearson type three distribution. *Journal of Hydrology* 101: 47-61.
- Wan Zin, W.Z., Jemain, A.A. & Ibrahim, K. 2009. The best fitting distribution of annual maximum rainfall in Peninsular Malaysia based on method of L-moment and LQ-moment. *Theoretical and Applied Climatology* 96(3-4): 337-344.
- Xiao, Y.Q., Li, Q.S., Li, Z.N., Chow, Y.W. & Li, G.Q. 2006. Probability distribution of extreme wind speed and its occurrence interval. *Engineering Structures* 28: 1173-1181.

Ani Shabri*

Department of Mathematics
Universiti Teknologi Malaysia
81300 Johor, Malaysia

Abdul Aziz Jemain
School of Mathematical Sciences
Universiti Kebangsaan Malaysia
43600 Bangi, Selangor Malaysia

*Corresponding author; email: ani@utm.my

Received: 13 August 2009

Accepted: 21 March 2010