TL-moments and L-moments Estimation of the Generalized Logistic Distribution

Ani Shabri (Corresponding author)
Faculty of Science, Universiti Teknologi Malaysia
81310 UTM Skudai, Johor Darul Takzim, Malaysia
Tel: 60-7-553-4238   E-mail: ani@mel.fs.utm.my

Ummi Nadiah Ahmad
Faculty of Science, Universiti Teknologi Malaysia
81310 UTM Skudai, Johor Darul Takzim, Malaysia
Tel: 60-1-7936-8096   E-mail: umminadiahahmad@gmail.com

Zahrahtul Amani Zakaria
Faculty of Informatics, Universiti Sultan Zainal Abidin
PO BOX 21300, Kuala Terengganu, Terengganu, Malaysia
Tel: 60-1-7974-2046   E-mail: amanizakaria@gmail.com

The research is financed by Kementerian Pengajian Tinggi (KPT) Malaysia and Ministry of Science, Technology and Innovation (MOSTI) Malaysia, under Vot 79346

Abstract
The generalized logistic (GLO) distribution has been used widely in extreme value event evaluation and also popular in hydrological risk analysis. In estimating the high return period events, censoring the data from below might be advantageous since the small floods are less significant to large ones, so the used of small floods can sometimes be only a nuisance value. In this paper the method of trimmed L-moments with one smallest value were trimmed (TLMOM1) was introduced as an alternative ways in estimating the flood for higher return period. TLMOM1 has an ability to reduce undesirable influence of small sample might have compared to former TL-moments (TLMOM) and L-moments (LMOM) method. The main objective of this study is to derive the TLMOM1 for GLO distribution. The performance of TLMOM1 was compared with LMOM and TLMOM through Monte Carlo simulation and stream flows data over station in Terengganu, Malaysia. The result shows that in certain cases, TLMOM1 is a better option as compared to LMOM and TLMOM in modelling those series.

Keywords: L-moments, TL-moments, Generalized Logistic Distributions, Parameter Estimation

1. Introduction
Flood events are part of nature which always exist and continue to exist. Flood, also known as deluge is a natural disaster that could diminish properties, plants and even cause thousands of people homeless. Since flood can cause a hazard towards human, so preventing flood are very important. However, in constructing the flood protection project, the information on flood magnitude and their frequencies are very important (Noto and Logia, 2009). Flood frequency analysis (FFA) is the most suitable approach for this case (Stedinger et al., 1992).

The frequency analysis is the estimation of how often a specified event will occur (Hosking and Wallis, 1997). The probability for future events can be predicted by fitting the past observations to selected probability distributions. However, the data analysis often requires estimation of parameters for the probability distributions. L-moments (LMOM) is a recent development in mathematical statistics which facilitates the estimation process in frequency analysis (Noto and Logia, 2009). LMOM which introduced by Hosking (1990), possess several advantages over any other product moments. The ratio estimators of location, scale and shape are nearly unbiased, for any probability distribution which the observation arise. It is also are more robust against the extreme values and exist whenever the mean of the distribution exists. Hosking (1990) presents the LMOM estimators for some distribution and in some cases, he shows that LMOM gives a better fit compared to other methods.

The method of LMOM has been used rapidly by many researchers; see for example Parida and Moalafi (2008), Kumar et al. (2003) and Betul Saf (2009). LMOM has been defined by Hosking (1990) as a linear combination of probability weighted moments (PWM’s). He developed the theories of LMOM from the order statistics. Let $X_1, \ldots, X_r$ be a conceptual
random sample of size $r$, and $X_{1,r} \leq X_{2,r} \leq \ldots \leq X_{r,r}$ denote the corresponding order statistics. The $r^{th}$ L MOM define by Hosking (1990) is:

$$
\lambda_r = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E(X_{r-k,r})
$$

(1)

$E(X_{r,r})$ can be written as

$$
E(X_{r,r}) = \frac{r!}{(i-1)!(r-i)!} \int_0^1 x(F)F^{i-1}(1-F)^{r-i-1} dF
$$

(2)

In order to select a reliable design quantile, the statistical method used in parameter estimation are very important. However, the standard estimation methods such as maximum likelihood method and least square method can be influenced by the outliers, so a robust estimation method which developed to reduce the influence of outliers on the final estimates should be used (Elamir and Seheult, 2003). As an alternative to this problem, a predetermine percentage of the extreme values from the sample needs to be trimmed before estimating the mean and standard deviation from untrimmed sample values. In 2003, Elamir and Seheult introduced an extension of LMOM called trimmed L-moments (TLMOM) where they trim one smallest and largest value from the conceptual sample. They denote this as $\lambda_r^{(1:2)}$:

$$
\lambda_r^{(1:2)} = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E(X_{r-k+1,r+1})
$$

(3)

From Eq. 3, clearly TLMOM reduce to LMOM when $t_1 = 0$ and $t_2 = 0$. TLMOM equation with $t_1 = 1$ and $t_2 = 1$ are defined by Elamir and Seheult (2003) as $\lambda_r^{(1:1)}$:

$$
\lambda_r^{(1:1)} = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E(X_{r+1-k,r+2})
$$

(4)

Related issues in TLMOM method is the choice of amount of trimming. An annual maximum series of floods over a period is often treated as a random sample from a specified homogeneous probability distribution such as generalized logistic (GLO) distribution and others. The purpose of analyzing annual maximum series of floods usually is to predict the magnitude of flood of relatively large return period events. Usually, when analyzing the arid or semiarid regions, many very low or even zero annual maximum flows occur (Bhattarai, 2004). As a solution, Cunnane (1987) and Wang (1990) suggested that by censoring the data from below might be advantageous since it might be that smaller sample values have only a nuisance value in the context of upper quantile estimation.

In this paper, we proposed the trimmed L-moments with one smallest value were trimmed (TLMOM1) from the conceptual sample. TLMOM1 assign zero weight on smallest extreme value which give them an ability to reduce undesirable influence of small sample might have. Estimation of the GLO distribution by using TLMOM1 is formulated. The performance of the TLMOM1 was compared with LMOM and TLMOM through Monte Carlo simulation and the stream flows data. The $r^{th}$ TLMOM1, $\lambda_r^{(1:0)}$ can be defined as:

$$
\lambda_r^{(1:0)} = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E(X_{r+1-k,r+1})
$$

(5)

2. Generalized Logistic Distribution

The generalized logistic (GLO) distribution is widely being used in extreme value event evaluation, mostly in recent years. GLO and generalized extreme value (GEV) distribution were found to be an appropriate distributions for extreme flood events in Sarawak region (Lim and Lye, 2003) and United Kingdom (Ashkar et. al, 2006). The probability density function of GLO distribution was given by:

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

(6)
Where $\zeta$ is a location, $\alpha$ is a scalar, and $k$ is a shape parameter. The corresponding cumulative distribution function is:

$$F(x) = \left[ 1 + \left\{ 1 - k \left( \frac{x - \zeta}{\alpha} \right)^{\frac{1}{k}} \right\} \right]^{-1}$$  \hspace{1cm} (7)

The value of $x$ can be in the range of $\zeta + \frac{\alpha}{k} \leq x < \infty$ for $k < 0$ and $-\infty \leq x < \zeta + \frac{\alpha}{k}$ for $k > 0$. The quantile function for GLO distribution can be written as:

$$x(F) = \zeta + \frac{\alpha}{k} \left[ 1 - \left\{ 1 - \frac{F}{1-F} \right\}^{k} \right]$$  \hspace{1cm} (8)

3. Methodology

3.1 LMOM of the GLO Distribution

The theoretical LMOM for real valued random variable $x$ with quantile function $x(F)$ are derived from the expectations of order statistics (Hosking, 1990). The order statistics of sample size $r$ are formed by ascending order $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$. From Eq. 1, the first four LMOM can be derived as follow:

$$\lambda_1 = E(X_{1:1})$$  \hspace{1cm} (9)

$$\lambda_2 = \frac{1}{2} E(X_{2:2} - X_{1:2})$$  \hspace{1cm} (10)

$$\lambda_3 = \frac{1}{3} E(X_{3:3} - 2X_{2:3} + X_{1:3})$$  \hspace{1cm} (11)

$$\lambda_4 = \frac{1}{4} E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4})$$  \hspace{1cm} (12)

In particular, $\lambda_1$, $\lambda_2$, $\lambda_3$ and $\lambda_4$ are population measure of the location, scale, skewness and kurtosis respectively. The sample of LMOM can be estimated unbiasedly from the sample order statistics (Asquith, 2007):

$$l_r = \frac{1}{r} \binom{n}{r} \sum_{i=0}^{r-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} \binom{i-1}{r-1-k} \binom{n-i}{k} X_{i:n}$$  \hspace{1cm} (13)

The LMOM of the GLO distribution is obtained by substituting the Eq. 8 into Eqs. 9-12. The first four sample LMOM of the GLO distribution are:

$$\lambda_1 = \zeta + \frac{\alpha}{k} \left( 1 - \Gamma(1-k)\Gamma(k+1) \right)$$  \hspace{1cm} (14)

$$\lambda_2 = \alpha \Gamma(1-k)\Gamma(k+1)$$  \hspace{1cm} (15)

$$\lambda_3 = -\alpha k \Gamma(1-k)\Gamma(k+1)$$  \hspace{1cm} (16)

$$\lambda_4 = \frac{\pi \alpha k (5k^2) + 1}{6 \sin(\pi k)}$$  \hspace{1cm} (17)

The L-skewness $\tau_3$ and L-kurtosis $\tau_4$ of GLO distribution can be defined as:

$$\tau_3 = \frac{\lambda_3}{\lambda_2} = -k$$  \hspace{1cm} (18)

$$\tau_4 = \frac{\lambda_4}{\lambda_2} = \frac{5k^2}{6} + \frac{1}{6}$$  \hspace{1cm} (19)
From Eq. 14, Eq. 15 and Eq. 18, the LMOM parameter of $\alpha, \zeta$ and $k$ can be estimated. The samples estimate for LMOM are:
\[
\alpha = \frac{l_2}{\Gamma(1-k)\Gamma(k+1)} \tag{20}
\]
\[
\zeta = l_1 + \frac{1}{k}(l_2 - \alpha) \tag{21}
\]
\[
k = -t_3 \tag{22}
\]

### 3.2 TLMOM1 of the GLO Distribution

The sample size difference is represented by a 'trim' level. The conceptual sample size will be increase from $r$ to $r + 1$ where 1 is our trimming value. Based on Eq. 5, the first four of TLMOM1 can be derived as follows:
\[
\lambda_1^{(1,0)} = E(X_{2:2}) \tag{23}
\]
\[
\lambda_2^{(1,0)} = \frac{1}{2}E(X_{3:3} - X_{2:3}) \tag{24}
\]
\[
\lambda_3^{(1,0)} = \frac{1}{3}E(X_{4:4} - 2X_{3:4} + X_{2:4}) \tag{25}
\]
\[
\lambda_4^{(1,0)} = \frac{1}{4}E(X_{5:5} - 3X_{4:5} + 3X_{3:5} - X_{2:5}) \tag{26}
\]

Elamir and Seheult (2003) present the following estimator for sample TLMOM as:
\[
\rho_r^{(1:0)} = \frac{1}{nr} \sum_{i=1}^{n} \sum_{r=0}^{n} (-1)^k \left( \begin{array}{c} r - 1 \\ k \end{array} \right) \left( \begin{array}{c} i - 1 \\ r - t_1 - k - 1 \end{array} \right) \left( \begin{array}{c} n - i \\ k + t_2 \end{array} \right) X_{i:r} \tag{27}
\]

Based on Eq. 27, the sample of $r^{th}$ TLMOM1, $(\rho_r^{(1,0)})$ can be derived as
\[
\rho_r^{(1,0)} = \frac{1}{nr} \sum_{i=1}^{n} \sum_{r=0}^{n} (-1)^k \left( \begin{array}{c} r - 1 \\ k \end{array} \right) \left( \begin{array}{c} i - 1 \\ r - k \end{array} \right) \left( \begin{array}{c} n - i \\ k \end{array} \right) X_{i:r} \tag{28}
\]

By substituting Eq. 8 into Eqs. 23-26, the first four TLMOM1 of GLO distribution can be computed as follows:
\[
\lambda_1^{(1,0)} = \zeta - \frac{\alpha(1-k)\Gamma(k+1)-1}{k} + \alpha(1-k)\Gamma(k+1) \tag{29}
\]
\[
\lambda_2^{(1,0)} = \frac{-3\alpha(1-k)\Gamma(k+1)}{2} + \frac{3\alpha(1-k)\Gamma(k+1)(k-3)}{4} \tag{30}
\]
\[
\lambda_3^{(1,0)} = \frac{10\alpha}{3k} + \frac{10\alpha(1-k)\Gamma(k+1) - 1}{3k} + \frac{8\alpha(1-k)\Gamma(k+1)(k-3)}{3} - \frac{5\alpha(4-k)\Gamma(k+1)}{9k} + 2\alpha(1-k)\Gamma(k+1) \tag{31}
\]
\[
\lambda_4^{(1,0)} = \frac{25\alpha(4-k)}{\Gamma(k+1)} - \frac{8k - 10\alpha(1-k)\Gamma(k+1) - 1}{k} - \frac{5\alpha(1-k)\Gamma(k+1)}{2} - \frac{25\alpha(1-k)\Gamma(k+1)(k-3)}{4} - \frac{10\alpha}{k} - \frac{35\alpha(5-k)\Gamma(k+1)}{96k} \tag{32}
\]
The TL-Skewness $\tau_3^{(1,0)}$ and TL-Kurtosis $\tau_4^{(1,0)}$ for GLO distribution are defined as:

$$\tau_3^{(1,0)} = \frac{\lambda_3^{(1,0)}}{\lambda_2^{(1,0)}} = \frac{4}{27} - \frac{20k}{27}$$  \hspace{1cm} (33)

$$\tau_4^{(1,0)} = \frac{\lambda_4^{(1,0)}}{\lambda_2^{(1,0)}} = \frac{35k^2}{72} - \frac{5k}{24} + \frac{5}{36}$$  \hspace{1cm} (34)

The parameter of $\alpha$, $\zeta$ and $k$ of TLMOM1 can be estimated as follows:

$$\alpha = \frac{4F_1^{(1,0)}}{-6\Gamma(1-k)\Gamma(k+1) - 3\Gamma(1-k)\Gamma(k+1)(k-3)}$$  \hspace{1cm} (35)

$$\zeta = F_1^{(1,0)} + \frac{\alpha\Gamma(1-k)\Gamma(k+1)-1}{k} - \alpha\Gamma(1-k)\Gamma(k+1)$$  \hspace{1cm} (36)

$$k = \frac{4 - 27t_3}{20}$$  \hspace{1cm} (37)

4. Monte Carlo Simulations

In this study, Monte Carlo simulation study is conducted to compare the performance of LMOM, TLMOM and TLMOM1 in the sense of bias and root mean square error (RMSE) for different sample size $n$ and shape parameter $k$. In Monte Carlo simulation, the value of parameter of scalar and location is set as 0 and 1 (Shabri and Jemain, 2007) with shape parameter $k$ is between $(-0.4, 0.4)$ (Wang, 1990). In each simulation, 10 000 sample were generated for sample size 15, 25, and 50. The bias and RMSE value were computed using Eq. 38 and Eq. 39 as follows:

$$\text{Bias} = \frac{1}{N} \sum_{i=1}^{N} (x_i - y_i)$$  \hspace{1cm} (38)

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - y_i)^2}$$  \hspace{1cm} (39)

The bias value has been obtained for quantiles $x(F)$, $F = 0.90$, $F = 0.98$ and $F = 0.99$, i.e; $Q_{40}$, $Q_{50}$ and $Q_{100}$. The bias obtained using LMOM, TLMOM and TLMOM, were plotted against the shape parameter $k$ as illustrated in Figure 1-3. Since all sample size lead to an almost similar results, only results for sample size 25 are shown here. From Figure 1, the result shows that for quantile estimate $x(F = 0.90)$, TLMOM produce smaller bias value at $-0.4 \leq k \leq 0.1$ while at $k \geq 0.2$, LMOM produce smaller bias value. For quantile $x(F = 0.98)$, TLMOM1 are almost unbiased for positive $k$ values. However, as $k$ goes negative, TLMOM1 results in greater bias than LMOM. For quantile $x(F = 0.99)$, TLMOM1 results in smaller bias value for all shape parameter $k$ except at $k \leq -0.3$.

The results of RMSE obtained using LMOM, TLMOM and TLMOM1 of quantile $x(F)$, $F = 0.90$, $F = 0.98$ and $F = 0.99$, were plotted against the shape parameter $k$ as illustrated in Figure 4-6. For quantile $x(F = 0.90)$, using TLMOM1 gives smaller RMSE value for all shape parameter $k$, except at $0.1 \leq k \leq 0.2$. For quantile $x(F = 0.98)$ and $x(F = 0.99)$, TLMOM1 gives larger RMSE value for $-0.2 \leq k \leq 0.1$, but have smaller RMSE value for other $k$ values compared to LMOM and TLMOM approach.

5. Data Analysis

The example are given in the following to show the effect of using different trimming value. The data used in this study comes from Sg. Kemaman at Rantau Panjang in Terengganu, Malaysia, with 34 annual maximum stream flows data. The data has been provided for this study by Department of Irrigation and Drainage, Ministry of Natural Resources and Environment, Malaysia, covering from 1976 until 2009. Table 1 shows the parameter estimates by using LMOM, TLMOM and TLMOM1, while observed and computed frequency curves are plotted in Figure 7.

The comparison was made between the method of LMOM (no data was trimmed), TLMOM (one smallest and one largest data was trimmed) and TLMOM1 (one smallest data was trimmed) to see the effectiveness of these methods in the real life. Figure 7 shows that TLMOM1 give more accurate estimation for the high quantile estimation compared to LMOM and TLMOM method. However, on the estimation of low flow quantile, LMOM and TLMOM method give more accurate
estimation. It is because, TLMOM1 gives zero weight on smallest sample value, which make TLMOM1 to be more suitable in characterizing the upper part of the distribution and larger events in the sample.

6. Conclusions
TLMOM is the robust version of LMOM where TLMOM trimmed a predetermined percentage of the extreme values from the sample. The expressions are derived to relate LMOM, TLMOM and TLMOM1 to GLO distribution. The Monte Carlo simulations was carried out to see the performance of LMOM, TLMOM and TLMOM1 involving various sample size \( n \) and different values of shape parameter \( k \) for different quantile estimators \( x(F) \).

From the results, it can be concluded that based on the RMSE and bias value obtained from the simulations, the LMOM and TLMOM1 are seems to perform better than TLMOM method. However, the strength of the method lies on the value of shape parameter \( k \). For certain \( k \) value, LMOM are performed better while at the other \( k \) values, TLMOM1 gives better result. So it is important to choose the right \( k \) values in order to produce a better estimation.

The methods of LMOM, TLMOM and TLMOM1 were applied to annual maximum stream flows series of Sg. Kemaman at Rantau Panjang in Terengganu, Malaysia. The results obtained shows that TLMOM1 gives more accurate estimation in estimating the high flow quantile compared to LMOM and TLMOM method. This study has demonstrated that conventional LMOM is not optimal for the certain cases of GLO distribution. The new derivation of TLMOM1 estimation shows a high efficiency at certain range of \( k \) as compared to LMOM. The simplicity and good performance of this method make it an attractive option in estimating the quantile of GLO distribution.

Acknowledgment
Special thanks to Kementerian Pengajian Tinggi Malaysia and Ministry of Science, Technology and Innovation (MOSTI) Malaysia for funding this research under Vot 79346. We thank Department of Irrigation and Drainage, Ministry of Natural Resources and Environment, Malaysia for providing the data for this study. Lastly, thanks are given to Universiti Teknologi Malaysia.

References


Table 1. Parameter estimates using LMOM, TLMOM and TLMOM1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>LMOM</th>
<th>TLMOM</th>
<th>TLMOM1</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>127.342</td>
<td>135.631</td>
<td>151.626</td>
</tr>
<tr>
<td>ζ</td>
<td>342.341</td>
<td>340.606</td>
<td>353.715</td>
</tr>
<tr>
<td>k</td>
<td>-0.167</td>
<td>-0.176</td>
<td>-0.025</td>
</tr>
<tr>
<td>10 Years</td>
<td>680.378</td>
<td>704.534</td>
<td>696.369</td>
</tr>
<tr>
<td>50 Years</td>
<td>1040.000</td>
<td>1099.000</td>
<td>974.050</td>
</tr>
<tr>
<td>100 Years</td>
<td>1222.000</td>
<td>1301.000</td>
<td>1093.000</td>
</tr>
</tbody>
</table>

Figure 1: Bias of $x(F = 0.90)$ estimator, fitting the GLO distribution to generated GLO samples.
Figure 2: Bias of $x(F = 0.98)$ estimator, fitting the GLO distribution to generated GLO samples.

Figure 3: Bias of $x(F = 0.99)$ estimator, fitting the GLO distribution to generated GLO samples.
Figure 4: RMSE of $x(F = 0.90)$ estimator, fitting the GLO distribution to generated GLO samples.

Figure 5: RMSE of $x(F = 0.98)$ estimator, fitting the GLO distribution to generated GLO samples.
Figure 6: RMSE of $x(F = 0.99)$ estimator, fitting the GLO distribution to generated GLO samples.

Figure 7: Fitting the GLO distribution to annual maximum flows at Sg. Kemaman, Terengganu, Malaysia