Development of L-Moment Based Models for Extreme Flood Events

1Sattar Chavoshi Borujeni and 2Wan Nor Azmin Sulaiman

1Faculty of Environmental Studies, Universiti Putra Malaysia,
43400 UPM Serdang, Selangor, Malaysia
1Soil Conservation and Watershed Research Center of Iran, Iran
E-mail: chsatar@yahoo.com

ABSTRACT

Accurate estimation of flood frequency is needed in many hydraulic designs such as dams, culverts and urban drainage systems. Different approaches were presented which use conventional moments to extract order statistics such as mean, standard deviation, skewness and kurtosis. Due to problems arising from data quality, such as short record and outliers, conventional moments are problematic. Hosking and Wallis (1997) developed L-moments which are linear combinations of order statistics. The main advantage of L-moments over conventional moments is that they suffer less from the effects of sampling variability. They are more robust to outliers and virtually unbiased for small samples. In this study the theory of L-moment and its advantage over conventional moments is discussed. Then its application on hydrology is addressed and finally as a case study, L-moment based method of regional flood frequency in central basin of Iran is mentioned. The L-moments have been used for parameter estimation, homogeneity testing and selection of the regional distribution. Records of peak floods, is analyzed using five distributions: generalized logistic, generalized extreme value, lognormal, Pearson type III and generalized Pareto. Each of these three-parameters distributions are estimated by the L-moment method. The discordancy index and homogeneity testing show that 5 out of the 7 study sites belong to a homogenous region. Based on the L-moment ratios diagram and the goodness-of-fit measure, the three-parameter lognormal distribution is identified as the most appropriate distribution in the homogeneous study region. The regional peak flood estimates for each return period are obtained based on this distribution. It is concluded that L-moment is an effective approach on hydrological statistics study in Iran, where conventional methods is problematic due to data shortage.

Keywords: Flood frequency, L-moment, Probability weighted method

INTRODUCTION

The hydrologic system embodies all of the physical processes that are involved in the conversion of precipitation to streamflow, as well as physical characteristics of the watershed and atmosphere that influence runoff generation. Flood and drought are two main features of hydrology which affect the human life. A significant aspect of flood hydrology is the
estimation of the magnitude of stream flow at various locations in a watershed resulting from a given precipitation input. Measure might be simply the magnitude associated with a particular point in time (as in flow forecasting), magnitude associated with a non-frequency based design flood (e.g., standard project or probable maximum), magnitude associated with duration (e.g., value that is exceeded, or not exceeded, of the time), or magnitude associated with a particular exceedence or non-exceedence frequency. These are the components of any design flood. Planning, designing and operation of any water resources projects such as dams, spillways, road and railway bridges, culverts, urban drainage systems, flood plain zoning and economic evaluation of flood protection projects are based on estimation of design flood.

Floods can be estimated through the use of several methods dependent upon the structure being designed or analyzed and the size of the watershed. Approaches to flood hydrology include deterministic, probabilistic, conceptual or parametric analysis. Such methods include the use of rational methods, frequency and regional analysis and hydrograph development. The reasonable estimation of flood has been remained one of the main problematic issues where hydrological data and information are limited. This is typical case in many basins of Iran, where most rivers are ungauged. This problem is intensified due to shortness of records, incomplete records, and inaccuracy of flow rating curves. In such cases, regionalization can be very helpful in pooling flood data such that design flood estimations can be made at ungauged basins. In this paper the application of L-moment theory in regional flood frequency analysis of North-Karoon catchment is presented. The main objective of this research is providing flood quantiles with different return periods in the study area, where faces with mentioned problems (Chavoshi et al., 1998).

Regional flood frequency analysis

One of the most important areas of flood hydrology is flood frequency analysis which involves the estimation of distributional parameters and the extrapolation of cumulative distribution functions to generate extreme flood values. The estimation of the probability distribution is affected by many complicating factors including unknown underlying flood distribution, short record lengths and different physical mechanisms of flood generation which lead to different underlying probability distributions. Regional flood frequency analysis, RFFA, allows estimating the T-year flood magnitude at any stream location within a region. It attempts to respond to the need for flood estimation in ungauged basins and for improving the at-site estimate
Development of L-Moment Based Models for Extreme Flood Events

by using the available flood data within a region. Thus, it enables flood quantile estimates for any site in a region to be expressed in terms of flood data recorded at all gauging sites in that region, including those at the specific site (Atiem and Harmancioglu, 2006). The basic idea behind regional flood frequency analysis is to make use of similarities in the characteristics of floods at different sites in a region. Consequently, regional homogeneity is an important requirement and a critical issue in such an analysis.

A homogeneous region is formed by a group of catchments that are believed to have similarities in climatic and physiographic characteristics and therefore similarity in the hydrologic flood response. Similarity between catchments is often defined, using physiographic parameters, climatic variables, and/or flood statistics, but not in terms of flood generating mechanisms (Abida and Ellouze, 2006).

The main objective of flood frequency analysis is to estimate flood quantiles at a specific site. Many procedures have been developed for flood frequency analysis purposes include: the method of residuals (MOR), the canonical correlation analysis (CCA), the region-of-influence (ROI) approach and its extensions, the hierarchical approach and its extension to ROI framework and cluster analysis. The method used for estimating the magnitude of flood quantiles depends on the availability of data, and on the form of the distribution and the estimation procedure used (Akhmad, 1993). Recent researches on flood suggest regional analysis to make flood frequency estimates at sites that have no flow records or to improve estimates at sites with relatively short records.

In such cases, floods information is transferred from gauged to ungauged sites within a pre-specified hydrologically homogeneous region. The statistical estimation of design floods by regionalization was first presented by Dalrymple,(1960) then developed by others. It involves two major steps of classifying sites into similar hydrological regions, and regional estimation of flood quantiles at the site of interest. Stedinger and Tasker,(1985) developed the method of residuals which is used by the USGS for regionalization. Bhaskar et al., (1989) and Bhaskar and O’Connor,(1989) compared the results by the method of residuals and by the clustering algorithms for data from Kentucky. Although results defined by cluster analysis were different with those obtained by the method of residuals, they show hydrologic response much better than the USGS regions. A review of different methods of index flood was conducted by Bocchiola et al., (2003).
L-MOMENTS APPROACH

The approach based on the theory of L-moments was firstly proposed by Wallis, (1989) then developed by Hosking and Wallis, (1997). L-moments are summary statistics for probability distributions and data samples. They are analogous to ordinary moments, because of providing measures of location, dispersion, skewness, kurtosis, and other aspects of the shape of probability distributions or data samples, but are computed from linear combinations of the ordered data values. L-moments may be applied in four steps of the regional frequency analysis including screening of the data, identification of homogeneous regions, choice of a frequency distribution and estimation of the frequency distribution (Hosking and Wallis, 1997). The main advantages of L-moments over conventional (product) moments are that they are able to characterize a wider range of distributions, and (when estimated from a sample) are less subject to bias in estimation and more robust to the presence of outliers in the data. The latter is because ordinary moments (unlike L-moments) require inversion of the data which causes disproportionate weight to be given to the outlying values. The identification of a distribution from which the sample was drawn is more easily achieved (particularly for skewed distributions) using L-moments than conventional moments. Basically, L-moments are linear functions of probability weighted moments (PWMs).

\[ \beta_r = E\{X[F(x)]^r \} \]  \hspace{1cm} (1)

Where \( F(x) \) is the cumulative distribution function of \( x \). The first four L-moments expressed as linear combinations of PWM are:

\[ \lambda_1 = \beta_0 \]  \hspace{1cm} (2)
\[ \lambda_2 = 2\beta_1 - \beta_0 \]  \hspace{1cm} (3)
\[ \lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \]  \hspace{1cm} (4)
\[ \lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \]  \hspace{1cm} (5)
\[ \tau_2 = \frac{\lambda_2}{\lambda_4} \]  \hspace{1cm} (6)
\[ \tau_3 = \frac{\lambda_3}{\lambda_2} \]  \hspace{1cm} (7)
\[ \tau_4 = \frac{\lambda_4}{\lambda_2} \]  \hspace{1cm} (8)
The L-mean, $\lambda_1$, is a measure of central tendency and the L-standard deviation, $\lambda_2$, is a measure dispersion. Their ratio, $\frac{\lambda_2}{\lambda_1}$, is termed the L-coefficient of variation, $\tau_2$. The ratio $\frac{\lambda_3}{\lambda_2}$ is referred to as $\tau_3$ or L-skewness, while the ratio $\frac{\lambda_4}{\lambda_2}$ is referred to as $\tau_4$ or L-kurtosis.

A Homogeneity Test

In an attempt to regional frequency analysis, many studies have been carried out to identify homogeneous regions. There are three principal methods commonly used to delineate a study region into homogeneous regions. The first is based on geographical or administratively defined regions such as provincial, national, rivers, valley and latitude/longitude boundaries. The second is based on the similarity of basin and climate characteristics such as geology, land use, drainage characteristics and rainfall/runoff similarity. The third is based on flood characteristics such as the homogeneity test of Dalrymple, (1960), similarity of the coefficient of variation (CV) test within the region (Cunnane, 1987) or a heterogeneity measure based on L-moments (Hosking, 1993).

Discordance Test

A test of discordance is an L-moments-based test to determine whether a particular site, which doesn’t appear to belong to the group of ($T_3$, $T_4$) points on the L-moment diagram, should be removed from the study area. In other words, a point which is far from the center of the diagram can be discarded. Hosking and Wallis, (1993) defined the discordance measure statistic, D-statistic, as:

$$\bar{U} = \frac{1}{N} \sum_{i=1}^{N} U_i$$  \hspace{1cm} (9)

$$S = \frac{1}{N} \sum_{i=1}^{N} (U_i - \bar{U})(U_i - \bar{U})^T$$  \hspace{1cm} (10)

$$D_1 = \frac{1}{3} (U_1 - \bar{U})^T S^{-1} (U_1 - \bar{U})$$  \hspace{1cm} (11)
Where $N$ is the total number of sites and $U_i$ is defined as a vector containing the L-moment ratios for site $i$. $\bar{U}$ and $S$ are the group averages and sample covariance matrix of $U_i$, respectively. Generally, sites with D-statistics greater than 3 are considered to be discordant from the rest of the region (Wallis, 1989).

**Heterogeneity Test**

For testing the regional homogeneity, a test statistic $H$, termed as heterogeneity measure was proposed by Hosking and Wallis, (1993). It compares the inter-site variations in sample L-moments for the group of sites with what would be expected of a homogeneous region. The inter-site variation of L-moment ratio is measured as the standard deviation of the at-site L-CV’s weighted proportionally to the record length at each site. To establish what would be expected of a homogeneous region, simulations are used. A number of 500 data regions are generated based on the regional weighted average statistics using a four parameter distribution i.e Kappa distribution. The inter-site variation of each generated region is computed and the mean (µ_r) and standard deviation σ_r of the computed inter-site variation is obtained. Then, heterogeneity measure $H$ is computed as below (Kumar et al., 2003).

\[
H = V - \frac{\mu_v}{\delta_v}
\] (12)

and

\[
V_1 = \sum_{i=1}^{N} \left\{ n_i \left[ \left( \text{Lcv}_i - \bar{\text{Lcv}} \right)^2 \right]^{\frac{3}{2}} \right\} / \sum_{i=1}^{N} n_i
\] (13)

\[
V_2 = \sum_{i=1}^{N} \left\{ n_i \left[ \left( \text{Lcv}_i - \bar{\text{Lcv}} \right)^2 + \left( \tau_{3i} - \bar{\tau}_3 \right)^2 \right]^{\frac{3}{2}} \right\} / \sum_{i=1}^{N} n_i
\] (15)

\[
V_3 = \sum_{i=1}^{N} \left\{ n_i \left[ \left( \tau_{3i} - \bar{\tau}_3 \right)^2 + \left( \tau_{4i} - \bar{\tau}_4 \right)^2 \right]^{\frac{1}{2}} \right\} / \sum_{i=1}^{N} n_i
\] (16)
The criteria established through Monte Carlo experiments by Hosking and Wallis, (1993) for assessing heterogeneity of a region is as follows:

- If $H < 1$, Region is acceptably homogeneous
- If $1 \leq H < 2$, Region is possibly heterogeneous
- If $H \geq 2$, Region is definitely heterogeneous

**Goodness-of-Fit Test for Selecting the Best Distribution**

This test is employed to select one of various unimodal distributions and to estimate its parameters. The goodness-of-fit criterion, $Z$-statistic, is defined as:

$$Z = \left( T_4 - \overline{T}_4 + \beta_4 \right) / \delta_4 \quad (17)$$

$$\beta_4 = \frac{1}{N_{sim}} \sum_{m=1}^{N_{sim}} \left( T_{4m} - T_4 \right) \quad (18)$$

$$\delta_4 = \sqrt{\frac{1}{N_{sim} - 1} \sum_{m=1}^{N_{sim}} \left( \overline{T}_{4m} - T_4 \right)^2 - N_{sim} \beta_4^2} \quad (19)$$

Where $\beta_4$ and $\delta_4$ are the bias and standard deviation of $T_4$ respectively. $N_{sim}$ is the number of simulated regional data sets generated using a kappa distribution (Adamowski, 2000). The goodness-of-fit measure, $Z$, judges how well the simulated L-skewness and L-kurtosis values obtained from the observed data (Isameldin et al., 2006). A probability distribution with the smallest value of $|Z|$ is considered the best choice among possible distributions. At the significance level of $\alpha = 0.10$, the critical value of $Z$ is 1.64, i.e., if a probability distribution whose $|Z| \leq 1.64$, then it is assessed to be an acceptable distribution for representing sample data at $\alpha = 0.10$.

**L-moments Ratio Diagram**

An L-moment ratio diagram of L-kurtosis versus L-skewness compares sample estimates of the dimensionless ratios $T_3$ and $T_4$ with their population counterparts for a range of statistical distributions. L-moment diagrams are useful for discerning grouping of sites with their similar flood frequency
behavior and identifying the statistical distribution likely to adequately describe this behavior (Pearson, 1991).

APPLICATION

Data Set and Some Basic Characteristics of Study Area

North – Karoon basin, located between 49°34’ to 51°47’ longitudes and 31°18’ to 32°40’ latitude, with 14476 km² area is the main sub-basin of the Great Karoon basin. It has a humid to semi-humid climate, high annual precipitation and associated discharge. There are six main rivers in the region including Behesht-Abad, Koohrang, Sabz-Kooh, Vanak, Bazoft and Lordegan (Figure 1). Most part of the region is occupied with high-elevated mountains. The minimum, maximum and average elevation of the study area is 780m, 4221m and 2279m, respectively. The longest river in the catchment with 137 km length has divided it into western and eastern parts. Main annual precipitation is due to the humid currents of Mediterranean, which affects the region for eight months. The long-term maximum and minimum annual rate of precipitation is about 1600mm and 300mm, respectively. The average annual precipitation of the region is 707mm while about 75% of the study area receives more than 500mm annual rainfall and only 25% of the catchment has more than 800mm annual rainfall. The mean annual volume of precipitation is 10.2 billion cubic meters, which varies between 6.8 to 12.4. Mean annual evaporation is 1620mm while in some years it gets to 2200mm in south of the region. There are a number of 30 hydrometric sites in the study area each of which with more than five years annual peak flood data. Among them, sites which are independent i.e. not subjected to upper catchments or any practices were selected for this study (Table 1). As the table below shows the area ranges between 397 to 34221 square kilometers. Moreover there is 1180m difference between the elevations of sites in the study area.

<table>
<thead>
<tr>
<th>Site</th>
<th>River</th>
<th>Area (km²)</th>
<th>Elevation (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zarinderakht</td>
<td>Khan Mirza</td>
<td>397</td>
<td>1770</td>
</tr>
<tr>
<td>Koohesookhteh</td>
<td>Kiar</td>
<td>2909</td>
<td>1980</td>
</tr>
<tr>
<td>Dezak</td>
<td>Biregan</td>
<td>630</td>
<td>2160</td>
</tr>
<tr>
<td>Tangedarkesh</td>
<td>Jooneghan</td>
<td>899</td>
<td>2000</td>
</tr>
<tr>
<td>Godarkabk</td>
<td>Agh Bolagh</td>
<td>716</td>
<td>2150</td>
</tr>
<tr>
<td>Marghak</td>
<td>Bazoft</td>
<td>34221</td>
<td>980</td>
</tr>
<tr>
<td>Lordegan</td>
<td>Lordegan</td>
<td>374</td>
<td>1650</td>
</tr>
</tbody>
</table>
L-moment Software

In order to apply the L-moment method for analysis of data in the study area, the Fortran Package programmed by Hosking, (1996) was used. This package has the ability to be applied for different stages of regional flood frequency analysis including screening of data, delineation of homogeneous region, tests of regional homogeneity, selection and estimation of regional frequency distribution and estimation of flow magnitude.

i. Checking on Homogeneity

Several methods have been proposed to identify homogeneous basins in the study area. Most of them use catchment’s attributes to delineate pooled groups. In order to delineate a pooling group, all sites that have a high similarity with the site of interest are grouped together. The Euclidean distance in a multidimensional attribute space is widely used in similarity measures. The test of homogeneity of the study area has been studied using several methods such as Cluster Analysis and Andrew’s Curve (Chavoshi et al., 1998).

ii. Discordancy and Heterogeneity Test

The discordancy measure is considered as a means of screening analysis, and the aim is to identify those sites that are grossly discordant with the group as a whole (Isameldin et al., 2006). Discordancy measure for each site of the basin have been calculated and found between 0.30 to 2.08. The data range shows no discordancy for the study sites. However the values of different heterogeneity measures, H1, H2 and H3 are found 5.93, 2.64 and 0.48, respectively. Since the first two heterogeneity measures, H1 and H2, are more than 2, deal with heterogeneity of the studied sites, two suspected sites were removed and the process repeated. The results for the five remained sites confirm that the rest of the sites may be considered as homogeneous region (Table 2).
Figure 1: Location of North Karoon Basin in Iran (above) and sites in the study area (below) of variation, L-skewness and L-kurtosis.
Development of L-Moment Based Models for Extreme Flood Events

TABLE 2: The L-moment’s properties and discordance measure.

<table>
<thead>
<tr>
<th>Site</th>
<th>N</th>
<th>Name</th>
<th>L-CV</th>
<th>L-SKEW</th>
<th>L-KURT</th>
<th>D(I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>Zarinderakht</td>
<td>0.4671</td>
<td>0.2864</td>
<td>0.1404</td>
<td>0.81</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>Koohesookhteh</td>
<td>0.3117</td>
<td>0.2744</td>
<td>0.1091</td>
<td>1.32</td>
</tr>
<tr>
<td>3</td>
<td>17</td>
<td>Dezak</td>
<td>0.4617</td>
<td>0.3993</td>
<td>0.2227</td>
<td>0.52</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>Tangedarkesh</td>
<td>0.2318</td>
<td>-0.0026</td>
<td>0.0927</td>
<td>1.28</td>
</tr>
<tr>
<td>5</td>
<td>21</td>
<td>Godarkabk</td>
<td>0.3226</td>
<td>0.3564</td>
<td>0.2551</td>
<td>1.06</td>
</tr>
</tbody>
</table>

Weighted means: 0.3594  0.2728  0.1728 -

Parameters of regional Kappa distribution

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5539</td>
<td>0.5447</td>
<td>-0.0659</td>
</tr>
</tbody>
</table>

iii. Homogeneity Testing using the Heterogeneity Measure

The heterogeneity measure compares the between-site variations in sample L-moments for the group of sites with what would be expected for a homogeneous region (Isameldin et al., 2006). Discordancy measure for each site of the basin has been calculated. The values of different heterogeneity measures $H_1$, $H_2$ and $H_3$ are found 1.94, 0.84 and 0.28, respectively. Therefore based on $H_2$ and $H_3$, the study region demonstrates acceptable homogeneity.

iv. Distribution Selection using the Goodness-of-Fit Measure

After confirming the homogeneity of the study region, an appropriate distribution needs to be selected for the regional frequency analysis. In other words, in a homogeneous region all sites should have the same population L-moments. The selection was carried out by comparing the moments of the candidate distributions to the average moments statistics derived from the regional data. The best fit to the observed data indicate the most appropriate distribution (Yongqin et al., 2006). A number of five three-parameter distributions, i.e. Generalized Logistic (GL), Generalized Extreme Value (GEV), Generalized Pareto (GP), General Normal (LNIII) and Pearson Type III (PEIII) were fitted to the region. The value of ZDIST statistic for the study area for each three-parameter distribution is presented in the table below. It can be seen that all the candidates are acceptable; however LNIII is the most appropriate.
TABLE 3: Z-statistics for various distributions

<table>
<thead>
<tr>
<th>Distribution</th>
<th>L-KURTOSIS</th>
<th>Z-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEN. LOGISTIC</td>
<td>L-KURTOSIS= 0.229</td>
<td>Z VALUE = 0.88</td>
</tr>
<tr>
<td>GEN. EXTREME VALUE</td>
<td>L-KURTOSIS= 0.199</td>
<td>Z VALUE = 0.35</td>
</tr>
<tr>
<td>LOG NORMAL III</td>
<td>L-KURTOSIS= 0.181</td>
<td>Z VALUE = 0.02</td>
</tr>
<tr>
<td>PEARSON TYPE III</td>
<td>L-KURTOSIS= 0.150</td>
<td>Z VALUE = -0.54</td>
</tr>
<tr>
<td>GEN. PARETO</td>
<td>L-KURTOSIS= 0.123</td>
<td>Z VALUE = -1.03</td>
</tr>
</tbody>
</table>

v. L-moment Ratio Diagram

In an attempt to assign the sites to regions i.e. a set of sites whose flood frequency distributions are approximately the same, the diagram of L-moment ratio were constructed (Figure 2). An advantage of L-moment diagrams is that one can compare the fit of several distributions using a single graphical instrument. Figure 2 compares the theoretical relationship between L-kurtosis and L-skewness for the Exponential, Uniform, Normal, Gumbel, Generalized Logistic (GL), Generalized Extreme Value (GEV), Log-normal III (LNIII), Pearson Type III (PEIII), Generalized Pareto (GP) and lower bound of Wakeby. Figure 2 is a plot of L-moments of stream flow in real space; hence, it is not possible to represent the Log-Pearson Type 3 distribution. In this diagram, sites are illustrated as the circle points. The L-moment ratios diagram is also a very effective tool for distribution selection. It shows that the point for regional average L-moments $\tau_3 = 0.2728$ and $\tau_4 = 0.1728$, lies closest to the LNIII, which support our selection.

vi. Regional Flood Quantile Estimation

The next step in regional flood frequency is to estimate flood quantiles in the region. In this paper flood quantiles for each distribution is presented at the 90 percent level (Table 4).

TABLE 4: Quantile estimates with different probability accepted at the 90% level

<table>
<thead>
<tr>
<th>Probability</th>
<th>0.010</th>
<th>0.020</th>
<th>0.050</th>
<th>0.100</th>
<th>0.200</th>
<th>0.500</th>
<th>0.900</th>
<th>0.950</th>
<th>0.990</th>
<th>0.999</th>
</tr>
</thead>
<tbody>
<tr>
<td>GL</td>
<td>0.014</td>
<td>0.084</td>
<td>0.203</td>
<td>0.321</td>
<td>0.479</td>
<td>0.845</td>
<td>1.799</td>
<td>2.277</td>
<td>3.753</td>
<td>7.330</td>
</tr>
<tr>
<td>GEV</td>
<td>0.069</td>
<td>0.126</td>
<td>0.223</td>
<td>0.323</td>
<td>0.466</td>
<td>0.834</td>
<td>1.851</td>
<td>2.325</td>
<td>3.614</td>
<td>6.090</td>
</tr>
<tr>
<td>LNIII</td>
<td>0.110</td>
<td>0.154</td>
<td>0.234</td>
<td>0.322</td>
<td>0.456</td>
<td>0.828</td>
<td>1.878</td>
<td>2.343</td>
<td>3.523</td>
<td>5.521</td>
</tr>
<tr>
<td>PEIII</td>
<td>0.188</td>
<td>0.207</td>
<td>0.254</td>
<td>0.319</td>
<td>0.437</td>
<td>0.820</td>
<td>1.918</td>
<td>2.360</td>
<td>3.360</td>
<td>4.754</td>
</tr>
<tr>
<td>GP</td>
<td>0.239</td>
<td>0.248</td>
<td>0.275</td>
<td>0.322</td>
<td>0.423</td>
<td>0.811</td>
<td>1.957</td>
<td>2.375</td>
<td>3.200</td>
<td>4.096</td>
</tr>
<tr>
<td>WAKEBY</td>
<td>0.162</td>
<td>0.180</td>
<td>0.230</td>
<td>0.308</td>
<td>0.448</td>
<td>0.829</td>
<td>1.894</td>
<td>2.362</td>
<td>3.480</td>
<td>5.164</td>
</tr>
</tbody>
</table>

*The most appropriate distribution in the study area
Moreover estimated parameters of each distribution are presented (Table 5). The regional parameters of the WAKEBY distribution have been included in the table because it has five parameters which are more than most of the common distributions so it can attain a wider range of distributional shapes than the common distributions. This makes the WAKEBY distribution particularly useful for simulating artificial data for use in studying the robustness, under changes in distributional form of methods of data analysis (Kumar et al., 2003).

**TABLE 5: Parameters of the distribution**

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEN. LOGISTIC</td>
<td>$\xi = 0.845$, $\alpha = 0.317$, $k = -0.273$</td>
</tr>
<tr>
<td>GEN. EXTREME VALUE</td>
<td>$\xi = 0.668$, $\alpha = 0.440$, $k = -0.154$</td>
</tr>
<tr>
<td>LOG NORMAL III</td>
<td>$\xi = 0.828$, $\alpha = 0.557$, $k = -0.568$</td>
</tr>
<tr>
<td>PEARSON TYPE III</td>
<td>$\mu = 1.000$, $\sigma = 0.692$, $\gamma = 1.640$</td>
</tr>
<tr>
<td>GEN. PARETO</td>
<td>$\mu = 0.230$, $\sigma = 0.880$, $\gamma = 0.143$</td>
</tr>
<tr>
<td>WAKEBY</td>
<td>$\xi = 0.145$, $\alpha = 1.161$, $\beta = 4.570$, $\gamma = 0.630$, $\delta = 0.026$</td>
</tr>
</tbody>
</table>

**Figure 2:** L-moment ratios of some common statistical distributions in the study area
CONCLUSION

In Iran, a number of studies have been carried out for estimation of design floods for various purposes by different organizations. The insufficient number of sites and its uneven spatial distribution, short record of flood, outliers and the years of no-flow causes these studies problematic. Furthermore, available data may not represent to the basin flood due to the changes in watershed characteristics, such as changing vegetation, and land use and urbanization. Hence, regionalization methods are applied to make estimates of flood statistics at ungauged catchments using regional characteristics. In cases where no point data are available for flood estimation, data may be used from gauged neighboring catchments or data from catchments with similar hydrologic. Previous study in North-Karoon (Chavoshi et al., 1998) addresses these problems and recommends new methodology to overcome them.

The method of L-moments was employed to identify the possible regional probability distributions of peak discharge in North-Karoon basin. Discordance measure for study sites indicates no discordancy; however the values of different heterogeneity measures show that the region is heterogeneous. So, two suspected sites were removed and the process repeated. The results for the remained sites confirm that the rest of the sites are homogeneous. In order to select the parent distribution for the study area, L-moment ratio diagram was applied. The Z-statistic was also applied to select a possible distribution type of peak discharge in the region. The possible distribution types identified by the Z-statistic are Generalized Logistic, Generalized Extreme Value, Generalized Pareto, General Normal (LNIII) and Pearson Type III. Both goodness-of-fit analysis and L-moment ratio diagram analysis indicated that the three-parameter Log Normal distribution is suitable for flood frequency analysis in the study area. Therefore, the flood quantiles for different return periods were obtained which is especially useful for water resources planning, design, and management in ungauged watersheds. The methodology used in this study can be adopted for other regions provided that sufficient flood records are available.

ACKNOWLEDGEMENTS

Most of the flow data used in this study was supplied by Water Resources Research Company (TAMAB), Iran. The useful comments of two anonymous reviewers are gratefully acknowledged.
REFERENCES


