International Journal of Engineering

Journal Homepage: www.ije.ir

Assessment of Goodness of Fit Methods in Determining the Best Regional Probability Distribution of Rainfall Data

B. Amirataee*, M. Montaseri, H. Rezaei

Department of Water Engineering, Urmia university, Urmia

PAPER INFO

Paper history:
Received 18 November 2013
Received inrevised form 03 June 2014
Accepted 26 June 2014

Keywords:
Probability Plot Correlation Coefficient
L-moments
Goodness of Fit

ABSTRACT

One of the main steps in regional rainfall analysis is to determine the most appropriate of several potentially possible probability distributions of rainfall data. For this purpose, the chi-square, the Kolmogrov-Smirnov and the Probability Plot Correlation Coefficient (PPCC) methods as goodness of fit tests are usually used. Recently, L-moment ratio diagrams have been recommended to verify the goodness of fit of various probability distributions to regional hydrological data such as rainfall. Therefore, the PPCC and L-moment procedure were applied to examine the most appropriate probability distributions of regional rainfall data investigation with 95% acceptance regions in north west of Iran. For this purpose, 50 years of monthly and annual rainfall data records at 12 synoptic stations were applied based on different evaluating criteria. The results of both PPCC procedure and L-moment diagram indicate that Pearson type three probability distribution is the best probability distribution for fitting rainfall data in north west of Iran, while the L-moment approach is able to test fitness of many samples using a single diagram.

***doi**: 10.5829/idosi.ije.2014.27.10a.07*

1. INTRODUCTION

A key step in planning water resource systems is to determine the probability distribution in the statistical analysis of hydrological time series in various time scales (e.g. annually or monthly) as single site and multiple sites [1-4]. Probability distribution functions that generally are used for recorded monthly and annual data (rainfall and streamflow), include Normal, two parameter Log-Normal, three parameter Log-Normal, Pearson type III and Log-Pearson type III [5]. Studies show that there is no specific method for selecting appropriate probability distributions for hydrological data. Thus, choosing the best statistical distribution is the most important factor in frequency analysis. Therefore, different distributions must be used and then, the most appropriate distribution of data should be selected [6]. Generally, selecting the appropriate probability distribution is based on goodness of fit tests. The procedures of goodness of fit investigate the consistence of observational data with probability distribution. These methods include chi-square, Kolmogrov-Smirnov, Standard Error Estimation, Probability Plot Correlation Coefficient (PPCC), and Modified Anderson-Darling Test (AD) [7, 8]. One of the major drawbacks of Chi-Square method is that it requires abundant data [9]. Among the various methods, PPCC is believed to be the most simple and powerful method for selecting the best single site probability distribution of data [5]. PPCC test uses two different probability distributions for comparison: a) graphical method (graphical data fitting with probability distribution) and, b) numerical criteria (serial correlation between observed data and frequency coefficient of probability distribution). This method is not limited to the number of samples and thus it is independent of probability distribution function [10]. Although analytical methods, such as L-moments and Maximum Likelihood Estimations, used to fit the probability

*Corresponding Author's Email: babak.amirataee@gmail.com (B. Amirataee)

distribution of observed data are more efficient methods than graphical probability plots, many researchers use the probability plot in engineering judgment as required [10]. To evaluate the performance of PPCC test on other tests, lots of studies (Filliben[11], Looney and Gulledge, [12], Fill and Stedinger, [13]) have been done for normal and Gumbel distributions.

Since the rainfall is a regional phenomenon and all stations in a region are linked together, assessment of regional frequency analyses must be necessarily undertaken. Thereupon, another way to determine the best probability distribution for a proposed region is using L-moments ratio diagrams (Hosking [14], Hosking and Wallis [15], Stedinger, et al.[5]). Several investigators (e.g. Schaefer [16], Pearson [17], Vogel et al.[18], Chow and Watt [19], Önöz and Bayazit [20], Vogel and Wilson [21]) have used L-moments methods for selection of probability distribution of their study areas. One sensible advantage of L-moments methods over other methods is that they can present the data fitting with different distributions on the diagram as a regional indicator. Önöz and Bayazit [20] and Ben-Zvi and Azmon [22] concluded that the L-moment diagram did not help to identify the best probability distribution. Liouet al.[23] also believed that the number of samples was effective in estimating the parameter uncertainty. To solve the problem, acceptance regions were presented at 95% significance level to better detect the most probable distribution based on goodness of fit test for normal and Gumbel distributions using stochastic simulation methods. Furthermore, Wu et al.[24] proposed acceptance region of L-moments based on goodness of fit test for Pearson III distribution**.**

In this study, monthly and annual rainfall data of north west of Iran for a period of 50 years are assessed via both PPCC and L-moments methods as single and regional stations based on different evaluating criteria.

2. DATA AND METHODS

2. 1. Study Area As shown in Figure 1, time series of monthly and annual rainfall data of twelve synoptic stations located in the north west of Iran are used in this study. These rainfall data series have a 50-year period. Furthermore, general characteristics of these stations and the main parameters of these data are depicted in Table 1. Regarding the study area, the average annual rainfall varies from 1755 mm in the station of Bandar-anzali to 235 mm in Tehran station. It must be stated that these station cover semi-arid to wet climate. However, the majority of these stations do not have striking range of annual rainfall. Furthermore, an average annual rainfall of 280 mm (Tabriz station) to 503 mm (Khorramabad station) are placed in semi-arid climate [25]. In order to use rainfall data for analysis, these researchers firstly evaluated homogeneity, randomness, and stationary of these data via using appropriate statistical tests [26]. The Double Mass Curve method was used to test the homogeneity of data; the results showed homogeneity with a linear correlation coefficient of 0.99 [27]. Nonparametric Spearman's rank correlation method was used to evaluate the stationary of data. Finally, randomness was assessed via using run test [26, 28]. The results of tests indicated that randomness and stationary of data placed between the critical points for all the stations.

2. 2. Probability Plot Correlation Coefficient (PPCC) This method has been developed by Filliben[11] as a simple and powerful method to choose the best probability distribution function. Using PPCC and regarding the probability distribution function of the observed data, the value of observed data sorted in descending state (x_i) against the corresponding frequency coefficient (k_i) is plotted in a diagram, the correlation coefficient between the observed data and frequency coefficient is also calculated and the probability distribution function which has the highest correlation will be selected as the most appropriate probability distribution. Chow Equation (1) is used to estimate the value of a hydrological variable corresponding to a specific return period deviation of observed data, in which $\bar{x}$ and s_x are the mean and standard deviation, respectively. k_i and x_i are frequency coefficient and theoretical value of hydrological variables corresponding to frequency coefficient (k_i), whose value is related to the return period and the probability distribution function [29].

$$x_i = \bar{x} + k_i s_x \tag{1}$$

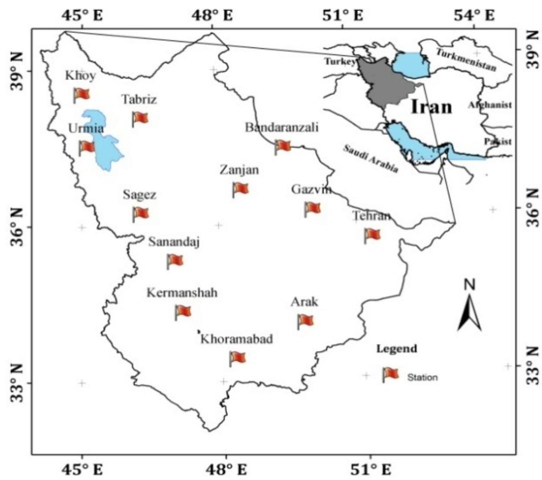

Figure 1. Locations of rainfall stations used in this study

TABLE 1. Characteristics of stations.

			Geographic coordinates		Statistical properties of annual rainfall series (1960-2010)			
No.	**Station**	**Elevation (m)**	**Latitude**	**Longitude**	**Mean (mm)**	**Coefficient of Variation**	**Skewness**	**Serial correlation Lag-1**
1	Arak	1708	34° 06′	49° 46′	332.26	0.30	0.32	-0.10
2	Urmia	1316	37° 32′	45° 05′	332.16	0.30	0.86	0.25
3	Bandar-anzali	-26	37° 28′	49° 28′	1755.46	0.19	0.76	-0.15
4	Tabriz	1361	38° 05′	46° 17′	283.22	0.30	0.95	0.36
5	Tehran	1191	35° 41′	51° 19′	235.43	0.30	0.12	-0.15
6	Khoramabad	1148	33° 26′	48° 17′	503.11	0.24	0.07	0.05
7	Khoy	1103	38° 33′	44° 58′	291.97	0.28	0.42	0.24
8	Zanjan	1663	36° 41′	48° 29′	304.77	0.26	0.00	-0.04
9	Sagez	1523	36° 15′	46° 16′	487.59	0.27	0.55	0.16
10	Sanandaj	1373	35° 20′	47° 00′	454.24	0.27	0.34	-0.01
11	Gazvin	1279	36° 15′	50° 03′	316.95	0.27	0.33	-0.27
12	Kermanshah	1319	34° 21′	47° 09′	449.92	0.27	0.59	0.11

2. 2. 1. Frequency Coefficient of Normal Distribution Joiner and Rosenblatt [30] provided a simple approximate equation for determining the standard normal distribution frequency coefficients, k_n, as a function of accumulated probability (p), that the estimated equation is as follows [31]:

$$k_n = 4.91(p^{0.14} - (1-p)^{0.14}) \tag{2}$$

This level is acceptable for $0.01 \le p \le 0.99$.

Abramowitz and Stegun [32] offered the following equations which are widely used:

$$k_n = w - \frac{2.5155 + 0.8028w + 0.0103w^2}{1 + 1.4327w + 0.1892w^2 + 0.0013w^3} \tag{3}$$

$$w = [\ln(\frac{1}{\sqrt{p^2}})]^{0.5} \qquad ,0 < p \le 0.5 \tag{4}$$

In the case of $p > 0.5, (1-p)$ is placed instead of p in Equation (4) and k_n obtained from Equation (3) is multiplied by a negative sign. Reported error for Equation (3) is less than 0.00045.To calculate the accumulated probability (p), several empirical relationships have been proposed that Blom formula (Equation (5)) is known as the most appropriate equation for Normal distribution, Log-Normal, and Log-Pearson (III) [10, 33].

$$p_i = \frac{i - 0.375}{n + 0.25} \tag{5}$$

Regarding the above equation, n is the total of observed data and $i$represents the number of rows.

2. 2. 2. Frequency Coefficient of Pearson Type III Distribution Loucks, et al. [27] proposed Wilson-Hilferty conversion for this purpose, which can convert the Normal frequency coefficient (k_n) into Pearson type III frequency coefficient (k_p), as follows:

$$k_g = \frac{2}{\delta}(1 + \frac{\delta \times k_n}{6} - \frac{\delta^2}{36})^2 - \frac{2}{\delta} \tag{6}$$

In which, k_g represents frequency coefficient of Pearson type III distribution, δ stands for Skewness and k_n shows the frequency coefficient of Normal distribution.

Wilson-Hilferty equation accuracy reduces in higher data coefficient of skewness and serial correlation coefficient. Wilson–Hilferty equation can be written according to standard normal frequency coefficient by k_g as follows:

$$k_g = \frac{6}{\delta}((\frac{k_g \times \delta + 2}{2})^{0.5} + \frac{\delta^2}{36} - 1) \tag{7}$$

2. 2. 3. Frequency Coefficient of Gumbel Distribution Vogel and Kroll [10] used the PPCC test for Gumbel distribution. Chow [19] provided the following equations for Gumbel frequency coefficient:

$$k_e = -\frac{\sqrt{6}}{\pi}[0.5772 + \ln(\ln(\frac{1}{1-p}))] \tag{8}$$

$$p_i = \frac{i - 0.44}{n + 0.12} \tag{9}$$

2. 3. L-moments L-moments are linear combinations of order statistics that are not sensitive to outliers and are unbiased for small samples of observed data. Therefore, their application to determine the best distribution function and parameter estimation seems appropriate [15]. The L-moment of each probability distribution is defined as follows:

$$L_1 = \beta_0 \tag{10}$$

$$L_2 = 2\beta_1 - \beta_0 \tag{11}$$

$$L_3 = 6\beta_2 - 6\beta_1 + \beta_0 \tag{12}$$

$$L_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \quad (13)$$

$$\beta_r = n^{-1} \sum_{j=r+1}^{n} \binom{j-1}{r} \binom{n-1}{r}^{-1} X_{(j,n)}, \quad r = 0, n-1 \quad (14)$$

Furthermore, L-moment ratios are defined as follows:

$$\tau_2 = L-CV = \frac{L_2}{L_1} \quad (15)$$

$$\tau_3 = L-Skewness = \frac{L_3}{L_2} \quad (16)$$

$$\tau_4 = L-Kurtosis = \frac{L_4}{L_2} \quad (17)$$

L-moments ratio diagrams indicate L-moment coefficient of variations (L-CV) against the L-moment coefficient of skewness (L-Skew) for two-parameter distributions and the L-moment coefficient of kurtosis (L-Kurtosis) against the L-moment coefficient of skewness (L-Skew) for three-parameter distributions. Theoretical values of two-parameter distributions [21] and three-parameter distributions [14] are obtained by the following polynomials whose coefficients are also presented in Table 2 and Table 3.

$$\tau_2 = A_0 + A_1\tau_3^1 + A_2\tau_3^2 + \ldots + A_7\tau_3^7 \quad (18)$$

$$\tau_4 = A_0 + A_1\tau_3^1 + A_2\tau_3^2 + \ldots + A_8\tau_3^8 \quad (19)$$

2. 4. Evaluation Criteria Regarding the present study, probability plots as well as two methods of the correlation coefficient and standard error are applied between observed and theoretical probability distributions values. Standard error equation is as follows:

$$S_e = \sqrt{\frac{\sum_1^n (x_i - x_j)^2}{n-1}} \quad (20)$$

In which, x_i stands for observed data, x_j shows the theoretical value of probability distribution and n is the number of data.

3. RESULTS AND DISCUSSION

In this study, the best regional distribution of annual and monthly rainfall data from twelve synoptic stations located in the north west of Iran are examined using a PPCC test and L-moments method. Probability distribution functions intended for observed monthly and annual rainfall data include the Normal distributions (N), two-parameter Log-Normal(LN2), three-parameter Log-Normal (LN3), Pearson type III (P3) and Log-Pearson type III (LP3) [5].

TABLE 2. Coefficients of polynomial approximations of two-parameter distributions.

Coeff.	Lognormal 2 (LN2)	Gamma (GAM)
A_0	0	0
A_1	1.16008	1.74139
A_2	-0.05325	0
A_3	0	-2.59736
A_4	-0.10501	2.09911
A_5	0	0
A_6	-0.00103	-0.35948
A_7	0	0

TABLE 3. Coefficients of polynomial approximations of three-parameter distributions.

Coeff.	Lognormal (LN3)	Pearson III	Normal	Gumbel
A_0	0.12282	0.12240	-	-
A_1	0	0	-	-
A_2	0.77518	0.30115	-	-
A_3	0	0	-	-
A_4	0.12279	0.95812	-	-
A_5	0	0	-	-
A_6	-0.13638	-0.57488	-	-
A_7	0	0	-	-
A_8	0.11368	0.19383	-	-
L-Skewness	-	-	0	0.16990
L-Kurtosis	-	-	0.12266	0.15004

TABLE 4. Correlation coefficients obtained from the PPCC test for annual rainfall data (maximum PPCC are underlined).

Station	Annual				
	N	P(3)	LN(2)	LN(3)	LP(3)
Arak	0.9900	<u>0.9929</u>	0.9859	0.9927	0.9922
Urmia	0.9684	0.9894	0.9917	0.9918	<u>0.9923</u>
Bandaranzali	0.9798	<u>0.9956</u>	0.9940	0.9952	0.9951
Tabriz	0.9721	<u>0.9969</u>	0.9955	0.9969	0.9968
Tehran	0.9928	0.9933	0.9842	0.9938	<u>0.9940</u>
Khoramabad	0.9959	<u>0.9960</u>	0.9850	0.9850	0.9959
Khoy	0.9889	0.9935	0.9927	0.9929	<u>0.9949</u>
Zanjan	0.9939	<u>0.9939</u>	0.9746	0.9748	0.9906
Sagez	0.9792	0.9872	0.9892	0.9874	<u>0.9895</u>
Sanandaj	0.9833	<u>0.9865</u>	0.9713	0.9851	0.9800
Gazvin	0.9923	0.9953	0.9916	0.9941	<u>0.9958</u>
Kermanshah	0.9873	<u>0.9972</u>	0.9961	0.9967	0.9967

TABLE 5. Correlation coefficients obtained from the PPCC test for monthly rainfall data (Urmia station) (maximum PPCC are underlined.

	Monthly				
Month	**N**	**P(3)**	**LN(2)**	**LN(3)**	**LP(3)**
Oct.	0.8570	0.9584	0.9695	0.9864	0.9841
Nov.	0.9421	0.9954	0.9288	0.9944	0.9843
Dec.	0.9034	0.9970	0.9915	0.9961	0.9951
Jan.	0.9719	0.9895	0.8793	0.9893	0.9810
Feb.	0.9827	0.9936	0.9721	0.9936	0.9907
Mar.	0.9614	0.9958	0.9756	0.9950	0.9951
Apr.	0.9848	0.9913	0.9776	0.9906	0.9940
May.	0.9425	0.9839	0.9600	0.9917	0.9908
Jun.	0.9301	0.9815	0.9449	0.9762	0.9773
Jul.	0.7291	0.9714	0.9536	0.9500	0.9573
Aug.	0.7165	0.9724	0.8622	0.8622	0.8866
Sep.	0.7881	0.9695	0.9364	0.9363	0.9394

TABLE 6. Standard errors obtained from the PPCC test for annual rainfall data (minimum standard errors are underlined).

Station	Annual				
	N	**P(3)**	**LN(2)**	**LN(3)**	**LP(3)**
Arak	13.87	11.80	15.74	12.07	11.69
Urmia	24.71	14.48	14.72	14.69	14.72
Bandaranzali	66.24	33.63	38.22	35.91	33.31
Tabriz	19.84	7.07	7.63	6.98	6.56
Tehran	8.52	8.23	13.69	8.38	8.40
Khoramabad	11.15	11.01	19.43	19.43	11.26
Khoy	12.13	9.36	11.33	11.07	9.51
Zanjan	8.73	8.73	14.67	8.75	9.46
Sagez	26.58	20.97	21.84	21.47	21.30
Sanandaj	22.22	20.07	21.73	19.91	22.74
Gazvin	10.73	8.47	11.59	8.50	8.32
Kermanshah	19.48	9.48	9.86	9.55	9.08

TABLE 7. Standard errors obtained from the PPCC test for monthly rainfall data (Urmia station) (minimum standard errors are underlined.

	Monthly				
Month	**N**	**P(3)**	**LN(2)**	**LN(3)**	**LP(3)**
Oct.	18.00	9.71	12.36	11.94	11.70
Nov.	10.75	3.78	6.09	4.33	7.73
Dec.	10.27	3.22	3.34	3.25	3.23
Jan.	4.35	2.79	3.82	2.76	6.99
Feb.	3.00	1.85	3.15	1.96	1.65
Mar.	8.97	3.13	4.83	3.05	3.48
Apr.	5.35	4.08	7.59	4.13	4.73
May.	12.57	6.71	10.08	8.59	5.00
Jun.	4.77	2.47	3.93	3.21	6.69
Jul.	8.03	2.80	4.23	4.23	18.18
Aug.	4.04	1.40	2.09	2.09	19.59
Sep.	4.87	1.89	2.77	2.76	17.45

PPCC test correlation coefficients and standard error values for annual of all and monthly rainfall data of e.g., Urmia station are shown in Tables 4 to 7. Box plot diagrams of the correlation coefficients and standard errors for 25%, minimum, maximum and 50% levels in different months and in all stations are also shown in Figures 2 to 5, respectively. According to the mentioned figures and the results of the correlation coefficients and standard errors (Tables 4 to 7) for the annual and monthly rainfall data with five distribution of Normal, two-parameter Log-Normal, three-parameter Log-Normal, Pearson type III and Log Pearson type III, it has been concluded that the majority of high correlation coefficients and also most of the less standard errors for monthly and annual data can be seen in the Pearson III distribution. Therefore and based on PPCC method, the most probable distribution of monthly and annual data for mentioned stations is chosen as Pearson type III distribution. Probability plot of the observed annual data for dominant frequency coefficient (Pearson type III) for the Urmia station is also presented in Figure 6.

Figures 7 to 12 show the L-kurtosis against the L-skewness for three-parameter distribution and L-moment coefficient of variations (L-CV) against the L-skewness two-parameter distributions in both annual and monthly levels. Based on the L-moment method, Normal and Gumbel distribution is defined as a point. Probability plots based on 95% acceptance region for some probability distributions that have developed recently [23, 24], as well as correlation coefficients and standard errors between the values of the observed and theoretical values of the probability distribution have been used to select the best probability distribution based on the L-moment method. Correlation coefficients and standard error values are presented in Table 8. Based on the L-moment diagrams and correlation coefficients and standard error values, it can be seen that by this method, the Pearson type III distribution is the best statistical distribution of rainfall data in the area of study.

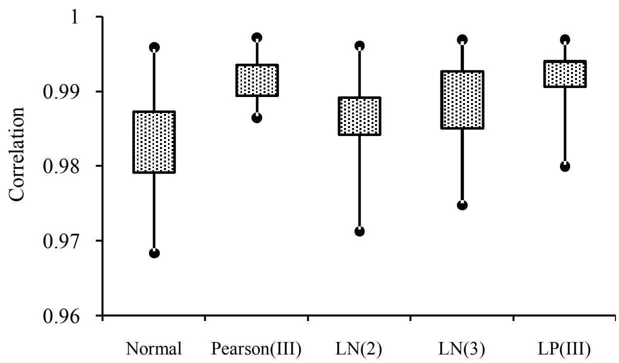

Figure 2. Correlation coefficient of annual data in all stations

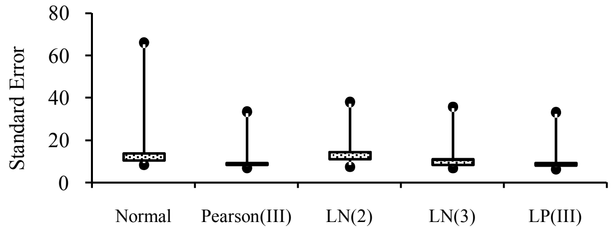

Figure 3. Standard errors of annual data in all stations.

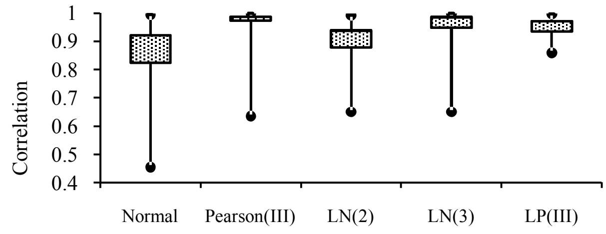

Figure 4. Correlation coefficients of monthly data in all stations.

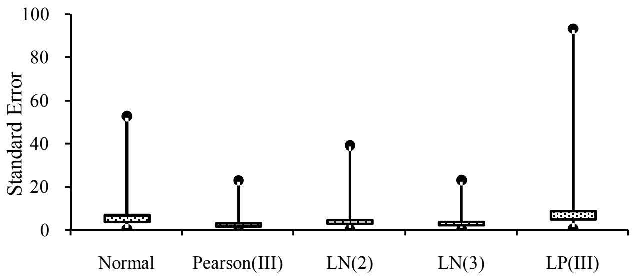

Figure 5. Standard errors of monthly data in all stations.

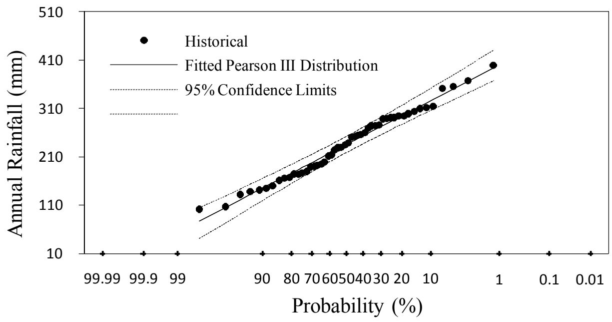

Figure 6. Pearson type III probability plots of annual data for Urmia station with 95% confidence limits.

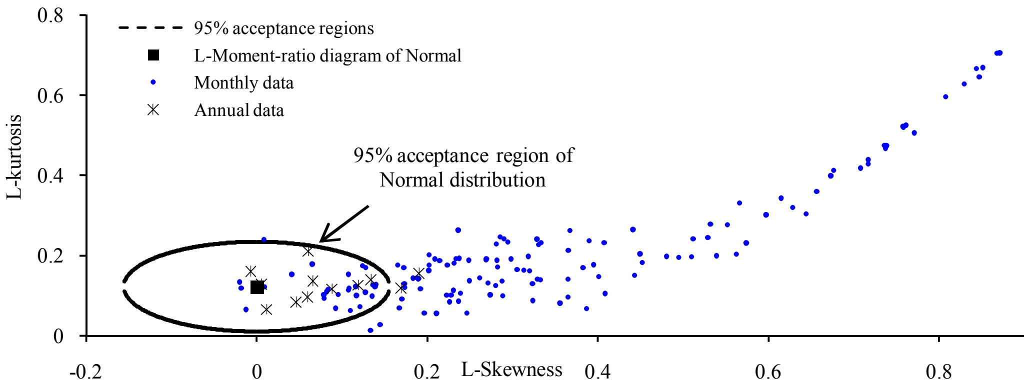

Figure 7. Relationship between L-Skewness and L-kurtosis with 95% acceptance region of Normal distribution

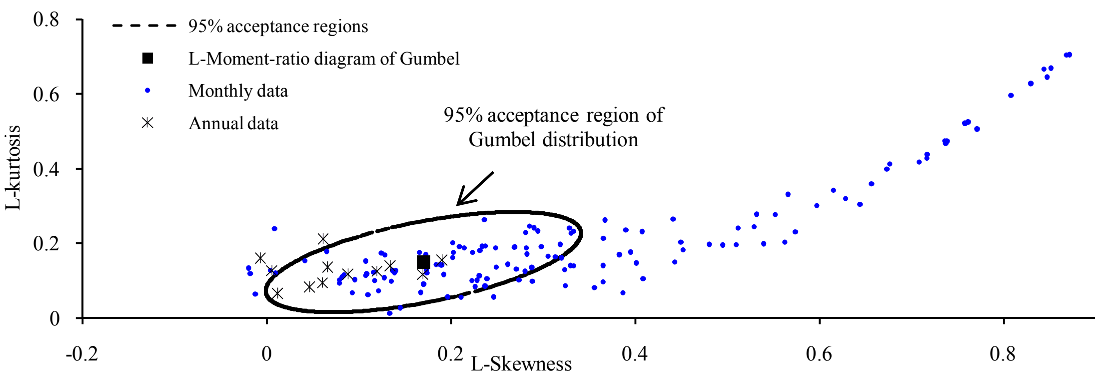

Figure 8. Relationship between L-Skewness and L-kurtosis with 95% acceptance region of Gumbel distribution

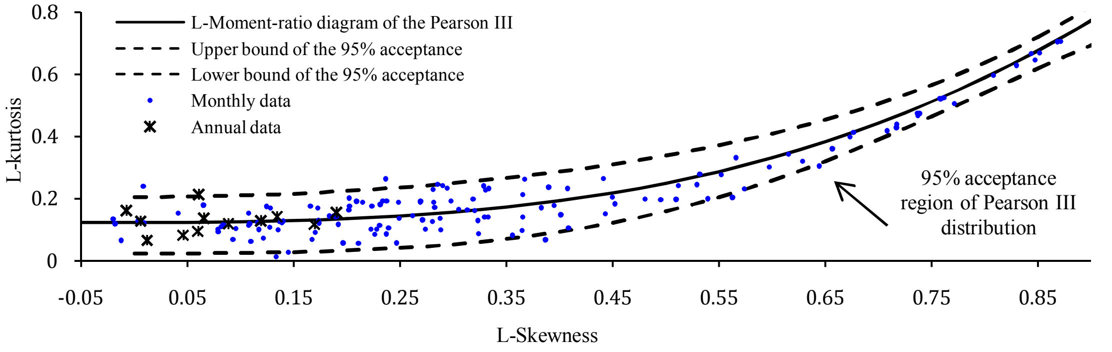

Figure 9. Relationship between L-Skewness and L-kurtosis with 95% acceptance region of Pearson III distribution

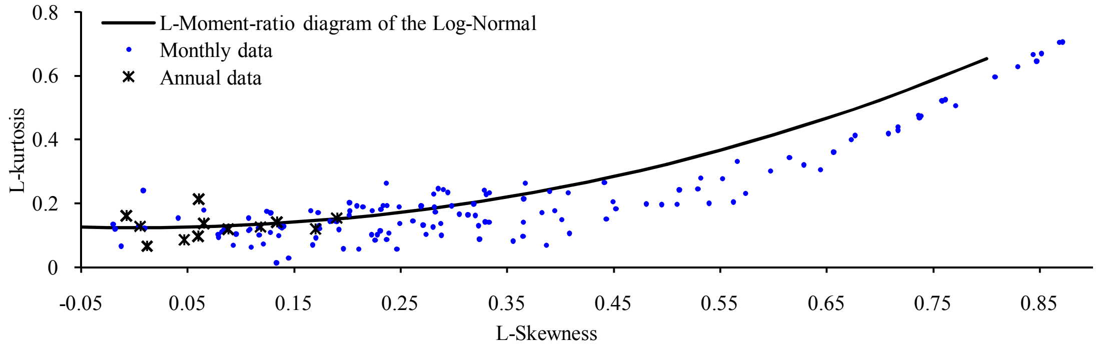

Figure 10. Relationship between L-Skewness and L-kurtosis for three-parameter Log-Normal distribution

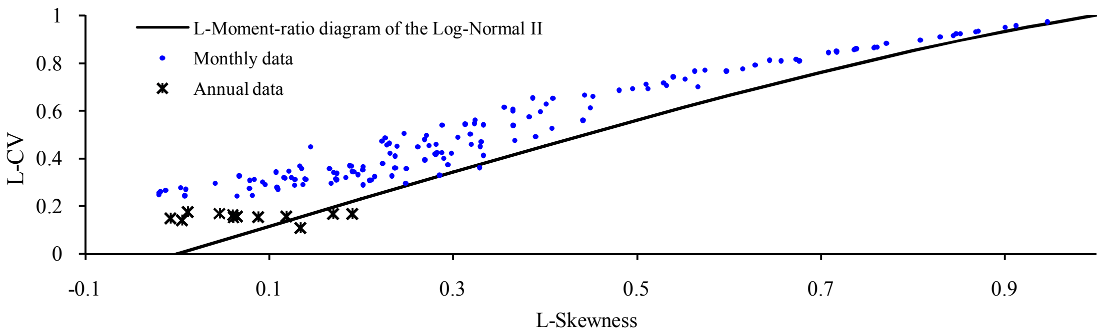

Figure 11. Relationship between L-Skewness and L-CV for two-parameter Log-Normal distribution

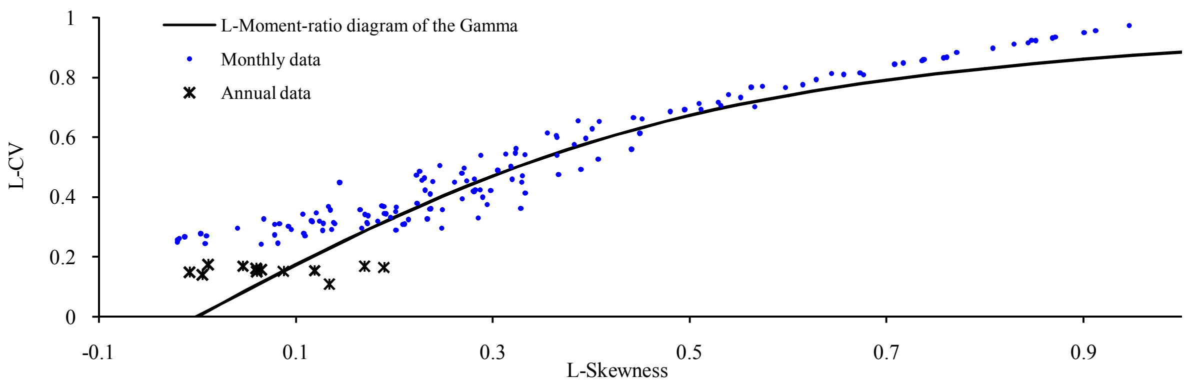

Figure 12. Relationship between L-Skewness and L-CV for Gamma distribution

TABLE 8. Correlation coefficients and standard errors for five probability distributions (maximum correlations and minimum standard errors are underlined).

Monthly data						
P3	**LN3**	**LN2**	**Gamma**	**Normal**	**Gumbel**	**Evaluation criteria**
0.9793	0.9489	0.9725	0.9418	_	0.0000	**R**
0.0496	0.0740	0.1475	0.0973	0.2820	0.1800	**RMSE**
Annual data						
P3	**LN3**	**LN2**	**Gamma**	**Normal**	**Gumbel**	**Evaluation criteria**
0.17707	0.17706	-0.0247	-0.0291	_	0.0000	**R**
0.03777	0.03773	0.0996	0.1102	0.1388	0.0582	**RMSE**

4. CONCLUSION

In this study, PPCC and L-moment methods were used to determine the best probability distribution of monthly and annual rainfall data in the north west of Iran. Results showed the superiority of Pearson type III distribution in this region. Meanwhile and regarding the PPCC method, the values of frequency factor of probability distributions and correlation coefficient between these values and the observed values were determined for each month. Although this method resulted in similar results with L-moment method and despite the simplicity and efficiency in determining the best probability distribution, this method was very time consuming because there were large number of stations. Thus, the authorsdo not recommend it for regional analyses. Among these, L-moment method used a diagram to determine the best distribution in an area. Therefore, L-moment method is recommended in order to use the time series models in a region rather than a single site PPCC method to determine the best probability distribution.

5. REFERENCE

1. Adeloye, A. and Montaseri, M., "Preliminary streamflow data analyses prior to water resources planning study/analyses préliminaires des données de débit en vue d'une étude de planification des ressources en eau", ***Hydrological Sciences Journal***, Vol. 47, No. 5, (2002), 679-692.
2. Samani, N. and Raeissi, E., "Stochastic synthesis of droughts for reservoir storage design", ***International Journal of Engineering***, Vol. 6, No. 1, (1993), 59-63.
3. Yazdani, A., Shahpari, A. and Salimi, M., "The use of monte-carlo simulations in seismic hazard analysis in tehran and surrounding areas", ***International Journal of Engineering-Transactions C: Aspects***, Vol. 25, No. 2, (2012), 159.
4. Motaei, A., Niaki, S. and Fard, N., "Bayes interval estimation on the parameters of the weibull distribution for complete and censored tests", ***International Journal of Engineering-Transactions C: Aspects***, Vol. 26, No. 9, (2013), 985.
5. Stedinger, J.R., "Frequency analysis of extreme events", ***in Handbook of Hydrology.***, Vol., No., (1993).
6. Heo, J.-H., Kho, Y.W., Shin, H., Kim, S. and Kim, T., "Regression equations of probability plot correlation coefficient test statistics from several probability distributions", ***Journal of Hydrology***, Vol. 355, No. 1, (2008), 1-15.
7. Amirataee, B. and Montaseri, M., "Evaluation of l-moment and ppcc method to determine the best regional distribution of monthly rainfall data (case study: Northwest of iran)", ***Journal of Urban and Environmental Engineering (JUEE)***, Vol. 7, No. 2, (2013).
8. Shin, H., Jung, Y., Jeong, C. and Heo, J.-H., "Assessment of modified anderson–darling test statistics for the generalized extreme value and generalized logistic distributions", ***Stochastic Environmental Research and Risk Assessment***, Vol. 26, No. 1, (2012), 105-114.
9. Maidment, D.R., "Handbook of hydrology, McGraw-Hill Inc., (1992).
10. Vogel, R.M. and Kroll, C.N., "Low-flow frequency analysis using probability-plot correlation coefficients", ***Journal of water resources planning and management***, Vol. 115, No. 3, (1989), 338-357.
11. Filliben, J.J., "The probability plot correlation coefficient test for normality", ***Technometrics***, Vol. 17, No. 1, (1975), 111-117.
12. Looney, S.W. and Gulledge Jr, T.R., "Use of the correlation coefficient with normal probability plots", ***The American Statistician***, Vol. 39, No. 1, (1985), 75-79.
13. Fill, H.D. and Stedinger, J.R., "L moment and probability plot correlation coefficient goodness-of-fit tests for the gumbel distribution and impact of autocorrelation", ***Water Resources Research***, Vol. 31, No. 1, (1995), 225-229.
14. Hosking, J.R., "L-moments: Analysis and estimation of distributions using linear combinations of order statistics", ***Journal of the Royal Statistical Society. Series B (Methodological)***, Vol., No., (1990), 105-124.
15. Hosking, J.R.M. and Wallis, J.R., "Regional frequency analysis: An approach based on l-moments, Cambridge University Press, (2005).
16. Schaefer, M., "Regional analyses of precipitation annual maxima in washington state", ***Water Resources Research***, Vol. 26, No. 1, (1990), 119-131.
17. Pearson, C., "Application of l-moments to maximum river flows", ***The New Zealand Statistician***, Vol. 28, No. 1, (1993), 2-10.
18. Vogel, R.M., McMahon, T.A. and Chiew, F.H., "Floodflow frequency model selection in australia", ***Journal of Hydrology***, Vol. 146, No., (1993), 421-449.
19. Chow, K.C.A. and Watt, W.E., "Practical use of the l-moments: Stochastic and statistical methods in hydrology and environmental engineering", Vol. 1 (ed. by K. W. Hipel),

Kluwer Academic Publishers, Boston, Massachusetts, USA,, No., (1994), 55-69.

20. Önöz, B. and Bayazit, M., "Best-fit distributions of largest available flood samples", ***Journal of Hydrology***, Vol. 167, No. 1, (1995), 195-208.

21. Vogel, R.M. and Wilson, I., "Probability distribution of annual maximum, mean, and minimum streamflows in the united states", ***Journal of Hydrologic Engineering***, Vol. 1, No. 2, (1996), 69-76.

22. Ben-Zvi, A. and Azmon, B., "Joint use of l-moment diagram and goodness-of-fit test: A case study of diverse series", ***Journal of Hydrology***, Vol. 198, No. 1-4, (1997), 245-259.

23. Liou, J.-J., Wu, Y.-C. and Cheng, K.-S., "Establishing acceptance regions forl-moments based goodness-of-fit tests by stochastic simulation", ***Journal of Hydrology***, Vol. 355, No. 1, (2008), 49-62.

24. Wu, Y.-C., Liou, J.-J., Su, Y.-F. and Cheng, K.-S., "Establishing acceptance regions for l-moments based goodness-of-fit tests for the pearson type iii distribution", ***Stochastic Environmental Research and Risk Assessment***, Vol. 26, No. 6, (2012), 873-885.

25. McMahon, T.A., Peel, M.C., Vogel, R.M. and Pegram, G.G., "Global streamflows–part 3: Country and climate zone characteristics", ***Journal of Hydrology***, Vol. 347, No. 3, (2007), 272-291.

26. Kottegoda, N.T. and Rosso, R., "Probability, statistics, and reliability for civil and environmental engineers", (1997).

27. Loucks, D.P., Stedinger, J.R. and Haith, D.A., "Water resource systems planning and analysis, Prentice-Hall., (1981).

28. McGhee, J.W., "Introductory statistics", ***West Publishing Co., New York, USA,***(1985).

29. Manzer, D.E. and Barnett, M.P., "Analysis by high speed digital computer in design of water resources systems", ***Harvard University Press***, (1966).

30. Joiner, B.L. and Rosenblatt, J.R., "Some properties of the range in samples from tukey's symmetric lambda distributions", ***Journal of the American Statistical Association***, Vol. 66, No. 334, (1971), 394-399.

31. Fennessey, N. and Vogel, R.M., "Regional flow-duration curves for ungauged sites in massachusetts", ***Journal of Water Resources Planning And Management***, Vol. 116, No. 4, (1990), 530-549.

32. Abramowitz, M. and Stegun, I.A., "Handbook of mathematical functions: With formulas, graphs, and mathematical tables, Courier Dover Publications, (1972).

33. Arnell, N.W., Beran, M. and Hosking, J., "Unbiased plotting positions for the general extreme value distribution", ***Journal of Hydrology***, Vol. 86, No. 1, (1986), 59-69.

Assessment of Goodness of Fit Methods in Determining the Best Regional Probability Distribution of Rainfall Data

B. Amirataee, M. Montaseri, H. Rezaei

Department of Water Engineering, Urmia university, Urmia

PAPER INFO

Paper history:
Received 18 November 2013
Received inrevised form 03 June 2014
Accepted 26 June 2014

Keywords:
Probability Plot Correlation Coefficient
L-moments
Goodness of Fit

چکیده

یکی از مراحل اساسی در تحلیل منطقه‌ای بارندگی، تعیین مناسبترین توزیع احتمالی داده‌های بارندگی از میان توزیع‌های احتمالی مختلف می‌باشد. بدین منظور، به طور معمول روش‌های کی دو، کلموگروف- اسمیرنف و ضریب همبستگی برازش احتمالاتی (PPCC)به عنوان روش‌های آزمون نکویی برازش مورد استفاده قرار می‌گیرد. در سال‌های اخیر، روش گشتاور خطی (L-moment)برای تعیین نکویی برازش توزیع‌های احتمالی مختلف در تحلیل منطقه‌ای داده‌های هیدرولوژیکی مانند بارندگی پیشنهاد و مورد استفاده قرار گرفته است. لذا در این مطالعه دو روش PPCC و L-moment برای ارزیابی و تعیین مناسب‌ترین توزیع احتمالی داده‌ها در فاصله اطمینان ۹۵٪ در تحلیل منطقه‌ای بارندگی شمال‌غرب ایران بکار گرفته شده است. بدین منظور داده‌های ۵۰ ساله بارندگی ماهیانه و سالیانه در ۱۲ ایستگاه سینوپتیک در منطقه مطالعاتی بر اساس معیارهای ارزیابی مختلف، مورد استفاده قرار گرفته است. نتایج حاصل از دو روش PPCC و L-moment نشان می‌دهد که توزیع پیرسون نوع سوم مناسب‌ترین توزیع احتمالی داده‌های بارندگی در منطقه بوده، درحالیکه روش L-moment دارای مزیت ارزیابی و تست تطبیق تعداد زیادی از سری داده‌ها با توزیع‌های احتمالی با استفاده از یک دیاگرام است.

***doi**: 10.5829/idosi.ije.2014.27.10a.07*