Civil and Environmental Research
ISSN 2224-5790 (Paper) ISSN 2225-0514 (Online)
Vol.8, No.12, 2016

Selection of Best Fitted Probability Distribution Function for Daily Extreme Rainfall of Bale Zone, Ethiopia

Asrat Fikre
Water Resources Research Center, Arba Minch University, PO box 21, Arba Minch, Ethiopia

Abstract
Different methods are employed to predict extreme rainfall in hydrology. The most common among these methods is the Probability Distribution method. In this study in order to identify suitable probability distribution for estimating of daily extreme rainfall series for different record period and class type of 18 rain gauge stations had been collected from Bale zone of Oromia region, Ethiopia. After homogeneity and consistency test, data has been analyzed by easy fit 5.5 standard software and Microsoft excel for predicting extreme values. Four probability distribution functions (Normal, Log Normal, Log Pearson type III and Gumbel extreme value type I) were employed and three goodness of fit tests (chi-sqare test, correlation coefficient and coefficient of determination) were applied to select the best probability distribution function for the study area. According to easy fit and Microsoft excel output Log Pearson type III is the most fitting for daily extreme rainfall.
Keywords: Extreme rainfall, Probability Distribution, Log Pearson Type III, Goodness of fit, Bale zone

1. Introduction

Extreme environmental events, such as floods, rainstorms, high winds and droughts, have severe consequences on environments. Planning for weather-related emergencies such as design of engineering structures, reservoir management, pollution control, and insurance risk calculations, all rely on knowledge of the frequency of these extreme events. The assessment of extreme precipitation is an important problem in hydrologic risk analysis and design. Extreme rainfall events could cause significant damage to agriculture, ecology and infrastructure, disruption to human activities, injuries and loss of lives (Einfalt *et al.,* 1998).

Hydrology cannot determine time of phenomenon of occurrence such as floods or discharge but can investigate previous events occurrence procedure and obtain their mean probability of occurrence. Calculation of mean probability of occurrence or floods mean return periods can help to solve many problems. Frequency analysis of floods and precipitation extreme values, the magnitude of this phenomenon and also their frequency give appropriate information for different analysis such as determination of risk criterions and reliability in the design of structures. This analysis provides this possible until the frequency value of events that are more than their observational value estimated during the period of data record. This estimate can be expressed using the concept of event return period (Hadian *et al*., 2011).

Distribution fitting is the procedure of selecting a statistical distribution that best fits to a data set generated by some random process. Among the frequency distributions normal distribution, lognormal distribution, Log-Pearson type III distribution (LPT3), and Gumbel extreme value type I distribution (GEVT1)

Hydrologic frequency analysis poses problem in that it must have a sufficiently long record of hydrologic data (daily maximum rainfall) and must test the goodness-of-fit (GoF) of these hydrologic data sets to evaluate appropriate distributions prior to use. Generally, three testing mechanisms were utilized to assess the distribution assumptions of GoF. These are chi-square test, correlation coefficient and coefficient of determination.

2. Materials and Methodology

2.1. Description of the Study Area

The study was carried out in Bale zone of Oromia National Regional State, Ethiopia. it is located between 5011'03''– 8009'27''N Latitude and 38012'04''– 42012'47''E longitude. The lowest and highest altitude of the zone is extended from 300 m around south east Rayitu and Gura Damole district to 4377 m in Tulu Dimtu mountain above sea level. The mean annual rainfall varies from 400mm on extreme lowland up to 1200mm on highlands. The rainfall increases from south, east and south-east lowlands toward the highlands. The maximum and minimum temperatures are 25oC and 10oC, respectively. (PSEBZ, 2010 - 2011).

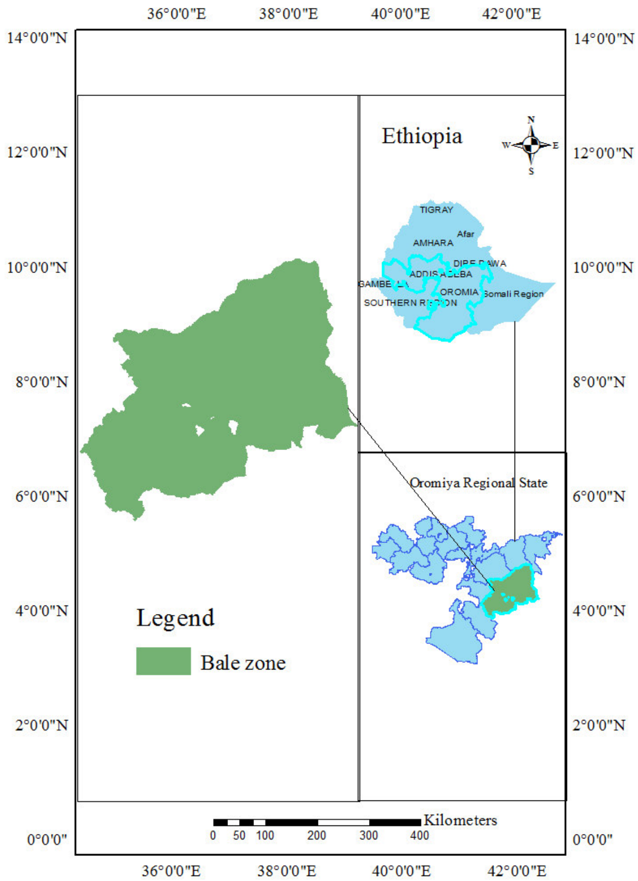

Fig. 1. Location map of the study

2.2. Data Evaluation and Analysis

Before using the data for analysis purpose stations were evaluated for their data adequacy and functionality. The evaluations showed that all stations in the studied area were active during the data collection and have maximum record length of 44 years, minimum record length of 14 years and with average of 26.22 years. Data were analyzed by Microsoft Excel sheet and Easy fit 5.5 standard software.

2.3. Data Processing

2.3.1. Identification and estimation of missed data

Rainfall data often has significant portion of the historic record missing needs to be estimated. Accordingly, the historical daily rainfall data of each considered station was checked for its missing value of the considered record years. To estimate the missed rainfall values, the Inverse Distance Weighting method (Simanton and Osborn, 1980) which is the most commonly used for estimation of missing data has been used. This method is also widely used and recommended by the United States Army Corps of Engineers (USACE, 2000). The inverse-distance (reciprocal-distance) weighting method for estimation of missing value of an observation, θm, using the observed values at other stations is given by:

$$\theta_m = \frac{\sum_{i=1}^{n} \theta_i d_{m,i}^{\ -k}}{\sum_{i=1}^{n} d_{m,i}^{\ -k}} \qquad (2.1)$$

where, θ_m is the observation at the base station m, n is the number of neighboring stations, θi is the observation at station i, d_m, is the distance from the location of station i to station m, and k is referred to as the friction distance (Vieux, 2001) that ranges from (1–6), the common value 2, was employed for this study.

2.3.2. Consistency analysis of the data set

The consistency of the data set of the given stations was checked by the double mass-curve method with reference to their neighborhood stations. The double mass curve was plotted by using the annual cumulative total rainfall of the base station as ordinate and the average annual cumulative total of neighboring stations as abscissa.

$$P_a = \frac{M_a}{M_o} P_o \tag{2.2}$$

where, Po is the observed value

Pa is the adjusted value

Mo is the slope of the double mass curve corresponding to the value, to which the observed values are being adjusted.

Ma corrected slope of the double mass curve

2.4. Fitting Data to the Probability Distribution Functions

Frequency analysis techniques were employed to analyze the annual maximum rainfall data. Fitting the theoretical probability distribution to the observed data was done by Weibull's plotting position, (Tao et al, 2002).

Normal Distribution: Values of standard normal deviate (Z) for exceedence of probability were interpolated by using Table 1.

Table 1. Values of the Standard Normal Deviate for the Cumulative Normal Distribution.

Exceedence probability %	Return period	Z
50	2	0.00
20	5	0.8416
10	10	1.2816
4	25	1.7507
2	50	2.0538
1	100	2.3264
0.2	500	2.8782

Log Normal Distribution: According to (Suresh, 2005) Annual maximum values were arranged in the descending order of magnitude and transformed in to logarithms then, assigning a rank m with 1 for the highest value, after that '*Z*','*W*' and other parameters were estimated using equations (2.3_ 2.8).

$$Z = K_T = w - \left[\frac{(2.516 + 0.8028w + 0.0103w^2)}{(1 + 1.4328w + 0.1893w^2 + 0.0013w^3)}\right] \tag{2.3}$$

where w is intermediate variable which is calculated using the formula:

$$w = \left[\ln \frac{1}{p^2}\right]^{1/2} \quad (0 < p \leq 0.5) \tag{2.4}$$

where p is the probability of exceedence. When p> 0.5, 1-p is substituted for p and the value of Z which is computed is given a negative sign.

Kite (1977) as cited by Suresh (2005) suggested that when Cs was zero then K_T = Z, if not K_T would be approximated by:

$$K_T = Z + (Z^2 - 1)k + \frac{1}{3}(Z^3 - 6Z)k^2 + Zk^4 + \frac{1}{3}k^5 \tag{2.5}$$

where, $k = {}^{C_s}/_6$ (2.6)

$Y_T = Yn + K_T S_y$ (2.7)

$X_T = 10^{Y_T}$ (2.8)

Gumbel Extreme Value Type I Distribution: By Subramanya (1996), fitting data to Gumbel EVI distribution was achieved by using equation 2.9 and 2.10, respectively and lastly $X_T = \bar{x} + K_T{*}S_n$ were obtained.

$$K_T = \frac{Y_T - Y_n}{S_n} \tag{2.9}$$

where, Y_T is the reduced variate which is given as:

$$Y_T = -\ln\left[\ln\left(\frac{T}{T-1}\right)\right] \tag{2.10}$$

2.5. Testing the Goodness of Fit of Data to Probability Distribution

In order to select the best fitted model probability distribution function for stations as well the zone were tested by the three parameters of goodness of fit, chi-square test (χ^2), correlation coefficient(r) and coefficient of determination (R^2). The assessment of the probability distribution models were based on the total test score obtained from all the tests. Test scores ranging from zero to four (0-4) was awarded to each distribution model based on the criteria that the distribution with the highest total score was or were chosen as the best distribution model for the data of a particular station. Later on, model which was selected repeatedly for each station was selected as best fit model for the zone. In general, the distribution best supported by a test was awarded a score of four; the next best was awarded three, and so on in descending order. A distribution was awarded a zero (0) score for a test if the test indicates that there was a significant difference between the rainfall values estimated by the distribution model and the observed rainfall data. For every test category, overall ranks of each distribution were obtained by summing the individual point rank at each of the 18 stations (Adegboye and Ipinyomi, 1995).

3. RESULTS AND DISCUSSION

3.1. Manipulation of Missing and Consistency Test

Rainfall data often has significant portion of the historic record missing needs to be estimated before using it for further analysis and missing value were estimated by Inverse Distance Weighting method. The consistency of the data set of the given stations was checked by the double mass-curve method with in-reference to their neighborhood stations and the result reveals data are consistent.

3.2. Record Length, Average Annual Total and Annual Daily Maximum

The recorded length of stations varied according to their establishment and functionality. The maximum, minimum and average recorded length of stations from the observed was 44, 14, and 26 years, respectively. Rainfall measured daily at a fixed time in the morning and expressed in depth (mm), average annual total, annual daily maximum, average annual daily maximum rainfall and recorded length of the stations are presented in Table 2. Relatively, high rainfall coverage in the northern and central districts and low rainfall was record in western districts.

Table 2. Record length, average annual total, annual daily maximum and average annual daily maximum rainfall of different stations.

Name of station	Length of record (years)	Average annual total rainfall (mm)	Annual daily maximum (mm)	Average annual daily Maximum rainfall (mm)
Abissa	18	732.56	84.9	47.87
Agarfa	44	1103.47	142.8	63.50
Angetu	17	853.41	86.9	53.33
Belle	16	902.04	90.4	43.78
Beletu	14	823.78	88.8	39.62
Bidre	16	874.34	94.0	62.48
D/Mena	22	1007.66	109.3	62.29
D/sebro	44	1199.63	141.5	57.24
Dinsho	32	1332.39	88.4	48.27
Gassera	44	1122.17	108.1	56.14
Ginir	43	1036.59	200.0	72.19
Goro	32	894.36	83.0	52.41
Jara	20	982.58	120.0	53.46
Rira	19	737.18	48.0	28.60
Robe	28	819.17	112.3	47.11
Sewena	16	939.52	90.9	39.37
Sinana	31	894.24	85.5	45.33
Sofumor	16	661.10	106.0	53.06
Average	26.22	939.79	104.49	

3.3. Selection and Comparison of the Probability Distribution Functions

A key step in frequency analysis of precipitation involves selection of a suitable and stable distribution for representing precipitation depth to investigate the extremes (Hansonl and Vogel, 2008). Probability distribution model determine and verifies the best distribution function for the studied area. It depends mainly on characteristics of data. The selection of appropriate model depends mainly on the characteristics of available data at the appropriate site (Ewemoje and Ewemooje, 2011).

3.3.1. Normal probability distribution function

The standard normal deviate (Z) value of exceedence probability for the annual maximum rainfall data of Abissa station were interpolated and derived as shown in Table 3. The result shows that the standard normal variate of records decrease with increase in plotting probability and obtained extreme value using the normal distribution function shows linear proportionality with the standard normal variate.

Table 1. Extreme value derived by normal distribution for Abissa station

RF order	Rank (m)	P	P (%)	Z	Z*Sn	XT
84.9	1	0.05	5.26	1.65	23.25	71.12
75.9	2	0.11	10.53	1.26	17.71	65.58
65.5	3	0.16	15.79	1.03	14.45	62.32
54.5	4	0.21	21.05	0.81	11.43	59.3
50.9	5	0.26	26.32	0.66	9.35	57.22
47.1	6	0.32	31.58	0.52	7.27	55.14
45.4	7	0.37	36.84	0.37	5.20	53.06
44.4	8	0.42	42.11	0.22	3.12	50.98
43.9	9	0.47	47.37	0.07	1.04	48.91
42.8	10	0.53	52.63	0	0	47.87
41.6	11	0.58	57.89	0	0	47.87
41.5	12	0.63	63.16	0	0	47.87
40.8	13	0.68	68.42	0	0	47.87
40.8	14	0.74	73.68	0	0	47.87
36.6	15	0.79	78.95	0	0	47.87
36.3	16	0.84	84.21	0	0	47.87
35.2	17	0.89	89.47	0	0	47.87
33.5	18	0.95	94.74	0	0	47.87
$\overline{X}_N$=47.87	Sd	14.08	Cv (%)	29.41		

n=18, p=plotting probability, $\overline{X}_N$=mean, S_n=standard deviation, Z= standard normal deviate variate, X_T=extreme drived value

3.3.2. Log normal probability distribution function

The standard normal variable (Z) value of exceedence probability for the annual maximum rainfall data of sample station(Abissa) was calculated and presented in Table 4 . The result shows that the standard normal variable of the records has direct relation with recurrence interval (inverse relation with plotting probability) and extreme value obtained using the Log normal distribution function shows linear proportionality with the standard normal variate.

Table 2. Standard normal variable (Z) and Extreme value derived by Log normal distribution function for Abissa station.

Rainfall order	Log RF	Rank (m)	P	P^2	W	Z	Y_T	X_T
84.9	1.93	1	0.05		2.43	1.62	1.84	69.41
75.9	1.88	2	0.11		2.12	1.25	1.80	63.30
65.5	1.82	3	0.16		1.92	1.00	1.77	59.47
54.5	1.74	4	0.21		1.77	0.80	1.75	56.58
50.9	1.71	5	0.26		1.63	0.63	1.73	54.21
47.1	1.67	6	0.32		1.52	0.48	1.72	52.15
45.4	1.66	7	0.37		1.41	0.34	1.70	50.31
44.4	1.65	8	0.42		1.32	0.20	1.69	48.62
43.9	1.64	9	0.47		1.22	0.07	1.67	47.03
42.8	1.63	10		0.47	1.22	-0.07	1.66	45.50
41.6	1.62	11		0.42	1.32	-0.20	1.64	44.01
41.5	1.62	12		0.37	1.41	-0.34	1.63	42.53
40.8	1.61	13		0.32	1.52	-0.48	1.61	41.03
40.8	1.61	14		0.26	1.63	-0.63	1.60	39.47
36.6	1.56	15		0.21	1.77	-0.80	1.58	37.82
36.3	1.56	16		0.16	1.92	-1.00	1.56	35.98
35.2	1.55	17		0.11	2.12	-1.25	1.53	33.80
33.5	1.53	18		0.05	2.43	-1.62	1.49	30.83
Yn	1.67	Sy	0.11					

Yn=mean of variate, Sy=standard deviation of the variate

3.3.3. Log Pearson type III probability distribution function

The standard normal variable (Z) and frequency factor (K_T) value for exceedence probability for the annual maximum rainfall data of Abissa station was calculated and presented as shown in Table 5. The result shows that the standard normal variable of the records decrease with decrease in recurrence interval and extreme value obtained using the Log Pearson distribution function shows linear proportionality with the standard normal variate.

Table 3. Standard normal variable (Z) and Extreme value derived by Log Pearson type III distribution for Abissa station.

RF order	Log RF	Rank (m)	P	P^2	W	Z	K_T	Y_T	X_T
84.9	1.93	1	0.05		2.43	1.60	1.84	1.87	73.37
75.9	1.88	2	0.11		2.12	1.24	1.28	1.80	63.67
65.5	1.82	3	0.16		1.92	0.99	0.93	1.77	58.36
54.5	1.74	4	0.21		1.77	0.80	0.67	1.74	54.72
50.9	1.71	5	0.26		1.63	0.63	0.46	1.72	51.95
47.1	1.67	6	0.32		1.52	0.47	0.29	1.70	49.70
45.4	1.66	7	0.37		1.41	0.33	0.13	1.68	47.81
44.4	1.65	8	0.42		1.32	0.19	-0.01	1.66	46.16
43.9	1.64	9	0.47		1.22	0.06	-0.14	1.65	44.69
42.8	1.63	10		0.47	1.22	-0.06	-0.25	1.64	43.42
41.6	1.62	11		0.42	1.32	-0.19	-0.37	1.63	42.18
41.5	1.62	12		0.37	1.41	-0.33	-0.48	1.61	41.01
40.8	1.61	13		0.32	1.52	-0.47	-0.59	1.60	39.89
40.8	1.61	14		0.26	1.63	-0.63	-0.70	1.59	38.81
36.6	1.56	15		0.21	1.77	-0.80	-0.81	1.58	37.72
36.3	1.56	16		0.16	1.92	-0.99	-0.93	1.56	36.62
35.2	1.55	17		0.11	2.12	-1.24	-1.06	1.55	35.43
33.5	1.53	18		0.05	2.43	-1.60	-1.23	1.53	34.02
Yn	1.67	Sn	0.11	Cs	1.18	K	0.20		

Yn=mean of variate, Sn=standard deviation of variate, Cs=coefficient of skewness, K=kurtosis

3.3.4. Gumbel EVI probability distribution function

The reduced variate value for exceedence probability for the annual maximum rainfall data of Abissa station was calculated and presented as shown in Table 6. The result shows that the reduced variate value of the records decreases with decrease in recurrence interval (increase in plotting probability) and extreme value obtained using Gumbel distribution function shows linear proportionality with the reduced variate.

Table 4. The reduced variate value of exceedence probability for the annual maximum rainfall data of Abissa station using Gumbel EVI distribution function.

RF order	Rank (m)	P	T	Y_T	K_T	X_T
84.9	1	0.05	19.00	2.92	2.29	80.17
75.9	2	0.11	9.50	2.20	1.60	70.45
65.5	3	0.16	6.33	1.76	1.18	64.59
54.5	4	0.21	4.75	1.44	0.88	60.29
50.9	5	0.26	3.80	1.19	0.64	56.84
47.1	6	0.32	3.17	0.97	0.43	53.92
45.4	7	0.37	2.71	0.78	0.25	51.34
44.4	8	0.42	2.38	0.60	0.08	49.00
43.9	9	0.47	2.11	0.44	-0.07	46.84
42.8	10	0.53	1.90	0.29	-0.22	44.79
41.6	11	0.58	1.73	0.15	-0.36	42.82
41.5	12	0.63	1.58	0.00	-0.49	40.88
40.8	13	0.68	1.46	-0.14	-0.63	38.95
40.8	14	0.74	1.36	-0.29	-0.77	36.97
36.6	15	0.79	1.27	-0.44	-0.92	34.89
36.3	16	0.84	1.19	-0.61	-1.08	32.61
35.2	17	0.89	1.12	-0.81	-1.27	29.93
33.5	18	0.95	1.06	-1.08	-1.53	26.32
mean 47.87	Sd	14.12	yn	0.52	sn	1.05

P=plotting probability, T=recurrence interval, K_T= frequency factor, X_T=extreme derived value

As shown from the result comparison of probability distribution function of stations, the variate of stations records decrease with recurrence interval and extreme value obtained shows linear proportionality with the standard normal variate.

3.4. Testing Goodness of Fit of Data to Probability Distribution Functions

The fitness of different probability distribution functions for obtained extreme values with the observed values were tested by chi-square, correlation coefficient and coefficient of determination. The distribution function that gave high score from the three tests was selected as the best probability distribution function for that station and the cumulative scores of stations represents for the zone.

3.4.1. Chi square test (X^2)

Comparison of the recorded data and the corresponding values obtained by each of the probability distribution functions was made by calculating x^2 and comparing it with tabulated x^2 at 5% significance level and degree of freedom (v)=m-p-1, in this case for Abissa station v=6. As the calculated Chi square was found to be less than that of the tabulated value, there is no significance difference between the observed and predicted ones. The one with the least value of x^2 was selected as the best fit model. From Table 7, it could be stated that the Log Pearson type III distribution function having least value of calculated x^2 (5.59) could be assumed as best fit model and be assigned with 4 points, followed by Gumbel EVI (8.06) could be the second fit and assigned with 3 point. Whereas normal distribution function as weak model with calculated x^2 value of 26.58 assigned with 1 point.

3.4.2. Correlation coefficient test

The relationships between the observed and predicted rainfall data for 18 years record of Abissa station for different probability distribution function were developed (Table 7). For the purpose of comparison, 4 point was assigned for r value which was closest to 1, and so on. For Abissa station, Log Pearson type III was the best fit and assigned as 4, normal distribution function could be the second and assigned 3, Gumbel EVI, the third and assigned 2 and the least was Log normal 1. In this test also Log Pearson Type III was selected as the best fit probability distribution function followed by Normal distribution.

3.4.3. Coefficient of determination test

Based on the values in Table 7, the closer R^2 to 1, the better the regression equation "fits" to the data. Here Log Pearson type III fits best assigned 4 point, normal with 3 point, Gumbel EVI with 2 point and Log normal distribution function fits least which assigned 1 as per R^2 obtained.

Table 7. Chi-square, correlation coefficient and coefficient of determination test of goodness of fit for Abissa station.

No.	Observed	Normal	Lognormal	Log Pearson III	GEVI
1	84.9	71.12	69.41	73.37	80.17
2	75.9	65.58	63.30	63.67	70.45
3	65.5	62.32	59.47	58.36	64.59
4	54.5	59.30	56.58	54.72	60.29
5	50.9	57.22	54.21	51.95	56.84
6	47.1	55.14	52.15	49.70	53.92
7	45.4	53.06	50.31	47.81	51.34
8	44.4	50.98	48.62	46.16	49.00
9	43.9	48.91	47.03	44.69	46.84
10	42.8	47.87	45.50	43.42	44.79
11	41.6	47.87	44.01	42.18	42.82
12	41.5	47.87	42.53	41.01	40.88
13	40.8	47.87	41.03	39.89	38.95
14	40.8	47.87	39.47	38.81	36.97
15	36.6	47.87	37.82	37.72	34.89
16	36.3	47.87	35.98	36.62	32.61
17	35.2	47.87	33.80	35.43	29.93
18	33.5	47.87	30.83	34.02	26.32
Total	861.60	954.43	852.05	839.51	861.60
Mean	47.87	53.02	47.34	46.64	47.87
Sd	14.12	7.25	10.49	10.53	14.53
	X^2 cal.	26.58	9.09	5.59	8.06
	X^2 tabulated	12.59	12.59	12.59	12.59
	Test score	1	2	4	3
	r-value	0.97	0.94	0.98	0.95
	Test score	3	1	4	2
	R^2	0.94	0.88	0.96	0.90
	Test score	3	1	4	2
	Total test score	7	4	12	7

**Log Pearson type III is selected as best fit distribution model,*Normal and GEVI as second fit

From the results of four frequency distribution models applied in this study, it could be concluded that

Log-Pearson type III distribution was the best fit distribution for Abissa station and also Bale Zone, which accounted 72.22% of the total station number, followed by Log normal and Gumbel distribution each accounted 11.11%, and Gumbel and Log-Pearson distribution accounted 5.56% and no station fitted with Normal distribution. The results of stations with the three goodness of fit test have been summarized as follow.

S/N	Distribution	Test result			Score			Total	Fitted Model
		x^2 test	r-test	R^2-test	x^2 test	r-test	R^2-test		
Abissa	Normal	26.5846	0.9675	0.9360	1	3	3	7	LPIII
	Log normal	9.0947	0.9354	0.8750	2	1	1	4	
	Log pearson	5.5886	0.9778	0.9562	4	4	4	12	
	Gumbel	8.0639	0.9498	0.9021	3	2	2	7	
Agarfa	Normal	166.6642	0.9604	0.9223	1	1	1	3	LPIII
	Log normal	20.3362	0.9894	0.9788	2	2	2	6	
	Log pearson	15.5870	0.9951	0.9902	3	4	4	11	
	Gumbel	10.2638	0.9899	0.9798	4	3	3	10	
Angetu	Normal	27.8688	0.8855	0.7841	1	1	1	3	Gumbel
	Log normal	2.0990	0.9817	0.9834	2	2	2	6	
	Log pearson	2.1000	0.9917	0.9835	3	3	3	9	
	Gumbel	0.8078	0.9928	0.9856	4	4	4	12	
Belle	Normal	39.3189	0.9747	0.9501	1	3	3	7	LPIII
	Log normal	8.2116	0.9700	0.9409	2	1	1	4	
	Log pearson	5.9665	0.9857	0.9717	4	4	4	12	
	Gumbel	7.0031	0.9707	0.9423	3	2	2	7	
Beletu	Normal	40.9347	0.9693	0.9395	1	3	3	7	LPIII
	Log normal	13.3814	0.9469	0.8966	3	2	2	7	
	Log pearson	9.2536	0.9794	0.9593	4	4	4	12	
	Gumbel	15.4802	0.9440	0.8911	2	1	1	4	
Bidre	Normal	32.2807	0.8987	0.8076	1	1	1	3	LPIII
	Log normal	3.5584	0.9835	0.9673	3	3	3	9	
	Log pearson	3.5686	0.9837	0.9677	2	4	4	10	
	Gumbel	3.1350	0.9813	0.9629	4	2	2	8	
D/Mena	Normal	58.4473	0.9421	0.8876	1	1	1	3	GEVI
	Log normal	3.4717	0.9938	0.9877	3	3	3	9	
	Log pearson	3.4727	0.9929	0.9858	2	2	2	6	
	Gumbel	1.7316	0.9942	0.9884	4	4	4	12	
D/sebro	Normal	183.0664	0.9729	0.9465	1	3	3	7	LPIII
	Log normal	32.6663	0.9696	0.9401	3	2	2	7	
	Log pearson	14.7437	0.9881	0.9764	4	4	4	12	
	Gumbel	49.2933	0.9654	0.9320	2	1	1	4	
Dinsho	Normal	68.5005	0.9448	0.8926	1	1	1	3	Lognormal
	Log normal	4.9078	0.9883	0.9767	2	4	4	10	
	Log pearson	4.8043	0.9882	0.9766	3	3	3	9	
	Gumbel	3.6127	0.9881	0.9763	4	2	2	8	
Gassera	Normal	134.0407	0.9676	0.9362	1	1	1	3	LPIII
	Log normal	12.3132	0.9864	0.9730	2	3	3	8	
	Log pearson	9.7119	0.9868	0.9738	4	4	4	12	
	Gumbel	11.3979	0.9860	0.9721	3	2	2	7	
Ginir	Normal	232.9372	0.9465	0.8958	1	1	1	3	LPIII
	Log normal	27.5268	0.9750	0.9505	3	3	3	9	
	Log pearson	16.2647	0.9867	0.9735	4	4	4	12	
	Gumbel	41.3952	0.9708	0.9425	2	2	2	6	
Goro	Normal	33.7512	0.8797	0.7739	1	1	1	3	LPIII
	Log normal	4.5050	0.9811	0.9626	3	3	3	9	
	Log pearson	3.5977	0.9849	0.9701	4	4	4	12	
	Gumbel	22.6959	0.9759	0.9524	2	2	2	6	
Jara	Normal	83.0233	0.9714	0.9435	1	1	1	3	LPIII
	Log normal	12.4952	0.9810	0.9624	3	3	3	9	
	Log pearson	8.9225	0.9912	0.9824	4	4	4	12	
	Gumbel	14.2908	0.9762	0.9529	2	2	2	6	
Rira	Normal	25.2388	0.9426	0.8886	1	1	1	3	Lognormal
	Log normal	2.8850	0.9847	0.9696	2	4	4	10	
	Log pearson	2.7843	0.9846	0.9695	3	3	3	9	
	Gumbel	1.9601	0.9844	0.9691	4	2	2	8	
Robe	Normal	71.8401	0.9327	0.8699	1	1	1	3	LPIII
	Log normal	19.3926	0.9373	0.8786	3	2	2	7	
	Log pearson	10.3598	0.9747	0.9501	4	4	4	12	
	Gumbel	19.4073	0.9429	0.8891	2	3	3	8	
Sinana	Normal	44.3509	0.9276	0.8604	1	1	1	3	LPIII and GEVI
	Log normal	3.9030	0.9812	0.9627	3	2	2	7	
	Log pearson	4.2652	0.9832	0.9667	2	4	4	10	
	Gumbel	3.1477	0.9821	0.9646	4	3	3	10	
Sewena	Normal	47.0954	0.9388	0.8813	1	1	1	3	LPIII
	Log normal	9.7734	0.9589	0.9195	2	3	3	8	
	Log pearson	8.7872	0.9656	0.9325	4	4	4	12	
	Gumbel	9.0495	0.9556	0.9132	3	2	2	7	
Sofumor	Normal	35.1993	0.9354	0.8750	1	1	1	3	LPIII
	Log normal	10.2098	0.9373	0.8785	2	2	2	6	
	Log pearson	7.4841	0.9640	0.9293	4	4	4	12	
	Gumbel	8.0220	0.9455	0.8939	3	3	3	9	

4. Conclusion

Probability distribution model determines and verifies the best distribution function for an area. It depends mainly on record period and characteristics of data. The outcome of frequency distributions was tested on three goodness-of-fit tests; Chi-square, correlation coefficient and coefficient of determination. Among the four frequency distributions applied in this study, log-Pearson type III was the best fitted distribution for Bale zone, which scored 72.22% of the total station number, followed by the Gumbel EVI and log normal distribution, which both accounted 11.11%.

Reference

Adegboye, O. S. and R. A. Ipinyomi. 1995. Statistical Tables for class work and Examination. Tertiary publications Nigeria Limited, Ilorin, Nigeria.

Einfalt, T. 1998. Correction and preparation of continuously measured rain gauge data: a standard method in North Rhine-Westphalia, *Water Science and Technology*, 37(11) 12-22.

Ewemoje, T.A. and O.S. Ewemoojee, 2011. Best distribution and plotting positions of daily maximum flood estimation at Ona River in Ogun-oshun river basin, Nigeria. Agricultural Engineering International: CIGR Journal, Vol.13, Manuscript No. 1380.

Hadian, M. Habibnezhad roshan, M. and Omidvar, A., 2011, Evaluation of different statistical distributions in instantaneous p eak discharges with different return periods in Tajen watershed. 6th National science and watershed engineer in g conference. 1-8.

Hansonl, L.S. and R. Vogel, 2008. The Probability Distribution of Daily Rainfall in the United States. Tufts Univ., Medford, MA 02155.

Simanton, J. R. and H. B. Osborn, 1980. Reciprocal-distance estimate of point rainfall. Journal of Hydraulic Engineering, Div. 106 (HY7), 1242–1246.

Subramanya, M.A. 1996. Goodness of fittest for extreme value distribution. Biometrika 84: 583-588.

Suresh, K.A. 2005. Trends in extreme daily precipitation indices in India. International journal of climatology. 24:457-466.

Tao, D.Q, V.T. Nguyen, and A. Bourque, 2002. On selection of probability distributions for representing extreme precipitations in Southern Quebec. Annual Conference of the Canadian Society for Civil Engineering. 5th -8th June 2002. pp 1-8.

United States Army Corps of Engineers (USACE), 2000. Hydrologic Modeling System HEC-HMS, Technical reference manual. Hydrologic Engineering Center, Davis, California.

Vieux, B. E. 2001. Distributed Hydrologic Modeling Using GIS. Water Science and Technology Library. Kluwer. Amsterdam.