Environment & Ecology 35 (1B) : 434—439, January—March 2017
Website: environmentandecology.com ISSN 0970-0420

Identifying Best Probability Distribution for Events of Maximum Rainfall–A Case Study of Karnal, Haryana

Manoj Kumar, Anurag, Diwan Singh, Raj Singh

Received 6 April 2016 ; Accepted 11 May 2016 ; Published online 22 June 2016

Abstract The present study has been undertaken to fit best probability distribution of rainfall of Karnal District representing eastern climatic zone of Haryana state. For the purpose, daily rainfall data of 47 years was analyzed and processed to identify the maximum rainfall received on daily, monthly, seasonal and annual basis. The analysis showed that the maximum daily rainfall among the years ranged between 38 mm (1980) to 229 mm (2009) indicating a very large variation during the period of study. The standard deviation and coefficient of variation were also computed. Different goodness of fit tests were carried out to select the best fit probability distribution on the basis of highest rank with minimum value of test statistic. The probability distribution was identified using maximum overall score based on sum of individual point score obtained from different selected goodness of fit tests. Generalized extreme value distribution was found to be best fit probability distribution for most of the periods. The results clearly established that the analytical procedure devised and tested in this study may be suitability applied for the identification of the best fit probability distribution of weather parameters.

Keywords Maximum rainfall, Probability distributions, Goodness-of-fit tests, Pearson type distribution, Log-normal.

Dr M. Kumar*
Assistant Professor, CCSHAU, College of Agriculture, Kaul (Kaithal), India

Dr Anurag
Assistant Scientist, Deptt. of Agricultural Meteorology, CCS HAU, Hisar, India

Dr D. Singh
Principal Scientist, Deptt. of Agricultural Meteorology, CCS HAU, Hisar, India

Dr R. Singh
Principal Scientist, Deptt. of Agricultural Meteorology, CCS HAU, Hisar, India
e-mail: m25424553@gmail.com
*Correspondence

Introduction

A good understanding of the pattern and distribution of rainfall is of vital importance for water resource management of a country. Analysis of rainfall data strongly depends on its distribution pattern over time and it has great a influence on agricultural production. It has long been a topic of interest in the fields of climatology to find a probability distribution that provides a good fit to daily rainfall. Several studies have been conducted in India and abroad on rainfall analysis and best fit probability distribution function such as Normal, Log-normal, gumbel, weibull and Pearson type distribution were identified.

Salami [1] studied the meteorological data for Texas and found that Gumbel distribution fits adequately for both evaporation and temperature data, while for precipitation data log-Pearson type III distribution confirms to be more accurate. Takara et al. [2] analyzed the extreme events and revealed that

hydrological extremes sometimes do not fit well to the theoretical extreme-value distribution such as the Bumbel and generalized extreme value distributions. Lee [3] indicated that log-Pearson type III distribution fits for 50% of total station number for the rainfall distribution characteristics of Chia-Nan plain area.

Bhaskar et al. [4] observed the frequency analysis of consecutive days peaked rainfall at Banswara, Rajasthan, India, and found gamma distribution as the best fit distribution. Kwaku et al. [5] revealed that the log-normal distribution was the best fit probability distribution for one to five consecutive days' maximum rainfall for Accra, Ghana. Olofintoye et al. [6] examined that 50% of the total station number in Nigeria follows log-Pearson type III distribution for peak daily rainfall, while 40% and 10% of the total station follows Pearson type III and log-Gumbel distribution, respectively.

The present study was planned to establish the methodology for identifying the best fit probability distribution on the basis of three types goodness of fit tests for eastern agro-climatic zone of the state where water requirements are high owing to cultivation or rice crop. The Karnal district which represents the agro climatic properties of eastern agro-climatic zone was chosen for the purpose. The maximum rainfall data was used to select a best fit probability distribution for rainfall. The geographic location of Karnal district lies on the western bank of the river Yamuna, which forms its eastern boundary and separates Haryana from Uttar Pradesh. Karnal District lies between 29°25´ and 30° North and 76°30´ and 77°10´ East. Its height above mean sea level is around 253 meters [7].

Materials and Methods

The present study is based on time series data of maximums among daily rainfall in a year, season, month and week. The maximum daily, weekly, monthly, seasonal and annual rainfall data of 47 years (1966 to 2013) were collected from observatory of Central Soil Salinity Research Institute, Karnal. The normal annual rainfall of the district is 582 mm with 32 rainy days in a year. About 82.39% of the annual rainfall is recorded during the southwest monsoon from July to September. August is the wettest month of the year with an average of 9.0 rainy days and 221.5 mm rainfall. Annual maximum rainfall of 1404 mm and minimum of 255 mm were observed in the years 1998 and 1987 respectively. The daily rainfall data were processed to identify the maximum rainfall received on any one day (24 h duration), in any week (7 days), in a month (4 weeks), in a season (monsoon season i.e. June to September as 82% of annual rainfall received in this season) and in a year (365 days period). Various probability distributions namely normal, lognormal (2P, 3P), gamma (2P, 3P), generalized gamma (3P, 4P), log-gamma, weibull (2P, 3P), Pearson 5 (2P, 3P), Pearson 6 (3P, 4P), log-Pearson 3, generalized extreme value were fitted and evaluated by using the following systematic steps.

Step I : Fitting the probability distribution

The description of various probability distribution functions viz. density function, range and the parameters involved are presented in Table 1.

Step II : Testing the goodness of fit

The goodness of fit tests viz. Kolmogorov–Smirnov test and Anderson–Darling test were used along with the chi-square test at α (0.01) level of significance for the selection of the best fit probability distribution.

Step III : Identification of best fit probability distribution.

Results and Discussion

The methodology presented above was applied to the 47 years weather data in which maximum rainfall (mm) events were taken from daily rainfall. The maximum daily rainfall ranged from 38 to 229 mm during the study period. The data was classified into 23 data sets as 1 annual (January to December), 1 seasonal (June to September), 4 months of rainy season (i.e. June, July, August and September) and 17 weeks (from Standard Meteorological Week no. 23 to 39) to

Table 1. Summary of statics for maximum daily rainfall.

Study period		Mean	SD	Skew-ness	CV	Max	Min	Quar-tile Q1	Quar-tile Q3
Annual	1 Jan-31 Dec	99.248	40.288	1.09	0.40594	229	38	69	120
Seasonal	1 Jun-30 Sep	97.781	40.457	1.551	0.41376	229	38	68.25	
Jun	1 Jun-30 Jun	37.948	22.444	0.59443	0.59143	98	12	22.4	48.45
Jul	1 Jul-31 Jul	71.942	50.052	1.3091	0.69574	229	5.8	36.35	99.35
Aug	1 Aug-31 Aug	65.692	37.351	0.60473	0.56858	158	0	35.1	94.075
Sep	1 Sep-30 Sep	46.34	33.701	0.5141	0.72726	124.4	0	15.375	67.1
1 week	4 Jun-10 Jun	8.2458	16.747	3.1424	2.031	88.2	0	0	11.05
2 week	11 Jun-17 Jun	14.006	19.949	2.126	1.4243	98	0	0	22.475
3 week	18 Jun-24 Jun	13.135	17.814	1.4502	1.3562	70	0	0	25.15
4 week	25 Jun-1 Jul	21.26	22.477	1.15	1.0572	79	0	3.2	36.55
5 week	2 Jul-8 Jul	25.344	38.862	2.2794	1.5334	165.8	0	0.3	32.8
6 week	9 Jul-15 Jul	37.8	43.878	2.1745	1.1608	229	0	4.1	5.2
7 week	16 Jul-22 Jul	30.027	37.233	2.4854	1.24	100.43	0	2.35	37.9
8 week	23 Jul-29 Jul	30.577	36.373	2.1088	1.1895	159.8	0	5.55	43.7
9 week	30 Jul-5 Aug	34.663	36.004	1.2176	1.0387	140.2	0	7.55	46.65
10 week	6 Aug-12 Aug	26.833	29.634	2.1596	1.1044	152	0	8.3	33.7
11 week	13 Aug-19 Aug	30.398	32.921	1.8012	1.083	158	0	6.825	45.25
12 week	20 Aug-26 Aug	26.33	33.543	1.6766	1.2738	124.8	0	1.7	35.75
13 week	27 Aug-2 Sep	22.521	27.938	1.493	1.2405	121	0	0	39.8
14 week	3 Sep-9 Sep	23.165	31.474	1.4663	1.3587	103.56	0	0	40.7
15 week	10 Sep-16 Sep	21.148	29.195	1.6602	1.3805	120.2	0	0	34.55
16 week	17 Sep-23 Sep	6.6583	13.883	2.6717	2.085	57.6	0	0	4.95
17 week	24 Sep-30 Sep	8.2896	22.834	3.6886	2.7546	124.4	0	0	2.2

study the distribution pattern at different levels. The summary statistics (mean, standard deviation, skewness coefficient, coefficient of variation, maximum and minimum values of daily maximum rainfall) are presented in Table 1. It is obseved that mean of maximum daily rainfall of all years is 99.3 mm, seasonal mean value is 97.8 mm. The seasons of monthly and weekly values ranged from 37.9—71.9 mm and 8.2—

Table 2. Study period wise first ranked probability distribution using goodness of fit tests.

Study period		Kolmogorov Smirnov		Anderson Darling		Chi-Square	
		Distribution	Statis-tics	Distribution	Statis-tics	Distribution	Statistics
Annual	1 Jan-31 Dec	Log normal	.06115	Log Pearson 3	.17337	Gen. Gamma	0.14786
Seaso-nal	1 Jun-30 Sep	Log normal	.06256	Log Pearson 3	0.20168	Gen. Gamma	0.4011
Jun	1 Jun-30 Jun	Gen. Extreme	.0741	Gen. Extreme	0.24015	Gamma	0.62061
Jul	1 Jul-31 Jul	Log Pearson 3	.05193	Log Pearson 3	0.17845	Pearson 6	0.55541
Aug	1 Aug-31 Aug	Log Pearson 3	.05424	Log Pearson 3	0.14333	Pearson 6	0.87318
Sep	1 Sep-30 Sep	Gen. Gamma (4P)	.07823	Gen. Extreme Value	0.54376	Gen. Gamma 4P	0.43922
1 week	4 Jun-10 Jun	Log normal	.31123	Pearson 6	-6.4889	Gen. Extreme Value	14.693
2 week	11 Jun-17 Jun	Gen. Extreme Value	0.22708	Gen. Extreme Value	2.6949	Normal	5.884
3 week	18 Jun-24 Jun	Gen. Extreme Value	0.22231	Gen. Extreme Value	2.9553	Gen. Extreme Value	6.4497
4 week	25 Jun-1 Jul	Gen. Extreme Value	0.1117	Gen. Extreme Value	.95025	Gen. Extreme Value	.47341
5 week	2 Jul-8 Jul	Gen. Extreme		Gen. Extreme		Gen. Extreme	

Table 2. Continued.

Study period		Kolmogorov Smirnov Distribution	Statis-tics	Anderson Darling Distribution	Statis-tics	Chi-Square Distribution	Statistics
		Value	0.15695	Value	1.43	Value	2.8428
6 week	9 Jul-15 Jul	Gen. Extreme		Gen. Extreme		Gen. Extreme	
		Value	0.11291	Value	.7145	Value	3.5071
7 week	16 Jul-22 Jul	Gen. Extreme		Gen. Extreme		Gen. Extreme	
		Value	0.12178	Value	.75757	Value	3.609
8 week	23 Jul-29 Jul	Gen. Gamma		Gen. Extreme			
		(4p)	0.0882	Value	.75503	Gamma	1.1012
9 week	30 Jul-5 Aug	Gen. Extreme		Gen. Extreme			
		Value	0.10881	Value	0.83787	Gen. Gamma	2.159
10 week	6 Aug-12 Aug	Gen. Extreme		Gen. Extreme			
		Value	0.07612	Value	.45907	Gamma	0.51205
11 week	13 Aug-19 Aug	Gen. Extreme		Gen. Extreme			
		Value	0.9647	Value	0.55914	Weibull	0.49948
12 week	20 Aug-26 Aug	Gen. Extreme		Gen. Extreme		Gen. Extreme	
		Value	0.1359	Value	1.1737	Value	2.763
13 week	27 Aug-2 Sep	Gen. Extreme		Gen. Extreme		Gen. Extreme	
		Value	0.18118	Value	2.1432	Value	6.6409
14 week	3 Sep-9 Sep	Gen. Extreme		Gen. Extreme		Gen. Extreme	
		Value	0.16957	Value	2.3093	Value	5.8423
15 week	10 Sep-16 Sep	Gen. Extreme		Gen. Extreme		Gen. Extreme	
		Value	0.16827	Value	1.9813	Value	7.2851
16 week	17 Sep-23 Sep	Gen. Extreme				Gen. Extreme	
		Value	0.28211	Weibull	-7.5802	Value	10.573
17 week	24 Sep-30 Sep	Normal	0.37841	Weibull	-1.3423	Normal	23.265

37.8 mm, respectively. The maximum daily rainfall in a year / monsoon season was 229 mm during different months (i.e. June, July, August and September) of monsoon period, the maximum daily rainfall ranged from 98—229 mm. The weekly maximum daily rainfall ranged from 70—229 mm. The minimum among the

Table 3. Parameters of the best fitted distribution.

Study period		Distribution	Parameters		
Annual	1 Jan–31 Dec	Lognormal	σ = 0.38964	μ = 4.5217	
		Log-Pearson 3	α = 4203.0	β = 0.00607	γ = –21.006
		Gen. Gamma	k =1.0245	α = 6.3612	β = 16.354
Seasonal	1 Jun–30 Sep	Lognormal	=0.39196	= 4.5052	
		Log-Pearson 3	= 337.79	= 0.02155	= –2.7749
		Gen. Gamma	k =1.0286	= 6.1659	= 16.739
Jun	1 Jun–30 Jun	Gen. Extreme Value	k = –0.10392	σ = 19.852	μ = 28.355
		Gamma	α = 2.8589	β = 13.274	
Jul	1 Jul–31 Jul	Log-Pearson 3	α = 95.897	β = –0.07235	γ = 10.983
		Pearson 6	α_1= 3.1943	α_2 = 8.0491	β = 159.19
Aug	1 Aug–31 Aug	Log-Pearson 3	α = 5.292	β = 0.29631	γ = 5.562
		Pearson 6	α_1= 2.7836	α_2 = 8821.8	β = 2.0769E+5
Sep	1 Sep–30 Sep	Gen. Gamma	k = 4.0742	α = 0.21883	β = 104.2
		Gen. Extreme Value	k = –0.08196	σ = 29.73	μ = 31.423
1 week	4 Jun–10 Jun	Normal	σ = 16.747	μ = 8.2458	
		Weibull	α = 0.25448	β = 1.1713	
		Gen. Extreme Value	k = 0.64078	σ = 3.0883	μ = 1.121

Table 3. Continued.

Study Period		Distribution	Parameters
2 week	11 Jun–17 Jun	Gen. Extreme Value	k = 0.40414 σ = 8.0118 μ = 4.1184
		Normal	σ = 19.949 μ = 14.006
3 week	18 Jun–24 Jun	Gen. Extreme Value	k = 0.37997 σ = 7.8561 μ = 3.9353
4 week	25 Jun–1 Jul	Gen. Extreme Value	k = 0.22053 σ = 13.511 μ = 9.736

Table 3. Continued.

Study period		Distribution	Parameters
5 week	2 Jul–8 Jul	Gen. Extreme Value	k = 0.48495 σ = 12.54 μ = 6.6793
6 week	9 Jul–15 Jul	Gen. Extreme Value	k = 0.27124 σ = 22.709 μ = 16.472
7 week	16 Jul–22 Jul	Gen. Extreme Value	k = 0.32116 σ = 17.168 μ = 12.233
8 week	23 Jul–29 Jul	Gen. Gamma	k = 0.36909 α = 5.1339 β= 0.25284
		Gen. Extreme Value	k = 0.36194 σ = 15.821 μ= 12.745
		Gamma	α = 0.70671 β = 43.267
9 week	30 Jul–5 Aug	Gen. Extreme Value	k = 0.21816 σ = 21.641 μ= 16.285
		Gen. Gamma	k = 0.9585 α = 0.96901 β= 37.397
10 week	6 Aug–12 Aug	Gen. Extreme Value	k = 0.29221 σ = 14.657 μ= 12.494
		Gamma	α = 0.81989 β = 32.728
11 week	13 Aug–19 Aug	Gen. Extreme Value	k = 0.24223 σ = 18.353 μ = 14.092
		Weibull	α = 0.24405 β = 21.643
12 week	20 Aug–26 Aug	Gen. Extreme Value	k = 0.35764 σ = 15.227 μ = 9.3235
13 week	27 Aug–2 Sep	Gen. Extreme Value	k = 0.30365 σ = 14.313 μ = 8.1973
14 week	3 Sep–9 Sep	Gen. Extreme Value	k = 0.39198 σ = 13.504 μ = 6.9359
15 week	10 Sep–16 Sep	Gen. Extreme Value	k = 0.40823 σ = 11.958 μ = 6.2565
16 week	17 Sep–23 Sep	Gen. Extreme Value	k = 0.67349 σ = 2.2597 μ = 0.82868
		Weibull	α = 0.29197 β = 1.2107
17 week	24 Sep–30 Sep	Normal	σ = 22.834 μ = 8.2896
		Weibull	α = 0.28446 β = 0.94792

maximum daily rainfall was 38 mm for annual. The maximum value of coefficient of variation was observed in the seventeen week which indicates a large fluctuation in the rainfall data set and minimum value of coefficient of variation 0.406 was observed for the whole year which shows that fluctuation was minimum for the whole year. The test statistics D, A^2 and χ^2 for each data set were computed for sixteen probability distributions and the probability distribution having the first rank along with their test statistic is presented in Table 2. It was observed that Lognormal distribution using Kolmogorov Smirnov test, log-Pearson 3 value using Anderson Darling test and Gen. Gamma using Chi-square test obtained the first rank for maximum daily Annual rainfall as well as for seasonal rainfall. Thus the three probability distributions were identified as the best fit based on these three tests independently. In Table 3 parameters of the distribution was identified.

References

1. Salami AW (2004) Prediction of the annual flow regime along Asa River using probability distribution models. AMSE Modeling, measurement and control, Lyon, France 65 : 41—56.
2. Takara K, Kobayashi K (2013) Frequency analysis of extreme events for disaster management. J Disaster Res 8 : 147—148.
3. Lee C (2005) Application of rainfall frequency analysis on studying rainfall distribution characteristics of Chia Nan plain area in Southern Taiwan. J Crop Environ and Bioinform 2 : 31—38.
4. Bhaskar SR, Bansal AN, Chhajed N, Purohit RC (2006) Frequency analysis of consecutive day's maximum rainfall at Banswara, Rajasthan, India. ARPN

J Eng Appl Sci 1 : 64—67.
5. Kwaku XS, Duke O (2007) Characterization and frequency analysis of one day annual maximum and two to five consecutive days maximum rainfall of Accra, Ghana. ARPN J Eng Appl Sci 2 : 27—31.
6. Olofintoye OO, Sule BF, Salami AW (2009) Best-fit probability distribution model for peak daily rainfall of selected cities in Nigeria. NY Sci J 2 : 1—12.
7. Singh D, Singh R, Anurag, Chander Shekhar, Rao VUM, Singh S (2010) Agro climatic atlas of Haryana, Tech Bull No. 15. Dep Agric Meteorol CCS HAU, Hisar (India).