

Statistical distribution of rainfall in Uttarakhand, India

Vikram Kumar¹ · Shanu² · Jahangeer¹

Received: 20 August 2016 / Accepted: 29 June 2017 / Published online: 7 July 2017
© The Author(s) 2017. This article is an open access publication

Abstract Understanding of rainfall is an important issue for Uttarakhand, India which having varied topography and due to that extreme rainfall causes quick runoff which warns structural and functional safety of large structures and other natural resources. In this study, an attempt has been made to determine the best-fit distribution of the annual series of rainfall data for the period of 1991–2002 of 13 districts of Uttarakhand. A best-fit distribution such as Chi-squared, Chi-squared (2P), exponential, exponential (2P), gamma, gamma (3P), gen. extreme value (GEV), log-Pearson 3, Weibull, Weibull (3P) distributions was applied. Comparisons of best distributions were based on the use of goodness-of-fit tests such as Kolmogorov–Smirnov, Anderson–Darling, and Chi squared. Results showed that the Weibull distribution performed the best with 46% of the total district, while the second best distribution was Chi squared (2P) and log-Pearson. The results of this study would be useful to the water resource engineers, policy makers and planners for the agricultural development and conservation of natural resources of Uttarakhand.

Keywords Best-fit distribution · Anderson–Darling · Chi square · Kolmogorov–Smirnov

Introduction

Rainfall is considered as the main source of domestic water for living as well as for agriculture in the Uttarakhand State. The state has varied topography, leading to great variation in rainfall spatially and temporally. An analysis of the rainfall series of the state would enhance the management of water resources as well as its optimum use. One of the challenging tasks with rainfall data is to deal with interpreting past records of rainfall events in terms of future probabilities of occurrence. Therefore, the understanding of the rainfall distribution that causes flood might play an important role for the sustainable development and conservation of natural resources of the state. Estimating a statistical distribution which gives a better fit to annual rainfall has long been a concerning topic for hydrologists, meteorologists, and other water resource personnel. The understanding of rainfall distribution is important for stochastic modeling, rainfall frequency analysis, and rainfall trend analysis. The aim of such a study is not so much to explore the properties of rainfall, but to use the rainfall sequences as inputs for another modeling to understand the hydrologic processes (Buishand 1978). Distribution fitting is a way of selecting a statistical distribution that best fits available records. It is also possible to calculate return periods using various probability distributions (Upadhaya and Singh 1998).

In Costa Rica, normal distribution was the best-fitted probability distribution for annual rainfall (Waylen et al. 1996). According to Abdullah and Al-Mazroui (1998), the gamma distribution was the best for annual rainfall of Saudi Arabia. Gamma distribution also has been applied in Africa for monitoring drought (Husak et al. 2006). Rainfall distribution was also well fitted by log-Pearson type III distribution in Texas (Salami 2004) and (Lee 2005) in

✉ Vikram Kumar
25.vikram@gmail.com

¹ Department of Hydrology, IIT Roorkee, Roorkee, Uttarakhand 247667, India

² Department of Electrical Engineering, MNIT Jaipur, Jaipur, Rajasthan 302017, India

China. Naghavi and Yu (1995) had chosen the general extreme value (GEV) distribution for Louisiana and similar distribution was opted by (Pilon et al. 1991) in Ontario and (Alahmadi et al. 2014) in Saudi Arabia. Using Kappa distribution (Park and Jung 2002) generated rainfall quantile maps for Korea. Distributions by the Pearson and log-Pearson method were used in Golestan in Iran (Osati et al. 2010). Alghazali and Alawadi (2014) found that there was no suitable distribution across Iraq to describe rainfall. The annual rainfall in Sudan was best fitted by normal and gamma probability distribution (Mohamed and Ibrahim 2015). A Libyan monthly rainfall distribution was best fitted by gamma probability distribution function (Sen and Eljadid 1999). Sharma and Singh (2010) used the daily rainfall series and found that lognormal and gamma distribution was best fitted. Tao et al. (2002) propose a systematic procedure to compare the performance of different probability distributions. GEV distribution provides a good fit to the monthly rainfall data in Bangladesh (Ghosh et al. 2016). Four probability distributions: normal, log-normal, log-Pearson type III and Gumbel were applied over Pakistan to find the best-fit probability distribution of yearly rainfall recorded at 24 h (Amin et al. 2016).

The choice of a best-fit distribution also has been applied to a discharge series in the USA (Benson 1968; Vogel et al. 1993), the UK (NERC 1975), Australia (McMahon and Srikanthan 1981), and Turkey (Haktanir 1991). A review of the selection of the best distribution was given by (Curmane 1989). The rainfall runoff behavior of the Tagwai dam in Nigeria was fitted by various probability distributions and the normal distribution for the yearly daily rainfall, and log-Gumbel distribution was the most appropriate for the prediction of the yearly maximum daily runoff (Olumide et al. 2013). Phien and Ajirajah (1984) assessed the log-Pearson III distribution to the flood and maximum rainfall series. Frequency analysis of consecutive day's rainfall in Rajasthan, India, was studied (Bhakar et al. 2006) and gamma distribution was found to be the best fit for the region and the corresponding return period was estimated using the gamma function. Sabarish et al. (2017) found that log-Pearson III distribution is a best-fit probability distribution for 1-day maximum rainfall over the southern part of India. Different probability distributions have mixed results at different locations and times (Lairenjam et al. 2016). Lognormal and Gumbel EV1 distribution were adopted for discharge in the Uttarakhand region (Kamal et al. 2016).

The various distributions are mainly applied and outcomes from it help to determine the risk, uncertainty and money loss. Tao et al. (2002) stated that various probability distributions have been developed to study the distribution of rainfall. However, the choice of a best suitable distribution is still a major challenge in hydrologic practice, since

there is no general agreement as to which distribution should be used for the annual rainfall series. The selection of an appropriate distribution depends mainly on the characteristics of available rainfall data at the particular site. Hence, it is necessary to evaluate many available distributions in order to evaluate best suitable distribution that could offer true extreme rainfall. This study aimed to determine the best-fit distribution for the annual rainfall data in different districts of Uttarakhand and to evaluate their parameters.

Data and analysis

The monthly rainfall data for all districts of Uttarakhand were obtained for the period 1901–2002 from the meteorological data tool of website IWP (2015), (http://www.indiawaterportal.org/met_data/) Uttarakhand State is mainly known for two different mountainous regions, namely Kumaon and Garhwal. Most of the parts of the state is under forest cover and main rivers like Ganga and Yamuna originate from this state. All 13 districts of the state were selected for this study, which includes Almora, Bageshwar, Chamoli, Champawat, Dehradun, Haridwar, Nainital, Pauri, Pithoragarh, Rudraprayag, Tehri, Udham Singh Nagar and Uttarkashi as shown in Fig. 1. The basic characteristics such as population and geographical area of the state are given in Table 1.

1. Map of Uttarakhand
2. Annual time series of rainfall
3. Average annual precipitation of Uttarakhand from 1901 – 2002
4. Seasonal variation of rainfall



Fig. 1 Map of Uttarakhand

Table 1 Characteristics of Uttarakhand districts

District name	Population (as of 2011)	Area (km ²)	Density (per km ²)
Almora	621,972	3083	202
Bageshwar	259,840	2302	113
Chamoli	391,114	7520	51
Champawat	259,315	1781	146
Dehradun	1,659,860	3088	550
Haridwar	1,927,029	2360	817
Nainital	955,128	3860	247
Pauri Garhwal	686,572	5399	127
Pithoragarh	485,993	7100	68
Rudraprayag	236,857	1890	125
Tehri Garhwal	616,409	4080	151
Udham Singh Nagar	1,648,367	2908	567
Uttarkashi	329,686	8016	41

Table 2 Summary of statistics for annual rainfall (1901–2002)

District	Minimum (mm)	Maximum (mm)	Mean (mm)	Standard deviation (mm)
Almora	575.80	2199.30	1132.43	268.76
Bageshwar	626.38	2325.89	1220.74	276.96
Chamoli	615.78	2082.46	1131.34	245.64
Champawat	554.73	2426.77	1240.55	296.32
Dehradun	460.91	1545.15	894.15	193.48
Haridwar	406.70	1374.72	821.12	179.36
Nainital	533.40	2073.77	1071.94	258.55
Pauri Garhwal	481.10	1837.18	995.01	224.82
Pithoragarh	631.99	2306.05	1244.52	264.86
Rudraprayag	579.61	2044.70	1110.44	241.01
Tehri Garhwal	512.03	1859.77	1021.70	221.75
Udham Singh Nagar	509.52	1881.72	1004.73	233.76
Uttarkashi	557.08	1810.58	1009.83	216.82

From Table 2, it can be inferred that the annual rainfall of the Uttarakhand State is spatial and there is a wide variation in annual rainfall amount. The average maximum annual rainfall of 2426.77 mm occurred in Champawat, whereas the lowest average annual rainfall 406.70 mm occurred in Haridwar. The time series graph of different districts shows the positive or negative correlation. The time series shows two different peaks as shown in Fig. 2, the year 1936 and 1980, which are characterized by different statistical behaviors.

Methodology One of the major concerns in rainfall record is with interpreting past rainfall data in terms of future probabilities of occurrences. A large number of probability distribution methods have been applied in different regions and found to be useful for rainfall

distribution. The best-fit probability distribution in the present case was evaluated using the following procedure.

Step I: Fitting the probability distribution The probability distributions, viz. chi squared, chi squared (2P), exponential, exponential (2P), gamma, gamma (3P), gen. extreme value (GEV), log-Pearson 3, Weibull, Weibull (3P), were applied to find out the best-fit probability distribution.

Chi-squared distribution The Chi-square distribution is given by:

$$f(x; n) = \frac{\left(\frac{x}{2}\right)^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2\Gamma\left(\frac{n}{2}\right)}, \quad (1)$$

where the variable $x \geq 0$ and the parameter n , the number of degrees of freedom, is a positive integer.

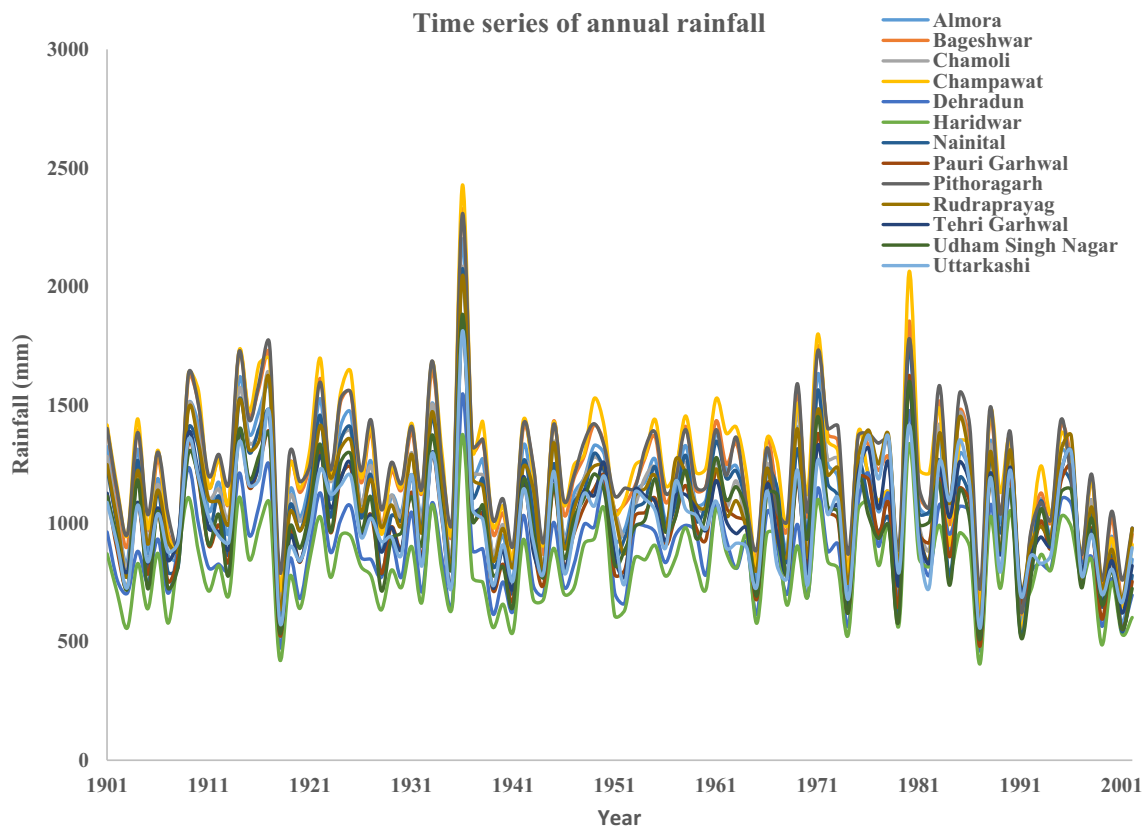


Fig. 2 Annual time series of rainfall

Exponential distribution The exponential distribution is given by:

$$f(x; n) = \frac{1}{\alpha} e^{-\frac{x}{\alpha}}, \quad (2)$$

where the variable x as well as the parameter α is a positive real quantity.

Gamma distribution The gamma distribution is given by:

$$f(x; a, b) = a(ax)^{b-1} e^{-ax} / \Gamma(b), \quad (3)$$

where the parameters a and b are positive real quantities as is the variable x .

Generalized extreme value (GEV) distribution The class of GEV distributions is very flexible with the tail shape parameter ξ (and hence the tail index defined as $\alpha = \xi^{-1}$) controlling the shape and size of the tails:

$$F_{\xi}(x) = \exp\left(-\left(1 + \xi x\right)^{-1/\xi}\right) \quad \text{with} \quad 1 + \xi x > 0, \quad (4)$$

$$\xi \neq 0.$$

The standardized GEV distribution, in the form of von Mises (1936) incorporates a location parameter μ and a scale parameter σ , in addition to the tail shape parameter, ξ , and is given by:

$$F_{\xi, \mu, \sigma}(x) = \exp\left(-\left(1 + \xi \frac{(x - \mu)}{\sigma}\right)^{-1/\xi}\right). \quad (5)$$

Log-Pearson type 3 distribution The log-Pearson 3 distribution is complicated, as it has two interacting shape parameters (Griffis and Stedinger 2007). Similar to GEV, it uses three parameters, location (μ), scale (σ) and shape (γ). A problem arises with LP3 as it has a tendency to give low upper bounds of the precipitation magnitudes, which is undesirable (Curmane 1989):

$$f(x) = \frac{(\xi - x)^{\alpha-1} e^{-(\xi-x)/\beta}}{\beta^{\alpha} \Gamma(\alpha)}, \quad (6)$$

where

$$\xi = \mu - 2\sigma/\gamma.$$

Weibull distribution The Weibull distribution is given by:

$$f(x; \eta, \sigma) = \frac{\eta}{\sigma} \left(\frac{x}{\sigma}\right)^{\eta-1} e^{-\left(\frac{x}{\sigma}\right)^{\eta}}, \quad (7)$$

where the variable x and the parameters η and σ are all positive real numbers.

Step II: Testing the goodness of fit The goodness-of-fit tests, namely, Kolmogorov–Smirnov, Anderson–Darling,

Table 3 Distribution type and their parameters

#	Distribution	Parameters					
		Almora	Bageshwar	Chamoli	Champawat	Dehradun	
1	Chi squared	$v = 1132$	$v = 1220$	$v = 1131$	$v = 1240$	$v = 894$	
2	Chi squared (2P)	$v = 35586, \gamma = -34454.0$	$v = 37796, \gamma = -36575.0$	$v = 29729, \gamma = -28598.0$	$v = 43260, \gamma = -42020.0$	$v = 18429, \gamma = -17536.0$	
3	Exponential	$\lambda = 8.8306E-4$	$\lambda = 8.1917E-4$	$\lambda = 8.8391E-4$	$\lambda = 8.0609E-4$	$\lambda = 0.001$	
4	Exponential (2P)	$\lambda = 0.002, \gamma = 575.8$	$\lambda = 0.001, \gamma = 626.38$	$\lambda = 0.0019, \gamma = 615.78$	$\lambda = 0.00146, \gamma = 554.73$	$\lambda = 0.002, \gamma = 460.91$	
5	Gamma	$\alpha = 17.75, \beta = 63.78$	$\alpha = 19.427, \beta = 62.84$	$\alpha = 21.21, \beta = 53.33$	$\alpha = 17.53, \beta = 70.78$	$\alpha = 21.36, \beta = 41.87$	
6	Gamma (3P)	$\alpha = 34.95, \beta = 45.06, \gamma = -442.51$	$\alpha = 32.47, \beta = 48.18, \gamma = -343.77$	$\alpha = 26.47, \beta = 47.45, \gamma = -124.7$	$\alpha = 37.04, \beta = 48.18, \gamma = -544.3$	$\alpha = 31.21, \beta = 34.40, \gamma = -179.72$	
7	Gen. extreme value	$k = -0.24, \sigma = 255.37, \mu = 1034.7$	$k = -0.22, \sigma = 261.63, \mu = 1117.9$	$k = -0.23, \sigma = 234.64, \mu = 1039.5$	$k = -0.23, \sigma = 277.49, \mu = 1132.1$	$k = -0.21, \sigma = 182.68, \mu = 820.06$	
8	Log-Pearson 3	$\alpha = 21.67, \beta = -0.05, \gamma = 8.14$	$\alpha = 29.1, \beta = -0.04, \gamma = 8.33$	$\alpha = 54.14, \beta = -0.03, \gamma = 8.63$	$\alpha = 17.0, \beta = -0.06, \gamma = 8.11$	$\alpha = 40.58, \beta = -0.03, \gamma = 8.18$	
9	Weibull	$\alpha = 5.13, \beta = 1219.1$	$\alpha = 5.42, \beta = 1310.3$	$\alpha = 5.67, \beta = 1212.0$	$\alpha = 5.09, \beta = 1335.9$	$\alpha = 5.63, \beta = 959.36$	
10	Weibull (3P)	$\alpha = 2.59, \beta = 739.39, \gamma = 473.83$	$\alpha = 2.61, \beta = 765.6, \gamma = 538.78$	$\alpha = 2.53, \beta = 657.0, \gamma = 547.16$	$\alpha = 2.74, \beta = 862.2, \gamma = 470.7$	$\alpha = 2.73, \beta = 555.35, \gamma = 399.16$	
#	Distribution	Parameters					
		Pauri Garhwal	Haridwar	Nainital	Pithoragarh	Rudraprayag	
1	Chi squared	$v = 995$	$v = 821$	$v = 1071$	$v = 1244$	$v = 1110$	
2	Chi squared (2P)	$v = 24900, \gamma = -23905.0$	$v = 15875, \gamma = -15055.0$	$v = 32950, \gamma = -31878.0$	$v = 34550, \gamma = -33306.0$	$v = 28613, \gamma = -27503.0$	
3	Exponential	$\lambda = 0.001$	$\lambda = 0.0012$	$\lambda = 9.3289E-4$	$\lambda = 8.0352E-4$	$\lambda = 9.0054E-4$	
4	Exponential (2P)	$\lambda = 0.002, \gamma = 481.1$	$\lambda = 0.002, \gamma = 406.7$	$\lambda = 0.00186, \gamma = 533.39$	$\lambda = 0.00163, \gamma = 631.99$	$\lambda = 0.0018, \gamma = 579.61$	
5	Gamma	$\alpha = 19.59, \beta = 50.80$	$\alpha = 20.96, \beta = 39.18$	$\alpha = 17.189, \beta = 62.362$	$\alpha = 22.078, \beta = 56.369$	$\alpha = 21.229, \beta = 52.309$	
6	Gamma (3P)	$\alpha = 42.63, \beta = 34.16, \gamma = -461.61$	$\alpha = 65.12, \beta = 22.11, \gamma = -618.81$	$\alpha = 44.511, \beta = 38.448, \gamma = -639.41$	$\alpha = 30.47, \beta = 47.538, \gamma = -203.95$	$\alpha = 23.779, \beta = 49.136, \gamma = -57.985$	
7	Gen. extreme value	$k = -0.24, \sigma = 215.48, \mu = 913.5$	$k = -0.25, \sigma = 174.96, \mu = 755.57$	$k = -0.25764, \sigma = 248.69, \mu = 980.1$	$k = -0.204, \sigma = 247.49, \mu = 1144.1$	$k = -0.216, \sigma = 229.5, \mu = 1019.1$	
8	Log-Pearson 3	$\alpha = 21.35, \beta = -0.05, \gamma = 7.95$	$\alpha = 17.92, \beta = -0.05, \gamma = 7.65$	$\alpha = 16.574, \beta = -0.062, \gamma = 7.967$	$\alpha = 42.761, \beta = -0.033, \gamma = 8.515$	$\alpha = 64.98, \beta = -0.027, \gamma = 8.762$	
9	Weibull	$\alpha = 5.38, \beta = 1069.2$	$\alpha = 5.45, \beta = 883.09$	$\alpha = 4.9907, \beta = 1156.1$	$\alpha = 5.8273, \beta = 1331.0$	$\alpha = 5.6918, \beta = 1189.3$	
10	Weibull (3P)	$\alpha = 2.77, \beta = 657.01, \gamma = 408.61$	$\alpha = 2.98, \beta = 552.38, \gamma = 327.29$	$\alpha = 2.6507, \beta = 724.35, \gamma = 426.37$	$\alpha = 2.7224, \beta = 764.17, \gamma = 562.82$	$\alpha = 2.6041, \beta = 662.05, \gamma = 521.41$	
#	Distribution	Parameters					
		Tehri Garhwal	Udhm Singh Nagar	Uttarkashi			
1	Chi squared	$v = 1021$	$v = 1004$	$v = 1009$			
2	Chi squared (2P)	$v = 24214, \gamma = -23193.0$	$v = 26954, \gamma = -25950.0$	$v = 23142, \gamma = -22133.0$			

Table 3 continued

#	Distribution	Parameters		
		Tehri Garhwal	Udham Singh Nagar	Uttarkashi
3	Exponential	$\lambda = 9.7876E-4$	$\lambda = 9.9529E-4$	$\lambda = 9.9027E-4$
4	Exponential (2P)	$\lambda = 0.002, \gamma = 512.03$	$\lambda = 0.002, \gamma = 509.52$	$\lambda = 0.002, \gamma = 557.08$
5	Gamma	$\alpha = 21.23, \beta = 48.13$	$\alpha = 18.47, \beta = 54.39$	$\alpha = 21.69, \beta = 46.55$
6	Gamma (3P)	$\alpha = 30.53, \beta = 39.83, \gamma = -194.44$	$\alpha = 68.72, \beta = 28.01, \gamma = -920.38$	$\alpha = 14.58, \beta = 56.69, \gamma = 183.32$
7	Gen. extreme value	$k = -0.21, \sigma = 209.63, \mu = 937.54$	$k = -0.29, \sigma = 229.15, \mu = 924.48$	$k = -0.18, \sigma = 203.07, \mu = 923.95$
8	Log-Pearson 3	$\alpha = 40.22, \beta = -0.03, \gamma = 8.31$	$\alpha = 14.92, \beta = -0.06, \gamma = 7.82$	$\alpha = 228.24, \beta = -0.01, \gamma = 10.15$
9	Weibull	$\alpha = 5.67, \beta = 1094.8$	$\alpha = 5.14, \beta = 1082.4$	$\alpha = 5.75, \beta = 1081.6$
10	Weibull (3P)	$\alpha = 2.75, \beta = 642.09, \gamma = 449.07$	$\alpha = 2.75, \beta = 673.21, \gamma = 404.22$	$\alpha = 2.45, \beta = 562.22, \gamma = 510.82$

and Chi-square test were used at 5% significance level for the selection of the best-fit distribution. The best-fitted distribution is selected based on the minimum error produced, which is evaluated by the following techniques:

(a) Kolmogorov–Smirnov test (K–S)

The Kolmogorov–Smirnov (K–S) test is a goodness-of-fit statistic that compares an empirical distribution function (F_x), with a specified distribution function (F_y). Many times test is used as an alternative to the Chi-square goodness-of-fit test. The Kolmogorov–Smirnov statistic (D) can be computed as:

$$D = \max |F_x(x) - F_y| \tag{8}$$

which measures the distance between the empirical distribution function F_x and the specified distribution F_y . Obviously, a large difference indicates an inconsistency between the observed data and the statistical model.

(b) Anderson–Darling test (A–D)

The Anderson–Darling (A–D) test was introduced by Anderson and Darling to place more weight or discriminating power at the tails of the distribution. This can be important when the tails of the selected theoretical distribution are of practical significance. The test statistic (AD) is defined as:

$$AD = -n - \frac{1}{n} \sum_{i=1}^n (2i - 1)(\ln(x_i) + \ln(1 - (x_{(n+1-i)}))), \tag{9}$$

where $\{x_{(1)} < \dots < x_{(n)}\}$ is the ordered (from smallest to the largest element) sample of size n , and $F(x)$ is the underlying theoretical cumulative distribution to which the sample is compared. The null hypothesis that $\{x_{(1)} < \dots < x_{(n)}\}$ comes from the underlying distribution $F(x)$ is rejected if AD is larger than the critical value AD_α at a given significance level (α).

(c) Chi-square (χ^2) test

It is a technique that checks if a specific distribution of a certain observed event’s frequency in a sample is suitable for that sample or not. Using O to define “observed count” and E to define “expected count”, the Chi-square test statistic is calculated by:

$$\chi^2 = \frac{\sum (O - E)^2}{E} \tag{10}$$

The null hypothesis states that there is no significant difference between the expected and observed frequencies. The alternative hypothesis states that they are different.

Step III: Identification of best-fit probability distribution
The goodness-of-fit test mentioned above was fitted to the

Table 4 Goodness of fit summary

District	Champawat				Dehradun				Pauri Garhwal									
	Kolmogorov–Smirnov		Anderson–Darling		Chi squared		Kolmogorov–Smirnov		Anderson–Darling		Chi squared		Kolmogorov–Smirnov		Anderson–Darling		Chi squared	
	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank
Chi squared	0.34	9	263.32	10	73.74	7	0.36	9	262.16	10	166.67	8	0.37	9	258.66	10	178.00	8
Chi squared (2P)	0.06	5	0.31	1	3.28	2	0.05	6	0.24	1	2.21	2	0.04	1	0.31	1	6.13	6
Exponential	0.42	10	28.08	9	235.19	9	0.43	10	28.79	9	247.08	9	0.44	10	29.32	9	249.60	9
Exponential (2P)	0.27	8	14.50	8	74.12	8	0.29	8	15.32	8	74.24	7	0.27	8	14.42	8	67.27	7
Gamma	0.06	6	0.51	5	7.00	6	0.05	4	0.34	4	3.58	5	0.07	7	0.37	4	3.82	1
Gamma (3P)	0.05	2	0.38	2	5.40	4	0.05	2	0.26	2	2.79	4	0.06	6	0.35	3	3.82	2
Gen. extreme value	0.06	4	4.24	7	N.A		0.05	5	4.14	7	N/A		0.05	2	4.19	7	N/A	
Log-Pearson 3	0.06	3	0.40	3	4.27	3	0.05	3	0.28	3	3.75	6	0.06	5	0.35	2	4.79	4
Weibull	0.04	1	0.42	4	0.67	1	0.04	1	0.46	6	2.05	1	0.06	3	0.64	6	4.43	3
Weibull (3P)	0.06	7	0.58	6	6.20	5	0.06	7	0.44	5	2.41	3	0.06	4	0.38	5	4.88	5
District	Champawat				Dehradun				Pauri Garhwal									
Distribution	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank
Chi squared	0.33	9	276.67	10	135.28	8	0.34	9	189.30	10	139.84	9	0.35	9	225.53	10	129.93	8
Chi squared (2P)	0.06	2	0.44	1	3.44	2	0.04	1	0.22	4	1.83	5	0.06	4	0.31	1	7.06	4
Exponential	0.42	10	28.13	9	247.66	9	0.45	10	29.37	9	253.54	10	0.43	10	28.79	9	253.31	9
Exponential (2P)	0.30	8	17.64	8	89.69	7	0.29	8	16.19	8	79.36	8	0.29	8	16.89	8	84.27	7
Gamma	0.08	6	0.59	5	4.01	3	0.04	3	0.22	3	2.60	7	0.06	3	0.45	4	5.12	3
Gamma (3P)	0.07	5	0.46	2	4.03	4	0.04	2	0.19	1	1.47	2	0.06	2	0.35	2	4.54	2
Gen. extreme value	0.06	3	4.36	7	N/A		0.04	5	0.22	5	1.71	4	0.07	6	4.25	7	N/A	
Log-Pearson 3	0.07	4	0.50	3	5.26	6	0.04	4	0.20	2	1.61	3	0.06	5	0.36	3	7.76	6
Weibull	0.05	1	0.54	4	2.74	1	0.05	7	0.58	7	0.87	1	0.05	1	0.47	5	3.35	1
Weibull (3P)	0.08	7	0.74	6	4.72	5	0.05	6	0.28	6	2.44	6	0.07	7	0.49	6	7.22	5
District	Haridwar				Nainital				Pithoragarh									
Distribution	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank	Statistic	Rank
Chi squared	0.32	9	177.04	10	135.14	9	0.34	9	262.23	10	80.79	8	0.37	9	240.09	10	183.99	9
Chi squared (2P)	0.03	1	0.18	1	2.05	5	0.05	2	0.36	1	3.05	2	0.07	6	0.33	4	1.91	2
Exponential	0.44	10	29.05	9	213.14	10	0.42	10	27.76	9	237.82	9	0.45	10	29.82	9	236.87	10
Exponential (2P)	0.29	8	16.70	8	64.97	8	0.26	8	14.59	8	78.60	7	0.31	8	17.35	8	84.10	8

Table 4 continued

District	Haridwar			Nainital			Pithoragarh											
	Kolmogorov–Smirnov Anderson–Darling Chi squared			Kolmogorov–Smirnov Anderson–Darling Chi squared			Kolmogorov–Smirnov Anderson–Darling Chi squared											
	Statistic	Rank	Statistic Rank	Statistic	Rank	Statistic Rank	Statistic	Rank	Statistic Rank									
Gamma	0.06	7	0.32	6	2.09	6	0.07	7	0.66	6	8.81	6	0.06	3	0.33	3	2.22	3
Gamma (3P)	0.05	6	0.21	2	1.47	2	0.06	5	0.46	3	5.81	4	0.06	2	0.29	1	2.64	5
Gen. extreme value	0.04	2	0.25	5	1.6534	4	0.05	3	4.32	7	N/A		0.07	5	0.37	5	2.50	4
Log-Pearson 3	0.04	4	0.21	3	1.62	3	0.06	4	0.49	4	5.79	3	0.07	4	0.30	2	1.81	1
Weibull	0.04	3	0.40	7	4.33	7	0.04	1	0.40	2	1.49	1	0.06	1	0.74	7	6.01	7
Weibull (3P)	0.05	5	0.25	4	0.73	1	0.07	6	0.65	5	6.61	5	0.08	7	0.47	6	2.64	6
District	Rudraprayag			Tehri Garhwal			Udham Singh Nagar											
Distribution	Kolmogorov–Smirnov Anderson–Darling Chi squared			Kolmogorov–Smirnov Anderson–Darling Chi squared			Kolmogorov–Smirnov Anderson–Darling Chi squared											
Chi squared	0.37	9	257.46	10	160.90	9	0.34	9	226.29	10	157.21	9	0.34	9	237.39	10	124.99	8
Chi squared (2P)	0.05	3	0.25	1	1.58	4	0.05	5	0.29	3	5.52	4	0.06	2	0.43	2	3.35	2
Exponential	0.45	10	29.29	9	257.54	10	0.45	10	29.36	9	254.26	10	0.42	10	28.28	9	234.83	9
Exponential (2P)	0.27	8	15.67	8	70.68	8	0.30	8	16.94	8	86.60	8	0.27	8	14.84	8	78.86	7
Gamma	0.06	6	0.28	5	2.26	6	0.05	4	0.29	4	4.76	2	0.08	7	0.82	6	4.81	5
Gamma (3P)	0.06	5	0.27	4	1.56	3	0.04	1	0.27	2	6.02	5	0.07	4	0.56	3	4.55	3
Gen. extreme value	0.05	1	0.29	6	1.11	1	0.05	3	0.30	5	6.15	6	0.07	5	4.39	7	N/A	
Log-Pearson 3	0.06	4	0.26	3	1.70	5	0.04	2	0.27	1	5.18	3	0.07	3	0.58	4	4.97	6
Weibull	0.07	7	0.63	7	4.48	7	0.06	7	0.64	7	3.02	1	0.05	1	0.43	1	3.17	1
Weibull (3P)	0.05	2	0.25	2	1.40	2	0.05	6	0.34	6	6.39	7	0.08	6	0.70	5	4.72	4
District	Uttarkashi			Uttarkashi			Uttarkashi											
Distribution	Kolmogorov–Smirnov			Anderson–Darling			Chi squared											
Chi squared	0.39	9	0.39	9	242.73	10	132.76	9	0.47	6	5.30	2	234.21	10	57.58	8	8.62	5
Chi squared (2P)	0.07	6	0.07	6	0.47	6	5.30	2	29.41	9	14.21	8	0.38	3	10.02	7	7.3338	4
Exponential	0.46	10	0.46	10	29.41	9	234.21	10	14.21	8	57.58	8	0.38	3	10.02	7	7.3338	4
Exponential (2P)	0.26	8	0.26	8	14.21	8	57.58	8	0.38	3	10.02	7	0.41	5	7.3338	4		
Gamma	0.05	2	0.05	2	0.41	5	7.3338	4	0.37	2								
Gamma (3P)	0.06	5	0.06	5	0.41	5	7.3338	4	0.37	2								
Gen. extreme value	0.05	3	0.05	3	0.37	2												

Table 4 continued

District	Uttarkashi		Anderson–Darling		Chi squared	
	Kolmogorov–Smirnov		Statistic		Statistic	
	Statistic	Rank	Statistic	Rank	Statistic	Rank
Log-Pearson 3	0.05	4	0.40	4	10.02	6
Weibull	0.09	7	1.03	7	5.06	1
Weibull (3P)	0.05	1	0.32	1	6.86	3

rainfall data of the study area. The test statistic was computed and tested at ($\alpha = 0.05$) level of significance. Accordingly, the ranking of different probability distributions was marked based on minimum test statistic value. The description of various probability distribution functions regarding probability density function, range and parameters are as shown in Tables 3 and 4.

Results and discussion

Analysis of rainfall data

Analysis of rainfall data plays an important role for any water resource planning as well as for hydrological modeling. The mean monthly rainfall data for all districts of Uttarakhand for 102 years (1901–2002) were used for the present study. Figure 3 displays the annual rainfall behavior recorded for whole Uttarakhand for the duration 1901–2002. The average annual rainfall recorded is 1069 mm for the whole duration. During this period, the highest amount of rainfall was about 1982.15 mm in 1936, whereas the lowest amount of rainfall recorded was about 559.98 mm during 1987. The dark line in the figure represents the annual average rainfall. If the annual rainfall in a year departs from the average annual rainfall by greater than or equal to 25%, then it is declared as a drought (meteorological drought) year (Subramanya 2008). On the basis of 25% departure from the average annual rainfall, 30.4% times there were dry years. There was sufficient rainfall from July to September to meet evapotranspiration demand and vice versa from the October to June.

To identify the seasonal rainfall distribution, the whole year was divided into three periods, namely monsoon (June–September), post-monsoon (October–February) and pre-monsoon (March–May). Figure 4 shows the seasonal variation of the rainfall. This reveals that the area receives about 82% of the total annual rainfall during the monsoon season, 10% during the post-monsoon season, and 8% during the pre-monsoon season. It indicates that more than 82% of rainfall occurs in the monsoon season, and in the remaining 8 months the crop suffers from moisture stress. Therefore, it is necessary to predict the expected rainfall to design a water conservation system. Based on the drought criteria, the years 1903, 1918, 1941, 1944, 1965, 1974, 1979, 1987, 1991 and 2001 can be characterized as drought years for Uttarakhand State.

Probability distribution

To understand the best distribution of monthly rainfall, data for all districts of Uttarakhand for 102 years (1901–2002) were analyzed. The probability analysis of

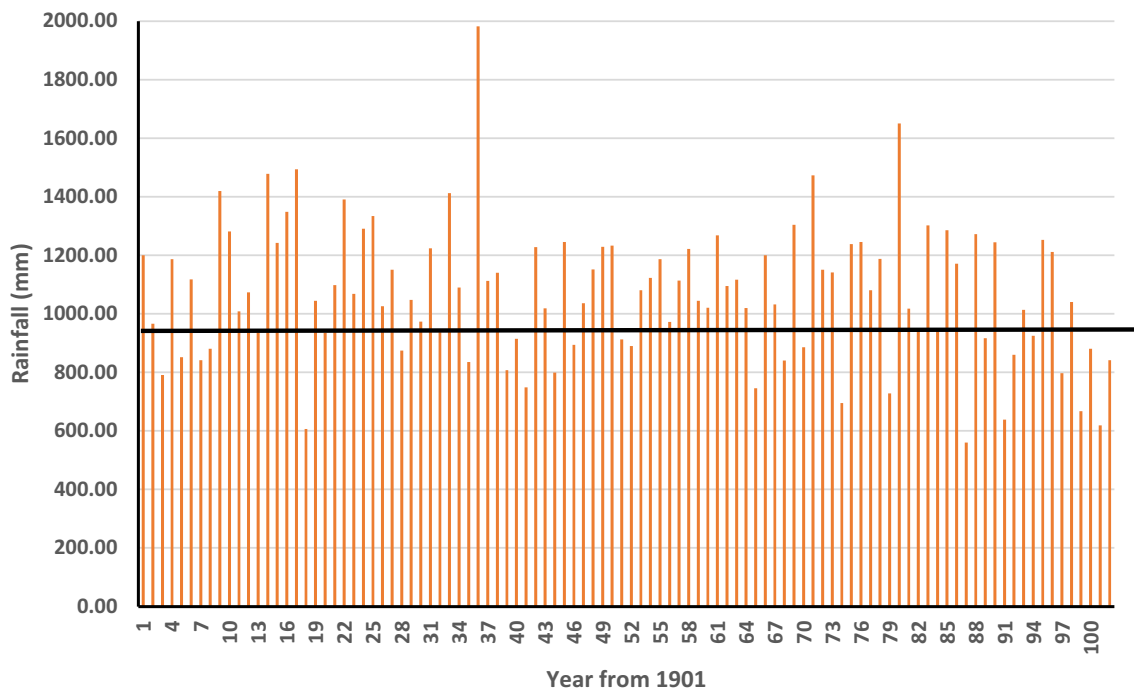
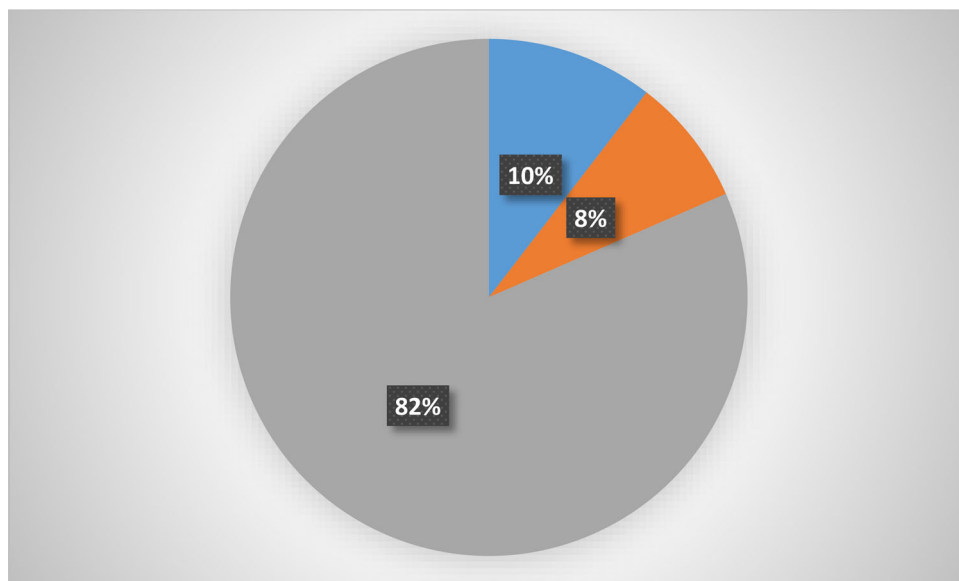


Fig. 3 Average annual precipitation of Uttarakhand from 1901 to 2002

Fig. 4 Seasonal variation of rainfall



monthly rainfall series data was carried out. A best-fit distribution, such as chi squared, chi squared (2P), exponential, exponential (2P), gamma, gamma (3P), gen. extreme value (GEV), log-Pearson 3, Weibull, Weibull (3P), was applied. Table 3 shows the distribution parameters for a different distribution. To obtain the best-fit distribution to this rainfall series, goodness-of-fit tests such as Kolmogorov–Smirnov, Anderson–Darling, and Chi squared were applied. The assessment of the best probability distribution was based on the total rank obtained

from all the tests. Ranks ranging from zero to ten (0–10) are given to each distribution model based on the criteria that the distribution(s) with the highest total score is or are chosen as the best distribution model(s) for the data of a particular district. According to the goodness-of-fit test, it was found that Weibull distribution best fitted the rainfall distribution for Almora, Bageshwar, Nainital and Udham Singh Nagar districts, Chi-squared (2P) distribution best fit for Chamoli, Champawat and Haridwar, Gamma (3P) distribution best fitted Dehradun and Pauri Garhwal, log-

Pearson 3 best fitted Pithoragarh and Tehri Garhwal and Weibull (3P) distribution best fitted Uttarkashi. The general extreme value distribution was best fitted during the monsoon (June–September). This best distribution is used to define the risk and uncertainty associated with modeling and planning of water resources. It also allows us to improve valid models which could protect us from time and economy loss.

Conclusion

The Uttarakhand State is facing the problem of quick translation of rainfall to surface runoff because of slope and faces the problems of landslides. Thus to cope up with these issues, an organized calculation of probability distribution to understand and selection of the best-fit probability distribution on an annual series of rainfall data for a period of 1991–2002 of 13 districts of Uttarakhand was made. The choice of best probability distribution could also be used to influence decisions relating to local economics and hydrologic safety systems. Annual rainfall series of all districts of the state were fitted by Chi-squared, Chi-squared (2P), exponential, exponential (2P), gamma, gamma (3P), gen. extreme value (GEV), log-Pearson 3, Weibull, Weibull (3P) distributions and comparisons of best distributions were done based on the use of goodness-of-fit tests such as Kolmogorov–Smirnov, Anderson–Darling, and Chi squared.

The goodness-of-fit test analysis indicated that Weibull, Weibull (3P), Chi-squared (2P), gamma, and log-Pearson (3P) distributions were suitable for 31, 15, 24, 15, and 15% of the stations, respectively. This study could provide a basis for choosing the best probability distribution for individual districts and corresponding distribution parameters. Further to this seasonal rainfall distribution, it reveals that the area receives about 82% of the total annual rainfall during the monsoon season, 10% during the post-monsoon season, and 8% during the pre-monsoon season. This preliminary result will help the water resource planner in hydrological modeling and the policy maker to frame general guidelines for the best use of rainfall for Uttarakhand.

Acknowledgements The authors would like to thank the anonymous referees for contributing insightful remarks and constructive suggestions, which led to a substantially improved manuscript. The authors would like to thank Shaktibala for valuable inputs.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abdullah MA, AL-Mazroui MA (1998) Climatological study of the southwestern region of Saudi Arabia. I. Rainfall analysis. *Clim Res* 9(3):213–223
- Alahmadi F, Abd Rahman N, Abdul Razzak M (2014) Evaluation of the best fit distribution for partial duration series of daily rainfall in Madinah, Western Saudi Arabia. In: *Evolving water resources system: understanding, predicting and managing water society interactions proceedings of ICWRS, Bologna, Italy*
- Alghazali NO, Alawadi DA (2014) Fitting statistical distributions of monthly rainfall for some Iraqi stations. *Civ Environ Res* 6(6):40–46
- Amin MT, Rizwan M, Alazba AA (2016) A best-fit probability distribution for the estimation of rainfall in northern regions of Pakistan. *Open Life Sci* 11(1):432–440
- Benson MA (1968) Uniform flood frequency estimating methods for federal agencies. *Water Resour Res* 4(5):891–908
- Bhakar SR, Bansal AN, Chhajed N, Purohit RC (2006) Frequency analysis of consecutive days maximum rainfall at Banswara. *ARPN J Eng Appl Sci* 1(3):64–67
- Buishand TA (1978) Some remarks on the use of daily rainfall models. *J Hydrol* 36(3):295–308
- Curmane C (1989) Statistical distributions for flood frequency analysis. Operational Hydrology Report no. 33, WMO no. 718, World Meteorological Organization, Geneva, Switzerland
- Griffis VW, Stedinger JR (2007) Log-Pearson Type 3 distribution and its application in flood frequency analysis. I: Distribution characteristics. *J Hydrol Eng* 12(5):482–491
- Ghosh S, Manindra KR, Soma CB (2016) Determination of the best fit probability distribution for monthly rainfall data in Bangladesh. *Am J Math Stat* 6(4):170–174
- Haktanir T (1991) Statistical modeling of maximum flows in Turkish rivers. *Hydrol Sci J* 36(4):367–389
- Husak GJ, Michaelson J, Funk C (2006) Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. *Int J Climatol* 27(1):935–944
- IWP (2015). http://www.indiawaterportal.org/met_data/. Accessed 22 March 2016
- Kamal V, Saumitra M, Singh P, Sen R, Vishwakarma CA, Sajadi P, Asthana H, Rena V (2016) Flood frequency analysis of Ganga river at Haridwar and Garhmukteshwar. *Appl Water Sci* 7(4):1979–1986
- Lairenjam C, Huidrom S, Bandyopadhyay A, Bhadra A (2016) Assessment of probability distribution of rainfall of North East Region (NER) of India. *J Res Environ Earth Sci* 2(9):12–18
- Lee C (2005) Application of rainfall frequency analysis on studying rainfall distribution characteristics of Chia-Nan plain area in Southern Taiwan. *J Crop Environ Bioinform* 2:31–38
- McMahon TA, Srikanthan R (1981) Log-Pearson type 3 distribution—is it applicable to flood frequency analysis of Australian streams? *J Hydrol* 52(1):139–147
- Mohamed TM, Ibrahim AAA (2015) Fitting probability distributions of annual rainfall in Sudan. *J Sci Technol* 17(2):34–39
- Naghavi B, Yu FX (1995) Regional frequency analysis of extreme precipitation in Louisiana. *J Hydraul Eng* 121(11):819–827
- NERC (1975) Flood Studies Report. Natural Environment Research Council, London
- Olumide BA, Saidu M, Oluwasesan A (2013) Evaluation of best fit probability distribution models for the prediction of rainfall and runoff volume (Case Study Tagwai Dam, Minna-Nigeria). *Int J Eng Technol* 3(2):94–98
- Osati K, Mohammed M, Karimi B, Naghi S, Mobaraki J (2010) Determining suitable probability distribution models for annual

- precipitation data (A case study of Mazandaran and Golestan provinces). *J Sustain Dev* 3(1):159–168
- Park JS, Jung HS (2002) Modeling Korean extreme rainfall using a Kappa distribution and maximum likelihood estimate. *Theor Appl Climatol* 72(1):55–64
- Phien HN, Ajirajah TJ (1984) Applications of the log-Pearson type-3 distributions in hydrology. *J Hydrol* 73(3):359–372
- Pilon PJ, Adamowski K, Alila Y (1991) Regional analysis of annual maxima precipitation using L-moments. *Atmos Res* 27(1):81–92
- Sabarish RM, Narasimhan R, Chandhru AR, Suribabu CR, Sudharsan J, Nithiyantham S (2017) Probability analysis for consecutive-day maximum rainfall for Tiruchirapalli City (south India, Asia). *Appl Water Sci* 7(2):1033–1042
- Salami AW (2004) Prediction of the annual flow regime along Asa River using probability distribution models. *AMSE Period* 65(2):41–56
- Sen Z, Eljadid AG (1999) Rainfall distribution function for Libya and rainfall prediction. *Hydrol Sci J* 44(5):665–680
- Sharma MA, Singh JB (2010) Use of probability distribution in rainfall analysis. *N Y Sci J* 3(9):40–49
- Subramanya K (2008) *Engineering hydrology*. Tata Mc-Graw-Hill Publishing Company Limited, New Delhi, pp 155–160
- Tao DQ, Nguyen VT, Bourque A (2002) On selection of probability distributions for representing extreme precipitations in Southern Quebec. *Ann Conf Can Soc Civ Eng* 1–8.
- Upadhaya A, Singh SR (1998) Estimation of consecutive days maximum rainfall by various methods and their comparison. *Indian J Soil Conser* 26(2):193–201
- von MR (1936) 'La distribution de la plus grande de n valeurs'. *Rev Math de l'Union Interbalkanique* 1:141–160
- Vogel RM, Thomas WO, McMahon TA (1993) Flood-flow frequency model selection in the Southwestern United States. *J Water Resour Plan Manag (ASCE)* 119(3):353–366
- Waylen PR, Quesada ME, Caviedes CN (1996) Temporal and spatial variability of annual precipitation in Costa Rica and Southern Oscillation. *Int J Climatol* 14(2):173–193