



Daily maximum rainfall estimation by best-fit probability distribution in the source region of Indus River

Muhammad Rizwan^{1,2} · Lubna Anjum³ · Qaisar Mehmood⁴ · Junaid Nawaz Chauhdary^{5,6,7} · Muhammad Yamin³ · Muhammad Awais⁸ · Muhammad Ansir Muneer⁹ · Muhammad Irfan¹⁰

Received: 9 November 2021 / Accepted: 13 December 2022 / Published online: 24 December 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

Abstract

Extreme events of precipitation can be guessed from best-fit probability distribution which is found through frequency analysis. The choice of best-fit probability distribution from several available distributions is a major problem. The goal of this research was the estimation of daily maximum precipitation using best-fitted probability distribution for observed data of 50 stations of the source region of Indus River from 1961 to 2015. Nine commonly used probability distributions were applied and methods of moments were used to find the parameters of applied distributions. Three goodness-of-fit tests were employed and the best-fitted probability model was selected whose sum of values from these goodness-of-fit tests was minimum. Generalized extreme value was selected as the best-fitted probability distribution on 54% of the rainfall stations, followed by log–Pearson type 3 (14% of the stations), Gamma (12% of the stations), Weibull type 3 (12% of the stations), Weibull (4% of the stations), log–normal (2% of the stations), and extreme value type 1 (2% of the stations). Then, using the best-fitted probability model at each of the rainfall station, daily maximum rainfall was estimated against different return periods. The models to minimize the threats of flooding and damages can be developed using the results of this study.

1 Introduction

For developing countries like Pakistan, agriculture plays the part of heart in economy with 21% contribution to gross domestic product (GDP) and with 3.2% annual growth (Govt.-of-Pakistan 2008). Agriculture in Pakistan is a major user of water. Precipitation provides water for agricultural, for livestock, and also for human use. In tropical countries, rainfall is the vital natural input source for the production of crops. The spatio-temporal variation, occurrence, and

distribution of precipitation are unpredictable in nature. Future probabilities of occurrence from the interpretation of past records of rainfall are the main problem in hydrology. Analysis and determination of daily maximum rainfall by probability distributions help in proper management and utilization of water resources. According to Bhakar et al. (2008), historical data can be used to find rainfall against different return periods through frequency analysis which helps to calculate occurrence probability of extreme rainfall events and can be used to estimate annual maximum precipitation

✉ Muhammad Rizwan
rizwan514@lzb.ac.cn

¹ Department of Civil Engineering, Swedish College of Engineering and Technology, Rahim Yar Khan 64200, Pakistan

² National Tibetan Plateau Data Center, State Key Laboratory of Tibetan Plateau Earth System, Resources and Environment, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100101, China

³ Faculty of Agricultural Engineering and Technology, University of Agriculture, Faisalabad 38040, Pakistan

⁴ Department of Agriculture (Field Wing), Government of Punjab, Lahore, Pakistan

⁵ Water Management Research Center, University of Agriculture Faisalabad, Faisalabad 38000, Pakistan

⁶ Research Centre of Fluid Machinery Engineering and Technology, Jiangsu University, Zhenjiang 212013, Jiangsu, China

⁷ Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, IN 47907-2093, USA

⁸ College of Earth & Environmental Sciences, University of the Punjab, Lahore, Pakistan

⁹ Adaptive Research Farm, Bhaun, Chakwal, Punjab, Pakistan

¹⁰ PMU, Agriculture Department, Government of Punjab, Lahore, Pakistan

for different return periods using the rainfall data (Bhakar et al. 2006).

The quantity and pattern of rainfall for specific place are significant parameters which affect management of water resources, flood protection, tourism, forestry, and agriculture. The floods are mainly caused by extreme rainfall events. The damages, as a result of floods and storms, can be minimized by accurate approximation of expected precipitation along with proper design of hydraulic structures. According to Tao et al. (2002), the data on extreme rainfall events with high return periods is prerequisite for operation and control of water resource policies, to lessen the damages caused by floods, and for the safe designing of hydrologic structures such as dams and urban drainage systems. Bhakar et al. (2008) reported that for good economic return, the farmers should be aware of the prediction of rainfall through scientific methods and systematic planning of their crops. Thus, many problems associated with water management can be solved through probability analysis of rainfall data. George and Kolappadan (2002) reported the prediction of rainfall of various quantities and frequencies by using probability distributions. As there are spatial and temporal variations in rainfall, projected rainfall during various return periods are estimated using different probability distributions. The projected rainfall which might be more or less than the recorded value is estimated using the best-fitted probability model. Precipitation data frequency analysis has been accomplished for various return periods (Barkotulla et al. 2009; Bhakar et al. 2006; Vivekanandan 2012).

In engineering practice, the main problem is to select proper distribution model. The main criterion for the selection of proper distribution model for a particular site is the available data of precipitation. Assessment of existing distribution models is required to choose appropriate distribution model to get the precise estimate of extreme rainfall. According to Tao et al. (2002), several probability models have been developed for the description of distribution for extreme precipitation values at a particular site. Anil (2000) and Singh (2001) reported that log-normal distribution fitted best to 24-h annual maximum precipitation in India. According to Bhakar et al. (2008), Gumbel distribution fitted best to monthly maximum rainfall in India. In Iran, generalized extreme value distribution and Pearson type 3 distribution were the best-fitted probability distributions for monthly maximum rainfall (Eslamian and Feizi 2007). In lower parts of northern areas of Pakistan, log-Pearson type 3 probability distribution fitted best to 24-h annual maximum rainfall values (Amin et al. 2016). Weibull, log-normal, and Pearson type 5 were the best-fitted distributions to monsoon, pre-monsoon, and winter seasons while for annual, summer, and post-monsoon seasons, normal distribution was selected as the best-fitted distribution, based on the study of 24-h monthly, seasonal, and annual maximum

precipitation for Sagar Island which is on the continental shelf of Bay of Bengal (Mandal and Choudhury 2015). According to Lee (2005), log-Pearson type 3 was selected as the best-fitted probability model to 50% of stations of Chia-Nan plain area in Southern Taiwan. For 24-h maximum rainfall, log-Pearson type 3 was selected as the best-fit model in Nigeria (Ogunlela 2001). According to Olofintoye et al. (2009), log-Pearson type 3 probability model fitted well to half of the stations in Nigeria and Pearson type 3 probability model fitted well to 40% of stations for 24-h maximum rainfall. According to Kwaku and Duke (2007), log-normal was selected as the best-fitted probability model for maximum precipitation of 1–5 consecutive days at in Ghana. For monthly maximum rainfall, Gamma distribution was selected as the best-fitted probability model in the arid areas of Libya (ŞEN and Eljadid 1999).

Most of the population in Pakistan depends on agriculture for their food and fiber requirements as agriculture is the major contributor in the country's economy. Water is essential for agriculture and precipitation provides water for agricultural production (Adnan and Khan 2009). Agriculture, biological diversity, and ecosystem are directly affected by extreme events of precipitation. It is therefore vital to estimate expected events of extreme rainfall to minimize the risk factors in the long-term measures of saving property and lives. At present, very limited work have been done in Pakistan to find the best-fitted probability distributions. No study has been carried out in SRIR which contributes water to rivers in Pakistan.

The current research was conducted in the study area that is the lifeline for the Indus Basin as it provides water for the inhabitants and for agriculture of this basin. Previously, only one study of this kind was conducted in the southwestern part of SRIR using rainfall data of only six rainfall stations and only four probability distributions. But, this study was conducted in the whole source region of Indus River and rainfall data of 50 stations was used. The objective of this research was the selection of appropriate probability model for daily maximum rainfall from 1965 to 2015 and to calculate the expected rainfall against the return period of 5, 10, 15, 20, 25, 50, and 100 years. The predicted amounts of rainfall will be helpful for making policies and developing plans to minimize the damages and risks of flooding from extreme rainfall events.

2 Materials and methods

2.1 Study area and data collection

This research was conducted in SRIR (Fig. 1) where elevation ranges from 193 to 8062 m above sea level. The Indus basin with a drainage area of about 1.08 million km² is

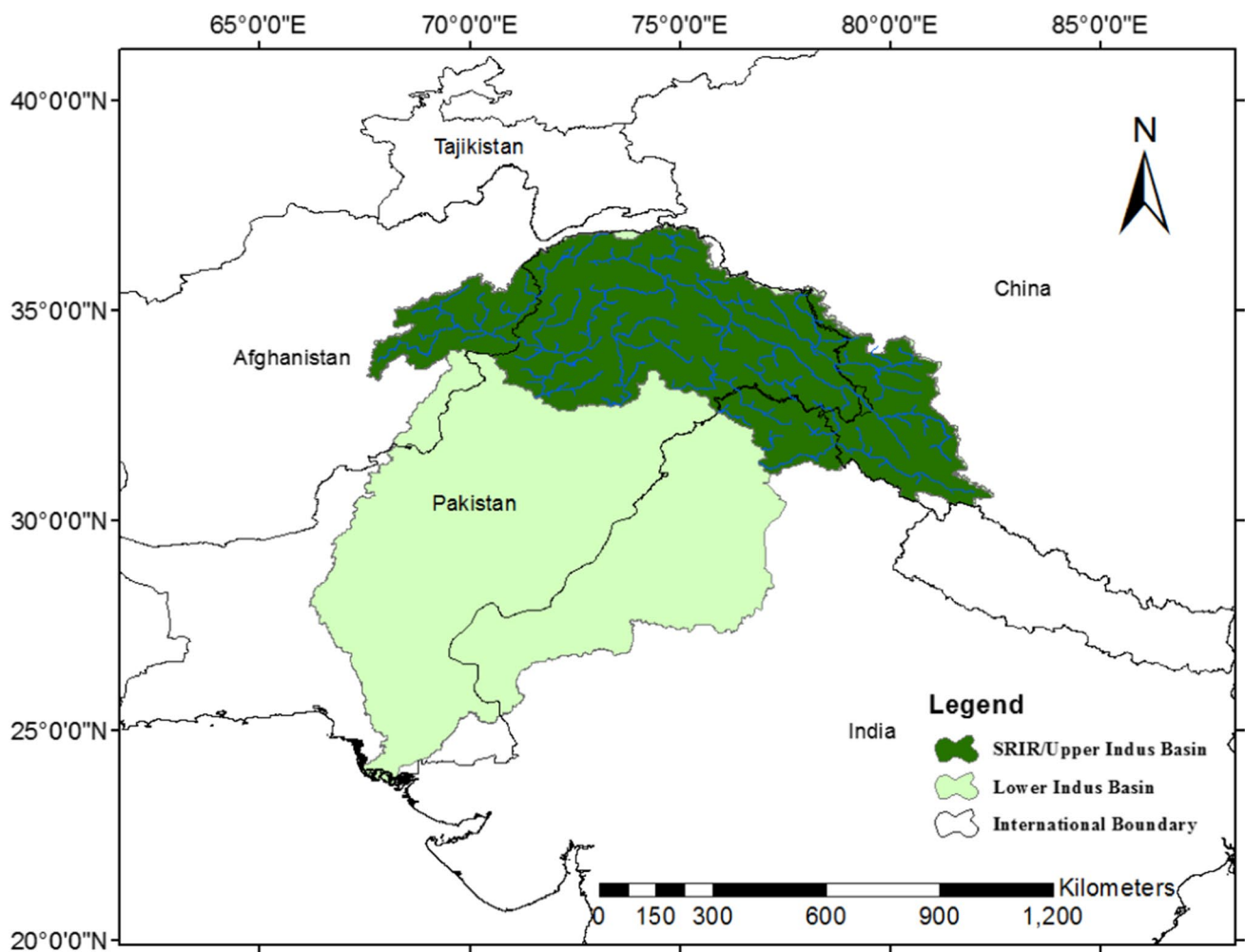


Fig. 1 Location map of study area

among the largest trans-boundary river basins with the major part being in Pakistan, i.e., 56% while 26.6% of area lies in India, 10.7% lies in China, and 6.7% lies in Afghanistan (Wolf et al. 1999). Upper Indus river basin is a distinctive area with multifarious climate (Lutz et al. 2016), diverse physical and geographical features, and divergent hydrological systems (Hasson et al. 2017). Indus River initiates from Mansarovar Lake in the Third Pole and flows through Northern Pakistan and ends its journey into the Arabian Sea. Source region of Indus River, also called Upper Indus Basin (UIB), is positioned in the mountainous range of Tibetan Plateau, Himalaya, Karakoram, and Hindu–Kush in the global range of 32.48° to 37.07° N and 67.33° to 81.83° E (Hasson et al. 2017; Khattak et al. 2011; Lutz et al. 2016). There are about 11,000 glaciers in these mountainous ranges (Hasson et al. 2017); with a glacier surface area of about 22,000 km², it becomes one of the most glaciated areas in the world (Bajracharya and Shrestha 2011). Indus River and its tributaries (Jhelum, Chenab, Ravi, Satluj, and Kabul)

flow through this region. Jhelum River is among the largest tributary of Indus River. The drainage area of Jhelum River Basin is approximately 33,435 km² and it lies in the global range of 33 to 35° N and 73 to 75.62° E. The drainage point of Jhelum River Basin is the Mangla reservoir which is the largest reservoir in Pakistan after the Tarbela reservoir. There are five rivers/tributaries (Jhelum, Kanshi, Poonch, Neelum, and Kunhar) which flow through this basin and add water to Mangla reservoir. Jhelum River Basin and its tributaries get water from the southern Himalaya and some parts of Pir Panjal mountains in Kashmir. The digital elevation model of SRIR and location of meteorological stations and gridded stations are shown in Fig. 2.

Twenty-four-hour annual maximum rainfall from 28 meteorological stations and 22 gridded stations of the study area during the period of 1961–2015 were used in this study. Meteorological stations are located within the boundary of Pakistan and rainfall data of these stations are obtained from Pakistan Meteorology Department. Gridded

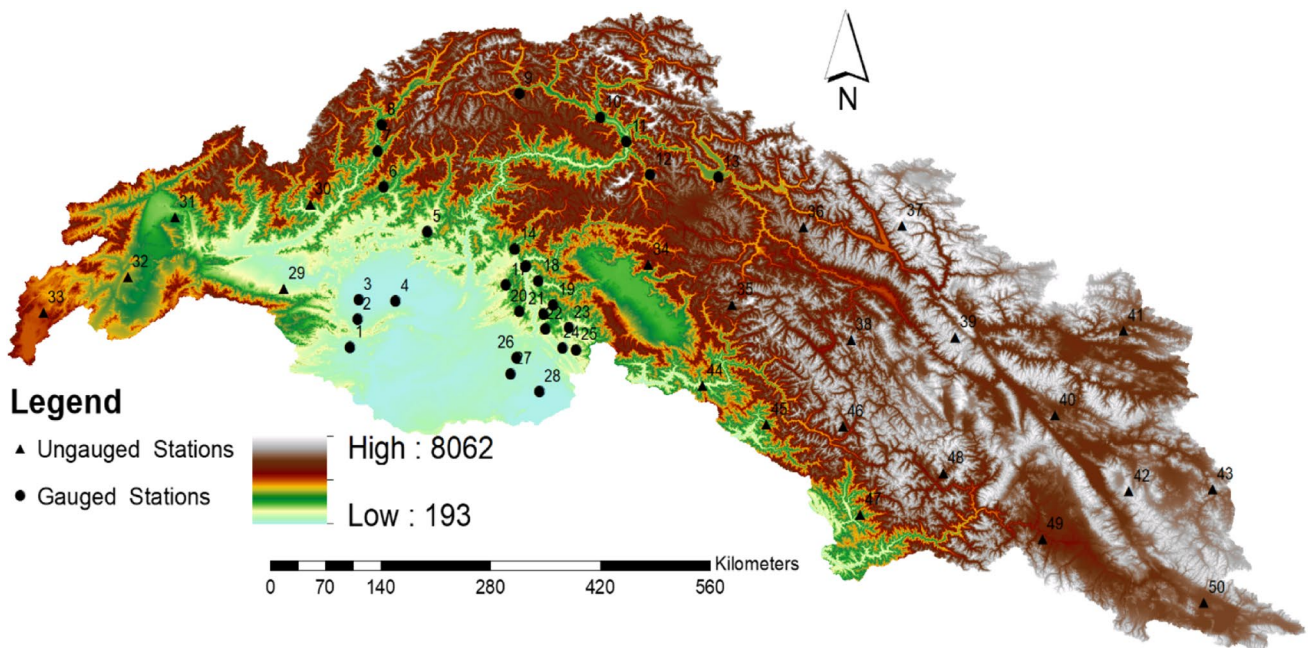


Fig. 2 Study area, showing elevation and location of rainfall stations

stations are located outside the boundary of Pakistan and rainfall data of these stations are extracted from APHRODITE. A summary of statistics along with basic information about the rainfall stations is shown in Table 1. Most of the stations (36 No.) of this region are highly skewed as their coefficient of skewness is greater than 1 while 14 stations are moderately skewed.

2.2 Probability distributions

A probability distribution is a method for showing the probable outcomes that a random variable may have along with their probabilities. The choice of proper probability model is very vital for selection of best-fitted model for a specific area. Most generally used probability distributions for the analysis of extreme precipitation are used in this study. These distributions are described in Table 2. Out of nine distributions, five have two parameters while four have three parameters. A distribution with more parameters is generally expected to give better results but its estimation method will be more difficult.

2.3 Goodness-of-fit test

The fitness of probability distributions is usually checked by the statistics of goodness-of-fit tests. Three most generally used goodness-of-fit tests are used in this study and are described below:

Kolmogorov–Smirnov test

This test finds whether the sample is from assumed continuous distribution. The statistic (D) of this test is the maximum vertical change between theoretical distribution and empirical distribution (Conover 1998). It is defined as:

$$D_n = \max |F_n(x) - F(x)| \quad (1)$$

The hypothesis that the data follow the selected probability model is rejected if the calculated value of D is higher than its critical value that is 0.17981 at $\alpha=0.05$.

Anderson–Darling test

It is a common statistical test to find whether the dataset follows the given probability model. The Anderson–Darling test statistics (A^2) is calculated using Eq. (2):

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) [\ln F(X_i) + \ln(1-F(X_{n-i+1}))] \quad (2)$$

The hypothesis that data comes from a particular probability model is rejected if the calculated value of A^2 is higher than its critical value that is 2.5018 at $\alpha=0.05$.

Chi-squared test

This test is based on the difference between observed and expected values of each class. Best-fitted distribution

Table 1 Statistics of 24-h annual maximum rainfall and characteristics of rainfall stations' location

	Stations	Statistics of 24-h annual maximum rainfall				Lat. (dd)	Lon. (dd)	Elev. (m)
		Avg. value (mm)	Standard deviation	Coeff. of variation	Skewness			
1	Kohat	55.5	45.1	0.8	2.7	33.5	71.5	466
2	Cherat	63.4	41.6	0.7	2.6	33.8	71.6	1372
3	Peshawar	58.7	41.4	0.7	2.9	34.0	71.6	327
4	Risalpur	68.2	42.5	0.6	2.4	34.0	72.0	315
5	Saidu sharif	71.1	37.0	0.5	2.1	34.7	72.4	961
6	Dir	77.8	32.9	0.4	2.0	35.2	71.9	1375
7	Darosh	49.8	24.8	0.5	2.6	35.6	71.8	1464
8	Chitral	45.7	27.0	0.6	2.8	35.9	71.8	1498
9	Gupis	34.1	34.9	1.0	2.6	36.2	73.4	2156
10	Gilgit	26.5	15.1	0.6	0.7	35.9	74.3	1460
11	Bunji	26.6	13.0	0.5	0.8	35.7	74.6	1372
12	Astore	44.8	18.9	0.4	1.0	35.3	74.9	2168
13	Sakardu	38.6	69.3	1.8	6.7	35.3	75.7	2317
14	Balakot	110.7	56.2	0.5	1.2	34.6	73.4	995
15	Domel	85.8	43.0	0.5	1.5	34.4	73.5	686
16	Muzafarabad	95.1	38.7	0.4	1.2	34.4	73.5	702
17	Kakul	80.1	33.7	0.4	2.6	34.2	73.3	1308
18	Gari Dupata	82.9	38.8	0.5	1.3	34.2	73.6	814
19	Bagh	71.4	41.0	0.6	0.8	34.0	73.8	1138
20	Murree	104.1	48.5	0.5	1.3	33.9	73.4	2168
21	Rawalkot	78.8	46.8	0.6	1.6	33.9	73.7	1677
22	Plandri	98.5	52.0	0.5	0.9	33.7	73.7	1402
23	S. Kakuta	84.0	44.2	0.5	0.8	33.7	74.0	915
24	Kotli	83.4	41.1	0.5	0.9	33.5	73.9	614
25	Khandar	74.2	51.9	0.7	1.7	33.5	74.1	1067
26	Kallar	78.7	43.8	0.6	2.0	33.4	73.4	518
27	Gujjar Khan	72.8	42.7	0.6	2.1	33.3	73.3	457
28	Mangla	80.8	36.2	0.4	0.6	33.1	73.6	283
29	Afgh1	30.7	18.8	0.6	2.5	34.1	70.7	926
30	Afgh2	52.7	27.6	0.5	2.2	35.0	71.0	3029
31	Afgh3	24.3	15.4	0.6	3.0	34.9	69.5	1994
32	Afgh4	23.2	16.4	0.7	3.7	34.3	68.9	2564
33	Afgh5	19.3	15.2	0.8	4.0	33.9	68.0	3114
34	Kash2	42.4	19.2	0.5	1.6	34.4	74.9	3653
35	Kash3	28.1	18.0	0.6	3.8	34.0	75.8	4565
36	Kash5	15.4	9.7	0.6	1.4	34.8	76.7	4974
37	Kash6	13.5	10.2	0.8	2.2	34.8	77.8	5887
38	Kash4	10.9	7.6	0.7	2.4	33.6	77.2	5413
39	Kash7	9.3	7.4	0.8	2.7	33.6	78.4	5634
40	China1	8.6	4.4	0.5	0.8	32.8	79.5	4639
41	China2	7.7	4.2	0.6	1.4	33.7	80.3	5449
42	China4	16.3	7.7	0.5	0.9	32.0	80.4	5577
43	China3	12.2	5.6	0.5	1.2	32.0	81.3	5114
44	Kash1	61.0	23.8	0.4	0.6	33.1	75.5	1597
45	India1	93.0	33.3	0.4	0.7	32.7	76.2	2726
46	India2	35.6	19.9	0.6	1.9	32.7	77.1	5279
47	India3	43.0	13.3	0.3	1.0	31.8	77.3	1897
48	India4	23.2	8.0	0.3	0.5	32.2	78.3	5111
49	India5	24.7	11.6	0.5	1.9	31.5	79.4	4397

Table 1 (continued)

Stations	Statistics of 24-h annual maximum rainfall				Lat. (dd)	Lon. (dd)	Elev. (m)
	Avg. value (mm)	Standard deviation	Coeff. of variation	Skewness			
50 India6	23.3	13.3	0.6	1.4	30.9	81.2	4591

Table 2 Probability distributions and their parameters

Distribution	Probability density function and cumulative distribution function	Parameters and range
Exponential	$f(x) = \lambda \exp(-\lambda(x - \gamma))$ $F(x) = 1 - \exp(-\lambda(x - \gamma))$	λ : scale parameter γ : location parameter $(0 < x < \infty)$
Extreme value type 1 (EV1)	$f(x) = \frac{1}{\delta} \exp(-z - \exp(-z))$ $F(x) = \exp(-\exp(-z))$ Where $z = \frac{x-\mu}{\delta}$	δ : scale parameter μ : location parameter $(-\infty < x < \infty)$
Log-normal (LN)	$f(x) = \frac{\exp\left(-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right)}{x\sigma\sqrt{2\pi}}$ $F(x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right)$	σ : continuous-parameter μ : continuous-parameter ϕ : Laplace Integral $(0 < x < \infty)$
Normal	$f(x) = \frac{\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)}{\sigma\sqrt{2\pi}}$ $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$	σ : scale parameter μ : location parameter ϕ : Laplace Integral $(-\infty < x < \infty)$
Weibull (W)	$f(x) = \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left(-\left(\frac{x}{\beta}\right)^\alpha\right)$ $F(x) = 1 - \exp\left(-\left(\frac{x}{\beta}\right)^\alpha\right)$	α : shape parameter β : scale parameter $0 < x < \infty$
Gamma	$f(x) = \frac{(x-\gamma)^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} \exp\left(-\frac{x-\gamma}{\beta}\right)$ $F(x) = \Gamma\left(\frac{x-\gamma}{\beta}, \alpha\right)$	α : shape parameter β : scale parameter γ : location parameter Γ : gamma function $0 < x < \infty$
Generalized extreme value (GEV)	$f(x) = \frac{1}{\delta} \exp\left(-\left(1 + kz\right)^{-\frac{1}{k}}\right) \left(1 + kz\right)^{-1 - \frac{1}{k}}$ $F(x) = \exp\left(-\left(1 + kz\right)^{-\frac{1}{k}}\right)$ Where $z = \frac{x-\mu}{\delta}$	k : shape parameter δ : scale parameter μ : location parameter $(-\infty < x < \infty)$
Log-Pearson type 3 (LP3)	$f(x) = \frac{1}{x \beta \Gamma(\alpha)} \left(\frac{\ln(x)-\gamma}{\beta}\right)^{\alpha-1} \exp\left(-\frac{\ln(x)-\gamma}{\beta}\right)$ $F(x) = \frac{\Gamma\left(\frac{\ln(x)-\gamma}{\beta}, \alpha\right)}{\Gamma(\alpha)}$	α : shape parameter β : scale parameter γ : location parameter $0 < x < \infty$
Weibull 3P (W3P)	$f(x) = \frac{\alpha}{\beta} \left(\frac{x-\gamma}{\beta}\right)^{\alpha-1} \exp\left(-\left(\frac{x-\gamma}{\beta}\right)^\alpha\right)$ $F(x) = 1 - \exp\left(-\left(\frac{x-\gamma}{\beta}\right)^\alpha\right)$	α : shape parameter β : scale parameter γ : location parameter $0 < x < \infty$

was selected by comparing each distribution’s chi-square value and selecting the function that gives the smallest chi-square value (Agarwal et al. 1988). The statistic of this test is defined as:

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i} \tag{3}$$

where O_i is the observed number and E_i is the expected number of cases in class i . The critical value for the chi-squared test is 11.07 at $\alpha = 0.05$.

2.4 Return period (T)

Calculation of return period is the main task of frequency analysis. Exceedance probability is inverse of return

Table 3 Sum of values of goodness-of-fit tests for probability distributions used in the study

Sr. no	Rainfall stations	Exponential	Gamma	GEV	EV1	LN	LP3	N	W	W3P
1	Astore	27	11	6	21	23	8	14	15	10
2	Bagh	24	22	12	14	18	12	16	6	11
3	Balakot	26	18	4	15	7	7	25	17	16
4	Bunji	27	11	12	19	14	13	17	12	10
5	Cherat	22	19	6	8	14	R	18	17	24
6	Chitral	24	12	7	15	12	12	27	11	15
7	Darosh	26	13	7	16	8	11	25	17	13
8	Dir	27	8	4	19	13	6	24	20	14
9	Domel	25	6	11	7	18	17	14	20	14
10	Gari Duptata	27	11	16	13	14	12	19	16	7
11	Gilgit	19	4	15	21	14	10	26	19	7
12	Gujar Khan	27	10	8	14	10	14	22	16	14
13	Gupis	20	14	5	22	6	7	27	13	21
14	Kakul	23	10	3	15	9	10	22	22	12
15	Khandar	13	18	10	13	12	7	21	9	15
16	Kohat	16	15	3	23	8	7	25	17	12
17	Kotli	27	14	9	16	23	9	16	6	15
18	Kallar	24	9	9	16	16	8	27	17	10
19	Mangla	22	13	6	7	11	R	12	23	14
20	Muzafarabad	14	9	10	10	18	14	24	25	11
21	Murree	25	19	5	13	6	10	26	16	15
22	Peshawar	22	16	3	9	7	R	16	23	12
23	Plandri	27	14	11	17	17	12	18	10	9
24	Rawalkot	27	12	5	17	6	8	24	19	18
25	Risalpur	22	7	6	13	9	R	13	23	15
26	Kakuta	25	23	7	15	21	6	20	8	10
27	Saidu Sharif	27	13	3	17	6	9	24	16	20
28	Sakardu	14	16	6	23	7	13	22	12	18
29	Afgh1	22	15	5	21	9	4	27	18	14
30	Afgh2	16	11	4	22	13	5	27	23	14
31	Afgh3	21	14	8	23	5	9	25	14	16
32	Afgh4	20	13	6	23	7	10	27	12	17
33	Afgh5	20	9	6	24	9	7	27	17	16
34	China1	24	6	16	21	12	10	20	21	5
35	China2	25	14	9	14	10	3	26	17	17
36	China3	21	7	12	14	15	14	27	16	9
37	China4	27	6	12	21	15	10	22	19	3
38	India1	27	10	9	7	12	11	24	21	14
39	India2	26	12	5	18	7	6	25	19	17
40	India3	27	4	11	16	18	12	19	23	5
41	India4	27	10	7	20	16	15	18	16	6
42	India5	23	14	5	17	7	6	27	21	15
43	India6	23	13	5	18	11	4	25	17	19
44	Kash1	27	8	5	20	23	8	18	10	16
45	Kash2	27	6	8	17	15	7	24	21	10
46	Kash3	26	12	3	21	9	6	24	17	17
47	Kash4	17	14	3	23	9	6	27	21	15
48	Kash5	22	11	5	20	10	4	27	21	15
49	Kash6	20	12	4	23	11	7	25	22	11
50	Kash7	16	15	4	24	7	7	27	20	15

R = probability distribution is rejected.

Bold italic values = best-fitted probability distribution.

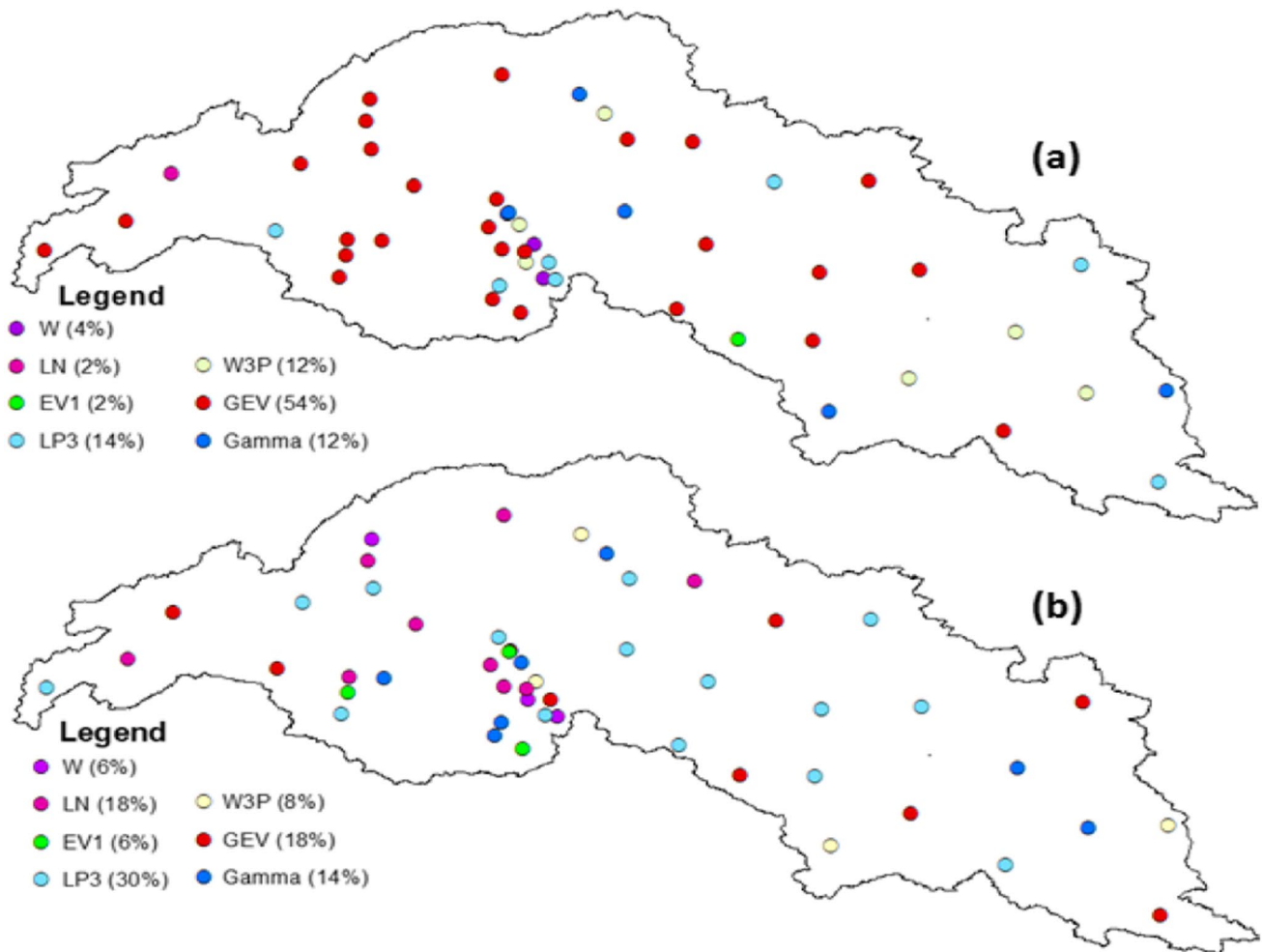


Fig. 3 Best-fit (a) and second best-fit (b) probability distributions at rainfall stations of SRIR

period. A 5-year return period has an exceedance probability of $1/5 = 0.2$ in EACH YEAR. If a variable whose value (x) becomes equal or higher than the event with value (x_T) once in T years, then exceedance probability (P) of the variable in a given year is given by:

$$P(x \geq x_T) = \frac{1}{T} \quad (4)$$

$$T = \frac{1}{P} \quad (5)$$

3 Results and discussion

3.1 Best-fit probability distribution

Twenty-four-hour annual maximum precipitation from 50 rainfall stations in SRIR from 1961 to 2015 was used in this

study. All the rainfall stations are located at different elevations (Fig. 2). Fourteen stations are located at an elevation of 0–1000 m and the average value of their daily maximum rainfall was 75.6 mm, and twenty stations are located at an elevation of 1001 to 3000 m with average rainfall of 57 mm while sixteen stations are located at an elevation of 3001 to 6000 m, and their average daily maximum rainfall was 21.4 mm. Average amount of daily maximum rainfall is decreasing with increasing elevation in the study area.

Nine probability distributions were used in this research work; five distributions have 2 parameters while four distributions have 3 parameters which were calculated using methods of moments. Three goodness-of-fit tests were applied to find the most suited probability distribution at each of the rainfall station in the study area. These goodness-of-fit tests were used following the procedure adopted by numerous authors in previous studies (Adegboye and Ipinoyomi 1995; Chowdhury et al. 1991; Leavenworth and Grant 2000). The test statistics were calculated using Eq. (1) for the Kolmogorov–Smirnov

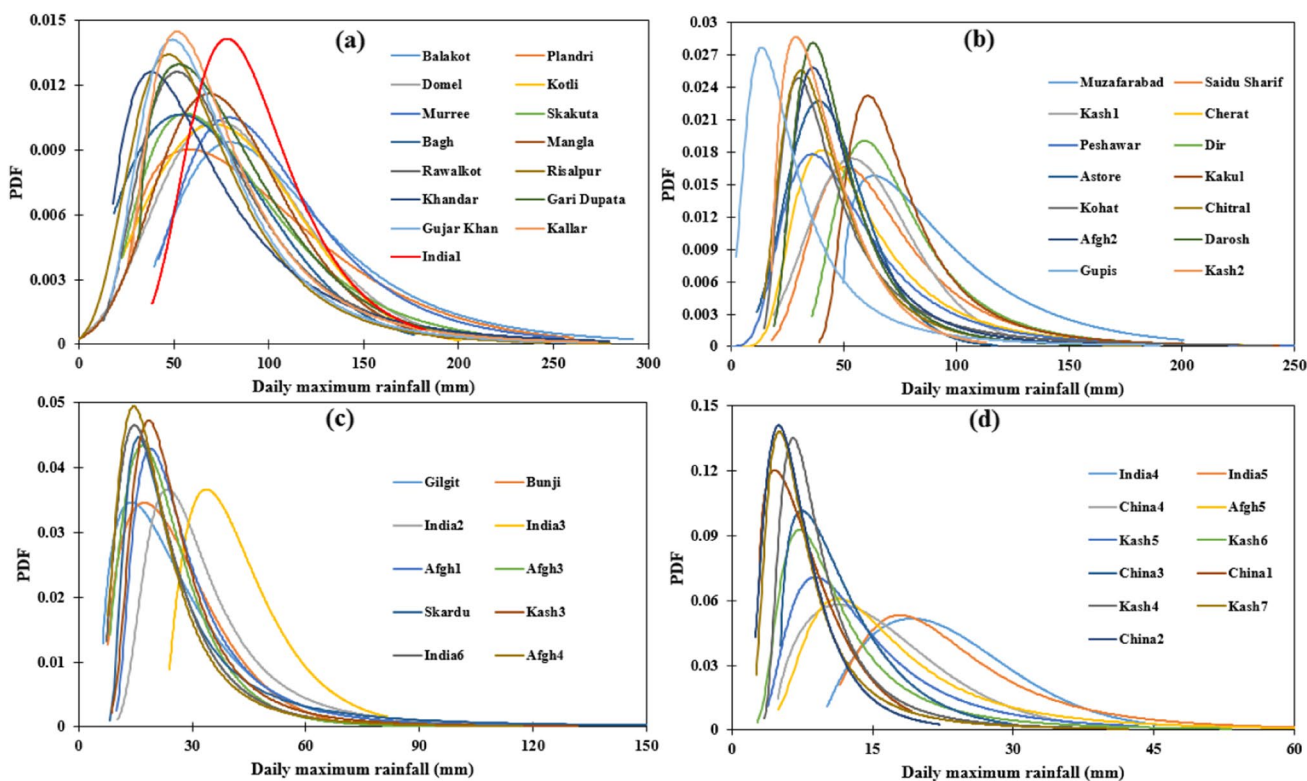


Fig. 4 Probability density functions at rainfall stations of SRIR. **a** Daily maximum rainfall ≥ 250 mm. **b** Daily maximum rainfall 150–250 mm. **c** Daily maximum rainfall 100–150 mm. **d** Daily maximum rainfall ≤ 60 mm

test, Eq. (2) for the Anderson–Darling test, and Eq. (3) for the chi-squared test. The critical values were obtained from table at significance level of 0.05. The hypothesis that the rainfall data follow the probability distribution is rejected if the calculated value of test statistics is more than its critical value. These goodness-of-fit tests were given the value of 1 for best-fitted and 9 for least-fitted probability distribution. The values of all goodness-of-fit tests used in this study were added for all the rainfall stations and are shown in Table 3. It is clear from Table 3 that only log–Pearson type 3 is rejected at four stations, i.e., Cherat, Mangla, Peshawar, and Risalpur. The best-fitted probability model was selected based on total minimum scores of values from used goodness-of-fit test and is shown by gray color in Table 3. The best-fitted probability models based on these results for all rainfall stations are shown in Fig. 3a. As can be seen from Fig. 3a, generalized extreme value (GEV) is the best-fitted model at 54% of the stations in the source region of Indus River (SRIR). Log–Pearson type 3 (LP3) is the best-fitted distribution at 14% of the stations, followed by Gamma (12%), Weibull 3P (12%), Weibull (4%), log–normal (2%), and extreme value type 1 (2%) in the source region of Indus River. According to Khudri and Sadia (2013), GEV was also the best-fitted probability model

for annual maximum rainfall in Bangladesh. Eslamian and Feizi (2007) also reported GEV and Pearson type 3 as the best-fitted probability models for maximum monthly values of rainfall in Iran. Second best-fitted models are shown in Fig. 3b. Log–Pearson type 3 (LP3) is the second best-fit probability distribution on most of the station in the study area.

Probability density function and cumulative distribution function at 50 rainfall stations of the study area were calculated using the best-fitted model of that station and results are presented in Figs. 4 and 5. Fifteen rainfall stations whose daily maximum rainfall is greater than or equal to 250 mm have lower probability, i.e., less than 1.5% (Fig. 4a), fourteen stations with daily maximum rainfall of 150–250 mm have probability up to 3% (Fig. 4b), and ten stations with daily maximum rainfall of 100–150 mm have probability up to 5% (Fig. 4c) while eleven rainfall stations with daily maximum rainfall less than equal to 60 mm have highest probability, i.e., up to 15% (Fig. 4d). The cumulative probability for fifteen rainfall stations is about 1 with daily maximum rainfall equal to or less than 250 mm (Fig. 5a), the cumulative probability is about 1 for fourteen stations with rainfall less than or equal to 150 mm (Fig. 5b), and the cumulative probability is about 1 for ten stations with daily maximum rainfall less than equal to

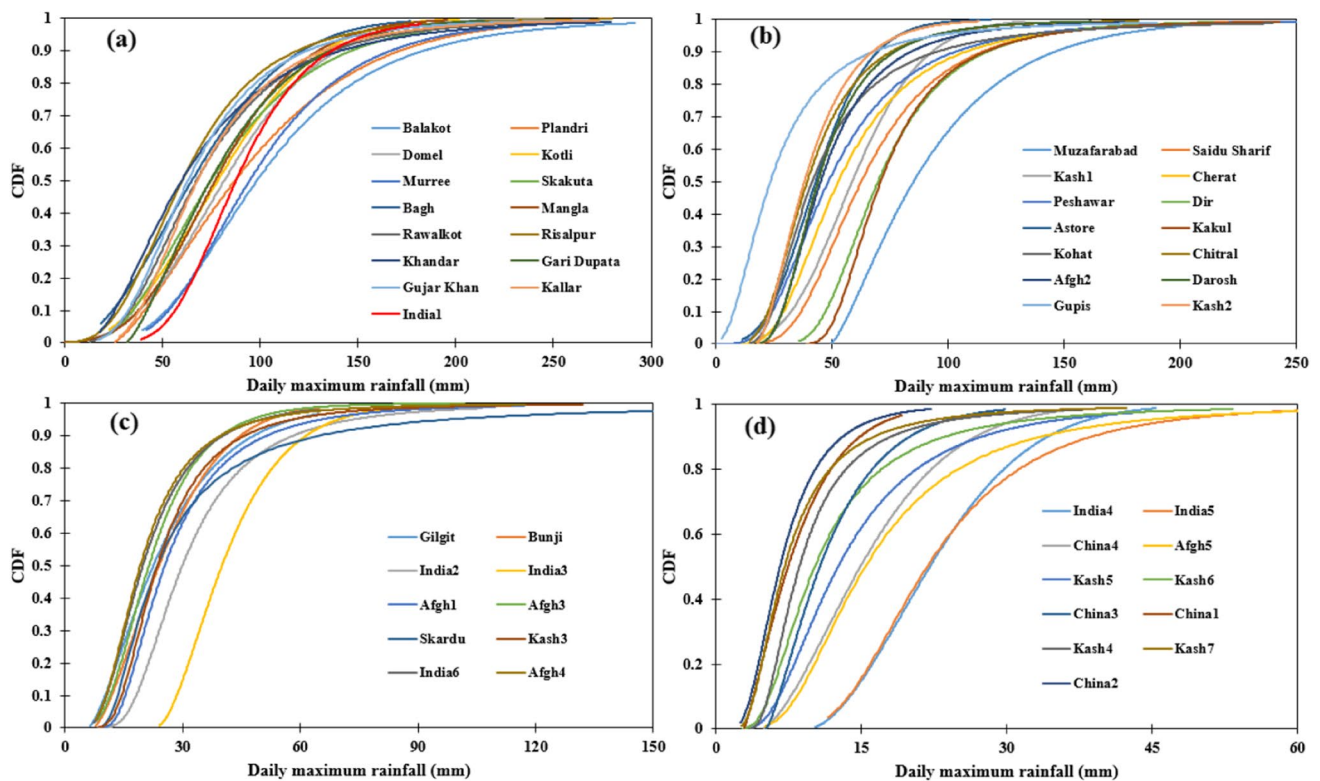


Fig. 5 Cumulative distribution functions at rainfall stations of SRIR. **a** Daily maximum rainfall ≥ 250 mm. **b** Daily maximum rainfall 150–250 mm. **c** Daily maximum rainfall 100–150 mm. **d** Daily maximum rainfall ≤ 60 mm

100 mm (Fig. 5c) while cumulative probability for eleven stations is about 1 with daily maximum rainfall less than equal to 60 mm (Fig. 5d).

3.2 Expected rainfall against various return periods

Daily maximum rainfall against the return periods of 5, 10, 15, 20, 25, 50, and 100 years is estimated using the best-fitted probability model at each of the rainfall station of the study area and the rainfall estimates are shown in Fig. 6. For the 5-year return period, daily maximum rainfall of less than 50 mm is estimated at 21 stations and 50–100 mm is estimated at 15 stations while more than 100 mm is estimated at 14 stations of the study area. For the 20-year return period, daily maximum rainfall of less than 50 mm is estimated at 12 stations, 50–100 mm is estimated at 12 stations, and 100–150 mm is estimated at 11 stations while more than 150 mm is estimated at 15 stations of the study area.

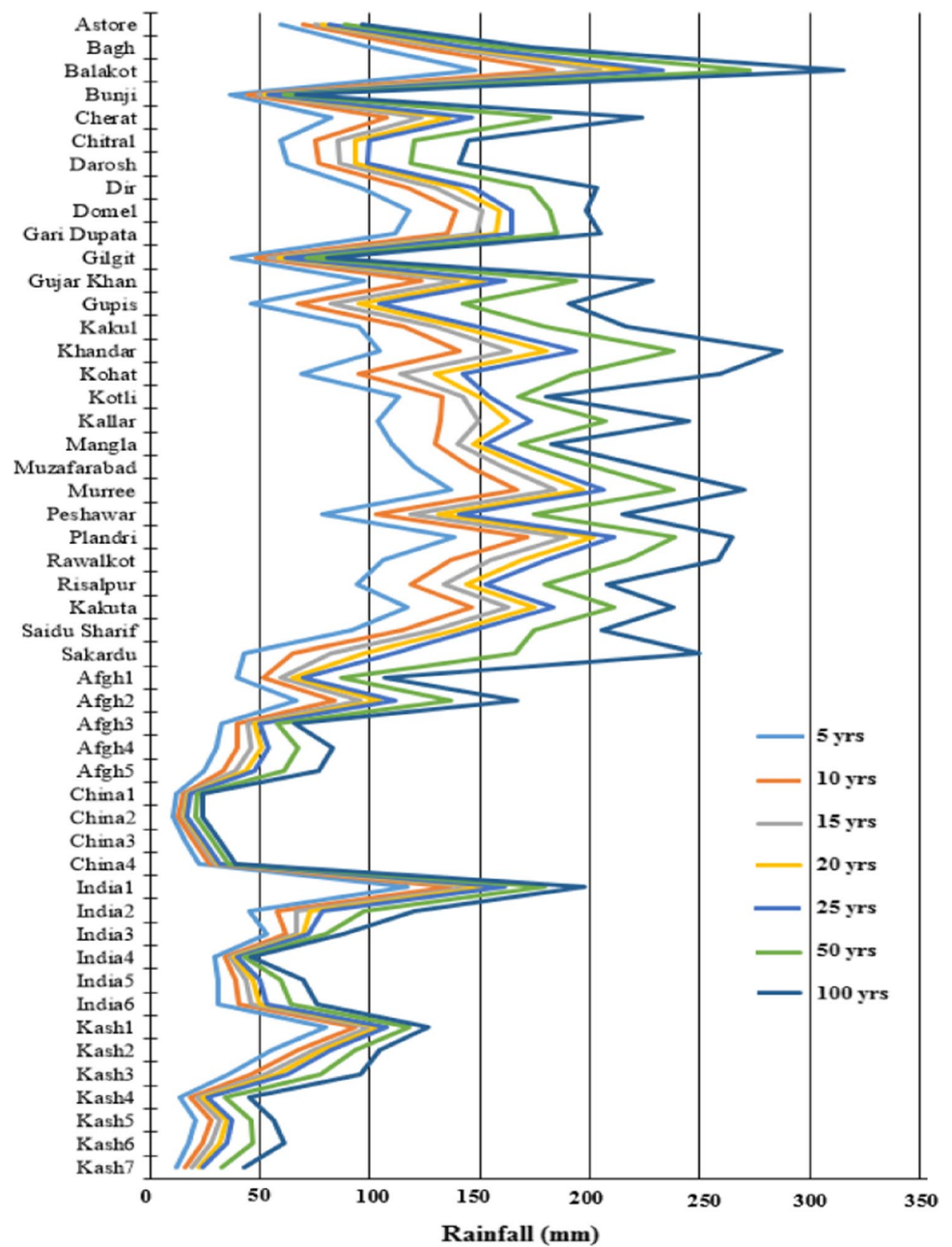
The spatial distributions of estimated rainfall against return periods of 5, 10, 15, 20, 25, 50, and 100 years in source region of Indus River are shown in Fig. 7. Daily values of annual maximum rainfall ranged from 12 mm in the

eastern part of SRIR to 186 mm in the middle of source region of Indus River against return period of 10 years (Fig. 7b). Annual maximum rainfall ranged from 17 to 233 mm and 20 to 272 mm against return period of 25 and 50 years respectively (Fig. 7e and f). Amount of expected rainfall is increasing from the eastern side toward the middle and from the western side toward the middle of SRIR. For the 100-year return period, daily values of annual maximum rainfall ranged from 23 to 317 mm (Fig. 7g).

4 Conclusions

Source region of Indus River is the area providing water to Indus River and its tributaries. Daily maximum rainfall observations from 1961 to 2015 of 50 rainfall stations of SRIR were used in this study to find the best-fitted probability model out of nine distributions used. Three goodness-of-fit tests were employed to find the most suited probability model. Each probability model is given the value of 1 to 9 based on the results of the goodness-of-fit tests. The values of these goodness-of-fit tests were added and best-fitted

Fig. 6 Rainfall estimates against different return periods using best-fitted probability distribution at rainfall stations in SRIR



probability distribution was selected based on the minimum sum of values. Daily maximum rainfall estimated using the non-best-fitted probability distributions could give high or low values as compared to the actual ones which will have adverse effects on the safety of hydrologic structures.

Generalized extreme value probability model was selected as best-fitted distribution on twenty-seven stations of the source region of Indus River (SRIR), followed by log–Pearson type 3 distribution on seven stations, Gamma

distribution on six stations, Weibull 3P distribution on six stations, Weibull distribution on two stations, log–normal distribution on one station, and extreme value type 1 distribution on one station. It is clear from the findings of this study that generalized extreme value distribution fitted best to more than 50% stations of SRIR.

The more applied result of this study was the predicted rainfall against different return periods for all the rainfall stations of the study area. The best-fitted probability

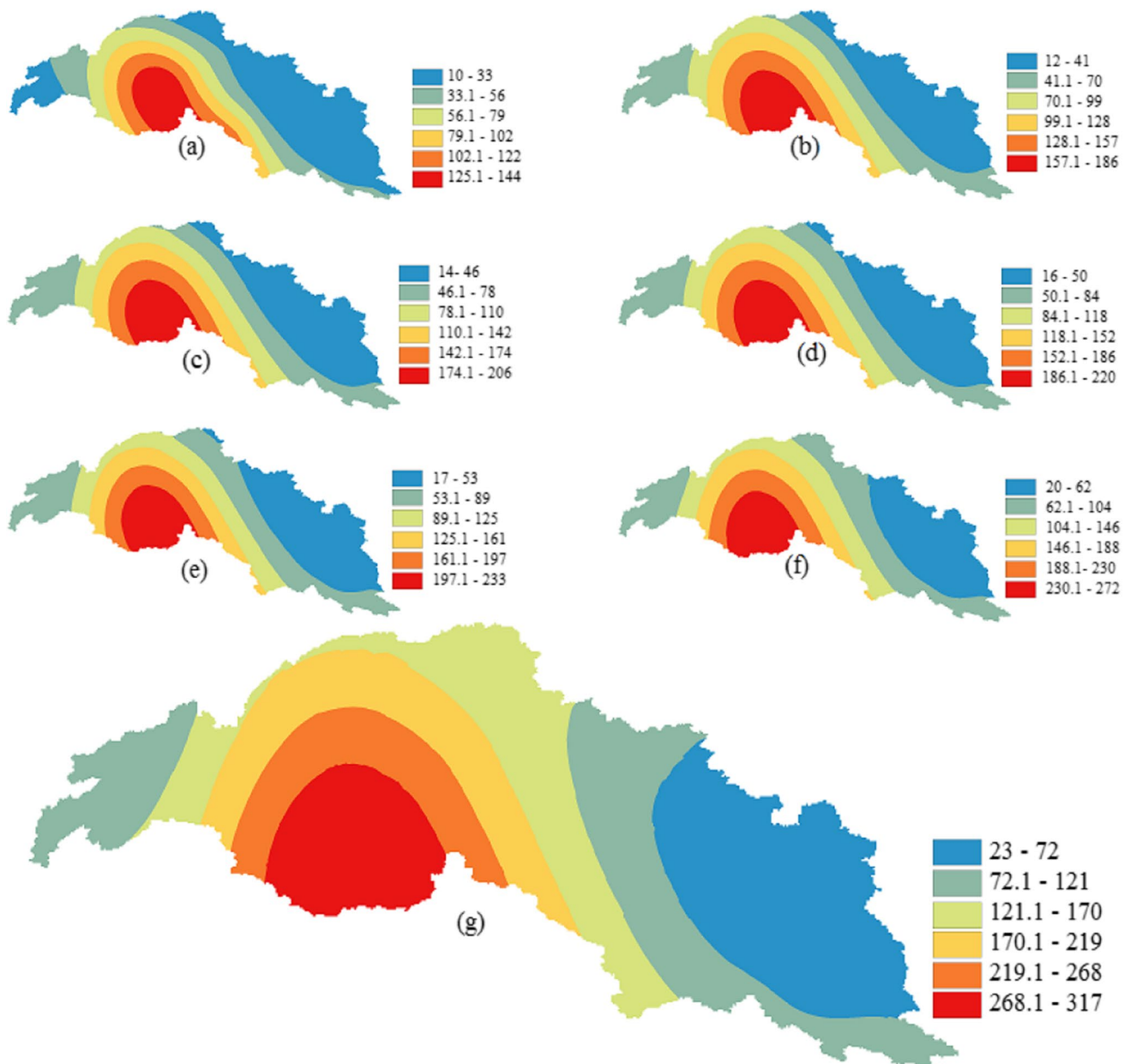


Fig. 7 Spatial distribution of rainfall against return period of 5 years (a), 10 years (b), 15 years (c), 20 years (d), 25 years (e), 50 years (f), and 100 years (g)

distributions, e.g., GEV, LP3, Gamma, W3P, W, LN, and EV1, were used for the estimation of rainfall against return periods of 5, 10, 15, 20, 25, 50, and 100 years. The spatial distribution of estimated daily values of annual maximum rainfall shows that the middle part of SRIR has the highest values, i.e., 186 mm for the 10-year return period and 317 mm for the 100-year return period. The eastern side of the study area has the lowest values of 24-h annual maximum rainfall, i.e., 12 mm and 23 mm for 10- and 100-year return period respectively.

A policymaker can use 25 or 50-year rainfall estimates for making policies or risk analysis of 25- or 50-year plan to minimize the threat and reparations from extreme events of rainfall and flooding.

Author contribution All authors are involved in the intellectual part of this paper. Muhammad Rizwan, Lubna Anjum, and Muhammad Awais designed the research. Muhammad Rizwan conducted the research and wrote the manuscript. Junaid Nawaz Chauhdary helped in the data arrangement while Muhammad Yamin helped in data analysis. Qaisar,

Ansir, and Irfan helped revise the manuscript and also provided many suggestions. All the authors have read and approved the final manuscript.

Data availability Data are available from the authors upon request.

Code availability R (language) codes are used in this research for data handling which can be provide upon request.

Declarations

Ethics approval The authors confirm that we have fully complied with ethical standards. No participation of human or animal involvement.

Consent to participate The authors declare that there is no human or animal participant in the study.

Consent for publication The authors give their consent to the publication of all details of the manuscript including texts, figures, and tables.

Conflict of interest The authors declare no competing interests.

References

- Adegboye O, Ipinoyomi R (1995) Statistical tables for class work and examination. Tertiary publications Nigeria Limited, Ilorin, Nigeria
- Adnan S, Khan AH (2009) Effective rainfall for irrigated agriculture plains of Pakistan. *Pakistan Journal of Meteorology* 6:61–72
- Agarwal M, Katiyar V, Babu R (1988) Probability analysis of annual maximum daily rainfall of UP, Himalaya. *Indian J Soil Conserv* 16:35–42
- Amin M, Rizwan M, Alazba A (2016) A best-fit probability distribution for the estimation of rainfall in northern regions of Pakistan. *Open Life Sciences* 11:432–440. <https://doi.org/10.1515/biol-2016-0057>
- Anil K (2000) Prediction of annual maximum daily rainfall of Ranchauri (Tehri Garhwal) based on probability analysis. *Indian J Soil Conserv* 28:178–180
- Bajracharya SR, Shrestha BR (2011) The status of glaciers in the Hindu Kush-Himalayan region. International Centre for Integrated Mountain Development (ICIMOD)
- Barkotulla M, Rahman M, Rahman M (2009) Characterization and frequency analysis of consecutive days maximum rainfall at Boalia, Rajshahi and Bangladesh. *J Dev Agric Econ* 1:121–126
- Bhakar S, Bansal AK, Chhajed N, Purohit R (2006) Frequency analysis of consecutive days maximum rainfall at Banswara, Rajasthan, India. *ARPN J Eng Appl Sci* 1:64–67
- Bhakar S, Iqbal M, Devanda M, Chhajed N, Bansal AK (2008) Probability analysis of rainfall at Kota. *Indian J Agric Res* 42:201–206
- Chowdhury JU, Stedinger JR, Lu LH (1991) Goodness-of-fit tests for regional generalized extreme value flood distributions. *Water Resour Res* 27:1765–1776
- Conover WJ (1998) Practical nonparametric statistics. John Wiley & Sons
- Eslamian SS, Feizi H (2007) Maximum monthly rainfall analysis using L-moments for an arid region in Isfahan province. *Iran J Appl Meteorol Climatol* 46:494–503
- George C, Kolappadan C (2002) Probability analysis for prediction of annual maximum daily rainfall of Periyar basin in Kerala. *Indian J of Soil Cons* 30:273–276
- Govt.-of-Pakistan (2008) Federal Bureau of Statistics, Ministry of Finance. Islamabad
- Hasson S, Böhner J, Lucarini V (2017) Prevailing climatic trends and runoff response from Hindukush–Karakoram–Himalaya, upper Indus Basin. *Earth System Dynamics* 8:337–355
- Khattak MS, Babel M, Sharif M (2011) Hydro-meteorological trends in the upper Indus River basin in Pakistan. *Climate Res* 46:103–119
- Khudri MM, Sadia F (2013) Determination of the best fit probability distribution for annual extreme precipitation in Bangladesh. *Eur J Sci Res* 103:391–404
- 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
- Leavenworth RS, Grant EL (2000) Statistical quality control. Tata McGraw-Hill Education
- Lee C (2005) Application of rainfall frequency analysis on studying rainfall distribution characteristics of Chia-Nan plain area in Southern Taiwan. *Crop Environ Bioinf* 2:31–38
- Lutz AF, Immerzeel WW, Kraaijenbrink PD, Shrestha AB, Bierkens MF (2016) Climate change impacts on the upper Indus hydrology: sources, shifts and extremes. *PLoS ONE* 11:e0165630
- Mandal S, Choudhury B (2015) Estimation and prediction of maximum daily rainfall at Sagar Island using best fit probability models. *Theoret Appl Climatol* 121:87–97
- Ogunlela A (2001) Stochastic analysis of rainfall events in Ilorin, Nigeria. *J Agric Res Dev* 1:39–50
- Olofintoye O, Sule B, Salami A (2009) Best-fit probability distribution model for peak daily rainfall of selected Cities in Nigeria. *New York Science Journal* 2:1–12
- Şen Z, Eljadid AG (1999) Rainfall distribution function for Libya and rainfall prediction. *Hydrol Sci J* 44:665–680
- Singh R (2001) Probability analysis for prediction of annual maximum rainfall of Eastern Himalaya (Sikkim mid hills). *Ind J Soil Conserv* 29:263–265
- Tao D, Nguyen V, Bourque A (2002) On selection of probability distributions for representing extreme precipitations in Southern Quebec. Annual conference of the Canadian society for civil engineering, 1–8
- Vivekanandan N (2012) Intercomparison of extreme value distributions for estimation of ADMR. *Int J Appl Eng Technol* 2:30–37
- Wolf AT, Natharius JA, Danielson JJ, Ward BS, Pender JK (1999) International river basins of the world. *Int J Water Resour Dev* 15:387–427

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.