ORIGINAL ARTICLE

CrossMark

Comparison of probability distributions in extreme value analysis of rainfall and temperature data

N. Vivekanandan[1]

Abstract

Estimation of rainfall and temperature for a desired return period is a prerequisite for planning, design and operation of various hydraulic structures and for evaluation of technical and engineering appraisal of large infrastructure projects. This can be computed through Extreme Value Analysis (EVA) by fitting probability distributions to the annual series of 1-day maximum rainfall, maximum and minimum temperature. This paper details the study on adoption of Extreme Value Type-1, Extreme Value Type-2, 2-parameter Log Normal and Log Pearson Type-3 (LP3) probability distributions in EVA of rainfall and temperature for Hissar. Based on the applicability, standard parameter estimation procedures such as method of moments, maximum likelihood method (MLM) and order statistics approach are used for determination of parameters of distributions. The adequacy on fitting of probability distributions used in EVA of rainfall and temperature is evaluated by goodness-of-fit (GoF) tests, viz. Anderson–Darling and Kolmogorov–Smirnov and diagnostic test using D-index. The GoF and diagnostic tests results suggest the LP3 (MLM) is better suited amongst four probability distributions adopted in EVA of rainfall and temperature for Hissar.

Keywords Anderson–Darling test · D-index · Kolmogorov–Smirnov test · Log Pearson Type-3 · Rainfall · Temperature

Introduction

Technical and engineering appraisal of large infrastructure projects such as nuclear, hydro and thermal power plants, dams, bridges and flood control measures needs to be carried out during the planning and formulation stages. In hydrological context, it is well recognized fact that, whatsoever extreme be the design loading, more severe conditions are likely to occur in nature. For assessing such phenomena, Extreme Value Analysis (EVA) of rainfall and temperature from the geographical region of the project is one of the basic requirements. Depending on the size and the design life of the structure, the estimated extreme events corresponding to particular return period is used. This can be computed by fitting of probability distribution to the series of Annual 1-day Maximum Rainfall (AMR), Annual Maximum Temperature (AMAXT) and Annual Minimum Temperature (AMINT) observed at the India Meteorological

Department observatory located in the vicinity of the project site. In general, the selection of suitable distribution plays a key role in estimating the extreme values.

Out of a number of available probability distributions, Extreme Value Type-1 (EV1), Extreme Value Type-2 (EV2), 2-parameter Log Normal (LN2) and Log Pearson Type-3 (LP3) are extensively used for Extreme Value Analysis (EVA). Based on the applicability, standard parameter estimation procedures viz., Methods of Moments (MoM) and Maximum Likelihood Method (MLM) are generally used for determination of parameters (Bobee and Ashkar 1991). In addition to MoM and MLM, Atomic Energy Regulatory Board (AERB 2008) guidelines described that the Order Statistics Approach (OSA) can also be considered for determination of parameters of EV1 and EV2 distributions. AERB guidelines also described that the OSA estimates are popular owing to less bias and minimum variance though number of methods are available for parameter estimation. AERB guidelines suggested the 1000-year return period Mean + SE (where Mean denotes the estimated value (x_T) of rainfall (or) temperature and SE the standard error), and value will be generally considered for arriving at a design value of rainfall and maximum temperature, whereas mean–SE value

✉ N. Vivekanandan
anandaan@rediffmail.com

1 Central Water and Power Research Station, Pune, Maharashtra, India

for minimum temperature for design purposes. In the recent past, numbers of studies have been carried out by researchers adopting probability distributions for EVA of rainfall and temperature.

Hughes et al. (2007) carried out statistical analysis using generalized extreme value (GEV) distribution and time series models for minimum and maximum temperatures in the Antarctic Peninsula. Varathan et al. (2010) found that the Gumbel (i.e. EV1) distribution is the best fitting distribution to analyse the annual maximum of rainfall of Colombo district. AlHassoun (2011) carried out a study on developing empirical formulae to estimate rainfall intensity in Riyadh region using EV1, LN2 and LP3. He concluded that the LP3 distribution gives better accuracy amongst three distributions studied in estimation of rainfall intensity. Baratti et al. (2012) carried out flood frequency analysis on seasonal and annual time scales for the Blue Nile River adopting EV1 distribution. Hasan et al. (2013) applied GEV distribution for modelling the AMAXT observed at 22 meteorological stations in Malaysia. Esteves (2013) applied EV1 distribution to estimate the extreme rainfall depths at different rain-gauge stations in southeast UK. Rasel and Hossain (2015) applied EV1 distribution for development of intensity duration frequency curves for seven divisions in Bangladesh. Vivekanandan (2015) applied EV1 and EV2 distributions for modelling the series of AMR, AMAXT, AMINT and annual hourly maximum wind speed (HMWS) for Kanyakumari region. He found that the EV1 (OSA) is a better-suited probability distribution for modelling the series of AMAXT, AMAXT and AMINT, whereas EV2 (OSA) for modelling of HMWS for Kanyakumari. Vivekanandan (2016) carried out statistical analysis of rainfall data using EV1 distribution for estimation of peak flood discharge for ungauged catchments of Bhul and Manjuhi Khad, Himachal Pradesh.

Generally, when different probability distributions are used for EVA, a common problem that arises is how to determine which model fits the best for a given set of data. This can be evaluated by goodness-of-fit (GoF) and diagnostic tests. GoF tests, viz. Anderson–Darling (A^2) and Kolmogorov–Smirnov (KS), are applied for checking the adequacy of fitting of probability distributions to the rainfall and temperature data (Zhang 2002). However, the results of GoF tests offered diverging inferences which lead to adopt D-index to aid the selection of suitable distribution for EVA of rainfall and temperature. Thus, there exist research efforts in assessing the extreme values of rainfall and temperature (maximum and minimum) and for aiding design parameter of interest and present work is an effort in this direction. This paper details the procedures involved in assessing the suitable probability distribution for estimation of rainfall and temperature though GoF and diagnostic tests with illustrative example.

Methodology

The objective of the study is to assess the adequacy of probability density function (PDF) for EVA of rainfall and temperature. In this context, various steps followed for data processing, validation and analysis include: (i) prepare the AMR data series from the daily rainfall data; (ii) prepare the AMAXT and AMINT data series from the hourly temperature data; (iii) select the PDFs for EVA (say, EV1, EV2, LN2 and LP3); (iv) select parameter estimation methods (say, MoM, MLM and OSA) wherever applicable; (v) select quantitative based GoF and diagnostic tests; and (vi) conduct EVA and analyse the results obtained thereof. The PDF and quantile estimator (x_T) of the distributions are presented in Table 1.

In Table 1, α, β and γ denote the location, scale and shape parameters of the distributions, respectively. For EV1 and EV2 distributions, the reduced variate (Y_T) for a given return period (T) is defined by $Y_T = -\ln(-\ln(1 - (1/T)))$, while in the mathematical representation of LN2 and LP3, K_P denotes the frequency factor corresponding to the probability of exceedance. The coefficient of skewness (CS) is CS = 0.0 for LN2, whereas CS is based on the log-transformed series of the observed data for LP3 (Rao and Hameed 2000). For the data series with AMINT, the value of Y_T will be read as $Y_T = -\ln(-\ln(1/T))$ and K_P the frequency factor corresponding to the probability of non-exceedance for LN2 and LP3 distributions.

Table 1 PDF and quantile estimator of probability distributions

Distribution	PDF	Quantile estimator
EV1	$f(x;\alpha,\beta) = \dfrac{e^{-(x-\alpha)/\beta}e^{-e^{-(x-\alpha)/\beta}}}{\beta}, \beta > 0$	$x_T = \alpha + Y_T\beta$
EV2	$f(x;\beta,\gamma) = \dfrac{\gamma}{\beta}\left(\dfrac{\beta}{x}\right)^{\gamma+1} e^{-\left(\frac{x}{\beta}\right)_{-\gamma}}, \beta > 0$	$x_T = \beta e^{(Y_T/\gamma)}$
LN2	$f(x;\alpha,\beta) = \dfrac{1}{\beta x\sqrt{2\pi}} \exp\left(-\dfrac{(\ln(x)-\alpha)^2}{2\beta^2}\right), -\infty < x < \infty, \beta > 0$	$x_T = e^{\alpha+K_P\beta}$
LP3	$f(x;\alpha,\beta,\gamma) = \dfrac{1}{\beta x\Gamma\gamma}\left(\dfrac{\ln(x)-\alpha}{\beta}\right)^{\gamma-1} e^{-\left(\frac{\ln(x)-\alpha}{\beta}\right)}, \beta, \gamma > 0$	$x_T = \text{Exp}((\alpha + \beta\gamma) + K_P\beta\sqrt{\alpha})$

Goodness-of-fit tests

Generally, A^2 test is applied for checking the adequacy of fitting of EV1 and EV2 distributions. The procedures involved in application of A^2 test for LN2 and LP3 are more complex though the utility of the test statistic is extended for checking the quantitative assessment. In view of the above, KS test is widely applied for the purpose of quantitative assessment. Theoretical descriptions of GoF tests statistic are as follows: A^2 test statistic is defined by:

$$A^2 = (-N) - (1/N) \sum_{i=1}^{N} \left\{ (2i-1)\ \ln(Z_i) + (2N+1-2i)\ \ln(1-Z_i) \right\}$$

(1)

Here, $Z_i = F(x_i)$ for $i = 1,2,3,\ldots,N$ with $x_1 < x_2 < \cdots < x_N$, $F(x_i)$ is the cumulative distribution function (CDF) of ith sample (x_i) and N is the sample size.

KS test statistic is defined by:

$$\text{KS} = \underset{i=1}{\overset{N}{\text{Max}}}\ (F_e(x_i) - F_D(x_i))$$

(2)

Here, $F_e(x_i)$ is the empirical CDF of x_i and $F_D(x_i)$ is the derived CDF of x_i by PDFs. In this paper, Weibull plotting position formula is used for computation of empirical CDF. The theoretical values of A^2 and KS tests statistic for different sample size (N) at 5% significance level are available in the technical note on 'goodness-of-fit tests for statistical distributions' by Charles Annis (2009).

Test criteria If the computed value of GoF tests statistic given by the distribution is less than that of theoretical value at the desired significance level then the distribution is found to be suitable for EVA of rainfall and temperature at that level.

Diagnostic test

Sometimes, the GoF test results would not offer a conclusive inference, thus posing a problem for the user in selecting a suitable PDF for their application. In such cases, a diagnostic test in adoption to GoF is applied for making inference. The selection of most suitable probability distribution for estimation of rainfall and temperature (maximum and minimum) is performed through D-index test (USWRC 1981), which is defined as below:

$$\text{D-index} = (1/\bar{x}) \sum_{i=1}^{6} |x_i - x_i^*|$$

(3)

Here, $\bar{x}$ is the average value of the observed data, whereas x_i ($i = 1$ to 6) and x_i^* are the six highest observed and corresponding estimated values by different PDFs. The distribution has the least D-index is considered as better-suited distribution for EVA of rainfall and temperature.

Application

In this paper, a study on EVA of rainfall and temperature adopting EV1, EV2, LN2 and LP3 probability distributions was carried out. MoM, MLM and OSA were used for determination of parameters of EV1 and EV2 distributions, whereas MoM and MLM for LN2 and LP3 distributions. The series of AMR derived from the daily rainfall data for the period 1969 to 2011 and the series of AMAXT and AMINT extracted from the hourly temperature data for the period 1977 to 2007 observed at Hissar observatory was used for EVA. From the scrutiny of the rainfall data, it was observed that the data of the year 2002 is not available. Similarly, from the scrutiny of the hourly temperature data, it was observed that the data for the period of 6 years (1978, 1979, 1981, 1983, 1987 and 1989) are missing. So, the missing data for the years are imputed by the series maximum value, i.e. 256.5 mm for AMR, 48.4 °C for AMAXT and 2.7 °C for AMINT as per AERB guidelines, and the entire data set is used for EVA. Table 2 gives the descriptive statistics of AMR, AMAXT and AMINT for Hissar.

Results and discussion

Based on the parameter estimation procedures of EV1, EV2, LN2 and LP3 distributions (Rao and Hameed 2000), computer codes were developed in FORTRAN language and used for EVA of rainfall and temperature. These programs compute the parameters of the distributions, estimated values (x_T) of rainfall and temperature with standard error (SE) for different return periods, GoF tests statistic and D-index values.

EVA of rainfall

The estimated 1-day maximum rainfall (ER) with SE obtained from EV1, EV2, LN2 and LP3 distributions are presented in Tables 3 and 4. The ER values obtained from EV1, LN2 and LP3 are used to develop the plots

Table 2 Descriptive statistics of AMR, AMAXT and AMINT for Hissar

Data series	Descriptive statistics			
	Average	SD	CS	CK
AMR	93.8 mm	56.4 mm	1.631	2.320
AMAXT	45.7 °C	1.2 °C	0.708	− 0.503
AMINT	5.3 °C	1.1 °C	− 0.465	− 0.511

SD standard deviation, *CS* coefficient of skewness, *CK* coefficient of kurtosis

Table 3 Estimates of 1-day maximum rainfall with SE adopting EV1 and EV2 distributions

Return period (year)	1-day maximum rainfall (mm) with SE (mm)											
	EV1						EV2					
	MoM		MLM		OSA		MoM		MLM		OSA	
	ER	SE	ER	SE	ER	SE	ER	SE	ER	SE	ER	SE
2	84.6	7.8	84.0	6.4	85.3	7.0	74.2	6.8	77.6	7.1	74.7	6.9
5	134.4	13.3	124.7	10.8	129.4	11.2	119.4	18.0	139.0	21.0	129.9	19.6
10	167.5	17.9	151.6	14.6	158.7	14.7	163.6	33.2	204.6	41.6	187.4	38.1
20	199.1	22.7	177.4	18.5	186.7	18.4	221.3	57.4	296.3	76.8	266.3	69.1
50	240.1	29.0	210.9	23.6	223.0	23.2	327.2	110.7	478.5	161.9	419.6	142.0
100	270.9	33.7	235.9	27.5	250.2	26.9	438.7	176.2	685.3	275.3	590.0	237.0
200	301.5	38.5	260.9	31.4	277.3	30.6	587.4	275.2	980.2	459.1	828.5	388.0
500	341.8	45.0	293.8	36.6	313.0	35.6	863.6	485.2	1571.7	883.2	1296.6	728.5
1000	372.4	49.7	318.7	40.6	340.0	39.3	1155.5	735.8	2245.7	1430.1	1818.9	1158.3

Table 4 Estimates of 1-day maximum rainfall with SE adopting LN2 and LP3 distributions

Return period (year)	1-day maximum rainfall (mm) with SE (mm)							
	LN2				LP3			
	MoM		MLM		MoM		MLM	
	ER	SE	ER	SE	ER	SE	ER	SE
2	80.4	6.8	81.0	6.6	79.3	8.9	80.2	8.7
5	128.4	12.6	126.8	12.0	126.7	17.5	127.4	17.0
10	163.9	18.7	160.2	17.6	163.2	26.4	162.8	25.7
20	200.5	26.1	194.4	24.2	202.3	37.5	199.9	36.0
50	251.6	37.8	241.6	34.6	259.0	55.7	252.3	52.6
100	292.8	47.9	279.3	43.7	306.5	72.3	295.0	67.6
200	336.3	59.5	318.9	54.0	358.5	91.8	340.8	84.7
500	397.8	76.8	374.5	69.3	436.0	122.9	406.3	111.2
1000	447.5	91.7	419.2	82.3	498.8	149.7	460.0	134.0

and presented in Fig. 1. For Hissar, it is noted that the ER (Table 3) obtained from EV2 (using MoM, MLM and OSA) are higher than the corresponding values of EV1 (MoM, MLM and OSA), LN2 (MoM and MLM) and LP3 (MoM and MLM); and thus the plots of EV2 were not clubbed with the plots of EV1, LN2 and LP3. Figure 2 presents the plots of observed and estimated ER by EV2 (MoM, MLM and OSA) for Hissar.

From Fig. 1, it can be seen that the values of ER obtained from LP3 (MoM) is relatively higher than the corresponding values of EV1 (MoM, MLM and OSA), LN2 (MoM and MLM) and LP3 (MLM) for return periods vary from 50 years to 1000 years. Also, from Fig. 2, it can be seen that the values of ER by EV2 (MLM) are comparatively higher than the corresponding values of EV2 (MoM) and EV2 (OSA). From Fig. 1, it can also be seen that the fitted curves of ER values by EV1 and LN2 are in the form of linear trend. Similarly, from Fig. 2, it can also be seen that the fitted curves

are in the form of exponential trend while EV2 is adopted for EVA of rainfall.

EVA of temperature

The estimated values of maximum temperature (ET(Max)) with SE computed from EV1, EV2, LN2 and LP3 are presented in Tables 5 and 6. Similarly, the estimated minimum temperature (ET(Min)) with SE using EV1, LN2 and LP3 are presented in Table 7.

As there was no existence of OSA for LN2 and LP3, the EVA results of LN2 (OSA) and LP3 (OSA) are not presented in Tables 6 and 7. Similarly, from the EVA of AMINT data, it was observed that the EV2 is not found to be feasible for fitting, and therefore EVA results of EV2 are not presented in Table 7. The values of ET(Max) obtained from EV1, EV2, LN2 and LP3 are used to develop the plots and presented in Fig. 3.

Fig. 1 Plots of observed and estimated values of rainfall by EV1, LN2 and LP3 distributions
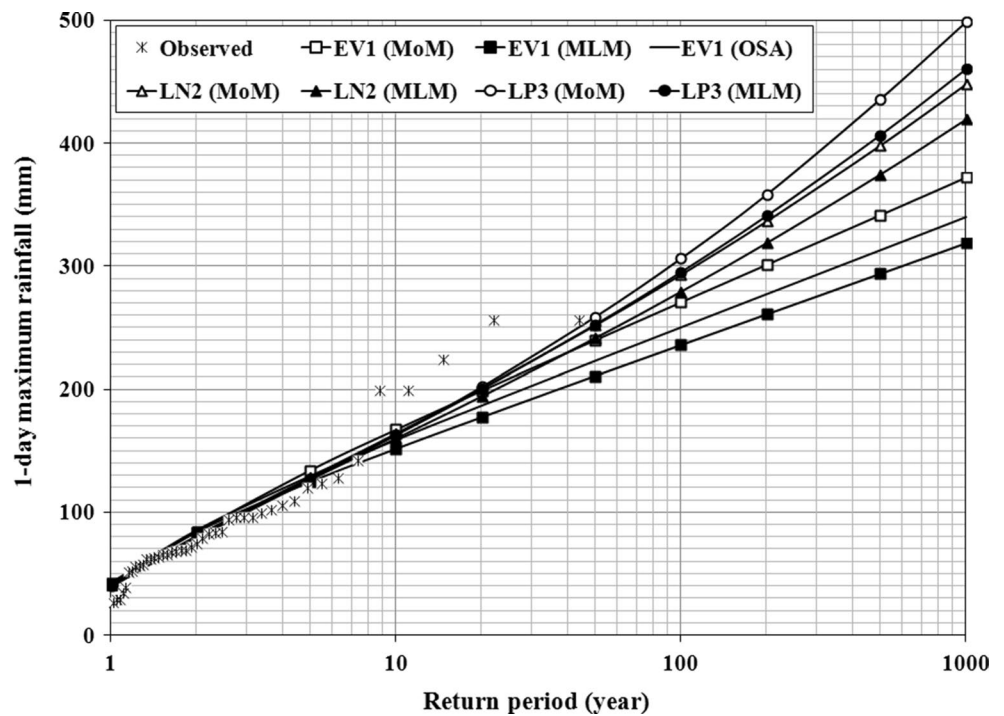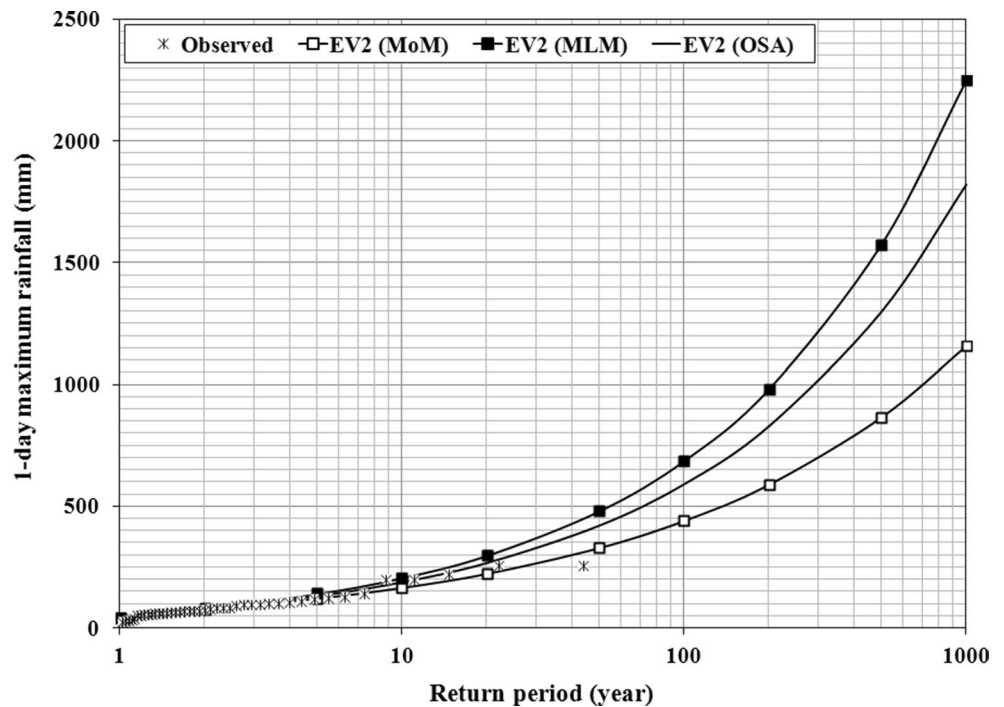

Fig. 2 Plots of observed and estimated values of rainfall by EV2 distribution

From Table 6, it is noted that there is no appreciable difference between the ET(Max) values given by LN2 (MoM and MLM) and LP3 (MoM and MLM). From Fig. 3, it can be seen that the values of ET(Max) obtained from EV2 (OSA) is relatively higher than the corresponding values of EV1 (MoM, MLM and OSA), EV2 (MoM and MLM), LN2 (MoM and MLM) and LP3 (MoM and MLM). Also,

from Fig. 3, it can be seen that the fitted curves by using the values of ET(Max) obtained from EV1, EV2, LN2 and LP3 are in the form of linear trend.

From Table 7, it is noticed that the values of ET(Min) by EV1 (MoM, MLM and OSA) vary between − 3.6 and 5.0 °C, whereas the corresponding values given by LN2 (MoM and MLM) and LP3 (MoM and MLM) vary

Table 5 Estimates of maximum temperature with SE adopting EV1 and EV2 distributions

| Return period (year)	Maximum temperature (°C) with SE (°C)											
	EV1						EV2					
	MoM		MLM		OSA		MoM		MLM		OSA	
	ET(Max)	SE	ET(Max)	SE	ET(Max)	SE	ET(Max)	SE	ET(Max)	SE	ET(Max)	SE
2	46.0	0.2	46.0	0.2	46.0	0.3	46.5	0.2	45.9	0.3	46.0	0.3
5	47.3	0.5	47.4	0.4	47.5	0.5	47.7	0.5	47.3	0.4	47.5	0.5
10	48.2	0.6	48.3	0.6	48.5	0.6	48.5	0.7	48.2	0.6	48.5	0.7
20	49.1	0.7	49.2	0.7	49.4	0.8	49.4	0.8	49.1	0.8	49.5	0.8
50	50.2	0.9	50.3	1.0	50.7	0.9	50.5	1.0	50.3	1.0	50.9	1.0
100	51.0	1.1	51.2	1.1	51.6	1.1	51.3	1.2	51.2	1.2	51.9	1.2
200	51.9	1.2	52.1	1.2	52.5	1.2	52.1	1.4	52.1	1.4	53.0	1.4
500	53.0	1.4	53.2	1.5	53.7	1.4	53.3	1.6	53.4	1.6	54.4	1.7
1000	53.8	1.6	54.1	1.6	54.6	1.6	54.1	1.9	54.3	1.9	55.5	1.9

Table 6 Estimates of maximum temperature with SE adopting LN2 and LP3 distributions

Return period (year)	Maximum temperature (°C) with standard error (°C)							
	LN2				LP3			
	MoM		MLM		MoM		MLM	
	ET(Max)	SE	ET(Max)	SE	ET(Max)	SE	ET(Max)	SE
2	46.2	0.3	46.2	0.3	46.1	1.6	46.1	1.5
5	47.5	0.4	47.5	0.3	47.5	1.9	47.4	2.0
10	48.2	0.4	48.2	0.4	48.3	2.3	48.2	2.3
20	48.8	0.4	48.8	0.4	48.9	2.7	48.9	2.6
50	49.5	0.5	49.4	0.5	49.7	3.2	49.7	3.1
100	49.9	0.6	49.8	0.6	50.2	3.6	50.2	3.5
200	50.3	0.7	50.2	0.7	50.7	3.9	50.8	3.8
500	50.8	0.7	50.7	0.7	51.4	4.3	51.4	4.3
1000	51.2	0.7	51.1	0.7	51.8	4.7	51.9	4.7

Table 7 Estimates of minimum temperature with SE adopting EV1, LN2 and LP3 distributions

Return period (year)	Minimum temperature (°C) with SE (°C)													
	EV1						LN2				LP3			
	MoM		MLM		OSA		MoM		MLM		MoM		MLM	
	ET(Min)	SE	ET(Min)	SE	ET(Min)	SE	ET(Min)	SE	ET(Min)	SE	ET(Min)	SE	ET(Min)	SE
2	3.7	0.2	5.0	0.2	5.0	0.2	4.6	0.2	4.6	0.3	4.4	0.4	4.4	0.4
5	2.5	0.4	3.7	0.4	3.7	0.4	3.6	0.2	3.5	0.3	3.4	0.3	3.4	0.4
10	1.6	0.5	2.8	0.6	2.9	0.5	3.1	0.2	3.0	0.2	3.1	0.4	3.0	0.4
20	0.8	0.7	1.9	0.7	2.1	0.7	2.8	0.2	2.7	0.3	2.8	0.4	2.7	0.4
50	− 0.2	0.9	0.8	0.9	1.1	0.9	2.5	0.2	2.3	0.2	2.6	0.4	2.5	0.5
100	− 1.0	1.1	0.0	1.1	0.3	1.0	2.3	0.2	2.1	0.2	2.4	0.4	2.3	0.4
200	− 1.8	1.2	− 0.9	1.2	− 0.5	1.1	2.2	0.3	2.0	0.3	2.3	0.4	2.2	0.4
500	− 2.9	1.3	− 2.0	1.4	− 1.5	1.3	2.0	0.3	1.8	0.3	2.2	0.4	2.1	0.5
1000	− 3.6	1.5	− 2.8	1.6	− 2.3	1.5	1.8	0.2	1.7	0.3	2.1	0.4	2.0	0.5

Fig. 3 Plots of observed and estimated values of maximum temperature by EV1, EV2, LN2 and LP3 distributions
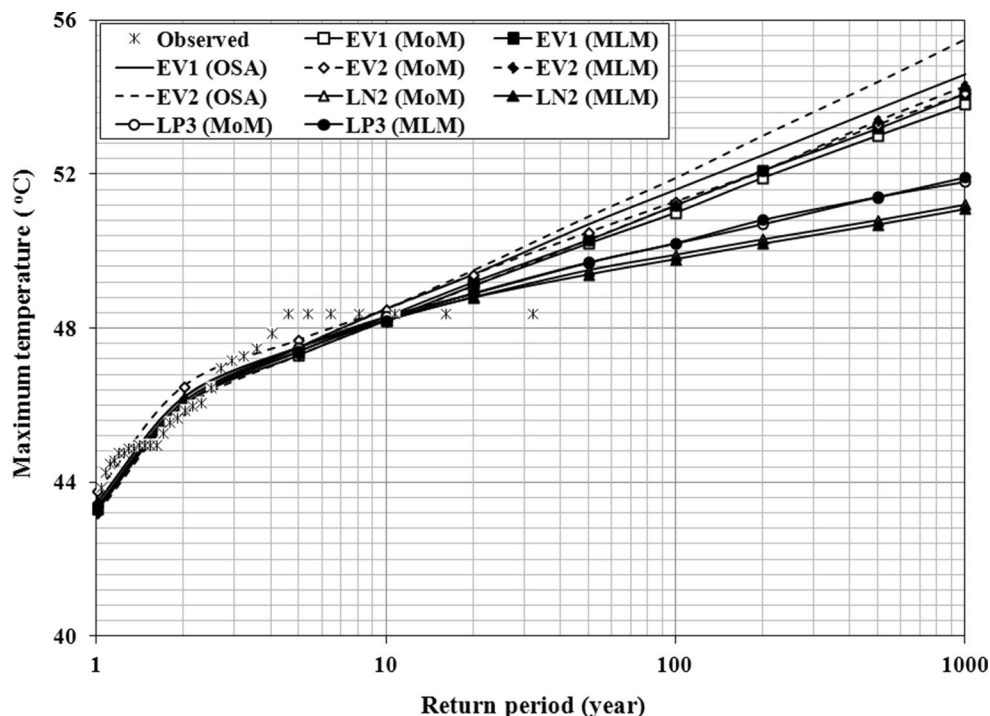
between 1.7 and 4.6 °C. By considering the amount of variation in ET(Min) values in lower tail region, the plots of observed and estimated values of minimum temperature by EV1 are presented in Fig. 4, whereas the plots of LN2 and LP3 are presented in Fig. 5.

Analysis based on GoF tests

The adequacy of fitting four different PDFs for EVA of rainfall and temperature was performed by adopting GoF tests, viz. A^2 and KS, as described earlier. The GoF tests results for Hissar are presented in Table 8.

Fig. 4 Plots of observed and estimated values of minimum temperature by EV1 distribution
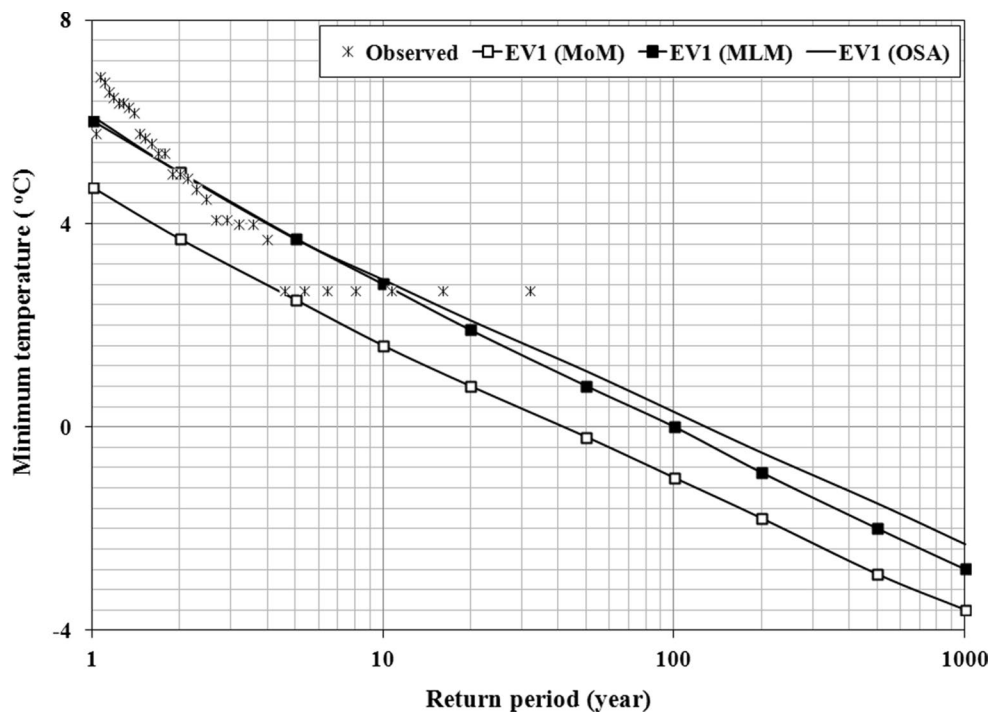
Fig. 5 Plots of observed and estimated values of minimum temperature by LN2 and LP3 distributions
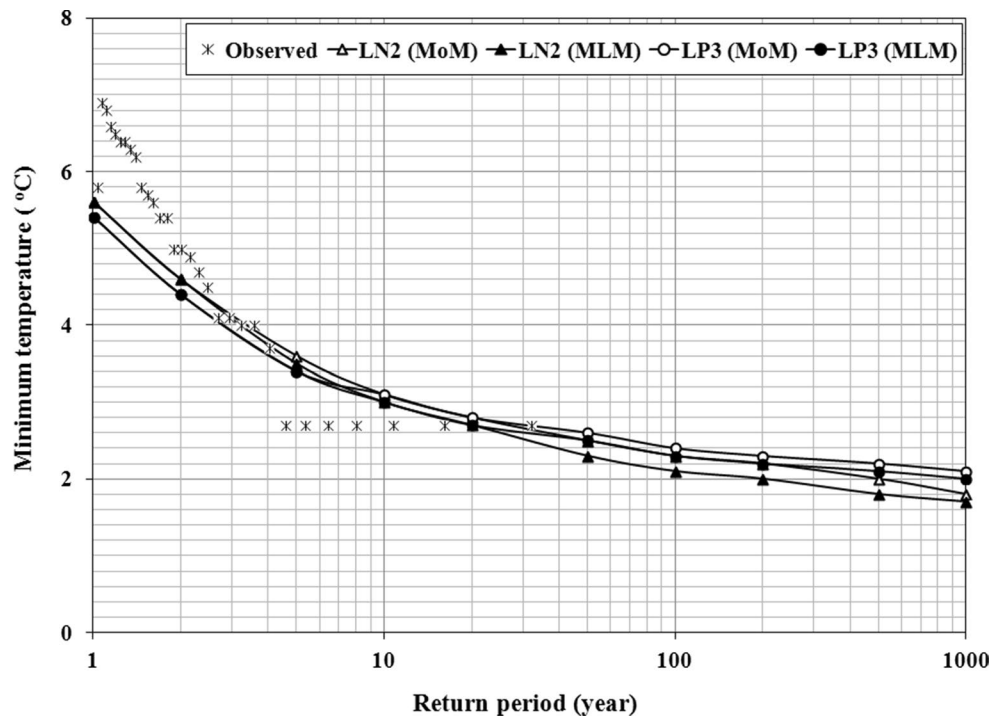

Table 8 Computed and theoretical values of GoF tests statistic adopting EV1, EV2, LN2 and LP3

GoF tests	Computed values of GoF tests statistic										Theoretical value at 5% level
	EV1			EV2			LN2		LP3		
	MoM	MLM	OSA	MoM	MLM	OSA	MoM	MLM	MoM	MLM	
AMR											
A^2	1.138	0.868	0.947	1.110	0.900	0.913	0.475	0.565	0.657	0.545	0.757
KS	0.105	0.079	0.070	0.106	0.104	0.122	0.083	0.088	0.079	0.097	0.203
AMAXT											
A^2	1.213	1.064	0.921	3.105	1.181	0.894	1.154	1.372	1.622	1.199	0.757
KS	0.123	0.118	0.104	0.011	0.128	0.101	0.165	0.182	0.162	0.174	0.238
AMINT											
A^2	2.740	2.698	3.365	NF	NF	NF	1.352	1.453	1.415	1.573	0.757
KS	0.188	0.518	0.482	NF	NF	NF	0.162	0.171	0.122	0.136	0.238

From the GoF tests results, the following observations were made:

(i)   A^2 test results confirmed the applicability of LN2 (MoM and MLM) and LP3 (MoM and MLM) for EVA of rainfall.

(ii) A^2 test results did not support the EV1 (MoM, MLM and OSA), EV2 (MoM, MLM and OSA), LN2 (MoM and MLM) and LP3 (MoM and MLM) for EVA of temperature.

(iii) KS test results suggested the EV1 (MoM, MLM and OSA), EV2 (MoM, MLM and OSA), LN2 (MoM and MLM) and LP3 (MoM and MLM) for EVA of rainfall and temperature of Hissar.

Analysis based on diagnostic test

A diagnostic test of D-index was used for assessing the adequate probability distribution for EVA of rainfall and temperature for Hissar. The D-index values of EV1, EV2, LN2 and LP3 distributions are computed and presented in Table 9.

From Table 9, it is noted that the D-index value computed from EV1 (MoM) is minimum when compared to the corresponding values of other probability distributions for the series of AMR. Also, from Table 9, it may be noted that the D-index values computed from LN2 (MoM) and LP3 (MLM) are minimum for the series AMAXT and AMINT, respectively.

Table 9 D-index values computed by EV1, EV2, LN2 and LP3 distributions

Data series	Computed values of D-index									
	EV1			EV2			LN2		LP3	
	MoM	MLM	OSA	MoM	MLM	OSA	MoM	MLM	MoM	MLM
AMR	1.602	2.588	2.138	2.380	5.828	4.104	1.737	2.222	1.816	1.972
AMAXT	0.099	0.103	0.110	0.119	0.104	0.117	0.069	0.078	0.076	0.078
AMINT	1.715	0.962	0.893	NF	NF	NF	0.562	0.462	0.447	0.403

Selection of probability distribution

Based on the findings obtained through GoF and D-index tests, the selection of probability distributions for EVA of rainfall and temperature for Hissar was made by considering the following points:

(i) For EVA of rainfall using the series of AMR, it is noted that:

  (a) The D-index value of EV1 (MoM) is minimum when compared with the corresponding values of other distributions. But, the A^2 test results did not suggest the EV1 (MoM, MLM and OSA) for EVA of rainfall.
  (b) By eliminating the D-index value of EV1 (MoM), it is observed that the D-index values computed by LN2 (MoM), LP3 (MoM) and LP3 (MLM) are following the value of 1.737, 1.816 and 1.972, respectively.

(ii) For EVA of temperature using the series of AMAXT, it is noted that:

  (a) The D-index values of LN2 (MoM), LP3 (MoM) and LP3 (MLM) are computed as 0.069, 0.076 and 0.078, respectively.
  (b) There is no difference between the D-index values of LN2 (MLM) and LP3 (MLM).

(iii) Generally, the EVA results obtained from MoM may not give satisfactory results as (i) it is difficult to assess exact information about the shape of a distribution by its moments of third and higher order; and (ii) the estimated parameters of distributions were biased in comparison to procedures.

(iv) By eliminating the D-index values of MoM, the D-index value of LP3 (MLM) is found as minimum; and hence LP3 (MLM) is considered as most appropriate distribution for EVA of rainfall and maximum temperature.

(v) For EVA of temperature using the series of AMINT, the D-index value computed by LP3 (MLM) is found as minimum when compared with the corresponding values of other distributions.

(vi) The study suggested the LP3 (MLM) distribution as better suited amongst four distributions adopted in EVA of rainfall and temperature (maximum and minimum).

Conclusions

The paper presents the study carried out for EVA of rainfall and temperature for Hissar by adopting four probability distributions and also assessing the adequacy of the probability distributions using GoF and diagnostic tests with applicable parameter estimation methods. The following conclusions were drawn from the study:

(i) It was found that the estimated extreme values by EV2 (MLM) for rainfall, EV2 (OSA) for maximum temperature and LP3 (MoM) for minimum temperature are higher than the corresponding values of other portability distributions used in EVA of rainfall and temperature.

(ii) Adequacy of fitting of probability distribution to the series of rainfall and temperature was evaluated by GoF tests (using A^2 and KS).

  (a) A^2 test results confirmed the applicability of LN2 (MoM and MLM) and LP3 (MoM and MLM) for EVA of rainfall.
  (b) A^2 test results did not support the EV1 (MoM, MLM and OSA), EV2 (MoM, MLM and OSA), LN2 (MoM and MLM) and LP3 (MoM and MLM) for EVA of annual maximum and minimum temperature.
  (c) KS test results suggested the EV1 (MoM, MLM and OSA), EV2 (MoM, MLM and OSA), LN2 (MoM and MLM) and LP3 (MoM and MLM) for EVA of rainfall and temperature of Hissar.

(iii) On the basis of evaluation of diagnostic test results, the study identified that the LP3 (MLM) distribution

is better suited for EVA of rainfall and temperature (maximum and minimum).

(iv) The study suggested that the 1000-year return period ER + SE value of rainfall of 594 mm, ET(Max) + SE value of maximum temperature of 56.6 °C and ET(Min)-SE value of minimum temperature of 1.5 °C given by LP3 (MLM) distribution could be used for design purposes while designing the hydraulic structures in the vicinity of Hissar region.

References

AERB (2008) Extreme values of meteorological parameters. Atomic Energy Regulatory Board (AERB) safety guide AERB/NF/SG/S-3

AlHassoun SA (2011) Developing an empirical formulae to estimate rainfall intensity in Riyadh region. J King Saud Univ Eng Sci 23(1):81–88

Baratti E, Montanari A, Castellarin A, Salinas JL, Viglione A, Bezzi A (2012) Estimating the flood frequency distribution at seasonal and annual time scales. Hydrol Earth Syst Sci 16(12):4651–4660

Bobee B, Ashkar F (1991) The Gamma family and derived distributions applied in hydrology. Water Resources Publications, Littleton

Charles Annis PE (2009) Goodness-of-Fit tests for statistical distributions. http://www.statisticalengineering.com/goodness.html. Accessed 5 Dec 2017

Esteves LS (2013) Consequences to flood management of using different probability distributions to estimate extreme rainfall. J Environ Manag 115(1):98–105

Hasan H, Salam N, Adam MB (2013) Modelling extreme temperature in Malaysia using generalized extreme value distribution. Int J Math Comput Natural Phys Eng 7(6):618–624

Hughes GL, Rao SS, Rao TS (2007) Statistical analysis and time-series models for minimum/maximum temperatures in the Antarctic Peninsula. Proc R Soc Ser A 463:241–259

Rao AR, Hameed KH (2000) Flood frequency analysis. CRC Publications, Washington

Rasel M, Hossain SM (2015) Development of rainfall intensity duration frequency equations and curves for seven divisions in Bangladesh. Int J Sci Eng Res 6(5):96–101

USWRC (1981) Guidelines for determining flood flow frequency. United States Water Resources Council (USWRC) Bulletin No. 17B

Varathan N, Perera K, Nalin (2010) Board of study in statistics and computer science of the postgraduate institute of science. University of Peradeniya, Sri Lanka

Vivekanandan N (2015) Modelling of annual extreme rainfall, temperature and wind speed Using OSA of EV1 and EV2 distributions. Int J Innov Res Comput Sci Technol 3(4):57–60

Vivekanandan N (2016) Statistical analysis of rainfall data and estimation of peak flood discharge for ungauged catchments. Int J Res Eng Technol 5(2):27–31

Zhang J (2002) Powerful goodness-of-fit tests based on the likelihood ratio. J R Stat Soc Ser B 64:281–294