International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2023; 8(4): 26-35

https://www.mathsjournal.com
Received: 11-04-2023
Accepted: 16-05-2023

Gaddala Prem
Ph.D. Student, Department of Agricultural Statistics, Uttar Banga Krishi Viswavidyalaya, West Bengal, India

M Gopinath Rao
Retired Professor, Department of Agricultural Statistics, Applied Mathematics and Computer Applications, University of Agricultural Sciences, Bangalore, Karnataka, India

Corresponding Author:
Gaddala Prem
Ph.D. Student, Department of Agricultural Statistics, Uttar Banga Krishi Viswavidyalaya, West Bengal, India

Evaluation of statistical models for rainfall at rars: Jagitial

Gaddala Prem and M Gopinath Rao

DOI: https://doi.org/10.22271/maths.2023.v8.i4a.1076

Abstract
In general, the amount of rainfall and its frequency impacts agricultural, ecological, hydrological, economic and living systems. Rainfall is one of the important factors in agriculture. It plays a major role in the growth and development of the crop starting from germination to harvesting. Change in pattern of rainfall will lead to unexpected responses from plant. Thus, farmers in India were facing problems, as most of the crops grown under rainfed condition. Since precipitation is unpredictable, the present study was undertaken to analyse the distribution pattern of rainfall and to know effect of other weather parameters like maximum and minimum temperature, relative humidity, evaporation etc. on rainfall. The latest 31 years secondary data pertaining to weather parameters was collected from RARS, Jagitial, Telangana. For the convenience of analysis data has been divided into 3 major data sets *viz.* Annual, Monsoon seasonal months and Standard Meteorological weeks. The results based on Kolmogorov Smirnov and RMSE tests indicated Exponential distribution was best fitted distribution for most of the data sets. Exponential distribution was best fit for annual and seasonal months, while for SMW's Exponential and Pearson type III were best fitted.

Keywords: Kolmogorov Smirnov, RMSE, Pearson type III

1. Introduction

Rainfall intensity, duration and its distribution play a major role in the growth of agriculture and other related sectors and the overall development of a country. The variability in rainfall affects the agricultural production, water supply, transportation, the entire economy of a region, and the existence of its people. In regions where the year-to-year variability is high, people often suffer great calamities due to floods or droughts. The damage due to extremes of rainfall cannot be avoided completely, a forewarning could certainly be useful and it's possible from analysis of rainfall data.

The prediction of rainfall at a particular place and time can be made by studying the behaviour of rainfall of that place over several years during the past. This behaviour is best studied by fitting a suitable distribution to the time series data on the rainfall (Kainth 1996) [5].

Alam MA, *et al.* (2018) [1], aimed to determine the best-fit probability distributions in the case of maximum monthly rainfall using 30 years of data (1984-2013) from 35 locations in Bangladesh by using different statistical analysis and distribution types. Generalized Extreme Value, Pearson type 3 and Log-Pearson type 3 distributions showed the largest number of best-fit results. Among the best score results, Generalized Extreme Value yielded the best-fit for 36% of the stations and Pearson type 3 and Log-Pearson type 3 each yielded the best-fit for 26% of the stations.

Młyński D, *et al.* (2019) [8], studied to determine the best probability distributions for calculating the maximum annual daily precipitation with the specific probability of exceedance. The Root Mean Square Error (RMSE), and the coefficient of determination (R^2) were used for assessing the fit of empirical and theoretical distributions. They identified that the GEV distribution - recommended for calculating the maximum daily precipitation with the specific probability of exceedance in the catchments of the upper Vistula Basin.

The study on statistical analysis of rainfall may help the government in agricultural planning and policy making, it helps farmers in contingency planning of the crop, adopting the farm production practices, management of the farm production etc. (Husak *et al.*, 2007) [4].

The present study focusses on fitting appropriate Statistical distributions for rainfall data. Thus to understand the distribution of rainfall in RARS, Jagitial region for better agricultural management.

2. Data and Study area

Telangana is classified as a semi-arid state and mean yearly rainfall varies from 677 mm to 998 mm (Waghaye, *et al.*, 2018) [9].

The present study was conducted to know the rainfall variability in and around the region of the Regional Agricultural Research Station, (RARS). The station is located at Polasa village of Jagitial district of Telangana State. It belongs to the Northern Telangana zone of Telangana. The geographical co- ordinates of the station are Latitude: 18-49' 40" North Longitude: 78-56' 45" East. The present study was based on the secondary data on weather parameters over a period of 30 years (1988-2018) which was collected from Regional Agricultural Research Station, Polasa, Jagitial.

3. Methodology

3.1 Fitting Probability distributions

Rainfall is highly variable in a given period. Hence, there is a need to study the distribution both on long term (annually or seasonally) as well as short term (monthly or weekly) basis. Among the weather parameters, amount of daily maximum rainfall (mm) was considered to fit appropriate probability distributions. The probability distributions *viz.*, log normal, Gamma, Generalized extreme value (GEV), Weibull (1 P, 2 P, 3 P), Gumbel, Generalized Pareto, Pearson type III and Log – Pearson type III were used to evaluate the best fit probability distribution for rainfall.

3.2 Description of parameters

3.2.1 Shape parameter

Shape parameters allow a distribution to take on a variety of shapes, depending on the value of the shape parameter. These distributions are particularly useful in modelling applications since they are flexible enough to model a variety of data sets. Examples of shape parameters are skewness and kurtosis.

3.2.2 Scale parameter

In probability theory and statistics, a scale parameter is a special kind of numerical parameter of a parametric family of probability distributions. The larger the scale parameter, the more spread out the distribution. In general, a scale parameter stretches or squeezes a graph. The examples of scale parameters include variance and standard deviation.

3.2.3 Location parameter

The location parameter determines the position of central tendency of the distribution along the x-axis. A location family is a set of probability distributions where μ is the location parameter. The location parameter defines the shift of the data. A positive location value shifts the distribution to the right, while a negative location value shifts the data distribution to the left. Examples of location parameters include the mean, median, and the mode.

A shape-dominated regime describes a pattern where the rainfall tends to be symmetrically distributed, indicating that drier-than-average events are as common as wetter-than-average events. Scale-dominated rainfall describes locations where the variance is quite large in comparison to the mean.

3.3 Testing for goodness of fit

The goodness of fit test measures the discrepancy between observed values and expected values. Kolmogorov- Smirnov test was used to test for the goodness of fit. In the present investigation, the goodness of fit test was conducted at $\alpha = 0.05$ level of significance. It was applied for testing the following hypothesis:

H_0: The maximum daily rainfall data follows a specified distribution.

H_1: The maximum daily rainfall data does not follow a specified distribution.

3.3.1 Kolmogorov- Smirnov test (K-S test)

This test was used to decide whether a sample comes from a hypothesized continuous PDF. The KS test compares the cumulative distribution functions of the theoretical distribution – the distribution described by the estimated shape and scale parameters – with the observed values and returns the maximum difference between these two cumulative distributions (Wilks, 1995) [10]. This maximum difference in cumulative distribution functions is frequently referred to as the KS statistic. It is based on the empirical distribution function *i.e.*, on the largest vertical difference between the theoretical and empirical cumulative distribution functions which is given as:

$$D = \max_{1 \le i \le n} \left(F\left(X_i - \frac{i-1}{n}, \frac{i}{n} - F(X_i)\right)\right)$$

Where, X_i = Random sample, I = 1, 2 … n.

$$CDF = F_n(X) = \frac{1}{n}\,[\,Number\ of\ observations\ \le x]$$

The frequency density at different levels of rainfall was calculated using the following formulae.

$$\text{Relative Frequency} = \frac{\text{Class frequency}}{\text{Total frequency}}$$

$$\text{Frequency Density} = \frac{\text{Relative frequency}}{\text{Class width}}$$

3.3.2 Root Mean Square Error (RMSE)

Root Mean square Error is the standard deviation of the residuals (Prediction errors). It tells about how concentrated the data is around the line best fit. Root Mean error is commonly used in climatology, forecasting and regression analysis to verify experimental results.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Z_{fi} - Z_{oi})^2}{N}}$$

Where,

$(Z_{fi} - Z_{oi})^2$ Are squared differences, and

N is the sample size.

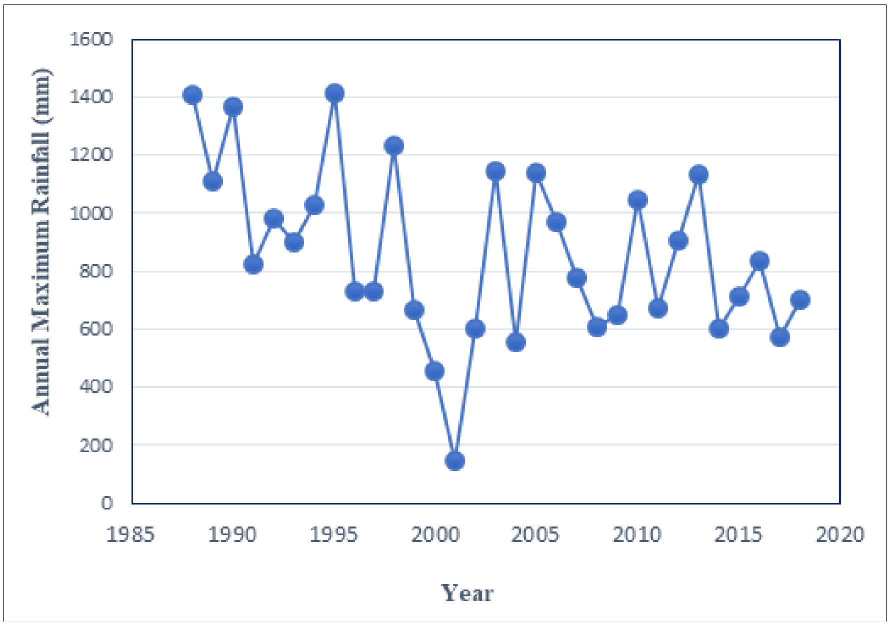


Fig 1: Variation in Annual rainfall (mm)

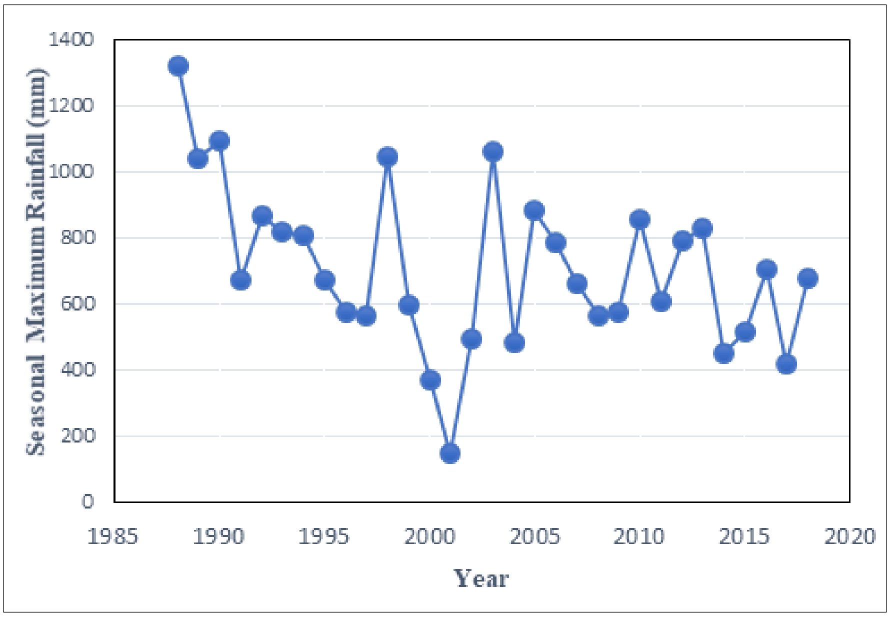


Fig 2: Variation in Monsoon season maximum daily rainfall (mm)

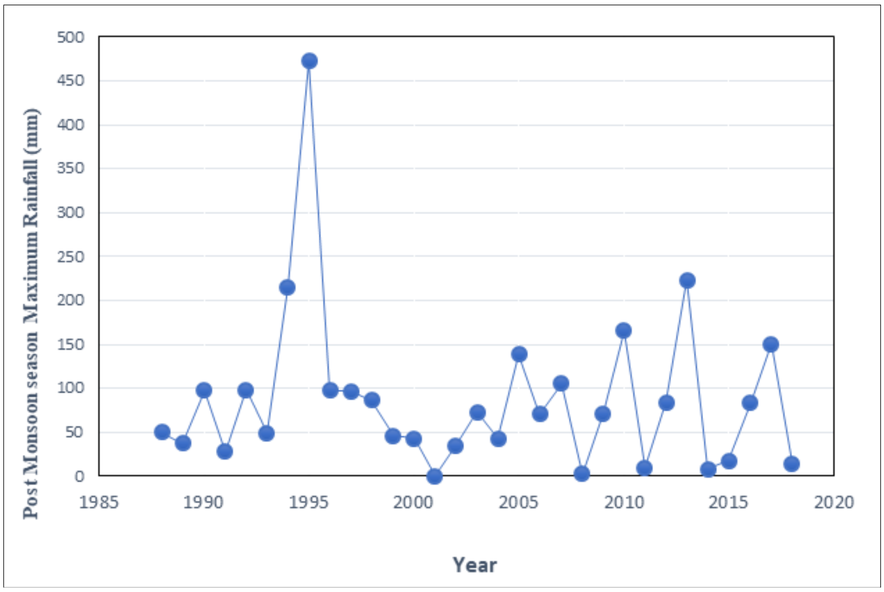


Fig 3: Variation in post monsoon season maximum daily rainfall (mm)

Table 1: Description of various probability description functions

Distribution	Probability density function	Range	Parameters
Log normal	$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x-\mu)^2}{2\sigma^2}\right)$	$-\infty \le \mu \le +\infty$ $\sigma > 0$	μ = Mean σ = Standard deviation
Pearson type III distribution	$f(x) = \frac{\lvert\alpha\rvert}{\Gamma(\beta)}[\alpha(x-\theta)]^{\beta-1}e^{-\alpha(x-\theta)}$	$\alpha \ne 0, \beta > 0$	θ = location parameter, α = scale parameter β = shape parameter
Log Pearson Type III distribution	$f(x) = \frac{1}{a\,x\,\Gamma(b)}\left(\frac{lnx-c}{a}\right)^{b-1} \exp\left[-\left(\frac{lnx-c}{a}\right)\right]$	$a > 0$ $b > 0$ $0 < c < lnx$	a = Scale parameter b = Shape parameter c = Location parameter
Exponential distribution	$f(x) = \lambda \exp(-\lambda x)$	$x > 0$ $\lambda > 0$	λ = Scale parameter
Gumbel distribution	$f(x) = \frac{1}{\beta}\, e^{-(x+e^{-z})}$	$\beta > 0$	μ = location parameter β = Scale parameter
Generalized Extreme value distribution	$f(x) = \frac{1}{\sigma}t(x)^{\xi+1}e^{-t(x)}$ $t(x) = \begin{cases} \left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi} & \xi \ne 0 \\ e^{-\left(\frac{x-\mu}{\sigma}\right)} & \xi = 0 \end{cases}$	$\sigma > 0$	μ = location parameter σ = Scale parameter ξ = Shape parameter
Weibull distribution	$f(x) = \beta x^{\beta-1}\exp(-x^{\beta})$	$x > 0, \beta > 0$	λ = Scale parameter β = Shape parameter
	$f(x) = \frac{\beta}{\gamma}\left(\frac{x}{\gamma}\right)^{\beta-1} e^{-\left(\frac{x}{\gamma}\right)^{\beta}}$	$x > 0, \beta, \gamma > 0$	
	$f(x) = \frac{\beta}{\gamma}\left(\frac{x-\mu}{\gamma}\right)^{\beta-1} e^{-\left(\frac{x-\mu}{\gamma}\right)^{\beta}}$	$x > \mu, \beta, \gamma > 0$	
Generalized Pareto distribution	$f(x) = \frac{1}{b}\left(1 - a\frac{x}{b}\right)^{\frac{1}{a-1}}$	$a \ne 0$	a = shape parameter b = scale parameter

4. Results and Discussion

The weather data of Regional Agricultural Research Station, Jagitial had been collected for 31 years (1988-2018) and were analysed by different statistical tools for achieving the previously stated objectives. The results of the study were shown under the following headings:

4.1 Fitting appropriate probability distributions for rainfall

Rainfall data was categorized into 24 sets *viz.*, 1 annual, 2 seasonal (Monsoon and Post Monsoon), Monsoon season months (Jun-Sep) and 17 Seasonal Standard Meteorological Weeks (23rd SMW – 39rd SMW) to study the distribution pattern of rainfall at different levels. The best-fitted distributions for rainfall on different sets of data were identified by using the Kolmogorov-Smirnov test and Root Mean Square Error tests.

4.1.1 Descriptive statistics

The graphical representation of taken three datasets is shown in Fig. 1, 2, 3. The annual highest precipitation was found to be 1415.1 mm, while for monsoon season highest precipitation recorded was 1319 mm. Similarly, for the Post Monsoon Season, it was observed to be 472.3 mm. The monthly precipitation for monsoon season ranged from 372.6 mm to 636.4 mm while weekly highest precipitation was ranged between 77.6 mm to 422.4 mm.

It was also noted that the lowest precipitation for annual and monsoon season was observed to be same (143 mm). And it ranged from 8.2 mm to 67.6 mm monthly. The lowest precipitation was observed to be 0 mm, for all the weeks. Mean annual rainfall was noted to be 859.21 mm whereas for overall seasonal months, it was 708.36 mm and for the Post Monsoon it was 87.62 mm. During the seasonal months, the mean precipitation ranged from 131.53 mm to 233.99 mm while for weekly, it varied from 14.86 mm to 64.36 mm. The value of CV for annual data was observed to be 34.55 per cent. During Monsoon season, the CV for rainfall varied by 34.72 per cent and in case of Standard Meteorological Weeks, the rainfall variation was seen to be ranging from 85.16 per cent to 162.45 per cent.

The asymmetry of the frequency distribution of data is shown by the coefficient of skewness. The coefficient of skewness for all the data sets ranged from 0.08 to 4.00 indicating positive skewness.

4.1.2 Fitting of probability distributions

The distributions used to estimate the best fit for rainfall were Log Normal, Exponential, Generalized Extreme Value (GEV), Weibull (1P, 2P, 3P), Gumbel, Pearson type III, Log Pearson type III, and Generalized Pareto. The goodness of fit for different probability distributions was tested using Root Mean Square Error and Kolmogorov – Smirnov (KS test). The test statistics along with *p*-values for each data set was computed for the above-mentioned probability distributions. Fig 1 shows the variation in Yearly rainfall from 1988 to 2018. It ranged from 149 mm to 1415.1 mm. The minimum annual rainfall was observed to be 149 mm which was during the year 2001 and the highest precipitation for annual data was noticed during the year 1995. And, the *Fig. 2* shows the variation in rainfall during the monsoon season, varying between 149 mm to 1319 mm. Similarly, *Fig.3* represents the variation in Post Monsoon season maximum daily rainfall. The highest peak was shown during the year 1995 which was about 417 mm and minimum was during the year 2001 when no rains were received.

4.1.3 Fitting of probability distributions

The distributions used to estimate the best fit for rainfall were Log Normal, Exponential, Generalized Extreme Value (GEV), Weibull (1P, 2P, 3P), Gumbel, Pearson type III (P – III), Log Pearson type III (LP – III), and Generalized Pareto (GP). Goodness of fit for different probability distributions was examined using three tests which are as follows:

1. Root Mean Square Error
2. II. Kolmogorov – Smirnov test (KS test)

The test statistic values along with the p – values for each data set was computed for the probability distributions mentioned earlier.
The results of analysis are grouped into three groups and presented as follows:

A. Annual and Seasonal datasets
This section consists of three datasets which include Annual, Monsoon and Post monsoon datasets. For all the data sets in this section the best fitted distribution as shown by KS test is Exponential distribution. The KS test statistic values of the three data sets were 0.10, 0.09 and 0.11 with P-Values of 0.93, 0.95, 0.84 respectively. Among all the three data sets Exponential distribution best describes the Monsoon data set which has the least KS test statistic with highest *p* – value.
Table 2 shows KS test values for eight distributions for annual and monsoon and post monsoon data sets and distributions considered are Log Normal, Exponential, GEV, Weibull (1P, 2P, 3P), Genialized Pareto and Gumbel. Pearson type III distributions was another distribution fitted well for monsoon data. Exponential, Weibull (3P), Log Pearson type III and Gumbel distribution fitted well for Post Monsoon data sets.
The graphs of the best fitted distributions were shown in the Fig. 4. From the graphs we can infer that one parameter Exponential distribution described the data well, which was also indicated by the graph shown in Fig. 4 (C).
In the Table 3 the parameter estimates of the best fitted probability distributions and the RMSE values were shown. The parameter of Exponential distribution for Annual, Monsoon and Post Monsoon were 0.0012, 0.0014, and 0.0114 respectively. The RMSE values of the best fitted distribution for this section were 907.49, 710.06, and 125.71 for the Annual, Monsoon and Post Monsoon season respectively.

B. Monsoon Seasonal months
The months considered under this section were June to September. KS test results showed that a total of nine distributions.

Table 2: Study Period wise fitted probability distributions using KS test

Study Period	Duration	Distribution	Statistic	P-Value
Annual	1 Jan - 31 Dec	Log Normal	0.19	0.24
		Exponential	0.10	0.93
		GEV	0.17	0.36
		Weibull	0.15	0.47
		Weibull (2P)	0.19	0.23
		GP(2P)	0.17	0.32
		Gumbel	0.20	0.18
Monsoon Season	1 June - 30 Sept	Log Normal	0.18	0.26
		Exponential	0.09	0.95
		GEV	0.17	0.31
		Weibull	0.21	0.15
		Weibull (2P)	0.20	0.15
		Pearson type III	0.18	0.26
		GP(2P)	0.19	0.20
		Gumbel	0.17	0.32
Post Monsoon Season	1 Oct - 31 Dec	Exponential	0.11	0.84
		Weibull (3P)	0.13	0.66
		Log - Pearson type III	0.19	0.22
		Gumbel	0.17	0.32
June	1 June - 30 June	Log Normal	0.15	0.47
		Exponential	0.12	0.81
		GEV	0.17	0.33
		Weibull	0.19	0.23
		Weibull (2P)	0.19	0.23
		Pearson type III	0.18	0.29
		Log - Pearson type III	0.19	0.21
		GP(2P)	0.25	0.05
		Gumbel	0.14	0.61
July	1 July - 31 July	Log Normal	0.54	0.45
		Exponential	0.10	0.92
		GEV	0.13	0.67
		Weibull	0.15	0.45
		Weibull (2P)	0.17	0.34
		Pearson type III	0.14	0.61
		Log - Pearson type III	0.18	0.24
		GP(2P)	0.15	0.47
		Gumbel	0.13	0.66
August	1 Aug - 31 Aug	Log Normal	0.14	0.54
		Exponential	0.15	0.53
		GEV	0.15	0.50
		Weibull (3P)	0.17	0.33
		Weibull (2P)	0.21	0.14
		Pearson type III	0.13	0.20
		Log - Pearson type III	0.18	0.29
		GP(2P)	0.14	0.5
		Gumbel	0.17	0.33

Study Period	Duration	Distribution	Statistic	p – value
September	1 Sept - 30 Sept	Log Normal	0.16	0.43
		Exponential	0.11	0.86
		GEV	0.20	0.16
		Weibull (3P)	0.17	0.31
		Weibull (2P)	0.24	0.06
		Pearson type III	0.19	0.24
		Log - Pearson type III	0.21	0.13
		Gumbel	0.18	0.26
23rd SMW	4 June - 10 June	Exponential	0.16	0.39
		Generalized Pareto	0.26	0.03
		Pearson Type III	0.15	0.46
		Log Pearson type III	0.18	0.26
24th SMW	11 June - 17 June	Exponential	0.17	0.35
		GEV	0.09	0.96
		Weibull (3P)	0.21	0.12
		Gumbel	0.20	0.15
		Pearson Type III	0.25	0.04
25th SMW	18 June - 24 June	Exponential	0.10	0.90
		GEV	0.17	0.30
		Weibull (3P)	0.23	0.08
		Gumbel	0.22	0.09
		Pearson Type III	0.11	0.84
		Log Pearson type III	0.19	0.22
26th SMW	25 June - 1 July	Exponential	0.12	0.78
		GEV	0.16	0.41
		Weibull (3P)	0.10	0.89
		Gumbel	0.14	0.59
		Pearson Type III	0.12	0.76
		Log Pearson type III	0.21	0.12
27th SMW	2 July - 8 July	Exponential	0.11	0.87
		GEV	0.13	0.63
		Weibull	0.14	0.58
		Gumbel	0.15	0.49
		Pearson Type III	0.21	0.13
		Log Pearson type III	0.28	0.01
28th SMW	9 July - 15 July	Exponential	0.09	0.98
		GEV	0.13	0.70
		Weibull	0.14	0.54
		Gumbel	0.15	0.48
		Pearson Type III	0.18	0.28

Note: High P-Value indicates best fit

Study Period	Duration	Distribution	Statistic	P-Value
29th SMW	16 July - 22 July	Exponential	0.11	0.86
		GEV	0.17	0.30
		Weibull	0.19	0.24
		Gumbel	0.14	0.58
		Pearson Type III	0.24	0.07
		Log Pearson type III	0.20	0.16
		Weibull (2P)	0.12	0.79
30th SMW	23 July - 29 July	Exponential	0.19	0.23
		GEV	0.13	0.71
		Weibull	0.22	0.09
		Gumbel	0.15	0.45
		Pearson Type III	0.16	0.41
		Log Pearson type III	0.16	0.43
31st SMW	30 July - 5 Aug	Exponential	0.12	0.74
		GEV	0.14	0.61
		Weibull	0.19	0.22
		Gumbel	0.21	0.12
		Pearson Type III	0.19	0.23
		Log Pearson type III	0.23	0.08
32nd SMW	6 Aug - 12 Aug	Exponential	0.14	0.61
		GEV	0.18	0.28
		Weibull	0.15	0.46
		Gumbel	0.16	0.40
33rd SMW	13 Aug - 19 Aug	Exponential	0.15	0.47
		GEV	0.20	0.18

		Weibull	0.19	0.22
		Gumbel	0.19	0.23
		Pearson Type III	0.17	0.35
		Log Pearson type III	0.19	0.20
34th SMW	20 Aug - 26 Aug	Exponential	0.15	0.51
		GEV	0.16	0.42
		Weibull	0.18	0.28
		Gumbel	0.13	0.69
		Log Pearson type III	0.23	0.07
35th SMW	27 Aug - 2 Sept	Exponential	0.19	0.23
		Pearson Type III	0.24	0.06
		Log Pearson type III	0.24	0.05
36th SMW	3 Sept - 9 Sept	Exponential	0.16	0.38
		GEV	0.17	0.31
		Gumbel	0.13	0.64
		Pearson Type III	0.17	0.33
		Log Pearson type III	0.21	0.13
37th SMW	10 Sept - 16 Sept	Exponential	0.17	0.34
		Pearson Type III	0.17	0.35
38th SMW	17 Sept - 23 Sept	Exponential	0.18	0.24
		Pearson Type III	0.15	0.49
39th SMW	24 Sept - 30 Sept	Exponential	0.21	0.12

Note: High P-Value indicates best fit

Fitted for each month under this section. The distribution that did not fit any of the data set under this section was Weibull (1P) distribution. KS test showed that Exponential distribution described well for June, June and September seasonal months, while for August the best fitted distribution was Log Normal distribution with test statistic 0.14 and P-Value 0.54. The test statistic values of Exponential distribution for the month of June, July and September were 0.12, 0.10 and 0.11 with p – values of 0.81, 0.92 and 0.86 respectively.

The parameter estimates of the best fitted distributions were given in the Table 3. The lambda (λ) were 0.0076, 0.0043 and 0.0069 for June, July and September respectively. The parameter estimate of the month of August was 0.0051. Along with parameter estimates the RMSE values of the best fitted distributions were also shown under the same Table. 3. For the 4 monsoon seasonal months the least RMSE value was observed for the month of June month which was of 157.80. The RMSE values for other three months were 265.54, 228.37 and 168.82 respectively. The best fitted distribution graphs for the monsoon months were shown in Fig. 4 (D-G).

C. Standard Meteorological Weeks

In this section Standard Meteorological weeks from 23 to 39 are considered. These weeks are in the months of June to September. KS test showed that Exponential, Generalized Extreme Value, Weibull (3P), Gumbel, Pearson type III distributions fitted for most of the data sets. For 37th, 38th and 39th Standard Meteorological Weeks only Exponential, and Pearson type III distributions fitted well. The KS test showed that Exponential distribution was the best fitted for majority of the Standard Meteorological Weeks. For few other weeks which showed other distributions as the best fitted distributions. In those Pearson type III distribution was the best fitted distribution for 23rd, 37th and 38th SMW's. Similarly, Generalized Extreme Value was the best fitted distribution for 24th and 30th SMW's. Likewise, for 26th and 36th SMW's the best fitted distributions were Weibull (3P) and Gumbel distribution.

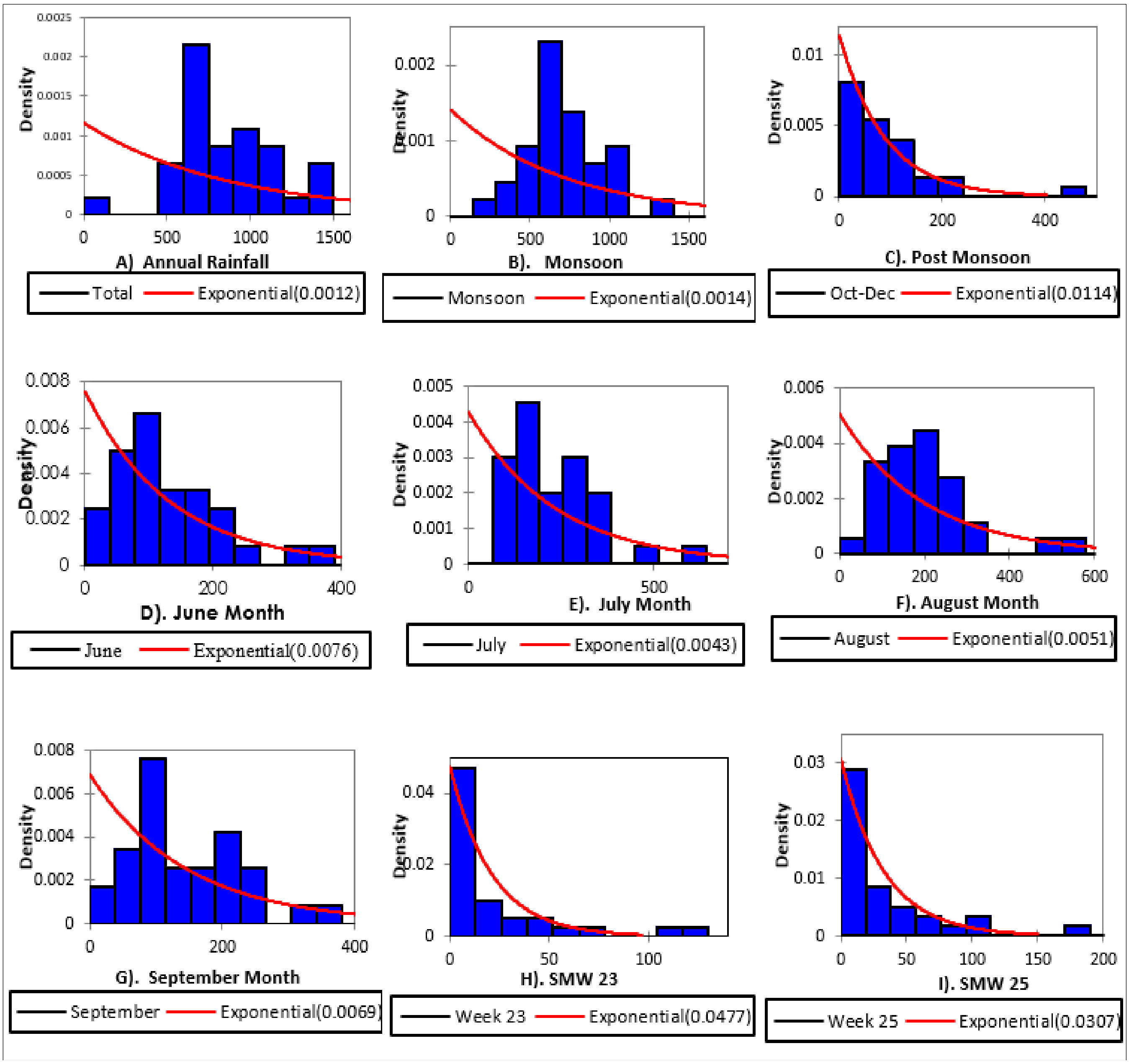
From all the distributions that fitted well for the SMW's Exponential distribution described well for 28th SMW which had the least KS test statistic value of 0.09 with a P-Value of 0.98. Along with this SMW, it also described the other SMW's like 25th, 27th and 29th which had the test statistic values of 0.10, 0.11 and 0.11 with a P-Values of 0.90, 0.87 and 0.86 respectively. Generalized Extreme Value distribution described 24th SMW well which also showed least test statistic value of 0.09 with a P-Value of 0.96. For 26th SMW Weibull (3P) also showed a least KS test statistic value of 0.10 with a *p* – value of 0.89.

The RMSE values and the parameter estimates for best fitted distributions for the selected Standard Meteorological Weeks were also shown in the Table 3. The RMSE values for the SMW's were low when compared to the RMSE values of the above two sections. The RMSE values under this section ranged between 33-92. The least RMSE value was observed for 39th SMW for Exponential distribution which was of 33.77. The best fitted distribution graphs for these SMW's were shown under the Fig. 4 (H - X). From the given graphs we can infer that for SMW 23, 25, 27, 32, 35, 37 the particular distributions described.

Table 3: Parameters estimates of best fitted Probability distributions

Study Period	Range	Distribution	Location	Scale	Shape	RMSE value
Annual	1 Jan - 31 Dec	Exponential	λ = 0.0012			907.49
Monsoon	1 June - 30 Sept	Exponential	λ = 0.0014			710.06
Post Monsoon	1 Oct - 31 Dec	Exponential	λ = 0.0114			125.71
June	1 June - 30 June	Exponential	λ = 0.0 076			157.80
July	1 July - 31 July	Exponential	λ = 0.0043			265.54
August	1 Aug - 31 Aug	Log Normal	λ = 0.0051			228.37

September	1 Sept - 30 Sept	Exponential	λ = 0.0069			168.82
23 SMW	4 June - 10 June	Exponential	λ = 0.0477			38.23
24 SMW	11 June - 17 June	GEV	μ = 13.7636	σ = 17.1372	ξ =0.6831	70.45
25 SMW	18 June - 24 June	Exponential	λ = 0.0307			51.66
26 SMW	25 June - 1 July	Weibull 3P	β =1.0703	γ = 40.0512	μ =-4.1893	49.58
27 SMW	2 July - 8 July	Exponential	λ = 0.0227			59.93
28 SMW	9 July - 15 July	Exponential	λ = 0.0195			68.51
29 SMW	16 July - 22 July	Exponential	λ = 0.0207			65.33
30 SMW	23 July - 29 July	GEV	μ = 38.5240	σ = 37.1452	ξ = 0.1115	83.95
31 SMW	30 July - 5 Aug	Exponential	λ = 0.0202			65.15
32 SMW	6 Aug - 12 Aug	Exponential	λ = 0.0193			91.32
33 SMW	13 Aug - 19 Aug	Exponential	λ = 0.0303			51.83
34 SMW	20 Aug - 26 Aug	Gumbel	μ = 26.7241	β = 31.9772		66.55
35 SMW	27 Aug - 2 Sept	Exponential	λ = 0.0223			66.70
36 SMW	3 Sept - 9 Sept	Gumbel	μ = 20.9452	β = 24.6865		48.42
37 SMW	10 Sept - 16 Sept	Exponential	λ = 0.0325			52.44
38 SMW	17 Sept - 23 Sept	Exponential	λ = 0.0246			76.64
39 SMW	24 Sept - 30 Sept	Exponential	λ =0.0458			33.77



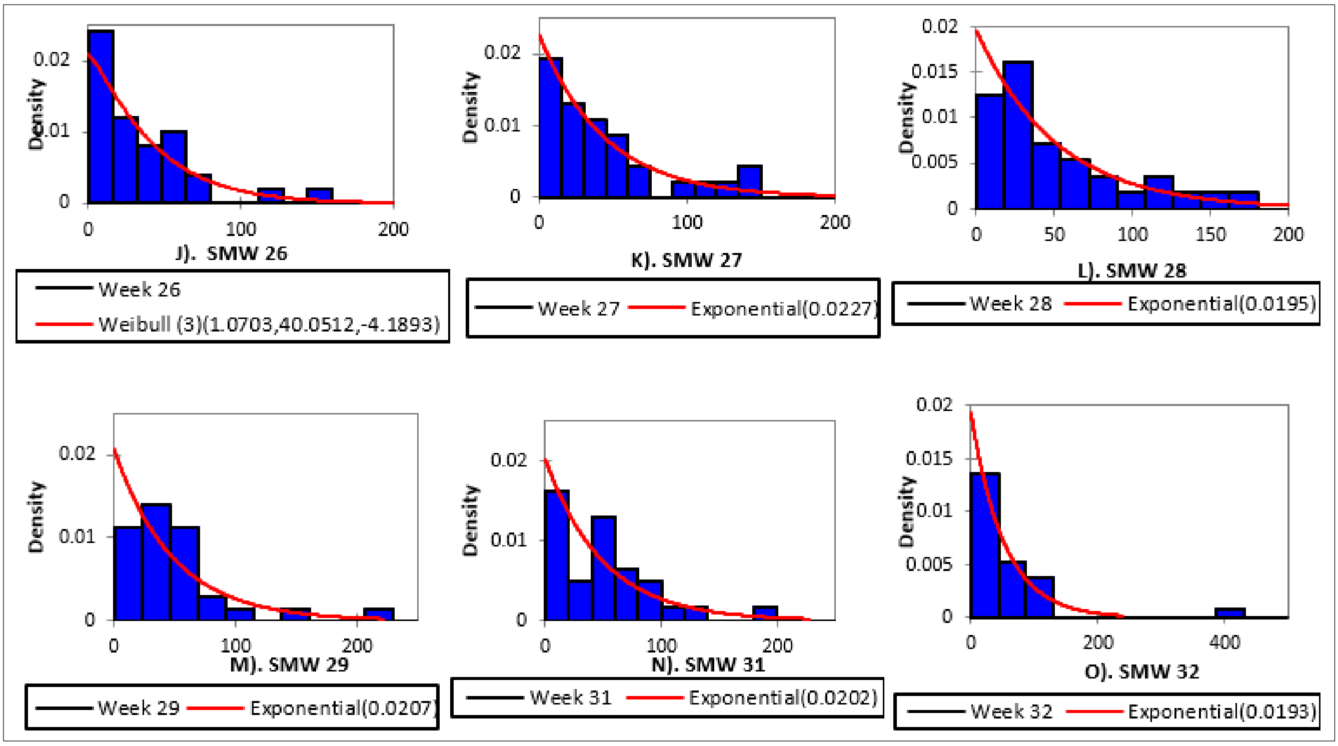


Fig 4: Graphs of Best Fitted Distributions by KS test and Anderson Darling Test

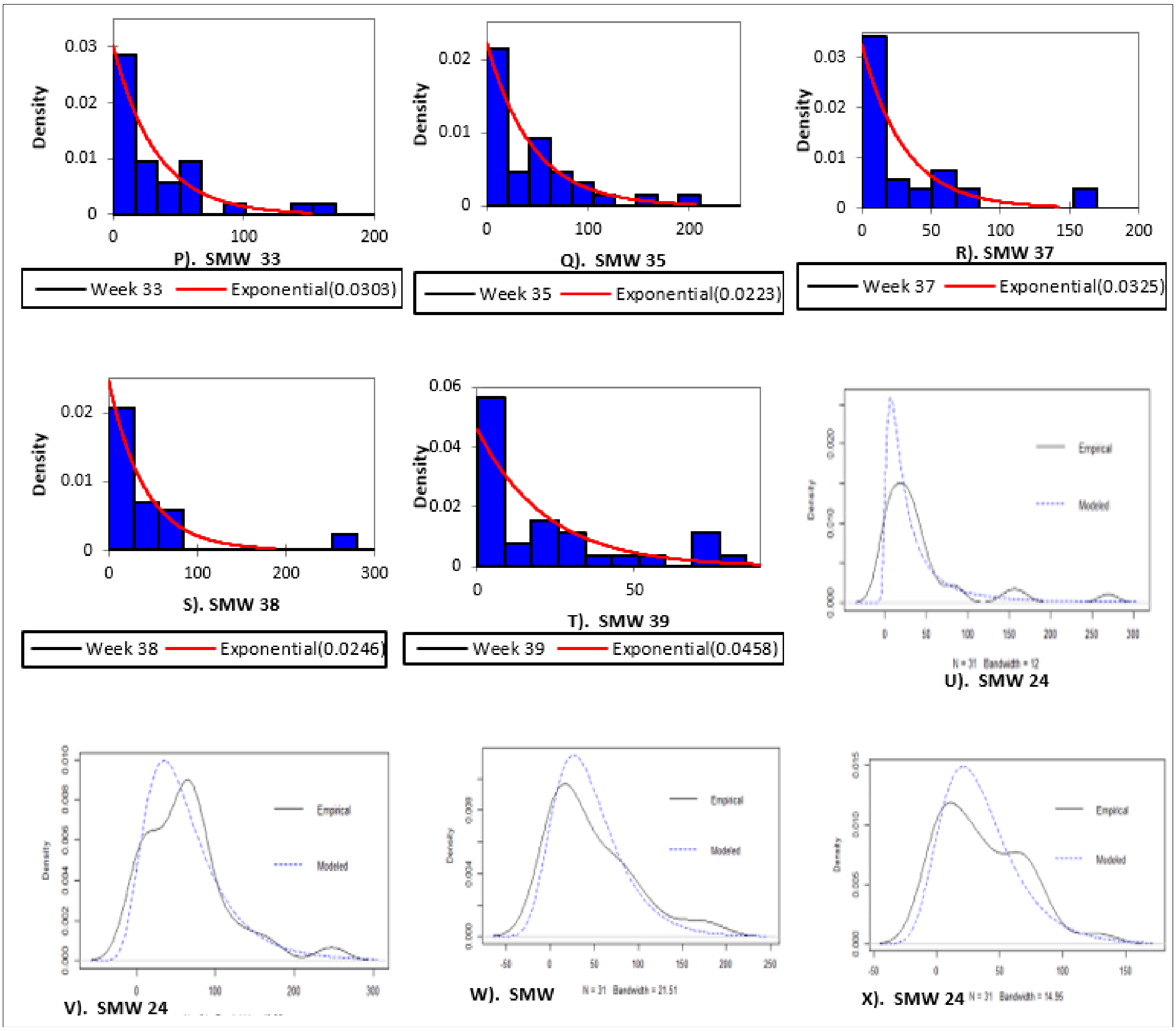


Fig 5: Graphs of best fitted distributions by KS test and Anderson darling test

Data sets well. Exponential distribution is a better fit than other distribution for different sets of rainfall data of Jagitial research station. Baghel, *et al*. (2019) [2] fitted seven different distribution in their work on rainfall data analysis and Log Normal and Gumbel distribution fitted well in majority of data sets. Bhoomika Raj (2015) [3] who analysed rainfall data of GKVK, Bengaluru trying nine different distribution found Weibull (3P) distribution fitting in majority of data.

Krishnamurthy (2017) [6] in his study tried 11 probability distribution for rainfall data and Log Normal and three parameter Weibull distribution fitted well for three stations *viz,* Hiriyur, Bengaluru and Mandya. Rainfall data analysed in search of best fitted distribution August month and the three parameter Weibull distribution was best fitted distribution for 26th Standard Meteorological Week.

In the present study Exponential, GEV, Pearson type III, Weibull (3P) and Gumbel distribution fitted well as based on KS test. Exponential distribution was the best fit in many cases. Fig. 4. (C), (K), (O), and (U) amply demonstrate that Exponential distribution better described the data sets. Pearson type III is describing the data sets next to Exponential distribution.

5. Conclusion

This study was attempted to understand the distribution pattern of rainfall at Regional Agricultural Research Station, Jagitial. The test used to find out the best fitted distribution are KS test and RMSE value. The results revealed that Exponential distribution widely fitted for almost all the data sets as shown by KS test. Along with Exponential distribution, Weibull (3P) and Gumbel distributions were also the next best fitted distributions for Annual and Seasonal data sets. While, for the monsoon season months Log Normal and Gumbel distributions also described the data sets well. In case of SMW's the first half of Standard Meteorological Weeks showed the best results of KS test.

6. References

1. Alam MA, Farnham EKC, Yuan J. Best-fit probability distributions and return periods for maximum monthly rainfall in Bangladesh. Climate. 2018;6(1):9-30.
2. Baghel H, Mittal HK, Singh PK, Yadav KK, Jain S. Frequency Analysis of rainfall data using probability distribution models. Int. J. Current Microbiol and Appl. Sci. 2019;8(6):1390-1396.
3. Bhoomika Raj R. Evaluation of statistical models for climatic characterization of GKVK Station. M.Sc. Thesis, Univ. Agric. Sci., Bengaluru; c2015.
4. Husak GJ, Michaelsen J, Funk C. Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. Int. J. Climatol. 2007;27(7):935-944.
5. Kainth GS. Weather and Supply Behaviour in Agriculture: An Econometric Approach. Daya Books; c1996.
6. Krishnamurthy KN. Statistical models for climatic characterization of some selected zones of Karnataka. Lulu Publications U.S.A. ISBN No: 97871-387-32730-0; c2017.
7. Mishra AK. Effect of rain gauge density over the accuracy of rainfall: A case study over Bangalore, India. Springer Plus. 2013;2(1):1-7.
8. Młyński D, Wałęga A, Petroselli A, Tauro F, Cebulska M. Estimating maximum daily precipitation in the upper Vistula basin, Poland. Atmosphere. 2019;10(2):43.
9. Waghaye AM, Rajwade YM, Randhe RM, Kumari N. Trend analysis and change point detection of rainfall of Andhra Pradesh and Telangana, India. J. Agrometeorology. 2018;20(2):160-163.
10. Wilks DS. Statistical Methods in the Atmospheric Sciences: An Introduction. Academic Press: San Diego, CAL; c1995.