Int.J.Curr.Microbiol.App.Sci (2018) *7*(4): 3053-3062

EXCELLENT PUBLISHERS

International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7706 **Volume 7 Number 04 (2018)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2018.704.346

Statistical Distribution of Seasonal Rainfall Data for Rainfall Pattern in TNAU1 Station Coimbatore, Tamil Nadu, India

S. Lavanya[1*], M. Radha[1] and U. Arulanandu[2]

[1]Agricultural Statistics, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India
[2]Statistics, Anbil Dharamalingam Agricultural College and Research Institute, Trichy, Tamil Nadu, India

**Corresponding author*

Keywords

Seasonal rainfall, Statistical distributions, Goodness of fit–test, Mann Kendall trend test, Sen's slope estimator

Article Info

Accepted:
23 March 2018
Available Online:
10 April 2018

A B S T R A C T

Rainfall plays an important role in agricultural production system. Studying the rainfall distribution through the statistical models is also essential in agricultural field. In this paper, an attempt has been made to show the pattern of rainfall in TNAU1 Station Coimbatore through the Statistical distributions. In the study, the different statistical distributions are fitted for seasonal rainfall data and the best fit is determined. The historical rainfall data were collected for 34 years (1982 -2015) from the Meteorological Observatory AC & RI Coimbatore. The statistical distributions like Normal, Log-normal, Weibull, Gamma, Logistic, Exponential and Generalized Extreme value (GEV) are used for three different seasons such as Kharif, Rabi and summer. For Parameter estimation the Method of Maximum Likelihood are used here. The Comparisons of best distributions by goodness of fit-tests such as Kolmogorov–Smirnov, Anderson Darling test and Chi square test are also made in this paper. The trend test such as Mann Kendall trend test and Sen's slope estimation were performed for studying the rainfall pattern. The Maximum rainfall occurred in the Rabi season with 9.028 (mm) and minimum rainfall in the summer with 0.371 (mm).

Introduction

The primary source of agricultural production for most of the country is rainfall. Efficient Utilization of rainfall improves the crop growth and development. About 80% of the world and 60% of Indian Agriculture is rain dependent.

Changing rainfall pattern effect is directly felt on cropping pattern. Considering the importance and issues of rainfed farming, the study was undertaken in Coimbatore region of Tamil Nadu for determining the rainfall pattern which will be useful in crop planning. The two and three parameter distribution Normal, log-Normal, Weibull, Gamma, Logistic, Exponential, GEV (Generalized extreme values) were used here to study the best fit. The rainfall seasons of Tamil Nadu are Kharif (June to September), Rabi (October to January) and Zaid or summer (February to May). The selection of best fitted distribution is the key in determining rain-fall pattern.

Probability and frequency analysis of rainfall data enables us to determine the expected rainfall at various chances (Bhakar *et al.*, 2008). The rainfall pattern decides the cultivation of crops, their varieties, adoption of cultural operations and harvesting of excess rain water of any region (Sinhabab, 1977; Budhar *et al.*, 1987; Thalur, 1998, Kar *et al.*, 2004). Kumar *et al.*, (2007) conducted a study to identify best fit trend of weekly rainfall data of 45 years (1955-1999) for the western Uttar Pradesh at various probability levels and expected rainfall frequency for crop planning and management to work out the irrigation period required for the crops and found 70% probability level was useful for planning of kharif crops and to decide proper time for various agriculture operations. Rajendran *et al.*, (2016) conducted a study to analyze the nature of distribution and frequency of rainfall of 32 years (1982-2013) for Dharmapuri district. Rainfall frequency analysis was done using Weibull's method and resulted that annual average rainfall of 938.1mm can be expected to occur once in 2.5 years at a probability of 40%. Fisher (1924) studied the influence of rainfall on the yield of wheat in Rothamasted. He showed that it is the distribution of rainfall during a season rather than its total amount which influence the crop yield.

Materials and Methods

Coimbatore region comprise of 10 Taluks with Latitudes N 11° 0' and 16.4016' and Longitudes E 76° 57' and 41.8752' (Fig. 1). The total geographical area of the district is 642.12 km2 (net area sown is 6.0%). The normal annual rainfall of North East Monsoon is 328.9mm and South West Monsoon 189.8mm. Thus the major part of rainfall comes from the North-east Monsoon. The Yearly rainfall (mm) data of 34 years (1982-2015) were collected from Meteorological Observatory AC and RI, Coimbatore and it was converted by arithmetical ways to seasonal rainfall data. The best fit probability distribution was evaluated by using the goodness of fit test such as Kolmogorov - Smirnov test, Anderson darling test and chi-square test at 5% level of significance.

Statistical distributions

Probability distributions are widely used for understanding the rainfall distribution and computation of assumed rainfall (Rajendran *et al.*, (2016). In this study the distributions such as normal, log-normal, Weibull, gamma, logistic, Exponential, extreme value distributions are considered for selection of appropriate distributions. These are the class of continuous probability distributions.

Normal distribution

f (x;μ,σ)= x > 0; with μ as the mean, σ as the standard deviation and follows X~N(μ,σ2).

Log-normal distribution

f(x;μ,σ)=, x > 0; with parameters μ as the location parameter, σ as the scale parameter and follows loge X ~N(μ,σ2).

Weibull distribution

f(x; c, α, μ)= x≥0,c,α>0, c as the shape parameter,α as the scale parameter and μ is the location parameter.

Gamma distribution

f(x; μ, θ) =, with λ as the shape parameter,β as the scale parameter, μ as the location parameter.

Logistic distribution

f(x; λ, s) = with parameters a belongs R and b > 0.

Exponential distribution

f(x, θ) = ;with parameter θ > 0.

Generalized extreme value distribution

f(x; μ, σ) = with parameters μ belongs R and σ > 0,k=shape parameter,μ is the location parameter, σ is the scale parameter.

Fitting the distribution and their parameters were estimated using EASY FIT software.

Goodness of fit test

Chi-square test

The chi-square test enables us to test whether the data agrees the observed and theoretical frequencies (Prof. Karl Pearson in 1900). It is calculated using the formula:

$$\chi^2 = \sum_{i-1}^{n} \left[\frac{(O_i - E_i)^2}{E_i} \right] \sim \text{(n-1) d.f.}$$

Where,

O_i= Observed frequency in the ith cell,
E_i= Expected frequency in the ith cell,
'i'=number of observations (1,2,…k)

The expected frequency can be computed by: $E_i = Np_i$,

Where, N= total number of observations, P= probability of the ith cell.

Kolmogorov –Smirnov test

The Kolmogorov –Simonov (K-S) test is a goodness of fit test used to determine whether an underlying probability distribution differs from a hypothesized distribution when given a finite data set (Guizani, 2010). It is based on empirical distribution function.

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

Where,

X_i = Random sample, i=1,2….n.
CDF=$F_n(X)$ = [number of observations ≤ x].

Anderson Darling test

The Anderson-Darling test (Stephens, 1974) is used to test if a sample of data comes from a population with a specific distribution.

The test is based on the distance, the EDF function is given by

$$A^2 = n \int_{-\infty}^{\infty} (F_n(x) - F(x))^2 w(x) dF(x),$$

Where, w(x) is the weight function, F_n, cumulative distribution function.

This test gives more weight to the tails than the Kolmogorov-Smirnov test.

The goodness of fit-test was obtained using the EASYFIT software.

Trend analysis

The trend analysis was used to detect the monotonic increasing and the decreasing pattern in the data and Sen's slope estimator for estimating the magnitude of the pattern.

Mann Kendall trend test

The Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable over time. It is a non-parametric test for detecting the trend.

The Mann-Kendall test statistic S is calculated using the formula that follows:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sign\ (T_j - T_i)$$

$$\text{Sign}(T_j - T_i) = \begin{cases} 1 \text{ if } T_j - T_i > 0 \\ 0 \text{ if } T_j - T_i = 0 \\ -1 \text{ if } T_j - T_i < 0 \end{cases}$$

Where T_j and T_i are the seasonal values in years j and i, j>i, respectively.

This statistics represents the number of positive differences minus the number of negative differences for all the differences considered. For large samples (N>10), the test is conducted using a normal approximation (Z statistics) with the mean and the variance as follows:

$$Z = \begin{cases} = \frac{s-1}{\sqrt{\text{var}(s)}} \text{ if } s > 0 \\ = 0 \text{ if } s = 0 \\ = \frac{s+1}{\sqrt{\text{var}(s)}} \text{ if } s < 0 \end{cases}$$

The presence of a statistically significant trend is evaluated using the Z value. A positive value of Z indicates an upward trend and its negative value a downward trend. The statistic Z has a normal distribution.

To test for either an upward or down-ward monotone trend (a two-tailed test) at α level of significance, H0 is rejected if the absolute value of Z is greater than Z1-α/2, where Z1-α/2 is obtained from the standard normal cumulative distribution tables. The Z values were tested at 0.05 level of significance.

Sen's Slope estimator

Theil-Sen estimator of a set of two-dimensional points (X_i, Y_i) is the median m of the slopes $(Y_j-Y_i)/(X_j-X_i)$ determined by pairs of samples point given by Theil (1950).

To estimate the true slope of an existing trend (as change per year) the nonparametric method Sen's slope is used. This method can be used when the trend can be assumed to be linear. This procedure was developed by Sen (1968).

$f(t) = Qt + B \rightarrow (1)$

Where, Q is the slope, B is a constant and t is the time.

To get the slope estimate, Q in the equation (1) we first calculate the slope of all data value pairs

$$Q_i = \frac{x_j - x_k}{j - k} \quad i = 1, 2 \ldots k. j>k \rightarrow (2)$$

If there are n values X_j in the time series we get as many as N = n(n-1)/2 slope estimates Qi.

The Sen's estimator of slope is the median of these N values of Qi. The N values of Q_i are ranked from the smallest to the largest and the Sen's estimator is

$$Q = \begin{cases} Q_{\frac{N+1}{2}} & \text{if } N \text{ is Odd} \\ \frac{1}{2}\left(Q\frac{N}{2} + Q\frac{N+1}{2}\right) & \text{if N is even} \end{cases} \rightarrow (3)$$

Then the B values are estimated using the formula $B = X_i - Qt_i. \rightarrow (4)$

The Mann Kendall and Sen's slope estimation were analyzed using XLSTAT 17 Software package.

Results and Discussion

Statistical distributions are widely used for fitting the appropriate distributions widely for the rainfall data and also for studying the rainfall pattern prevailing for the crop management.

Table.1 Summary statistics for seasonal rainfall (1982-2015)

Seasons	Minimum(mm)	Maximum(mm)	Mean(mm)	Standard deviation	Skewness	Kurtosis	CV
Kharif	0.593	2.963	1.460	0.541	0.7945	0.151	0.370
Rabi	0.825	**9.028**	3.239	1.488	1.5755	4.448	**0.459**
Summer	**0.371**	2.842	1.506	0.681	0.2217	-1.030	0.452

Table.2 Estimated parameters of the statistical distributions

Distribution	Seasons	Parameters
Exponential	Kharif	□ =0.68477
	Rabi	□ =0.30875
	Summer	□ =0.66416
Exponential (2P)	Kharif	□ =1.1534 □ =0.59333
	Rabi	□ =0.41431 □ =0.82517
	Summer	□ =0.88099 □ =0.37057
Gamma	Kharif	□ =7.2992 □ =0.20007
	Rabi	□ =4.7391 □ =0.68344
	Summer	□ =4.8946 □ =0.30762
Gamma (3P)	Kharif	□ =3.9648 □ =0.27324 □ =0.37699
	Rabi	□ =6.2919 □ =0.56662 □ =-0.32738
	Summer	□ =7.6324 □ =0.25127 □ =-0.41434
Gen. Extreme Value	Kharif	k=-0.02875 □ =0.4493 □ =1.2134
	Rabi	k=-0.12016 □ =1.2318 □ =2.66
	Summer	k=-0.18942 □ =0.66078 □ =1.2302
Logistic	Kharif	□ =0.29801 □ =1.4603
	Rabi	□ =0.82027 □ =3.2388
	Summer	□ =0.37522 □ =1.5057
Lognormal	Kharif	□ =0.36582 □ =0.31295
	Rabi	□ =0.47311 □ =1.0725
	Summer	□ =0.51556 □ =0.2911
Lognormal (3P)	Kharif	□ =0.33959 □ =0.387 □ =-0.0986
	Rabi	□ =0.27297 □ =1.5893 □ =-1.8506
	Summer	□ =0.20264 □ =1.1849 □ =-1.8267
Normal	Kharif	□ =0.54052 □ =1.4603
	Rabi	□ =1.4878 □ =3.2388
	Summer	□ =0.68057 □ =1.5057
Weibull	Kharif	□ =3.2142 □ =1.5796
	Rabi	□ =2.5383 □ =3.4925
	Summer	□ =2.2064 □ =1.6694
Weibull (3P)	Kharif	□ =1.7927 □ =-1.0418 □ =0.53282
	Rabi	□ =1.8232 □ =2.9186 □ =0.63673
	Summer	□ =2.0176 □ =1.4615 □ =0.20979

Table.3 Goodness of Fit – Summary for kharif season

s.no	Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Exponential	0.37048	11	6.5824	11	30.642	11
2	Exponential (2P)	0.21949	10	2.8947	10	8.4489	10
3	Gamma	0.07422	6	0.1541	2	0.19885	6
4	Gamma (3P)	0.07095	2	0.16062	4	0.09518	3
5	Gen. Extreme Value	**0.06553**	**1**	**0.14998**	**1**	0.0974	4
6	Logistic	0.13284	9	0.44028	9	2.0355	9
7	Lognormal	0.07172	4	0.17114	6	0.09442	2
8	Lognormal (3P)	0.07114	3	0.16301	5	**0.09149**	**1**
9	Normal	0.11797	8	0.4303	7	1.6638	8
10	Weibull	0.08734	7	0.43541	8	0.86603	7
11	Weibull (3P)	0.07236	5	0.16019	3	0.10286	5

Table.4 Goodness of Fit – Summary for Rabi season

s.no	Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Exponential	0.30017	11	5.5298	11	22.63	11
2	Exponential (2P)	0.23061	10	3.8285	10	4.8824	10
3	Gamma	0.11801	8	0.57906	6	1.4711	7
4	Gamma (3P)	0.10618	4	0.4915	4	**0.03754**	**1**
5	Gen. Extreme Value	0.09329	2	0.44857	2	0.58766	6
6	Logistic	**0.09019**	**1**	**0.41701**	**1**	0.22855	4
7	Lognormal	0.14378	9	0.82024	9	3.6768	9
8	Lognormal (3P)	0.10526	3	0.45804	3	0.03836	2
9	Normal	0.11271	5	0.69273	8	0.45813	5
10	Weibull	0.1145	6	0.55398	5	0.06479	3
11	Weibull (3P)	0.11701	7	0.65818	7	2.4361	8

Table.5 Goodness of Fit – Summary for summer season

s.no	Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Exponential	0.25784	11	4.4569	11	14.736	11
2	Exponential (2P)	0.20488	10	2.7612	10	4.6301	10
3	Gamma	0.09278	7	0.42892	7	1.1549	7
4	Gamma (3P)	0.08123	5	0.28497	4	0.51382	2
5	Gen. Extreme Value	**0.06623**	**1**	**0.21349**	**1**	0.8672	3
6	Logistic	0.09804	8	0.47208	8	1.6901	8
7	Lognormal	0.12216	9	0.55148	9	2.363	9
8	Lognormal (3P)	0.08404	6	0.29292	5	0.96056	4
9	Normal	0.08107	4	0.30573	6	1.0345	5
10	Weibull	0.08066	3	0.23745	2	**0.40469**	**1**
11	Weibull (3P)	0.07784	2	0.26952	3	1.1273	6

Table.6 Annual and seasonal trend of the TNAU1 station, Coimbatore

Season	Mann Kendall trend test(Z)	Sen's slope estimator(Q)	b (regression Coefficient)	Trend
Kharif	-0.030	-0.003	-0.0034	Downward
Rabi	0.048	0.007	-0.0122	Upward
Ziad or summer	0.187	0.021	0.0203	Upward
Annual	0.112	0.008	0.0023	Upward

Fig.1 Coimbatore district Map

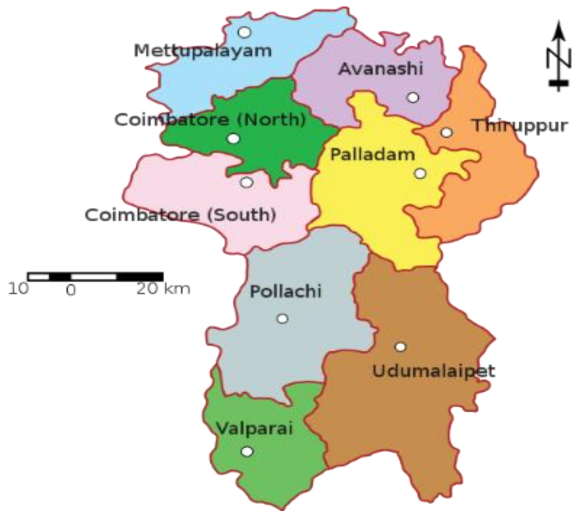


Fig.2 Probability density curve of all the distributions for kharif season

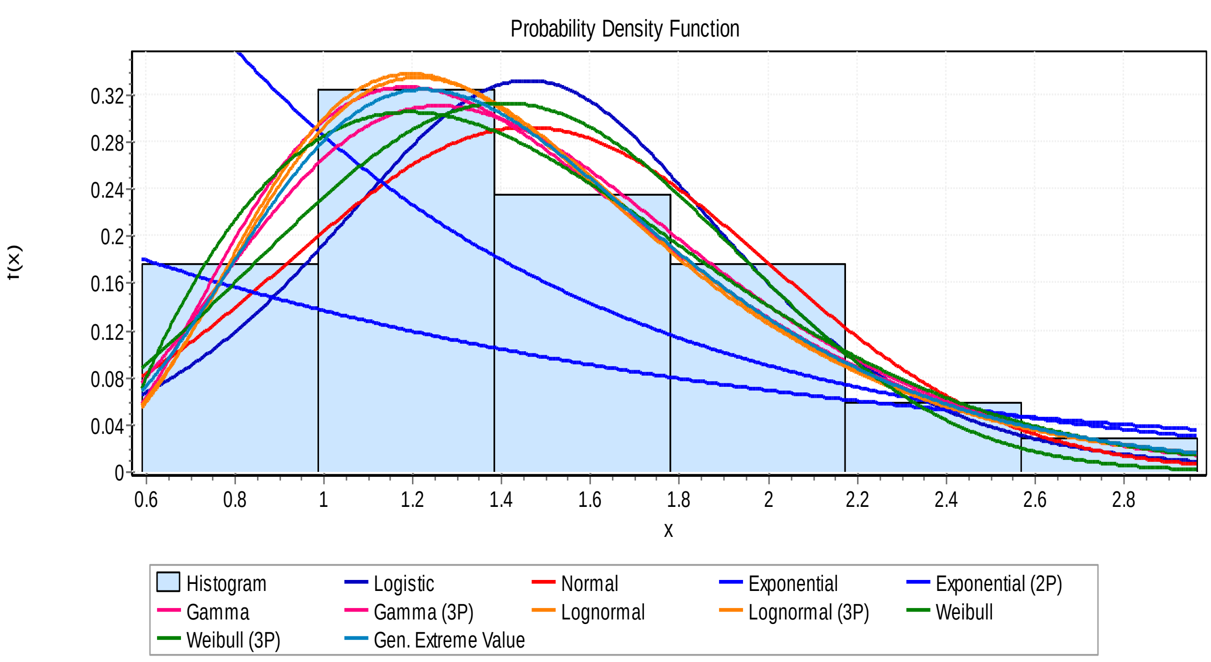


Fig.3 Probability density curve of all the distributions for Rabi season

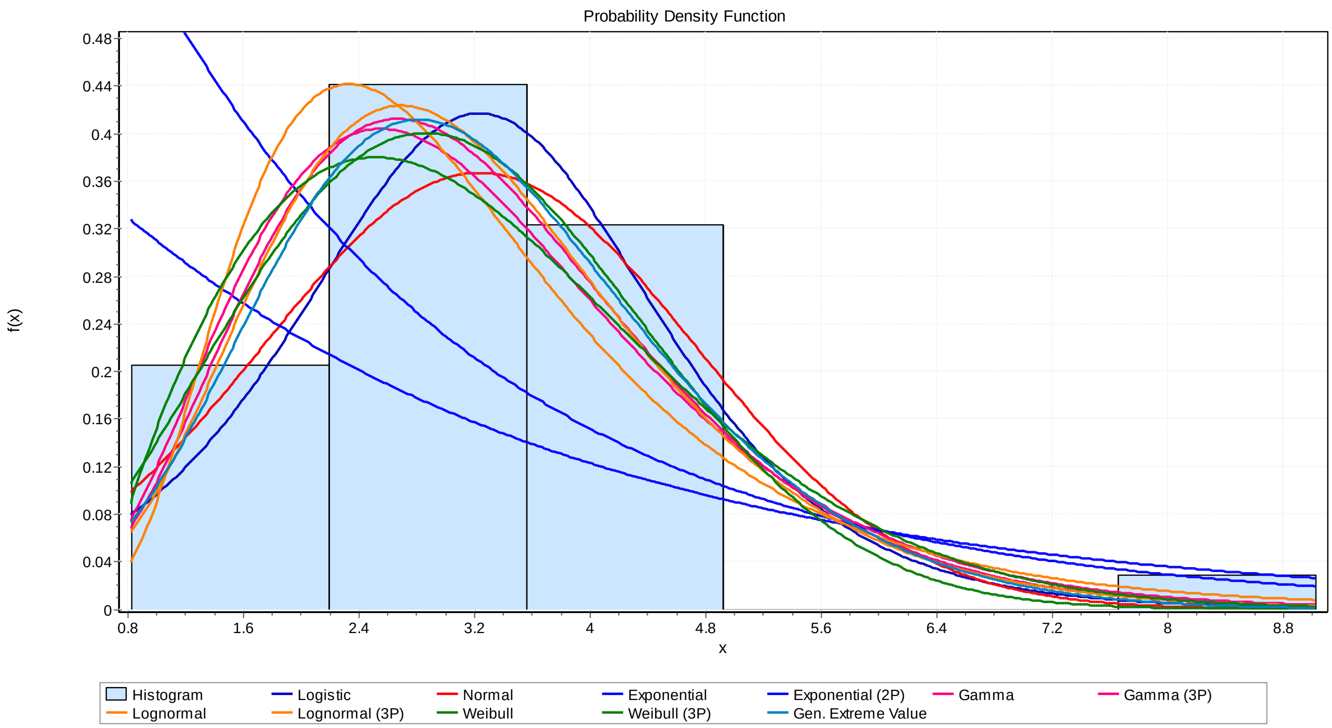
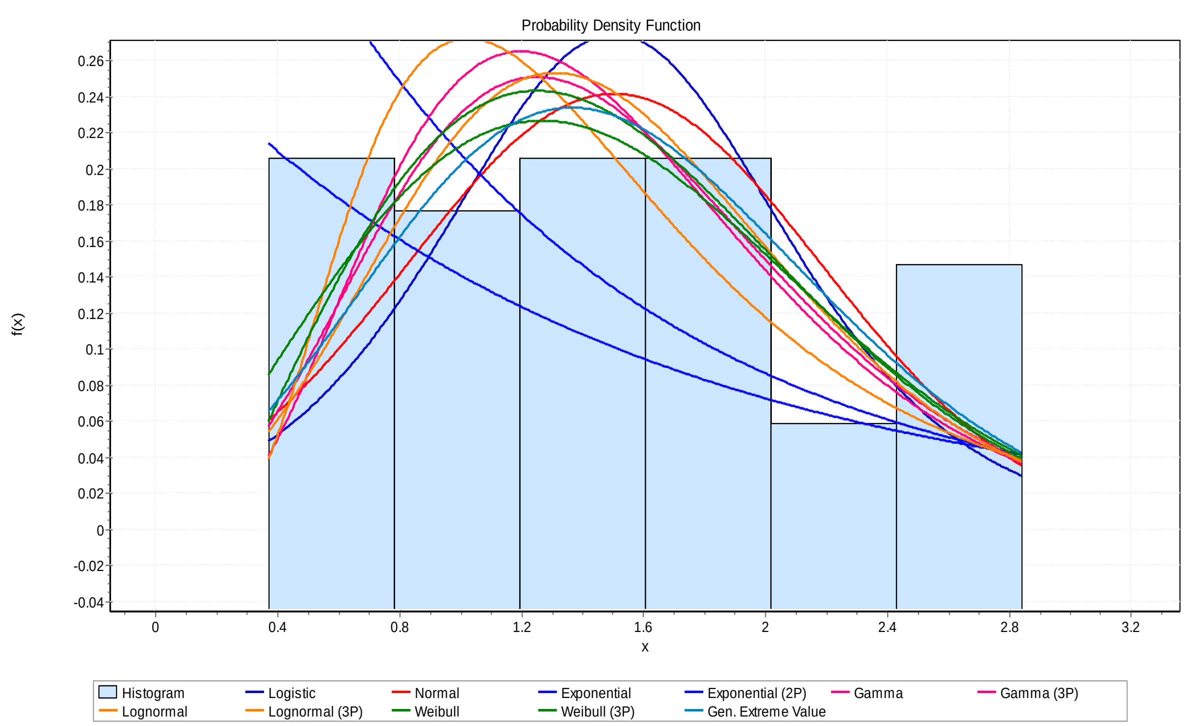


Fig.4 Probability density curve of all the distributions for Ziad season



Seven statistical distributions such as Normal, log-normal, exponential, logistic, Weibull, Gamma and GEV distribution for all the three seasons (Kharif, Rabi and Ziad or summer) are fitted in this study for best appropriate distribution. Under Kharif season, the results of goodness of fit-test of Kolmogorov Smirnov test and Anderson Darling test are best fit for GEV distribution and under chi-square test log-normal distribution were found as the best fit. Next for Rabi season, it was found best fit for logistic distribution on Kolmogorov-Smirnov test and gamma distribution under Chi-square test. Then for summer season, GEV distribution was found best fit under Kolmogorov-Smirnov test and Anderson Darling test, whereas Weibull distribution was found best fit under chi-square test. The results of above discussion showed that for three distributions Kolmogorov smirnov and Anderson darling test showed the similar results. So, GEV distribution was considered as the best fit distribution for most of the seasons. For the Rainfall pattern the non-parametric test, such as Mann Kendall test and Sen's slope estimator were employed to study monotonic trends. The trend value (Z) for kharif season gave the decreasing trend (-0.030) and for Rabi season with increasing trend (0.048) and for Summer season with increasing trend (0.187). Thus by understanding the behavior of the rainfall accordingly we can plan the crop management and planning. Identifying the distribution amount of monthly rainfall data has a wide range of applications in agriculture, hydrology, engineering design and climate research.

References

Al-suhili, P. R. H., and Khanbilvardi, P. R. (2014). Frequency Analysis of the Monthly Rainfall Data at Sulaimania, *3*(5), 212–222.

Drápela, K., and Drápelová, I. (2011). Application of Mann-Kendall test and the Sen ' s slope estimates for trend detection in deposition data from Bílý Kříž (Beskydy Mts., the Czech Republic) 1997 – 2010, *4*(2), 133–146.

Engineering, W. C., and Nadu, T. (2016). Analysis of rainfall trend using Mann – Kendall test and the Sen'S, *6*(2), 131–138.

Ghosh, S., Roy, M. K., and Biswas, S. C. (2016). Determination of the Best Fit Probability Distribution for Monthly Rainfall Data in Bangladesh, *6*(4), 170–174. https://doi.org/10.5923/j.ajms.20160604.05

Golroudbary, V. R., Zeng, Y., Mannaerts, C. M., and Su, Z. B. (2016). Attributing seasonal variation of daily extreme precipitation events across The Netherlands. *Weather and Climate Extremes*, *14*(November), 56–66. https://doi.org/10.1016/j.wace.2016.11.003

Kumar, V., and Jahangeer, S. (2017). Statistical distribution of rainfall in Uttarakhand, India. *Applied Water Science*, *7*(8), 4765–4776. https://doi.org/10.1007/s13201-017-0586-5

Kwarteng, A. Y., Dorvlo, S., and Vijaya, G. T. (2009). Analysis of a 27-year rainfall data (1977 – 2003) in the Sultanate of Oman, *29*(9), 605–617. https://doi.org/10.1002/joc

Longobardi, A., and Villani, P. (2009). Trend analysis of annual and seasonal rainfall time series in the Mediterranean area. https://doi.org/10.1002/joc

Mayooran, T., and Laheetharan, A. (2014). The Statistical Distribution of Annual Maximum Rainfall in Colombo District, *15*(2), 107–130.

Mohamed, T. M., Abd, A., and Ibrahim, A. (2016). Fitting Probability Distributions of Annual Rainfall in Sudan, *17*(2).

Murumkar, A. R., and Arya, D. S. (2014).

Trend and Periodicity Analysis in Rainfall Pattern of Nira Basin, Central India, (March), 60–70.

Obaid, N., Alghazali, S., Adnan, D., and Alawadi, H. (2014). Fitting Statistical Distributions of Monthly Rainfall for Some Iraqi Stations, *6*(6), 40–47.

Of, J., and Engineering, C. (2014). International. *International Journal of Civil Engineering and Technoogy*, *5*(3), 100–106.

Rajendran, V., and Venkatasubramani, R. (2017). Frequency analysis of rainfall deviation in Dharmapuri district in Tamil Nadu, *46*(8), 1710–1714.

Salmi, T., Maatta, A., Anttila, P., Ruoho-Airola, T., and Amnell, T. (2002). *Detecting Trends of Annual Values of Atmospheric Pollutants by the Mann-Kendall Test and Sen's Solpe Estimates the Excel Template Application MAKESENS*. *Finnish Meteorological Institute, Air Quality Research*. https://doi.org/456-789X

Sharma, M. A., and Singh, J. B. (2010). Use of Probability Distribution in Rainfall Analysis, *3*(9), 40–49.

Statistical Analysis of Rainfall Trend for Volta Region in Ghana. (n.d.).

Submitted to Vasantrao Naik Marathwada Krishi Vidyapeeth In partial fulfillment of the Requirements for the Degree of. (2017), *402*.

Taylor, P., Kumar, V., Jain, S. K., Singh, Y., Kumar, V., Jain, S. K., and Singh, Y. (2010). Analysis of long-term rainfall trends in India Analysis of long-term rainfall trends in India, *55*(4), 37–41.

Yue, S., and Hashino, M. (2007). Probability distribution of annual, seasonal and monthly precipitation in Japan Probability distribution of annual, seasonal and monthly precipitation in Japan, *52*(5).

How to cite this article:

Lavanya, S., M. Radha and Arulanandu, U. 2018. Statistical Distribution of Seasonal Rainfall Data for Rainfall Pattern in TNAU1 Station Coimbatore, Tamil Nadu, India. *Int.J.Curr.Microbiol.App.Sci.* 7(04): 3053-3062. doi: https://doi.org/10.20546/ijcmas.2018.704.346