EARTH SCIENCES
RESEARCH JOURNAL
Earth Sci. Res. J. Vol. 29, No. 3 (September, 2025): 363 - 378

UNIVERSIDAD NACIONAL DE COLOMBIA

METEOROLOGY

Fitting of Statistical Distribution in Assessing Rainfall Patterns Across Various Zones of Assam, India

Tanusree Deb Roy[1], Dipanjali Ray[1]*, Subhankar Dutta[2], and Sebul Islam Laskar[3]
1. Department of Statistics, Assam University, Silchar, Assam, India.
2. Department of Mathematics, Bioinformatics and Computer Applications, Maulana Azad National Institute of Technology Bhopal, Madhya Pradesh, India
3. India Meteorological Department (IMD), New Delhi, India
* Corresponding author: dipanjali.ray@aus.ac.in

ABSTRACT:

The maximum and minimum projected rainfall for a specific time of the year, taking into account a particular level of probabilities, makes up the probable rainfall, which is a great meteorological parameter of information. The purpose of this study is to assess the performance of probability distributions using various goodness-of-fit tests in order to develop a standard for selecting them in various zones of Assam based on monthly rainfall data from 1985-2022. Ten different distributions such as Gumbel, Weibull, Gamma, Logistic, Exponential, log-Normal, Pearson-0, Pearson-I, Pearson-III and Pearson-V distributions have been considered in the study. The maximum likelihood estimation approach was used to estimate the parameters associated with these distributions. The best fitted probability distribution is determined using a variety of goodness of fit techniques, including the Bayesian information criteria (BIC), Akaike information criterion (AIC), and Kolmogorov-Smirnov (K-S) test. Although not one distribution fits the rainfall data perfectly for every month, the Pearson-I distribution typically fits the data better than the other distributions most of the time, according to the goodness of fit tools. The data has been collected from the National Data Centre (NDC), India Meteorological Department (IMD), Pune.

Keywords: AIC; BIC; K-S test; Maximum Likelihood Estimation; Precipitation.

Ajuste de la distribución estadística para evaluar los patrones de precipitaciones en diversas zonas de Assam, India

RESUMEN

La precipitación máxima y mínima proyectada para una época específica del año, con un nivel particular de probabilidades, constituye la precipitación probable, que es un gran parámetro meteorológico de información. El propósito de este estudio es evaluar el rendimiento de las distribuciones de probabilidad utilizando varias pruebas de bondad de ajuste con el fin de desarrollar un estándar para luego aplicarlo en varias zonas de Assam con base en datos mensuales de precipitación de 1985 a 2022. En el estudio se han considerado diez distribuciones diferentes como Gumbel, Weibull, Gamma, Logística, Exponencial, log-Normal, Pearson-0, Pearson-I, Pearson-III y Pearson-V. El enfoque de estimación de máxima verosimilitud se utilizó para estimar los parámetros asociados con estas distribuciones. La distribución de probabilidad mejor ajustada se determina utilizando una variedad de técnicas de bondad de ajuste, incluyendo los criterios de información bayesianos (BIC), el criterio de información de Akaike (AIC) y la prueba de Kolmogorov-Smirnov (K-S). Aunque ninguna distribución se ajusta perfectamente a los datos de precipitación para todos los meses, la distribución Pearson-I suele ajustarse mejor a los datos que las demás distribuciones la mayor parte del tiempo, según las herramientas de bondad de ajuste. Los datos se recopilaron del Centro Nacional de Datos (NDC) del Departamento Meteorológico de la India (IMD) en Pune.

Palabras clave: AIC; BIC; prueba K-S; precipitación máxima proyectada; precipitation.

Record
Manuscript received: 30/03/2025
Accepted for publication: 15/09/2025

How to cite this item:
Deb Roy, T., Ray, D., Dutta, S., & Laskar, S. I. (2025). Fitting of Statistical Distribution in Assessing Rainfall Patterns Across Various Zones of Assam. *Earth Sciences Research Journal, 29(*3), 363-378. https://doi.org/10.15446/esrj.v29n3.119576

ISSN 1794-6190 e-ISSN 2339-3459
https://doi.org/10.15446/esrj.v29n3.119576

1 Introduction:

The identification of a probability distribution that precisely matches rainfall data has garnered interest in several disciplines, including hydrology and meteorology. In precipitation research, one of the most prominent topics is determining the most suitable distribution for rainfall data. Various forms of precipitation include drizzle, snow, graupel, sleet, and freezing drizzle, with rainfall being the largest type. By employing several probability distributions, it is feasible to predict the amount of precipitation rather precisely. When managing water resources and identifying the dry and wet seasons, this is crucial information for agricultural planners. Hydrological data sets can be fitted with a variety of probability distributions (Ozonur et al., 2021, Mandal and Choudhury, 2014).

Numerous researchers have employed the probability distribution function to examine rainfall in various dimensions. Thom (1958) has analysed the meteorological properties of the Gamma distribution, Bobee and Robitaille (1977) demonstrated that Pearson III outperforms Log Pearson III for flood data across global stations, including 11 stations in Canada, while Swift and Schreuder (1981) evaluated several distributions such as Log-normal, Gamma, Weibull, S_B (Special bound), and Beta on daily precipitation data, concluding that the S_B distribution best fits high-precipitation areas like the Southern Appalachian Mountains. Sarker et al. (1982) used the Gamma distribution for precipitation in India, Phien and Jivajirajah (1984) applied Log-Pearson III for flood analysis in Thailand, Kottegoda (1987) demonstrated Johnson SB (Special bound) for rainfall in various station of United Kingdom and South Africa, and Koutrouvelis and Canavos (1999) studied Pearson III for 100-year simulated floods data. Griffis and Stedinger (2007) studied Log-Pearson III for U.S. floods, Feng et al. (2007) modeled extreme precipitation in China using the Generalized Extreme Value distribution, and Olofintoye et al. (2009) compared several distributions, including Gumbel and Pearson, for rainfall in 20 Nigerian cities.

Deka et al. (2009) analysed the fitting of five three-parameter extreme value distributions i.e. Generalized Extreme Value, Generalized Logistic, Generalized Pareto, Log-Normal, and Pearson III using L-moments and LQ-moments for stations in North-East India maximum daily rainfall data for the period 1966 to 2007. Millington et al. (2011) compared Gumbel, Generalized Extreme Value, and Log-Pearson III distributions for the Upper Thames River, Mandal and Choudhury (2015) identified optimal rainfall distributions for Sagar Island, and Amin et al. (2016) assessed Normal, Log-Normal, Log-Pearson III, and Gumbel Max distributions for northern Pakistan. Kumar et al. (2017) tried to fit Exponential, Gamma, Weibull, Log-Pearson III, and Generalized Extreme Value distributions to rainfall data in Uttarakhand, Douka and Karacostas (2017) evaluated Generalized Pareto, Johnson S_B, Log-Gamma, and Log-Normal distributions for precipitation of Thessaloniki. Ghosh et al. (2016) assessed monthly rainfall in Bangladesh using Normal, Lognormal, Gamma, Weibull, Inverse Gaussian, and Generalized Extreme Value distributions. Mamoon and Rahman (2017) examined rainfall frequency in Qatar with fourteen distributions. Bhavyashree and Bhattacharyya (2018) analyzed monsoon rainfall across Karnataka, while Kurniawan (2019) focused on Jakarta's rainfall patterns.

Shaharudin et al. (2020) investigated various distributions for rainfall simulation, and Moccia et al. (2021) evaluated daily rainfall extremes in Lazio and Sicily, Italy, using Weibull, Gumbel, Frechet, Pareto, Gamma, and Lognormal distributions. Greece and Ozonur et al. (2021) tested eight distributions, including Weibull, Rayleigh, Gamma, Log-Normal, and Gumbel, for modelling rainfall data in central West Brazil. Mohamed and Adam (2022) identified the best-fit distribution for maximum rainfall in Somalia, while Chandran et al. (2023) assessed annual, monthly, and seasonal rainfall in Tamil Nadu. Karami et al. (2023) examined distribution models for annual and 24-hour maximum rainfall. Abreu et al. (2023) explored criteria for selecting suitable distributions in case of rainfall data, and Haseeb et al. (2025) compared nine distributions to determine the most appropriate fit for 42 years of rainfall data in Pakistan.

Rainfall varies significantly across different locations and times, making statistical distributions helpful in understanding its unpredictability. Examining rainfall data with various statistical models allows researchers to assess variability and identify normal and extreme patterns. Heavy rainfall can lead to flooding, while insufficient rain may lead to droughts. Statistical modelling helps estimate the probability of these extreme events, which is vital for designing flood barriers, building dams, and preparing for drought scenarios.

Assam remains highly vulnerable to floods and river basin–related disasters, which exert profound impacts on agriculture and livelihoods. Understanding rainfall patterns is therefore critical, as it can generate valuable insights for flood preparedness, agricultural planning, and regional resilience. However, a comprehensive distributional analysis of rainfall behaviour across Assam for the period 1985–2022 has not yet been undertaken. This study addresses this gap by examining rainfall patterns in the region, thereby contributing to improved climate risk assessment and resource management.

This study aims to enhance the understanding of rainfall variability in specific areas of Assam, India. The available climatological data will be analyzed to identify the most suitable probability distribution for rainfall in this region. In this study, ten different distributions have been taken into account to fit the rainfall data. To estimate the parameters of the chosen distributions, the maximum likelihood method has been considered. Before statistical analysis, the Kolmogorov-Smirnov (K-S) test has been considered to assess the adequacy of these distributions to the data based on the K-S distance and the associated p-value.

2 Materials and methods

2.1 Characterization and location of the study area

The North East region of India, which includes Assam is expected to be extremely vulnerable to the effects of climate change because of its delicate geo-ecological configuration, strategic location near an international border, the presence of the Eastern Himalayan ranges, transboundary river systems, the inhabitation of the ecosystem by people from various ethnic groups, and inherent socioeconomic disparities (ASTEC, 2011).

Depending on its geographical location, the whole Assam is divided into various zones. Namely, Lower Brahmaputra valley zone, Upper Brahmaputra valley zone, North Bank plain zone, Hills zone and Barak valley zone. Five stations have been selected from the five zones except hills zone on the basis of available data provided by meteorological department for the period of 1985-2022. The name of the stations are given below.

Table 1 represents the names of rainfall monitoring stations across various zones of Assam, where the best probability distribution function will be determined from a selection of ten probability distributions.

The geological location of the above stations are represented below

Table 1. Name of stations in respective zones

Name of Zones	Name of stations
Lower Brahmaputra Valley Zone	Guwahati, Dhubri
North Bank Plain Zone	Tezpur
Upper Brahmaputra Valley Zone	Dibrugarh
Barak Valley Zone	Silchar

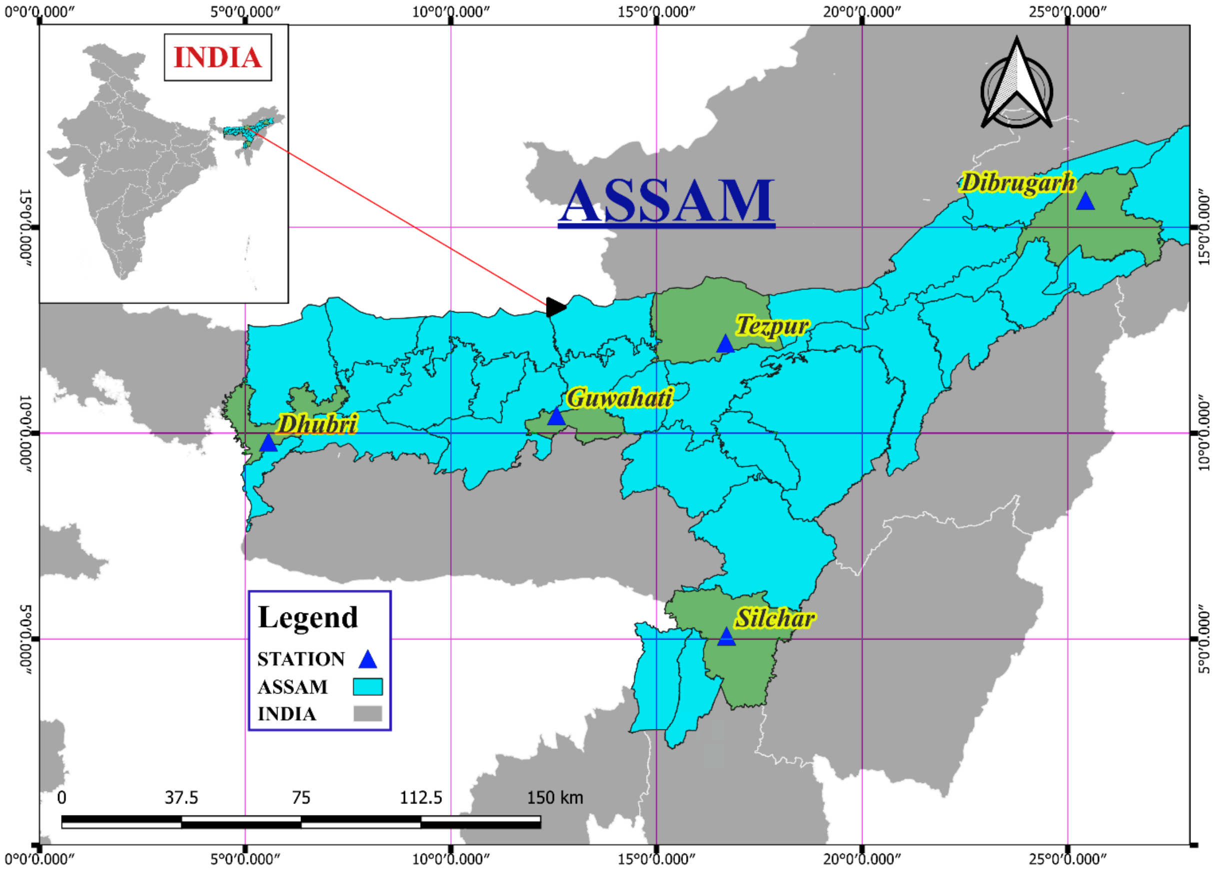

Figure 1. Study area in Assam

Figure 1 aims to represent the inferred geological direction of all selected stations spanning diverse zones of Assam created using QGIS 3.32.0 software. Assam's climate is mostly humid subtropical, with warm, muggy summers, intense monsoons, and moderate winters. The range of winter temperatures is 10^0 to 22^0 degrees Celsius. The summertime temperature ranges from 30^0 to 36^0 degrees Celsius. The months of July and August receive the most rainfall in Assam (29 percent of the total rainfall during the South West Monsoon), while September receives 24 percent of the total rainfall. The remaining 66% of the yearly rainfall falls during the Southwest Monsoon (June – September). The state Assam lies between $89^0 46^{/}$ - $96^0 01^{/}$E longitude and $24^0 03^{/}$- $27^0 58^{/}$ N latitude and covers an area of 78,438 km² (Guhathakurta et al., 2020)The longitude and latitude of remain stations are given below

Table 2 represents all the longitudes and latitudes of selected stations for this study.

Table 2. Longitude and latitude of selected stations from various zones of Assam

Name of stations	Longitude	Latitude
Guwahati	91° 44′ 45″ E	26° 10′ 20″ N
Dhubri	~89° 58′ 44″ E	~26° 1′ 20″ N
Tezpur	92° 47′ 59″ E	26° 37′ 59″ N
Dibrugarh	~94° 54′ 43″ E	~27° 28′ 22″ N
Silchar	~92° 47′ 52″ E	~24° 49′ 38″ N

2.2 Data

The data is collected from National Data Centre (NDC), India Meteorological Department (IMD), Pune (http://dsp.imdpune.gov.in/) for the period of 1985-2022. For the analysis, data is arranged monthly and also season wise, namely Winter (January, February), Pre-Monsoon (March, April, May), Monsoon (June, July, August, September) and Post-Monsoon (October, November and December).

2.3 Probability distribution

In this experiment, ten different probability distributions are considered to study the monthly rainfall data in five stations of various agro-climatic zones of Assam. Gumbel distribution (two-parameter) (Gumbel, 1941), Weibull distribution (two-parameter) (Weibull, 1939), Gamma distribution (two-parameter) (Bobee & Ashkar, 1991), Logistic distribution (two-parameter) (Jhonson et al., 1995), Exponential distribution (one-parameter) (Ferguson, 1964), log-Normal distribution (two-parameter) (Gaddum,1945), Pearson-0 distribution (two-parameter), Pearson-I distribution (four-parameter), Pearson-III distribution (three-parameter) and Pearson-V distribution (three-parameter) (Pearson, 1895) have been considered in the study. Suppose X is a random variable with shape parameter α, a, b, scale parameter β and location parameter μ then the probability density function (pdf) and their corresponding cumulative distribution function (cdf) are depicted in Table 3.

Table 3. Name of distributions with its pdf, cdf, domain and parameters

Distributions	pdf	cdf	Domain	Parameter
Exponential	$\alpha e^{-\alpha x}$	$1 - \alpha e^{-\alpha x}$	$x \in (0,\infty)$	$\alpha > 0$
Gumbel	$\frac{1}{\beta} e^{\left(-\frac{x-\mu}{\beta} - e^{-\frac{x-\mu}{\beta}}\right)}$	$e^{\left(-e^{-\frac{x-\mu}{\beta}}\right)}$	$x \in \mathbb{R}$	$\mu > 0$, $\beta > 0$
Weibull	$\frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{\left(-\frac{x}{\beta}\right)^{\alpha}}$	$1 - e^{\left(-\frac{x}{\beta}\right)^{\alpha}}$	$x \in (0,\infty)$	$\alpha > 0$, $\beta > 0$
Gamma	$\frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^{\alpha} \lceil \alpha}$	$\frac{1}{\lceil \alpha} \gamma\left(\alpha, \frac{x}{\beta}\right)$	$x \in (0,\infty)$	$\alpha > 0$, $\beta > 0$
Logistic	$\frac{e^{-\frac{x-\mu}{\beta}}}{s\left(1 + e^{-\frac{x-\mu}{\beta}}\right)^2}$	$\frac{1}{1 + e^{-\frac{x-\mu}{\beta}}}$	$x \in (-\infty,\infty)$	$\mu > 0$, $\beta > 0$
Log-Normal	$\frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(lnx-\mu)^2}{2\sigma^2}}$	$\frac{1}{2} + \frac{1}{2} erf\left(\frac{lnx - \mu}{\sigma\sqrt{2\pi}}\right)$ Where erf is the error function	$x \in (0,\infty)$	$\mu \in (-\infty,\infty)$, $\sigma > 0$
Pearson 0	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$	$x \in (-\infty,\infty)$	$\mu > 0$, $\sigma > 0$
Pearson I	$\frac{1}{\lvert\beta\rvert} \frac{\lceil(a+b)}{\lceil a \lceil b} \left(\frac{x-\mu}{\beta}\right)^{a-1} - \frac{x-\mu}{\beta}\Big)^{b-1}$	$\frac{1}{\lvert\beta\rvert} \frac{\lceil(a+b)}{\lceil a \lceil b} \int_{a}^{x} \left(\frac{x-\mu}{\beta}\right)^{a-1} \left(1 - \frac{x-\mu}{\beta}\right)^{b-1} dx$	$x \in [a,b]$ Where a &b are any real number	$a > 0$, $b > 0$ $\beta \neq 0$, $0 < \frac{x-\mu}{\beta} < 1$
Pearson III	$\frac{(x-\mu)^{a-1}}{\beta^{a} \lceil a} e^{-\frac{x-\mu}{\beta}}$	$\frac{1}{\beta^{a} \lceil a} \int_{\mu}^{x} (x-\mu)^{a-1} e^{-\frac{x-\mu}{\beta}} dx$	$x \in [\mu,\infty]$	$a > 0$, $\beta > 0$ $\frac{x-\mu}{\beta} > 0$
Pearson V	$\frac{\lvert\alpha\rvert^{a}}{\lceil a} \lvert x-\mu\rvert^{-a-1} e^{-\frac{\alpha}{x-\mu}}$	$\frac{\lvert\alpha\rvert^{a}}{\lceil a} \int_{0}^{x} \lvert x-\mu\rvert^{-a-1} e^{-\frac{\alpha}{x-\mu}} dx$	$x \in (0,\infty)$	$a > 0$, $\alpha,\beta > 0$ $\frac{\alpha}{x-\mu} > 0$

2.4 Maximum likelihood method

The parameter estimates of the distributions may be estimated using a variety of techniques. The maximum likelihood approach is the most used estimating technique. This approach is used in this work to estimate the parameters of any probability distribution that is taken into consideration. Due to its good asymptotic features, this approach is a widely used and recommended statistical estimating technique. It is widely acknowledged that it generates parameter estimates that are both statistically consistent and effective.

If we consider a sample $\{X_1, X_2, \ldots, X_n\}$ following distributions with PDF $f(x;\theta)$, where θ represents the set of parameters, then the log-likelihood function can be expressed as

$$L(x_1, x_2, \ldots, x_n;\ \theta) = \sum_{i=1}^{n} \log f(x_i;\theta).$$

By maximizing the log L, , the maximum likelihood estimates (MLEs) can be obtained. To obtain the MLEs, we have to solve $\frac{\partial \log L}{\partial \theta} = 0$, with respect to set of parameters. Sometimes these equations cannot be solved explicitly, then Newton-Raphson method is used (Gupta & Kapoor, 1997; Feng et al., 2007; Ozonur et al., 2021).

2.5 Goodness of fit tests and model selection criteria

To assess the appropriateness of the chosen probability distributions to the rainfall data, Kolmogorov-Smirnov (K-S) test has been considered initially. Additionally, Akaike's information criterion (AIC) and Bayesian information criterion (BIC) have been considered as the model selection criteria. Using K-S test, one can determine whether the data follows the considered distribution or not. In K-S test, the null hypothesis has been considered as the data follows the selected distribution whereas the alternative hypothesis yields that the data does not follow the selected distribution. K-S test statistic defines the maximum difference between sample and theoretical CDF, and it can be expressed as

$$D_n = \max_{1 \le k \le n} \left\{ \left| \frac{k}{n} - \hat{F}(x_{(k)}) \right|, \left| \hat{F}(x_{(k)}) - \frac{k-1}{n} \right| \right\},$$

Where $\hat{F}x(_{(k)})$ is the estimated value of the CDF, and $\{x_{(1)},\ldots,x_{(n)}\}$ are ascending ordered observations. If the test statistics $D_n > D_n(\alpha)$ then the null hypothesis will be rejected at α significance level, where $D_n(\alpha)$ represents the critical value of the K-S statistic (An & Cheng, 1996; Sharma and Singh, 2010; Mandal and Choudhury, 2014; Ozonur et al., 2021; Mamoon & Rahman, 2017; Moccia et al. 2021).

Further, the p-values of the goodness of tests are often used to assess the appropriateness of the above-mentioned distributions. By using the p-value technique, the null hypothesis is rejected if the p-value is less than the selected significance threshold. The null hypothesis is not rejected if the p-value is bigger than the significance threshold, and there is insufficient evidence to draw the conclusion that the data do not follow the given distribution. If it is determined that both of the distributions under comparison may be appropriate, then the distribution with the greater p-value fits the data more closely.

The effectiveness of the investigated pdfs for modelling the rainfall data is determined using the model selection criteria AIC and BIC after the appropriateness of the pdfs is determined based on the findings of the goodness of fit tests. The AIC and BIC can be obtained by using the following formulas:

$$AIC = -2\log L + 2m,$$

$$BIC = -2\log L + m\log n\ ,$$

where log L yields the value of the estimated log-likelihood function at the MLEs of the selected model, m is the number of parameters of the selected distribution, and n is the number of observed data (Akaike, 1969; Schwarz, 1978; Ozonur et al., 2021; Mohamad & Adam, 2021). These are analysed in software R version 4.1.1.

3. Result and discussion

In this section, the best fit distribution is elaborately discussed on total rainfall (mm) monthly and season wise for Guwahati, Silchar, Tezpur, Dibrugarh, Dhubri station and are given below

The above Table 4 describes the frame of the data in both monthly and season wise by introducing length of the data, mean, maximum (Max), minimum (Min), median, standard deviation (SD), coefficient of variation (CV), coefficient of skewness (CS) and coefficient of kurtosis (CK). Out of the twelve months, on April month variation is highest as SD is 122.444 and in season wise on Pre-Monsoon season shows highest variation (SD=124.134).

Similarly, the descriptive statistics have been calculated for the remaining station from which it has been observed that April (194.319) month and Pre-Monsoon (201.759) show highest SD in Silchar station. The month of June is showing highest SD for the remaining stations i.e. Tezpur (144.872), Dibrugarh (165.55) and Dhubri (281.681). Season wise, Monsoon shows highest SD for the remaining station Tezpur (118.590), Dibrugarh (149.852) and Dhubri (233.280).

Table 4. Descriptive statistics of rainfall in monthly and season wise data of Guwahati station for the period of 1985-2022

Month	Length	Mean	Min	Max	Median	SD	CV	CS	CK
January	38	11.387	0	78.100	7.900	15.104	132.648	2.416	7.599
February	38	19.840	0	96.100	14.700	19.812	99.850	1.707	3.744
March	38	52.180	3.800	152.100	42.450	40.544	77.703	0.891	-0.168
April	38	179.200	24.000	520.900	160.700	102.342	57.121	0.889	1.431
May	38	246.200	88.500	569.700	222.000	122.444	49.729	0.777	-0.120
June	38	301.600	104.100	573.000	307.200	117.190	38.853	0.262	-0.825
July	38	301.900	131.600	638.800	285.900	118.365	39.207	0.880	0.397
August	38	239.100	67.100	569.200	221.100	111.544	46.660	0.550	0.080
September	38	190.300	28.200	372.000	183.100	86.267	45.342	0.353	-0.786
October	38	115.540	7.200	354.400	117.750	81.680	70.691	0.868	0.618
November	38	11.900	0	88.100	4.450	18.381	154.463	2.367	6.005
December	38	4.989	0	29.000	0.600	8.014	160.626	1.554	1.155
Winter	76	15.610	0	96.100	11.000	18.006	115.332	2.023	5.272
Pre-Mon	114	159.189	3.800	569.700	135.100	124.134	77.978	1.008	0.755
Monsoon	152	258.200	28.200	638.800	242.100	117.810	45.626	0.623	0.177
Post-Mon	114	44.140	0	354.400	8.650	69.969	158.501	2.054	4.315

Table 5. Estimates of probability distribution fitting of total monthly rainfall in Guwahati station for January, February and March month

Estimates					
Months	**Distribution**	**Shape1**	**Shape2**	**Location**	**Scale**
January	**Gamma**	0.2525	-	-	0.0221
	Logistics	-	-	9.1336	6.9780
	Exponential	0.0878	-	-	-
	Gumbel	-	-	5.7015	8.4914
	Weibull	0.3485	-	-	5.0727
	Log-Normal	-	-	0.3758	4.7371
	Pearson 0	-	-	11.3868	14.9043
	Pearson I	0.4280	8.4427	0.0001	223.9029
	Pearson III	-	0.6068	0.0001	17.3831
	Pearson V	-	2.09E-01	1.00E-04	1.95E-14
February	**Gamma**	0.4177	-	-	0.0210
	Logistic	-	-	17.2762	9.9634
	Exponential	0.0504	-	-	-
	Gumbel	-	-	11.7239	12.7800
	Weibull	0.5721	-	-	15.6152
	Log-Normal	-	-	1.4234	3.7790
	Pearson 0	-	-	19.8421	19.5499
	Pearson I	0.7961	13.3457	0.0001	320.231
	Pearson III	-	0.9681	0.0001	16.1597
	Pearson V	-	3.1893	-10.2752	68.1229
March	**Gamma**	1.5351	-	-	0.0294
	Logistic	-	-	47.3156	22.5857
	Exponential	0.0191	-	-	-
	Gumbel	-	-	34.2173	28.9744
	Weibull	1.3049	-	-	56.5877
	Log-Normal	-	-	3.5949	0.9391
	Pearson 0	-	-	52.1789	40.0078
	Pearson I	0.7492	1.918	3.8	173.9859
	Pearson III	-	0.9922	3.8	47.0663
	Pearson V	-	36.4221	-166.352	7729.309

Table 5 represents all the estimates of parameters i.e. shape parameter α,a ,b, , scale parameter β and location parameter μ for January, February and March month considering rainfall data for the period of 1985-2022 in Guwahati station.

Table 6 represents probability distribution fitting of monthly total rainfall in Guwahati station monthly for January, February and March month. It elaborates parameters for each distribution and comparison criteria's i.e. -2logL (log likelihood), AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), K-S value and p-value. Fitting of best probability distribution on a data set may be characterised as finding the distribution with the lowest values of -2logL, AIC, BIC, and K-S values. Here p-value plays very significant role, since by having greater value than 0.05 represents the data follows the required distribution.

The findings of the K-S test indicate that the rainfall data for January at the Guwahati station does not follow the Gamma, Exponential, Log-Normal, or Pearson 0 and V distributions, as their p-values are less than 0.05. In contrast, the data does conform to the Logistic, Weibull, Pearson Type I, and Pearson Type III distributions since their p-values are greater than 0.05. Among the followed distribution by January month, Pearson I distribution has least -2logL (14.5150), AIC (22.1151), BIC (29.0654), K-S value (0.2105) and p-value is 0.0589, which is greater than 0.05 which indicates that the data points satisfy the null hypothesis i.e. the data points of January month follow Pearson I distribution. Likewise, interpretation can be elaborated for remaining months also.

Table 6. Criteria for comparison of probability distribution fitting of total monthly rainfall in Guwahati station for January, February and March

Criteria for Comparison						
Months	**Distribution**	**-2logL**	**AIC**	**BIC**	**KS-value**	**p-value**
January	**Gamma**	168.0214	172.0214	175.2966	0.2213	0.0482
	Logistic	302.6070	306.607	309.8821	0.2126	0.0643
	Exponential	260.8668	262.8668	264.5044	0.2479	0.0187
	Gumbel	289.4652	293.4652	296.7404	0.1681	0.2331
	Weibull	180.5210	184.521	187.7962	0.217	0.0557
	Log-Normal	197.4897	201.4897	204.7649	0.2482	0.0185
	Pearson 0	313.1651	317.1651	320.4403	0.2224	0.0392
	Pearson I	14.5150	22.1151	29.0654	0.2105	0.0589
	Pearson III	70.9679	76.9679	81.8807	0.2105	0.0589
	Pearson V	164.1624	170.1624	175.0751	0.7874	2.22E-16
February	**Gamma**	274.4866	278.4867	281.7618	0.2197	0.051
	Logistic	328.2530	332.2529	335.5281	0.1500	0.3589
	Exponential	303.0732	305.0733	306.7109	0.1052	0.7937
	Gumbel	317.9188	321.9187	325.1939	0.1173	0.6723
	Weibull	284.239	288.2391	291.5143	0.2051	0.0815
	Log-Normal	317.0596	321.0596	324.3348	0.3241	0.0006
	Pearson 0	333.7851	337.7851	341.0603	0.1557	0.2894
	Pearson I	255.954	263.954	270.5044	0.1200	0.6014
	Pearson III	297.6733	303.6733	308.586	0.1567	0.2775
	Pearson V	312.9186	318.9186	323.8314	0.1008	0.7974
March	**Gamma**	372.8106	376.8107	380.0859	0.0763	0.9676
	Logistic	388.5022	392.5022	395.7774	0.1271	0.5295
	Exponential	376.5556	378.5556	380.1932	0.1196	0.6061
	Gumbel	378.9424	382.9424	386.2176	0.0952	0.8484
	Weibull	372.7302	376.7303	380.0055	0.0723	0.9803
	Log-Normal	376.2848	380.2848	383.5600	0.0997	0.8085
	Pearson 0	388.2091	392.2091	395.4843	0.1630	0.2372
	Pearson I	354.6303	362.6303	369.1807	0.0914	0.8794
	Pearson III	370.4733	376.4733	381.3861	0.1124	0.6805
	Pearson V	381.5386	387.5386	392.4513	0.11345	0.6704

The criteria for fitting probability distribution functions, along with their best parameter estimates, have been thoroughly discussed above for each month, supported by statistical values to demonstrate their significance. The graphical representation of these distributions for any given month or season can be effectively illustrated using a histogram of the raw data from the corresponding period, overlaid with line plots for each fitted distribution. In line with this approach, the rainfall data for January of Guwahati station can be depicted graphically below, and this procedure can similarly be applied to the remaining months and seasons for the station to provide a comprehensive visual analysis.

Figure 2 represents the histogram with multiple line diagrams of ten selected probability distribution functions on raw rainfall data of January month in Guwahati station just to check the suitability of these selected distributions on rainfall data. It is distinctly observed that these distributions effectively characterize the rainfall data for January month of Guwahati station, and this approach can be systematically extended to encompass the remaining months and the rainfall records of other stations also.

Figure 3 represents a Q-Q plot, which shows how the quantiles of two distributions relate, indicating if the data follow a specific theoretical distribution or share the same distribution. Gumbel (orange) fits reasonably well in the middle but overestimates at higher quantiles, with points above the line. Logistic (blue) outperforms exponential and gamma, especially in the mid to upper range, though some deviations remain at the extremes. Weibull (brown) stays close to the line in lower to mid quantiles but diverges upward at higher values. P1/Pearson I (orange) aligns better than Log-Normal (LGN) and Pearson 0 (P0), but still deviates in mid-probabilities. P3/Pearson III (blue) remains closer to the line across most of the range, especially mid to upper probabilities, indicating a better fit. These visual representations can also be summarised by test statistics. This visual interpretation is nearly the same as the interpretation of Table 6, which shows various criteria for the comparisons of the selected distribution.

Figure 4 represents a P-P plot which compares the empirical cumulative distribution function of January data to a theoretical distribution's cumulative function. Its interpretation is similar to that of a Q-Q plot. The logistic distribution (blue) follows the reference line more closely, especially in the central part, but shows divergence in the tails. The Weibull distribution (brown) stays near the diagonal over a wide range, although it doesn't fit perfectly in the lower tail. P3/Pearson III (blue) remains closest to the diagonal throughout most of the range and provides the best fit overall among the second set of options.

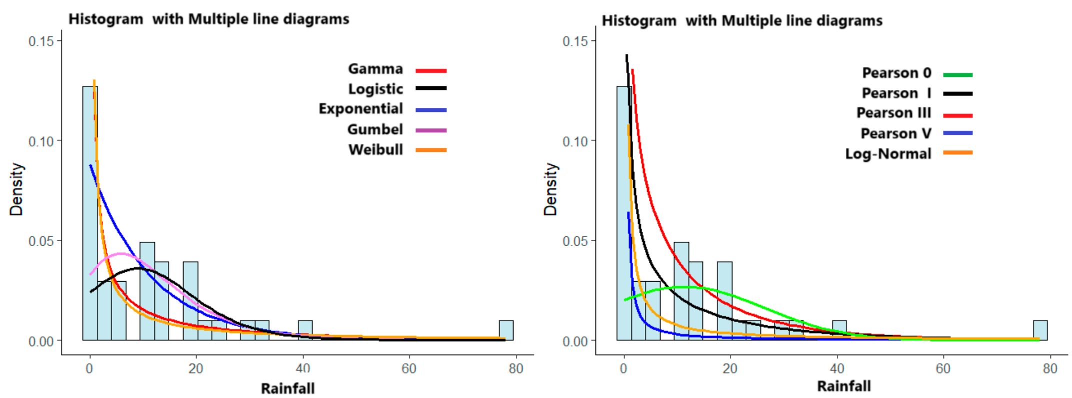

Figure 2. Histogram with multiple line diagrams of 10 selected distributions of January month of Guwahati station for the period of 1985-2022 on Rainfall

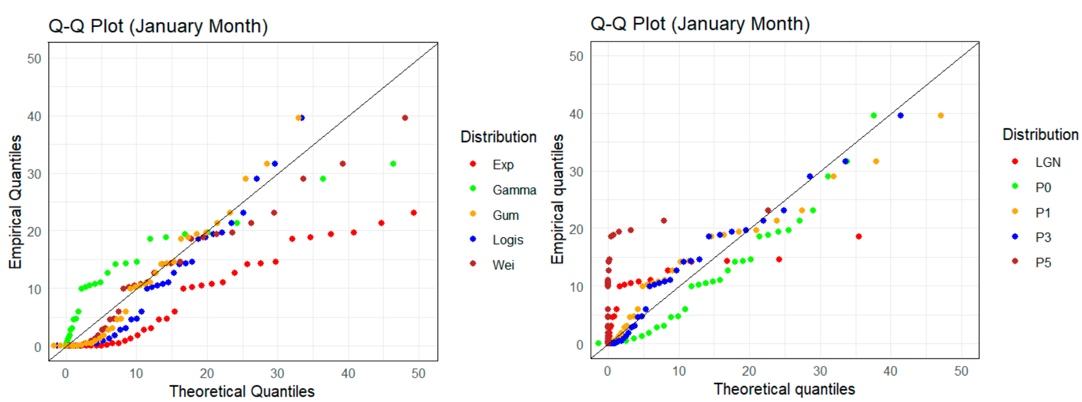

Figure 3. Q-Q Plot of January month at Guwahati Station

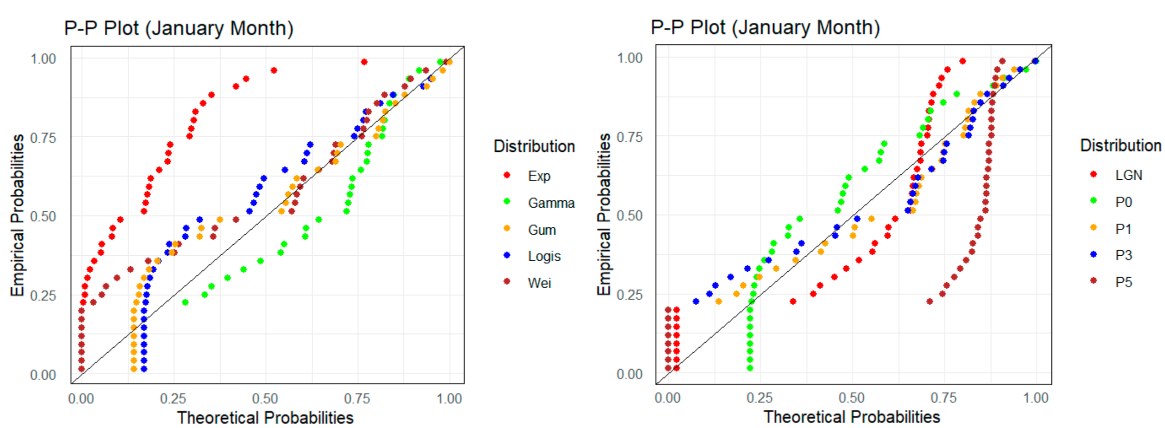

Figure 4. P-P Plot of January month at Guwahati Station

Table 7. Criteria for comparison of best probability distribution fitting of total monthly rainfall in Guwahati station for April, May, June, July, August, September, November and December

Station: Guwahati						
Months	**Distribution**	**-2logL**	**AIC**	**BIC**	**KS-value**	**p-value**
April	**Pearson I**	442.3452	450.3452	456.8956	0.2426	0.1870
May	**Pearson I**	454.6502	462.6502	469.2005	0.1205	0.5967
June	**Pearson I**	463.8386	471.8386	478.389	0.0786	0.9580
July	**Log-Normal**	462.3650	466.3650	469.6402	0.0596	0.9993
August	**Weibull**	461.8864	465.8865	469.1616	0.0726	0.9881
September	**Weibull**	443.2952	447.2953	450.5704	0.0639	0.9977
October	**Weibull**	431.2786	435.2786	438.5538	0.1565	0.3098
November	**Gamma**	178.507	182.507	185.7822	0.1778	0.1808
December	**-**	-	-	-	-	-

Table 8. Criteria for comparison of probability distribution fitting of total monthly rainfall in Guwahati station for Winter, Pre-Monsoon, Monsoon and Post-Monsoon season

Season wise (Guwahati Station)						
Months	**Distribution**	**-2logL**	**AIC**	**BIC**	**KS-value**	**p-value**
Winter	**Gamma**	448.0652	452.0652	456.7266	0.2239	0.0009
	Logistic	639.2842	643.2842	647.9457	0.1863	0.0102
	Exponential	569.7262	571.7261	574.0569	0.1737	0.0203
	Gumbel	615.1228	619.1229	623.7843	0.1373	0.1135
	Weibull	472.7544	476.7545	481.416	0.2188	0.0013
	Log-Normal	519.7514	523.7513	528.4128	0.2828	1.05E-05
	Pearson 0	654.0812	658.0812	662.7426	0.1913	0.0065
	Pearson I	267.4619	275.4619	284.7848	0.1578	0.0402
	Pearson III	374.2139	380.2139	387.2061	0.1627	0.0315
	Pearson V	593.8993	599.8993	606.8915	0.14447	0.0747
Pre-Mon	**Pearson I**	1371.428	1379.428	1390.372	0.0751	0.5159
Monsoon	**Pearson I**	1865.889	1873.889	1885.984	0.0469	0.8755
Post-Mon	**-**	-	-	-	-	-

From Table 6 and 7, the data points of February, March, April, May, June month fit best the Pearson I distribution with least -2logL, AIC, BIC and K-S value with greater 0.05 p-value, but in July month, both Pearson I and Log-normal distribution fits well, but it is more favored to choose Log-normal distribution since K-S value is lowest in case of Log-normal distribution than Pearson I distribution. Similarly, August, September and October month follow Weibull distribution. November month follows Gamma distribution and December month doesn't follow any distribution i.e. the p-value for each distribution satisfies the alternative hypothesis of K-S test. It means none of 10 distributions follow the required distribution i.e. all are having p-values less than 0.05. The criteria for determining the best-fitting probability distribution season wise are presented in Table 8.

Table 8 also briefly explains the quantitative values of criteria's of fitting best distribution in season wise i.e. Winter (January, February), Pre-Monsoon (March, April, May), Monsoon (June, July, August, September) and Post-Monsoon season (October, November, December) respectively. It is observed that the rainfall data of January and February month i.e. Winter season follows Gumbel and Pearson V distribution with greater than 0.05 p-value, but depending on K-S distance or value, Gumbel fits better with least k-S value. Similarly, Pre-Monsoon follows Pearson I, Monsoon follows Pearson I distribution and Post-Monsoon does not satisfy the null hypothesis of K-S test.

Table 9, 10, 11 and 12 represents selected criteria's of fitting of distribution with least quantitative value for each month and seasons for Silchar, Tezpur, Dibrugarh and Dhubri respectively.

From Table 5 to 12, it is observed that among 10 distributions, 34.48%, 22.48% and 12.06% of rainfall data from 1985-2022, follows Pearson type 1, Weibull and Gumbel distribution respectively considering net stations in monthly wise. Similarly, 50%, 18.75% and 12.5% follows Pearson type 1, Weibull and Gumbel distribution in season wise also.

This analysis provides a refined understanding of probability distribution functions for monthly and seasonal rainfall, essential for shaping advanced agricultural strategies. By employing these models, farmers can accurately predict rainfall patterns, allowing precise adjustments in crop selection, planting schedules, and resource management. This precision enhances operational efficiency while promoting sustainable practices, ensuring resilience and long-term productivity in an increasingly unpredictable climate.

Table 9. Criteria for comparison of best probability distribution fitting of total monthly rainfall in Silchar station

Station: Silchar						
Criteria for comparison						
Months	**Distribution**	**-2logL**	**AIC**	**BIC**	**KS-value**	**p-value**
January	**Gumbel**	272.3906	276.3906	279.6658	0.2059	0.0795
February	**Pearson I**	279.4234	287.4234	293.9738	0.1315	0.4859
March	**Weibull**	451.964	455.964	459.2392	0.0805	0.9495
April	**Pearson I**	487.2584	495.2584	501.8087	0.1787	0.1558
May	**Weibull**	506.2132	510.2132	513.4883	0.0797	0.9534
June	**Pearson II**	501.6652	509.5792	516.1295	0.0798	0.9530
July	**Pearson V**	486.2718	492.2718	497.1846	0.1070	0.7369
August	**Logistic**	491.0006	495.0006	498.2758	0.0917	0.9061
September	**Pearson III**	485.5180	491.5180	496.4307	0.1172	0.6308
October	**Gumbel**	453.9594	457.9594	461.2346	0.1388	0.4564
November	**Gumbel**	349.366	353.3666	356.6412	0.2133	0.0628
December	**Weibull**	408.8964	412.8965	416.1716	0.1858	0.1447
Season wise						
Winter	**Pearson I**	280.6986	288.6986	295.3529	0.1282	0.5027
Pre-Mon	**Pearson I**	294.8297	302.8297	309.5852	0.1200	0.5192
Monsoon	**Logistic**	1995.745	1999.745	2005.793	0.0424	0.9467
Post-Mon	-	-	-	-	-	-

Table 10. Criteria for comparison of best probability distribution fitting of total monthly rainfall in Tezpur station

Station: Tezpur						
Months	**Distribution**	**-2logL**	**AIC**	**BIC**	**KS-value**	**p-value**
January	**Pearson I**	74.2633	82.2633	88.8137	0.1861	0.1263
February	**Pearson I**	293.1259	301.1259	307.6762	0.0677	0.9901
March	**Pearson I**	332.5493	340.5493	347.0996	0.0943	0.8563
April	**Gamma**	411.9816	445.9816	449.2568	0.0917	0.8775
May	**Weibull**	446.7120	450.7119	453.9871	0.0962	0.873
June	**Pearson 0**	484.9906	488.9906	492.2658	0.0835	0.9334
July	**Gamma**	454.6172	458.6172	461.8924	0.0989	0.8511
August	**Pearson I**	452.3756	460.3756	466.9259	0.1027	0.7793
September	**Pearson I**	444.3329	452.3329	458.1402	0.1215	0.6280
October	**Pearson I**	396.8594	404.8594	411.4098	0.0813	0.9453
November	**Pearson I**	94.5341	102.5341	108.8682	0.218	0.0555
December	**Pearson I**	46.8992	54.8992	61.3428	0.1351	0.4684
Season wise						
Winter	**Pearson I**	399.0063	407.0063	416.3293	0.1032	0.3674
Pre-Mon	**Pearson I**	1352.836	1360.836	1371.78	0.0678	0.6446
Monsoon	**Gumbel**	1879.527	1833.527	1889.575	0.0531	0.7835
Post-Mon	**Pearson I**	588.7287	596.7287	607.5668	0.1084	0.1387

Table 11. Criteria for comparison of best probability distribution fitting of total monthly rainfall in Dibrugarh station

Station: Dibrugarh						
Months	**Distribution**	**-2logL**	**AIC**	**BIC**	**KS-value**	**p-value**
January	**Weibull**	314.4238	318.4238	321.6989	0.1380	0.4637
February	**Weibull**	385.9224	389.9224	393.1976	0.0857	0.9427
March	**Pearson I**	381.6493	389.6493	396.1997	0.1235	0.5656
April	**Weibull**	314.4238	318.4238	321.6989	0.138	0.4637
May	**Pearson III**	481.0075	487.0075	491.9202	0.1004	0.8015
June	**Logistic**	493.8898	497.8899	501.1651	0.0856	0.9205
July	**Pearson I**	466.1879	474.1879	480.7382	0.0748	0.9726
August	**Logistic**	487.8244	491.8243	495.0995	0.1525	0.3394
September	**Pearson I**	463.8983	471.8983	478.4487	0.1112	0.6937
October	**Pearson I**	411.8757	419.8757	426.426	0.1434	0.3784
November	**Gumbel**	372.4454	376.4454	379.7205	0.1657	0.2477
December	**Weibull**	190.1	194.1	197.3219	0.1949	0.1200
Season wise						
Winter	**Weibull**	722.0576	726.0576	730.719	0.0617	0.9341
Pre-Mon	**Gamma**	1412.07	1416.0700	1421.542	0.0532	0.9341
Monsoon	**Pearson III**	1953.289	1957.337	1963.337	0.0373	0.9786
Post-Mon	**Weibull**	992.6616	996.6616	1002.081	0.1123	0.1214

Table 12. Criteria for comparison of best probability distribution fitting of total monthly rainfall in Dhubri station

Station: Dhubri						
Months	**Distribution**	**-2logL**	**AIC**	**BIC**	**KS-value**	**p-value**
January	-	-	-	-	-	-
February	**Gamma**	112.5378	316.5378	319.813	0.2032	0.0864
March	**Gumbel**	401.4268	405.4268	408.7019	0.1792	0.1740
April	**Gumbel**	460.5102	464.5102	467.7853	0.1120	0.7264
May	**Logistic**	477.3674	481.3674	484.6426	0.1262	0.5802
June	**Weibull**	503.5918	507.5918	510.867	0.0749	0.9225
July	**Pearson V**	486.2718	492.2718	497.1846	0.10702	0.7369
August	**Gumbel**	494.8362	498.8362	502.143	0.0944	0.8872
September	**Gamma**	482.6098	486.6098	489.8849	0.1099	0.7477
October	**Weibull**	454.9794	458.9793	462.9793	0.1485	0.3715
November	**Pearson I**	124.111	132.111	138.6614	0.1323	0.4784
December	-	-	-	-	-	-
Season wise						
Winter	-	-	-	-	-	-
Pre-Mon	**Gumbel**	1452.097	1461.569	1461.569	0.0779	0.4926
Monsoon	**Weibull**	1501.223	1505.223	1510.696	0.0536	0.8979
	Pearson III	1499.329	1505.329	1513.537	0.0556	0.8519
Post-Mon	-	-	-	-	-	-

4. Conclusion

In this study ten probability distribution functions were considered to fit the monthly and season wise rainfall data of five zones of Assam.

From the analysis, the results reveal the statistical frame of data set for each station. Considering all the months and seasons of five stations, the month of June for Dhubri station shows the highest SD (281.681) and Monsoon (233.280) season of Dhubri gives peak of SD which indicates there is a higher deviation in rainfall data in the month of June and in Monsoon (June, July, August, September) season amongst the five stations. The criteria of best fitting of selected distributions are elaborately explained along with the estimation of parameter. The Weibull distribution fits best in the month of January and Pearson V fits best in Winter season considering ten distributions with least -2logL, AIC, BIC, K-S value with a p-value greater than 0.05 in Guwahati station. Likewise, the best fitting distribution along with its criteria of fitting and parameter estimation are explained monthly and season wise for each station.

By taking into consideration all the stations in case of best fitting distribution, Pearson type 1, Weibull and Gumbel distribution covers 34.48 percent, 22.48 percent and 12.06 percent respectively for the twelve months. For season wise, the percentage area of Pearson type 1 is 50 percent, 18.75 percent for Weibull and Gumbel distribution covers 12.5 percent. This paper elaborates the fitting of distribution on rainfall data in various zones of Assam which indicates there is also a possibility to get more precise form of distribution by proposing new generalized form of distribution to fit rainfall data in future. Accurate rainfall predictions are vital for estimating river discharge and reservoir inflows, enabling early flood warnings. Short-term forecasts are especially important for managing stormwater drainage. For example, in areas like Guwahati and Silchar, Assam farmers rely heavily on monsoon rainfall. Seasonal weather forecasts guide their decisions on sowing, transplanting paddy, and selecting short-duration crops. Understanding the probability of heavy rain allows farmers to adopt preventive measures against waterlogging, crop damage, pests, and outbreaks diseases etc.

Acknowledgements

The authors would like to thank the India Meteorological Department (IMD), Pune for supplying the required study's data. Authors would like to thank the editor and the referees for their support and assistance to improve the manuscript.

Author Contributions

All authors contributed to the study conception and design. Material preparation was performed by Tanusree Deb Roy and Subhankar Dutta, data collection was done by Sebul Islam Laskar and analysis was performed by Dipanjali Ray. The first draft of the manuscript was written by Dipanjali Ray and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

The authors declare that no funds were received during the preparation of this manuscript.

Data Availability

The data has been collected from National Data Centre (NDC), India Meteorological Department (IMD), Pune (http://dsp.imdpune.gov.in/).

References:

Akaike, H. (1969), Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, 21, 243-247.

Abreu, C. M., De Souza, A, Lyra, B. G., De Oliveira-Junior, J. F., Pobocikova, I., De Almeida, T. L., De Souza Fraga, M., Aristone, F., & Cecilio, R. A. (2023). Assessment and characterization of the monthly probabilities of rainfall in Midwest Brazil using different goodness-of-fit tests as probability density functions selection criteria. *Theoretical and Applied Climatology,* 151, 491-513. https://doi.org/10.1007/s00704-022-04286-z

ASTEC (2011). Assam's strategy and action plan on climate change—Recommendations—First draft—ASTEC (2011). India Water Portal Hindi. https:// https://www.indiawaterportal.org/articles/assams-strategy-and-action-plan-climate-change-recommendations-first-draft-astec-2011

Amin, M., Rizwan, M., & Alazba, P. (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

Bobee, B., & Ashkar, F. (1991). *The Gamma Family and Derived Distributions Applied in Hydrology.* Water Resources Publication, Littleton, Colorado.

Bobee, B. B., & Robitaille, R. (1977). The use of the Pearson type 3 and log Pearson type 3 distributions revisited. Water Resources Research, 13(2), 427-443. https://doi.org/10.1029/WR013i002p00427

Bhavyashree, S., & Bhattacharyya, B. (2018). Fitting Probability Distributions for Rainfall

Analysis of Karnataka, India. *International Journal of Current Microbiology and Applied Sciences,* 7(03), 1498-1506. https://doi.org/10.20546/ijcmas.2018.703.178

Chandran, S., Selvan, P., Namitha, M. R., Mishra, P., & Kumar, V. (2023). Probability analysis and rainfall forecasting using ARIMA model. MAUSAM Journal, 74(4), 1081-1092. https://doi.org/10.54302/mausam.v74i4.805

Deka, S., Borah, M., & Kakaty, S. C. (2009). Distributions of Annual Maximum Rainfall Series of North-East India. *European Water*, 3–14.

Douka, M., & Karacostas, T. (2018). Statistical analyses of extreme rainfall events in Thessaloniki, Greece. Atmospheric Research, 208, 60-77. https://doi.org/10.1016/j.atmosres.2017.08.025

Feng, S., & Nadarajah, S., & Hu, Q. (2007). Modeling Annual Extreme Precipitation in China Using the Generalized Extreme Value Distribution. *Journal of the Meteorological Society of Japan,* 85, 599–613.

Ferguson, T. S. (1964). A Characterization of the Exponential Distribution. *The Annals of Mathematical Statistics,* 35(3), 199–1207

Gaddum, J. H. (1945). Lognormal Distributions. Nature, 156(3964), 463-466. https://doi.org/10.1038/156463a0

Griffis, V. W., & Stedinger, J. R. (2007). Log-Pearson Type 3 Distribution and Its Application in Flood Frequency Analysis. I: Distribution Characteristics. *Journal of Hydrologic Engineering,* 12(5), 482–491. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:5(482)

Ghosh, S., Roy, K. M., & Biswas, C. S. (2016). Determination of the Best fit Probability Distribution for Monthly Rainfall Data in Bangladesh. *American Journal of Mathematics and Statistics*, 6(4), 170-174. DOI: 10.5923/j.ajms.20160604.05

Gupta, S. C., & Kapoor, V. K. (1997). *Fundamentals of Mathematical Statistics*. Sultan Chand and Sons, New Delhi: 11.23-12.2

Gumbel, E. J. (1941). The return period of flood flows. *The Annals of mathematical Statistics,* 12, 163-190.

Haseeb, F., Ali, S., Ahmed, N., Alarifi, N., & Youssef, Y. M. (2025). Comprehensive Probabilistic Analysis and Practical Implications of Rainfall Distribution in Pakistan. *Atmosphere,* 16(2), 122. https://doi.org/10.3390/atmos16020122

Johnson, N. L., Kotz. S., & Balakrishnan, N. (1995). *Continuous Univariate Distributions*. Wiley, New York, 2nd. Edition.

Karami, H., Ghazvinian, H., & Dadrasajirlou, Y. (2023). Application of statistical and geostatistical approaches in temporal and spatial estimations of rainfall. *Journal of Water and Climate Change,* 14(5), 1696-1722. https://doi.org/10.2166/wcc.2023.

Kottegoda, N. T. (1987). Fitting Johnson S_B curve by the method of maximum likelihood to annual maximum daily rainfalls. *Water Resources Research,* 23(4), 728–732

Koutrouvelis, I. A., & Canavos, G. C. (1999). Estimation in the Pearson type 3 distribution. *Water Resources Research,* 35(9), 2693–2704.

Kumar, V., Khan, S., & Jahangeer, J. (2017). Statistical distribution of rainfall in Uttarakhand, India. *Applied Water Science*, 7, 4765-4777. https://doi.org/10.1007/s13201-017-0586-5

Kurniawan, V. (2019). *Distribution fitting on rainfall data in Jakarta*. First International Conference of Construction, Infrastructure, and Materials, 650. DOI: 10.1088/1757-899X/650/1/012060

Mamoon, A. A. & Rahman, A. (2017). Selection of the best fit probability distribution in rainfall frequency analysis for Qatar. *Natural Hazards,* 86, 281-296. https://doi.org/10.1007/s11069-016-2687-0

Mandal, S., & Choudhury, B. (2015). Estimation and prediction of maximum daily rainfall at Sagar Island using best fit probability models. *Theoretical and Applied Climatology*, 121, 87-97. https://doi.org/10.1007/s00704-014-1212-1

Millington, N., Das, S., & Simonovic, S. (2011). *The Comparison of GEV, Log-Pearson Type 3 and Gumbel Distributions in the Upper Thames River Watershed under Global Climate Models (Water Resource Report no. 077)*. Department of Civil and Environmental Engineering, The University of Western Ontario.

Moccia, B., Mineo, C., Ridolfi, E., Russo, F., & Napolitano, F. (2021). Probability distributions of daily rainfall extremes in Lazio and Sicily, Italy, and design rainfall inferences. Journal of Hydrology: Regional Studies, 33, 100771. https://doi.org/10.1016/j.ejrh.2020.100771

Mohamed, J., & Adam, B. M. (2022). Modeling of magnitude and frequency of extreme rainfall in Somalia. *Modelling Earth Systems and Environment,* 8, 4277-4294. https://doi.org/10.1007/s40808-022-01363-0

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.

Ozonur, D., Pobocikova, I., & Souza, A. (2021). Statistical analysis of monthly rainfall in Central West Brazil using probability distributions. *Modeling Earth Systems and Environment*, 7, 1979-1989. https://doi.org/10.1007/s40808-020-00954-z

Pearson, K. (1895). Continuous to the mathematical theory of evolution, II: Skew variation in homogeneous materials. *Philosophical Transactions of the Royal Society, London*, 186, 343-414.

Phien, H. N., & Jivajirajah, T. (1984). Applications of the log Pearson type-3 distribution in hydrology. *Journal of Hydrology,* 73(3), 359–372.

Sarker, R. P., Biswas, B. C., & Khambete, N. N. (1982). Probability analysis of short period rainfall in dry farming tract in India. *MAUSAM Journal,* 33(3), 269-284.

Schellander, H., Lieb, A., & Hell, T. (2019). Error Structure of Metastatistical and Generalized Extreme Value Distributions for Modeling Extreme Rainfall in Austria. *Earth and Space Science,* 6(9), 1616-1632. https://doi.org/10.1029/2019EA000557

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics,* 6, 461-464

Sharma MA, & Singh JB (2010). Use of Probability Distribution in Rainfall Analysis. *New York Science Journal,* 3(9)

Shaharudin, S. M., Ahmad, N., Mohamed, N. S., & Mahdin, H. (2020). Fitting statistical

Distribution of extreme rainfall data for the purpose of simulation. *Indonesian Journal of Electrical Engineering and Computer Science,* 18(3), 1367-1374. http://doi.org/10.11591/ijeecs.v18.i3.pp1367-1374

Swift, L. W., & Schreuder, H. T. (1981). Fitting Daily Precipitation Amounts Using the SB Distribution. *Monthly Weather Review,* 109(12), 2535–2540.

Thom, H. C. S. (1958). A note on the Gamma distribution. *Monthly Weather Review,* 86(4), 117–122.

Weibull, W. (1939). A Statistical Theory of the strength of Materials. Ingeniors Ventenskaps Akademiens Handlingar, Royal Swedish Institute for Engineering Research, Stockholm Sweden, 151

Appendix A: Parameter Estimates of Guwahati Station

Table A. Estimates of probability distribution fitting of total monthly rainfall in Guwahati station for remaining months

Months	Distribution	Shape1	Shape2	Location	Scale
April	**Pearson I**	0.7542	2.1747	24.0000	522.8001
May	**Pearson I**	0.9553	3.1823	88.5000	642.4579
June	**Pearson I**	1.2658	1.8853	100.8439	496.6607
July	**Log-Normal**	-	-	5.6387	0.3776
August	**Weibull**	2.3195	-	-	270.3478
September	**Weibull**	2.3941	-	-	214.9069
October	**Weibull**	1.3993	-	-	126.3508
November	**Gamma**	0.2627	-	-	0.022
December	**-**	-	-	-	-

Table B. Estimates of probability distribution fitting of total monthly rainfall in Guwahati station for Winter, Pre-Monsoon, Monsoon and Post-Monsoon season

Months	Distribution	Shape1	Shape2	Location	Scale
Winter	**Gamma**	0.3085	-	-	0.0197
	Logistic	-	-	12.9754	8.8017
	Exponential	0.0640	-	-	-
	Gumbel	-	-	8.4552	10.9346
	Weibull	0.4250	-	-	9.4203
	Log-Normal	-	-	0.5238	4.3783
	Pearson 0	-	-	15.6144	17.8897
	Pearson I	0.5949	10.2068	0.0001	266.7931
	Pearson III	-	0.7298	0.0001	17.0653
	Pearson V	-	1.5629	-3.9785	14.8985
Pre-Mon	**Pearson I**	0.9917	4.8964	3.8000	911.8026
Mon	**Pearson I**	3.4559	11.5491	1.8968	1113.133
Post-Mon	**-**	-	-	-	-

In Table A, the parameters for April to December at the Guwahati station are listed. Similarly, in Table B, each parameter for each season i.e. Winter (January, February), Pre-Monsoon (March, April, May) and Monsoon (June, July, August, September) are mentioned except Post-Monsoon season (October, November and December), since Post-Monsoon season does not follow any one of the 10 selected distributions with the rainfall data in Guwahati Station. For the winter season, parameter estimates are presented for all ten selected distributions, whereas for the other seasons, only the parameters of the best-fitted distribution are reported.

Appendix B: Parameter Estimates of Silchar Station, Tezpur Station, Dibrugarh Station and Dhubri Station in monthly and season wise data.

Table C: Estimates of best probability distribution fitting of total monthly rainfall in Silchar Station

Station: Silchar					
Estimates					
Months	**Distribution**	**Shape1**	**Shape2**	**Location**	**Scale**
January	**Gumbel**	-	-	3.9666	6.6712
February	**Pearson I**	0.3509	0.9882	0.0001	0.0652
March	**Weibull**	1.3377	-	-	162.4325
April	**Pearson I**	0.8518	1.8127	23.2000	802.9113
May	**Weibull**	2.2891	-	-	468.592
June	**Pearson II**	1.7622	-	109.7549	803.5272
July	**Pearson V**	-	5.6109	492.3355	120.8108
August	**Logistic**	-	-	406.7817	85.6269
September	**Gamma**	-	6.4457	-	0.0173
October	**Gumbel**	-	-	136.06	82.2392
November	**Gumbel**	-	-	14.0195	19.143
December	**Weibull**	-	0.4074	-	95.1627
Season wise					
Winter	**Pearson I**	0.3621	1.0518	0.0001	659.5552
Pre-Mon	**Pearson I**	0.3588	1.0380	0.0001	657.548
Monsoon	**Logistic**	-	-	443.0995	96.6872
Post-Mon	-	-	-	-	-

Table D. Estimates of best probability distribution fitting of total monthly rainfall in Tezpur Station

Station: Tezpur					
Months	**Distribution**	**Shape1**	**Shape2**	**Location**	**Scale**
January	**Pearson I**	0.2866	1.2953	0.0001	101.5330
February	**Pearson I**	0.6298	2.1496	0.0001	110.4932
March	**Pearson I**	0.6321	0.8496	0.9000	109.7770
April	**Gamma**	3.4355	-	-	0.0205
May	**Weibull**	3.1960	-	-	283.2683
June	**Pearson 0**	-	-	303.8474	142.9531
July	**Gamma**	9.4320	-	-	0.0308
August	**Pearson I**	0.9232	1.3029	97.5000	425.9379
September	**Pearson I**	0.9534	2.3509	85.5000	475.2478
October	**Pearson I**	0.8165	1.6495	12.0000	254.0389
November	**Pearson I**	0.5354	3.7419	0.0001	147.7994
December	**Pearson I**	0.3106	3.5235	0.0001	133.2491
Season wise					
Winter	**Pearson I**	0.3837	1.6437	0.0001	103.9733
Pre-Mon	**Pearson I**	0.9077	2.0101	0.9000	482.7833
Monsoon	**Gumbel**	-	-	220.4292	103.0446
Post-Mon	**Pearson I**	0.3455	1.7339	0.0001	278.8994

Table E. Estimates of best probability distribution fitting of total monthly rainfall in Dibrugarh Station

Station: Dibrugarh					
Months	**Distribution**	**Shape1**	**Shape2**	**Location**	**Scale**
January	**Weibull**	1.2982	-	-	26.0708
February	**Weibull**	1.2670	-	-	65.7468
March	**Pearson I**	0.6223	0.8717	28.8000	211.9777
April	**Weibull**	1.2982	-	-	26.0708
May	**Pearson III**	-	3.3481	29.1147	82.5268
June	**Logistic**	-	-	396.9677	90.1618
July	**Pearson I**	4.572	4.7715	639.4236	245.1121
August	**Logistic**	-	-	373.2674	82.0850
September	**Pearson I**	0.9131	1.3561	128.1000	500.7443
October	**Pearson I**	0.8149	1.5532	30.6000	289.403
November	**Gumbel**	-	-	13.5589	22.0851
December	**Weibull**	1.9822	-	-	5.7701
Season wise					
Winter	**Weibull**	1.0804	-	-	44.029
Pre-Mon	**Gamma**	2.3473	-	-	0.0110
Monsoon	**Pearson III**	-	-	397.9645	149.3592
Post-Mon	**Weibull**	0.4773	-	-	32.7641

Table F. Estimates of best probability distribution fitting of total monthly rainfall in Dhubri Station

Station: Dhubri					
Months	**Distribution**	**Shape1**	**Shape2**	**Location**	**Scale**
January	-	-	-	-	-
February	**Gamma**	0.2849	-	-	0.0048
March	**Gumbel**	-	-	31.0939	37.3340
April	**Gumbel**	-	-	141.1324	88.0224
May	**Logistic**	-	-	324.4702	73.7592
June	**Weibull**	3.1264	-	-	579.8261
July	**Pearson VII**	-	5.6109	492.3355	120.8100
August	**Gumbel**	-	-	248.8648	137.3011
September	**Gamma**	6.4457	-	-	0.0173
October	**Weibull**	1.8043	-	-	204.1932
November	**Pearson I**	0.3539	0.7547	0.0001	94.8928
December	-	-	-	-	-
Season wise					
Winter	-	-	-	-	-
Pre-Mon	**Gumbel**	-	-	120.9820	117.9553
Monsoon	**Weibull**	2.649	-	-	494.7737
	Pearson III	-	25.8363	458.5850	34.6120
Post-Mon	-	-	-	-	-

Table C, D, E and F represents the estimates of parameters of best distribution among ten selected distribution considering rainfall data for each month and season wise for Silchar, Tezpur, Dibrugarh and Dhubri station respectively.