Journal of Mathematical Problems, Equations and Statistics

E-ISSN: 2709-9407
P-ISSN: 2709-9393
JMPES 2025; 6(1): 148-156

www.mathematicaljournal.com
Received: 02-01-2025
Accepted: 04-02-2025

Upendra Kumar Sahoo
Department of Statistics, Visva-Bharati, Santiniketan, West Bengal, India

Tirthankar Ghosh
Department of Statistics, Visva-Bharati, Santiniketan, West Bengal, India

Corresponding Author:
Upendra Kumar Sahoo
Department of Statistics, Visva-Bharati, Santiniketan, West Bengal, India

Forecasting regional level rainfall of West Bengal

Upendra Kumar Sahoo and Tirthankar Ghosh

Abstract
Accurate forecasting of Indian Summer Monsoon Rainfall (ISMR) at the district level remains a significant challenge, particularly in India. This study focuses on district-level ISMR rainfall in West Bengal from 1901 to 2020. Various probability distributions, including Gamma, Generalized Extreme Value, Weibull, and Log-Normal, were fitted to the data at the district scale. Among these, the Weibull distribution emerged as the best-fit model, effectively capturing rainfall patterns in 94% of districts, as indicated by p-values greater than 0.05. Furthermore, categorical rainfall forecasting (above normal, normal, and below normal) based on the Weibull distribution demonstrated high accuracy, exceeding 85% for all districts except Darjeeling. These findings highlight the suitability of the Weibull distribution for district-level rainfall forecasting in West Bengal, offering valuable insights for agricultural planning and climate risk management.

Keywords: ISMR, Gamma, Generalized Extreme Value, Weibull and Log Normal

Introduction

The climate of a region is the average weather over 30-year time period. In weather study, major parameters of interest are rainfall, temperature, humidity, sea surface temperature etc. Rainfall plays a crucial role in the study of weather and climate. The rainfall pattern changes of a region, over a considerable period of time, indicates change in the general climatic conditions of that region. Therefore, the trend (long-term upward/downward movement) analysis of rainfall of a region over the years is necessary to assess if there is any change in the climate with regard to rainfall in that region. In this chapter concentration has been put to study the rainfall pattern change in West Bengal and its districts over the period 1901-2020. Categorical predictions of rain at district levels have also been carried out here.

Researchers have focused on trend analysis, fitting different distributions, and developed statistical models for forecasting Indian Summer Monsoon Rainfall (ISMR) over the country and state levels. Suman Kumar Kundu and Tarun Kumar Mondal (2019) [1] analysed long-term annual and seasonal rainfall trends in West Bengal using 102 years (1901 to 2002) monthly rainfall data from 18 rainfall stations. The study results showed a significant increasing trend at five stations in annual rainfall, six stations in the monsoon season, and eight stations in the post-monsoon season.

Dash *et al.* (2007) [2] showed a decreasing tendency in summer monsoon rainfall over Indian landmass and an increasing trend during pre-monsoon and post-monsoon season months.

N Subash and A K Sikka analysed the mean monthly rainfall over thirty-five meteorological subdivisions and five homogeneous regions of India from 1904 to 2003 (2014) [3]. In this study, they described no direct relationship between increasing rainfall and increasing maximum temperature monthly and seasonally over the meteorological subdivisions of India. There is a possible relationship between rainfall trends and temperature over India's homogeneous regions.

A rainfall pattern analysis at the annual and seasonal scale at the district level of West Bengal from 1991 to 2000 was done using Mann-Kendall and Theil-Sen's slope estimate by (Verma *et al.*, 2022) [4]. The study showed that the state's southern districts had a decreasing rainfall trend in both annual and monsoon seasons, but rainfall showed an increasing trend in the summer season. In the post-monsoon season, rainfall has shown a downward trend, and in the winter season rainfall exhibited an upward trend in most of the state's districts during that period.

Rainfall patterns have also been explained through the fitting of probability distributions by the researchers.

Monthly rainfall data from 13 Iraqi stations during the period 1970 to 2010 were studied using the fitting of different probability distributions such as Normal, Gamma, and Weibull (Alghazali N. *et al.*, 2014) [5].

The monthly rainfall data of all districts of Uttarakhand from 1901 to 2002 were fitted to several probability distributions, such as chi-squared, chi-squared (2 parameters), exponential, exponential (2 parameters), gamma, gamma (3 parameters), gen. extreme value (GEV), log-Pearson 3, Weibull and Weibull (3 parameters) (Kumar V., 2017) [6]. The study results show that 46, 24, 15, and 15% of districts are best fitted by Weibull, chi-squared, gamma, and log-Pearson (3 parameters) distributions, respectively.

Machine learning techniques such as artificial neural network (ANN), support vector machine (SVM), decision tree (DT), and random forest (RF) have been used by researchers to classify ISMR. K means clustering and map-to-map correlation methods are used to classify Indian summer monsoon rainfall using the data set from 1871-1990 (Kulkarni, *et al.*, 1992) [7]. ANN has been used to classify rainfall variable data in climatological stations in Cyprus for a long period from 1917 to 1995 (Michaelides *et al.*, 2001) [8].

A comparative study among Deep Neural Networks (DNN), Random Forest (RF), and Support Vector Machines (SVM) was done for forecasting rainfall (Sangiorgo *et al.*, 2020) [9]. Reanalysis data was used in this study to downscale 26 variables from the National Centre for Environmental Prediction (NCEP). This study shows that DNN and RF models are more efficient than the SVM model.

SVM, K-nearest neighbors, RF, and Decision Tree (DT) have been used to classify rainfall data of the districts in Andhra Pradesh (Raju *et al.*, 2017) [10]. They found that SVM produces better predictions than other methods. Backpropagation networks (BPN) and Radial Basis Function Networks (RBFN) have been used to classify rainfall data for flood prediction in Kuching City, Malaysia (Chai *et al.*, 2017) [11]. They used six meteorological parameters of daily data to classify the rainfall into four classes, from lighter to very heavy rainfall. This study has shown that BPN is better performed than RBFN.

A binary classification of rainfall time series using a machine-learning technique was done for the state of Maharashtra (Hundurkar *et al.*, 2022) [12]. SVM and ANN models were developed using pressure, humidity, and temperature data from ISMR. The accuracy of the model was 82.1 and 82.8%, respectively. However, no such study has been reported on the Indian state of West Bengal so far.

In West Bengal, agricultural production is heavily dependent on the ISMR. Therefore, it is very crucial that a district-wise rainfall pattern analysis is undertaken and possible categorical predictions are made. In this chapter, the ISMR rainfall data of 120 years (1901-2020) are taken for such an analysis, and categorical predictions of the ISMR at the district level of West Bengal have been considered. Accounts of the error of such prediction have also been computed and reported in the results section. The article is organised in the following manner. Section 2 presents the study area and data sources. The methodology of the study is described in Section 3. The results and discussion are stated in Section 4. Finally, the conclusion made in Section 5.

Study Area and Data Sources

Study Area

The Indian state of West Bengal is situated in the eastern part of India. It spreads over 21.30^0 to 27.30^0 N and 85.30^0 to 89.45^0 E geographically. The study area is shown in Figure 1. West Bengal covers 88,752 square km area and is about 3% of the total area of India. In the East, it is bounded by Bangladesh, and the Bay of Bengal is in the South. Odisha and Jharkhand have covered the southeast and west directions of the state. In the northwest, it is bounded by Bihar and Nepal, and in the north by Sikkim and Bhutan. Assam bounds it in the northeast. West Bengal is a part of the Ganga basin, which is very close to the Bay of Bengal. The entire West Bengal, except a small portion in the west and the hilly region of the north, is very fertile and it is famous for its paddy, potato, jute and vegetable productions.

Data Sources

According to IMD, the ISMR starts on 1st June and ends on 31st September in India. In this study, we selected monthly (June, July, August, and September) and Seasonal (sum of June, July, August, and September) rainfall data for 19 districts of West Bengal. The necessary cleaning of the data has been carried out before the actual analysis. The above data were obtained from the India Water Resources Information System (India WRIS) (https://indiawris.gov.in/wris/rainfall). The data length was from 1901 to 2020.

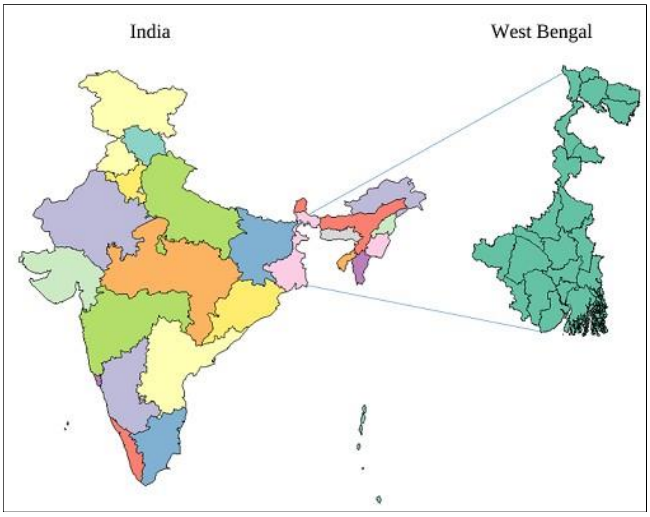

Fig 1: Maps of Considered Study Area

Methodology

The entire data of 120 years have been divided into four-time segments, each of 30 years. The 1st, 2nd, 3rd, and 4th segments contained the rainfall data for the periods 1901-1930, 1931-1960, 1961-1990, and 1991-2020, respectively. The descriptive statistics study of the time segments of rainfall in West Bengal has been carried out. The rainfall data was fitted to a Gamma, Weibull, Generalized extreme value (GEV), and Log Normal distribution with shape and scale parameters. Then, the data is classified into below normal, normal, and above normal by Weibull distribution. In this model, the data from 2001 to 2020 are used to predict the rainfall category in each district of West Bengal. This model was developed using 100×4 data points for rainfall in each district.

Gamma Distribution

The gamma distribution is a continuous probability distribution. The Probability density function (pdf) of a random variable x having the gamma distribution is given by

$$f(x;\alpha,\beta)=\frac{x^{\alpha-1}e^{\frac{-x}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)};x>0;\alpha>0,\beta>0 \quad (1)$$

Where:

- α is the shape parameter,
- β is the scale parameter,
- Γ (α) is the Gamma function evaluated at α.

Generalized extreme value Distribution (GEV)

GEV is a family of continuous probability distributions developed within Gumbel, Frechet, and Wuibell distributions. The probability density function of the distribution with the parameters location (−∞ < μ < ∞), scale (σ > 0), and shape (−∞ < ξ < ∞) in the following. The GEV distribution is used to model the distribution of the maximum and minimum of a large number of independent, identically distributed random variables. The probability density function

$$f(x;\mu,\sigma,\xi)=\frac{1}{\sigma}\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-\left(1+\frac{1}{\xi}\right)}\exp\left(-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right) \quad (2)$$

Where

- μ is the location parameter,
- σ is the scale parameter,
- ξ is the shape parameter,
- The support of the distribution depends on the value of ξ, it can be the real line (when ξ = 0, bounded below (when ξ > 0) and bounded above (when ξ < 0).

Weibull Distribution

In statistics and probability theory, Weibull distribution is a continuous probability distribution used to model various data with non-negative values. The pdf of the distribution is given by

$$f(x;\lambda,k)=\begin{cases}\frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1}e^{-(x/\lambda)^{k}} & x\geq 0\\ 0 & x<0\end{cases} \quad (3)$$

Where

- λ is the scale parameter,
- k is the shape parameter.

Log Normal Distribution

The log normal distribution is also continuous probability distribution. A random variable is said to follow a log normal if logarithm of the variable follows normal distribution. The probability density function of log normal distribution of a random variable is given by

$$f(x;\mu,\sigma)=\begin{cases}\frac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\frac{(lnx-\mu)^2}{2\sigma^2}\right) & x>0\\ 0 & x\leq 0,\end{cases} \quad (4)$$

Where

- μ is the mean,
- σ is the standard deviation of the logarithm of the variable.

In this case, logarithmic transformed rainfall data were fitted to the normal distribution.

Results

Descriptive Statistics

The mean, standard deviation (SD), and coefficient of variation (cv) in ISMR data for 19 districts of West Bengal in 30-year intervals from 1901 to 2020 are calculated. The descriptive statistics results of four segments are depicted in Table 1. In Bankura, Darjeeling, Purulia, Nadia, Maldah, Dakshin Dinajpur, and Birbhum districts, rainfall variation increased over the segments. The rainfall variation decreased in Bardhaman, South 24 Parganas, Uttar Dinajpur, North 24 Parganas, Kolkata, Hugli, Howarh and Jalpaiguri. There is no change in rainfall variation in West Midnapore, East Midnapore, and Coochbehar.

Annual Average Rainfall

The mean, standard deviation (SD), and coefficient variation (CV) of annual rainfall in the 19 districts in the state of West Bengal from 1901 to 2020 were calculated. The results shown in Table 2. The annual average rainfall, sd and cv of West Bengal are 1792.46, 201.16, and 11.22 respectively. The range of annual average rainfall and standard deviation in considered districts varied from 1331.16 mm (Purulia) to 3772.3 mm (Jalpaiguri) and 224.42 (Purulia) to 636.15(Jalpaiguri), respectively. In Darjeeling, Jalpaiguri and Coochbehar, the average annual rainfall is more than 3000 mm during the considered period. Except for the above stations/districts, the other districts ranged from 1300 mm to 2000 mm. The only Uttar Dinajpur gained 2147.48 mm of rainfall. The rainfalls vary from 13 (Darjeeling) to 27% (Kolkata) from their annual average rainfall in the state.

Table 1: Mean, Standard Deviation (SD) and Coefficient Variation (CV) of ISMR and Annual rainfall for 19 districts of West Bengal during 1901-1930, 1931-1960, 1961-1990 and 1991-2020

Year	1901-1930			1931-1960			1961-1990			1991-2020		
Districts	**Mean**	**SD**	**CV**	**Mean**	**SD**	**CV**	**Mean**	**SD**	**CV**	**Mean**	**SD**	**CV**
Bankura	1039.23	191.7	18.4	1014.37	179.3	17.7	1065.42	214.5	20.1	1073.08	191.1	17.8
Bardhaman	1021.21	180.3	17.7	977.27	164.8	16.9	1069.74	262.7	24.6	998.92	247.1	24.7
Birbhum	1031.50	205.5	19.9	1020.31	204.5	20	1069.54	204.8	19.2	1029.45	286.2	27.8
Dakshin Dinajpur	1481.71	316.7	21.4	1437.71	208.6	14.5	1290.90	297.7	23.1	1222.29	494.5	40.5
Darjeeling	2506.09	272.9	10.9	2506.03	335.6	13.4	2533.95	360	14.2	2724.67	439.5	16.1
East Midnapore	1148.85	240	20.9	1233.16	228.6	18.5	1299.2	253.4	19.5	1287.97	259.1	20.1
Howarh	1077.25	233.6	21.7	1224.17	278.8	22.8	1343.19	385.4	28.7	1196.65	258.7	21.6
Hoogly	973.62	182	18.7	1119.38	281.9	25.2	1187.95	273.3	23	1058.68	229.9	21.7
Jalpaiguri	3052.48	505	16.5	2989.59	438.1	14.7	2964.33	639.6	21.6	2752.39	471.8	17.1
Coochbehar	2305.6	391.5	17	2193.45	295.5	13.5	2321.31	485.3	20.9	2371.98	515.3	21.7
Kolkata	947.21	280.3	29.6	1157.98	347.9	30	1269.35	415.5	32.7	1079.18	315.5	29.2
Maldah	1161.94	310.7	26.7	1154.15	219.9	19.1	1146.45	306.7	26.8	1183.73	413.4	34.9
Murshidabad	1041.38	206.6	19.8	1028.13	185.3	18	1050.35	217.4	20.7	957.94	281.1	29.3
Nadia	989.87	148	15	973.4	191.2	19.6	1029.78	274.6	26.7	954.33	290.9	30.5

N24P	976.29	184.9	18.9	1107.87	256.9	23.2	1272.17	307.2	24.1	1170.25	251.4	21.5
Purulia	1050.73	188.8	18	1033.21	166.9	16.2	1047.42	217.2	20.7	1112.28	213.2	19.2
S24P	1071.43	230.6	21.5	1225.18	264.3	21.6	1349.93	364.9	27	1260.38	268.3	21.3
Uttar Dinajpur	1725.25	322.8	18.7	1789.03	287.8	16.1	1750.7	264.4	15.1	1662.72	381.7	23
West Midnapore	1123.07	193.5	17.2	1103.43	190	17.2	1181.76	206.6	17.5	1179.66	247.9	21

Table 2: Mean, Standard Deviation (SD) and Coefficient Variation (CV) of ISMR and Annual rainfall for 19 districts of West Bengal during 1991-2020

Districts	ANNUAL			ISMR		
	Mean	**SD**	**CV**	**Mean**	**SD**	**CV**
Bankura	1371.36	229.2	16.71	1048.02	193.48	18.46
Bardhaman	1362.44	266.12	19.53	1016.79	217.81	21.42
Birbhum	1360.4	268.31	19.72	1037.7	225.89	21.77
Dakshin Dinajpur	1747.78	389.8	22.3	1358.15	357.01	26.29
Darjeeling	3238.26	431.31	13.32	2567.68	364.24	14.19
East Midnapore	1680.16	302.29	17.99	1242.3	249.69	20.1
Howrah	1635.04	369.15	22.58	1210.31	306.19	25.3
Hoogly	1478.18	306.72	20.75	1084.91	254.53	23.46
Jalpaiguri	3772.3	636.15	16.86	2939.7	525.06	17.86
Coochbehar	3026.98	506.37	16.73	2298.09	430.13	18.72
Kolkata	1530.31	414.9	27.11	1113.43	358.99	32.24
Maldah	1475.65	338.38	22.93	1161.57	316.38	27.24
Murshidabad	1362.84	271.46	19.92	1019.45	225.55	22.12
Nadia	1354.96	288.4	21.28	986.85	232.42	23.55
N24P	1538.66	332.81	21.63	1131.65	272.75	24.1
Purulia	1331.16	224.32	16.85	1060.91	197.44	18.61
S24P	1655.76	363.3	21.94	1226.73	300.27	24.48
Uttar Dinajpur	2147.48	342.76	15.96	1731.93	316.6	18.28
West Midnapore	1540.16	253.64	16.47	1181.41	202.37	17.13

Districtwise summer monsoon rainfall

The descriptive statistics of annual ISMR were done from 1901 to 2020. The mean, sd, and cv are shown in Table 2. The range of average annual ISMR and standard deviation are varied from 986.85 (Nadia) mm to 2939.87 (Jalpaiguri) mm and 193.48 (Bankura) to 525.06 (Jalpaigiuri) respectively. The coefficient variation of annual ISMR is varied from 14 (Darjeeling) to 32% (Kolkata) from their average annual ISMR in the state. The figures have shown the annual ISMR district wise of West Bengal. The range of ISMRs are 700-1600, 500-1900, 500-1600, 300-2300, 1500-4000, 700-2000, 400-2700, 300-1800, 1500-4300, 1300-2800, 200-2400, 500-2200, 300-1600, 200-1500, 300-2300, 600-1700, 500-2800, 900- 2800 and 800-1900 mm in Bankura, Bardhaman, Birbhum, Dakshin Dinajpur, Darjeeling, East Midnapore, Howrah, Hoogly, Jalpaiguri, Coochbehar, Kolkata, Maldah, Murshidabad, Nadia, North 24 Parganas, Purulia, South 24 Parganas, Uttar Dinajpur and West Midnapore respectively during 1901-2020. The ISMR is increasing trend in Bankura, East Medinipur, Howrah, Hoogly, Coochbehar, South 24 Parganas, Purulia, North 24 Parganas, and Kolkata. In Dakshin Dinajpur, Jalpaiguri, Nadia, Uttar Dinajpur, and Murshidabad, annual ISMR showed a decreasing trend. There is no trend in Birbhum, West Midnapore, Maldah, and Bardhaman districts. We also have considered the decadal averages of ISMR for each district of West Bengal. The results have been shown in Figure 2. It is observed that in 13 out of 19 districts, the decadal average has decreased in the last two decades, i.e., 2000 to 2010 and 2010 to 2020. Increased decadal average has been observed in Darjeeling, Jalpaiguri, and Coochbehar districts in last two decades. However, in Kolkata, Uttar Dinajpur, and Birbhum the decadal average has increased in the last decade. Therefore, it may be said that in the last two decades, West Bengal has been receiving less ISMR on average. It may be an indication of change in the climate in the state with respect to ISMR.

Fitting of Probability Distribution

The results of fitting probability distributions, p-values and parameters are shown in Table 3. The chi-square test and p-values defined the probability distribution are fitted good or bad. If the p-value is greater than 0.05, then the probability distribution best fits the district. After fitting the probability distributions with the rainfall data of 19 districts, we observed that the gamma distribution fits 63% of the state's districts. The gamma distribution is fitting with the ISMR of Bankura, Birbhum, East Midnapore, Howrah, Coochbehar, Murshidabad, Nadia, North 24 Parganas, South 24 Parganas, Uttar Dinajpur, and West Midnapore. However, the distribution has not fitted the remaining districts, i.e., Bardhaman, Dakshin Dinajpur, Darjeeling, Hooghly, Jalpaiguri, and Kolkata rainfall data. The Weibull distribution fits 95% of the data of districts of West Bengal. Except for Howarh, the ISMR of all the districts have been fitted by Weibull distribution. All the districts except Howarh, South 24 Parganas's ISMR are fitted by the generalized extreme value (GEV) distribution. The GEV distribution has fitted 90% of the ISMR data of districts of the state. The log-normal distribution has fitted 84% of the ISMR of the stations/districts in the state. All the ISMR is fitted with the distribution except the ISMR of the districts Bardhaman, Dakshin Dinajpur, Darjeeling and Purulia. Among the above distributions, the Weibull distribution is the best-fitted with the ISMR data of the state's districts. The density, cumulative distribution function (cdf), P-P, and Q-Q plots of West Bengal and Bardhaman district of Weibull distribution are given in Figure 3 and Figure 4 respectively.

Table 3: Parameters of Gamma, Weibull, GEV and Log Normal Distributions for districts wise of West Bengal

Distribution	**Gamma**			**Weibull**			**GEV**				**Log Normal**		
Districts	α	β	P value	k	λ	P value	μ	σ	ξ	P value	μ	σ	P value
Bankura	30.2	0.03	0.21	5.6	1129.3	0.5	967.66	172.19	-0.13	0.09	6.94	0.18	0.14
Bardhaman	22.8	0.02	0	4.6	1104.6	0.19	925.58	191.78	-0.1	0.63	6.9	0.21	0
Birbhum	21	0.02	0.67	4.9	1128.5	0.55	948.38	211.93	-0.19	0.15	6.92	0.22	0.64
D Dinajpur	12.7	0.01	0.02	4.2	1491.2	0.59	1237.6	353.58	-0.25	0.47	7.17	0.3	0.01
Darjeeling	51	0.02	0.01	6.8	2727.5	0.07	2419.9	336.64	-0.14	0.49	7.84	0.14	0.01
E Midnapore	25.3	0.02	0.48	5.06	1344.7	0.21	1140.7	226.61	-0.14	0.4	7.1	0.2	0.53
Howrah	16.4	0.01	0.15	3.8	1326.4	0.05	1082.7	270.29	-0.09	0.02	7.07	0.25	0.26
Hoogly	17.9	0.02	0.03	4.4	1184.9	0.12	984.8	237.05	-0.17	0.25	6.96	0.24	0.07
Jalpaiguri	31.1	0.01	0.01	5.8	3159.5	0.2	2737.8	504.08	-0.21	0.28	7.97	0.18	0.1
Coochbehar	30	0.01	0.34	5.3	2478.7	0.14	2112.5	368.2	-0.08	0.91	7.72	0.18	0.42
Kolkata	9.5	0.01	0.04	3.2	1239.1	0.37	961.8	311.32	-0.1	0.95	6.96	0.34	0.09
Maldah	14.4	0.01	0.06	3.7	1281.3	0.19	1022.8	260.41	-0.05	0.93	7.02	0.26	0.16
Murshidabad	18.9	0.03	0.3	4.97	1108.5	0.62	938.63	225.71	-0.27	0.27	6.9	0.24	0.32
Nadia	16.6	0.02	0.32	4.5	1077.3	0.21	900.17	226.51	-0.23	0.11	6.86	0.26	0.23
N24P	17.6	0.02	0.19	4.14	1237.8	0.07	1018.2	245.55	-0.12	0.29	7	0.24	0.36
Purulia	29.5	0.03	0.05	6.57	1143.5	0.5	980.21	178.24	-0.14	0.37	6.95	0.18	0.05
S24P	18	0.02	0.49	3.83	1341.7	0.1	1101.6	257.28	-0.07	0	7.08	0.24	0.62
U Dinajpur	30	0.02	0.42	5.64	1864.1	0.29	1606.7	297.7	-0.17	0.61	7.44	0.18	0.5
W Midnapore	35.5	0.03	0.33	5.84	1267.8	0.16	1096.3	176.29	-0.1	0.62	7.06	0.17	0.36

Forecasting Model Results

As rainfall is a continuous variable, Gamma, Log Normal, Generalized extreme value (GEV), and Weibull continuous probability distributions are fitted in each district of West Bengal. The parameters and p-values are shown in Table 3. The Weibull probability distribution fits 94% of districts of West Bengal as the p-values are greater than 0.05. So, we predict the rainfall categories normal, below normal, and above normal in each district by using the Weibull distribution for 2000 to 2020. In the Bankura district, the model performs well with an accuracy of 85%, which is statistically significant given the p-value in Table 3. Kappa statistic measures the agreement between the predicted and actual classifications. The Kappa statistic of 0.72 suggests substantial agreement between predicted and actual classifications, implying that the model is reliable. The confusion matrix shows that most of the normal predictions are accurate (11 out of 14), with minor misclassifications between normal and below or above normal Table 4.1.

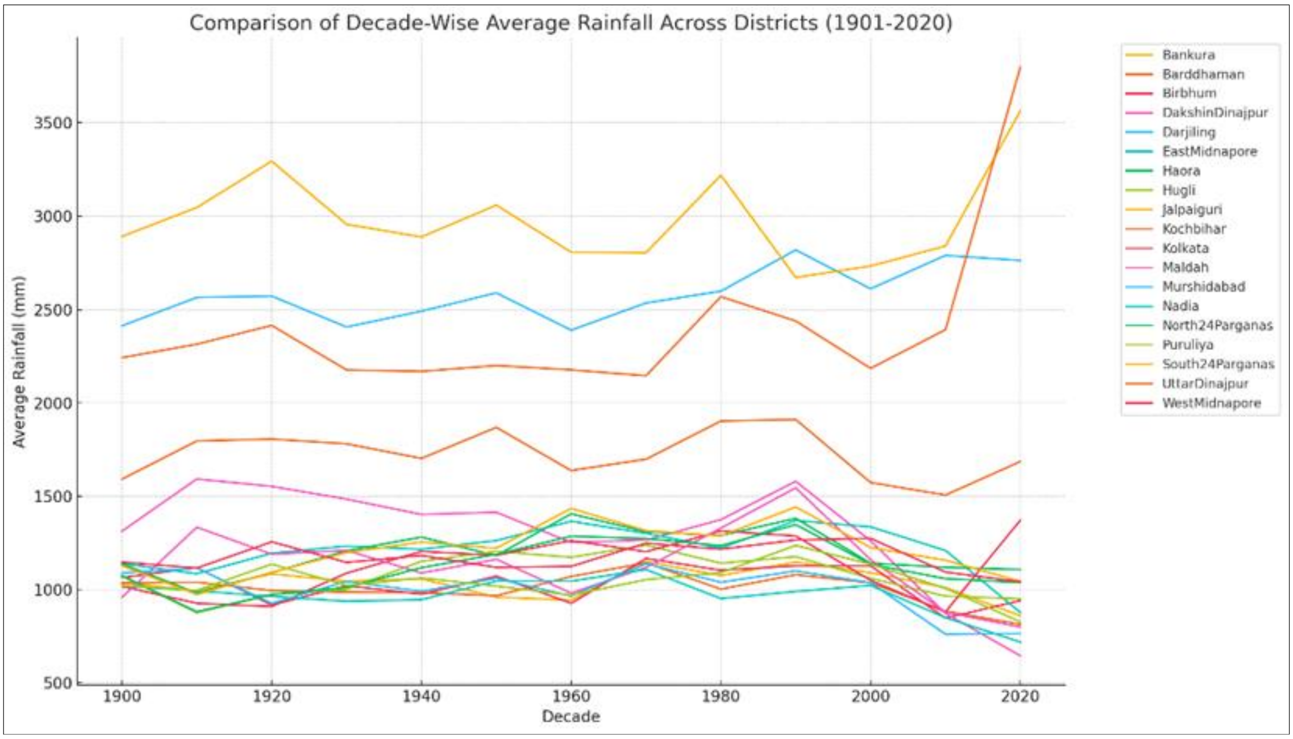

Fig 2: The decadal average ISMR for all disticts of West Bengal durimg 1901-2020

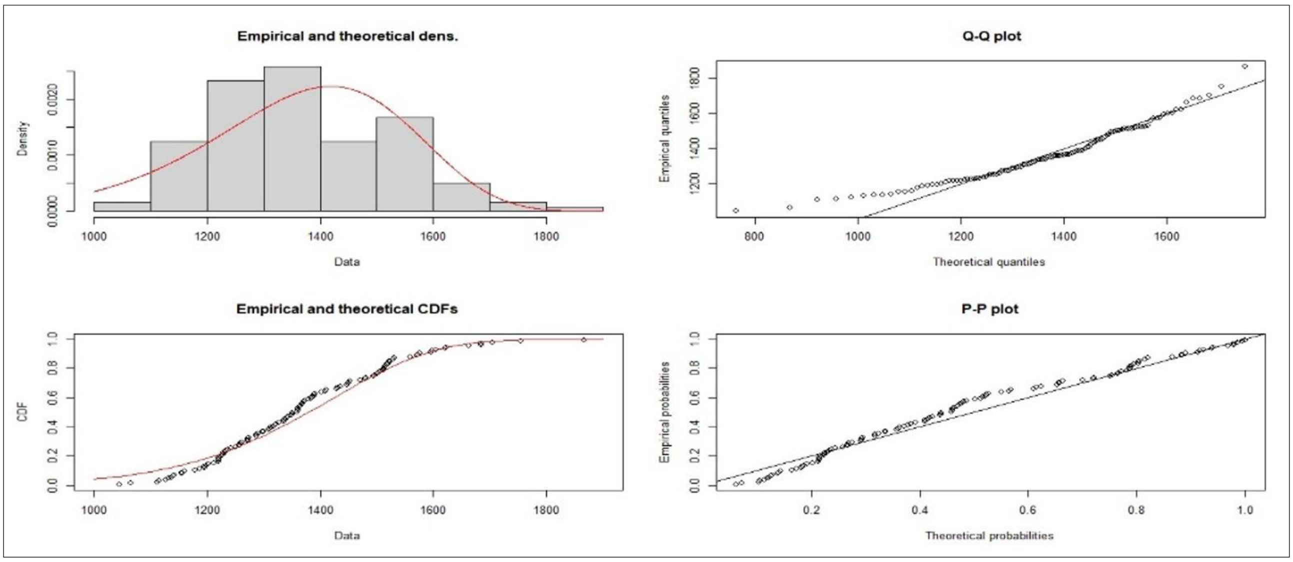

Fig 3: Density, CDF, P-P, and Q-Q plots of monsoon rainfall fitting with Weibull Distribution of West Bengal

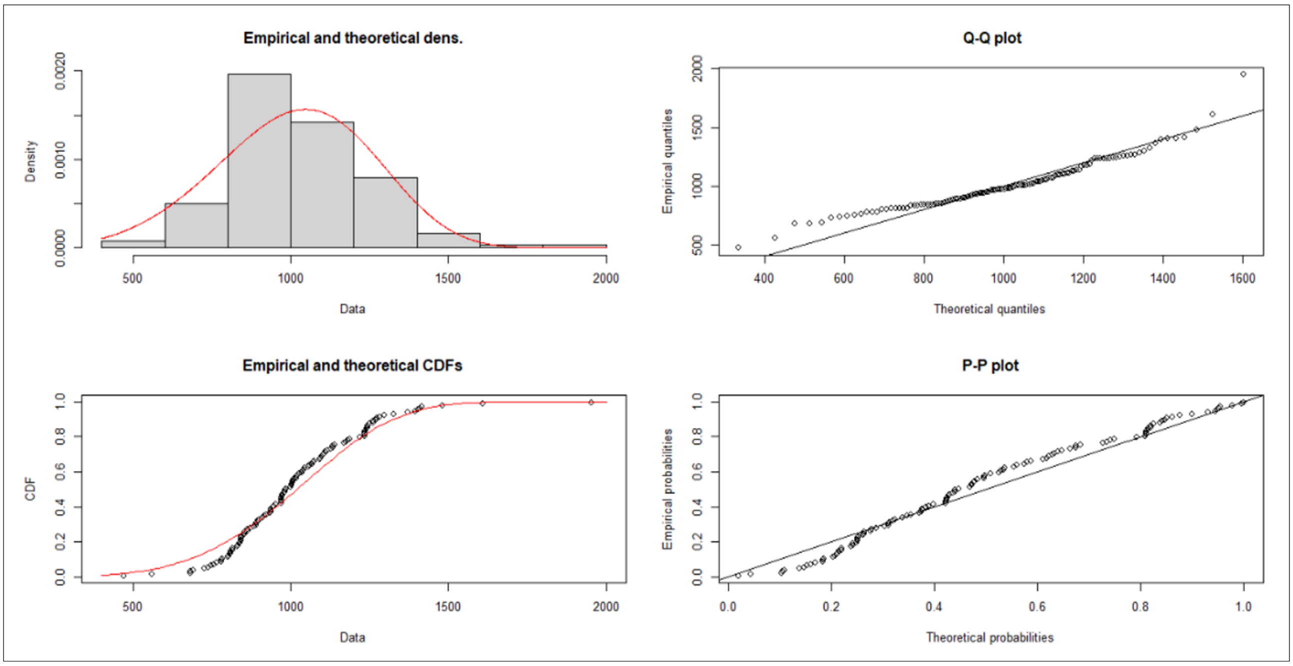

Fig 4: Density, CDF, P-P, and Q-Q plots of monsoon rainfall fitting with Weibull Distribution in Bardhaman Districts of West Bengal

The model showed strong performance with a 90% accuracy, which is statistically significant as indicated by the p-value in the Bardhaman district. The Kappa statistic of 0.83 points to a near-perfect agreement between the predicted and actual classifications, highlighting the model's high reliability. The confusion matrix showed that the majority of the below and above predictions are correct, with only slight misclassifications occurring in the normal category in the Table 4.2. The model demonstrated flawless performance with an accuracy of 100%, and the p-value confirms that this result is statistically significant. The Kappa statistic is 1. It indicated perfect agreement between predicted and actual classifications, underscoring the model's exceptional reliability. The confusion matrix showed complete accuracy across all categories, with no misclassifications in the Birbhum district 4.3.

In the Coochbehar district, the model has strong performance with a 90% accuracy, supported by the statistically significant p-value. The Kappa statistic of 0.82 indicates substantial agreement between the predicted and actual classifications, underscoring the model's high reliability. The confusion matrix showed that the majority of normal and above-normal predictions are correct, with only slight misclassifications in the below-normal category Table 4.4.

The model has strong performance with a 90% accuracy, which is statistically significant as indicated by the p-value in the Dakshin Dinajpur district. The Kappa statistic is 0.81. It pointed to substantial agreement between the predicted and actual classifications, highlighting the model's high reliability. The confusion matrix showed that the majority of below-normal and normal predictions are correct, with only slight misclassifications in the above-normal category 4.5. In the Darjeeling district, with a 75% accuracy rate, the model is performed satisfactorily and its p-value confirmed statistical significance. The Kappa statistic is 0.5726. It suggested a moderate level of agreement between predictions and actual results, indicating fair reliability. The confusion matrix showed accurate above-normal predictions, though some misclassifications occur in the normal and below-normal categories 4.6. The model performed impressively with a 95% accuracy, supported by a highly significant p-value. The Kappa statistic of 0.92 suggested nearly perfect agreement between predicted and actual classifications, indicating the model's high reliability. The confusion matrix of the model has shown that the majority of predictions for below-normal, normal, and above-normal are accurate, with minimal misclassifications in the East Midnapore district 4.7. The model performed well with an 85% accuracy, backed by a statistically significant p-value. The Kappa statistic is 0.66. It indicated substantial agreement between predicted and actual outcomes, highlighting the model's fair reliability in the Howrah district. The confusion matrix revealed accurate predictions for normal and above-normal, with a few errors in the below-normal category 4.8.

In the Hoogly district, with a 90% accuracy rate, the model is performed satisfactorily and its p-value confirmed statistical significance. The Kappa statistic is 0.66. It suggested a moderate level of agreement between predictions and actual results, indicating fair reliability. The confusion matrix of the model showed accurate above-normal and normal category predictions, though some misclassifications occurred in the below-normal categories 4.9. With an 85% accuracy rate, the model demonstrates good performance, and this result is statistically significant as shown by the p-value. The Kappa statistic of 0.74 pointed to substantial agreement between the predicted and actual classifications, highlighting the model's fair reliability. The confusion matrix of the model indicated that most predictions for below-normal and normal are accurate, though there are some misclassifications in the above-normal category in the Jalpaiguri district 4.10.

The model performs exceptionally well with 100% accuracy, supported by a significant p-value. The Kappa statistic 1 suggests perfect agreement between predicted and actual classifications, indicating the model's high reliability. The confusion matrix showed that the majority of predictions for below, normal, and above are correct, with no misclassifications in the Kolkata district 4.11. In the Malda district, the model showed excellent performance with a 95% accuracy, supported by a statistically significant p-value. The Kappa statistic 0.90 indicates almost perfect agreement between predicted and actual outcomes, highlighting the model's high reliability. The confusion matrix revealed accurate predictions for below-normal and normal categories with minimal misclassifications 4.12. The true accuracy of the model lies between 75% to 99%. The model exhibits strong performance with a 95% accuracy, supported by a statistically significant p-value. The Kappa statistic 0.91 indicates almost perfect agreement between predicted and actual categories, underscoring the model's high reliability. The confusion matrix showed accurate predictions for below-normal and normal, with minimal misclassifications with above-normal in the Murshidabad district 4.13. In Nadia district, the model has shown strong performance with an 85% accuracy, supported by a statistically significant p-value of 0.001. The Kappa statistic is 0.72. It suggested substantial agreement between predicted and actual categories, underscoring the model's reliability. The confusion matrix revealed accurate predictions for below-normal and normal, with some misclassifications in the normal category 4.14.

The model showed excellent performance with a 95% accuracy, supported by a statistically significant p-value in the North 24 Parganas districts. The Kappa statistic of 0.88 indicates almost perfect agreement between predicted and actual outcomes, underscoring the model's high reliability. The confusion matrix reveals accurate predictions for below-normal and normal, with minimal misclassifications in above-normal 4.15. The model performs very well with a 95% accuracy, supported by a significant p-value. The Kappa statistic of 0.92 suggested almost perfect agreement between predicted and actual classifications, indicating the model's high reliability. The confusion matrix showed that most predictions below-normal, normal, and above-normal predictions are correct, with no misclassifications between normal and above-normal in Purulia districts 4.16. Achieving 90% accuracy, the model performs well, as confirmed by the significant p-value. The Kappa value of 0.78 reflects substantial alignment between predictions and actual classifications, suggesting reliability. The confusion matrix showed that while below-normal and above-normal predictions are mostly accurate, there are minor misclassifications in the normal category in South 24 parganas 4.17.

Achieving 90% accuracy, the model performs well, as confirmed by the significant p-value. The Kappa value of 0.79 reflects substantial alignment between predictions and actual classifications, suggesting reliability. The confusion matrix showed that while belownormal and normal predictions are mostly accurate, there are minor misclassifications in the below-normal category in Uttar Dinajpur Table 4.18.

In the district West Midnapore, the model performed well with an 85% accuracy, supported by a significant p-value. The Kappa statistic is 0.75. It suggested substantial agreement between predicted and actual classifications, indicating the model's reliability. The confusion matrix of the model showed that the majority of below-normal and above-normal predictions are correct, with some misclassifications in normal category 4.19.

Table 4.1: Bankura District

Bankura		Reference		
		Below	**Normal**	**Above**
Predicted	Below	3	0	0
	Normal	1	11	2
	Above	0	0	3

Table 4.2: Bardhaman District

Bardhaman		Reference		
		Below	**Normal**	**Above**
Predicted	Below	7	0	0
	Normal	1	9	1
	Above	0	0	2

Table 4.3: Birbhum District

Birbhum		Reference		
		Below	**Normal**	**Above**
Predicted	Below	8	0	0
	Normal	0	7	0
	Above	0	0	5

Table 4.4: Coochbehar District

Coochbehar		Reference		
		Below	**Normal**	**Above**
Predicted	Below	2	0	0
	Normal	2	11	0
	Above	0	0	5

Table 4.5: Dakshin Dinajpur District

Dakshin Dinajpur		Reference		
		Below	**Normal**	**Above**
Predicted	Below	10	0	0
	Normal	0	8	2
	Above	0	0	0

Table 4.6: Darjeeling District

Darjeeling		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	1	0	0
	Normal	1	8	4
	Above	0	0	6

Table 4.7: East Midnapore District

E Midnapore		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	4	0	0
	Normal	1	8	0
	Above	0	0	7

Table 4.8: Howrah District

Howrah		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	1	0	0
	Normal	3	13	0
	Above	0	0	3

Table 4.9: Hoogly District

Hoogly		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	5	0	0
	Normal	0	12	2
	Above	0	0	1

Table 4.10: Jalpaiguri District

Jalpaiguri		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	7	1	0
	Normal	0	9	2
	Above	0	0	1

Table 4.11: Kolkata District

Kolkata		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	6	0	0
	Normal	0	12	0
	Above	0	0	2

Table 4.12: Maldah District

Maldah		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	10	0	0
	Normal	0	7	1
	Above	0	0	2

Table 4.13: Murshidabad District

Murshidabad		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	8	0	0
	Normal	1	10	0
	Above	0	0	1

Table 4.14: Nadia District

Nadia		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	6	0	0
	Normal	2	10	1
	Above	0	0	1

Table 4.15: North 24 Parganas District

N24P		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	3	0	0
	Normal	0	14	1
	Above	0	0	2

Table 4.16: Purulia District

Purulia		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	5	1	0
	Normal	0	8	0
	Above	0	0	6

Table 4.17: South 24 Parganas District

S24P		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	2	0	0
	Normal	1	13	1
	Above	0	0	3

Table 4.18: Uttar Dinajpur District

U Dinajpur		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	7	0	0
	Normal	2	11	0
	Above	0	0	0

Table 4.19: West Midnapore District

W Midnapore		**Reference**		
		Below	**Normal**	**Above**
Predicted	Below	5	0	0
	Normal	2	9	1
	Above	0	0	3

Conclusion

There are 9 districts that have an increased trend from 1901 to 2020 in West Bengal and 5 districts showed a decreasing trend. In decadal average ISMR, Coochbehar, Jalpaiguri and Darjeeling have increased for the last two decades. The classification model results have shown the categories of rainfall in the forecasting result section. The districts that gained a large percentage of above-normal categories, these districts may affected by floods. The large percentage of rainfall categories below normal that the districts gain may be experienced by drought. A higher percentage of normal categories of rainfall for the districts is a better condition for the agricultural sector.

References

1. Kundu SK, Mondal TK. Analysis of long-term rainfall trends and change point in West Bengal, India. Theoretical and Applied Climatology. 2019;138:1647-1666.
2. Dash SK, Jenamani RK, Kalsi S, Panda SK. Some evidence of climate change in twentieth-century India. Climatic Change. 2007;85(3):299-321.
3. Subash N, Sikka A. Trend analysis of rainfall and temperature and its relationship over India. Theoretical and Applied Climatology. 2014;117:449-462.
4. Verma A, Ponnusamy K, Singh A, Biswas S, Dixit A, Sendhil R. Rainfall pattern analysis to depict the effects of climate change in West Bengal. Environment and Ecology. 2022;40(2):410-417.

5. Alghazali N, Alawadi D, *et al.* Fitting statistical distributions of monthly rainfall for some Iraqi stations. Civil and Environmental Research. 2014;6(6):40-46.
6. Kumar V, Shanu, Jahangeer. Statistical distribution of rainfall in Uttarakhand, India. Applied Water Science. 2017;7:4765-4776.
7. Kulkarni A, Kripalani R, Singh S. Classification of summer monsoon rainfall patterns over India. International Journal of Climatology. 1992;12(3):269-280.
8. Michaelides SC, Pattichis CS, Kleovoulou G. Classification of rainfall variability by using artificial neural networks. International Journal of Climatology. 2001;21(11):1401-1414.
9. Sangiorgio M, Barindelli S, Guglieri V, Biondi R, Solazzo E, Realini E, Venuti G, Guariso G. A comparative study on machine learning techniques for intense convective rainfall events forecasting. In: Theory and Applications of Time Series Analysis: Selected Contributions from ITISE 2019. Springer; 2020. p. 305-317.
10. Raju KHP, Sandhya N, Mehra R. Supervised SVM classification of rainfall datasets. Indian Journal of Science and Technology. 2017;10(15):1-6.
11. Chai SS, Wong WK, Goh KL. Rainfall classification for flood prediction using meteorology data of Kuching, Sarawak, Malaysia: Backpropagation vs radial basis function neural network. International Journal of Environmental Science and Development. 2017;8(5):385-388.
12. Hudnurkar S, Rayavarapu N. Binary classification of rainfall time-series using machine learning algorithms. International Journal of Electrical and Computer Engineering. 2022;12(2).
13. Kumar A. Prediction of annual maximum daily rainfall of Ranichauri (*Tehri Garhwal*) based on probability analysis. Indian Journal of Soil Conservation. 2000;28:178-180.
14. Singh RK. Probability analysis for prediction of annual maximum rainfall of Eastern Himalaya (*Sikkim mid hills*). Indian Journal of Soil Conservation. 2001;29:263-265.
15. Kumar R, Singh RD, Sharma KD. Water resources of India. Current Science. 2005;89(5):794-811.
16. Sharma MA, Singh JB. Use of probability distribution in rainfall analysis. New York Science Journal. 2010;3(9):40-49.