Probability Analysis of Rainfall in Dhemaji Region, Assam, India

Nandana Pegu[1] Prof.R.K. Malik[2]
[1]M. Tech. Student [2]Professor
[1,2]Department of Civil Engineering
[1,2]Amity School of Engineering and Technology, Amity University Haryana, Gurgaon, India

Abstract— The monthly rainfall data of Dhemaji region, Assam, India for the period of 35 years (1980-2014) had been analyzed using probability distribution methods viz Normal, Log normal, Log Pearson-III and Gumbel distribution. These probability distribution methods were applied to estimate the expected monthly, seasonal and annual rainfall for the period of 1980-2014. Weibull`s plotting position was used for computation of observed monthly, seasonal and annual rainfall separately in 1.03, 1.05, 1.11, 1.25, 1.33, 2, 4, 5, 10, 20 and 25 years return periods. The Observed and expected values were compared using Chi Square Test (ℵ^22) test. The result indicated that Gumbel distribution method was the best fit to predict annual rainfall for different return periods. In case of seasonal analysis, Log-Pearson-III was found to be the best fit. However, the best fit probability distribution of monthly data was found to be different for different months. Log-Pearson-III was determined as the best fit probability distribution for the months of June and July. In case of the month of August, the lowest Chi square value (10.24) obtained by Log normal distribution and for the month of September; the best probability distribution selected by Chi square test was Normal distribution. As per Gumbel distribution, annual rainfall of 2437.76mm can be expected with 90% probability in 1.11 year return period while annual rainfall of 4796.47mm can be expected with 1% probability in return period of 100 years. The magnitude of 2960.18 mm annual rainfall with 50% probability can be expected in every 2 years which is approaching to mean annual of 3050mm. Again the seasonal rainfall of 2121.82 can be expected with 50% probability in same return periods as estimated by Log Pearson-III. In monsoon season, the expected rainfall during July is higher as compared to other monsoon months. On this month, precipitation of 295.22mm is expected with 90% probability in 1.11 years. In every 2 years return period, July month is expected more shower of 741.92mm with 50% probability as compared to other monsoon months June (508.9mm), August (533.18 mm) and September (449.19 mm). The highest average rainfall observed in July was 636.81mm .The developed regression model also indicated the best fitted curve of each monsoon months, seasonal and annual scattered observed rainfalls with R^2value more than 0.90, except in case of July, it was 0.78.

Key words: Monthly, Seasonal, Annual Rainfall, Return Period, Frequency, Probability Distribution

I. INTRODUCTION

The liquid form of precipitation is generally termed as rainfall. It is the most essential substance in terms of fresh water and regarding agricultural concern. The magnitude and concentration of rainfall in different durations influences the production of agriculture and thus determines the economy of the country. On the other hand its adverse effect causes flood or draughts which are responsible for loss of fertile lands, life, properties and crops, eventually degradation of overall economy of the country.

In hydrologic science, the importance of analysis of rainfall of a particular region or area is numerous. It helps in estimating of future trend and probable magnitude of rainfall at different time periods. It also determines the intensity of occurrence in different durations and thus it helps in adopting require strategy in project formulation and mitigation of the extreme events. It also especially helps in designing of flood plain management, flood moderation works, earthen embankment constructions, soil & water conservation and crop planning to help the farmer to receive better economic returns(Agarwal et al.,1988). Frequency analysis using statistical probability distribution can help in predicting occurrence of future events. (Kumar and Kumar. 1980). The frequency and magnitude of rainfall in Dhemaji region varies with time and spaces. The strength of rainfall changes from month to month and year to year. In general, the amount of rainfall is high in Dhemaji region based on topography. However, the climate variation due to human activities such massive deforestations in Himalaya regions, construction of roads, bridges and houses without proper planning in the region are also noticeable issues which result in differential magnitude of rainfall and extreme event like flood. Frequency analysis for rainfall data in North east was conducted by many scholars. Deka et al., (2009) analyzed distributions of annual maximum rainfall series of North East India using data for the period of 1966 to 2007 of nine distantly located stations in North east India. They found the goodness of fit test as the most appropriate distribution for describing the annual maximum rainfall series for the maximum no of stations in North east India .Jain et al. (2012) made an analysis of rainfall and temperature trends in Northeast India covered 7 states Assam, Meghalaya, Nagaland, Manipur West Bengal , Sikkim and Gangetic West Bengal. The magnitude of trend in a time series was estimated using nonparametric method, Sens`s estimator (sen1968) and Mann-Kendall test. His analysis also showed trend of rainfall with large variability in magnitude in different seasons and months. In India and abroad, frequency analysis using different probability distributions such as Normal, Log normal, Log pearso-III, Gumbel, Gamma distributions etc. has been performed by many scholars like Moaely et al., 1970; Bhargava et al., 1971; Phien and Ajirajah 1984; Biswas and Khambete 1989; Duan et al.,1995; Elijadid 1999; Topaloglu 2002; Salami 2004; Chin-Yu Lee 2005; Baskar et al., 2006; Kwaku et al., 2007; Hanson et al., 2008; Nemichandrappa et al., 2009. Mohammad et al., 2010; Kumar (2011); Singh et al., 2012; Roy, M 2013; Goyal 2014.

The objective of analysis of rainfall in the present paper, are use of different probability distributions viz Normal, Log normal, Log pearson-III and Gumbel distributions to evaluate the expected monthly (June ,July, August and September), seasonal (total of June to

September) and annual rainfalls of Dhemaji region at different return periods and corresponding probability levels. Secondly, to test the best probability distribution for estimating the expected precipitation in different return periods by Chi square test and to develop regression model between dependent variate (rainfall) and independent variate (months, years) by fitting suitable regression equation to observed data.

II. MATERIALS AND METHODS

Dhemaji region is located in the eastern most part of Assam at $27^0 28' 49''$ & $94^0 32' 58''$ E and bounded by Arunachal Himalaya to its north & east, Lakhimpur district in the west and the Brahmaputra in the south. The region receives highest rainfalls in the month of July during monsoon periods of June to September. Physiographically, the region is divided into foot hill zone near Arunachal Himalaya, flood plain zone near Brahmaputra and low lying swampy area. Topographically it is a very heavy rainfalls and geographically unstable zone. Heavy rainfalls, and accompanying floods, water logging and erosion damages lands, crops and properties every years. The climate of the region is on and average warm and temperate. The total annual rainfall of this region range from 2600mm to 3200 mm and annual average temperature is 23.5^0C. The relative humidity varies from 73 to 90 percent. The minimum temperature receive during winter is 5.9^0C. The four seasons observed in study area are: i)Pre- monsoon season: From March to May, ii)Summer season (Monsoon): From June to September, iii)Post monsoon: October & November and iv) Winter: December to February. The soil is coarse loamy near Himalaya foothills. In flood plain zone, the soil of the area is loamy sand/ sandy loam soils. About 60% of the people in the region under study is agriculture. Main crops is rice and it is influenced by seasonal rainfalls.

A. Data collections and analysis

Monthly rainfall data of Dhemaji region, Assam for the period of 1980-2014 was collected from main district, Dhemaji Water Resource department, Assam for statistical analysis and to estimate the expected monthly(June to September), seasonal and annual rainfalls for different return periods. The statistical parameters such as mean, standard deviation, co-efficient of variation and coefficient of skewness were computed herein in order to describe the variability of 35 years rainfalls of Dhemaji. The occurrence of sample data(rainfall) were represented graphically.

B. Return period

Return period or recurrence interval is the average interval of time within which any extreme event of a given magnitude will be equaled or exceeded at least once (Patra, 2001 & Singh et al., 2012). Weibull`s plotting position formula (Chow, 1964) were used to determine the return periods.

Weibull plotting position formula (Chow,1964) by arranging rainfall data in descending order giving their respective rank as

$$T= (n+1)/m \quad (1)$$

Where, n= total number of years of record,

And m= rank of observed rainfall values arranged in descending order.

Weibull`s plotting position formula was used for evaluation of observed monthly, seasonal and annual rainfall at return periods 1.03, 1.05, 1.11, 1.25, 1.33, 2, 5, 10, 20, 25 years.

C. Frequency factors

Frequency or probability distribution helps to relate the magnitude of extreme hydrological events such as floods, droughts and heavy storms with their number of occurrence such that their chances of occurrence with time can be predicted (Singh et al., 2012). Chow (1951) general frequency formula is used in this study work to get the observed rainfall statistically. The formula used for evaluation of expected values or frequency of occurrence of rainfalls were expressed in term of frequency factor K_τ.

If, X_τ = the expected value of event or rainfall corresponding to return period, T

$$X_\tau = \bar{X} + K_\tau\, \sigma \quad (2)$$

Equation (2) is the frequency factor equation for calculating the expected value of rainfall corresponding to return period, T

Where,

K_τ= frequency factor, $\bar{X}$= arithmetic mean of all the rainfalls in the annual series, σ = standard deviations, T = return period

The Frequency factor, K_τ depends on return period, T and assumed frequency distribution K_τ can be calculated by Normal distribution, Log normal distribution, Log Pearson type-III and Gumbel distribution.

D. Normal distribution

For normal distribution, the Frequency Factor, K_τ is determined by the following relation:(Chow, 1988)

$$K_\tau = (X_\tau - \mu)/\sigma \quad (3)$$

This is the same as the standard normal variate z. The value z corresponding to an exceedance of probability of p, (p=1/T) can be calculated by finding the value of an intermediate variable w:

$$w = [\ln(\tfrac{1}{p^2})]^{1/2}, (0<P \leq 0.50) \quad (4)$$

Calculation of z using the approximation,

$$z = w - \frac{2.515517 + 0.802853w + 0.010328w^2}{1 + 1.432788w + 0.189269w^2 + 0.001308w^3}] \quad (5)$$

When p>0.5, 1-p is substituted for p in equation (4) and the value of z computed by equation no (5) is given a negative sign (Bhakar et.al., 2006). The Frequency factor, K_τ for the normal distribution is equal to z.

1) Log normal distribution

For log normal distribution, the normal values are transformed into logarithmic form and that is Y=lnX, is normally distributed [the value of variate X was replaced by its natural logarithm].Here Y was the magnitude of rainfalls and X was the time in months or years.

$$Y = \ln X \quad (6)$$

The expected value of rain fall `$X_\tau$` at return period, T can be calculated by the following relation:

$$X_\tau = \exp(Y_\tau) \quad (7)$$

$$\text{Where, } Y\tau = \bar{Y}(1 + CvyK\tau) \quad (8)$$

Where, $\bar{Y}$ is the mean and Cvy is the coefficient of variation of Y. and

$$K_\tau = (y_\tau - \mu_\tau)/\sigma \quad (9)$$

The value of Frequency factor, K_τ can be determined by using equation no (5) above.

E. Log Pearson III distribution

At last, the value of variate (expected rainfall), X_τ at different return periods is computed by using the following relationship

$$\log X = \log\overline{X} + K_\tau.\sigma_{\log X} \quad (10)$$

Hence,

$$X_\tau = \text{Antilog}X \quad (11)$$

Where,

$\log\overline{X}$ = mean of logarithmic values of observed rainfall and $\sigma_{\log X}$ = standard deviation of these observed rainfall

K_τ= Frequency factor and it is taken from Benson (1968) corresponding to co-efficient of skewness(Cs) of transformed variate as

$$K_\tau = \frac{2}{C_s}[\{(z-\frac{C_s}{6})\frac{C_s}{6}+1\}^3-1] \quad (12)$$

F. Gumbel distribution

Here, X_τ is determined as

$$x_T = \overline{X} + K\sigma_x \quad (13)$$

$$\text{Or, } X_\tau = \overline{X}(1+CvK_\tau) \quad (14)$$

Where,

X_τ= expected rainfall, $\overline{X}$= mean of observed rainfall, Cv= co-efficient of variation=$\sigma_x/\overline{X}$

K_τ= frequency factor and it can be determined by the following relation,

$$K_\tau = -\frac{\sqrt{6}}{\pi}\{0.5772+\ln[\ln(T/(T-1)]\}. \quad (15)$$

G. Testing goodness of fit of the probability distribution by Chi square test

Chi square test for goodness of fit was used to decide the best fit distribution among those four probability distributions.

$$\aleph^2 = \sum_{i=1}^{k}(O_i - E_i)^2/E \quad (16)$$

Where, O_i is the observed rainfall, E_i is the expected rainfall, i is number of observations (1,2,3…………..k). The Chi degree of freedom (d.f) will equal to (N-k-1).

K= number of classes, N=number of parameters in the theoretical distribution. As per Agrawal et.al. 1988, the best probability distribution function is one which gives lowest Chi square value after comparing Chi square values of different probability distributions.

If X^2cal> X^2tab for (N-k-1),d.f; Then the difference value between observed and expected values is considered to be significant.

H. Regression Model

Linear logarithmic regression model used herein is expressed by

$$Y_i = a + b\log X \quad (18)$$

Where *a*= constant, *b*= predictors (slope)

The least squares method is the most common method for fitting a regression line. This method calculates the best fitting line for the observed data by minimizing the sum of the squares of the vertical deviations from each data point to the line. The ranges of coefficient of determination (R^2)from 0 to 1.

I. Results and Discussion

The monthly rainfall data for 35 years (1980-2014) is classified into three different data sets. These are 1 monthly (i.e. June, July, August and September), 1 seasonal (i.e. total of June to September) and 1 annual (12 months) and is presented in table1. The statistical parameters such as mean, standard deviation, co-efficient of skewness, maximum and minimum rainfall each month (June –Sept), seasonal and annual is shown in table3. It has been observed that maximum rainfall of 1101.7 mm was observed in June(1998). Sharp cyclic increasing and decreasing magnitude of precipitation was observed in every year especially during the month of June, to September. The strength and concentration accelerates from June to July. It remained moderate during August and gradually decreased in September. The average rainfall during monsoon season (seasonal) was about 2182.763mm.

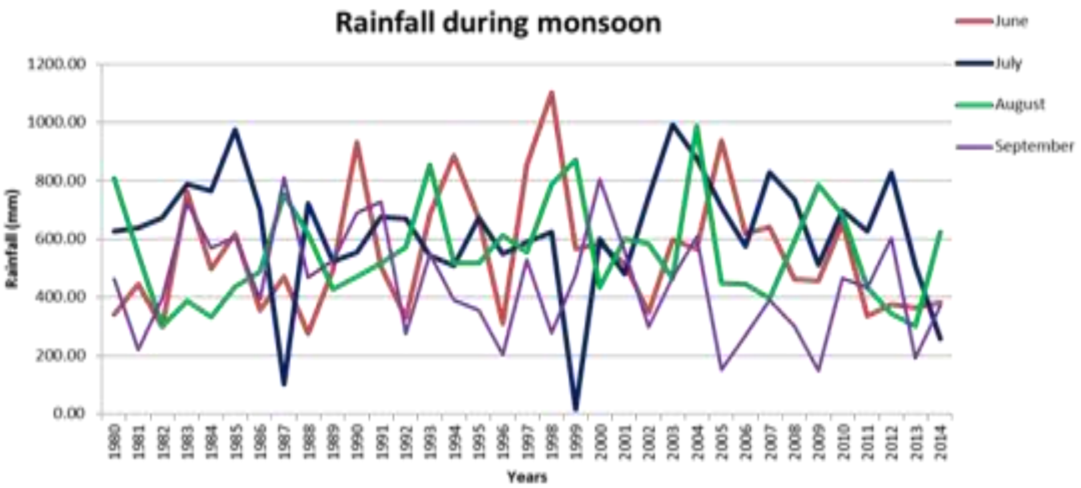

Fig 1: Fluctuation of monthly rainfall of Dhemaji region during (1980-2014)

Heavy precipitation occurred during the month of July. However, there was one occasion in July-1999, where relatively very low precipitation of 16.1mm was recorded during the period of 1980-2014. There was one occasion in June-1998, where highest magnitude of rainfall of 1101.7 mm was observed. The maximum value of co-efficient of variation was observed in September which indicated a high fluctuation in the rainfall data set on that month.

Years	Rainfall(mm)				Seasonal(total of June to Sept)	Annual
	June	July	Aug	Sept	Rainfall(mm)	Rainfall(mm)
1980	340.8	625.7	808.2	462.7	2237.4	2990.5
1981	446.4	638	548.05	221.08	1853.53	2589.57
1982	298.37	673.43	302.19	395.23	1669.22	2339.7
1983	762.87	789.36	388.4	725.9	2666.53	3306.24
1984	496.4	764	333.8	570	2164.2	3083.4
1985	618.2	975.6	437.4	606.1	2637.3	3561.4
1986	355	702.8	488.9	392.5	1939.2	2520.6
1987	470	103.1	754.7	811.1	2138.9	2823
1988	274.4	721.9	619.3	469.7	2085.3	2700.8
1989	494.4	523	428.4	524.9	1970.7	2635.72
1990	932.4	555.8	471.5	690.6	2650.3	3330.9
1991	505.4	676.06	518.5	728	2427.96	3155.76
1992	329.5	671	571.2	275.2	1846.9	2742
1993	685.7	543.8	854.2	548.7	2632.4	3629.4
1994	886.7	510	517.6	389.3	2303.6	3346.8
1995	673	674.3	517.2	356.2	2220.7	3231.4
1996	309.5	548.1	612.5	203.7	1673.8	3026.3
1997	856	590.8	554.8	528.4	2530	3245.3
1998	1101.7	625.3	789.4	277	2793.4	3875.9
1999	567.8	16.1	870.9	476.1	1930.9	2700.3
2000	583.8	600.4	433.4	805.3	2422.9	3055.3
2001	516.8	480	601.5	561.1	2159.4	2901.5
2002	346.2	745.6	584.3	299.7	1975.8	2793
2003	598.7	991.9	466.2	468.2	2525	3531
2004	567.5	878.9	986	606	3038.4	5011.4
2005	939.61	709.66	449.67	152.93	2251.87	3670.12
2006	617.5	576.25	446.2	269.14	1909.09	2934.57
2007	641.19	828.56	395.86	390.81	2256.42	3307.9
2008	462.11	737.35	591.14	301.05	2091.65	2861.83
2009	456.17	511.39	784.7	148.83	1901.09	2544.11
2010	657.23	699.36	685.06	466.79	2508.44	3552.42
2011	336.45	627.91	432.95	432.95	1830.26	2482.66
2012	377.34	827.92	345.92	605.13	2156.31	3132.51
2013	363.57	508.18	300.64	190.45	1362.84	1942.33
2014	382.25	256.73	625.21	370.8	1634.99	2249.13

Table 1: Monthly (June to Sept), seasonal (total June to Sept) and annual rainfall (mm)

Description	Probability density functions	Range	Parameters
Normal	$f(x)=\frac{1}{\sigma\sqrt{2\pi}}.\exp[\frac{-(X-\mu)^2}{2\sigma^2}]$	$-\infty < X < \infty$	X= Variables(Rainfall) $\bar{X}=\mu$,Mean or Average S = σ , Standard deviations
Log Normal(Y=lnX)	$f(x)=\frac{1}{\sigma_y\sqrt{2\pi}}.\exp[-\frac{(y-\mu_y)^2}{2\sigma_y}]$	$0<X<\infty$	$\mu_{y)}=\bar{Y}=Mean$ $\sigma_y=s_y$, Standard deviation
Log Pearson III(Y=lnX)	$f(x)=f_o[(1-\frac{x}{\alpha})^C.\exp(-\frac{C.x}{2})]$ $f_o=\frac{N}{\alpha}.\frac{C^{c+1}}{C^c[c+1)]}$	$0<X<\infty$	**C=4/β-1** $\alpha=\frac{C}{2}(\mu3/\mu2)$ β=(μ3/μ2)^2 **μ2=Variance** **μ3=Third moment about the mean** **Cs= Co-efficient of skewness** **N=length of records** **σ=Standard deviation**
Gumbel distribution	$f(x)=\frac{1}{\alpha}\exp[-((x-\mu)/\alpha)-\exp(-(x-\mu)/\alpha)$	$-\infty < X < \infty$	$\mu=\bar{X}+0.5772\,\alpha$, $\alpha=S\sqrt{6}/\pi$

Table 2: Probability distribution`s density function:

Study Periods	Parameters					
	Mean	Standard deviation	Coefficient of variation	Skewness co-efficient	Maximum	Minimum
June	550.027	208.037	0.378	0.1947	1101.7	274.4
July	636.810	201.999	0.317	-3.8849	991.9	16.10
August	557.597	171.555	0.308	-0.2409	986	300.64
September	449.188	181.251	0.404	-1.7342	811.1	148.83
Seasonal(monsoon)	2182.763	371.160	0.170	-0.3331	3038.4	1362.84
Annual	3051.565	556.516	0.182	0.1936	5011.4	1942.33

Table 3: Statistical parameters of monthly, seasonal and annual rainfall

Serial no	Proba bility	Return periods	Observed rainfall (mm)					
	%	T years	June	July	Aug	Sept	Season al	Annual
1	97	1.03	274.4	16.1	300.64	148.83	1362.84	1942.33
2	95	1.05	293.23	34.74	315.3	152.05	1576.67	2183.39
3	90	1.11	322	304.57	353.05	199.28	1668.89	2435.01
4	80	1.25	349.13	311.77	412.52	266.23	1869.38	2604.95
5	75	1.33	363.57	543.8	433.4	299.7	1909.09	2700.3
6	50	2	505.4	638	518.5	462.7	2159.4	3026.3
7	25	4	657.23	737.35	625.21	570	2508.44	3330.9
8	20	5	721	781.69	712.64	621.25	2513.2	3346.4
9	10	10	887.12	885.56	837.21	721.79	2706.12	3686.6
10	5	20	1009.99	938.55	917.35	783.24	2867.43	4296.02
11	4	25	1038.55	948.76	935	796.54	2908.36	4483.62
12	2	50						
13	1	100						
14	0.5	200						

Table 4: Observed rainfalls at different probability levels.

Expected rainfall(mm) by Normal distribution						Expected rainfall(mm) by Log normal distribution					
June	July	Aug	Sept	Seasonal	Annual	June	July	Aug	Sept	Seasonal	Annual
185.73	272.23	257.19	131.8	1532.82	2077.1	271.36	153.21	313.37	185.35	1587.11	2206.64
222.33	307.77	287.37	163.68	1598.12	2175	291.45	176.63	332.49	202.51	1641.86	2283.94
290.88	374.32	343.89	223.41	1720.42	2358.3	327.04	222.17	365.84	233.61	1734.2	2414.38
377.62	458.55	415.42	298.98	1875.17	2590.4	380.28	300.02	414.61	281.65	1863	2596.5
410.1	490.09	442.21	327.28	1933.12	2677.3	402.64	336.17	434.73	302.32	1914.24	2669
550.03	625.95	557.6	449.19	2182.76	3051.6	514.97	548.75	533.18	410.18	2151.56	3005.19
690.28	762.14	673.26	571.39	2433	3426.8	659.03	896.8	654.24	556.91	2418.98	3384.67
725.08	795.92	701.95	601.7	2495.08	3519.9	700.62	1013.03	688.31	600.81	2490.32	3486.01
816.67	884.86	777.48	681.5	2658.49	3764.9	823.07	1396.14	786.72	733.62	2688.31	3767.52
891.73	957.73	839.37	746.89	2792.39	3965.6	940.14	1819.45	878.48	865.12	2863.59	4016.97
914.32	979.67	858	766.57	2832.69	4026.1	977.27	1965.34	907.16	907.68	2916.75	4092.69
977.38	1040.89	910	821.51	2945.2	4194.8	1091.88	2451.02	994.58	1041.47	3074.49	4317.45
1034.09	1095.96	956.77	870.92	3046.37	4346.5	1206.39	2989.51	1080.37	1178.55	3223.63	4530.12
1085.98	1146.35	999.56	916.13	3138.96	4485.3	1321.68	3585.34	1165.34	1319.74	3366.43	4733.89

Table 5: Estimated expected rainfall at different probability levels by Normal and Log normal distributions

Expected rainfall(mm) by Log Pearson-III distribution						Expected rainfall(mm) by Gumbel distribution					
June	July	Aug	Sept	Seasonal	Annual	June	July	Aug	Sept	Seasonal	Annual
278.28	120.56	305.06	126.94	1557.3	2233.4	251.65	347.09	311.55	189.23	1650.43	2253.39
296.61	177.71	326.36	155.89	1621.18	2303.2	275.92	370.66	331.56	210.37	1693.73	2318.31
329.37	295.22	363.13	209.88	1726.92	2422.6	320.57	414.02	368.38	249.28	1773.39	2437.76
379.09	473.18	416.01	291.93	1870.09	2592.6	379.28	471.02	416.79	300.43	1878.13	2594.8
400.22	538.64	437.5	325.42	1925.63	2661.3	402.6	493.66	436.02	320.74	1919.73	2657.18
508.9	741.92	539.74	473.58	2172.36	2988.2	515.87	603.64	529.43	419.43	2121.82	2960.18
654.47	797.58	658.84	605.63	2430.08	3373.4	658.45	742.09	647.01	543.65	2376.2	3341.61
697.8	798.47	691.09	632.9	2495.34	3479.3	699.64	782.08	680.97	579.54	2449.69	3451.79
828.78	803.13	781.25	691.73	2669.09	3780	821.32	900.22	781.31	685.55	2666.77	3777.28
958.7	833.28	861.56	725.81	2814.24	4054.8	938.03	1013.55	877.56	787.23	2875	4089.5
1000.85	850.78	885.97	733.39	2856.74	4139.7	975.05	1049.5	908.09	819.49	2941.05	4188.54
1133.85	934.96	958.38	749.81	2978.8	4396.2	1089.1	1160.24	1002.14	918.85	3144.53	4493.63
1271	1075.28	1026.74	758.81	3088.89	4644.8	1202.31	1270.16	1095.49	1017.49	3346.51	4796.47
1413.35	1293.19	1091.98	763.29	3189.71	4888.5	1315.11	1379.68	1188.51	1115.76	3547.74	5098.21

Table 6: Estimated expected rainfall at different probability levels by Log Pearson III and Gumbel distributions

Chi Square statistic values in different probability distribution and return periods

Recurrence interval, Years	Months							
	June				July			
	Normal	Log normal	Log-Pearson-III	Gumbel	Normal	Log normal	Log-Pearson-III	Gumbel
1.03	42.33	0.03	0.05	2.06	240.98	122.70	90.51	315.64
1.05	22.61	0.01	0.04	1.09	242.21	113.98	115.02	304.44
1.11	3.33	0.08	0.16	0.01	13.00	30.56	0.30	28.93
1.25	2.15	2.55	2.37	2.40	46.98	0.46	55.06	53.84
1.33	5.28	3.79	3.36	3.78	5.89	128.25	0.05	5.09
2.00	3.62	0.18	0.02	0.21	0.23	14.52	14.56	1.96
4.00	1.58	0.00	0.01	0.00	0.81	28.35	4.55	0.03
5.00	0.02	0.59	0.77	0.65	0.25	52.83	0.35	0.00
10.00	6.08	4.98	4.11	5.27	0.00	186.72	8.46	0.24
20.00	15.68	5.19	2.74	5.52	0.38	426.50	13.30	5.55
25.00	16.88	3.84	1.42	4.14	0.98	525.83	11.28	9.67
Total	119.5684	21.2564	15.0597	25.1234	551.71	1630.70	313.44	725.39

Recurrence interval, Years	August				September			
	Normal	Log normal	Log-Pearson-III	Gumbel	Normal	Log normal	Log-Pearson-III	Gumbel
1.03	7.34	0.52	0.06	0.38	2.20	7.20	3.78	8.63
1.05	2.72	0.89	0.37	0.80	0.83	12.57	0.09	16.17
1.11	0.24	0.45	0.28	0.64	2.60	5.04	0.53	10.03
1.25	0.02	0.01	0.03	0.04	3.59	0.84	2.26	3.89
1.33	0.18	0.00	0.04	0.02	2.32	0.02	2.03	1.38
2.00	2.74	0.40	0.84	0.23	0.41	6.73	0.25	4.46
4.00	3.43	1.29	1.72	0.73	0.00	0.31	2.10	1.28
5.00	0.16	0.86	0.67	1.47	0.63	0.70	0.21	3.00
10.00	4.59	3.24	4.01	4.00	2.38	0.19	1.31	1.92
20.00	7.24	1.72	3.61	1.80	1.77	7.75	4.54	0.02
25.00	6.91	0.85	2.71	0.80	1.17	13.61	5.44	0.64
Total	35.57	10.24	14.35	10.91	17.91	54.96	22.55	51.42

Recurrence interval, Years	Seasonal				Annual			
	Normal	Log normal	Log-Pearson-III	Gumbel	Normal	Log normal	Log-Pearson-III	Gumbel
1.03	18.85	31.69	24.28	50.11	8.74	31.66	37.94	42.94
1.05	0.29	2.59	1.22	8.09	0.03	4.43	6.24	7.85
1.11	1.54	2.46	1.95	6.16	2.49	0.18	0.06	0.00
1.25	0.02	0.02	0.00	0.04	0.08	0.03	0.06	0.04
1.33	0.30	0.01	0.14	0.06	0.20	0.37	0.57	0.70
2.00	0.25	0.03	0.08	0.67	0.21	0.15	0.49	1.48
4.00	2.34	3.31	2.53	7.36	2.68	0.85	0.54	0.03
5.00	0.13	0.21	0.13	1.65	8.55	5.59	5.08	3.22
10.00	0.85	0.12	0.51	0.58	1.63	1.74	2.31	2.18
20.00	2.02	0.01	1.01	0.02	27.53	19.38	14.35	10.43
25.00	2.02	0.02	0.93	0.36	52.00	37.34	28.56	20.79
Total	28.61	40.47	32.78	75.10	104.13	101.71	96.19	89.66

Table 7: Chi square values for different probability distribution and return periods.

The observed rainfall is presented in table 4 and expected rainfall in different probability levels and corresponding return periods evaluated by Normal & Log normal, and Log Pearson-III & Gumbel distributions are presented in table, 5 and 6 respectively. The observed rainfalls calculated by plotting position at different return periods are graphically represented in fig.2, fig.3 respectively. The logarithmic regression values of monthly, seasonal and annual values are more than 0.9 (as shown in figs) except in case of the month of July is 0.7816.

The Chi square values presented in table 7 reveal that monthly (June to September), seasonal and annual rainfall followed the different distribution functions. According to Chi- square values (table-7), that Gumbel distribution is the best fit method for evaluation of expected annual rainfall. In case of seasonal analysis, Log-Pearson-III was found to be the best fitted. However, the best fit probability distribution of monthly data was found to be different for different months. Similar results were found by Singh, M.A. and Singh, J.B (2010) in their study of probability distribution in rainfall analysis of Pantnagar, India. Log-Pearson-III was determined as the best fitted probability distribution for the months of June and July. In case of the month of August the lowest Chi square value (10.24) obtained by Log normal distribution. However, Gumbel distribution was also acceptable for evaluation of expected rainfall of the month of August as the Chi value here found 10.911.For the month of September; the best fitted distribution determined was Normal distribution.

The annual rainfall of 4796.47mm can be expected with 1% probability in return period of 100 years. The magnitude of 2960.18 mm annual rainfall with 50% probability can be expected in every 2 years which is approaching to mean annual of 3050mm. As estimated by Log Pearson-III, the seasonal rainfall of 3088.89mm can be expected in 100 years with 1% probability and in every 2 year return periods, the magnitude of rainfall of 2121.82 is expected with 50% probability.

As per Chi square test and Log Pearson-III distribution, during the month of June, rainfall of 1271mm is expected with 1% probability in 100 years and in every 2 years return period with 50% probability, rainfall of 508.9mm can be expected. Using the same probability distribution, rainfall of 1075.29m is expected in 100 years with 1% probability and 50% probability of 741.92mm rainfall can be expected in every 2 years in the month of July. As compared to other months, the strength of rainfall was found higher during the month of July. The highest average rainfall of observed during the month of July was 636.81mm.

As selected, the log normal as the best fit probability distribution for the month of August, rainfall of 533.18 mm can be expected in every 2 years with 50% probability which is approaching to its average rainfall of 557.597mm. Normal distribution was found to be the best fit probability distribution for the month of September. As estimated, the expected rainfall in the month of September in every 2 year return periods is 449.19 mm with 50% probability. With 1% probability, rainfall of 870.92mm rainfall can be expected in 100 return periods during the month of September. Bhakar et al., 2006 recommended that amount of rainfalls expected in 2 to 100 years is sufficient return periods for soil and water conservation measures, construction of dams, irrigation and drainage works.

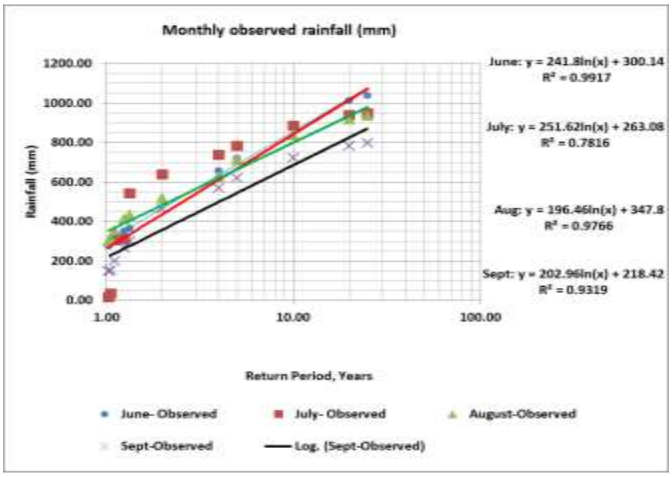

Fig. 2: Monthly observed rainfall at different return periods.

Logarithmic transformed linear regression model was developed from observed monthly, seasonal and annual rainfalls with different return periods. It had been found that the coefficient of determination R^2 for the months of June, August and September were 0.9917, 0.9766, 0.9319 respectively. But for the month of July, R^2 was found to be 0.7819. For seasonal and annual the value of R^2 were found to be 0.902 and 0.9601 respectively. These results indicated that about more than 90% of observed rainfalls are explainable by the estimated return periods except it was 78% in case of July.

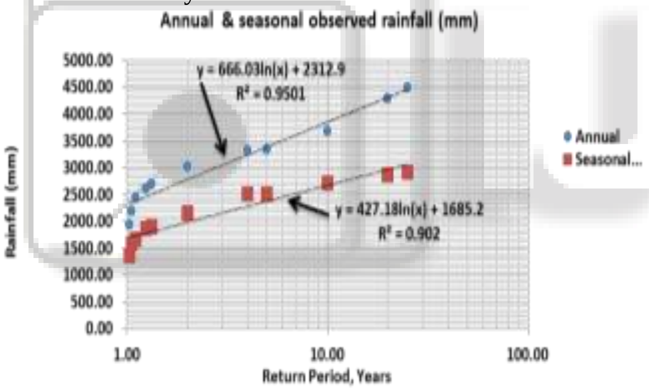

Fig 3: Annual & seasonal observed rainfall at different return periods.

III. CONCLUSION

The rainfall data of Dhemaji for the period of 1980-2014 had showed the large fluctuation of rainfall every months and years where its magnitude ranged from zero to maximum 1101.7 mm (1998). It had showed that the best fit probability distribution for each monsoon months, season and annual rainfalls were different. As per Chi-square test, Annual rainfall followed the Gumbel distribution while seasonal followed the log Pearson-III distributions. Besides, the best fit probability distribution were different for each months except the month June and July followed the Log Pearson-III distribution. Logarithmic transformed linear regression model have indicated R^2more than 0.9 in all months, season, annual except for the month of July is equal to 0.78. This analysis and prediction of expected monthly(June to Sept), seasonal and annual rainfall would help in water conservation measures, construction of dams, irrigation, adaptation of cropping plan and drainage works of Dhemaji region.

ACKNOWLEDGEMENT

I would like to acknowledge Executive Engineer Water Resource department of Dhemaji, Assam, India for providing the data for analysis.

REFERENCES

[1] Spiegel, M. R. and R. W. Boxer (1981). Schaum's outline of theory and problems of statistics in SI units. McGraw-Hill, New York.
[2] Final technical report to U.S. department of Interior Washington, D.C. 20240, Project no. 374103, Grant agreement no. G-866 (1984). Statistical analyses of spatial and temporal storm rainfall characteristics in Puerto Rico.
[3] Vogel, R. and D.E. McMartin (1991). Probability plot goodness of fit and skewness estimation procedures for the Pearson type-3 distribution. Water Resource Research, (27): 3149-3158.
[4] Davidson, R. and J. G. Mackinnon (1999). Regression models. Chapter 1.
[5] Tahal, CES, Org & Jps. How to analyze rainfall data. Hydrology project training module SWDP-12 (2002). Hydrology project World Bank & Government of Netherlands funded, (12): Doc, version 17/09/02.
[6] Kripalini, R.H. and P. Kumar (2004). Northeast monsoon rainfall variability over South peninsular India vis- `a- vis the Indian ocean dipole mode. International Journal of Climatology, (24): 1267-1282.
[7] Lee, C (2005). Application of rainfall frequency analysis on studying rainfall distribution characteristics of Chia- Nan Plain area in Southern Taiwan. Journal of Crop, Environment & Bioinformatics, (2): 31-38.
[8] Bhakar, S.R., A.K. Bansal, N. Chhajed and R.C. Purohit (2006). Frequency analysis of consecutive days maximum rainfall at Banswara, Rajasthan, India. ARPN Journal of Engineering and Applied Science, (1): 64-67.
[9] Suresh, R (2008). Watershed Hydrology. Standard publishers distributors, Delhi, India.
[10] Campbell, D. and S. Campbell (2008). Introduction to regression and data analysis. StatLab Workshop, pp: 1-15.
[11] Deka, S., M. Borah and S.C. Kakaty (2009). Distributions of annual maximum rainfall series of North east India. European Water, 27/28: 3-14.
[12] Longobardi, A. and P.Villani (2009). Trend analysis of annual and seasonal rainfall time series in the Mediterranean area. International Journal of Climatology, DOI: 10.1002.
[13] Mfc bulletin (2009). Flood and erosion-led displacement in Dhemaji and Dhakuwakhana: concerns over health entitlements of the people, R.N. 27565/76.
[14] Sharma, M.A. and J.B. Singh (2010). Use of probability distribution in rainfall analysis. New York Journal, 3(9).
[15] Hanson, L.S. and R. Vogel. The probability distribution of daily rainfall in the United States.
[16] Githui, F. W., A. Opere. and W. Bauwns. Staistical and trend analysis of rainfall and river discharge: Yala river basin, Kenya.

[17] Nemichandrappa, M., P. Ballakrishnan and S. Senthilvel (2010). Probability and confidence limit analysis of rainfall in Raichur region. Karnataka Journal of Agricultural Science, 23(5): 737-741.
[18] Chow, V.T., D. R Maidment and L.W. Mays (2010). Applied Hydrology. McGraw- Hill International Editions, New York.
[19] Benoft, K (2011). Linear regression models with logarithmic transformations. Methodology Institute, London School of Economics.
[20] Singh, B., D. Rajpurohit., A. Vasishth. and J. Singh (2012). Probability analysis for estimation of annual one day maximum rainfall of Jhalarapatan area of Rajasthan, India. Plant Archives, (12): 1093-1100.
[21] Jain, S.K., V. Kumar and M. Saharia (2012). Analysis of rainfall and temperature trends in north east India. International Journal of Climatology, DOI: 10.1002: 968-978
[22] Roy, M (2013). Time series, factors and impacts analysis of rainfall in north-eastern part in Bangladesh. International Journal of Scientific and Research publications, (3): Issue 8.
[23] Baishya, S.J. (2013). Vulnerability assessment and management of flood hazard in Baralia river (Bhairatolajan): a case study of Khopanikuchi village of Hajo revenue circle, Kamprup district, Assam. International Journal of Innovative in Science, Engineering Technology, (2): Issue 9.
[24] Das, L.M. and Z. Hussain Quershi (2014). Flood frequency analysis for Jiya Dhol rivere of Brahmaputra. Intenational Journal of Scienes, 14(2): 14-24.
[25] Aaranyak.org. (2015). Adjusting to floods on the Brahmaputra plains. The International Centre for Integrated Mountain Development, ICIMOD, Web. 31 May.
[26] Dhemaji.gov.in. (2015). Dhemaji at a glance. Web.31 May.
[27] Majumdar, P.P. Stochastic Hydrology. Indian Institute of Science, lecture-27.
[28] Phukan and S. Dev (2015). Floods- the annual mayhem in Assam.
[29] http://dhemaji.nic.in/physiography.htm. Physiography of Dhemaji District.
[30] Wikipedia, the free encyclopedia: http://en.wikipedia.org/wiki/Time_series.