International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 4 Issue 09, September-2015

Prediction of Maximum Monthly Rainfalls for Different Return Periods using Frequency Factor Analysis for Dhemaji Region in Assam, India

Nandana Pegu
[1] Former Post-Graduate Student ,
Department of Civil Engineering,
Amity School of Engineering and Technology,
Amity University Haryana, Gurgaon, India.

R. K. Malik
Professor of Hydrology and
Water Resources Engineering and Head,
Department of Civil Engineering,
Amity School of Engineering and Technology,
Amity University Haryana, Gurgaon, India.

Abstract **- Maximum monthly rainfall data of Dhemaji region in Assam, India for a period of 35 years (1980-2014) were analyzed. The expected maximum monthly rainfalls by using the frequency factor approach for probability distributions viz: Normal, Log-normal, Log-Pearson-III and Gumbel distributions for return periods of 2, 5, 10, 25, 50, 100 and 200 years were evaluated. The Weibull's plotting position method was used for estimating the observed rainfalls for different return periods. The observed and expected maximum rainfalls data were then compared by using goodness of fit by Chi-square test (χ^2). It was observed that the Log-normal distribution gave the best fit to predict the maximum monthly rainfall of Dhemaji region for different return periods. As per, Log-normal distribution, the maximum monthly rainfall of 760.0 mm is expected with 50% probability for a return period of 2 years and. maximum monthly rainfall of 1154.0 mm is expected with 1% probability for a return period of 100 years. The observed maximum monthly rainfalls were also predicted as a function of return period with the linear logarithmic regression model with coefficient of determination of 0.9092.**

Key words: Maximum monthly rainfall (MMR), Probability distribution, Return period, Dhemaji region.

INTRODUCTION

The heavy rainfalls and the accompanying floods, waterlogging, soil erosion and sedimentation are a perennial problems of Assam state in India as a whole and of Dhemaji region in particular. These problems result in severe damages of thousands of hectares of fertile lands, crops ,hydraulic structures and other infrastructures. The topography, global warming due the heavy deforestation in this Himalayan region, construction of the roads, bridges and the houses have also been contributing to the change of climate of this region and resulting in variations in magnitudes of rainfalls and extreme events like floods. It is, therefore, necessary to analyze the rainfall in this region which would enable to forecast the extreme hydrological events. In order to predict the probable frequency of rainfalls for different return periods, design engineers and hydrologists require maximum daily or maximum monthly rainfall for the design of small or big dams , bridges and proper crop planning (Agarwal et al., 1988, Nemichandrappa et al., 2010). Ray et al. (2013) performed the evaluation of annual maximum rainfall for Central Meghalaya (North -East), India using Normal, Log- normal, Pearson Type III and Gumbel type-I probability distributions. The analysis revealed that Gumbel distribution was the close fit to estimate the probable rainfall in Central Meghalaya for different return periods. Many other investigators had also studied the rainfall distributions in India and abroad by using frequency (probability) analyses using different probability distributions such as Normal, Log -normal, Log Pearson-III, and Gumbel, distributions by Bhargava et al., 1971; Phien and Ajirajah 1984; Duan et al.,1995; Elijet al. 1999; Topaloglu 2002; Baskar et al., 2006; Nemichandrappa et al., 2010. 2014; Jain et al., 2012; Sarkar et al., 2013; Roy,2013 Gilletal ,2013; Shiv Prasad,2015 and Pegu and Malik, 2015. The detailed review was presented by Jain and Kumar (2012). Annual maximum rainfall series of North East India using rainfall data for the period of 1966 to 2007 of nine distantly located stations was analyzed by Deka et al. (2009). Their analyses indicated that the most appropriate test for fitting the annual maximum rainfall series data for the maximum number of stations in North East India was the goodness of fit test.

This study was undertaken to predict the expected maximum monthly rainfall of Dhemaji region in Assam, India by using different probability distributions viz Normal, Log -normal, Log- Pearson-III and Gumbel distributions for different return periods and corresponding probability levels, to test the fitting performance of these distributions by Chi- square test and further to develop the suitable regression model. This analysis of prediction of expected maximum monthly rainfall would provide necessary information to the project planners and help in planning of water conservation measures, designing and construction of storage structures, planning of cropping strategies and drainage works of this region.

MATERIALS AND METHODS

The study region Dhemaji is situated in the eastern most part of Assam at $27^0 28' 49'' N$ and $94^0 32' 58''$ E. It is an area surrounded by Arunachal Himalaya to its north and east, and the Brahmaputra river towards its south. The highest magnitude of rainfalls occurs in this region in the month of July during the monsoon periods of June to September. The topography of the Dhemaji region is undulating. The gradient of this drainage system is high in north and eastern part and gradually decreases towards the western side and eventually meet the Brahmaputra river and its tributaries. The statistical analysis of maximum monthly rainfall data for the period of 35 years (1980-2014) of Dhemaji region in Assam, India was performed to predict the probable maximum monthly rainfall (MMR) for different return periods based on the theoretical probability distributions. The four most commonly used probability distributions viz Normal, Log - normal, Gumbel and Log Pearson-II distributions were used for the analysis. The best probability distribution was evaluated by using Chi- square test. The computed statistical parameters such as mean, standard deviation, co-efficient of variation and coefficient of skewness were used to judge the variability of maximum monthly rainfalls of Dhemaji region.

The monthly rainfall data of Dhemaji region for the peiod of 35 years (1980-2014) was collected from the Office of the Executive Engineer, Water Resources Department, Dhemaji, Assam. The magnitudes of the maximum monthly rainfalls were taken from these data for the purpose of analysis. Return period or recurrence interval (T) which is the interval of time within which any extreme event of a given magnitude will be equaled or exceeded at least once was calculated using Weibull`s plotting position method (Chow et al., 2010) by arranging rainfall data in the descending order and giving their respective rank as :

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

Where n is the total number of years of rainfall record and m is the rank of observed rainfall magnitudes arranged in the descending order. If P is the probability of exceedance of rainfall, then

$$P = 1/T \qquad (2)$$

Weibull`s plotting position method was used for the evaluation of the observed maximum monthly rainfall in mm for the return periods of 2, 5, 10, 25, 50, 100 and 200 years.

Frequency factor method for probability distributions:

Frequency or probability distributions help 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). The formula used for the evaluation of the expected values or frequency of occurrence of rainfalls were expressed in term of the frequency factor K_τ.If X_τ is the expected value of an event or rainfall corresponding to the return period T, then

$$X_\tau = \bar{X} + K_\tau\, \sigma \qquad (3)$$

Equation (3) is the frequency factor equation for calculating the expected value of the rainfall corresponding to the return period, T and $\bar{X}$ is the arithmetic mean of all the rainfalls in the annual series, σ is the standard deviations and T is the return period for the occurrence of an event of a particular magnitude. The frequency factor K_τ, depends on the return period T and the assumed frequency distribution. The values of for the K_τ were calculated for the Normal, Log- normal, Log Pearson type-III and Gumbel distributions.

Normal distribution

For the normal distribution, the frequency factor, K_τ was determined by the following relation as proposed by (Chow et al., 2010).

$$K_\tau = (X_\tau - \mu)/\sigma \qquad (4)$$

This is the same as the standard normal variate z as the frequency factor, K_τ for the normal distribution is equal to z. The value z corresponding to an exceedance of probability of p, (p = 1/T) was calculated by finding the value of an intermediate variable w as:

$$w = [(\tfrac{1}{p^2})]^{1/2}, \qquad (0 < P \leq 0.50) \qquad (5)$$

The value of z was estimated using the approximation as:

$$Z = w - \frac{2.515517 + 0.802853w + 0.010328w^2}{1 + 1.432788w + 0.189269w^2 + 0.001308w^3} \qquad (6)$$

When p > 0.5, 1-p is substituted for p in equation (5) and the value of the z computed by equation (6) is given negative sign (Bhakar et al., 2006).

Log- normal distribution

For the log- normal distribution, the values were transformed into logarithmic form and i.e.

$$Y = \ln X \qquad (7)$$

The expected value of the rainfall `$X_\tau$` at the return period T was calculated as :

$$X_\tau = \exp(Y_\tau) \qquad (8)$$

$$Y\tau = \bar{Y}\,(1 + Cvy\, K\tau) \qquad (9)$$

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

$$K_\tau = (y_\tau - \mu_\tau) / \sigma \quad (10)$$

The value of the frequency factor K_τ was determined by using equation (10) .

Log Pearson III distribution

The value of the variate (expected rainfall) X_τ at the different return periods was computed by using the relationship as :

$$\log X = \log \bar{\boldsymbol{X}} + K_\tau.\sigma_{\log X} \quad (11)$$

Where log $\bar{\boldsymbol{X}}$ is the mean of the logarithmic values of the observed rainfall and $\sigma_{\log X}$ is the standard deviation of these observed rainfall values. K_τ is the frequency factor and it was taken corresponding to the co-efficient of skewness (Cs) of the transformed variate (Benson,1968) as

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

Gumbel distribution

For this distribution, the value of the X_τ was determined as:

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

or

$$X_\tau = \bar{\boldsymbol{X}}\,(1 + Cv\, K_\tau) \quad (14)$$

Where X_τ is the expected rainfall, $\bar{\boldsymbol{X}}$ **is** the mean of the observed rainfall and Cv is the co-efficient of variation which is equal to $\sigma_x/\bar{\boldsymbol{X}}$. The frequency factor K_τ was determined by the following relation as:

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

Testing goodness of fit of probability distribution

The Chi-square test for goodness of fit was used to evaluate the best fit distribution among these probability distributions as:

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

Where O_i is the observed rainfall, E_i is the expected rainfall, i is number of observations (1, 2, 3…………..k). The degree of freedom (df) is equal to (N-k-1). K is the number of classes; N is the number of parameters in the theoretical distribution. Agrawal et al. (1988) reported that the best probability distribution function is one which gives lowest Chi-square value after comparing the Chi-square values for the different probability distributions. If $\chi^2_{Cal} > \chi^2_{tab}$ for (N-k-1) df, then the difference between the observed and expected values is considered to be significant.

Regression Model

The linear logarithmic regression model (Benoft, 2011) used herein is expressed as:

$$Y = a + b \log X \quad (17)$$

Where Y is the expected maximum monthly rainfall (mm) , X is the return period in years and a and b are the constants which were determined by the least squares method that gives the best fitting line for the observed data by minimizing the sum of the squares of the deviations between the estimated and the observed magnitudes of the rainfalls for each data point. The coefficient of determination (R^2) which is ratio of the explained variation to the total variation was determined as:

$$R^2 = \frac{\sum(Y_{est} - \bar{Y})^2}{\sum(Y - \bar{Y})^2} \quad (18)$$

Where Y_{est} and $\bar{Y}$ are the estimated and the average values of the dependent variable Y, respectively.

RESULTS AND DISCUSSION

The maximum monthly rainfall (MMR) data in mm extracted from 35 years rainfall data of Dhemaji region was analyzed using the probability distributions discussed above. The year-wise fluctuating trend of the maximum monthly rainfall are shown in Fig.1. The maximum rainfall of 1101.7 mm was observed in June, 1998 and the least maximum rainfall of 508.18 mm was recorded in July, 2013.

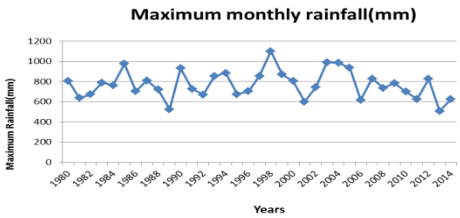

Fig1: Year-wise maximum monthly rainfall (mm)

The maximum rainfall was observed in the month of July though its strength and concentration started accelerating from the month of May and then June to July. Again its magnitude gradually decreased from August to September. The statistical parameters viz mean, standard deviation, co-efficient of variation and co-efficient of the skewness, are shown in Table 1. The statistical parameters presented in Table 1 were used for determination of probable MMR (mm) for the different probability distributions. The average rainfall during the monsoon season (total of June to September) was observed about 2182.7 mm. The average maximum monthly rainfall during the period of 1980-2014 was recorded 771.8 mm.

The observed MMR (mm) for the period of 35 years was predicted against different return periods are shown in Fig.2. The expected MMR (mm) for the different return periods of 2, 5, 10, 25, 50,100 and 200 years and their corresponding probability levels as analyzed by the Normal, Log- normal, Gumbel and LogPearsn-III distributions are presented in Table 2. The expected MMR (mm) for the different probability distributions are shown in Fig. 3. It was observed that the trend of the expected rainfall were the same as the observed rainfall. The best fit probability distribution among these probability distributions were determined by the Chi-square goodness of fit test. The results of the Chi-square test are shown in Table 3.

Table 1. Summary of statistical parameters for maximum monthly rainfall (mm)

Mean	771.8
Standard deviation	137.1
Coefficient of variations	0.177
Skewness coefficient	-0.17

Table 2: Observed and expected MMR (mm) at different return periods and probability levels

Probability, %	Return period, years	Observed rainfall (mm)	Expected rainfall (mm)			
			Normal	Log-normal	Gumbel	Log-Pearson-III
50	2	764.0	771.8	760.0	749.3	763.8
20	5	883.3	885.6	883.9	870.5	885.0
10	10	979.6	945.1	956.6	950.7	963.1
4	25	1054.6	1008.6	1040.7	1052.1	1029.4
2	50		1049.6	1198.8	1127.3	1080.6
1	100		1086.5	1154.0	1202.0	1127.9
0.5	200		1120.2	1206.8	1276.4	1172.4

Table 3: Chi-Square values at different probability levels for different probability distributions

Probability, %	Return period, years	Chi-Square values			
		Normal	Log-normal	Gumbel	Log-Pearson-III
50	2	0.080	0.021	0.287	0.000
20	5	0.006	0.000	0.187	0.004
10	10	1.259	0.556	0.879	0.739
4	25	2.097	0.187	0.006	0.617
2	50				
1	100				
0.5	200				
Total		3.442	0.764	1.359	1.359

The Chi-square test revealed that the log -normal distribution was found to be the best fit distribution with minimum Chi-square value of 0.764 as compared to the other distributions viz Normal, Gumbel and Log -Pearson-III distributions with Chi- square values of 3.442, 1.359 and 1.359, respectively. The analysis performed by log-normal distribution showed that the expected maximum monthly rainfall for the return period of 2 years with 50% probability
was 760.0 mm and the rainfall of 883.9, 956.6, 1040.7, 1098.8, 1154.0, 1206.8 mm were expected for the return periods of 5, 10, 25, 50,100 and 200 years, respectively. The probable maximum rainfall of 760.0 mm in return period 2 years was the most expected maximum monthly rainfall in the area under study which is approaching to its average rainfall of 771.8 mm.

The logarithmic transformed regression model (Fig. 2) was developed from the observed maximum monthly rainfall magnitude (Y) as a function of the return period (X) and is described as :

$$Y = 627.72 + 151.57 \ln (X) \quad (19)$$

The coefficient of determination (R^2) for the regression model relating maximum monthly rainfall
with return period was estimated to be 0.9092 which showed that there is a very close relationship.

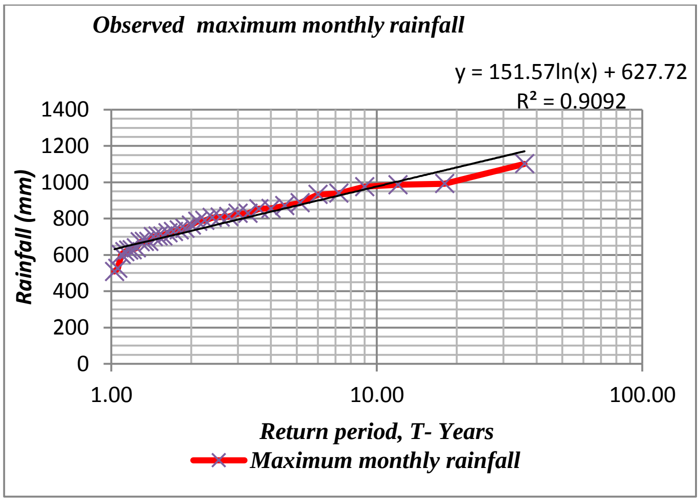

Fig. 2: Observed maximum monthly rainfall (mm) with return period in years (Weibull method).

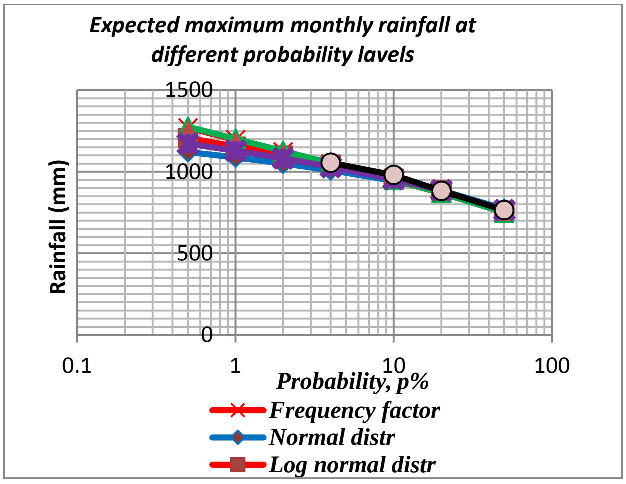

Fig. 3: Expected maximum monthly rainfall (mm) at different probability levels.

CONCLUSION

The maximum monthly rainfall (MMR) data of Dhemaji region in Assam, India for the period of 35 years (1980-2014) indicated large variations of rainfall every month and year whereas the highest magnitude of 1101.7 mm and the lowest MMR of 508.1 mm were observed in 1998 and 2013, respectively. It was observed that the MMR data of Dhemaji region fitted best to the Log-normal distribution. The probability analysis of the MMR of this region performed by log- normal distribution showed that the probable maximum monthly rainfall of 760.01 mm is expected in every 2 years return period with 50% probability and MMR of 1154 mm is expected in 100 years with 1% probability. The coefficient of determination for the linear logarithmic regression model for predicting the maximum monthly rainfall as a function of return period was estimated to be 0.9092.

ACKNOWLEDGEMENTS

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

REFERENCES

[1] Aggarwal, M.C., V.S. Katiyar and Ram Babu (1988) Probability analysis of annual maximum daily rainfall of U.P, Himalaya. Ind. J. Soil. Conservation. 16 (1): 35-42.
[2] Benoft, K (2011). Linear Regression Models with Logarithmic Transformations. Methodology Institute, London School of Economics.
[3] Benson, M.A. (1968) Uniform flood frequency estimating method for federal agencies. Water Resour. Res. 4 (5): 891-908.
[4] 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 J. Engineering and Appl. Sci. 1: 64-67.
[5] Bhargava, P.N. (1974) The influence of rainfall on crop production. Res. J. Jawaharlal Nehru Krishi Vishwavidyalaya. 32 (1 & 2) : 88-98.
[6] Chow, V.T., D. R Maidment and L.W. Mays (2010). Applied Hydrology. McGraw- Hill International Editions, New York.
[7] .Deka, S., M. Borah and S.C. Kakaty (2009). Distributions of annual maximum rainfall series of North east India. European Water. 27: 3-14.
[8] Duan, J., A.K. Sikka and G.E. Grant (1995) A comparison of stochastic models for generating daily precipitation at the H.J. Andrews Experiment Forest. North-West Sci. 69 (4) : 318-329.
[9] Gill. K.K., S.S.Kukal, S.S. Sandhu, and H. Brar (2013) Spatial and temporal variation of extreme rainfall events in Central Punjab. Intern. J. Appl. Engg. Res. 8 (15) : 1757-1764.
[10] Jain, S.K. and V. Kumar (2012) Trend analysis of rainfall and temperature data for India. Current. Sci. 102 (1) : 37-49.
[11] Jain, S.K., V. Kumar and M. Saharia (2012). Analysis of rainfall and temperature trends in north east India. Intern. J. Climatology, DOI: 10.1002: 968-978.
[12] Nemichandrappa, M., P. Ballakrishnan and S. Senthilvel (2010). Probability and confidence limit analysis of rainfall in Raichur region. Karnataka J. Agric.Sci. 23 (5): 737-741.
[13] Ray, L.I.P., P.K. Bora., V. Ram., A.K.Singh., N.J.Singh., R. Singh and S.M.Feroze (2013). Estimation of annual maximum rainfall for Central Meghalaya. Indian Journal of Hill Farming, 26(1):47-51.
[14] Roy, M (2013). Time series, factors and impacts analysis of rainfall in north-eastern part in Bangladesh. Intern. J. Scientific and Research 3 (8) : 267-282.
[15] Sarkar. S., D.K.Kundu and B.S.Mahapatra(2013). Rainfall probability analysis of the Western Odisha Plateau Region for Sisal. Journal of Agricultural Physics 13(1): 62-70.
[16] 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.
[17] Swami Shivprasad, S. T., A. Desmukh., G. Patil and S. Kahar(2015). Probability distribution analysis of rainfall data for Western Maharashtra region. Intern. J. Sci. and Res. 4 (5) : 1375-1379.
[18] Topalogu, F. (2002) Determining suitable probability distribution models for flow and precipitation series of the Seyhan River Basin. Turk. J. Agric. 2 (26) : 189-191.
[19] Pegu, N and R.K.Malik (2015). Probability analysis of rainfall Dhemaji region, Assam, India. Intern. J. Scientific Res. and Dev. 3 (4) : 3387 -3394.
[20] Phien, H.N. and T.J. Ajirajah (1984) Applications of Log-Pearson type-3 distributions in hydrology. J. Hydrology. 73 : 359-372.