Chapter 1

Introduction

Agriculture plays a vital role for the economic development of North East India. Every year most part of North East India has been effected by flood caused by heavy rainfall which damages crops and properties of people. Rainfall has a direct impact in the economy of this region. Also the study of extreme rainfall is very much useful for design of dam, road, bridge and hydrological planning. So proper estimation of extreme rainfall is necessary for economic development of this region. Tao et al. [47] reported that several probability models have been developed to describe the annual peak rainfall at single site. However, the choice of a suitable model is still one of the major problems in engineering practice since there is no general agreement as to which distribution or distributions should be used for the frequency analysis of extreme rainfalls. The selection of appropriate model depends mainly on the characteristics of available rainfall data at the particular site. Hence it is necessary to evaluate many available distributions in order to find a suitable model that could provide accurate extreme rainfall estimates.

1.1 Regional Frequency Analysis

If event frequencies are similar for the different observed quantities, then more accurate conclusions can be reached by analyzing all of the data samples together than by using only a single sample. In environmental application this approach is known as regional frequency analysis, because the data samples analyzed are typically observations of the same variable at a number of measuring sites within a suitably defined region (Hosking and Wallis [25]).

Generally, the estimation of extreme rainfall for a given frequency have been done by fitting a probability distribution to observations at a single site. In this approach the estimates suffer from sampling variability especially for estimation of return periods that exceeds the length of the observations (Hosking and Wallis [24, Cunnane [13]). Several studies (e.g. [30], [25]) have shown that regional frequency analysis gives more accurate estimates than at site frequency analysis. Deka S. et al. [17] analysed maximum rainfall data of 9 stations of North East India using at site procedure, but regional rainfall frequency analysis for this region is yet to be done.

1.2 Background of the Study

Natural disaster and severe weather events have a close link because all severe weather events, due to climate change or otherwise could and often lead to natural disasters that occur on varying time and space scales. Disaster may strike any country but the greatest burden falls on less developed countries and their highly populated regions. One of the common features of developing countries in the South East Asia is flash flood in urban areas during rainy season and acute shortage of water for domestic and agriculture uses during winter (Deka S. [15])

North East India, located at 80^0E and 21^0N, is one of the major disaster prone region of India because of their unique geographical locations and physical features. During summer major flood occur that often lead to disaster.

Statistical analysis of extreme events is usually done for predicting large return period. Thus the more relevant analysis is the upper quantiles of the distributions and the extreme samples (Wang [51], Shabri and Jemain [46]).

There are various parameter estimation techniques available in hydrological literature. Some of the commonly used methods are methods of maximum likelihood (MLM), method of moment (MOM) etc. However, all the methods are not suitable for all application also depend on the situation and data availability. In case of maximum rainfall frequency, data are not available for longer periods due some technical problems and measurement error. The MOM is mostly used because of its relative ease of application but it is generally not as efficient as MLM and it is too sensitive to upper quantiles of distribution (Vogel and Fennessey [50]). MLM sometimes under estimates and so causes large bias and variance of extreme upper quantile and does not always work well in small samples (Park [39]). Greenwood et al. [20] defined probability weighted moment (PWM) which gives more accurate parameter estimation from small samples than MLM and MOM.

Hosking [23] introduced L-moment method which is a linear combination of PWMs. L-moment method gives more accurate parameter estimation of a probability distribution than PWMs. L-moments can be defined for any random variable whose mean exists. L-moments are analogous to conventional moments. However, a distribution may be specified by its L-moment even if some of its conventional moments do not exist.

Wang [51] introduced LH-moments which are generalization of L-moments and linear combination of higher order statistics. LH-moments reflect more and more the characteristics of the upper part of distribution and larger events in data (Wang [51]).

Mudholkar and Hutson [34] defined another method known as LQ-moment which behaves like L-moment. They found that LQ-moment always exists and are often easy to calculate than L-moment.

Elamir and Sheult [18] introduced Trimmed L-moment which is known as TL-moment. It is a generalisation of L-moment. TL-moment assign zero weight to extreme observations, they are easy to calculate, their sample variances and covariances can be obtained in closed form and they are more robust than L-moments. Moreover, a population TL-moment can exist when the corresponding population L-moment and central moment does not exist.

Application of extreme value distribution to rainfall data have been investigated by several authors from different parts of the world. Shabri, A. B. et al. [43] used L-moment and TL-moment to analyse the maximum rainfall data of 40 stations of Selangor Malaysia. Comparison between the two approaches showed that the L-moments and TL-moments produced equivalent results. GLO and GEV distributions were identified as the most suitable distributions for representing the statistical properties of extreme rainfall in Selangor. Deka, S. et al. [17] fitted three extreme value distributions to 9 stations separately using LH moment of order zero to four and found that GPA distribution is the best fitting distribution for the majority of the stations in the North East Region of India. Also Deka, S. et al. [15] tried to determine the best fitting distribution to describe the annual series of maximum daily rainfall data for a period of 42 years of nine stations of the North East Region of India using at site approach. Five extreme value distributions were fitted using L-moment and LQ-moment. Generalised Logistic distribution is empirically proved to be the most appropriate distribution for the majority of the stations in the North East Region of India. The drawback of the study was at site approach for extreme rainfall frequency analysis does not give accurate results. Jou P.H. et al. [27] studied the monthly precipitation for five old rain gauged stations in Iran using parametric and nonparametric methods. The results showed that the monthly precipitation data fitted to the parametric methods much better than nonparametric method. Norbiato et al. [35] tried to characterize the severity of a flash flood generating storm on 29th August 2003 in the eastern Italian Alps which was characterized by extra ordinary rainfall. Regional frequency analysis based on the index variable method and L-moments are utilized to analyse annual maximum rainfall data for the region of north eastern Italy. It was found that the regional growth curves based on Kappa distribution may be useful for the region. Modarres [33] studied regional rainfall frequency analysis of Iran using L-moment. It was found that Log normal distribution describes the overall distribution of rainfall in Iran. Trefry et al. [48] used L-moments method to analyse annual maximum rainfall and partial duration rainfall data of 152 stations of the state of Michigan. It was found that GEV distribution is the best fit distribution for annual maximum rainfall data and GPA distribution is the best fit distribution for partial duration rainfall data. Olofintoye, O.O. et al. [37] analysed annual rainfall data of 54 years from 20 different stations of Nigeria using Gumbel, Log-Gumbel, Normal, Log-Normal, Pearson and Log-Pearson distribution. The result showed that the Log-Pearson Type III distribution is the best distribution. Koutsoyiannes, D. [28] performed an extensive empirical investigation using a collection of 169 of the longest available rainfall records worldwide each having 100-154 year of data. This verified that the Gumbel distribution is not an appropriate distribution for rainfall distribution while Extreme value distribution of type II (EV2) is an appropriate distribution. Zalina et al. [52] discussed the comparative assessment of eight candidate distributions in providing accurate and reliable maximum rainfall estimates for Malaysia. L-moments method is used to estimate the model parameters. It was found that GEV distribution is the most appropriate distribution for describing the annual maximum rainfall series in Malaysia. Tao et al. [47] proposes a systematic assessment procedure to compare the performances of different probability distributions in order to identify an appropriate model that could provide the most accurate extreme rainfall estimates at a particular site. Out of nine probability models, it was found that Wakeby, GNO and GEV models could provide the most accurate extreme rainfall estimates. However, the GEV was recommended as the most suitable distribution due to its theoretical basis for representing extreme value process and its relatively simple parameter estimation. Ogunlela [36] studied the stochastic analysis of rainfall event in Ilorin using probability distribution functions. He found that the log Pearson type III distribution is the best for describing peak daily rainfall data of Ilorin. Lee [29] studied the rainfall distribution characteristics of Chia-Nan plain area by using different statistical analyses such as normal distribution, log-normal distribution, extreme value type I distribution, Pearson type III distribution and Log-Pearson type III distribution. The result showed that the Log-Pearson type III distribution performed the best probability distribution occupying 50% of the total station number. Adamowski et al. [1] used L-moments method for regional rainfall frequency analysis of Canada and found that GEV distribution is the best fit distribution for rainfall frequency analysis of Canada. Rakhecha et al. [41] analysed the annual extreme rainfall series at 316 stations of India, covering 80 years of rainfall data for trend and persistence using standard statistical tests.

1.3 Data collection and study region

The region North East India is situated at east of 80^0 E and North of 21^0 N. The average annual rainfall in the North East India ranges from 2000-4000 mm with a maximum of 11000 mm in Cherapunjee.

For this study 12 distantly situated gauged stations of North East India viz. Agartala, Dhubri, Guwahati, Imphal, Itanagar, Jorhat, Lakhimpur, Lengpui, Mohanbari, Pasighat, Shillong, and Silchar are considered. Annual daily maximum rainfall data of these stations for a period of 30 years from 1984 to 2013 are considered for this study. Data are collected from Regional Meteorological centre, Guwahati. A brief description of these stations are presented in this section.

Agartala: Agartala is the capital of the state Tripura situated at $23^0$50’N and $91^0$16’E. Agartala lies on the bank of the Hoara river and is located 2 km away from the

Bangladesh boarder. It is the third largest city in North East India. It has a monsoon influenced humid subtropical climate with large amounts of rainfall almost all year. The city experiences long, hot and summers lasting from April to October. Average temperatures are around 28^0c, fluctuating with rainfall. There is a short, mild winter from November to March with average temperature around 18^0c. Annual average rainfall is around 2178mm. Total population of Agartala city is 404004 and total area is $92km^2$.

Dhubri: Dhubri is the headquarter of Dhubri district of Assam located at 26.02^0N and 90^0E. It is an old town in the bank of the rivers Brahmaputra and Gadadhar. From March to October the weather is hot with an average maximum temperature around 31.1^0C and winter from November to February with average maximum temperature around 25.5^0C. Annual average precipitation is around 3016mm. According to census 2011 population of Dhubri is 63,388.

Guwahati: Guwahati is the largest city of North East India located at $26^0$11'N and $91^0$44'E. It is situated at the south bank of the river Brahmaputra. It has a humid subtropical climate. From March to September the weather is hot and wet summer. From November to February it is dry winter. During summer the average maximum temperature is around 27.2^0C and during winter it is around 25.3^0C. The average total annual rainfall is around 1722mm. Total area of Guwahati city is 215 km^2 and according to census 2011 its population is 957352.

Imphal: The city Imphal is the capital of Indian state Manipur. It is located at 24.8074^0N and 93.9384^0E in extreme eastern India. Its elevation is 786 m from sea level. Imphal has a sub-tropical climate with cool, dry winters, a warm summer and a moderate monsoon season. The annual average rainfall is about 1320 mm of with June being the wettest month. According to census 2011 its population is 264985.

Itanagar: Itanagar is the capital city of Indian state Arunachal Pradesh situated at North East. The location of Itanagar is 27.1^0N and 93.62^0E. It has an average altitude

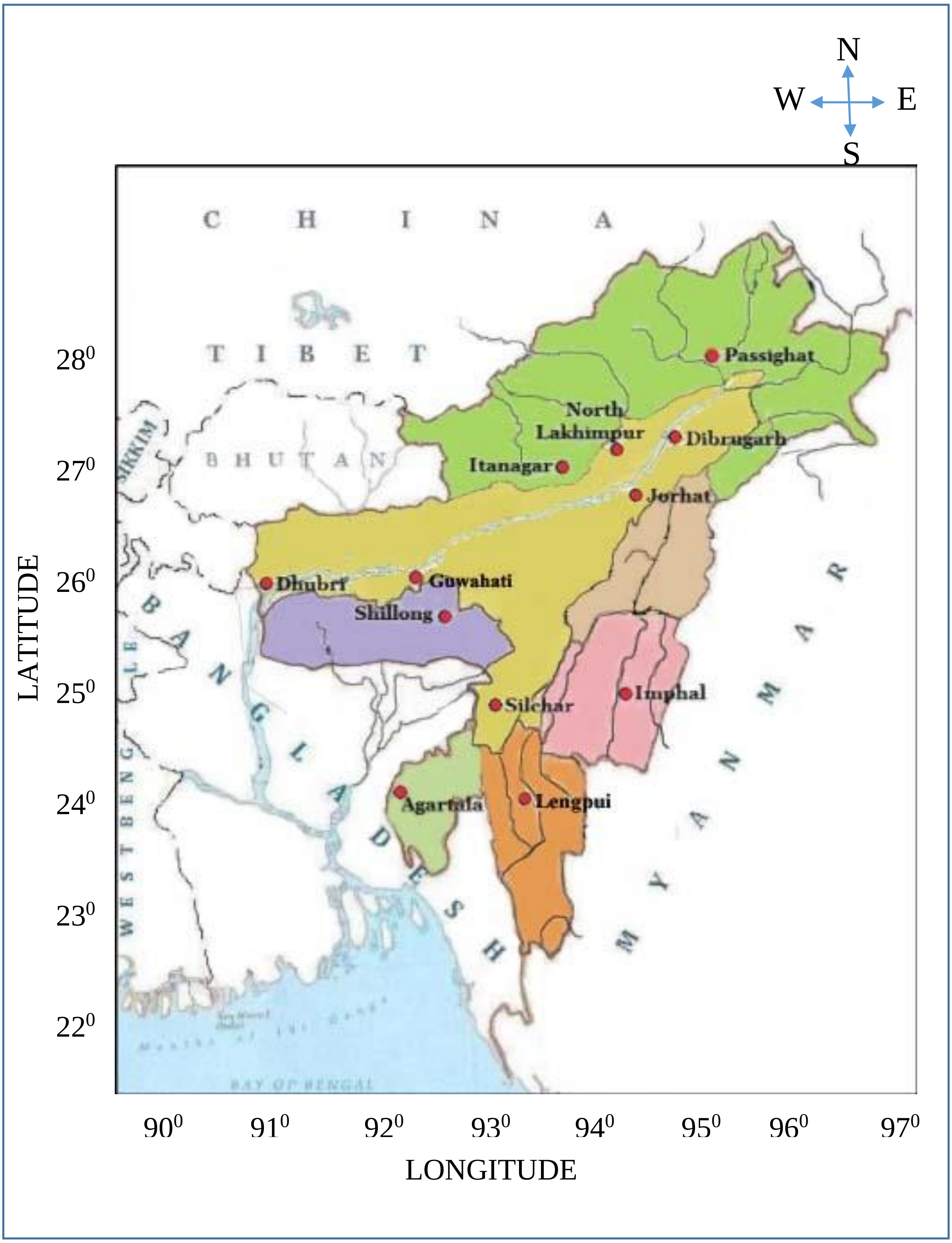

Figure. 1.1 Map showing the 12 gauged stations of North East India

of 750 metres. It has humid subtropical climate with dry, mild winters and warm, wet summers. Average annual rainfall of Itanagar is about 3302.09 mm. Its population is 59490 according cencus 2011.

Jorhat: Jorhat is a major city in the state of Assam, India. Its location is 26.75^{0}N and 94.22^{0}E. It has an elevation of 116 metres above the sea level. In the north of the district the Brahmaputra river forms the second largest riverine island of the world which is known as Majuli. Its weather is hot and humid during summer. Total area of the city is 9 km^2 and its population is 153259.

Lakhimpur: Lakhimpur is situated in the state of Assam in the Eastern part of India. It lies on the North bank of mighty river Brahmaputra. It is located at 27^{0}13$^/$60N and 94^{0}7$^/$0E. It has an elevation of 101 m above the sea level. Its city area is 15 km^2 and its population is 59814. The hottest month is July (29^{0}C) and coldest month is January (17^{0}C). Total annual rainfall is about 1551.3mm.

Lengpui: Lengpui is situated at Aizwal, the capital city of the State of Mizoram, India. Its location is 23.50^{0}N and 92.37^{0}E. Its elevation above the sea level is 1132m. It has a mild, subtropical climate due to its location and elevation. Average annual rainfall in Lengpui is 2161.4 mm. Area of Aizwal is 457km^2 and its population is 293416.

Mohanbari: Mohanbari is situated at 15 km North East from the city Dibrugarh, Assam India. Its location is 27^{0}28$^/$60N and 95^{0}01$^/$18E. It has a subtropical monsoon climate with mild winter, warm and humid summer. Area of Dibrugarh city is 66.14 km^2 and its population is 154019. The average annual rainfall of the Dibrugarh city in the North is 2760 mm while in the south, it is 1630 mm.

Passighat: Passighat is the headquarter of East Siang district of the state Arunachal Pradesh, India. It has an average elevation of 155 metres. Its co-ordinates are 28.07^{0}N and 95.33^{0}E. It has tropical humid climate during summer and dry mild winter. It has heavy rainfall every year during monsoon, starting from may until September. Its area is 14.60 km^2 and population is 24565. Annual total rainfall of Passighat is about 438834mm.

Shillong: Shillong is the capital and hill station of Meghalaya, India. It is the smallest city of India. Its location is 25.57^0N and 91.88^0E. It is situated at an average altitude of 4908 feet above sea level, with the highest point being Shillong peak at 6449 feet. The monsoon arrives in June and it rains almost until the end of August. Area of Shillong is 64.36 km^2 and its population is 143229. Annual total rainfall in Shillong is about 2167.4mm.

Silchar: Silchar is the headquarter of Cachar district in the state of Assam, India. Its location is 24.82^0N and 92.8^0E. It has an elevation of 22 m above the sea level. The climate is tropical by nature. Summer is hot, humid and interspersed with heavy rainfall and thunderstorms. Winter starts towards the end of November and lasts till February. The average annual rainfall is 3128.6 mm.

1.4 Objective

The main objectives of this study are as follows:

- Estimation of regional maximum rainfall frequency based on parametric methods.
- Comparison among the various methods of parameter estimation.
- To select the best fit probability distribution for maximum rainfall frequency estimation for the North East India
- To develop model for maximum rainfall quantile estimation for the North East India.

1.5 Organization of the Thesis

The thesis has seven chapters followed by appendix and bibliography. The thesis is organized as follows:

Chapter 2: In this chapter regional annual maximum rainfall analysis of North East India has been carried out using L-moment. Here theory of L-moments and regional rainfall frequency analysis using L-moments are covered. It is found that the PE3 distribution is the best fit distribution for the region. Parameters of PE3 distribution and quantile estimates are calculated. Regional rainfall frequency relationship for gauzed stations of the study region has been developed.

Chapter 3: In this chapter regional annual maximum rainfall analysis of North East India has been carried out using LQ-moment. Here theory of LQ-moments and regional rainfall frequency analysis using LQ-moments are elaborated. It is found that the GPA distribution is the best fit distribution for the region. Parameters of GPA distribution and quantile estimates are calculated. Regional rainfall frequency relationship for gauzed stations of the study region has been developed.

Chapter 4: In this chapter regional annual maximum rainfall analysis of North East India has been carried out using LH-moment. Here theory of LH-moments and regional rainfall frequency analysis using LH-moments are described in details. GPA distribution is selected as the best fit distribution for LH-moment of order 1 and GLO distribution is selected as the best fit distribution for other orders of LH-moments. Parameters the best fit distributions and quantile estimates are calculated for respective orders of LH-moment. Regional rainfall frequency relationship for gauzed stations of the study region has been developed.

Chapter 5: In this chapter regional annual maximum rainfall analysis of North East India has been carried out using TL-moment. Here theory of TL-moments and regional rainfall frequency analysis using TL-moments are discussed in details. It is found that the GPA distribution is the best fit distribution for the region. Parameters of GPA distribution and quantile estimates are calculated. Regional rainfall frequency relationship for gauzed stations of the study region has been developed.

Chapter 6: In this chapter a comparative study among the parameter estimation methods has been carried out. Monte Carlo simulation for relative root mean square (RRMSE) and relative bias (RBIAS) are used for comparison. Box plot, a graphical tool has been used to study the values of RRMSE and RBIAS. Six comparative studies

(a) Comparison between L-moment and LQ-moment,
(b) Comparison between L-moment and LH-moment,
(c) Comparison between LH-moment and TL-moment,

(d) Comparison between L-moment and TL-moment,

(e) Comparison between LQ-moment and TL-moment,

(f) Comparison between LQ-moment and LH-moment,

have been carried out. From comparison it is found that the GPA distribution designated by LH-moment of order 1 is the best fit distribution for maximum rainfall estimates of the region North East India. Also LH-moment of order 1 is significantly more efficient than other estimation methods used in this study.

Chapter 7: In this Chapter a brief conclusion and discussion of the results are described.

***