ANALYSIS OF EXTREME RAINFALL EVENT IN THE METEOROLOGICAL SUBDIVISION OF UTTAR PRADESH USING THREE-PARAMETER EXTREME VALUE DISTRIBUTIONS

A DISSERTATION
Submitted in partial fulfillment of the requirement for the award of the degree
Of

MASTER OF TECHNOLOGY
IN

HYDRAULICS AND WATER RESOURCES ENGINEERING

Submitted by

ASTHA YADAV

(2K18/HFE/05)

Under the Guidance of

Prof. VIJAY K. MINOCHA
Department of Civil Engineering
Delhi Technological University
Delhi

DEPARTMENT OF CIVIL ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY
(FORMERLY DELHI COLLEGE OF ENGINEERING)
DELHI – 110042
AUGUST -2020

CANDIDATE'S DECLARATION

I, ASTHA YADAV, Roll NO. 2K18/HFE/05 of M. Tech. (Hydraulics and Water Resources Engineering), hereby declare that the project Dissertation titled **"Analysis of Extreme Rainfall Event in the Meteorological Subdivision of Uttar Pradesh Using Three-Parameter Extreme Value Distributions"** which is submitted by me to the Department of Civil Engineering, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of technology is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma, Fellowship, or other similar title or recognition.

ASTHA YADAV
2K18/HFE/05
(M.TECH)

Date:
Place: DELHI

CERTIFICATE

I hereby certify that the Project Dissertation titled, **"Analysis of Extreme Rainfall Event in the Meteorological Subdivision of Uttar Pradesh Using Three-Parameter Extreme Value Distributions"** which is submitted by Astha Yadav, 2K18/HFE/05 of M. Tech (Hydraulics and Water Resources Engineering) Delhi Technological University, Delhi in fulfillment of the Major project of Master Of Technology, is a record of the project work carried out by him under my supervision and guidance.

The best of my knowledge, these works have not been submitted in part or full any Degree or Diploma to this University or elsewhere.

Supervisor

Prof. Vijay K. Minocha
Department of Civil Engineering
Delhi Technological University,
Delhi, 110042

Date:

Place: Delhi

ACKNOWLEDGMENT

First, I would like to thank my family. Without their love and support over the years, none of this would have been possible. They have always been there for me, and I am thankful for everything they have helped me achieve.

I would like to express my gratitude to all those who gave me the possibility to have this project. I want to thank the Department of Civil Engineering and Hydraulics and Water Resources Engineering of “DELHI TECHNOLOGICAL UNIVERSITY” for giving me such a golden opportunity to commence this project in the first instance. I have furthermore to thank “Prof. Vijay K. Minocha” who encouraged me to go ahead with my project and for his stimulating support.

My colleagues from the Civil Engineering Department supported me in my project work. I want to thank them for all their help, support, interest, and valuable hints. I am exceedingly obliged to my project guide, who looked closely at the report correcting both and offering suggestions for improvement.

ASTHA YADAV

(2K18/HFE/05)

ABSTRACT

Extreme rainfall is a global phenomenon occurring in almost every major country of the world that cause significant damage such as floods and erosion that can destroy infrastructure, human and animal life, disruptive economic activities, and related development. The forecasts of heavy rainfall help to implement strategies, and measures before they occur. In this study, we used statistics strategies to create models that could work to predict maximum rainfall in Uttar Pradesh. For this purpose, the annual maximum rainfall from 1979-2018 applies in the subdivision of Uttar Pradesh. Extreme value distribution GEV, GLO, GPA, and UEV considered analyzing extreme events. The parameters of the 1st three distribution are determined using the method of moment (MOM) and probability-weighted L- moment (PWM-L). The last distribution parameters are determined using the simple objective method (SO). Two methods used to analyze the best fit distribution among four distribution, i.e., graphical method (coefficient of determination, R^2), and goodness of fit test (GOF). Five different GOF tests apply in this study, i.e., RRMSE, RMSE, MAE, MADI, and PPCC. Basic time series analysis such as outliers test, normality test, homogeneity test, and stationarity tests performed to ensure that the information used is adequate and appropriate. The results obtained indicate that the GEV (PWM-L) was an appropriate method for the distribution of the annual maximum rainfall series in the west Uttar Pradesh subdivision and the GLO (PWM-L) which was an appropriate distribution method to analyze the series of the East Uttar Pradesh subdivision. The coefficient of determination (R^2) for the observed versus predicted rainfall based on the best fit model observed to be 0.9899 and 0.9865, respectively, West Uttar Pradesh subdivision and East Uttar Pradesh subdivision. Finally, the most appropriate distribution at each site applies to predict maximum rainfall at different return periods.

Keywords: outliers, stationarity test, homogeneity test, normality test, GEV, GLO, GPA, UEV distribution, method of moment, a probability-weighted moment with L-moment, simple objective method, R^2, GOF, maximum rainfall, Uttar Pradesh.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

N = No of observations

Min = Minimum rainfall

Max = Maximum rainfall

X = Variate

x'= Mean

S^2 = Variance

σ = standard deviation

C_V = coefficient of variation

C_S = skewness coefficient

C_K = kurtosis coefficient

L_{95} = Lower bound on the mean

U_{95} = Upper bound on the mean

AMS = Annual maximum series

PDF = probability density function

CFD = cumulative density function

GEV = generalized extreme value distribution

GLO = generalized logistic distribution

GPA = generalized Pareto distribution

UEV = Unified extreme value distribution

X_P = Quantile function for GEV, GLO, and GPA distribution

X_T = Quantile function for UEV distribution

α = scale parameter of GEV, GLO, and GPA distribution

k = Shape Parameter of GEV, GLO, and GPA distribution

μ = Location Parameter of GEV, GLO and GPA distribution

c = scale parameter of UEV distribution

k = shape parameter of UEV distribution

w = location parameter of UEV distribution

T = Return period

F = Frequency

MOM = method of moment

PWM-L = probability weighted moment with l- moment

SO= simple objective method

GOF= goodness of fit

RRMSE= relative root means square error

RMSE= root mean square error

MAE= maximum absolute error

MADI= mean absolute deviation index

PPCC= probability plot correlation coefficient

CHAPTER 1.0

INTRODUCTION

1.1 GENERAL [3,5,11, 16,18]

Water resource has become a prime concern for any development and planning, including flood analysis, food production, control, dam constructions, and sufficient water resource management. Rainfall is also an essential part of the hydrologic cycle, and the modification of its pathway can have a direct impact on the water source of the component, as it affects stream flows, soil moisture, and groundwater retention. Serious hydrological events cause severe damage to humans, animals, and injuries, such as floods and landslides. Excessive rainfall is defined as the average daily rainfall during the year or rainfall higher than 100mm in 24 hours. The state (U.P), with an area of 243,290 square kilometers, is the fourth largest area in India in terms of space, and is located in the middle of northeastern India and shares the country’s border with Nepal [3].the land covers the humid climate and has four different seasons. In the province winter starts in January and February, followed by summer between March and May and then the season between June and September. In Uttar Pradesh, summer is extreme, and temperatures fluctuate between 0^0c and 50^0c in dry, hot air [3].

Based on the advanced conditions of topography and climatologically over the subcontinent, the India Ministry of Environment has divided the country into 36 meteorological subdivisions [3]. The sub-divisional monthly rainfall data collected from the IMD during the quantity 1979-2018. There are 26 and 19 rain gauge stations, respectively, in East Uttar Pradesh subdivision and West Uttar Pradesh subdivisions.

Rainfall is a decrease in the amount of rainfall over a while. Rainfall patterns were always determined, including limitation of rainfall distribution, depth of low or high rainfall, and identification of wet or dry events on a particular day, a strategy and frameworks required to build overtime work. For example, a reservoir should be able to store the expected rainfall in the seed region [3]. If the dam is not large enough, there is a risk that water can pass through the dam, and water from the dam can occur in flood conditions. The rain that should be stored in the reservoir is likely to come from an overcrowded region and build up over several days. Although monsoon air travel determines the hot or rainy season, changes in the earth’s climate over the past several years have made the season unpredictable throughout the year [3]. The floods are a known danger to Uttar Pradesh due to the overflow of its major rivers such as Ganga, Yamuna, Rāmgangā, Gomti, Sharda, Ghaghara, Rapti, and Gandak [18]. The estimated annual losses due to floods in the province are 4.32 billion. Significant flood management efforts have been made to reduce this rainfall. Most of these floods occurred due to monsoon rains, flooding of dams, and flooding during the rainy season. The year 2010 saw one year of state floods.

Excess rainfall data should be weighed against the appropriate number distribution that provides the best of high rain patterns. Many distributions have been able to be used in excessive rainfall analysis, such as generalized extreme value distribution, generalized Pareto distribution, generalized logistics distribution, and unified extreme value distribution. The distribution of rainfall that has occurred in the province is classified as one of the worst events since the UP was discovered during a hot climate. Therefore, in this study, extreme value distribution is used to equal the daily rainfall in the west and east Uttar Pradesh subdivision.

1.2 BASIC ASSUMPTION

While doing the rainfall analysis in West Uttar Pradesh subdivision and East Uttar Pradesh subdivision, there are certain assumptions to follow because it provides us better accuracy & precision. Afterward, predicting the Rainfall with their corresponding return period.

Assumption:

1. There are no outliers in data sets.
2. Rainfall data should be statistically random, or it should not be dependent one over the other.

1.3 STATEMENT OF PROBLEM

Extreme high Rainfall is a dangerous hazard of nature because high Rainfall can lead to floods and landslides, which can threaten human life, disrupt transport. Excessive high Rainfall can severely affect both the environment and social lives. Also, extreme low Rainfall is a dangerous hazard of nature because extreme low Rainfall can lead to drought conditions, which even can threaten human life. In this study, we study only excessive-high Rainfall in West Uttar Pradesh and East Uttar Pradesh subdivision. This study aims to use mathematical and statistical techniques to make a model that can apply to forecast the high Rainfall in Uttar Pradesh.

1.3 OBJECTIVES OF THE STUDY

1.4.1 GENERAL OBJECTIVE

The objective of this thesis was to develop an extreme value model that can use to forecast the occurrence of excessive maximum Rainfall in West Uttar Pradesh and East Uttar Pradesh subdivision.

1.4.2 SPECIFICS OBJECTIVES

The specific objectives of this study are:

1. To estimate the statistical parameters such as mean, coefficient of variance, standard deviation, coefficient of kurtosis, coefficient of skewness, for annual maximum daily rainfall series.
2. To test the outliers, homogeneity, normality, stationary of time series data (annual maximum daily Rainfall).
3. To determine GEV, GLO, GPA distribution parameters using a method of the moment.

4. To determine the parameter of GEV, GLO, GPA distribution using probability-weighted moment with L-Moment (PWM-L).
5. To determine the parameters of Unified Extreme Value distribution using a simple objective method.
6. To forecast annual maximum Rainfall in West Uttar Pradesh and East Uttar Pradesh subdivision using GEV, GLO, GPA, and UEV.
7. To draw a graph between observed annual maximum daily rainfall and predicted annual maximum daily rainfall. Estimate the coefficient of determination for both cases. Based on R^2 values determined best – fit Distribution for West Uttar Pradesh subdivision and East Uttar Pradesh subdivision.
8. To determine the best fit distribution for both cases using various Goodness of fit test (RRMSE, MSE, MAE, MADI, and PPCC).
9. To predict values of return periods for the next 2, 5, 10, 20, and 50 years based on 40-year annual maximum daily rainfall data and using best fit extreme value distribution with the best parameter estimated method.

1.5 SCOPE OF THE STUDY [2, 5, 18]

Extreme rainfall data need to be modeled by a suitable statistical distribution. For this study, GEV, GLO, GPA, and UEV distribution were analyzed. The most suitable distribution for both cases is determined using different GOF test. The annual maximum daily rainfall data of 19 and 26 rain gauge stations in West Uttar Pradesh and East Uttar Pradesh, respectively, which have records for 40 years from the year 1979 – 2018, is analyzed.

Method of Moments, Probability Weighted Moment with L- Moment is applied to estimate the parameters of GEV distribution, GLO Distribution, GPA distribution, and Simple objective method used to determine the parameters of UEV distribution. To select the best Distribution, RRMSE, RASE, MAE, MADI, and PPCC were carried out. Using the best distribution predicted values of different return periods based on the 40-year annual maximum rainfall

1.6 SIGNIFICANCE OF THE STUDY [16, 18]

It is known that rainfall analysis is beneficial in managing water use to measure building capacity and helps predict extreme weather events. Rainfall data can apply to assess the risk of severe flooding and heavy rainfall and also crucial for drought conditions.

Also, a civil engineer can predict what should apply in the construction of retaining walls, bridges, and dam processes based on rainfall received over a while. It can also help our country with unnecessary costs and economic losses and avoid the risk of flooding.

Predictable return times have been calculated using GEV, GLO, GPA, and UEV distribution for 2, 5, 10, 20, and 50 years calculated based on 40 years of rainfall data. Therefore, the results can help engineers and measurement planners measure the ability to build structures that can survive under extremely adverse conditions.

CHAPTER 2.0

LITERATURE REVIEW

2.1 GENERAL

In frequency analysis of flood (basically caused by extreme Rainfall), the statistical approach is the most popular way to proceed with the active research within the desired direction many of them use this approach to search out the answer for his or her problem. The development in these areas uses the annual maximum daily rainfall series. In this, we discuss the review of the literature given below.

2.2 PREVIOUS LITERATURE

So many studies have conducted various researchers as follows:

Annual maximum series (AMS), it is a sequence of annual rainfall, with yearly rainfall defined as the maximum peak rainfall of the year of records. Many candidate have suggested distribution for annual maximum series includes Extreme Type I, Extreme Type II, Extreme Type III, generalized extreme value, and many more.

Bjorn H. Auestad, Andreas H. and A. Karlsen (2012), Modelling and analysis of daily rainfall data, in his study, they developed a statistical model for the regular rainfall measurements. Data series from Bergen the last 107 years and from Sviland, Rogaland the previous 115 years used. He used generalized linear models as a statistical tool for fitting the data to the model. Based on simulation studies, comparisons of the model estimated quantities, and corresponding data quantities, the model seems to fit a series of daily rainfall data very well.

Ho Ming Kang and Fadhilah Yusof (2012), Homogeneity tests on daily rainfall series in peninsular Malaysia, Homogeneity test performed on the regular rainfall series of stations in Damansara, Johor, and Kelantan. In his or her study, they used four methods of homogeneity test, i.e., Standard normal homogeneity test (SNHT), Buishand range test, Pettitt test, and Von Neumann ratio test [11]. And Daily rainfall data from 33 stations are obtained from J.P.S. for stations in Damansara (1998-2007), Johor (1996-2005) and Kelantan (1998-2007) taken for study [11].

T.O. Olatayo and A. I. Taiwo (2014), Statistical modeling and Prediction of rainfall time series data. In his study, they present tools for modeling and predicting the behavioral pattern in rainfall phenomena based on past observations. The study introduces three fundamentally different approaches for designing a model, the statistical method based on the autoregressive integrated moving average (ARIMA), the emerging fuzzy time series (FST) model and the non-parametric method (Theil's regression).

Tefaruk H. and H.C (2014) work on-trend, independence, stationarity, and homogeneity tests on the maximum rainfall series of standard durations recorded in Turkey. In his study, they used 73 years of annual maximum rainfall data. They used Mann-Kendall and linear regression for trend analysis, von-Neumann independence, Wald-Wolfowitz for stationarity, and Mann–Whitney for homogeneity test. It concluded that annual maximum rainfall series in Turkey could generally treat as independent and identically distributed random variables, allowing conventional intensity duration frequency calculations to perform by statistical frequency analysis.

Majid Javari (2016); trend and homogeneity analysis of rainfall in Iran. His study aims to check at trends and homogeneity through the analysis of rainfall variability patterns in Iran. During this study for homogeneity test analysis of monthly, seasonal, and annual rainfall employed in 140 stations within the 1975-2014 period. He used two homogeneity tests, i.e., ACE and VN tests, at a 5% significance level.

Susheel Kumar Patel and Subash (2018), a comparative study of trends in Rainfall over the metrological subdivision of Uttar Pradesh. In his research, they investigate trends over two subdivisions of Uttar Pradesh.

Uwimana Oliver and Joseph K. (2018) gave the idea of modeling maximum Rainfall using GEV: case study King City [9]. In his research, they compare Gumble distribution, Weibull Distribution, Frechet Distribution with generalized extreme value distribution.

Wenny Susanti and Arisman Adnan (2018), work on the analysis of extreme rainfall in Pekanbaru City using three-parameter GEV and GPA distribution. For his or her study, they used 17 years of annual maximum daily rainfall of Pekanbaru City and using GOF tests (RRMSE, RASE, PPCC) to select the best distribution.

M. Masereka, George M. Ochieng, and J.S (2018) Statistical analysis of annual maximum daily Rainfall for Nelspruit and its environs. For his or her study, they take 54 years of maximum yearly daily Rainfall (mm) data, Nelspruit. GEV, GLO, and Gumbel max distribution used for analysis.

Sushil k. Singh (2018) gave the concept of recent extreme value distribution, i.e., Unified Extreme Value Distribution (UEV). For his study, they used the annual maximum series of rivers in the United States for 52 years on the White River near Nora Indiana, 43 years on Sugar Creek at Crawfordsville, Indiana, and 48 years on the Tippecanoe River near Delphi, Indiana [1].

Rudolph and Ebierin A. Otuaro (2019) gave concept simple to use Microsoft Excel Template for estimation of the parameters of some selected probability distribution model by using of L-Moment [2]. In his study, they used generalized extreme value distribution, generalized logistics distribution, generalized Pareto Distribution, and take 40 years of annual maximum precipitation data, Nigeria.

Mahmood Hassan, Omar Hayat, and Zahra Noreen (2019) Selecting the most effective probability distribution for at-site flood frequency analysis; a study of Torne River [14]. For his or her research, they used generalized extreme value, three-parameter log-normal, generalized logistic, Pearson type-III, and Gumbel distributions for annual maximum steam flow at five gauging sites of Torne River in Sweden. And to estimate the parameters of distributions, maximum likelihood estimation, and L-moments methods used.

2.3 CONCLUSION

After going through the literature review, we mainly bear with the statistical approach during which we faced the different extreme value distributions which used namely GEV, GLO, GPA, and UEV distribution for predicting the future rainfall values with their corresponding return period and goodness-of-fit tests (RRMSE, RMSE, MAE, MADI, and PPCC) and graphical method (R^2) which is very helpful for considering the best distribution among all the distribution.

CHAPTER 3.0

METHODOLOGY

3.1 INTRODUCTION

In this chapter, the Rainfall is in the form of annual maximum daily rainfall data series of (1979-2018) years. Firstly the outliers in the data can be check using the Dixon test and Grubbs test and removed the outliers in datasets. Afterward, we arranged the rainfall data in ascending order and providing the position numbers as least Rainfall is having the first position and highest Rainfall having the last place afterward we check Homogeneity, Normality, Stationary of datasets. We are now using various extreme value distributions, namely: GEV, GLO, GPA, and UEV distribution predicting the Rainfall at Two IMD Subdivision of Uttar Pradesh. Using coefficient of determination (R^2), and various GOF tests (RRMSE, RMSE, MADI, MAE, and PPCC) predict best probability distribution and calculate rainfall values for different return periods (2, 5, 10, 20, and 50).

3.2 STUDY AREA [3]

Uttar Pradesh (U.P), with a total area of 243,290 square kilometers, is India's fourth-largest state in terms of land area, and it is situated in the Central northeast of India and shares an international boundary with Nepal (Fig 1) [3]. U.P has a humid subtropical climate and experiences four distinct seasons [3]. In province winter is starting in January and February is followed by summer between March and May and so the monsoon season between June and September [3]. Summers are extreme, with temperatures fluctuating between 0° C and 50° C with dry, hot winds [3].

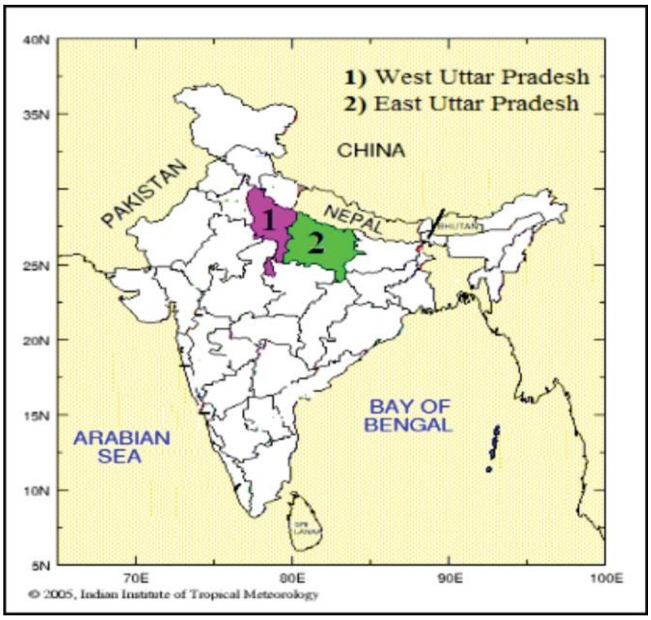

Fig 3.1 Two meteorological subdivisions of the U.P.

3.3 DATA AVAILABILITY [4]

Based on the climatologically and prevailing topography conditions over the sub-continent, the Indian department of meteorological has divided the country into 36 meteorological subdivisions. The sub-divisional monthly series of rainfall data were collected from the subdivision IMD excel file during the period 1979-2018. There are 26 and 19 rain gauge stations, respectively, in East UP and West UP subdivisions.

3.3.1 WEST UTTAR PRADESH SUBDIVISION

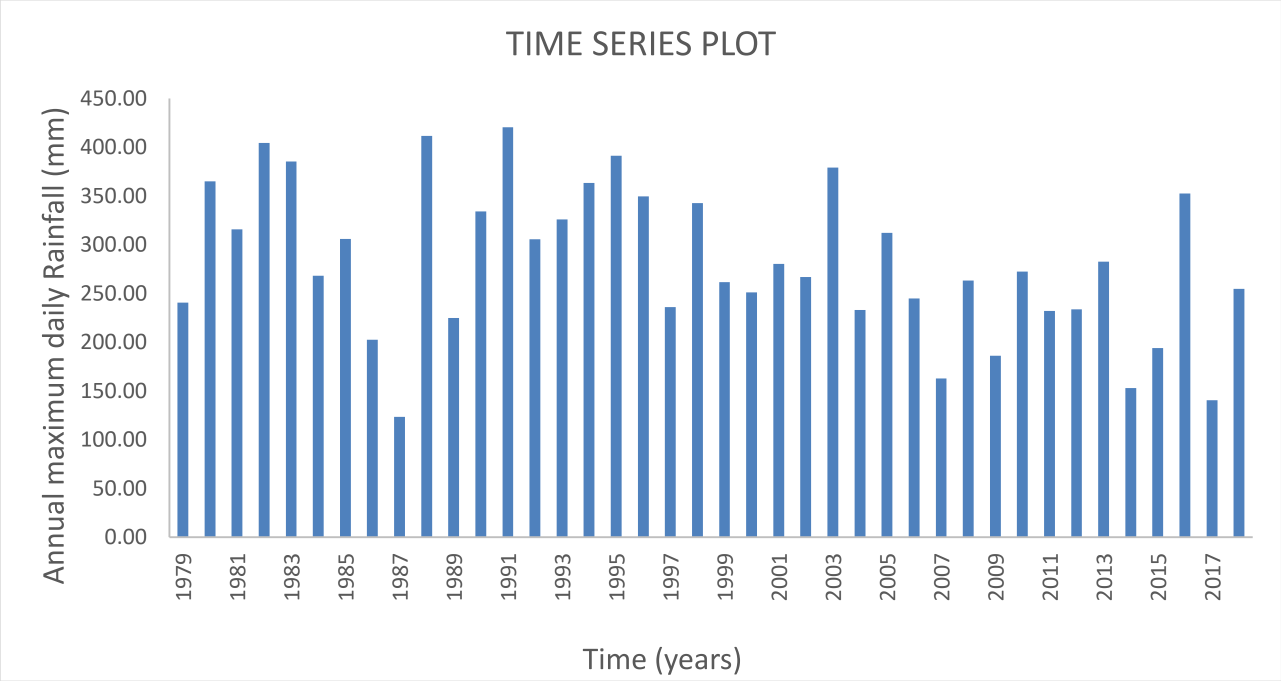

Fig 3.2 Annual maximum daily rainfall data series of (1979-2018) years of West U.P. Subdivision.

3.3.2 EAST UTTAR PRADESH SUBDIVISION

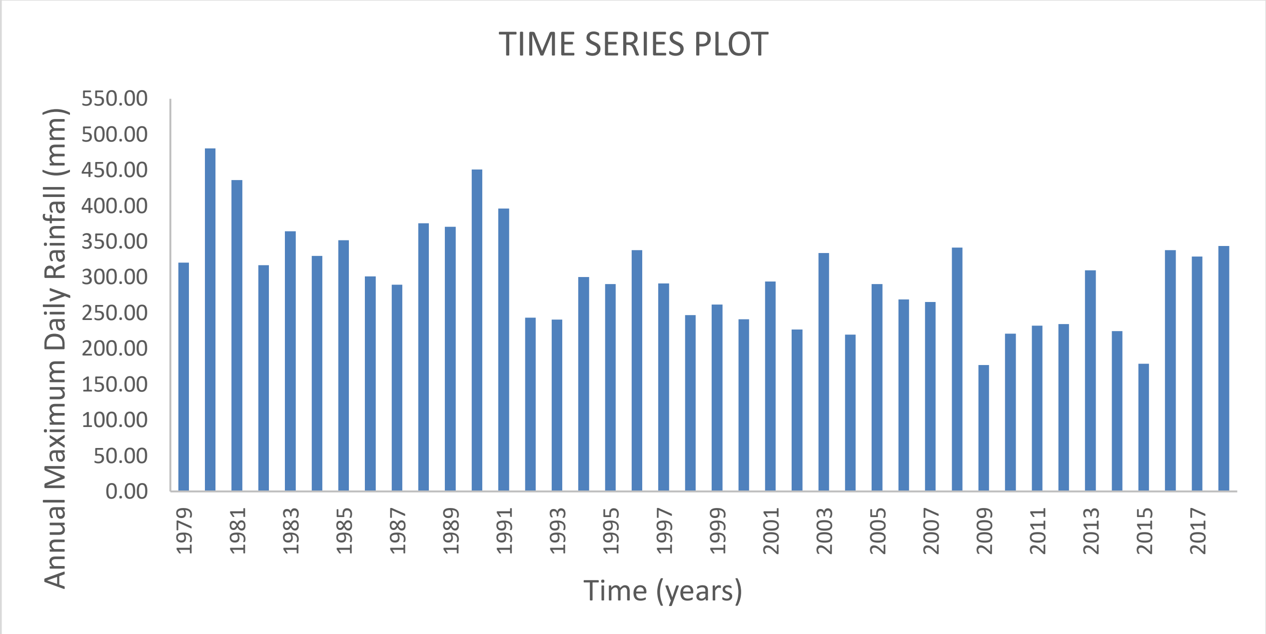

Fig 3.3 Annual maximum daily rainfall data series of (1979-2018) years of East U.P. Subdivision.

3.4 SCREENING OF DATASETS [6]

It is necessary to screen the data which you will analysis further; otherwise, your prediction of Rainfall is incorrect; the peak rainfall data used for frequency analysis should meet the following requirement are:

1. The data on Rainfall should be independent.
2. There are no outliers in data sets.
3. The sample size such that the population parameters can be estimates from it.

It is essential for a screening of rainfall data sets. Various outliers, stationarity, homogeneity, normality tests were performed using XLSTAT software.

3.4.1 TEST FOR OUTLIERS

We have two types of outliers test:

1. Dixon test
2. Grubbs test.

These test aims to check if data (annual maximum daily Rainfall) has any outliers. If any outliers find in data sets first find the valid reason behind it, the reason is essential then removes that outliers or reason is not necessary like tail end outlier than outliers considered for further analysis.

3.4.1.1 DIXON TEST

In statistics, the Dixon test applies to the identification and rejection of outliers in the dataset. The p-value has computed using 1000000 Monte Carlo simulations and a 99% confidence interval. Calculate z-scores value, identify outliers in the data set.

Test interpretation:

Ho: there is no outlier in the data set

Ha: the min or max value is an outlier in the data set

The result is based on computed p-value when the p-value is greater than the significance level alpha equal to 5%, and one cannot reject the null hypothesis Ho.

3.4.1.2 GRUBBS TEST

In statistics, the Grubbs test or the Grubbs test (named after Frank E. Grubbs, who published the analysis in 1950) is a test used to detect outliers in a data assumed to come from a normally distributed population. Calculated z-scores and identify outliers in the data set.

Test interpretation:

Ho: there is no outlier in a data set

Ha: there is precisely one outlier in the dataset

As the computed p-value is greater than the significance level alpha =0.05, one cannot reject the null hypothesis Ho.

Both the Dixon test and the Grubbs test were performed using XLSTAT software. Identify the outliers in annual maximum daily rainfall data set, give a specific reason, and then removed them from the data set. The reaming data set is used for further analysis.

3.4.2 STATIONARITY TEST [5, 6, 11, 14]

It is crucial to study the stationarity of rainfall data, the sequence of random variables represents the series of maximum Rainfall, and it must have constant properties through time. Statistical parameters such as mean, variance, and autocorrelation are constant over time. The result identifies that the series is stationary or not. In this thesis, three stationarity tests apply. KPSS test that considers as null hypothesis Ho that the series is stationary and a unit root tests such as ADF and PP test, for which the null hypothesis Ho that the unit root is present and series is not stationary. The three different stationary tests are given below:

1. Augmented Dickey-Fuller (ADF) test
2. Kwiatkowski Phillips, Schmidt and Shin (KPSS) test
3. Phillips-Perron test.

These test aims to check if data (annual maximum daily rainfall) in periods is stationary.

3.4.2.1 AUGMENTED DICKEY FULLER TEST

The ADF test is a unit root test for stationarity and the null hypothesis that a unit root is present in a time series sample.

Test interpretation:

Ho: there is a unit root for the series

Ha: there is no unit root for the series; it is stationary.

The result is based on the computed p-value; if the p-value is greater than the significance level alpha equal to 5%, one cannot reject the null hypothesis Ho.

As the computed p-value is greater than the significance level alpha = 0.05, one cannot reject the null hypothesis Ho.

3.4.2.2 KWIATKOWSKI PHILLIPS, SCHMIDT AND SHIN TEST

KPSS test applies for testing a null hypothesis that an observable time series is stationary around a deterministic trend against the alternative of a unit root.

Test interpretation:

Ho: the series is stationary

Ha: the series is not stationary

As the computed p-value is lower than the significance level alpha= 0.05, one should reject the null hypothesis Ho and accept the alternative hypothesis Ha.

3.4.2.3 PHILLIPS-PERRON TEST

The Phillips-Perron test is a unit root test. That is, it is used in time series analysis to test the null hypothesis that a time series integrated of order 1[6].

Test interpretation:

Ho: There is a unit root for the series

Ha: There is no unit root for the series. The series is stationary.

As the computed p-value is greater than the significance level alpha=0.05, one cannot reject the null hypothesis Ho.

3.4.3 NORMALITY TEST

Normality tests associated with the null hypothesis that the population from which a sample extracted follows a normal distribution. When the p-value linked to a normality test is lower than the risk alpha, the corresponding Distribution is significantly not-normal [6]. In this study, two types of normality test used:

1. Shapiro-Wilk test
2. Anderson –Darling test

3.4.3.1 SHAPIRO-WILK TEST

Shapiro-Wilk test is best suitable for samples of less than 5000 observations. This test rejects the hypothesis of normality when the p-value is less than or equal to 0.05.

Test interpretation:

Ho: The variable from which the sample extracted follows a Normal distribution [6].

Ha: The variable from which the sample extracted does not follow a Normal distribution [6].

3.4.3.2 THE ANDERSON – DARLING TEST

This test proposed by Stephens (1974) is a modification of the Kolmogorov-Smirnov test. It is suited to several distributions, including the normal distribution for cases where the parameters of the distribution are not known and have to be estimated [6].

Test interpretation:

Ho: the variable from which the sample extracted follows a Normal distribution.

Ha: the variable from which the sample extracted does not follow a Normal distribution.

As the computed p-value is greater than the significance level alpha=0.05, one cannot reject the null hypothesis Ho.

- OUTPUT- Plots associated to the Normality tests are:
  1. P-P Plots: P-P plots (for Probability –Probability) used to compare the empirical distribution function of a sample with that of a sample distribution of the same mean and variation .if the sample follows a normal distribution, the points lie along the first bisector of the plan.
  2. Q-Q Plots: Q-Q plots (for Quantile-Quantile) were used to compare the quantities of the sample with those of a sample distributed according to a normal distribution of the same mean and variance [6]. Suppose the sample follows a normal distribution, the points lie along the first bisector of the plan.

3.4.4 HOMOGENEITY TEST [12, 13, 14]

Homogeneity is a critical issue with detecting the variability of the data [13]. In general, when the data is homogeneous, it means that the measurements of the data are taken at a time with the same instruments and environments [12]. A homogeneity test used to determine a series may be homogeneous or time at which a change occurs. In this study, four homogeneity tests were used to analyze the test the homogeneity of the rainfall data.

1. Pettitt's test
2. Standard normal homogeneity test
3. Buishand's test
4. Von Neumann's ratio test

- These three tests are capable of detecting the year where the break occurs. Meanwhile, the VNR test is not able to give information on the year break because the analysis assumes the series not randomly distributed under the alternative hypothesis. These test aims to check data are homogeneous over time.

3.4.4.1 PETTITT'S TEST

Pettitt test is a nonparametric test that requires no assumption about the distribution data [13]
Test interpretation:

Ho: Data are homogeneous

Ha: There is a data at which change in the data

As the computed p-value is lower than the significance level alpha=0.05, one should reject the null hypothesis Ho, and accept the alternative hypothesis Ha.

3.4.4.2 STANDARD NORMAL HOMOGENEITY TEST

The SNHT test (Standard Normal Homogeneity Test) was developed by Alexanderson (1986) to detect a change in a series of rainfall data. The test applied to a series of ratios that compare the observations of a measuring station with an average of several stations. The proportions are then standardized.

Test interpretation:

Ho: Data are homogeneous

Ha: There is a data at which there is a change in the data

As the computed p-value is lower than the significance level alpha= 0.05, one should reject the null hypothesis Ho, and accept the alternative hypothesis Ha.

3.4.4.3 BUISHAND'S TEST

Buishand's test (1982) can be used on variables following any type of distribution. But its properties have been mainly studied for the normal case [6].

Test interpretation:

Ho: Data are homogeneous

Ha: There is a data at which there is a change in the data

As the computed p-value is lower than the significance level alpha= 0.05, one should reject the null hypothesis Ho, and accept the alternative hypothesis Ha.

3.4.4.4 VON NEUMANN'S RATIO TEST

It is a test that used the ratio of mean square successive (year to year) difference to the variance [13].
Test interpretation:

Ho: Data are homogeneous

Ha: There is a data at which there is a change in the data

As the computed p-value is greater than the significance level alpha = 0.05, one cannot reject the null hypothesis Ho.

- The results are categorized into three classes, which are useful, doubtful, and suspect according to the number of tests rejecting the null hypothesis.

1. Class A: Useful

   The series that rejects one or none null hypothesis under the four tests at a 5% significance level considered. Under this class, the series is grouped as homogeneous and can use for further analysis.

2. Class B: Doubtful

   The series that reject two null hypotheses of the four tests at a 5% significance level paced in this class. In this class, the series has an inhomogeneous signal and should critically be inspected before further analysis.

3. Class C: Suspect

   When there are three or all tests are rejecting the null hypothesis at a 5% significance level, then the series is classified into this category. In this class, the series can be deleted or ignored before further analysis.

3.5 DESCRIPTIVE STATISTICS [7, 8]

3.5.1 MEAN

It is the average of the numbers.

$$x' = \frac{\sum_{i=1}^{N}(X)}{N} \tag{3.1}$$

3.5.2 STANDARD DEVIATION

It is given as the square root of the variance expressed as:

$$\sigma = \left(\frac{\sum_{i=1}^{N}(X_i - x')^2}{N-1}\right)^{1/2} \tag{3.2}$$

3.5.3 COEFFICIENT OF VARIATION

It is equal to the ratio of the standard deviation to the mean.

$$C_V = \frac{\sigma}{x'} \tag{3.3}$$

3.5.4 SKEWNESS COEFFICIENT

The coefficient of skewness measures the skewness of a distribution. It is non-dimensional

It is a non-dimensional measure of the asymmetry of the frequency distribution fitted to the data. Its unbiased estimate is given by:

$$C_S = \frac{N\sum_{i=1}^{N}(X_i - x')^3}{(N-1)(N-2)\sigma^3} \tag{3.4}$$

3.5.5 KURTOSIS COEFFICIENT

The peakedness of the frequency distribution near its center measured by kurtosis coefficient, which expressed as:

$$C_K = \frac{N^2 \sum_{i=1}^{N} (X_i - x')^4}{(N-1)(N-2)(N-3)\sigma^4} \tag{3.5}$$

3.6 PROBABILITY DISTRIBUTION [15, 16, 17]

3.6.1 GENERALIZED EXTREME VALUE DISTRIBUTION (GEV) [2]

Generalized extreme value distribution is a three-parameter distribution, the PDF, CDF, and Quantile function of the Distribution is given by:

The probability density function (PDF) is given by:

$$F(X) = \frac{1}{\alpha}\left[1 - k\left(\frac{X-\mu}{\alpha}\right)\right]^{\frac{1}{k}-1} exp\left[-\left\{1 - k\left(\frac{X-\mu}{\alpha}\right)\right\}^{1/k}\right] \tag{3.6}$$

The cumulative density function (CDF) is given by:

$$F(X) = exp\left[-\left\{1 - k\left(\frac{X-\mu}{\alpha}\right)\right\}^{1/k}\right] \tag{3.7}$$

Range:

$$For\ k < 0 \quad \alpha > 0, \mu + \frac{\alpha}{k} \leq X < \infty$$

$$For\ k > 0 \quad -\infty \leq X \leq \mu + \frac{\alpha}{k}$$

The Quantile function (X_P) is given by:

$$X_P = \mu + \frac{\alpha}{k}[1 - (-ln(F))^k] \tag{3.8}$$

$$F = 1 - \frac{1}{T} \tag{3.9}$$

Where,

α = Scale Parameter

k= Shape Parameter

μ= Location Parameter

$X =$ Variate

T = Return period

F= Frequency

3.6.2 GENERALIZED LOGISTICS DISTRIBUTION [2]

Generalized logistics distribution is a three-parameter distribution and given by (Hosking, 1986), the PDF, CDF, and Quantile function of the Distribution is given by:

The probability density function (PDF) is provided by:

$$F(X) = \frac{1}{\alpha}\left[1 - k\left(\frac{X-\mu}{\alpha}\right)\right]^{\left(\frac{1}{k}-1\right)}\left[1 + \left\{1 - k\left(\frac{X-\mu}{\alpha}\right)\right\}^{1/k}\right]^{-2} \tag{3.10}$$

The cumulative density function (CDF) is given by:

$$F(X) = \left[1 + \left\{1 - k\left(\frac{X-\mu}{\alpha}\right)\right\}^{1/k}\right]^{-1} \tag{3.11}$$

Range: $For\ k < 0 \quad \mu + \frac{\alpha}{k} \leq X < \infty$

$$For\ k > 0 \quad -\infty < X \leq \mu + \frac{\alpha}{k}$$

The Quantile Function (X_P) is given by:

$$X_P = \mu + \frac{\alpha}{k}\left[1 - \left(\frac{1-F}{F}\right)^{k}\right] \tag{3.12}$$

$$F = 1 - \frac{1}{T} \tag{3.13}$$

Where,

α = Scale Parameter

k= Shape Parameter

μ= Location Parameter

X = Variate

T = Return period

3.6.3 GENERALIZED PARETO DISTRIBUTION [2]

Generalized Pareto Distribution is a three-parameter distribution, and the PDF, CDF, and Quantile function of the Distribution is given by:

The probability density function (PDF) is provided by:

$$F(X) = \frac{1}{\alpha}\left[1 - k\left(\frac{X-\mu}{k}\right)\right]^{\frac{1}{k}-1} \tag{3.14}$$

The cumulative density function (CDF) is given by:

$$F(X) = 1 - \left[1 - k\left(\frac{X-\mu}{k}\right)\right]^{1/k} \qquad (3.15)$$

Range:

$$For\ k < 0\ \ \mu \le X < \infty$$

$$For\ k > 0\ \ \mu \le X \le \mu + \frac{\alpha}{k}$$

The Quantile Function (X_P) is given by:

$$X_P = \mu + \frac{\alpha}{k}[1 - (1 - F)^k] \qquad (3.16)$$

$$F = 1 - \frac{1}{T} \qquad (3.17)$$

Where,

α = Scale Parameter

k= Shape Parameter

μ= Location Parameter

X= Variate

T = Return period

3.6.4 UNIFIED EXTREME VALUE DISTRIBUTION [1]

Unified extreme value distribution is a three-parameter distribution and given by (Sushil k. Singh, 2016), the CDF and Quantile function of the distribution is given by:

The cumulative density function (CDF) is provided by:

$$F(X) = exp\left[-\left(k\,\frac{X-w}{c}\right)^{\left(-\frac{1}{k}\right)}\right] \qquad (3.18)$$

The Quantile Function ($X_{.T.}$) is given by:

$$X_T = w + \frac{c}{k}exp\left[-kln\left\{ln\left(\frac{T}{T-1}\right)\right\}\right] \qquad k \ne 0 \qquad (3.19)$$

$$X_T = \mu_1 - cln\left[ln\left(\frac{T}{T-1}\right)\right] \qquad k = 0 \qquad (3.20)$$

$$T = \frac{1}{1-F} \qquad (3.21)$$

Where,

c = Scale Parameter

k= Shape parameter

w = Location Parameter

X= Variate

T = Return period

3.7 PARAMETER ESTIMATION METHOD

In this study, a two-parameter estimation method used for GEV, GLO, GPA distribution first method of moment, and second probability weighted moment with L- moment and for UEV distribution simple objective method is used.

3.7.1 METHOD OF MOMENTS [7, 8, 10]

The method of moments makes use of the fact that if all the moments of Distribution are known, then everything about the Distribution is known.

3.7.1.1 GENERALIZED EXTREME VALUE DISTRIBUTION [7, 10]

In this study, GEV parameters are estimated using Easy Fit 5.6 standard software.

The MOM estimators of the G.E.V. parameters are given by Stedinger et al. (1993) as:

$$\mu = x' + \frac{\alpha[\Gamma(1+k)-1]}{k} \tag{3.22}$$

$$\alpha = \frac{\sigma k}{\{[\Gamma(1+2k)]-[\Gamma(1+k)]^2\}^{\frac{1}{2}}} \tag{3.23}$$

$$C_S = sign(k)\frac{\left\{\left[-\Gamma(1+3k)+3\Gamma(1+k)\Gamma(1+2k)-2\left(\Gamma(1+k)\right)^3\right]\right\}}{\left[\Gamma(1+2k)-\left(\Gamma(1+k)\right)^2\right]^{\frac{3}{2}}} \tag{3.24}$$

Where,

A sign is '+' or '-' depending upon the value of k estimate

α = Scale Parameter

k= Shape Parameter

μ= Location parameter

C_S= Skewness

3.7.1.2 GENERALIZED LOGISTICS DISTRIBUTION [8]

The parameter estimates of the GLO distribution using MOM estimation method given as:

Region 1: -10< C_S <10;-1/3< k <1/3

$$k = \frac{2}{3\pi}\tan^{-1}(-0.59484C_S) \quad (3.25)$$

Region 2: 0< C_S <10; 1/3< k <1/2

$$k = \frac{1}{3\pi}\tan^{-1}(0.03688 - 0.29824C_S) + \frac{1}{2} \quad (3.26)$$

Region 3: -10< C_S <0;-1/2< k <-1/3

$$k = \frac{1}{3\pi}\tan^{-1}(0.036884 - 0.29824C_S) - \frac{1}{2} \quad (3.27)$$

$$\alpha = \frac{x'kC_V}{[g_2-(g_1)^2]^{1/2}} \quad (3.28)$$

$$\mu = x' - \frac{\alpha}{k}(1 - g_1) \quad (3.29)$$

$$g_1 = \Gamma(1+k)\Gamma(1-k) \quad (3.30)$$

$$g_2 = \Gamma(1+2k)\Gamma(1-2k) \quad (3.31)$$

Where,

x' =Mean

α = Scale Parameter

k = Shape Parameter

μ= Location parameter

C_S= Skewness

C_V =Coefficient of variation

3.7.1.3 GENERALIZED PARETO DISTRIBUTION [7]

The parameter estimates of the GP distribution using MOM estimation method given as:

$$C_S = \frac{2(1-k)(1+2k)^{1/2}}{(1+3k)} \quad (3.32)$$

$$\alpha = \sigma[(1+k)^2(1+2k)]^{1/2} \quad (3.33)$$

$$\mu = x' - \frac{\alpha}{(1+k)} \quad (3.34)$$

Where,

x'= Mean

σ =Standard deviation

α = Scale Parameter

k = Shape Parameter

μ= Location parameter

C_S= Skewness

3.7.2 PROBABILITY WEIGHTED MOMENT WITH L- MOMENT [2]

L-Moment Theory and Statistics

L-moments can obtain by considering linear combinations of the observation in a sample of data that has been arranged in ascending order [2]. Consider measurement of the shape of a distribution, given a small sample drawn from the Distribution. Denote by $X_{1:1} \leq X_2$:n≤ …… $X_{n:n}$ (Hosking & Wallis, 1997; Eregno, 2014).The necessary steps in the determination of L-Moment statistics described below:

Step One: Computation of probability-weighted moments of Distribution (PWMS)

Probabilities weighted moments needed for the calculation of L-moment. The data must first rank in ascending order of magnitude, after that; the following equations proposed by (Cunnane, 1989) can thus be applied [2]:

$$b_0 = \frac{1}{N}\sum_{i=1}^{n} X_{(i:n)} \tag{3.35}$$

$$b_1 = \frac{1}{N}\sum_{i=2}^{n} X_{(i:n)} \left[\frac{(i-1)}{(n-1)}\right] \tag{3.36}$$

$$b_2 = \frac{1}{N}\sum_{i=3}^{n} X_{(i:n)} \left[\frac{(i-1)(i-2)}{(n-1)n-2)}\right] \tag{3.37}$$

$$b_3 = \frac{1}{N}\sum_{i=4}^{n} X_{(i:n)} \left[\frac{(i-1)(i-2)(i-3)}{(n-1)(n-2)(n-3)}\right] \tag{3.38}$$

Where:

X_i represents the ranked annual maximum series in which X_1 is the smallest precipitation, and X_n is the largest. The parameters (b_0, b_1, b_2, b_3) can easily be determined by using the developed Microsoft excel algorithm.

Step Two: Computation of L-Moment Values

L-moment values are easily calculated in terms of the probability-weighted moment (PWMs). In particular, the first four L-moment values given as follows (Hosking & Wallis, 1997).

$$\lambda_1 = b_0 \tag{3.39}$$
$$\lambda_2 = (2b_1 - b_0) \tag{3.40}$$
$$\lambda_3 = (6b_2 - 6b_1 + b_0) \tag{3.41}$$
$$\lambda_4 = 20b_3 - 30b_2 + 12b_1 - b_0 \tag{3.42}$$

The parameters ($\lambda_1, \lambda_2, \lambda_3, \lambda_4$) can easily be determined by using the developed Microsoft excel algorithm that requires forty year's annual maximum daily rainfall data.

Step Three: Computation of L-Moment Ratio

L-Moment ratio used for expressing the parameter estimates are as follows (Hosking and Wallis, 1997) [2]. The parameters (τ_2, τ_3, τ_4) computed using the formula below (Hosking & Wallis, 1997; Gubareva and Gartsman, 2010).

L- Coefficient of variation $(\tau_2) = \lambda_1$ (3.43)

L-Skewness $(\tau_3) = \frac{\lambda_3}{\lambda_2}$ (3.44)

L-Kurtosis $(\lambda_4) = \frac{\lambda_4}{\lambda_3}$ (3.45)

The parameters (τ_2, τ_3, τ_4) can easily be determined by using the developed Microsoft excel algorithm that requires forty year's annual maximum daily rainfall data.

3.7.2.1 GENERALIZED EXTREME VALUE DISTRIBUTION

L-Moment parameter estimates equations

$$k = 7.8590C + 2.9554C^2 \tag{3.46}$$

$$C = \frac{2}{3+\tau_3} - \frac{ln2}{ln3} \tag{3.47}$$

$$\alpha = \frac{\lambda_2 k}{\Gamma(1+k)\Gamma(1-2^{-k})} \tag{3.48}$$

$$\mu = \lambda_1 + \frac{\alpha[\Gamma(1+k)-1]}{k} \tag{3.49}$$

3.7.2.2 GENERALIZED LOGISTICS DISTRIBUTION

L-Moment parameter estimates equations

$$k = -\tau_3 \tag{3.50}$$

$$\alpha = \frac{\lambda_2}{\Gamma(1+\kappa)\Gamma(1-\kappa)} \tag{3.51}$$

$$\mu = \lambda_1 + \frac{(\lambda_2-\alpha)}{k} \tag{3.52}$$

3.7.2.3 GENERALIZED PARETO DISTRIBUTION

L-Moment parameter estimates equations

$$k = \frac{(1-3\tau_3)}{(1+\tau_3)} \tag{3.53}$$

$$\alpha = \lambda_2[(k+1)(k+2)] \tag{3.54}$$

$$\mu = \lambda_1 - \lambda_2(2+K) \tag{3.55}$$

3.7.3 SIMPLE OBJECTIVE [1]

3.7.3.1 UNIFIED EXTREME VALUE DISTRIBUTION

Simple objective parameter estimates equation for unified extreme value distribution:

$$k = \frac{[E(XY) - E(X)E(Y)]}{E(X^2) - [E(X)]^2} \tag{3.56}$$

$$w = E(X) - \frac{E(Y)}{k} \tag{3.57}$$

$$c = \frac{1}{n-m} \sum_{i=m}^{n-1} \frac{k(X_i - w)}{exp\left[kln(lnF_i^{-1})^{-1}\right]} \tag{3.58}$$

$$Y_i = \frac{dx}{d[\ln(lnF^{-1})^{-1}]} \tag{3.59}$$

$$x = \frac{x_{i-1} + x_{i+1}}{2} \tag{3.60}$$

$$E(z) = \frac{1}{(n-m)} \sum_{i=m}^{n-1} z_i \tag{3.61}$$

Where,

E (z) = expected value of a variable z

n = no. of observations

3.7.4 GRAPHICAL METHOD

The Graphical Method draws a graph between observed annual maximum daily rainfall and predicted annual maximum daily rainfall based on different extreme value distribution and calculates the coefficient of determination (R^2), in which extreme value distribution has the highest value of R^2 a best-fit distribution for that case.

3.8 GOODNESS OF FIT [2]

The GOF statistical model used to find the most appropriate distribution for a given set of observations; in this study, five different GOF method applies, and which are given below:

1. Relative root means square error (RRMSE)
2. Root mean square error (RMSE)
3. Maximum absolute error (MAE)
4. Mean absolute deviation index (MADI)
5. Probability plot correlation coefficient (PPCC)

The first four methods involve the assessment of the difference between the observed values and the expected values under the assumed Distribution. In contrast, the fifth method consists of measuring the correlation between the ordered values and the associated expected values.

The overall goodness of fit of each Distribution is judged using a ranking scheme by comparing the three categories of test criteria based on the relative magnitude of the statistical test results. The Distribution with the lowest RRMSE, lowest RMSE, lowest MAE, lowest MADI, and highest PPCC was assigned a score of 4, the next given the rating 3, 2, while the worst given the score 1 [2]. The overall score of each Distribution was obtained by summing the individual point scores, and the Distribution with the highest total point score elected as the best-fit distribution model.

3.8.1 RELATIVE ROOT MEANS SQUARE ERROR

Relative root means the square error is a relative difference between values predicted by a model and values observed. The Distribution with the lowest RRMSE assigned score 4, the next was given the score 3, 2, while the worst given the score 1.

$$RRMSE = \left[\frac{\sum\left(\frac{X_i - Y_i}{X_i}\right)^2}{n-m}\right]^{\left(\frac{1}{2}\right)} \tag{3.62}$$

3.8.2 ROOT MEAN SQUARE ERROR

Root mean square error is a frequently used measure of the difference between values predicted by a model and values observed [2]. A lower RMSE is better than a higher one. The Distribution with the lowest RMSE assigned score 4, the next was given the score 3, 2, while the worst given the score 1.

$$RMSE = \left[\frac{\sum(X_i - Y_i)^2}{n-m}\right]^{\left(\frac{1}{2}\right)} \tag{3.63}$$

3.8.3 MAXIMUM ABSOLUTE ERROR

A maximum absolute error is the absolute maximum difference between observed and estimated values. The distribution with low MAE assigned score 4, the next given the rating 3, 2, while the worst given score 1.

$$MAE = Max(|X_i - Y_i|) \tag{3.64}$$

3.8.4 MEAN ABSOLUTE DEVIATION INDEX

The mean absolute deviation index is the relative absolute difference between observed and estimated values divided by the number of observations. The Distribution with the lowest MADI assigned score 4, the next given the rating 3, 2, while the worst given the score 1.

$$MADI = \frac{1}{N}\left|\frac{X_i - Y_i}{X_i}\right| \tag{3.65}$$

3.8.5 PROBABILITY PLOT CORRELATION COEFFICIENT

The validity of the probability scheme citation (filliben 1975) is a graphical way of identifying a distribution family status parameter that best describes a set of data [2]. The Distribution with the highest PPCC assigned score 4, the next given the rating 3, 2, while the worst given the score 1.

$$PPCC = \frac{\sum\left[\left(X_i - X'\right)\left(Y_i - Y'\right)\right]}{\left[\sum(X_i - X')^2 \sum(Y_i - Y')^2\right]^{\left(\frac{1}{2}\right)}} \tag{3.66}$$

3.9 CONCLUSION

In this chapter, using the statistical approach for the present study is carried out on two rainfall subdivision in Uttar Pradesh. Rainfall data of West Uttar Pradesh subdivision and East Uttar Pradesh subdivision collected in the form of annual maximum daily Rainfall of 40 years from the IITM datasheet. Before using various distributions first, you should check the outliers, homogeneity, normality, and stationarity of the data using various tests. The Method of Moment and Probability Weighted Moment with L- Moments use to calculate the parameter of 1st three distribution and UEV distribution parameters calculated by SO method. Their probability distribution function and their cumulative distribution function and found out the general expression used to evaluate the results in the form of tables and also using the graphical approach, which also gives the coefficient of determination for all the four Distribution, which discussed above. After getting all the results from the various Distribution than applying the Goodness of fit test (RRMSE, RMSE, MAE, MADI, PPCC) which used to compare all the Distribution or saying that the which Distribution is the best fit among these four distributions by using the various error values of all the Distribution.

CHAPTER 4.0

RESULTS AND DISCUSSIONS

4.1 GENERAL

The result of annual maximum daily rainfall data series (1979-2018) by using various distributions in the meteorological subdivision of Uttar Pradesh, i.e., west Uttar Pradesh and East Uttar Pradesh subdivision from the Uttar Pradesh region taken for study. At first, checked out that our rainfall data is independent or not, it's also saying that the rainfall data is homogeneous or not and any outliers are present or not. It is essential for further computation by using the statistical approach. Later on, applying the Flood frequency formulae using a different distribution used to compute or evaluate the Rainfall for the corresponding return period. The results of the statistics and Goodness of fit are discussed here.

4.2 SCREENING OF DATASETS

It is necessary to screen the data which you analyze further. Otherwise, your prediction of Rainfall is incorrect. The results of various screening of data sets are discussed below.

4.2.1 TEST FOR OUTLIERS

It is necessary to test for outliers of the data which you analyze further. Otherwise, your prediction of Rainfall is incorrect. All tests for outliers done in XLSTAT software and obtained results are shown given below in tables 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, and fig 4.1, 4.2, 4.3, 4.4.

4.2.1.1 CASE 1- WEST UTTAR PRADESH SUBDIVISION

There are 19 rain gauge stations in West Uttar Pradesh subdivision and 40 years (1979-2018) annual max Daily rainfall data are used for the analysis tests for outliers. Two tests of outliers used in this study, the first Dixon test and the second Grubbs test, both tests performed in XLSTAT software.

Table 4.1: Summary of Descriptive Statistics of West Uttar Pradesh annual maximum daily rainfall (1979-2018) data set

Variable	**Observations**	**Missing data**	**Obs. Without missing data**	**Minimum**	**Maximum**	**mean**	**Std. deviation**
Rainfall	40	0	40	123.700	420.500	281.910	77.669

4.2.1.1.1 DIXON TEST

Dixon test for outliers (two-tailed test)

Table 4.2: Summary of the Dixon test result values for case 1

R10 (Observed Value)	0.057
R10 (Critical value)	0.273
p-value (two-tailed)	0.829
Alpha	0.05

the p-value has computed using 1000000 Monte Carlo simulations for Dixon test, 99% confidence interval on the p-value]0.828, 0.830[

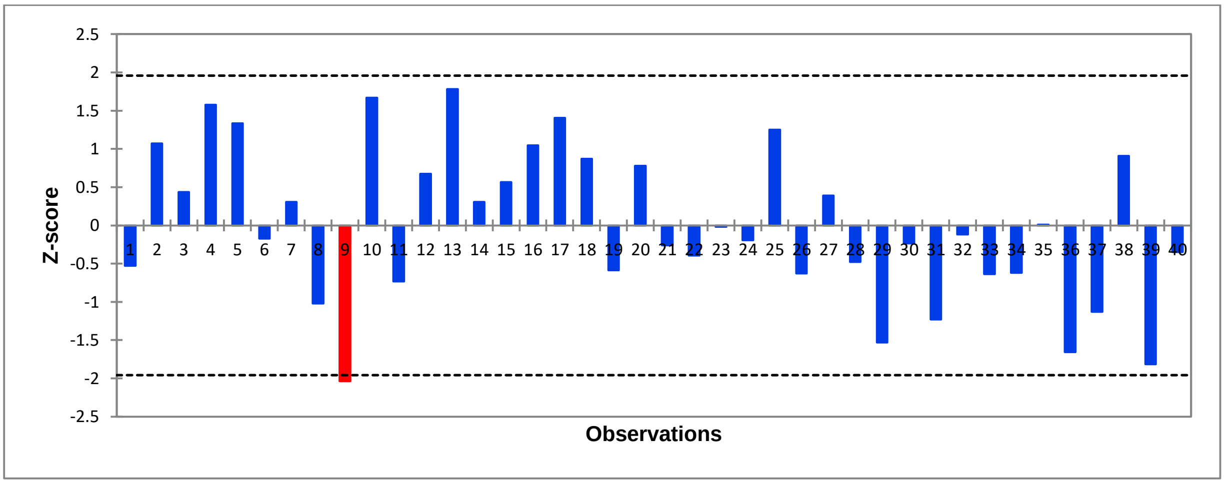

Fig 4.1: Z-scores of annual maximum daily rainfall data set of West Uttar Pradesh subdivision

- The result shows that One outlier is present in series, i.e., 123.700(Rainfall) - **-2.037(Z-Score)**, but it is because of tail end event or minimum value of rainfall in series. So it is not removed from data; all data considered for further analysis.
- As the computed p-value is greater than the significance level alpha (0.829>0.05), one cannot reject the null hypothesis Ho. The risk to reject the null hypothesis Ho while it is true is 82.95%. So, there is no outlier in the data.

4.2.1.1.2 GRUBBS TEST

Grubbs test for outliers (two-tailed test)

Table 4.3: Summary of Grubbs test result values for case 1

G (observed value)	2.037
G (critical value)	3.036
p-value (two- tailed)	1.000
alpha	0.05

99% confidence interval on the p-value] 1.000, 1.000[

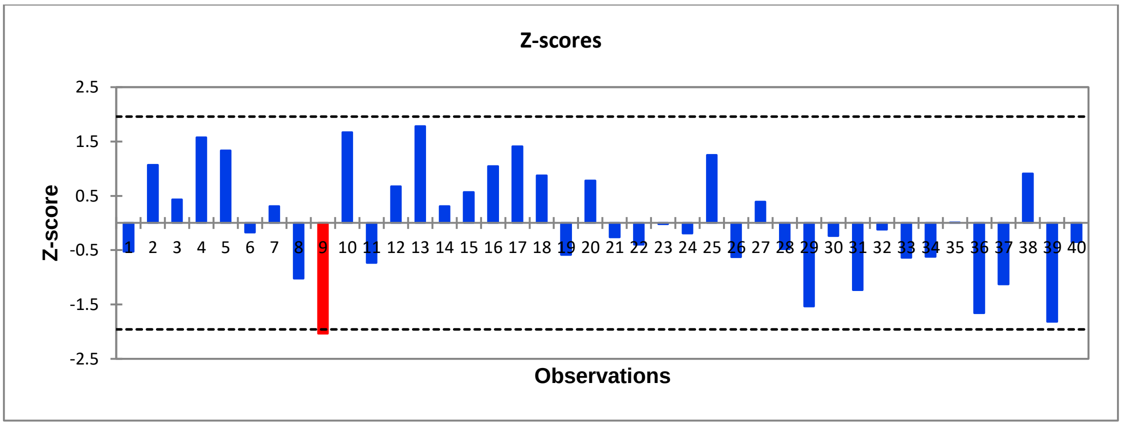

Fig 4.2: Z-scores of annual maximum daily rainfall data set of West Uttar Pradesh subdivision

- The result shows that One outlier is present in series, i.e., 123.700(Rainfall) - **-2.037(Z-Score)**, but it is because of tail end event or minimum value of rainfall in series. So it is not removed from data; all data consider for further analysis.
- As the computed p-value is greater than the significance level alpha (1.00>0.05), one cannot reject the null hypothesis Ho. The risk to reject the null hypothesis Ho, while it is true is 100% So, there is no outlier in the data.

- For both the test, there are no outliers in annual maximum daily rainfall data in West Uttar Pradesh subdivision, so all data used for analysis and forecasting of rainfall for different return periods.

4.2.1.2 CASE 2 – EAST UTTAR PRADESH SUBDIVISION

There are 26 rain gauge stations in East Uttar Pradesh subdivision and 40 years (1970-2018) annual max daily rainfall data are used for the analysis tests for outliers. Two tests of outliers used in this thesis, the first Dixon test, and the second Grubbs test, both tests are performed in XLSTAT software.

Table 4.4: Summary of Descriptive Statistics of East Uttar Pradesh annual maximum daily rainfall (1979-2018) data set

Variable	**observations**	**Obs. With missing data**	**Obs. Without missing data**	**minimum**	**Maximum**	**Mean**	**Std. deviation**
Rainfall	40	0	40	177.00	480.30	301.589	69.814

4.2.1.2.1 DIXON TEST

Dixon test for outliers (two-tailed)

Table 4.5: Summary of the Dixon test result values for case 2

R10 (Observed Value)	0.098
R10 (Critical value)	0.273
p-value (Two-tailed)	0.753
alpha	0.05

The p-value has computed using 1000000 Monte Carlo simulations for Dixon Test.

99% confidence interval on the p-value:] 0.752, 0.754[

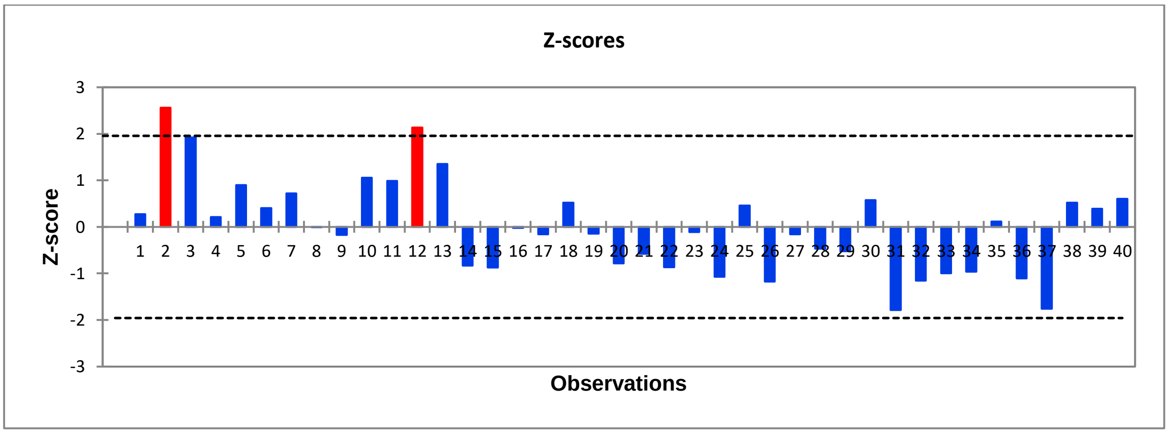

Fig 4.3: Z-scores of annual maximum daily rainfall data set of East Uttar Pradesh subdivision

- The result shows that Two outliers are present in series, i.e., 480.300(Rainfall) - **2.560(Z-Score)**, and 450.650 (Rainfall) – **2.135(Z-Score)**. But both are because of tail end event or maximum value of rainfall in series. So it is not removed from data; all data considered for further analysis.
- As the computed p-value is greater than the significance level alpha (0.753>0.05), one cannot reject the null hypothesis Ho. The risk to reject the null hypothesis Ho while it is true is 75.32%. So, there is no outlier in the data.

4.2.1.2.2 GRUBBS TEST

Grubbs test for outliers / Two-tailed test:

Table 4.6: Summary of Grubbs test result values case 2

G (Observed Value)	2.560
G (Critical value)	3.036
p-value (Two-tailed)	0.309
alpha	0.05

99% confidence interval on the p-value] 0.308, 0.310[

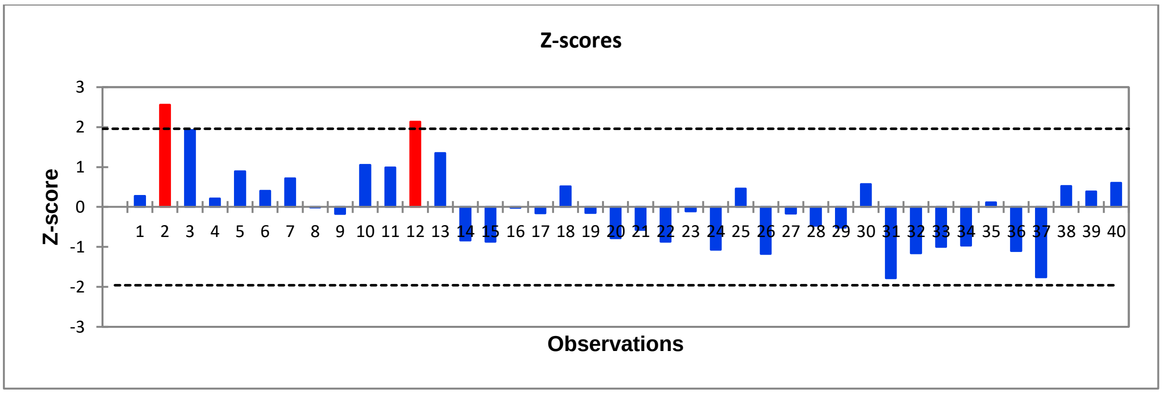

Fig 4.4: Z-scores of annual maximum daily rainfall data set of East Uttar Pradesh subdivision

- The result shows that Two outliers are present in series, i.e., 480.300(Rainfall) - **2.560(Z-Score)**, and 450.650 (Rainfall) – **2.135(Z-Score)**. But both are because of tail end event or maximum value of rainfall in series. So it is not removed from data; all data considered for further analysis.
- As the computed p-value is greater than the significance level alpha (0.309>0.05), one cannot reject the null hypothesis Ho. The risk to reject the null hypothesis Ho while it is true is 30.92%. So, there is no outlier in the data.

- For both the test, there are no outliers in annual maximum daily rainfall data in East Uttar Pradesh Subdivision. So data is used for analysis and forecasting of rainfall for different return periods.

4.2.2 STATIONARITY TEST

It is necessary to test the stationarity of the data, which you analyze further. Otherwise, your prediction of Rainfall is incorrect. All stationarity test is done in XLSTAT software, and obtained results are shown in given below tables 4.7, 4.8, 4.9, 4.10, 4.11, 4.12, 4.13, 4.14, 4.15, and 4.16.

4.2.2.1 CASE 1- WEST UTTAR PRADESH SUBDIVISION

There are 19 rain gauge stations in West Uttar Pradesh subdivision, and 40 years annual maximum daily rainfall data are used for the analysis of stationarity tests. Three stationarity tests used in this study, first Augmented Dickey-Fuller Test, second Kwiatkowski Phillips, Schmidt and shin Test, and last Phillips-Perron Test, all tests are performed in XLSTAT software.

Table 4.7: Summary of Descriptive Statistics of West Uttar Pradesh annual maximum daily rainfall data set

Variable	Observations	Obs. with missing data	Obs. without missing data	Minimum	Maximum	Mean	Std. deviation
RAINFALL	40	0	40	123.700	420.500	281.910	77.669

4.2.2.1.1 AUGMENTED DICKEY-FULLER TEST

Dickey-Fuller test (ADF (stationary) / k: 3 / RAINFALL):

Table 4.8: Summary of Dickey –Fuller test result values for case1

Tau (Observed value)	-3.359
Tau (Critical value)	-3.495
p-value (one-tailed)	0.045
alpha	0.05

4.2.2.1.2 KWIATKOWSKI PHILLIPS, SCHMIDT AND SHIN TEST

KPSS test (Level / Lag Short / RAINFALL):

Table 4.9: Summary of KPSS test result values for case 1

Eta (Observed Value)	0.732
Eta (Critical value)	0.446
p-value (one-tailed)	0.066
alpha	0.05

4.2.2.1.3 PHILLIPS-PERRON TEST

Phillips-Perron test (PP (no intercept) / Lag: Short / RAINFALL):

Table 4.10: Summary of Phillips-Perron test result values for case 1

Tau (Observed value)	-5.731
Tau (Critical value)	-2.939
p-value (one-tailed)	<0.0001
alpha	0.05

Table 4.11: Summary of stationarity test result p-values (Two-tailed)

	ADF	**KPSS**	**PP**
RAINFALL	0.045	0.066	<0.0001

- The results show that for – Significance level (%): 5, as the computed p-value of the ADF test, are lower than the significance level alpha (0.045<0.05), one should reject the null hypothesis Ho, and accept the alternative hypothesis, Ha. So ADF test results show that there is no unit root for the series, and series is stationary.
- As the computed p-value of the KPSS test is greater than the significance level alpha (0.066>0.05), one cannot reject the null hypothesis Ho, So the KPSS test result shows that the series is stationary.
- As the computed p-value of the PP test is lower than the significance level alpha (0.0001<0.05), one should reject the null hypothesis Ho, and accept the alternative hypothesis Ha. So PP test results show that there is no unit root for the series, and the series is stationary.

- All three tests show that the series is stationary, so the stationarity of our data was approved to make forecasting of extreme rainfall.

4.2.2.2 CASE 2- EAST UTTAR PRADESH SUBDIVISION

There are 26 rain gauge stations in East Uttar Pradesh subdivision, and 40 years annual maximum daily rainfall data are used for the analysis of stationarity tests. Three stationarity tests used in this study, first Augmented Dickey-Fuller Test, second Kwiatkowski Phillips, Schmidt and shin Test, and last Phillips-Perron Test, all tests are performed in XLSTAT software.

Table 4.12: Summary of Descriptive Statistics of East Uttar Pradesh annual maximum daily rainfall data set

Variable	**Observations**	**Obs. with missing data**	**Obs. without missing data**	**Minimum**	**Maximum**	**Mean**	**Std. deviation**
R	40	0	40	177.000	480.300	301.589	69.814

4.2.2.2.1 AUGMENTED DICKEY-FULLER TEST

Dickey-Fuller test (ADF (stationary) / k: 3 / R):

Table 4.13: Summary of Dickey –Fuller test result values for case 2

Tau (Observed value)	-3.057
Tau (Critical value)	-3.495
p-value (one-tailed)	0.026
alpha	0.05

4.2.2.2.2 KWIATKOWSKI PHILLIPS, SCHMIDT AND SHIN TEST

KPSS test (Level / Lag Short / R):

Table 4.14: Summary of KPSS test result values for case 2

Eta (Observed Value)	0.928
Eta (Critical value)	0.446
p-value (one-tailed)	0.563
alpha	0.05

4.2.2.2.3 PHILLIPS-PERRON TEST

Phillips-Perron test (PP (no intercept) / Lag: Short / R):

Table 4.15: Summary of Phillips- Perron test result values for case 2

Tau (Observed value)	-3.761

Tau (Critical value)	-2.939
p-value (one-tailed)	0.007
alpha	0.05

Table 4.16: Summary of stationarity test result p-values (Two-tailed)

	ADF	**KPSS**	**PP**
RAINFALL	0.026	0.563	0.007

- The results show that, for Significance level (%):5, as the computed p-value of the ADF test is lower than the significance level alpha ($0.026<0.05$), one should reject the null hypothesis Ho, and accept the alternative hypothesis Ha. So ADF test results show that there is no unit root for the series, and series is stationary.
- As the computed p-value of the KPSS test is greater than the significance level alpha ($0.563>0.05$), one cannot reject the null hypothesis Ho, So the KPSS test result shows that the series is stationary.
- As the computed p-value of the PP test is lower than the significance level alpha ($0.007<0.05$), one should reject the null hypothesis Ho, and accept the alternative hypothesis, Ha. So PP test results show that there is no unit root for the series, and the series is stationary.

- All three tests show that the series is stationary, so the stationarity of our data was approved to make forecasting of extreme rainfall.

4.2.3 NORMALITY TEST

It is essential to test the normality of the data, which you analyze further. Otherwise, your prediction of Rainfall is incorrect, all Normality test is done in XLSTAT software, and obtained results shown in given below tables 4.17, 4.18, 4.19, 4.20, and fig 4.5, 4.6, 4.7, 4.8.

4.2.3.1 CASE 1- WEST UTTAR PRADESH SUBDIVISION

There are 19 rain gauge stations in West Uttar Pradesh subdivision, and 40 years annual maximum rainfall data are used for the analysis of normality tests. The normality test was used in this study, i.e., the Shapiro-Wilk test and the Anderson-Darling test, which performed in XLSTAT software.

4.2.3.1.1 SHAPIRO-WILK TEST

Shapiro-Wilk test (RAINFALL):

Table 4.17: Shapiro-Wilk test result values for case 1

W	0.977
p-value (Two-tailed)	0.591
alpha	0.05

- ➢ The results show that the Significance level (%): 5, the computed p-value is higher than the significance level alpha (0.591>0.05), one cannot reject the null hypothesis Ho. The risk to reject the null hypothesis Ho while it is true is 59.10%, so the variable from which the sample extracted follows a Normal distribution.

4.2.4.1.2 ANDERSON-DARLING TEST

Anderson-Darling test (R):

Table 4.18: Anderson-Darling test result values for case 1

A^2	0.246
p-value (Two-tailed)	0.742
alpha	0.05

- ➢ The result shows that the Significance level (%): 5, the computed p-value is higher than the significance level alpha (0.742>0.05), one cannot reject the null hypothesis Ho. The risk to reject the null hypothesis Ho while it is true is 74.23%, so the variable from which the sample extracted follows a Normal distribution. The output p-p plot and Q-Q plot shown in fig 4.5, 4.6.
- ➢ **P-P plots**-show that the series follows a normal distribution, the points lie along the first bisector of the plan

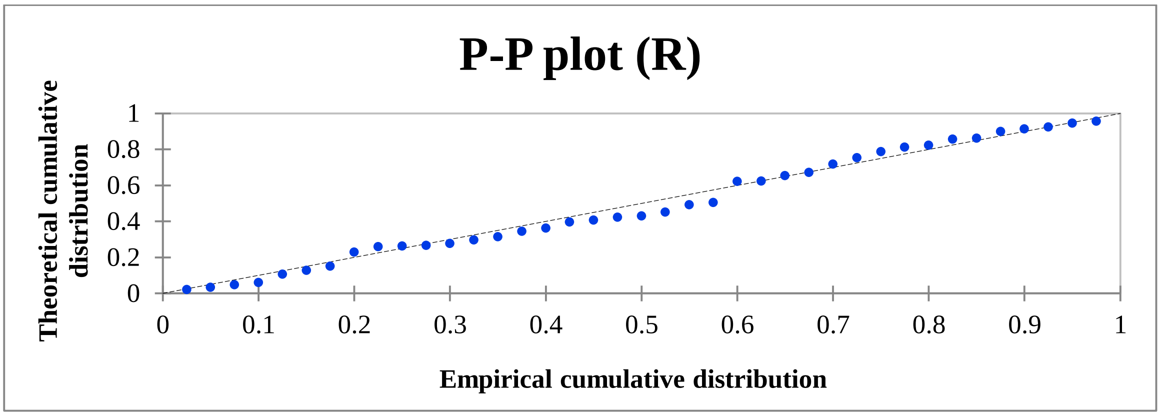

Fig 4.5: p-p plot (Rainfall) for case 1

- ➢ **Q-Q plots**-show that the series follows a normal distribution, the points lie along the first bisector of the plan.

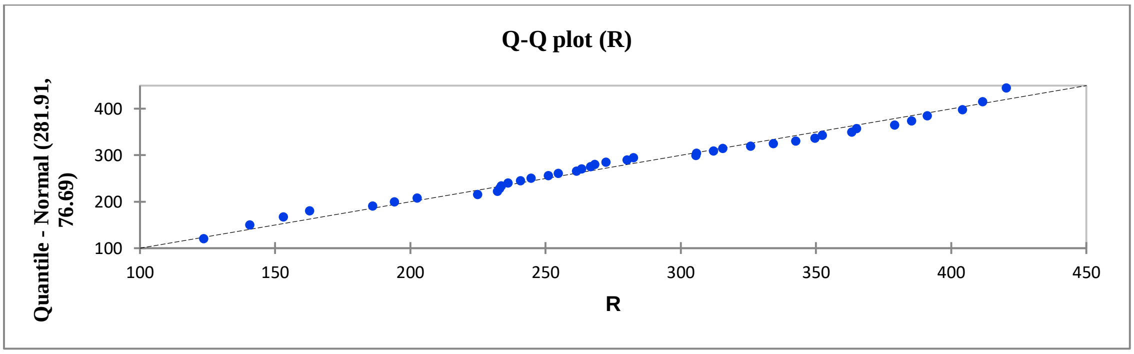

Fig 4.6: Normal Q-Q plot (Rainfall) for case 1

4.2.4.2 CASE 2- EAST UTTAR PRADESH SUBDIVISION

There are 26 rain gauge stations in East Uttar Pradesh subdivision, and 40 years of rainfall data are used for the analysis of normality tests. The normality test was used in this study, i.e., the Shapiro-Wilk test and the Anderson-Darling test, which performed in XLSTAT software.

4.2.4.2.1 SHAPIRO-WILK TEST

Shapiro-Wilk test (RAINFALL):

Table 4.19: Summary of Shapiro-Wilk test result for case 2

W	0.970
p-value (Two-tailed)	0.364
alpha	0.05

- The result shows that for Significance level (%): 5, as the computed p-value is greater than the significance level alpha (0.364>0.05), one cannot reject the null hypothesis Ho. The risk to reject the null hypothesis Ho while it is true is 36.41%, so the variable from which the sample extracted follows a Normal distribution.

4.2.4.1.2 ANDERSON-DARLING TEST

Anderson-Darling test (R):

Table 4.20: Anderson-Darling test result values for case 1

A^2	0.346
p-value (Two-tailed)	0.465
alpha	0.05

- The results show that for Significance level (%): 5, as the computed p-value is greater than the significance level alpha (0.465>0.05), one cannot reject the null hypothesis Ho. The risk to reject the null hypothesis Ho while it is true is 46.52%, so the variable from which the sample extracted follows a Normal distribution. The output p-p plot and Q-Q plot shown in fig 4.7, 4.8.
- **P-P plots**-show that the series follows a normal distribution, the points lie along the first bisector of the plan.

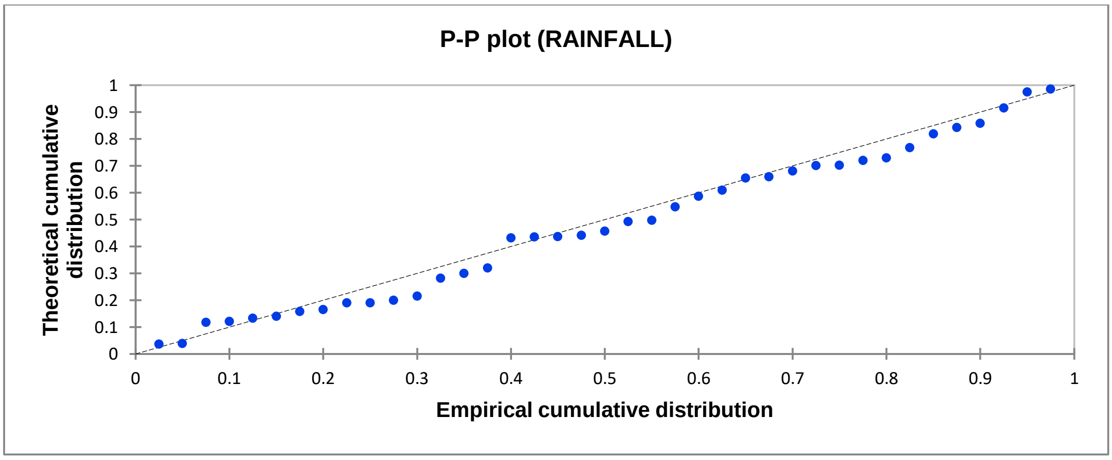

Fig 4.7: Normal p-p plot (Rainfall)

- **Q-Q plots-**show that the series follows a normal distribution, the points lie along the first bisector of the plan.

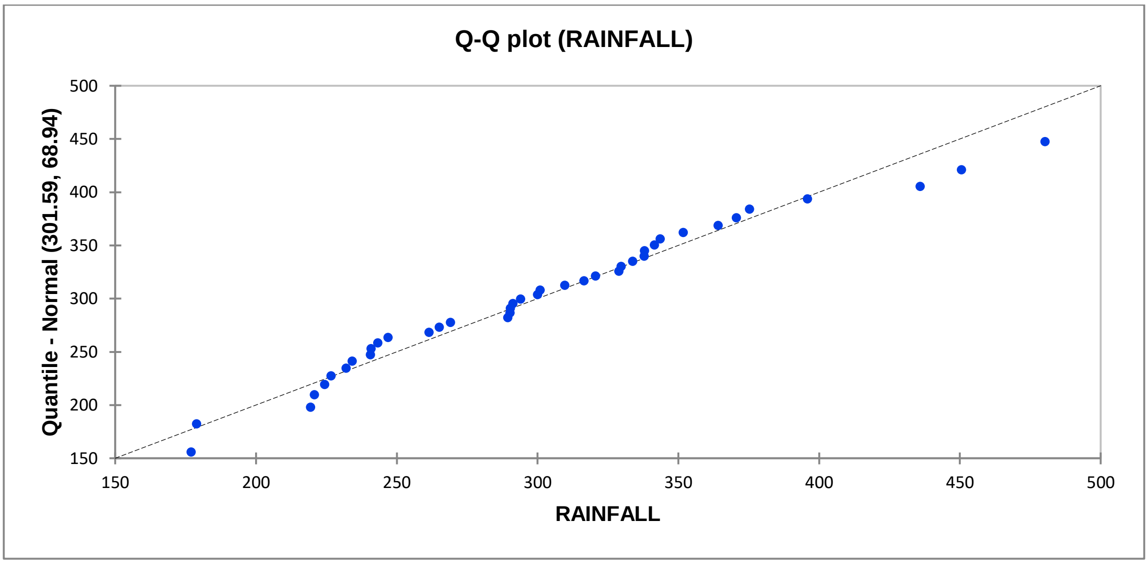

Fig 4.8: Normal Q-Q plots (Rainfall)

4.2.4 HOMOGENEITY TEST

It is necessary to test the homogeneity of the data, which you analyze further; otherwise, your prediction of rainfall is incorrect. All Homogeneity test is done in XLSTAT software and obtained results shown given below in tables 4.21, 4.22, 4.23, 4.24, 4.25, 4.26, 4.27, 4.28, 4.29, and 4.30.

4.2.3.1 CASE 1- WEST UTTAR PRADESH SUBDIVISION

There are 19 rain gauge stations in West Uttar Pradesh subdivision, and 40 years annual maximum rainfall data are used for the analysis of stationarity tests. Four stationarity tests were used in this study, first Pettit’s Test, second Standard Normal Homogeneity Test, third Buishand's Test, and last Von Neumann's test, all tests are performed in XLSTAT software.

4.2.3.1.1 PETTITT'S TEST

Pettitt's test (RAINFALL):

Table 4.21: Summary of Pettitt's test result values for case 1

K	215.000
t	2003
p-value (two tailed)	0.025
Alpha	0.05

The p-value has computed using 10000 Monte Carlo simulations, 99% confidence interval on the p-value] 0.021, 0.029[

4.2.3.1.2 STANDARD NORMAL HOMOGENEITY TEST

SNHT test (RAINFALL):

Table 4.22: Summary of SNHT test result values for case 1

T0	9.021
t	2005
p-value	0.059
Alpha	0.05

The p-value has computed using 10000 Monte Carlo simulations. Time elapsed: 0s, 99% confidence interval on the p-value:] 0.055, 0.063[

4.2.3.1.3 BUISHAND'S TEST

Buishand's test (RAINFALL):

Table 4.23: Summary of Buishand's test result values for case 1

R	9.251
p-value (Two-tailed)	0.084
alpha	0.05

The p-value has computed using 10000 Monte Carlo simulations. Time elapsed: 0s.

99% confidence interval on the p-value:] 0.077, 0.091[

4.2.3.1.4 VON NEUMANN'S RATIO TEST

Von Neumann's ratio test (RAINFALL):

Table 4.24: Summary of Von Neumann's ratio test result for case 1

N	1.807
p-value (Two-tailed)	0.275
alpha	0.05

The p-value has computed using 10000 Monte Carlo simulations, 99% confidence interval on the p-value] 0.269, 0.280[

The results show that for, Significance level (%): 5

Maximum time (s): 180

Number of simulations: 10000

Seed (random numbers): 4556916

- As the computed p-value of Pettitt's test is lower than the significance level alpha (0.025<0.05), one should reject the null hypothesis Ho, and accept the alternative hypothesis Ha. So there is a date at which there is a change in the data.
- As the computed p-value of the SNHT test is greater than the significance level alpha (0.059 >0.05), so SNHT test results show that data are homogenous.
- As the computed p-value of the Buishand test is greater than the significance level alpha (0.084>0.05), so Buishand test results show that data are homogenous.
- As the computed p-value of the Von Neumann test is greater than the significance level alpha (0.275>0.05), one cannot reject the null hypothesis Ho, so VNR test results show that data are homogeneous.

The obtained test results are classified by Wijngaard et al. theory [11]

1) Class A: Useful

   The series that rejects one or none null hypothesis under the four tests at a 5% significance level considered. Under this class, the series is grouped as homogeneous and can use for further analysis.

2) Class B: Doubtful

   The series that reject two null hypotheses of the four tests at a 5% significance level paced in this class. In this class, the series has an inhomogeneous signal and should be critically inspected before further analysis.

3) Class C: Suspect

   When there are three or all tests are rejecting the null hypothesis at 5% significance level, then the series is classified into this category. In this class, the series can be deleted or ignored before further analysis.

- Table 4.25 shows the results of the homogeneity tests for annual maximum rainfall in case 1. Based on the results, West Uttar Pradesh subdivision is homogenous since the null hypothesis for the SNHT, Buishand test, and VNR test not rejected at a 5% level of significance. Although the null hypothesis rejected in the Pettitt test, case 1 can still be considered as "useful" and can use for further analysis.

4.2.3.2 CASE 2- EAST UTTAR PRADESH SUBDIVISION

There are 26 rain gauge stations in East Uttar Pradesh subdivision, and 40 years annual max rainfall data are used for the analysis of stationarity test. Three stationarity tests used in this thesis, first Pettitt's Test, second Standard Normal Homogeneity Test, third Buishand's Test, all tests are performed in XLSTAT software.

4.2.3.2.1 PETTITT'S TEST

Pettitt's test (RAINFALL):

Table 4.26: Summary of Pettitt's test result values for case 2

K	279.000
t	1991
p-value (Two-tailed)	0.054
alpha	0.05

The p-value has computed using 10000 Monte Carlo simulation, 99% confidence interval on the p-value] 0.050, 0.059[

4.2.3.2.2 STANDARD NORMAL HOMOGENEITY TEST

SNHT test (RAINFALL):

Table 4.27: Summary of SNHT test result values for case 2

T0	17.361

T	1991
p-value (Two-tailed)	0.078
alpha	0.05

The p-value has computed using 10000 Monte Carlo simulations, 99% confidence interval on the p-value] 0.070, 0.080[

4.2.3.2.3 BUISHAND'S TEST

Buishand's test (RAINFALL):

Table 4.28: Summary of Buishand's test result value

R	14.305
p-value (Two-tailed)	0.031
alpha	0.05

99% confidence interval on the p-value] 0.026, 0.036[

4.2.3.2.4 VON NEUMANN'S RATIO TEST

Von Neumann's ratio test (RAINFALL):

Table 4.29: Summary of von Neumann's ratio test result value for case 2

N	1.116
p-value (Two-tailed)	0.185
alpha	0.05

The p-value has computed using 10000 Monte Carlo simulations. Time elapsed: 0s.

99% confidence interval on the p-value:] 0.180, 0.189[

Table 4.30: Summary of the homogeneity test result of the p-value

	Pettitt	**SNHT test**	**Buishand**	**von Neumann ratio**
RAINFALL	0.054	0.078	0.031	0.185

The result shows that for the Significance level (%): 5

Maximum time (s): 180

Number of simulations: 10000

Seed (random numbers): 369350839

- As the computed p-value of Pettitt's test is greater than the significance level alpha (0.054>0.05), one cannot reject the null hypothesis, Ho, so data are homogenous.
- As the computed p-value of the SNHT test is greater than the significance level alpha (0.078 >0.05), so SNHT test results show that data are homogenous.
- As the computed p-value of the Buishand test is lower than the significance level alpha (0.031<0.05), one should reject the null hypothesis Ho, and accept the alternative hypothesis Ha. So there is a date at which there is a change in the data.
- As the computed p-value of the Von Neumann ratio test is greater than the significance level alpha (0.185>0.05), one cannot reject the null hypothesis Ho, so VNR test results show that data are homogeneous.

- Table 4.30 shows the results of the homogeneity tests for annual maximum rainfall in case 2. Based on the results, East Uttar Pradesh subdivision is homogenous since the null hypothesis for the Pettitt test, SNHT, and VNR test does not reject at a 5% level of significance. Although the null hypothesis rejected in the Buishand test, case 2 can still be considered as “useful” and can use for further analysis.

4.3 DESCRIPTIVE STATISTICS

4.3.1 BASIC STATISTICS

Basics statistics, values and other statistics parameters value discussed below table 4.31 for both cases.

Table 4.31: Basic Statistical description of monthly Rainfall in West Uttar Pradesh Subdivision and East Uttar Pradesh Subdivision from 1979 to 2018

Statistics		**Case 1(West UP)**	**Case2 (East UP)**
Number of observations	N	40	40
Minimum	min	123.70	177.000
Maximum	max	420.50	480.300
1st Quartile	1st Q	233.65	242.700
3rd Quartile	3rd Q	344.37	338.975
Mean	*x*’	281.91	301.589
Variance	S^2	6032.519	4873.942
Standard deviation	σ	77.669	69.814
Coefficient of variation	C_V	0.275511	0.231486
Skewness coefficient	C_S	-0.06483	0.505543
Kurtosis coefficient	C_K	-0.68756	0.183983
Lower bound on the mean (95%)	L_{95}	257.07	279.261
Upper bound on the mean (95%)	U_{95}	306.75	323.916

4.3.2 L-MOMENT STATISTICS

L-moment statistics values are given below in table 4.32 for both cases.

Table 4.32: L-Moment statistics for case1 (West Uttar Pradesh Subdivision) and case2 (East Uttar Pradesh Subdivision)

L-Moment Statistics		**Case 1(West UP)**	**Case2 (East UP)**
Probability Weighted Moment Values	b_o	281.9100	301.59
	b_1	163.4067	170.5381
	b_2	116.4014	120.8559
	b_3	90.8429	94.2970
L-Moment Values	λ_1	281.91	301.59

	λ_2	44.9035	39.4875
	λ_3	0.12234	3.49535
	λ_4	3.78781	5.13166
L- moment ratios	L-CV	0.15928	0.13093
	L- Skewness	0.00272	0.08851
	L- Kurtosis	0.08435	0.12995

4.4 DISTRIBUTION PARAMETER

To estimate and compare the parameters of GEV, GLO, GPA distribution using the method of moment and probability Weighted Moment with L- Moment and Unified Extreme Value distribution parameters calculated using a simple objective method which all shown in table 4.33 and 4.34 respectively West Uttar Pradesh subdivision and East Uttar Pradesh subdivision.

4.4.1: CASE 1 WEST UTTAR PRADESH SUBDIVISION

Table 4.33: The Parameters of GEV, GLO, GPA, UEV using method of moment, probability-weighted moments with L-moments and simple objective method

Method	**Distribution**	**Shape parameter (k)**	**Scale parameter (α)**	**Location parameter (μ)**
MOM	GEV	0.15428	57.906	252.542
PWM-L	GEV	0.27975	75.06957	259.8059
MOM	GLO	-0.06888	42.08799	277.11
PWM-L	GLO	-0.00272	43.91309	246.4394
MOM	GPA	0.00654	78.68687	203.7344
PWM-L	GPA	0.98913	266.98653	147.68729
Simple Objective	UEV	-0.52383	74.10038	458.3881

4.4.2: CASE 2 EAST UTTAR PRADESH SUBDIVISION

Table 4.34: The Parameters of GEV, GAP, GLO, and UEV using MOM, PWM-L, and SO method

Method	**Distribution**	**Shape Parameter (k)**	**Scale Parameter (α)**	**Location Parameter (μ)**
MOM	GEV	-0.247000	47.74800	261.0700
PWM-L	GEV	0.131513	52.09239	258.7094
MOM	GPA	0.62013	169.2931	197.0952
PWM-L	GPA	0.67472	176.8808	195.9706
MOM	GLO	-0.39261	18.31237	287.26
PWM-L	GLO	-0.08852	38.92118	295.1906
Simple Objective	UEV	0.13095	34.3947	10.0801

4.5 OBSERVED AND PREDICTED ANNUAL MAXIMUM RAINFALL CALCULATED BY GEV, GPA, GLO, and UEV

By using four different extreme value distribution, GEV, GLO, GPA, and UEV distribution predict the maximum annual rainfall of 40 years for both cases. And for distribution parameter calculation using the Method of Moments and Probability Weighted Moment with L- Moment for GEV, GLO, GPA distribution, and for UEV distribution used simple objective method. All result is shown in table 4.35 and 4.36 respectively West Uttar Pradesh subdivision and East Uttar Pradesh subdivision.

4.5.1 CASE 1 WEST UTTAR PRADESH SUBDIVISION

Table 4.35: observed and predicted annual maximum rainfall calculated by GEV, GPA, GLO, and UEV using MOM, PWM-L, SO methods

Rank	**Observed annual maximum Rainfall (X_i)**	**Predicted annual maximum Rainfall (Y_i) calculated by GEV**		**Predicted annual maximum Rainfall (Y_i) calculated by GPA**		**Predicted annual maximum Rainfall (Y_i) calculated by GLO**		**Predicted annual maximum Rainfall (Y_i) based on UEV**
		MOM	PWM-L	MOM	PWM-L	MOM	PWM-L	Objective
1	123.70	168.33	140.81	205.68	154.20	140.01	85.26	177.13
2	140.70	182.75	162.56	207.67	160.71	164.05	116.52	205.98
3	153.10	192.54	177.01	209.71	167.23	179.07	135.33	224.34
4	162.80	200.30	188.28	211.81	173.75	190.30	149.04	238.20
5	186.20	206.90	197.73	213.96	180.27	199.43	159.98	249.52
6	194.20	212.74	206.00	216.18	186.79	207.22	169.18	259.21
7	202.60	218.06	213.45	218.46	193.32	214.08	177.19	267.74
8	225.00	223.00	220.29	220.80	199.84	220.29	184.33	275.43
9	232.30	227.64	226.66	223.22	206.37	225.99	190.83	282.46
10	233.20	232.07	232.67	225.71	212.90	231.30	196.83	288.97
11	233.80	236.31	238.40	228.29	219.43	236.31	202.44	295.07
12	236.20	240.43	243.89	230.95	225.97	241.08	207.74	300.82
13	240.80	244.43	249.20	233.71	232.51	245.66	212.78	306.28
14	244.80	248.36	254.36	236.56	239.05	250.08	217.62	311.50
15	251.10	252.22	259.39	239.52	245.59	254.39	222.30	316.52
16	254.80	256.05	264.33	242.60	252.13	258.61	226.85	321.37
17	261.60	259.85	269.20	245.80	258.68	262.77	231.30	326.06
18	263.50	263.64	274.02	249.14	265.23	266.88	235.68	330.64
19	266.80	267.44	278.80	252.62	271.79	270.97	240.00	335.10
20	268.30	271.26	283.57	256.26	278.35	275.06	244.30	339.48
21	272.50	275.11	288.34	260.09	284.91	279.17	248.58	343.78
22	280.30	279.01	293.13	264.10	291.47	283.31	252.88	348.03
23	282.60	282.97	297.95	268.34	298.04	287.51	257.21	352.23
24	305.70	287.02	302.82	272.81	304.61	291.80	261.59	356.40
25	305.90	291.16	307.76	277.55	311.19	296.19	266.05	360.55
26	312.20	295.43	312.80	282.60	317.77	300.70	270.61	364.69
27	315.70	299.83	317.95	287.99	324.36	305.39	275.31	368.84
28	326.00	304.41	323.24	293.78	330.95	310.27	280.17	373.01

29	334.30	309.19	328.69	300.03	337.54	315.40	285.23	377.22
30	342.60	314.22	334.36	306.82	344.15	320.83	290.56	381.49
31	349.70	319.54	340.28	314.25	350.76	326.63	296.20	385.83
32	352.50	325.22	346.51	322.46	357.37	332.90	302.24	390.27
33	363.30	331.35	353.12	331.63	364.00	339.76	308.79	394.84
34	365.10	338.04	360.22	342.03	370.63	347.38	315.99	399.58
35	379.20	345.46	367.94	354.01	377.27	356.03	324.07	404.54
36	385.50	353.88	376.49	368.17	383.93	366.11	333.36	409.79
37	391.30	363.70	386.20	385.47	390.60	378.29	344.43	415.46
38	404.30	375.68	397.66	407.75	397.29	393.88	358.32	421.72
39	411.70	391.44	412.07	439.07	404.00	415.84	377.41	428.93
40	420.50	415.83	432.87	492.42	410.75	453.88	409.24	438.04

4.5.2 CASE 2: EAST UTTAR PRADESH SUBDIVISION

Table 4.36: observed and predicted annual maximum rainfall calculated by GEV, GPA, GLO, and UEV using MOM, PWM-L, SO methods

Rank	**Observed annual maximum Rainfall (X_i)**	**Predicted annual maximum Rainfall (Y_i) calculated by GEV**		**Predicted annual maximum Rainfall (Y_i) calculated by GPA**		**Predicted annual maximum Rainfall (Y_i) calculated by GLO**		**Predicted annual maximum Rainfall (Y_i) based on UEV**
		MOM	PWM-L	MOM	PWM-L	MOM	PWM-L	Objective
1	177.00	207.56	184.11	201.24	200.30	251.58	172.70	231.26
2	179.00	214.88	196.73	205.43	204.67	255.15	193.53	237.32
3	219.50	220.21	205.33	209.66	209.07	257.83	206.69	241.65
4	220.80	224.67	212.17	213.93	213.51	260.09	216.60	245.22
5	224.50	228.62	218.00	218.25	217.99	262.10	224.70	248.34
6	226.70	232.26	223.17	222.61	222.52	263.96	231.64	251.18
7	232.10	235.70	227.89	227.02	227.08	265.70	237.79	253.84
8	234.20	238.99	232.28	231.48	231.69	267.36	243.36	256.36
9	240.70	242.18	236.42	235.99	236.34	268.96	248.49	258.78
10	240.90	245.31	240.36	240.55	241.04	270.53	253.29	261.14
11	243.30	248.40	244.16	245.17	245.79	272.07	257.82	263.45
12	247.00	251.48	247.84	249.85	250.59	273.60	262.15	265.73
13	261.50	254.57	251.43	254.59	255.45	275.13	266.32	268.00
14	265.20	257.67	254.95	259.40	260.36	276.66	270.36	270.26
15	269.10	260.81	258.42	264.27	265.33	278.20	274.29	272.53
16	289.50	264.00	261.87	269.22	270.37	279.76	278.16	274.82
17	290.30	267.25	265.29	274.24	275.47	281.35	281.97	277.14
18	290.40	270.59	268.72	279.34	280.64	282.98	285.75	279.50
19	291.30	274.02	272.15	284.53	285.88	284.65	289.52	281.90
20	294.00	277.56	275.61	289.80	291.21	286.38	293.30	284.37
21	300.10	281.24	279.10	295.18	296.61	288.16	297.09	286.91
22	301.00	285.08	282.65	300.65	302.10	290.02	300.93	289.54
23	309.70	289.09	286.25	306.24	307.69	291.97	304.84	292.26
24	316.60	293.31	289.94	311.95	313.39	294.02	308.82	295.10

25	320.6	297.77	293.73	317.78	319.19	296.19	312.91	298.08
26	328.90	302.51	297.63	323.76	325.11	298.50	317.13	301.21
27	329.70	307.58	301.67	329.88	331.16	300.98	321.51	304.53
28	333.80	313.04	305.87	336.18	337.35	303.66	326.09	308.06
29	337.90	318.96	310.27	342.67	343.70	306.57	330.91	311.85
30	338.10	325.43	314.91	349.36	350.22	309.78	336.03	315.95
31	341.60	332.59	319.83	356.29	356.94	313.34	341.51	320.42
32	343.60	340.59	325.10	363.49	363.89	317.37	347.44	325.36
33	351.80	349.67	330.80	371.00	371.09	321.98	353.95	330.88
34	364.20	360.17	337.04	378.87	378.59	327.37	361.21	337.16
35	370.70	372.57	344.00	387.19	386.44	333.83	369.48	344.45
36	375.40	387.70	351.91	396.05	394.74	341.86	379.15	353.14
37	395.90	406.95	361.19	405.62	403.60	352.33	390.89	363.94
38	436.00	433.12	372.58	416.15	413.22	367.00	406.00	378.16
39	450.65	472.88	387.68	428.14	423.97	390.33	427.43	398.88
40	480.30	550.03	411.36	442.80	436.73	439.12	464.98	436.53

4.6 GRAPHICAL METHOD

Using the graphical method, calculate the coefficient of determination (R^2) for all four extreme value distribution shown in fig 4.9 to 4.15 and 4.16 to 4.22 respectively for West Uttar Pradesh subdivision and East Uttar Pradesh subdivision.

4.6.1 CASE 1- WEST UTTAR PRADESH SUBDIVISION

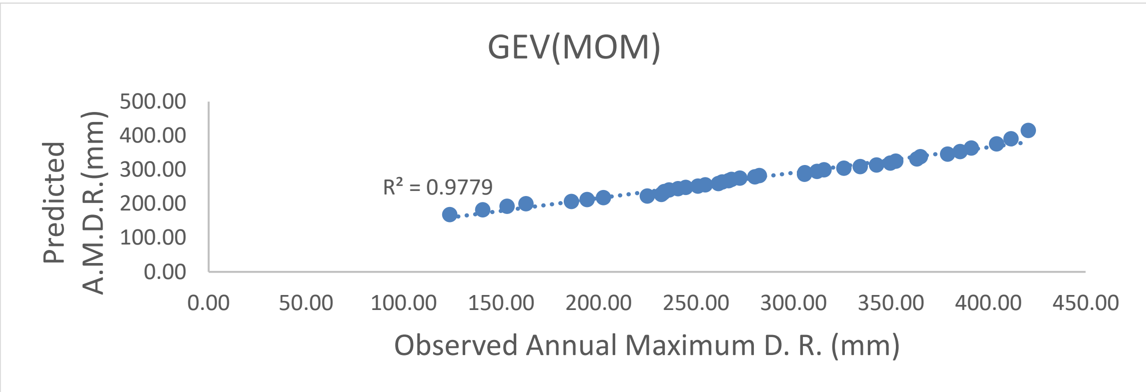

Fig 4.9: observed and predicted rainfall calculated by GEV (MOM) distribution

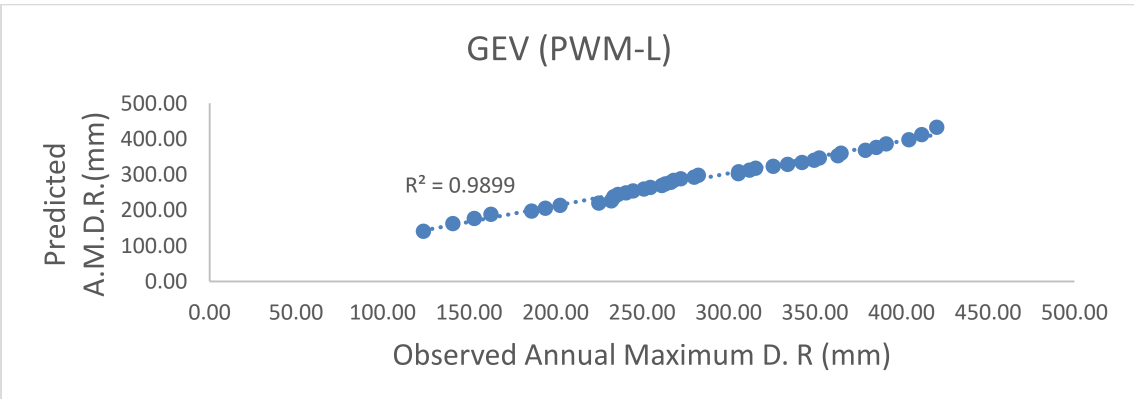

Fig 4.10: observed and predicted rainfall calculated by GEV (PWM-L) distribution

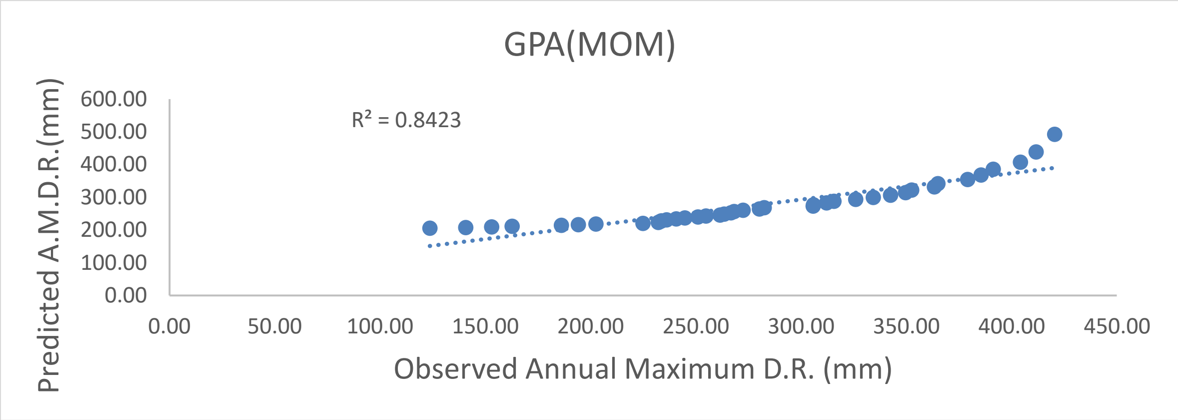

Fig 4.11: observed and predicted rainfall calculated by GPA (MOM) distribution

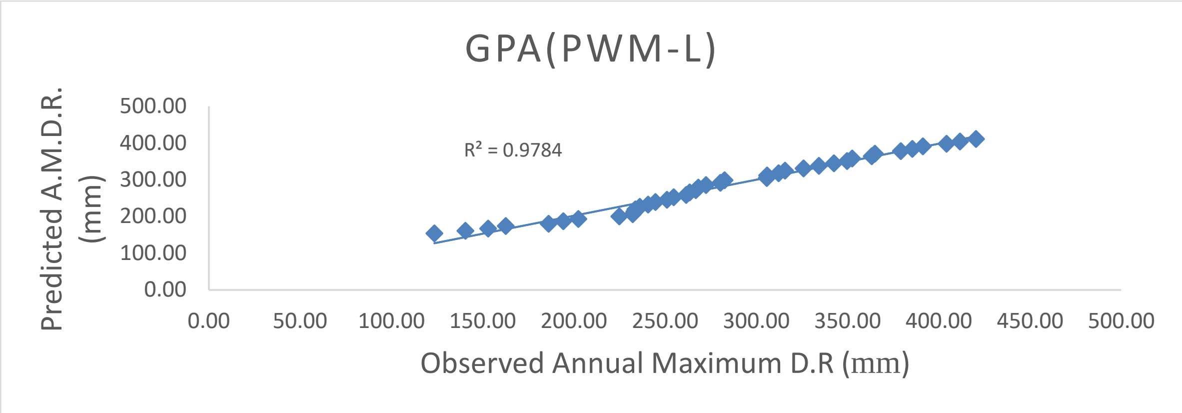

Fig 4.12: observed and predicted rainfall calculated by GPA (PWM-L) distribution

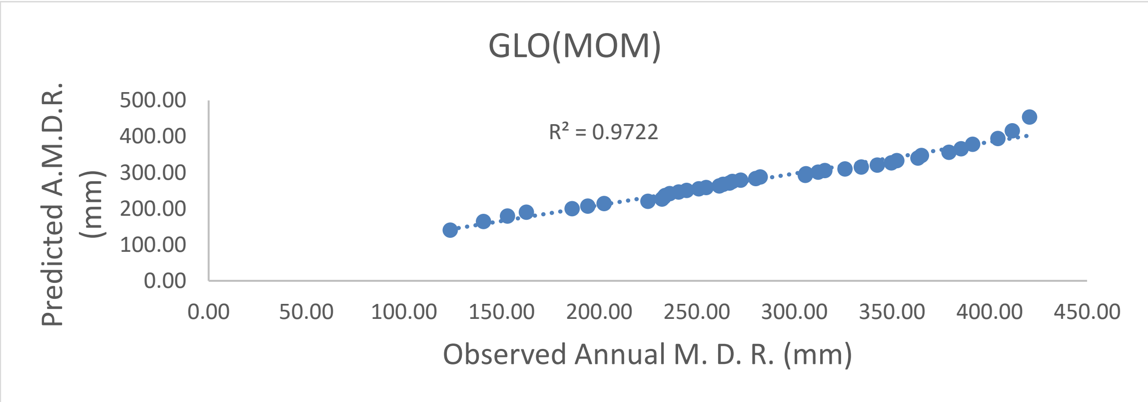

Fig 4.13: observed and predicted rainfall calculated by GLO (MOM) distribution

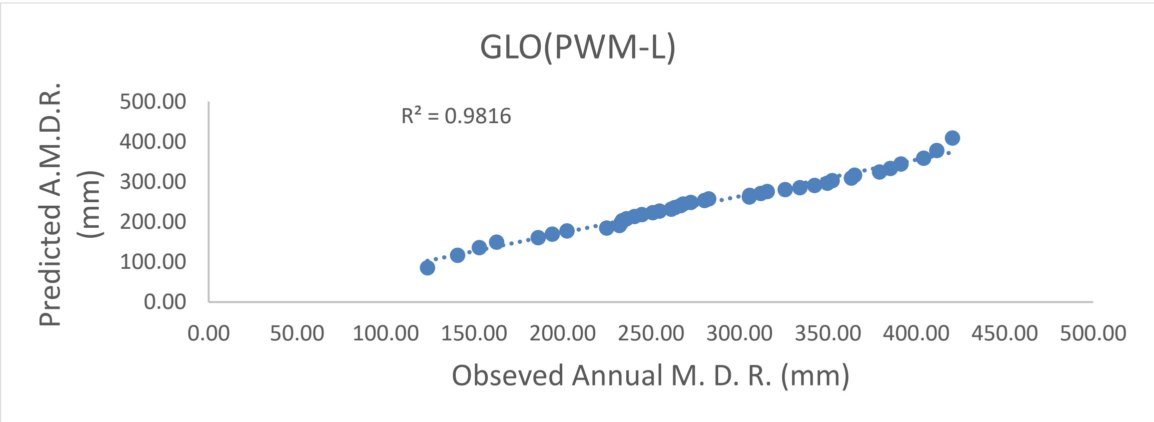

Fig 4.14: observed and predicted rainfall calculated by GLO (PWM-L) distribution

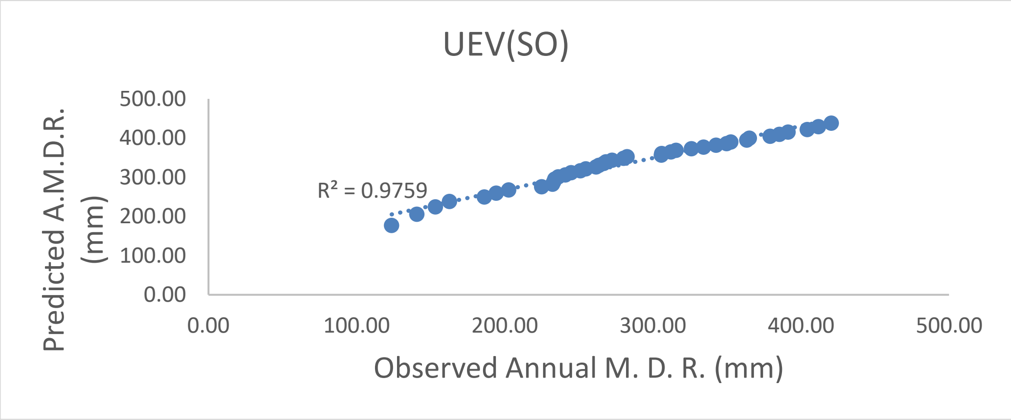

Fig 4.15: observed and predicted rainfall calculated by UEV (SO) distribution

- Fig 4.9 to 4.15 represents the graph between observed and predicted rainfall calculated by four distribution, and also show the value of the coefficient of determination for all Distribution. The computed coefficient of determination (R^2) value was observed to be 0.9899 for GEV with PWM-L, which is highest among all other distributions. Based on the computed R^2, it concluded that generalized Extreme Value distribution had the best fit of the annual maximum rainfall data for West Uttar Pradesh subdivision.

4.6.2 CASE 2 – EAST UTTAR PRADESH SUBDIVISION

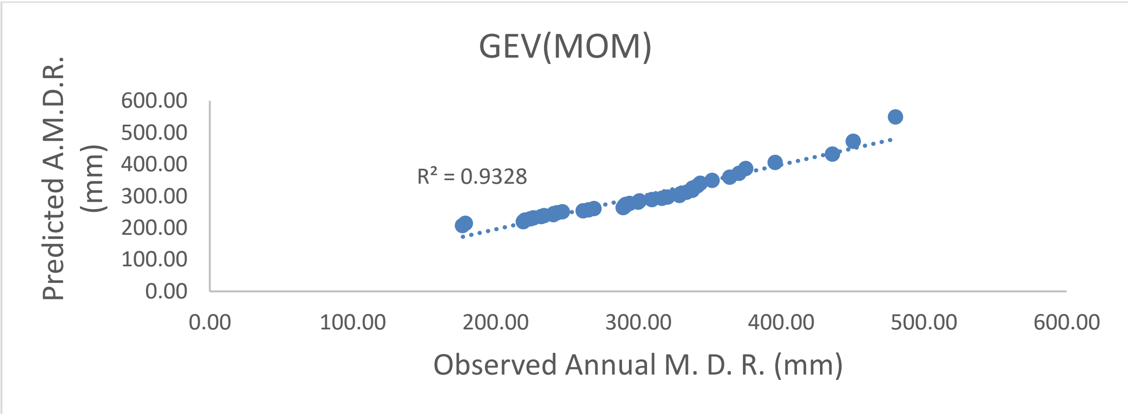

Fig 4.16: observed and predicted rainfall calculated by GEV (MOM) distribution

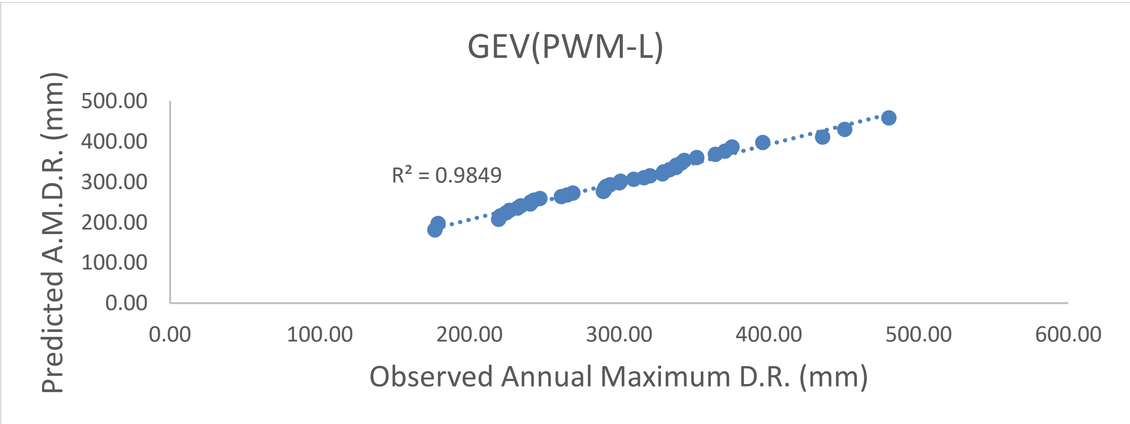

Fig 4.17: observed and predicted rainfall calculated by GEV (PWM-L) distribution

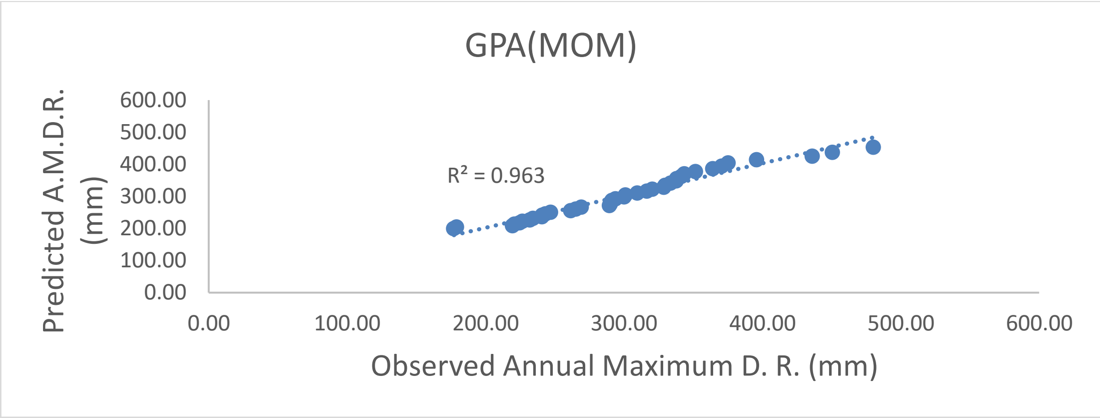

Fig 4.18: observed and predicted rainfall calculated by GPA (MOM) distribution

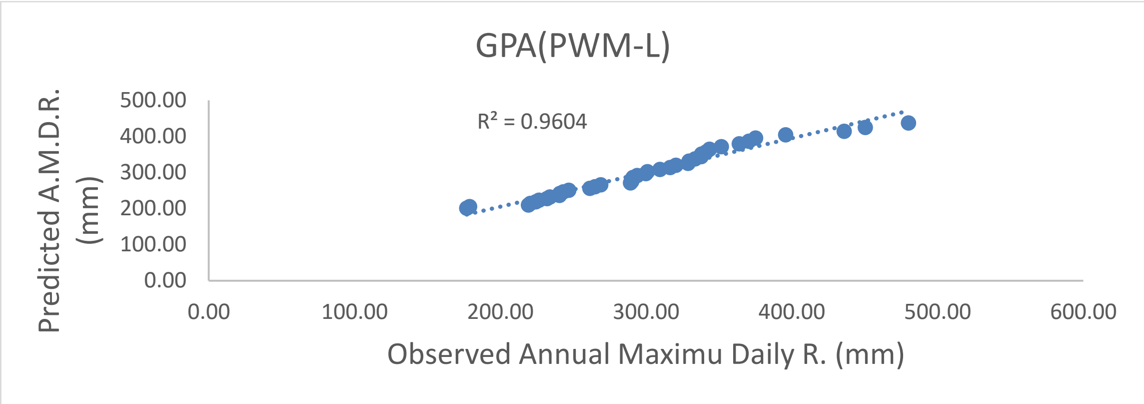

Fig 4.19: observed and predicted rainfall calculated by GPA (PWM-L) distribution

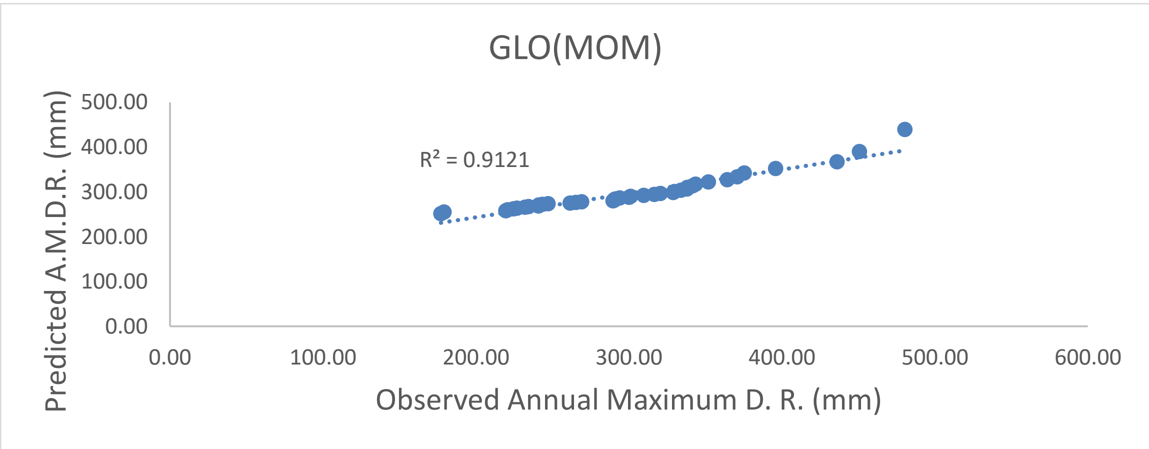

Fig 4.20: observed and predicted rainfall calculated by GLO (MOM) distribution

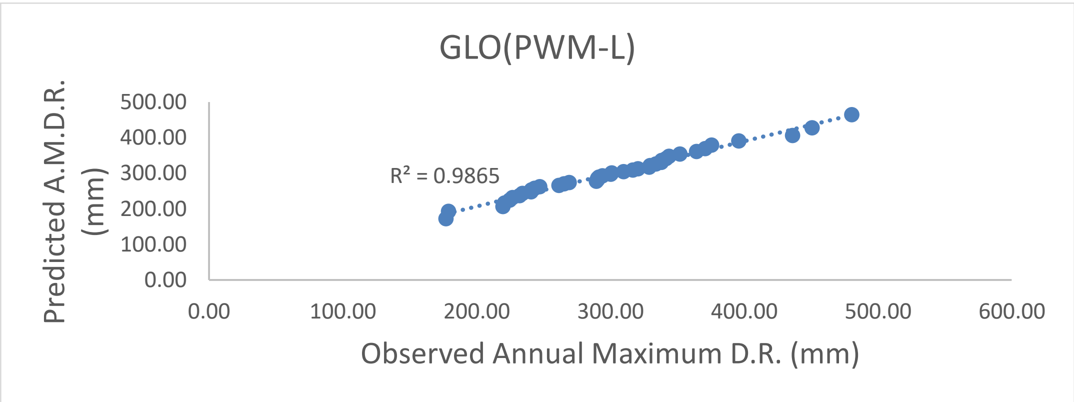

Fig 4.21: observed and predicted rainfall calculated by GLO (PWM-L) distribution

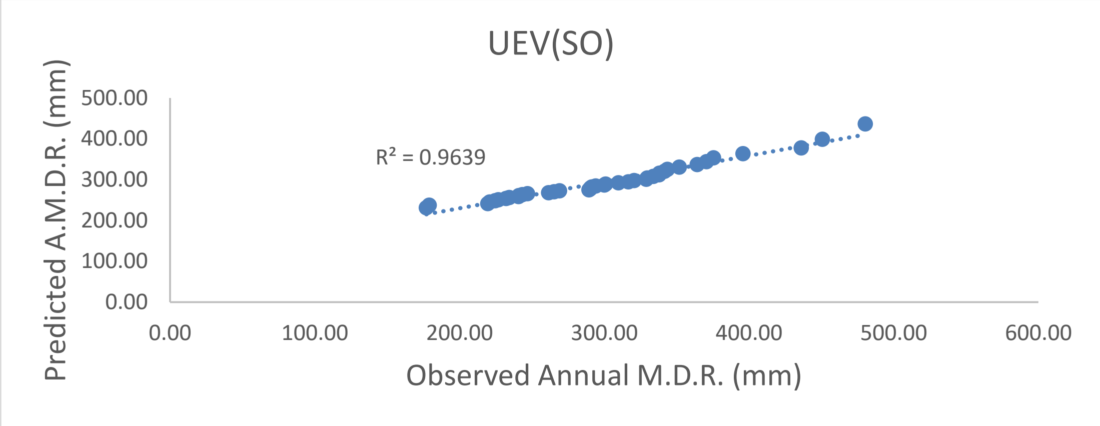

Fig 4.22: observed and predicted rainfall calculated by UEV (SO) distribution

- Fig 4.16 to 4.22 represents the graph between observed and predicted rainfall calculated by four distribution, and also the value of the coefficient of determination for all Distribution. The computed R^2 value was observed to be 0.9865 for GLO with PWM-L, which is highest among all other distributions. Based on the computed R^2, it concluded that generalized Logistics distribution had a fit of the annual maximum rainfall data for East Uttar Pradesh subdivision.

4.7 GOODNESS-OF-FIT TEST

In this study, the Five Goodness of Fit test was used. Relative root means square error (RRMSE), Root mean square error (RMSE), and Maximum absolute error (MAE), and Mean absolute deviation index (MADI), Probability plot correlation coefficient (PPCC). Estimate the value of five Goodness of fit test for GEV (MOM), GEV (PWM-L), GLO (MOM), GLO (PWM-L), GPA (MOM), GPA (PWM-L), UEV (SO) distribution results shown in table 4.37 and 4.39, respectively West Uttar Pradesh subdivision and East Uttar Pradesh subdivision.

Compare the results of the Goodness of Fit test based on the Distribution with the low RRMSE, low RMSE, low MAE, low MADI, and high PPCC rated 4, followed by negative rating 3, 2, while negative score 1. Comparisons, overall rating, and position results in Table 4.38 and 4.40, respectively West Uttar Pradesh subdivision and East Uttar Pradesh subdivision.

4.7.1 CASE 1 WEST UTTAR PRADESH SUBDIVISION

Table 4.37: Distribution of the Goodness-of-fit test for GEV, GLO, GPA, and UEV

GOF	METHOD	RMSE	RRMSE	MADI	MAE	PPCC
GEV	MOM	22.29657205	0.110298898	0.067598972	0.446345745	0.988876891
	PWM-L	11.43894593	0.059776043	0.040692985	0.254757889	0.994952476
GPA	MOM	32.03661269	0.172551235	0.105187567	0.82000000	0.917771761
	PWM-L	11.72619992	0.063909576	0.039582494	0.30500027	0.999128524
GLO	MOM	15.38240362	0.066498141	0.047726517	0.53375300	0.985985093
	PWM-L	38.68090778	0.140873374	0.129147027	0.551299286	0.990767194
UEV	OBJECT	57.22240906	0.255601071	0.21627072	0.753988392	0.98787978

Table 4.38: Total rank of different probability distribution models

GOF Test Criteria	Distribution Score/ Rank							
	GEV		GPA		GLO		UEV	
	MOM	**PWM-L**	**MOM**	**PWM-L**	**MOM**	**PWM-L**	**S.O(MOM)**	**S.O(PWM-L)**
RMSE	3	4	2	3	4	2	1	1
RRMSE	3	4	2	3	4	2	1	1
MADI	3	3	2	4	4	2	1	1
MAE	4	4	1	3	3	2	2	1
PPCC	4	3	1	4	2	2	3	1
Total Score/Rank	17	18(1st)	8	17	17	10	8	5

- Table 4.38 shows the results of distribution scoring for annual maximum rainfall in case 1. Based on the results, the highest score is 18 for GEV (PWM-L), so for West Uttar Pradesh subdivision, the best distribution model is generalized extreme value distribution with probability-weighted with L- moment parameters estimation method.

4.7.2 CASE 2 EAST UTTAR PRADESH SUBDIVISION

Table 4.39: Distribution of the Goodness-of-fit test for GEV, GLO, GPA, and UEV

	METHOD	**RMSE**	**RRMSE**	**MADI**	**MAE**	**PPCC**
GEV	MOM	19.92620541	0.067432862	0.04805377	0.69732618	0.965836384
	PWM-L	27.22768308	0.075536632	0.063341418	0.689385029	0.992420978
GPA	MOM	13.80886974	0.048172119	0.034094513	0.370000000	0.981323916
	PWM-L	14.29611715	0.047870618	0.033574011	0.435736177	0.980021635
GLO	MOM	35.89834044	0.142205643	0.107517427	0.761486704	0.955055813
	PWM-L	10.02769374	0.033551327	0.025312232	0.300037933	0.993204104
UEV	OBJECT	27.86777408	0.105635394	0.081861599	0.583233511	0.981792586

Table 4.40: the total rank of different probability distribution models

GOF Test Criteria	Distribution rank /score							
	GEV		GPA		GLO		UEV	
	MOM	**PWM-L**	**MOM**	**PWM-L**	**MOM**	**PWM-L**	**S.O(MOM)**	**S.O(PWM-L)**
RMSE	3	2	4	3	1	4	2	1
RRMSE	3	2	4	3	1	4	2	1
MADI	3	2	4	3	1	4	2	1
MAE	2	1	4	3	1	4	3	2
PPCC	2	3	3	1	1	4	4	2
Total Score/Rank	13	10	19	13	5	20(1st)	13	7

- Table 4.40 shows the results of distribution scoring for annual maximum rainfall in case 2. Based on the results, the highest score is 20 for GLO (PWM-L), so for East Uttar Pradesh subdivision, the best distribution model is Generalized Logistics Distribution, and the parameters estimation method is Probability Weighted Moment with L- Moment.

4.8 RETURN PERIODS

For West Uttar Pradesh subdivision, the return period 2,5,10,20,50 is forecast based on Generalized Extreme Value Distribution (Probability Weighted Moment with L- Moment) and for East Uttar Pradesh subdivision the return level 2,5,10,20,50 is forecast based on Generalized Logistics Distribution (Probability Weighted Moment with L-Moment), Both shown in table 4.41.

Table 4.41: Return periods for different return intervals for both case

Return periods (T)	**2**	**5**	**10**	**20**	**50**
Case 1	285.96	351.77	385.17	411.25	438.07
Case 2	295.19	352.60	389.59	426.11	476.03

CHAPTER 5.0

CONCLUSION

5.1 CONCLUSION

In this study we have discussed the results of annual maximum daily rainfall series data of meteorological subdivision of Uttar Pradesh, i.e., West Uttar Pradesh subdivision and East Uttar Pradesh subdivision taken for study purpose initially screened for outliers, stationarity, normality, homogeneity test, R^2, and Goodness of fit. Flood frequency formulae using different distributions have been calculated.

1. Obtained statistical parameters for annual maximum daily rainfall dataset of meteorological subdivision of Uttar Pradesh are:

   a.) WEST UTTAR PRADESH SUBDIVISION

   It observed that the parameter for the statistical point of view for series of West Uttar Pradesh subdivisions is a coefficient of variance 0.27551, mean 281.91, standard deviation 77.669, kurtosis coefficient -0.68756, and skewness coefficient -0.06483.

   b.) EAST UTTAR PRADESH SUBDIVISION

   It observed that the parameter for the statistical point of view for series of East Uttar Pradesh subdivisions is a coefficient of variance 0.23148, mean 301.589, standard deviation 69.814, kurtosis coefficient 0.18398, and skewness coefficient 0.505543.

2. Preliminary analysis of Hydrological data, i.e., Rainfall at the meteorological subdivision of Uttar Pradesh (West Uttar Pradesh subdivision and East Uttar Pradesh subdivision) check for outliers, stationarity, homogeneity, normality test in data sets.

   a.) Check for the outlier, based on the results it found that there are no outliers in annual maximum daily rainfall dataset at West Uttar Pradesh subdivision and East Uttar Pradesh subdivision.

   b.) Check for the stationarity test, based on the results it found that both series are stationary.

   c.) Check for normality test, based on the results it found that the annual maximum daily rainfall dataset at West Uttar Pradesh subdivision and East Uttar Pradesh subdivision followed Normal Distribution.

   d.) Check for homogeneity test, based on the results it found that both series are homogenous.

3. Based on the results for case 1, it found that the R^2 value of GEV (PWM-L) is 0.9899, which is the highest value among other Distribution. So Generalized Extreme Value Distribution (PWM-L) distribution well-suited distribution for West Uttar Pradesh subdivision. And for case 2, it observed that the R^2 value of GLO (PWM-L) is 0.9865, which is the highest value among other Distribution. So Generalized Logistics Distribution (PWM-L) well-suited distribution for East Uttar Pradesh subdivision.
4. Comparisons between the various methods of distribution by Goodness of fit test for case 1, found that GEV (PWM-L) obtained the highest ranking among others, i.e., 18. Therefore Generalized Extreme Value Distribution with (PWM-L) - distribution suitable for the partition of West Uttar Pradesh. And in the 2nd case, the GLO (PWM-L) gets the highest rank among others, e.g., 20. Thus the generalized logistics distribution with (PWM-L) the well-distributed distribution of East Uttar Pradesh is well divided.
5. In this study, a single distribution is not as effective as the proper distribution of both cases. Overall, the PWM-L rating system is ready to identify the appropriate distribution.
6. Has used the most appropriate distribution in each area, namely, GEV with (PWM-L) and GLO with (PWM-L), respectively, in West Uttar Pradesh and East Uttar Pradesh divisions and predicted heavy rainfall at a huge event, approximately 2, 5, 10, 20, and 50 return times.

REFERENCES

1. Singh K. Sushil (2018) "Unified Extreme Value Distribution", [DOI: 10.1061/ (ASCE) IR.1943-4774.0001232. © 2017 American Society of Civil Engineers].
2. Rudolph and Ebierin Akpoebidimiyen Otuaro (2019), "Simple to Use Microsoft Excel Template for Estimating the Parameters of Some Selected Probability Distribution Model by Method of L-Moment," [ISSN 2224-5790 (Paper) ISSN 2225-0514 (Online) Civil and Environmental Research].
3. Patel K. Susheel and subash (2018) "comparative study of trends in rainfall over metrological subdivisions of Uttar Pradesh," [ISSN: 2348-8808(Print), 2348-8867(Online), Journal of AgriSearch].
4. Rainfall data Excel sheet, 4-Subdivision IMD 2017 (rainfall datasheet report).
5. Bjorn H. Auestad, Andreas H., Hans A. Karlsen (2012), " Modeling and analysis of daily rainfall data" [DOI 10.1007/978-94-007-4153-9_40, © Springer Science Business Media Dordrecht 2012]
6. XLSTAT software, "user manual – statistical guide book."
7. Ojha C.S.P. "Book of Engineering Hydrology," [fifth Edition, chapters 6 and 7].
8. Rao R.A. and Hamed K.H., "Book of flood frequency analysis," [third edition, chapter 9].
9. T.O. Olatayo, and A. I. Taiwo (2014), "Statistical modeling and prediction of rainfall time series data" [Online ISSN: 0975-4172 & Print ISSN: 0975-4350]
10. EasyFit 5.6 Software, "user manual guide book."
11. Oliver U., K. Joseph (2018), "Modelling Extreme Maximum Rainfall Using Generalized Extreme Value Distribution: Case Study Kigali City," [International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064].
12. Javari M. (2016), "Trend and homogeneity analysis of precipitation in Iran," [Journal of climate].
13. Kang Ming and Yusof F. (2012), "Homogeneity tests on daily rainfall series in Peninsular Malaysia," [Int. J. Contemp. Math. Sciences, Vol. 7, 2012, no. 1, 9 – 22].
14. Tefaruk Haktanir and H. Citakoglu (2014), "Trend, independence, stationarity, and homogeneity tests on maximum rainfall series of standard durations recorded in Turkey" [doi:10.1061/(ASCE)HE.1943-5584.0000973, Journal of Hydrologic Engineering]
15. M. Masereka, George M. Ochieng, and J.S (2018), "Statistical analysis of annual maximum daily rainfall for Nelspruit and its environs," [Journal of Disaster Risk Studies ISSN: (Online) 2072-845X, (Print) 1996-1421].

16. Wenny Susanti and Arisman Adnan (2018), “the analysis of extreme rainfall in Pekanbaru City using three-parameter GEV, and GPA distribution” [Applied Mathematical Sciences, Vol. 12, 2018, no. 2, 69 – 80].
17. Mahmood Hassan, Omar Hayat, and Zahra Noreen (2019), "Selecting the best probability distribution for at-site flood frequency analysis; a study of Torne River," [S.N. Applied Sciences (2019) 1:1629 | https://doi.org/10.1007/s42452-019-1584-z].
18. Wikipedia, "the climate of Uttar Pradesh."
19. J. R. M. Hosking and J. R. Wallis, Regional Frequency Analysis: An Approach Based on L-Moments, Cambridge University Press, London, 1997. https://doi.org/10.1017/cbo9780511529443