“FREQUENCY ANALYSIS OF RAINFALL DATA OF SHIMLA AND DHARAMSHALA REGION”

A Thesis

Submitted in partial fulfillment of the requirements for the award of the degree of

MASTER OF TECHNOLOGY

IN

CIVIL ENGINEERING

With specialization in

ENVIRONMENTAL ENGINEERING

Under the supervision of

Dr. Ashish Kumar

&

Dr. Rajiv Ganguly

By

Nitish Kumar Sharma

(142759)

to

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY

WAKNAGHAT SOLAN – 173234

HIMACHAL PRADESH INDIA

June, 2016

CERTIFICATE

This is to certify that the work which is being presented in the project title **"FREQUENCY ANALYSIS OF RAINFALL DATA OF SHIMLA AND DHARAMSHALA REGION"** in partial fulfillment of the requirements for the award of the degree of Master of technology and submitted in Civil Engineering Department, Jaypee University of Information Technology, Waknaghat is an authentic record of work carried out by **Nitish Kumar Sharma** during a period from July 2015 to May 2016 under the supervision of **Dr. Ashish Kumar** and **Dr. Rajiv Ganguly** Associate Professor, Civil Engineering Department, Jaypee University of Information Technology, Waknaghat.

The above statement made is correct to the best of my knowledge.

Date: 26 May, 2016

Dr. Ashish Kumar	**Dr. Rajiv Ganguly**	……………………
Associate Professor	Associate Professor	**External Examiner**
Department of Civil Engineering	Department of Civil Engineering	
JUIT Waknaghat	JUIT Waknaghat	

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the dissertation entitled **"FREQUENCY ANALYSIS OF RAINFALL DATA OF SHIMLA AND DHARAMSHALA REGION"** in partial fulfillment of the requirements for the award of the degree of Master of technology and submitted in Civil Engineering Department, Jaypee University of Information Technology, Waknaghat is an authentic record of work carried out by **Nitish Kumar Sharma** during a period from July 2015 to May 2016 under the supervision of **Dr. Ashish Kumar** and **Dr. Rajiv Ganguly** Associate Professor, Civil Engineering Department, Jaypee University of Information Technology, Waknaghat.

I have not submitted the matter embodied in this dissertation for the award of any other degree.

Date: 26 May, 2016 **NITISH KUMAR SHARMA**

Place: Waknaghat, Solan, H.P. **142759**

ACKNOWLEDGEMENT

I take upon this opportunity endowed upon me by grace of the almighty, to thank all those who have been part of this endeavor.

I would like to thank my supervisors **'Dr. Ashish Kumar'** and **'Dr. Rajiv Ganguly'** for giving me the right direction to follow and proper guidance regarding the topic. Without their active involvement and the right guidance this would not have been possible.

I sincerely thank our Head of Department **'Prof. (Dr.) Ashok Kumar Gupta'** for giving me the chance as well as the support for all the time being.

Last but not the least, I heartily appreciate all those people who have helped me directly or indirectly in making this task a success. In this context, I would like to thank all the other staff members, both teaching and non-teaching, which have extended their timely help and eased my task.

NITISH KUMAR SHARMA

142759

ABSTRACT

The present study aims to evaluate the rainfall magnitude for different return periods and also to ascertain the type of probability distribution that best fits the rainfall data of Shimla & Dharamshala (H.P.), India. Probability and frequency analysis of rainfall data enables us to determine the expected rainfall at various chances. Such information can also be used to prevent floods and droughts, and applied to planning and designing of water resources related to engineering such as reservoir design, flood control work, soil and water conservation planning. The present study involves the 20 years of annual rainfall data are useful for the prediction of annual one day maximum rainfall and two to seven days consecutive days maximum rainfall corresponding to return period varying from 2 to 20 years are to be used for the economic planning, by design engineers and hydrologists, design of small and medium hydrologic structures and determination of drainage coefficient for agricultural fields. Various probability distributions (viz: Normal, Log Normal and Gamma distribution) and transformations are applied to estimate one day and two to seven consecutive days annual maximum rainfall of various return periods at Shimla and Dharamshala (H.P.) region. A maximum of 121mm in 1 day, 178.15 mm in 2 days, 203.6 mm in 3 days, 219.4 mm in 4 days, 230.23 mm in 5 days, 240.6 mm in 6 days and 257.5mm in 7 days is expected to occur at Shimla, Himachal Pradesh every 2 years. For a recurrence interval of 20 years, the maximum rainfall expected in 1 day, 2, 3, 4, 5, 6 and 7 days is 160.5 mm, 239.11 mm, 265.36 mm, 285.9 mm, 297.42 mm, 309.5 mm and 330.38 mm respectively. A maximum of 191.1 mm in 1 day, 262.7 mm in 2 days, 314.27 mm in 3 days, 370.24 mm in 4 days, 419.7 mm in 5 days, 477.4 mm in 6 days and 522.32 mm in 7 days is expected to occur at Dharamshala, Himachal Pradesh every 2 years. For a recurrence interval of 20 years, the maximum rainfall expected in 1 day, 2, 3, 4, 5, 6 and 7 days is 277.7 mm, 373.6 mm, 445.11 mm, 518.62 mm, 589 mm, 680.3 mm and 753.79 mm respectively. Various probability distributions and transformations are applied to estimate one day and two to seven consecutive days annual maximum rainfall of various return periods. Three commonly used probability distributions (viz: Normal, Log Normal and Gamma distribution) were tested by comparing the Chi-square, Kolmogorov-Smirnov and Anderson Darling values.

Keywords: rainfall, consecutive, normal, lognormal, gamma, distribution, probability, frequency, analysis.

LIST OF SYMBOLS

n_i	:	number of observation in interval i
n	:	total number of observation
m	:	number of interval
p	:	number of parameters used in fitting the proposed distribution
C_V	:	coefficient of variation
K_T	:	frequency factor
μ	:	mean of the sample
σ	:	standard deviation of the sample
y	:	ln x
x	:	a variable
μ_y	:	mean of y
σ_y	:	standard deviation of y
z	:	standard normal variable
v	:	degree of freedom
α	:	significance level
$T(\beta)$	:	gamma function
X_T	:	magnitude of the event having a return period t
$\overline{X}$	:	mean value of X

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 General

The primary source of water for agricultural production for most of the world is rainfall. Three main characteristics of rainfall are its amount, frequency and intensity, the values of which vary from place to place, day to day, month to month and also year to year. In general, three types of probability distributions (Normal, Log Normal and Gamma distribution) are commonly used to determine the best fit probability distribution using the comparison of Chi-square, Anderson-Darling and Kolmogorov–Smirnov values. Several applications in water resources engineering require appropriate estimation of rainfall depth and its return period from available historic data. Estimation of flood in watersheds, water balance studies, water management studies, rainwater harvesting, detention and retention pond design, evapotranspiration estimation, irrigation planning, etc. are some of the examples where rainfall provides a vital input to design and modeling. Planning and development of water resources at the local or regional level require comprehensive and reliable information of hydrological data of the area under investigation. Proper database is needed to assess the water availability of a region, the absence of which can lead to erroneous planning and design. Long period data can provide reliable water resource assessment. The procedure for estimating the frequency of occurrence of a hydrological event such as flood is known as frequency analysis. Though the nature of most hydrological events (such as rainfall) is erratic and varies with time and space, it is commonly possible to predict return periods using various probability distributions. Flood frequency analysis is a statistical tool to help engineers, hydrologists, and watershed managers to deal with this uncertainty. Flood frequency is utilized to determine how often a storm of a given magnitude would occur. It is an important tool for the building and design of the safest possible structures (e.g. dams, bridges, culverts, drainage systems etc.) because the design of such structures demands knowledge of the likely floods which the structure would have to withstand during its estimated economic useful life. Rainfall at a particular place is also known to be influenced by the results of its local/regional atmospheric and geomorphological environments. An important aspect in hydrology is to interpret the future probabilities of occurrence from past records of hydrologic events. Chi square, Anderson-Darling and Kolmogorov–Smirnov goodness of fit tests are used to judge the applicability of the distributions for modeling of the recorded rainfall data. The standard procedure for estimating the frequency of occurrence of hydrological event is frequency analysis. The objective of frequency analysis of hydrologic data is to relate the magnitude of extreme events to their frequency of occurrence using probability distributions. The extreme rainfall event involves the selection of a sequence of the maximum observations from the respective data series. Information of the amount, intensity and distribution of monthly or annual rainfall for the most important places in the world is generally available. Both flood and drought occurrence probabilities can be interpreted by using various statistical methods. Frequency analysis is used in hydrology as one such tool. In frequency analysis, the magnitude of extreme events is related to their frequency of occurrence through the use of probability distributions. More the magnitude of an extreme event, lesser is its frequency of occurrence. Frequency analysis is a statistical tool to help design engineers, hydrologists, and researchers to deal with this uncertainty. For such purposes, annual maximum flood peaks are used as input data series. In absence of a long record of discharge gauge data for any watershed, rainfall data series is used to interpret extreme storm event corresponding to certain return period and then find the discharge through various possible rainfall-runoff models. Long-term records of daily rainfall have been compiled for years; norms and standard deviations have been worked out; floods and droughts have been defined and climatic zones of potential evapotranspiration less precipitation have been mapped from rainfall patterns and crop studies. Efforts are being made to predict future trends in order to refine planning and the study of rainfall analyses will help us for future use like for various designs, agricultural purposes, useful for research scholars. The salient points in the practical application in the field of agriculture of data on effective rainfall as follows:

a) designing irrigation projects on a sound economic basis;
b) fixing cropping patterns and working out the irrigation requirements of crops;
c) operating irrigation projects efficiently from year to year;
d) preparing schedules of other farm operations in irrigated agriculture;
e) planning cropping patterns in unirrigated or rain fed areas;
f) designing drainage and land reclamation projects;
g) planning soil and water conservation programmes;
h) interpreting field experiments accurately;
i) classifying regions climatologically for agriculture.

1.2 Literature review

A number of research papers are available on the topic of frequency analysis of rainfall data. Some of the most important research papers in the context of rainfall have been reviewed here in the present study. Benson et al. (1968) carried out an annual one day maximum rainfall and two to five days consecutive days maximum rainfall corresponding to return period varying from 2 to 100 years. A maximum of 154.31mm in 1 day, 250.88mm in 2 days, 270.15mm in 3 days, 284.18mm in 4 days and 295.54mm in 5 days is expected to occur at Udaipur, Rajasthan every 2 years. For a recurrence interval of 100 years, the maximum rainfall expected in 1 day, 2, 3, 4 and 5 days is 773.6mm, 849.34mm, 874.19mm, 931.78mm and 957.89mm, respectively. The magnitudes of 1 day as well as 2 to 5 consecutive days annual maximum rainfall corresponding to 2 to 100 years return period were estimated using Gamma function. RamaRao et al. (1975) carried out an annual one day maximum rainfall and two to five consecutive days maximum rainfall corresponding to a return period of 2 to 100 years. Based on the best fit probability distribution a maximum of 84.05 mm in 1 day, 91.60 mm in 2 days, 100.40 mm in 3 days, 105.67 mm in 4 days and 109.47 mm in 5 days is expected to occur at Accra every two years. Similarly a maximum rainfall of 230.97 mm, 240.49, 272.77 mm, 292.07 mm and 296.54 mm is expected to occur in 1 day, 2, 3, 4 and 5 days respectively every 100 years. The results from the study could be used as a rough guide by engineers and hydrologists during the design and construction of drainage systems in the Accra metropolis as poor drainage has been identified as one of the major factors causing flooding in Accra. Sharda et al. (1985) carried out an probability distributions to predict rainfall status of various return period estimating one day and two to seven consecutive days annual maximum rainfall of Boalia, Rajshahi, Bangladesh. Based on the best fit probability distribution the maximum rainfall of 116.15, 161.09, 190.14, 205.96, 220.37, 234.66 and 245.21 mm was expected to occur 1, 2, 3, 4, 5, 6 and 7 days respectively at Boalia every two years. For a recurrence interval of 100 years, the maximum rainfall expected in 1, 2, 3, 4, 5, 6 and 7 days were 290.24, 406.49, 544.08, 558.56, 600.33, 631.28 and 633.89 mm respectively. Upadhaya et al. (1998) carried out an rainfall analysis by Gumbel’s extreme value distributions for southern Telangana. A maximum of mean annual 1 day and 2 to 5 day consecutive rainfall was expected to 87.27mm, 108.2 mm, 119.73 mm, 128.61mm and 137.30 mm respectively. A maximum of 80.1 mm in 1 day, 93.1 mm in 2 days, 111.18mm in 3 days, 119.97 mm in 4 days, and 128.13 mm in 5 days is expected to occur at Hyderabad every 2 years. For recurrence interval of 100 years maximum rainfall expected in 1 day ,2, 3, 4 and 5 days is of 224.92 mm, 395.9 mm, 281.83 mm, 292.42 mm, and 311.12 mm respectively. The magnitude of 1 days as well as 2 to 5 consecutive days annual maximum rainfall corresponding to 2 to100 years return periods were estimated by gumble's methods for extreme event frequency analysis. Deidda et al. (2006) carried out a rainfall intensity and frequency estimated for the Gandak basin, a region prone to high floods with an unrealized and unexplored hydro-potential. A maximum rainfall of 25.07 mm in 1-day, 37.09 mm in 2-day, 46.53 mm in 3-day, 55.8 mm in 4-day and 65.51 mm in 5 consecutive days for a return period of two years is likely to occur. Similarly, a precipitation maximum of 42.09, 65.83, 79.31, 86.96 and 94.03 mm in 1, 2, and 3, 4 and 5 consecutive days respectively is expected to occur for a return period of 100 years. Such estimates of maximum precipitation for different return periods are substantially facilitative in planning the safe and economic design of various engineered structures such as bridges, culverts, levees, canals and other irrigation and drainage works. Generally, design life of a hydraulic structure decides the period to be considered for design calculations of a structure. For example, the design life of a bridge can vary from 25 years to 100 years depending on the importance and size of the structure. Similar assumptions have to be made in the case of other hydraulic structures. Bhakar et al. (2008) carried out a study of annual one day maximum rainfall and two, three, four, five, six and seven consecutive day's maximum rainfall corresponding to a return period of 2 to 100 years has been conducted for the Panam dam, Gujarat, India. Based on the best fit probability distribution, a maximum of 125.56 mm in 1 day, 149.03 mm in 2 days, 168.49 mm in 3 days, 180.36 mm in 4 days, 191.64 mm in 5 days, 216.89 mm in 6 days and 213.61 mm in 7 days is expected to occur at Panam dam every two years. Similarly a maximum rainfall of 413.82 mm, 419.98 mm, 433.82 mm,456.98 mm, 470.87 mm, 473.33 mm and 520.87 mm is expected to occur in 1, 2,3,4,5,6 and 7 days respectively every 100 years. Flood frequency analysis is a statistical tool to help engineers, hydrologists, and watershed managers to deal with this uncertainty. Frequency analysis of rainfall data enables us to determine the rainfall for future use and also for the water resource planning and management like reservoirs. The results from the study could be used as a rough guide by engineers and hydrologists during the design and construction of drainage systems in the catchment area of Panam dam and computation of drainage co-efficient. The study of these research papers helps me a lot to evaluate the rainfall magnitude for different return periods and also to ascertain the type of probability distribution that best fits the rainfall data of Shimla and Dharamshala (H.P.), India.

1.3 Scope of study

Analysis of rainfall and determination of annual maximum daily rainfall would enhance the management of water resources applications as well as the effective utilization of water resources. Probability and frequency analysis of rainfall data enables us to determine the expected rainfall at various chances. Such information can also be used to prevent floods and droughts, and applied to planning and designing of water resources related to engineering such as reservoir design, flood control work, soil and water conservation planning. Useful for forecasting the floods to downstream towns and villages. Probability analysis of rainfall is necessary for solving various water management problems and to access the crop failure due to deficit or excess rainfall. The present study aims to evaluate the rainfall magnitude for different return periods and also to ascertain the type of probability distribution that best fits the rainfall data of Shimla (H.P.), India.

1.4 Objectives

The objectives for the rainfall analysis are as follows:

a) Probability analysis for one day to seven consecutive days annual maximum rainfall for Dharamshala and Shimla region, Himachal Pradesh.

b) Probability and frequency analysis of rainfall data enables us to determine the expected rainfall at various chances.

c) Such information can also be used to prevent floods and droughts, and applied to planning and designing of water resources related to engineering such as reservoir design, flood control work, soil and water conservation planning and design of small and medium hydraulic structures such as small dams, bridges, culverts drainage works, etc.

d) Analysis of rainfall and determination of annual maximum daily rainfall would enhance the management of water resources applications as well as the effective utilization of water resources. In the analysis of rainfall data we have to check the data homogeneity.

e) For analysis of rainfall data from a single series should ideally possess property of homogeneity - i.e. properties or characteristics of different portion of the data series do not vary significantly. Rainfall data for multiple series at neighbouring stations should ideally possess spatial homogeneity. Tests of homogeneity are required for validation purposes.

1.5 Limitations of the study

The limitations exist in measuring the accurate rainfall data are as follows:

- **a)** These limitations come from varieties of sources that influence the accuracy in recording the data.
- **b)** The physical obstructions -- such as trees, buildings, flow measurement structures and others located in the immediate vicinity of recording instrument influence the rainfall data.
- **c)** Also, changing weather conditions, such as wind speed or wind direction during the rainfall event, impact the rainfall data.
- **d)** The errors in data measurements could come from malfunction of instruments including power surge, power failure, mechanical mechanism and/or electronic system failure, and require re-calibration. However, as the rainfall duration increases from daily to monthly, seasonal or annual, the spatial variation also reduces significantly and becomes explainable through stronger correlation.
- **e)** To obtain geographically better representation of daily rainfall data, a denser rain gage network is needed. The limitation of our study is that we have only use 20 years of data.

CHAPTER 2

LITERATURE REVIEW

2.1 General

Rainfall intensities of various frequencies and durations are the basic inputs in hydrologic design. They are used, for example, in the design of storm sewers, culverts and many other structures as well as inputs to rainfall-runoff models. Precipitation frequency analysis is used to estimate rainfall depth at a point for a specified exceedance probability and duration. A number of research papers are available on the topic of frequency analysis of rainfall data for both different parts in India and also worldwide (Benson et al. (1968), RamaRao et al. (1975), Baheerathan et al. (1978), Kulandaivelu et al. (1984), Sharda et al. (1985), Dalabehra et al. (1993), Hirose et al. (1994), Duan et al. (1995), Upadhaya et al. (1998), Mohanty et al. (1999), Rizvi et al. (2001), Tao et al. (2002), Fowler et al. (2003), Lee et al. (2005), Deidda et al. (2006), Nadarajah et al. (2007), Bhakar et al. (2008), Olofintoye et al. (2009), Sharma et al. (2010), Suribabu et al. (2015) etc.). Only a few has been reviewed here in the following section.

2.2 Frequency analysis of rainfall data

Benson et al. (1968) carried out an annual one day maximum rainfall and two to five days consecutive days maximum rainfall corresponding to return period varying from 2 to 100 years. A maximum of 154.31mm in 1 day, 250.88mm in 2 days, 270.15mm in 3 days, 284.18mm in 4 days and 295.54mm in 5 days is expected to occur at Udaipur, Rajasthan every 2 years. For a recurrence interval of 100 years, the maximum rainfall expected in 1 day, 2, 3, 4 and 5 days is 773.6mm, 849.34mm, 874.19mm, 931.78mm and 957.89mm, respectively. The magnitudes of 1 day as well as 2 to 5 consecutive days annual maximum rainfall corresponding to 2 to 100 years return period were estimated using Gamma function. RamaRao et al. (1975) carried out an annual one day maximum rainfall and two to five consecutive days maximum rainfall corresponding to a return period of 2 to 100 years. Based on the best fit probability distribution a maximum of 84.05 mm in 1 day, 91.60 mm in 2 days, 100.40 mm in 3 days, 105.67 mm in 4 days and 109.47 mm in 5 days is expected to occur at Accra every two years. Similarly a maximum rainfall of 230.97 mm, 240.49, 272.77 mm, 292.07 mm and 296.54 mm is expected to occur in 1 day, 2, 3, 4 and 5 days respectively every 100 years. The results from the study could be used as a rough guide by engineers and hydrologists during the design and construction of drainage systems in the Accra metropolis as poor drainage has been identified as one of the major factors causing flooding in Accra. Sharda et al. (1985) carried out an probability distributions to predict rainfall status of various return period estimating one day and two to seven consecutive days annual maximum rainfall of Boalia, Rajshahi, Bangladesh. Based on the best fit probability distribution the maximum rainfall of 116.15, 161.09, 190.14, 205.96, 220.37, 234.66 and 245.21 mm was expected to occur 1, 2, 3, 4, 5, 6 and 7 days respectively at Boalia every two years. For a recurrence interval of 100 years, the maximum rainfall expected in 1, 2, 3, 4, 5, 6 and 7 days were 290.24, 406.49, 544.08, 558.56, 600.33, 631.28 and 633.89 mm respectively. Upadhaya et al. (1998) carried out an rainfall analysis by Gumbel's extreme value distributions for southern Telangana. A maximum of mean annual 1 day and 2 to 5 day consecutive rainfall was expected to 87.27mm, 108.2 mm, 119.73 mm, 128.61mm and 137.30 mm respectively. A maximum of 80.1 mm in 1 day, 93.1 mm in 2 days, 111.18mm in 3 days, 119.97 mm in 4 days, and 128.13 mm in 5 days is expected to occur at Hyderabad every 2 years. For recurrence interval of 100 years maximum rainfall expected in 1 day ,2, 3, 4 and 5 days is of 224.92 mm, 395.9 mm, 281.83 mm, 292.42 mm, and 311.12 mm respectively. The magnitude of 1 days as well as 2 to 5 consecutive days annual maximum rainfall corresponding to 2 to100 years return periods were estimated by gumble's methods for extreme event frequency analysis. Rizvi et al. (2001) carried out a study that includes rainfall probability analysis of previous 34 years rainfall data (1980-2013) with the prime objective for prediction of annual rainfall of Allahabad district. The observed values were computed by weibulls formula (1939). The annual rainfall values were estimated by proposed prediction models Viz. Gumbel and Log Normal (Chow 1964). The rainfall data in the above distribution and their corresponding rainfall events were estimated at 2.9, 11.4, 20.0, 40.0, 51.4, 60.0, 80.0 and 97.1 percent probabilities level. The goodness of fit was tested by Chi-square test. It clearly indicates that the Gumbel distribution was found to be best model for predicting the annual rainfall (mm). While Log Normal distribution is fairly close to the observed annual rainfall (mm). Deidda et al. (2006) carried out a rainfall intensity and frequency estimated for the Gandak basin, a region prone to high floods with an unrealized and unexplored hydro-potential. A maximum rainfall of 25.07 mm in 1-day, 37.09 mm in 2-day, 46.53 mm in 3-day, 55.8 mm in 4-day and 65.51 mm in 5 consecutive days for a return period of two years is likely to occur. Similarly, a precipitation maximum of 42.09, 65.83, 79.31, 86.96 and 94.03 mm in 1, 2, and 3, 4 and 5 consecutive days respectively is expected to occur for a return period of 100 years. Such estimates of maximum precipitation for different return periods are substantially facilitative in planning the safe and economic design of various engineered structures such as bridges, culverts, levees, canals and other irrigation and drainage works. Generally, design life of a hydraulic structure decides the period to be considered for design calculations of a structure. For example, the design life of a bridge can vary from 25 years to 100 years depending on the importance and size of the structure. Similar assumptions have to be made in the case of other hydraulic structures. Bhakar et al. (2008) carried out a study of annual one day maximum rainfall and two, three, four, five, six and seven consecutive day's maximum rainfall corresponding to a return period of 2 to 100 years has been conducted for the Panam dam, Gujarat, India. Based on the best fit probability distribution, a maximum of 125.56 mm in 1 day, 149.03 mm in 2 days, 168.49 mm in 3 days, 180.36 mm in 4 days, 191.64 mm in 5 days, 216.89 mm in 6 days and 213.61 mm in 7 days is expected to occur at Panam dam every two years. Similarly a maximum rainfall of 413.82 mm, 419.98 mm, 433.82 mm,456.98 mm, 470.87 mm, 473.33 mm and 520.87 mm is expected to occur in 1, 2,3,4,5,6 and 7 days respectively every 100 years. The results from the study could be used as a rough guide by engineers and hydrologists during the design and construction of drainage systems in the catchment area of Panam dam and computation of drainage co-efficient. Olofintoye et al. (2009) carried out an extreme value analysis of rainfall for Tiruchirapalli City in Tamil Nadu using 100 years of rainfall data. In this study, an extreme value analysis of rainfall for Tiruchirapalli City in Tamil Nadu was carried out using 100 years of rainfall data. Statistical methods were used in the analysis. The best-fit probability distribution was evaluated for 1, 2, 3, 4 and 5 days of continuous maximum rainfall. The goodness of fit was evaluated using Chi-square test. The results of the goodness-of-fit tests indicate that log-Pearson type III method is the overall best-fit probability distribution for 1-day maximum rainfall and consecutive 2-, 3-, 4-, 5- and 6-day maximum rainfall series of Tiruchirapalli. To be reliable, the forecasted maximum rainfalls for the selected return periods are evaluated in comparison with the results of the plotting position. Sharma et al. (2010) carried out a study for the analysis of rainfall data for western Maharashtra region. Monthly rainfall data for 35 years (1966-2000) measured by Indian Meteorological Department, Pune was used. It was found that, in Ahmednagar and Solapur district Beta distribution is the best fitted distribution for one month rainfall dataset, Dagum distributions for consecutive cumulative two months rainfall dataset, Dagum and Gen.Extreme distribution for consecutive cumulative three months rainfall dataset, Log Logistic and Gen.Extreme distribution for consecutive cumulative four and five months rainfall dataset, Log Logistic and Nakagami distributions for consecutive cumulative six months rainfall dataset and Log Logistic and Johnson SU distribution for consecutive cumulative seven months rainfall. Suribabu et al. (2015) carried out estimates that were made of magnitude/frequency relationships for daily rainfall amounts recorded at representative Nigerian stations. Probable maximum precipitation estimates were also made based on the Chow general frequency formula. Three commonly used probability distributions (Normal, Log Normal and Gamma distribution) were tested to determine the best fit probability distribution using the comparison of chi-square values. Results showed that the log-normal distribution was the best fit probability distribution for one day and two to seven consecutive days annual maximum rainfall for the region.

2.3 Concluding Remarks

The following summary is drawn on the basis of the literature review:

(a) The log-normal distribution is the best fit probability distribution for one day annual maximum as well as two, three and four consecutive days, while for five and seven consecutive days, gamma distribution and for six days normal distribution fits better for the region.

(b) The study of frequency analysis will be used as a rough guide by engineers and hydrologists to prevent floods and droughts, and applied to planning and designing of water resources related to engineering such as reservoir design, flood control work, soil and water conservation planning and design of small and medium hydraulic structures such as small dams, bridges, culverts drainage works, etc.

(c) Analysis of rainfall and determination of annual maximum daily rainfall would enhance the management of water resources applications as well as the effective utilization of water resources.

(d) The results of this study would be useful for agricultural scientists, decision makers, policy planners and researchers in order to identify the areas where agricultural development and construction of drainage systems takes place.

CHAPTER 3

MATERIALS AND METHODOLOGY

3.1 General

Probability and frequency analysis of rainfall data enables us to determine the expected rainfall at various chances. Such information can also be used to prevent floods and droughts, and applied to planning and designing of water resources related to engineering such as reservoir design, flood control work, soil and water conservation planning and design of small and medium hydraulic structures such as small dams, bridges, culverts drainage works, etc. Analysis of rainfall and determination of annual maximum daily rainfall would enhance the management of water resources applications as well as the effective utilization of water resources. This study is useful for agricultural scientists, decision makers, policy planners and researchers in order to identify the areas where agricultural development and construction of drainage systems takes place. The present study aims to evaluate the rainfall magnitude for different return periods and also to ascertain the type of probability distribution that best fits the rainfall data of Shimla and Dharamshala (H.P.), India. So in this chapter, various probability distributions and transformations were applied to estimate one day and two to seven consecutive days annual maximum rainfall of various return periods. Three commonly used probability distributions (viz: Normal, Log Normal and Gamma distribution) were tested by comparing the Chi-square, Anderson-Darling and Kolmogorov–Smirnov values.

3.2 Materials and Methods

Annual daily rainfall data for 20 years are collected from Indian Meteorological Department, Shimla (H.P.), India. The annual daily rainfall data of 20 years was used for analysis of probability distribution. The goodness of fit using Kolmogorov Smirnov, Anderson Darling and Chi-Square were tested. Three commonly used probability distributions normal, lognormal and gamma distribution were tested to determine the best fit probability distribution that describes the annual one day maximum and two to seven consecutive days maximum rainfall series by comparing observed values with tabulated Chi-square, Anderson-Darling and Kolmogorov–Smirnov values. It is the data analysis and simulation application allowing to probability distributions to apply the analysis results to make better decisions.

The types of processing considered in this module are:

a) Computation of basic statistics
b) Fitting of frequency distributions

3.3 Computation of basic statistics

Basic statistics are widely required for validation and reporting. The following are commonly used: arithmetic mean, skewness and kurtosis, median - the median value of a ranked series X_i, mode - the value of X which occurs with greatest frequency or the middle, value of the class with greatest frequency. From these selected values of exceedance probability or non-exceedance probability can be extracted, e.g. the daily rainfall which has been exceeded 1%, 5% or 10% of the time.

3.4 Fitting of frequency distributions

A common use of rainfall data is in the assessment of probabilities or return periods of given rainfall at a given location. Such data can then be used in assessing flood discharges of given return period through modeling or some empirical system and can thus be applied in schemes of flood alleviation or forecasting and for the design of bridges and culverts. Frequency analysis usually involves the fitting of a theoretical frequency distribution using a selected fitting method, although empirical graphical methods can also be applied. The fitting of a particular distribution implies that the rainfall samples of annual maxima were drawn. For each distribution one can obtain the following: Estimation of parameters of the distribution, a table of rainfalls of specified exceedance probabilities or return periods with confidence limits, results of goodness of fit tests, a graphical plot of the data fitted to the distribution.

3.5 Statistical analysis of data

The mean, variance, standard deviation, coefficient of variation and coefficient of skewness which describe the variability of rainfall were computed (Walck 2007). The mean, variance, standard deviation, coefficient of variation and coefficient of skewness are describe below by statistical formulas:

a) Mean

Mean represents measures of central tendency. Arithmetic mean is given by

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{1}$$

Where,

$\overline{X}$ = mean

X_i = variate

N = total number of observations

b) Standard deviation

This parameter, as a measure of variability is most adoptable to statistical analysis. It is computed by

$$S = \sqrt{\frac{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2}{(N-1)}} \tag{2}$$

Where, S = standard deviation

c) Coefficient of variation

This is a dimensionless dispersion parameter usually expressed as percent. It is given by

$$C_v = \frac{S}{\overline{X}} \tag{3}$$

Where, C_v = coefficient of variation

d) Coefficient of skewness

The coefficient of skewness was determined for each set of data from the formula given below:

$$C_s = \frac{N^2 M_3}{(N-1)(N-2)S^3} \tag{4}$$

Where,

C_s = coefficient of skewness

$$M_3 = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^3 \quad (5)$$

3.6 Testing the goodness of fit

Goodness of fit can be analyzed by comparing the theoretical and sample values of the relative frequency of the cumulative frequency function of a probability distribution (Dabral et al. 2009). In case of the relative frequency function, the Chi-square test is used.

3.6.1 Chi-Square test

Chi-square test is the commonly used test for testing the goodness of fit of empirical data. In applying the Chi-square goodness of fit test, the data is grouped into suitable frequency classes. The test compares the actual number of observations and the expected number of observations. The expected number of observations is obtained by considering the above discussed probability distributions one by one falling in the respective class intervals. The expected numbers are calculated by multiplying the expected relative frequency by total number of observations. The chi-square test statistic is computed from the equation:

$$\chi^2 = \sum \frac{(O - E)^2}{E} \quad (6)$$

χ^2 = the test statistic , O = observed frequencies

E = expected frequencies

The chi-square distribution has *(N - k -1)* degree of freedom. The best probability distribution function was determined by comparing Chi-square values obtained from each distribution and selecting the function that gives smallest Chi-square value.

The sample value of the relative frequency of interval *i* is calculated by the following equation:

$$f_s(x_i) = \frac{n_i}{n} \quad (7)$$

The theoretical value of the relative probability function is

$P(X_i) = F(X_i) - F(X_{i-1})$

The Chi-square test statistic χ_c^2 is given by the equation

$$\chi_c^2 = \sum_{i=1}^{m} n\left[\frac{\left(f_s(X_i) - P(X_i)\right)^2}{P(X_i)}\right] \tag{8}$$

Where, $n\, f_s\,(x_i) = n_i$ is the observed number of occurrences in interval i, $n\,P\,(x_i)$ is the corresponding expected number of occurrences in interval i

The χ^2 distribution functions are tabulated in many statistics texts. In the χ^2 test,

$v = m - p - 1$

Where, v = degree of freedom,

m = number of intervals,

p = number of parameters used in fitting the proposed distribution

A confidence level is chosen for the test, it is often express as 1-α, where 'α' is termed as the significant level. A typical value for the confidence level is 95 per cent. The null hypothesis for the test is that the proposed probability fits the data adequately. This hypothesis is rejected if the value of χ_c^2 is larger than a limiting value, $\chi^2_{\nu,\ 1-\alpha}$ (which is determined from the χ^2 distribution with ν degree of freedom at 5 % level of significance. Otherwise it was rejected.

3.6.2 Kolmogorov-Smirnov test

In statistics , the Kolmogorov Smirnov test (*K–S* test or *KS* test) is a non-parametric test of the equality of continuous, one-dimensional probability distributions that can be used to compare a sample with a reference probability distribution (one-sample K–S test), or to compare two samples (two-sample *K–S* test). The empirical distribution function F_n for n observations, X_i is defined as

$$F_n(x) = \frac{1}{N}\sum_{i=1}^{n} I\{x_i \leq x\} \tag{9}$$

where, $I\{x_i \leq x\}$ is the indicator function, equal to 1 if and equal to 0 otherwise. The Kolmogorov–Smirnov statistic for a given cumulative distribution function $F(x)$ is:

$T = sup_x \mid F^*(x) - S(x) \mid$

where, sup_x is the supre mum of the set of distances.

3.6.3 Anderson Darling test

The Anderson-Darling Test will determine if a data set comes from a specified distribution, in our case, the normal distribution. The test makes use of the cumulative distribution function. The Anderson-Darling statistic is given by the following formula:

$$A^2 = \left[\frac{\sum_{i=1}^{n} (2i-1)(\ln(u_i) + \ln(1-u_{n+1-i}))}{n} - n \right] \quad (10)$$

where , $u_i = F(x_i)$, n = sample size

$F(X_i)$ = cumulative distribution function for the specified distribution

i = the i^{th} sample when the data is sorted in ascending order

$AD = -n - S$

Summation term in the Anderson-Darling equation:

$$S = (2i - 1)[\ln F(X_i) + \ln(1 - F(X_{n-i+1}))] \quad (11)$$

3.7 Probability analysis using Frequency factors

Three commonly used probability distributions namely Normal, Log Normal and Gamma distribution were tested by comparing the Chi-square, Anderson-Darling and Kolmogorov–Smirnov values.

Normal distribution

The normal distribution, a two parameter distribution, has been identified as the most important distribution of continuous variables applied to symmetrically distributed data. The figure for normal distribution curve is shown below in figure 3.1:

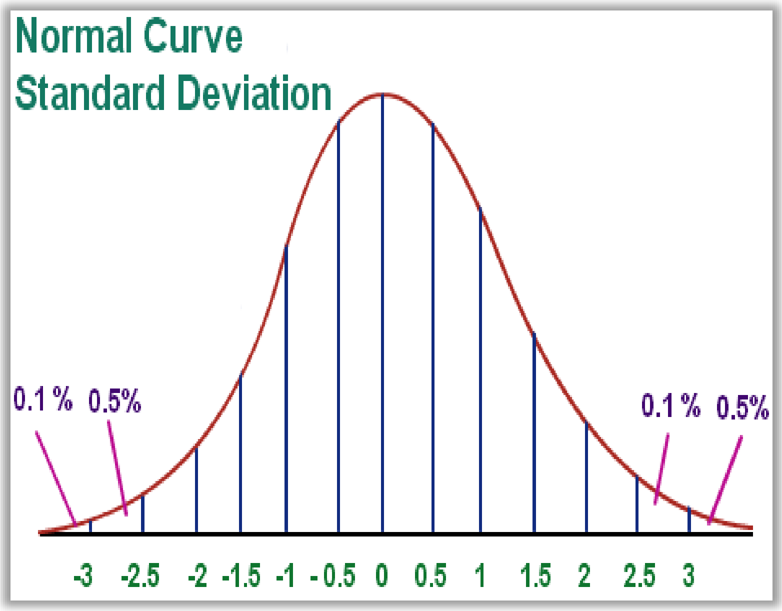

Fig. 3.1: Normal distribution curve (Walck 2007)

Log normal distribution

Most hydrological continuous random variables are noted to be asymmetrically distributed hence it is computationally advantageous to transform the distribution to a lognormal distribution. The transformation is usually achieved by taking the logs of the variables in question. A random variable x is said to follow a lognormal distribution if the logarithm (usually natural logarithm) of x is normally distributed. The figure for log normal distribution curve is shown below in figure 3.2:

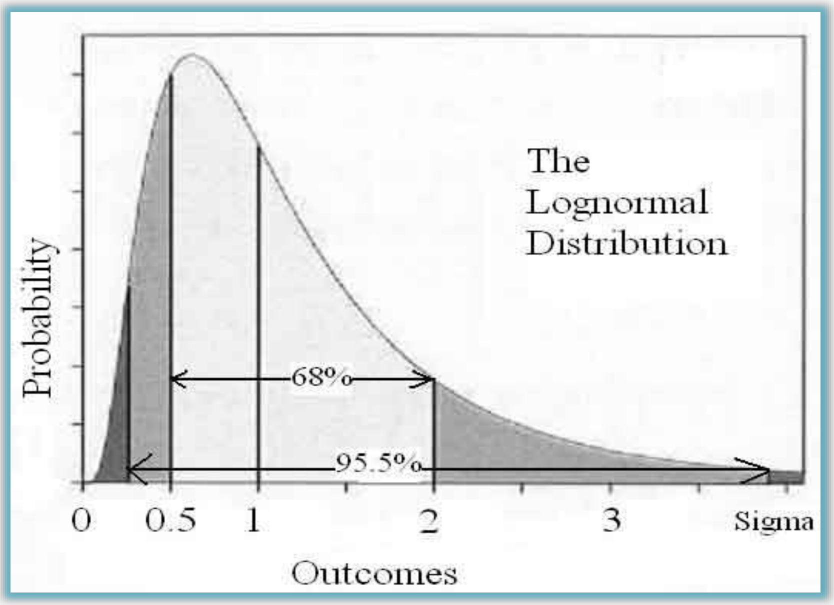

Fig. 3.2: Lognormal distribution curve (Walck 2007)

Gamma distribution

In hydrology the gamma distribution has an advantage of having only positive values. The gamma function is defined such that the total area under the density function is unity. The figure for gamma distribution curve is shown below in figure 3.3:

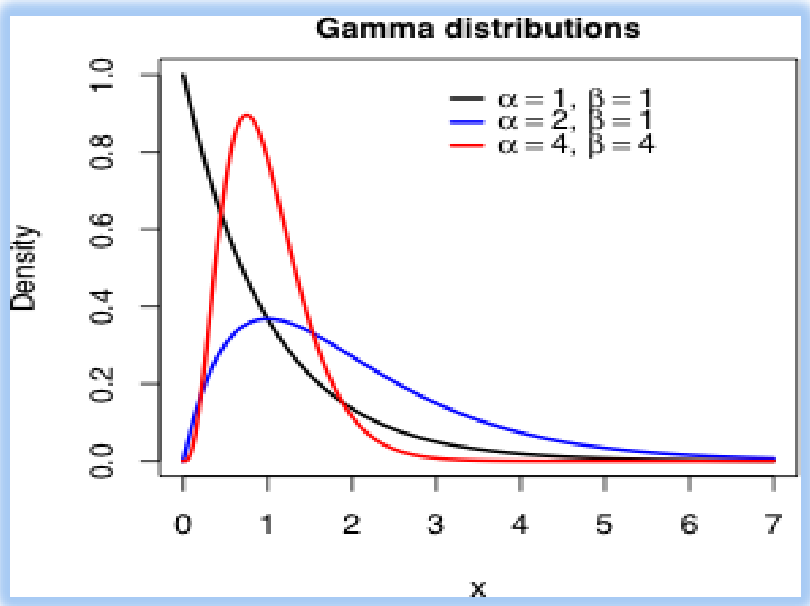

Fig. 3.3: Gamma distribution curve (Walck 2007)

i. Normal & Log-Normal	*CHOW METHOD*
ii. Gamma	*HANN METHOD*

3.8 Frequency analysis using frequency factors

For **Normal** and **Log Normal** distribution, the frequency factor can be expressed by the following equation (Chow, 1988)

$$K_T = \frac{X_T - \mu}{\sigma} \qquad (12)$$

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

Where,

$$w = \left[\ln\left(\frac{1}{p^2} \right) \right]^{\frac{1}{2}} \qquad (0 < p \le 0.5) \tag{13}$$

Then calculating z using the equation

$$z = w - \left[\frac{(2.515517 + 0.802853w + 0.010328w^2)}{(1 + 1.432788w + 0.189269w^2 + 0.001308w^3)} \right] \tag{14}$$

When $p > 0.5$, $1\text{-}p$ is substituted for p in equation (10) and the value of z is computed by equation (11) is given a negative sign. The frequency factor K_T for the normal distribution is equal to z, as mentioned above.

In case of **Gamma** distribution, frequency analysis was done by the method as described by Hann (1994) as shown in table 3.1:

Table 3.1: Description of various probability distribution functions (Walck 2007).

Distribution	Probability density function	Range	Equation for the parameters in terms of the sample moment
Normal	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	-α ≤ x ≤ α	$\mu = x$ $\sigma = S_x$
Log Normal	$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-\mu_y)^2}{2\sigma_y^2} \right)$	x > 0	$\mu_y = y$ $\sigma_y = S_y$
Gamma	$f(x) = \frac{\lambda^\beta x^{\beta-1} e^{-\lambda x}}{T(\beta)}$	x ≥ 0	$\lambda = \frac{\bar{x}}{S^2 x}$ $\beta = \left(\frac{\bar{x}}{S_x} \right)^2$

Values of 1 day, 2, 3, 4, 5, 6 and 7 consecutive days maximum rainfall can be estimated statistically through the use of the Chow (1988) general frequency formula. The formula expresses the frequency of occurrence of an event in terms of a frequency factor, K_T , which depends upon the distribution of particular event investigated. Chow (1988) has shown that many frequency analyses can be reduced to the form

$$\boldsymbol{X_T = X (1 + C_V K_T)} \text{ where,} \tag{15}$$

C_V : coefficient of variation

K_T : frequency factor

X : mean value of X

X_T : magnitude of the event having a return period t

μ : mean of the sample

σ : standard deviation of the sample.

CHAPTER 4

ANALYSIS OF DATA AND DISCUSSION OF RESULTS

4.1 General

Annual daily rainfall data for 20 years (1992-2012) were collected from Indian Meteorological Department, Shimla (H.P.), India. The total data included in this study were 32900 rainfall values in mm for 1- day and the total effected data for 20 years was 16000 rainfall values. In previous chapter, we defined various methods to calculate frequency analysis of rainfall data. Using above methods we can analyse the data in this chapter and so in this chapter, various probability distributions were applied to estimate one day and two to seven consecutive days annual maximum rainfall of Shimla and Dharamshala region. Three commonly used probability distributions (viz: Normal, Log Normal and Gamma distribution) were tested by comparing the Chi-square, Anderson-Darling and Kolmogorov–Smirnov values and we can also check the type of distribution which is best fit for the region.

4.2 Results and Discussion

Annual one day maximum rainfall and two to seven days consecutive days maximum rainfall corresponding to return period varying from 2 to 20 years will be used by design engineers and hydrologists for the economic planning, design of small and medium hydrologic structures and determination of drainage coefficient for agricultural fields. For Shimla region, a maximum of 121mm in 1 day, 178.15 mm in 2 days, 203.6 mm in 3 days, 219.4 mm in 4 days, 230.23 mm in 5 days, 240.6 mm in 6 days and 257.5mm in 7 days is expected to occur at Shimla, Himachal Pradesh every 2 years. For a recurrence interval of 20 years, the maximum rainfall expected in 1 day, 2, 3, 4, 5, 6 and 7 days is 160.5 mm, 239.11 mm, 265.36 mm, 285.9 mm, 297.42 mm, 309.5 mm and 330.38 mm respectively. For Dharamshala region, a maximum of 191.1 mm in 1 day, 262.7 mm in 2 days, 314.27 mm in 3 days, 370.24 mm in 4 days, 419.7 mm in 5 days, 477.4 mm in 6 days and 522.32 mm in 7 days is expected to occur at Dharamshala, Himachal Pradesh every 2 years. For a recurrence interval of 20 years, the maximum rainfall expected in 1 day, 2, 3, 4, 5, 6 and 7 days is 277.7 mm, 373.6 mm, 445.11 mm, 518.62 mm, 589 mm, 680.3 mm and 753.79 mm respectively. The magnitudes of 1 day as well as 2 to 7 consecutive days annual maximum rainfall corresponding to 2 to 20 years return period were estimated using Gamma function. Various probability distributions and transformations were applied to estimate one day and two to seven consecutive days annual maximum rainfall of various return periods. Three commonly used probability distributions (viz: Normal, Log Normal and Gamma distribution) were tested by comparing the Chi-square value. Gamma distribution is found to be best fit for the region. Table 4.1 and Table 4.2 gives the 1 day and consecutive days maximum rainfall for different return periods at Shimla and Dharamshala region as determined by selected distribution. This study is useful for agricultural scientists, decision makers, policy planners and researchers in order to identify the areas where agricultural development and construction of drainage systems takes place. Such information can also be used to prevent floods and droughts, and applied to planning and designing of water resources related to engineering such as reservoir design, flood control work, soil and water conservation planning and design of small and medium hydraulic structures such as small dams, bridges, culverts drainage works, etc.

Table 4.1: 1 day as well as consecutive days maximum rainfall for various return periods at Shimla region.

S. No.	Return Period	1 Day	2 Days	3 Days	4 Days	5 Days	6 Days	7 Days
1	1.25	117.4	172.75	198.13	213.5	224.27	234.5	251.04
2	2	121	178.15	203.6	219.4	230.23	240.6	257.5
3	5	130.5	192.81	218.45	235.4	246.39	257.2	275.03
4	10	144.5	214.42	240.35	258.9	270.2	281.6	300.86
5	20	160.5	239.11	265.36	285.9	297.42	309.5	330.38

Table 4.2: 1 day as well as consecutive days maximum rainfall for various return periods at Dharamshala region.

S. No.	Return Period	1 Day	2 Days	3 Days	4 Days	5 Days	6 Days	7 Days
1	1.25	183.5	252.9	302.68	357.09	404.8	459.4	501.81
2	2	191.1	262.7	314.27	370.24	419.7	477.4	522.32
3	5	211.9	289.4	345.74	405.93	460.4	526.2	577.99
4	10	242.6	328.7	392.11	458.52	520.4	598.1	660.03
5	20	277.7	373.6	445.11	518.62	589	680.3	753.79

In Table 4.3, the mean value of one-day annual maximum rainfall at Shimla is found to be 98.9mm with standard deviation and coefficient of variation of 25.06 and 57.19 respectively. The coefficient of skewness is 1.541. For 2 to 7 days consecutive annual maximum rainfall range values for mean, standard deviation, coefficient of variation and coefficient of skewness are 144.2 - 216.9 mm , 38.581 - 46.131, 46.239 - 31.127 and -0.041 - 0.3089. It is observed that all distribution fitted function significantly. In Table 4.4, the mean value of one-day annual maximum rainfall at Dharamshala is found to be 142.9 mm with standard deviation and coefficient of variation of 54.8 and 51.34 respectively. The coefficient of skewness is 1.1. For 2 to 7 days consecutive annual maximum rainfall range values for mean, standard deviation, coefficient of variation and coefficient of skewness are 201 – 393.4 mm , 70.17 – 146.5, 41.65 – 30.47 and 0.726 – 1.593. The values of table 4.3 and 4.4 were obtained from appendix I & II.

Table 4.3: Statistical parameters of annual 1 day to 7 consecutive days maximum rainfall at Shimla region.

Parameters	**1 day**	**2 days**	**3 days**	**4 days**	**5 days**	**6 days**	**7 days**
Maximum (mm)	177.1	219.9	244.7	244.7	257.5	261.9	283.3
Minimum (mm)	44.7	71.7	93.9	103.3	112.4	119.5	120.9
Mean (mm)	98.9	144.2	169.2	182.3	192.8	202.3	216.9
Standard deviation (mm)	25.06	38.581	39.091	42.11	42.53	43.574	46.131
Coefficient of variation	57.19	46.239	40.67	37.121	34.558	32.604	31.127
Coefficient of skewness	1.541	0.3089	0.2278	0.0857	0.2913	0.0557	-0.041

Table 4.4: Statistical parameters of annual 1 day to 7 consecutive days maximum rainfall at Dharamshala region.

Parameters	**1 day**	**2 days**	**3 days**	**4 days**	**5 days**	**6 days**	**7 days**
Maximum (mm)	279.6	334.4	397.2	475.8	574.0	691.6	770.4
Minimum (mm)	84.1	104.8	128.7	156.9	203	218.4	240.9
Mean (mm)	142.9	201.0	241.4	287.6	325.5	364.4	393.4
Standard deviation (mm)	54.8	70.17	82.81	93.91	107.1	128.4	146.5
Coefficient of variation	51.34	41.65	37.18	34.56	32.77	31.47	30.47
Coefficient of skewness	1.1	0.726	0.886	0.785	1.12	1.279	1.593

4.2.1 Chi-Square test

The data presented in Table 4.5 and 4.6 reveals that 1 day, 2, 3, 4, 5, 6 and 7 consecutive days maximum rainfall followed Normal, Log Normal and Gamma distribution. The data presented in Table 4.5 and Table 4.6 revealed that the computed Chi-square values for three probability distribution i.e. Normal, Log Normal and Gamma are found to be less than the critical value of Chi-square at 95% confidence level for 1 day as well as consecutive days maximum rainfall series. Gamma distribution gave minimum value of χ^2 for annual 1 day and 2 to 7 days consecutive maximum rainfall series at Shimla region whereas Log Normal distribution gave minimum value of χ^2 for annual 1 day and 2 to 7 days consecutive maximum rainfall series at Dharamshala region. The statistical comparison by Chi-square test for goodness of fit clearly shows that Gamma distribution and Log Normal is the best fitting representative function for rainfall frequency analysis at Shimla and Dharamshala region.

Table 4.5: Chi-square values for different distributions at Shimla

Consecutive days	Normal	Log Normal	Gamma
1 Day	2.64	4.77	2.32
2 Days	2.98	7.32	1.03
3 Days	3.78	8.65	1.38
4 Days	3.94	7.63	3.34
5 Days	6.58	3.35	1.59
6 Days	2.60	1.98	1.88
7 Days	5.73	1.48	2.41

Table 4.6: Chi-square values for different distributions at Dharamshala

Consecutive days	Normal	Log Normal	Gamma
1 Day	1.20	4.83	2.08
2 Days	1.04	2.19	6.59
3 Days	3.67	1.86	1.53
4 Days	1.68	1.14	2.44
5 Days	3.78	1.23	7.12
6 Days	1.82	2.34	2.64
7 Days	6.35	3.72	5.09

Various probability distributions and transformations are applied to estimate one day and two to seven consecutive days annual maximum rainfall of various return periods. Three commonly used probability distributions (viz: Normal, Log Normal and Gamma distribution) were tested by comparing the Chi-square, Anderson-Darling and Kolmogorov–Smirnov values. Such information can also be used to prevent floods and droughts, and applied to planning and designing of water resources related to engineering such as reservoir design, flood control work, soil and water conservation planning and design of small and medium hydraulic structures such as small dams, bridges, culverts drainage works, etc. The graphical plot as shown in figure 4.1 and figure 4.2 below shows one day annual maximum, 2 to 7 consecutive day's annual maximum rainfall data was fitted to consecutive days. The results of this study would be useful for agricultural scientists, decision makers, policy planners and researchers in order to identify the areas where agricultural development and construction of drainage systems takes place.

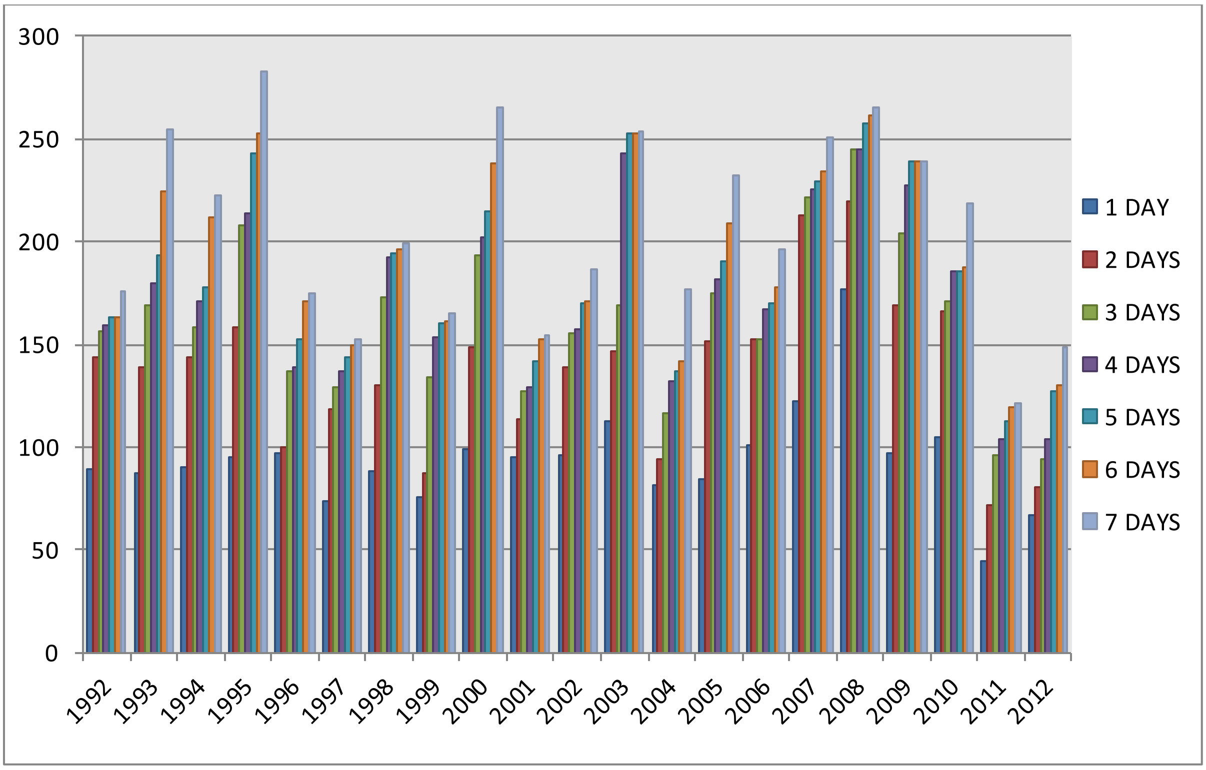

Fig. 4.1: One day annual maximum and 2 to 7 consecutive day's annual maximum rainfall data at Shimla

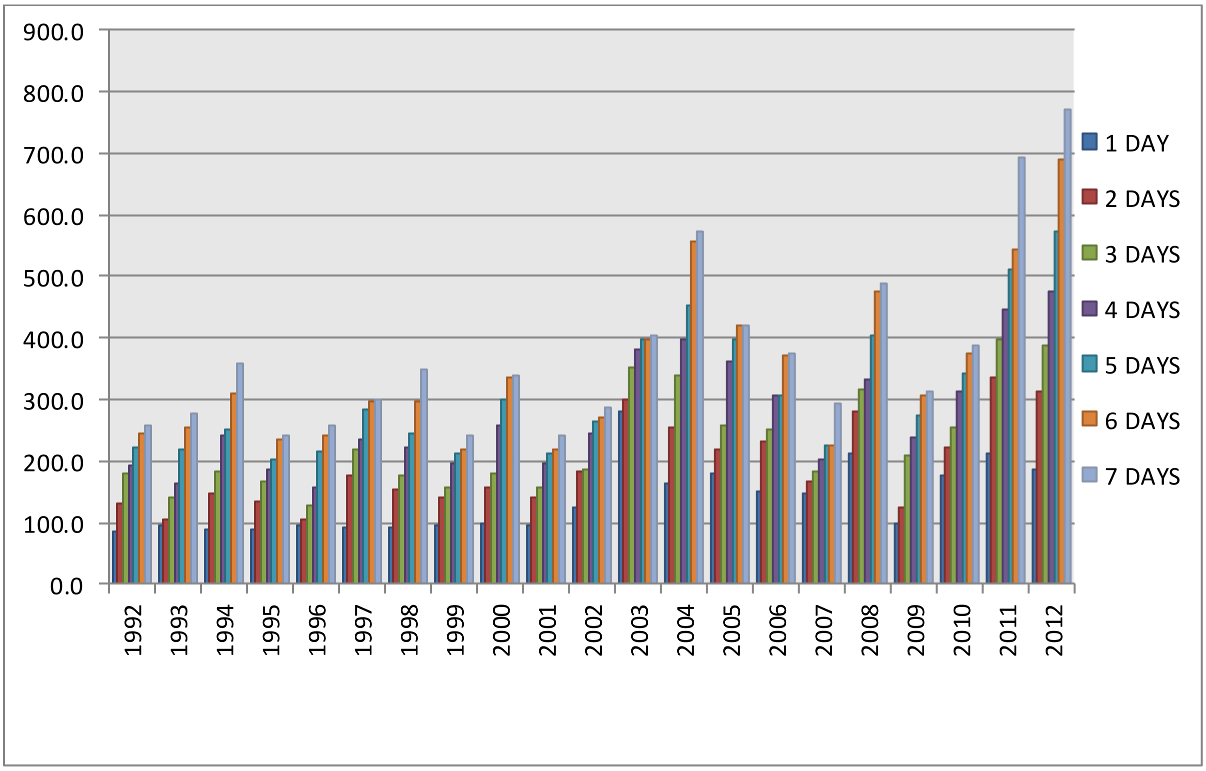

Fig. 4.2: One day annual maximum and 2 to 7 consecutive day's annual maximum rainfall data at Dharamshala

The graphical plot as shown in figure 4.3 and figure 4.4 below shows Chi-square values for different distribution at Shimla and Dharamshala region. Gamma distribution gave minimum value of χ^2 for annual 1 day and 2 to 7 days consecutive maximum rainfall series at Shimla region whereas Log Normal distribution gave minimum value of χ^2 for annual 1 day and 2 to 7 days consecutive maximum rainfall series at Dharamshala region. The statistical comparison by Chi-square test for goodness of fit clearly shows that Gamma distribution and Log Normal is the best fitting representative function for rainfall frequency analysis at Shimla and Dharamshala region.

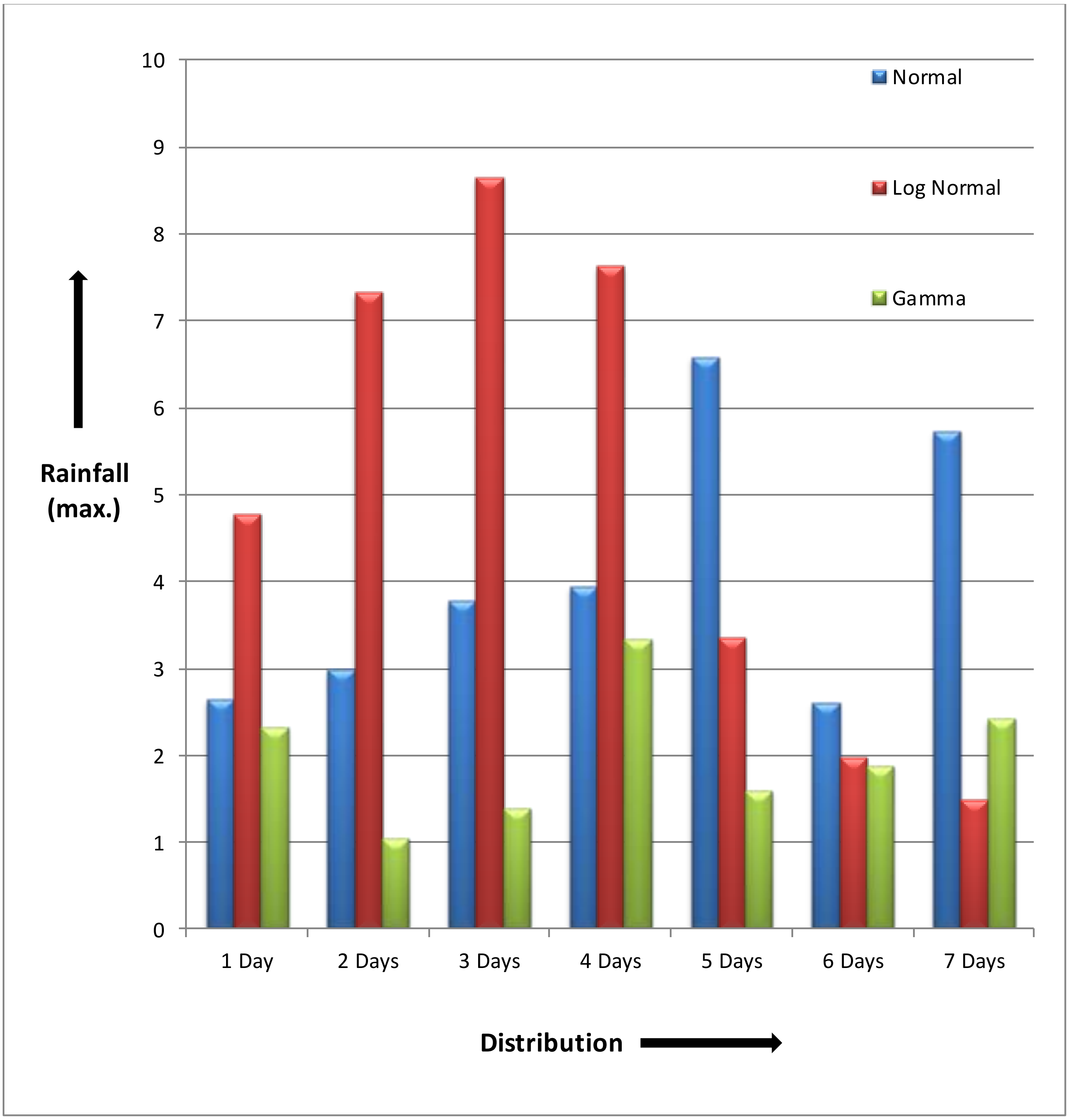

Fig. 4.3: Graphical plot showing Chi-square values for different distributions at Shimla

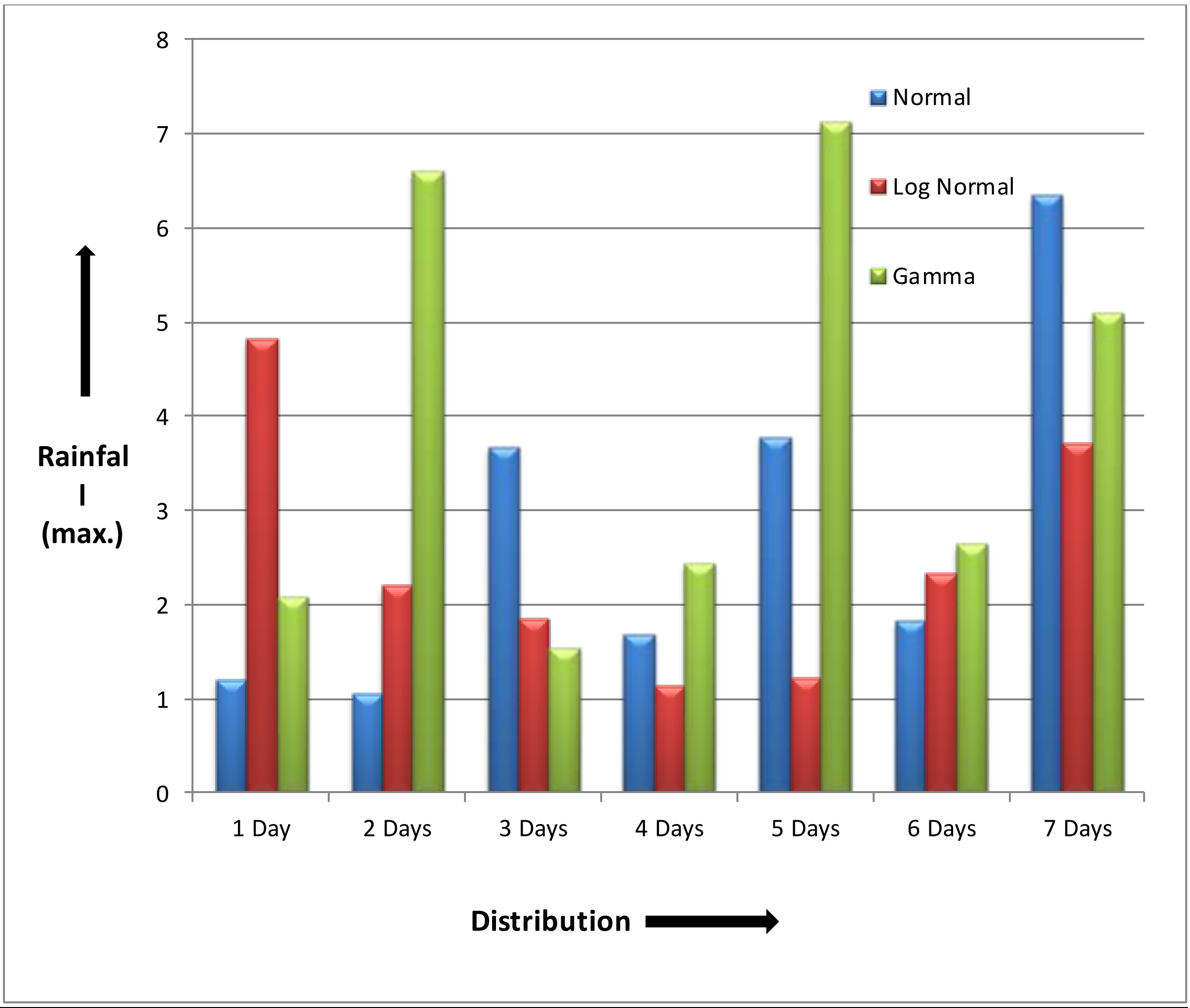

Fig. 4.4: Graphical plot showing Chi-square values for different distributions at Dharamshala

4.2.2 Kolmogorov-Smirnov test

The data presented in Table 4.7 and 4.8 reveals that 1 day, 2, 3, 4, 5, 6 and 7 consecutive days maximum rainfall followed Normal, Log Normal and Gamma distribution. The data presented in Table 4.7 and Table 4.8 revealed that the computed K-S values for three probability distribution i.e. Normal, Log Normal and Gamma are found to be less than the critical value of K-S at 95% confidence level for 1 day as well as consecutive days maximum rainfall series. Normal distribution gave minimum value of K-S for annual 1 day and 2 to 7 days consecutive maximum rainfall series at Shimla region whereas Gamma distribution gave minimum value of K-S for annual 1 day and 2 to 7 days consecutive maximum rainfall series at Dharamshala region. The statistical comparison by K-S test for goodness of fit clearly shows that Normal distribution and Gamma is the best fitting representative function for rainfall frequency analysis at Shimla and Dharamshala region.

Table 4.7: K-S values for different distributions at Shimla

Consecutive days	Normal	Log Normal	Gamma
1 Day	1.08	1.82	1.6
2 Days	2.94	3.68	2.88
3 Days	1.06	2.82	1.02
4 Days	3.59	4.64	1.32
5 Days	1.18	1.16	2.04
6 Days	2.61	2.92	3.32
7 Days	2.826	3.16	5.04

Table 4.8: K-S values for different distributions at Dharamshala

Consecutive days	Normal	Log Normal	Gamma
1 Day	1.05	2.16	1.82
2 Days	3.89	1.24	3.49
3 Days	2.16	3.96	2.16
4 Days	1.26	2.22	1.34
5 Days	2.08	1.32	2.82
6 Days	3.09	4.62	2.96
7 Days	4.28	5.16	3.72

The graphical plot as shown in figure 4.5 and figure 4.6 below shows K-S values for different distribution at Shimla and Dharamshala region. Normal distribution gave minimum value of K-S for annual 1 day and 2 to 7 days consecutive maximum rainfall series at Shimla region whereas Gamma distribution gave minimum value of K-S for annual 1 day and 2 to 7 days consecutive maximum rainfall series at Dharamshala region. The statistical comparison by K-S test for goodness of fit clearly shows that Normal distribution and Gamma is the best fitting representative function for rainfall frequency analysis at Shimla and Dharamshala region.

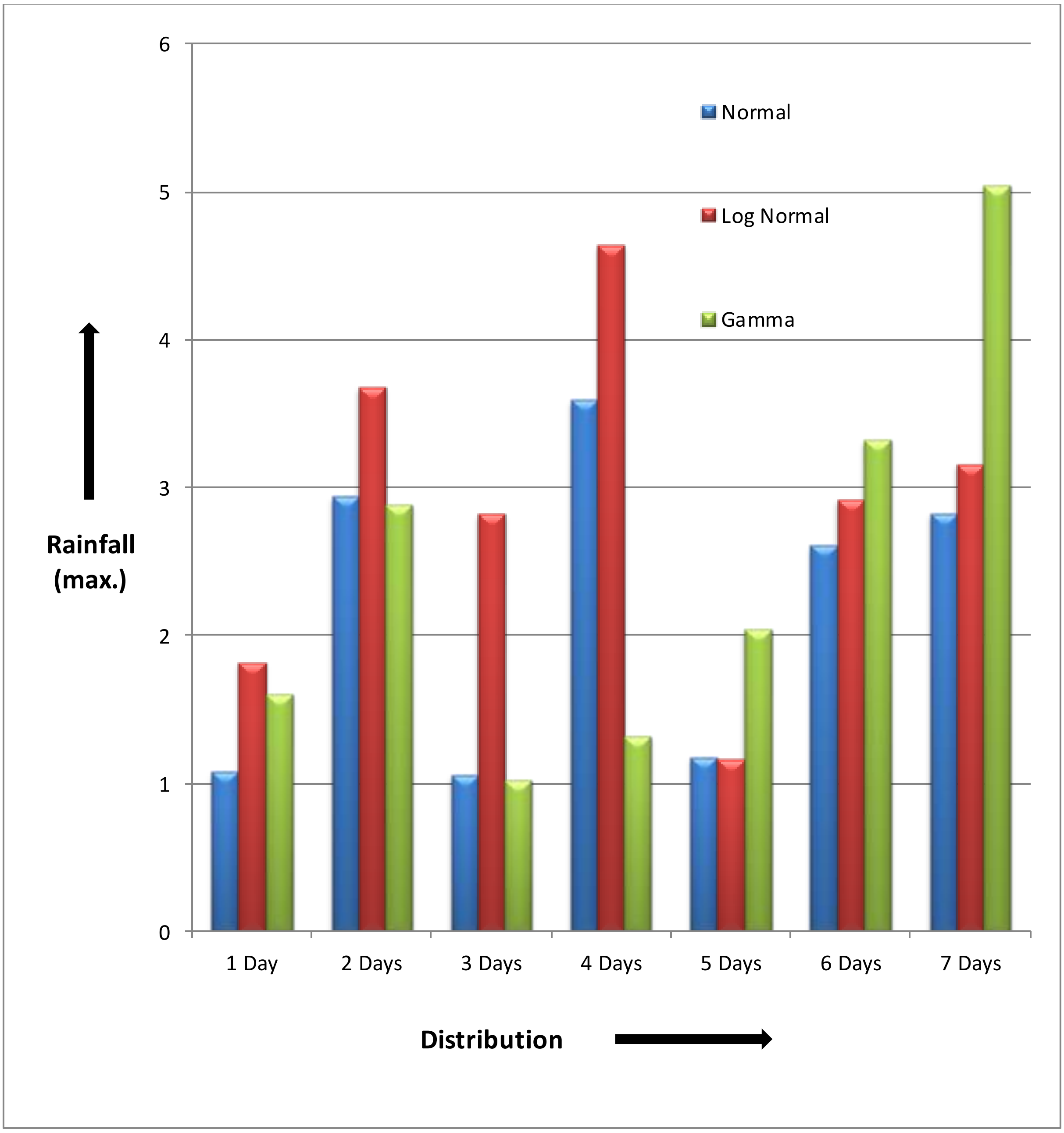

Fig. 4.5: Graphical plot showing K-S values for different distributions at Shimla

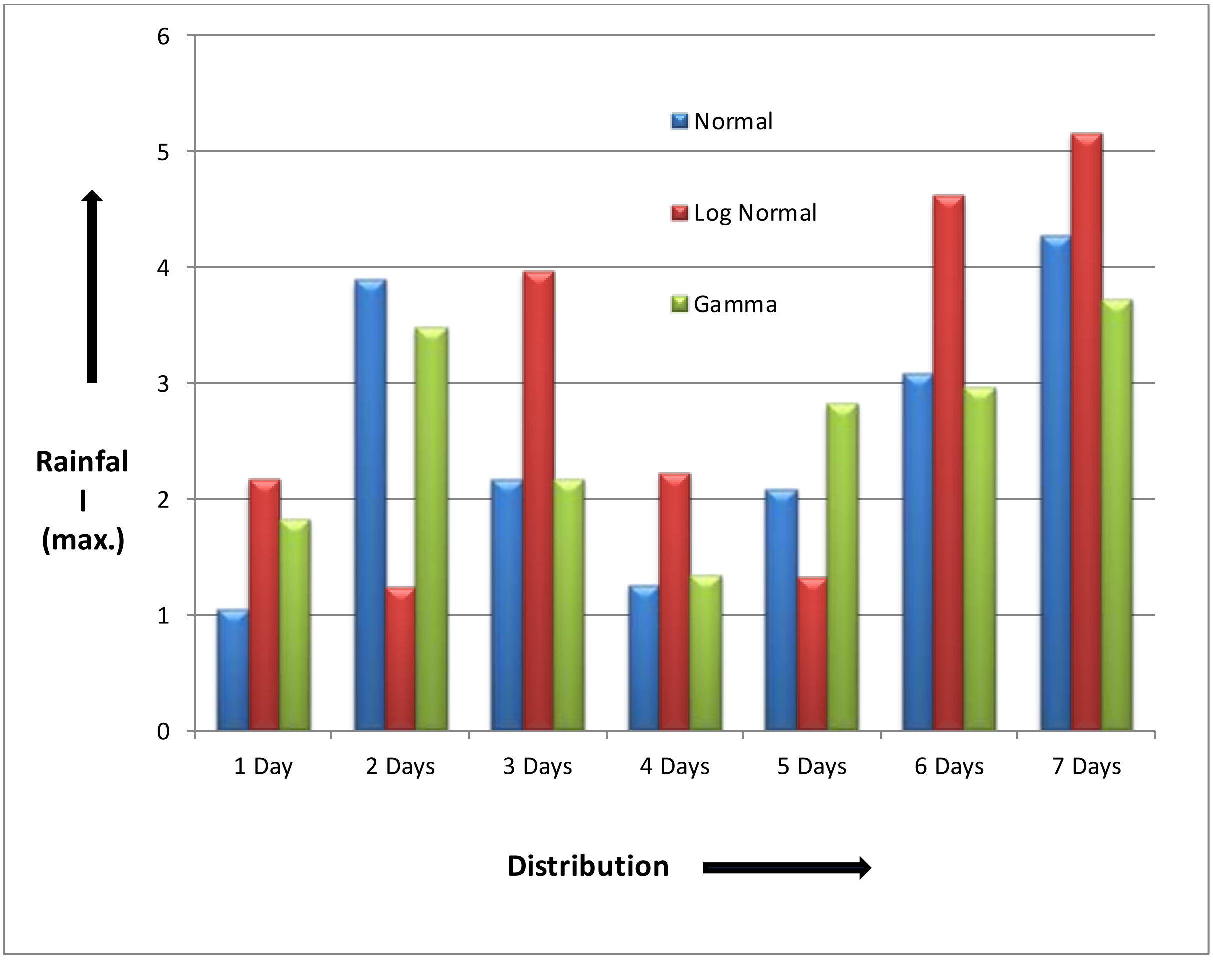

Fig. 4.6: Graphical plot showing K-S values for different distributions at Dharamshala

4.2.3 Anderson Darling test

The data presented in Table 4.9 and 4.10 reveals that 1 day, 2, 3, 4, 5, 6 and 7 consecutive days maximum rainfall followed Normal, Log Normal and Gamma distribution. The data presented in Table 4.9 and Table 4.10 revealed that the computed AD values for three probability distribution i.e. Normal, Log Normal and Gamma are found to be less than the critical value of AD at 95% confidence level for 1 day as well as consecutive days maximum rainfall series. Gamma distribution gave minimum value of AD for annual 1 day and 2 to 7 days consecutive maximum rainfall series at Shimla region whereas Log Normal distribution gave minimum value of AD for annual 1 day and 2 to 7 days consecutive maximum rainfall series at Dharamshala region. The statistical comparison by AD test for goodness of fit clearly shows that Gamma distribution and Log Normal is the best fitting representative function for rainfall frequency analysis at Shimla and Dharamshala region.

Table 4.9: AD values for different distributions at Shimla

Consecutive days	Normal	Log Normal	Gamma
1 Day	2.25	1.51	3.47
2 Days	3.96	2.61	1.69
3 Days	1.81	3.25	2.06
4 Days	1.46	2.11	1.39
5 Days	5.18	1.04	3.57
6 Days	4.26	3.98	2.81
7 Days	2.15	4.29	2.06

Table 4.10: AD values for different distributions at Dharamshala

Consecutive days	Normal	Log Normal	Gamma
1 Day	2.84	1.12	3.78
2 Days	3.67	2.43	2.92
3 Days	2.11	3.92	4.76
4 Days	1.42	1.78	1.47
5 Days	2.69	2.91	3.99
6 Days	4.78	3.24	4.89
7 Days	5.6	1.48	1.24

The graphical plot as shown in figure 4.7 and figure 4.8 below shows AD values for different distribution at Shimla and Dharamshala region. Gamma distribution gave minimum value of AD for annual 1 day and 2 to 7 days consecutive maximum rainfall series at Shimla region whereas Log Normal distribution gave minimum value of AD for annual 1 day and 2 to 7 days consecutive maximum rainfall series at Dharamshala region. The statistical comparison by AD test for goodness of fit clearly shows that Gamma distribution and Log Normal is the best fitting representative function for rainfall frequency analysis at Shimla and Dharamshala region.

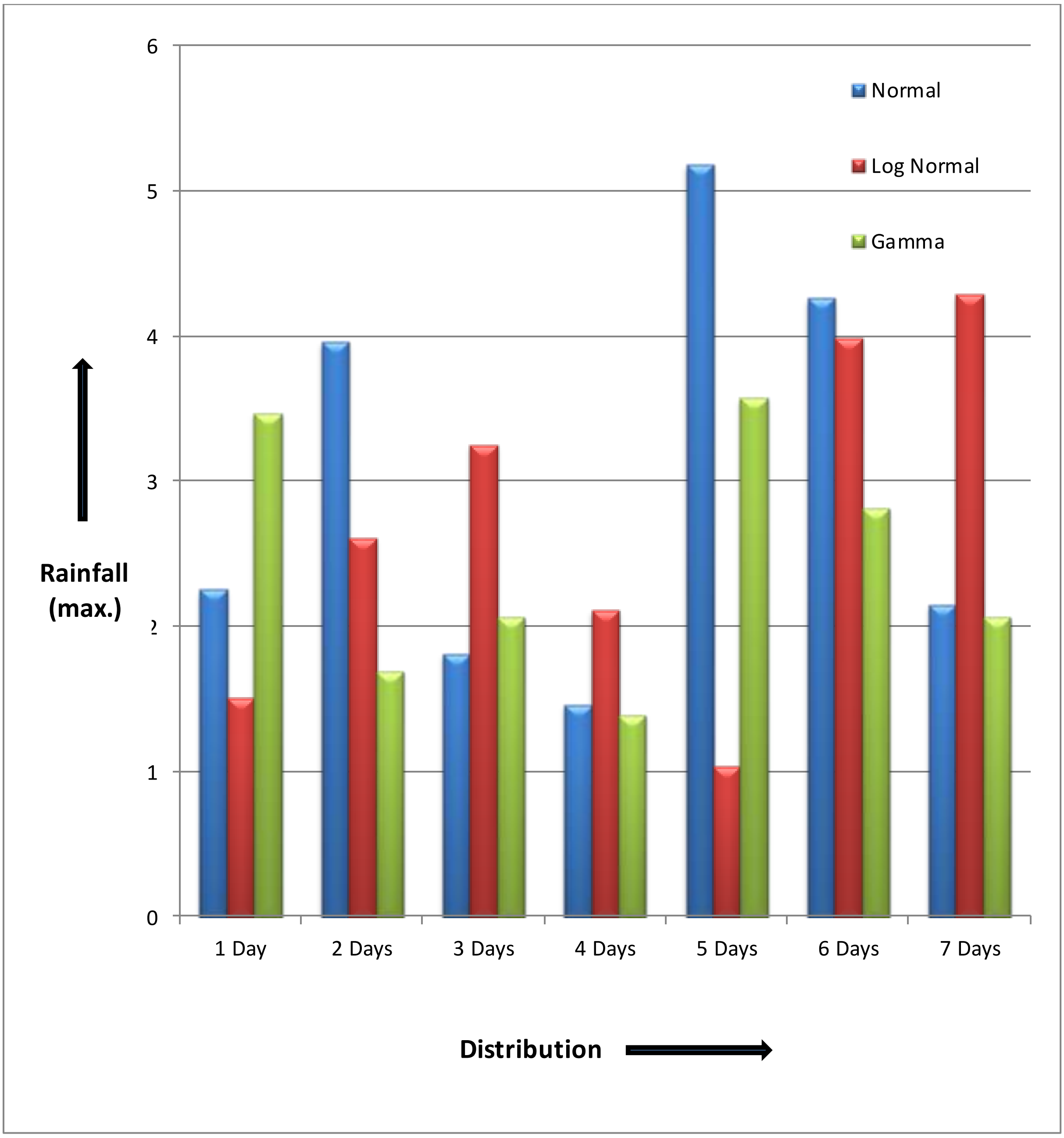

Fig. 4.7: Graphical plot showing AD values for different distributions at Shimla

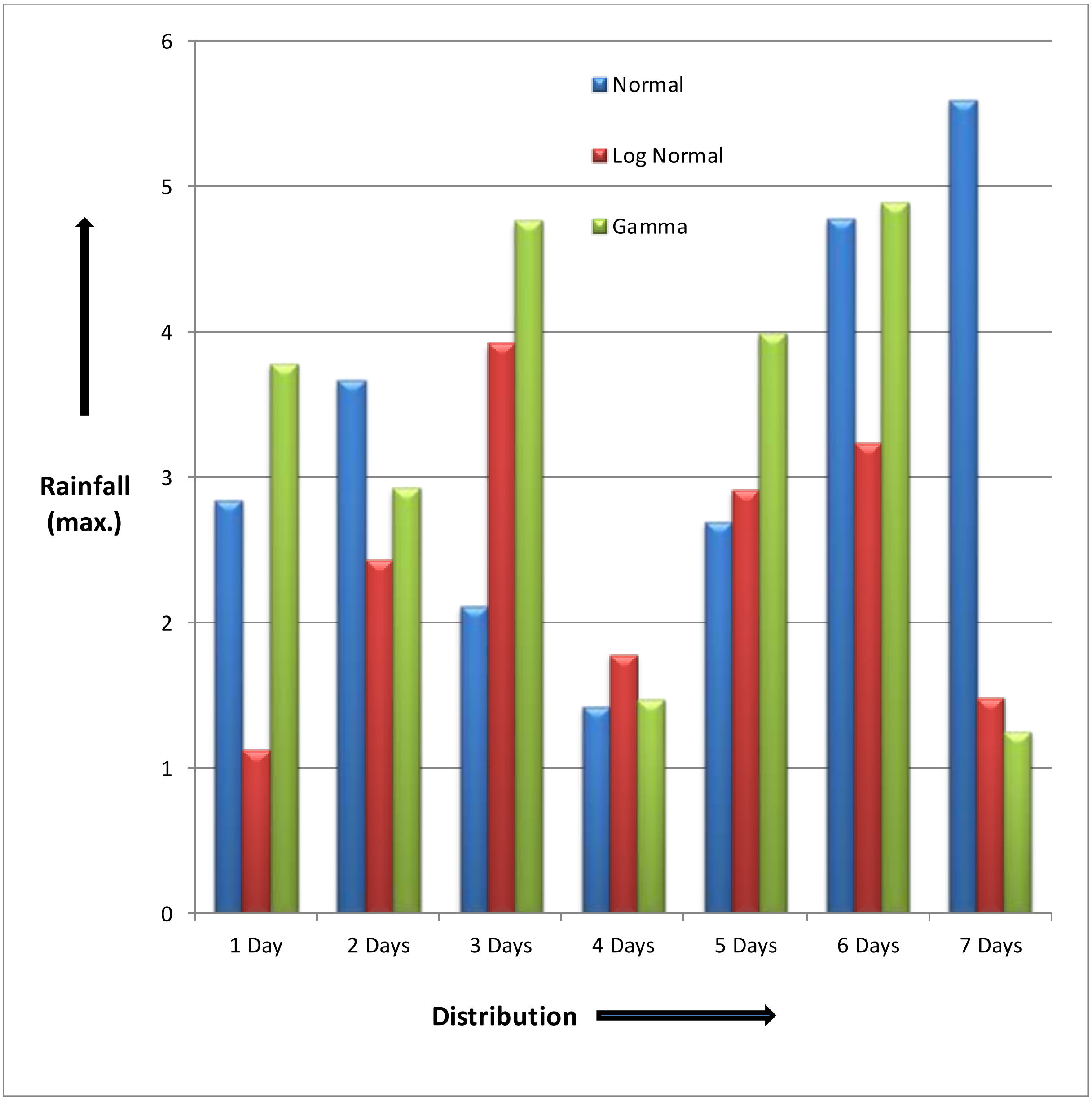

Fig. 4.8: Graphical plot showing AD values for different distributions at Dharamshala

CHAPTER 5

CONCLUSIONS

a) The main objective of present study was to do the rainfall analysis for Shimla and Dharamshala region. For this purpose the rainfall data of 20 years (1992-2012) were collected from Indian Meteorological Department, Shimla (H.P.), India. The log-normal distribution is the best fit probability distribution for one day annual maximum as well as two, three and four consecutive days, while for five and seven consecutive days, gamma distribution and for six days normal distribution fits better for the region. The study of frequency analysis will be used as a rough guide by engineers and hydrologists to prevent floods and droughts, and applied to planning and designing of water resources related to engineering such as reservoir design, flood control work. The results of this study would be useful for agricultural scientists, decision makers, policy planners and researchers in order to identify the areas where agricultural development and construction of drainage systems takes place.

b) The present study aims to evaluate the rainfall magnitude for different return periods and also to ascertain the type of probability distribution that best fits the rainfall data of Shimla and Dharamshala region (H.P.), India. From the above results, Gamma distribution is found to be best fit for the region. In order to achieve the objectives in present study, the following tests were performed.

c) On the basis of these tests, the results will be helpful in soil and water conservation planning and design of small and medium hydraulic structures such as small dams, bridges, culverts drainage works, etc. A common use of rainfall data is in the assessment of probabilities or return periods of given rainfall at a given location. Such data can then be used in assessing flood discharges of given return period through modeling or some empirical system and can thus be applied in schemes of flood alleviation or forecasting. Frequency analysis usually involves the fitting of a theoretical frequency distribution using a selected fitting method, although empirical graphical methods can also be applied. The fitting of a particular distribution implies that the rainfall samples of annual maxima were drawn. For each distribution one can obtain the following: Estimation of parameters of the distribution, a table of rainfalls of specified exceedance probabilities or return periods with confidence limits, results of goodness of fit tests, a graphical plot of the data fitted to the distribution.

d) For Shimla region, the mean value of one-day annual maximum rainfall is found to be 98.9 mm with standard deviation and coefficient of variation of 25.06 and 57.19 respectively. The coefficient of skewness is 1.541. For 2 to 7 days consecutive annual maximum rainfall range values for mean, standard deviation, coefficient of variation and coefficient of skewness are 144.2 - 216.9 mm , 38.581 - 46.131, 46.239 - 31.127 and -0.041 - 0.3089. It is observed that all distribution fitted function significantly. A maximum of 121mm in 1 day, 178.15 mm in 2 days, 203.6 mm in 3 days, 219.4 mm in 4 days, 230.23 mm in 5 days, 240.6 mm in 6 days and 257.5 mm in 7 days is expected to occur at Shimla, Himachal Pradesh every 2 years. . For a recurrence interval of 20 years, the maximum rainfall expected in 1 day, 2, 3, 4, 5, 6 and 7 days is 160.5 mm, 239.11 mm, 265.36 mm, 285.9 mm, 297.42 mm, 309.5 mm and 330.38 mm respectively.

e) For Dharamshala region, the mean value of one-day annual maximum rainfall is found to be 142.9 mm with standard deviation and coefficient of variation of 54.8 and 51.34 respectively. The coefficient of skewness is 1.1. For 2 to 7 days consecutive annual maximum rainfall range values for mean, standard deviation, coefficient of variation and coefficient of skewness are 201 – 393.4 mm , 70.17 – 146.5, 41.65 – 30.47 and 0.726 – 1.593. It is observed that all distribution fitted function significantly. A maximum of 191.1 mm in 1 day, 262.7 mm in 2 days, 314.27 mm in 3 days, 370.24 mm in 4 days, 419.7 mm in 5 days, 477.4 mm in 6 days and 522.32 mm in 7 days is expected to occur at Dharamshala, Himachal Pradesh every 2 years. For a recurrence interval of 20 years, the maximum rainfall expected in 1 day, 2, 3, 4, 5, 6 and 7 days is 277.7 mm, 373.6 mm, 445.11 mm, 518.62 mm, 589 mm, 680.3 mm and 753.79 mm respectively.

REFERENCES

Daily annual rainfall data (1992-2012) of Shimla and Dharamshala region, H.P., India, Indian Meteorological Department , Shimla.

Benson, M. A. (1968). "Uniform flood frequency estimating methods for federal agencies." Water Resources Research, 4 (5), 891-908.

RamaRao, B. V., Kavi, P. S. and Sridharan, P. C. (1975). "Study of rainy days and wet spells at Bijapur." Annual Arid Zone, 14(4), 371-372.

Baheerathan, V. R. and Shaw, E. M. (1978). "Rainfall Depth-Duration-Frequency Studies for Sri Lanka." Journal of Hydrology, 37, 223-239.

Kulandaivelu, R. (1984). "Probability analysis of rainfall and evolving cropping system for Coimbatore." Mausam, 5(3), 257-258.

Sharda, V. N. and Bhushan, L. S. (1985). "Probability analysis of annual maximum daily rainfall for Agra." Indian Journal of Soil Conservation, 13(1), 16-20.

Hirose, H. (1994). "Parameter Estimation in the Extreme-Value distributions using the Continuation Method." Information Processing Society of Japan, 35, 9.

Duan, J., Sikka, A. K. and Grant, G. E. (1995). "A comparison of stochastic models for generating daily precipitation at the H.J. Andrews Experiment Forest." Northwest Science, 69(4), 318-329.

Upadhaya, A. and Singh, S. R. (1998). "Estimation of consecutive day's maximum rainfall by various methods and their comparison." Indian Journal of Soil Conservation, 26(2), 1993-2001.

Mohanty, S., Marathe, R. A. (1999). "Probability analysis of annual maximum daily rainfall for Amravati." Indian Journal of Soil Conservation, 43(1), 15-17.

Rizvi, R. H., Singh, R., Yadav, R. S., Tewari, R. K., Dadhwal, K. S. and Solanki, K. R. (2001). "Probability analysis of annual maximum daily rainfall for Bundelkhand region of Uttar Pradesh." Indian Journal of Soil Conservation, 29(3), 259-262.

Tao, D. Q., Nguyen, V. T., and Bourque, A. (2002). "On selection of probability distributions for representing extreme precipitations in Southern Quebec." Annual Conference of the Canadian Society for Civil Engineering, 12(2), 1-8.

Fowler, H. and Kilsby, C. (2003). "A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000." International Journal of Climatology, 23, 1313-1334.

Lee, C. (2005). "Application of rainfall frequency analysis on studying rainfall distribution characteristics of ChiaNan plain area in Southern Taiwan." Journal of Crop, Environment & Bioinformatics, 29(2), 31-38.

Deidda, R. and Puliga, M. (2006). "Sensitivity of goodness-off fit statistics of rainfall data rounding off." Physics and Chemistry of the Earth, 31(1), 1240-1251.

Nadarajah, S. and Choi, D. (2007). "Maximum daily rainfall in South Korea." Journal of civil engineering, 35(4), 311–320.

Bhakar, S. R., Iqbal, M., and Bansal, A. K. (2008). "Probability analysis of rainfall at Kota." Indian Journal of Agricultural Research, 42(1), 201 -206.

Olofintoye, O. O, Sule, B. F. and Salami, A. W. (2009). "Best–fit Probability distribution model for peak daily rainfall of selected Cities in Nigeria." New York Science Journal, 2(3), 1554-2000.

Sharma, M. A. and Singh, J. B. (2010). "Use of Probability Distribution in Rainfall Analysis." New York Science Journal, 40-49.

Suribabu, C. R., Sudarsan, J. S. and Ramanan, S. R. (2015). "A study on consecutive day maximum rainfall for disaster mitigation in Tiruchirapalli City, India." Jordan Journal of Civil Engineering, 9(2), 197–202.

APPENDIX I

Final rainfall data set of 20 years of Shimla region

S.NO	1 DAY	2 DAYS	3 DAYS	4 DAYS	5 DAYS	6 DAYS	7 DAYS
1992	89.4	143.3	156.7	158.9	163.1	163.2	176
1993	87.7	138.9	168.9	179.5	193.5	224.4	254.4
1994	90.2	143.9	158.4	171.4	177.8	211.7	222.3
1995	95	158.1	207.5	214.2	243.3	252.6	283.3
1996	97	100.1	137.1	139.2	152.4	171	175
1997	73.7	118.1	128.8	136.5	143.5	149.1	152.8
1998	88.3	129.7	173.1	192.1	194.3	196.2	199.6
1999	75.5	87.2	133.7	153.8	160.2	161.4	165.3
2000	99.3	148.3	193.8	201.8	214.9	238.3	265.1
2001	95.1	113.8	127.6	128.7	142.2	152.8	154.6
2002	95.7	139.1	155.5	157.3	170.4	171	186.4
2003	112.5	146.2	168.6	242.7	252.3	253.1	253.5
2004	81.1	94	116.6	132.1	137	141.5	177
2005	84.4	151.1	174.5	181.9	190.1	209.3	232.7
2006	101	152.1	152.1	166.8	170.4	178.1	196.7
2007	122.2	212.9	221.1	225.4	229.7	234	250.5
2008	177.1	219.9	244.7	244.7	257.5	261.9	264.9
2009	96.8	168.7	204.4	227	238.8	238.8	238.8
2010	104.6	165.9	170.9	185.1	185.1	187.5	218.8
2011	44.7	71.7	96.2	103.3	112.4	119.5	120.9
2012	66.6	80.3	93.9	103.5	127.6	130.4	149

APPENDIX II

Final rainfall data set of 20 years of Dharamshala region

S.NO	1 DAY	2 DAYS	3 DAYS	4 DAYS	5 DAYS	6 DAYS	7 DAYS
1992	84.1	130.0	180.4	191.7	221.5	245.2	256.5
1993	94.7	104.8	140.9	163.3	218.8	256.1	277.5
1994	90.1	145.7	181.9	242.6	252.3	311.0	357.0
1995	88.8	134.9	167.0	185.2	203.0	233.6	240.9
1996	95.0	106.4	128.7	156.9	216.1	243.1	259.3
1997	91.6	175.8	218.9	235.4	283.0	297.5	300.5
1998	91.9	154.0	175.4	222.4	243.8	297.4	349.3
1999	96.6	141.2	158.1	196.8	212.9	218.4	241.2
2000	98.8	156.4	181.2	259.2	298.6	334.6	337.6
2001	96.6	141.2	158.1	196.8	212.9	218.4	241.2
2002	123.0	182.8	187.2	245.2	265.2	271.6	288.2
2003	279.6	298.6	351.8	381.6	397.6	397.6	403.2
2004	165.0	254.2	340.2	397.8	453.8	556.6	572.0
2005	178.8	218.9	259.3	362.1	396.3	419.9	419.9
2006	150.2	233.2	252.8	305.4	307.2	372.0	373.8
2007	145.6	168.0	182.0	202.2	224.6	226.4	295.0
2008	212.4	279.4	317.6	333.0	402.8	474.0	489.4
2009	97.8	125.0	207.4	238.8	274.6	306.0	312.8
2010	176.8	223.0	254.2	312.4	341.2	373.6	388.8
2011	213.4	334.4	397.2	447.0	509.8	542.6	693.8
2012	186.4	312.4	388.6	475.8	574.0	691.6	770.4