RAINFALL DATA GENERATION FOR SAMARU, ZARIA USING STATISTICAL PARAMETERS

BY

JAMIU MUSA

DEPARTMENT OF AGRICULTURAL ENGINEERING,
FACULTY OF ENGINEERING,
AHMADU BELLO UNIVERSITY,
ZARIA, NIGERIA

SEPTEMBER, 2015

RAINFALL DATA GENERATION IN SAMARU, ZARIA USING STATISTICAL PARAMETERS

BY

Jamiu MUSA, B.ENG (ABU) 2008

M.Sc./ENG/13746/2010-2011

A DISSERTATION SUBMITTED TO THE SCHOOL OF POSTGRADUATE STUDIES, AHMADU BELLO UNIVERSITY, ZARIA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF SCIENCE DEGREE IN AGRICULTURAL ENGINEERING (SOIL AND WATER OPTION)

DEPARTMENT OF AGRICULTURAL ENGINEERING,

FACULTY OF ENGINEERING,

AHMADU BELLO UNIVERSITY, ZARIA, NIGERIA

SEPTEMBER, 2015

DECLARATION

I declare that to the best of my knowledge that the work in this Dissertation entitled "**RAINFALL DATA GENERATION IN SAMARU, ZARIA USING STATISTICAL PARAMETERS ''** has been carried out by me in the Department of Agricultural Engineering under the supervision of Professor O.J. Mudiare and Professor A.A. Ramalan..All the information derived from the literature has been duly acknowledged in the text and a list of references provided. No part of this dissertation has been previously presented for another degree at this or any other University/Institution.

Jamiu, MUSA _____ _____

M.Sc./ENG/13746/2010-2011 Signature Date

CERTIFICATION

This dissertation entitled **"RAINFALL DATA GENERATION IN SAMARU,ZARIA USING STATISTICAL PARAMETERS "** by Jamiu, MUSA meets the regulations governing the award of degree of Master of Science (Agricultural Engineering) of the Ahmadu Bello University, Zaria and is approved for its contribution to scientific knowledge and literary presentation.

Prof. O. J. Mudiare Chairman, Supervisory Committee	_____ Signature	_____ Date
Prof. . A. A. Ramalan Member, Supervisory Committee	_____ Signature	_____ Date
Prof. M. L. Sulaiman Head of Department	_____ Signature	_____ Date
Prof. Kabir Bala Dean, School of Postgraduate studies	_____ Signature	_____ Date

DEDICATION

This work is dedicated to my wonderful late parents, my late brother, Mallam Tiamiyu Musa,to my wife Hajia Risiqat and to my children, Summayah and Mubaraq for their sacrifice, love, encouragement, spiritual and financial support throughout my years of study.

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to the Almighty Allah for his protection and provision for me throughout my course of study. I am also grateful to my late parents, Mr and Mrs Musa for all they did for me since birth. I would like to appreciate and give special gratitude to my major supervisor, Professor O.J.Mudiare for his supervision, advice and guidance throughout the period of this work. I am indebted to him more than it can be verbally expressed. God will bless you and grant all your heart desires

It is worthwhile to express my since gratitude to my Project co-supervisor, Professor A.A. Ramalan for his advice, supervision, crucial roles and contributions.I pray that God will reward you abundantly.

I will forever be grateful to all staff of the Department of Agricultural Engineering starting from the Head of Department, Prof. M.L.Suleman, Prof. D.D. Yusuf, Prof. M.A Oyebode, Dr.U.S Muhammed, Prof. H.E.Igbadun, Dr. M.A.Gwarzo, Prof.A.El-Okene Dr. M.Isiaka, Prof.S.Z Abubakar and Dr. M.K.Othman for their efforts to bring this work to reality. My sincere gratitude also goes to all my colleagues in the Department namely Engr. M. Abdul-Salam, Engr.Habib, Engr. A. Afolabi and others for their concern throughout the period of this work. May God bless you all.

I am also grateful to my wife and children, for their love, support and patience. May Allah bless you all abundantly.I am equally indebted to my colleagues in Leventis Agric. Training School,Yaba,Abuja in person of former and present acting Principal, Mr.M.Awolaja and Mrs Agnes Yusuf together with Mr. O.C.Igheckpe,Head of Engineering Department(Leventis Agric.Training School)

I am equally thankful to all my friends, Francis, Nda Abdullahi and Stephen of Institute for Agricultural Research (IAR) metereology unit for all their efforts in making this research a reality.Thanks to you all.

ABSTRACT

Estimation of rainfall for a desired return period is one of the pre-requisites for any design purpose at a particular site, which can be achieved by probabilistic approach. This study is aimed at using the statistical parameters of long-term observed rainfall data to generate monthly rainfall depth and estimate rainfall values at different return periods for hydrological and agricultural planning purposes. The daily rainfall data of 60 years from 1953 to 2012 were collected from the meteorological station situated at the Institute for Agricultural Research (IAR), Ahmadu Bello University (ABU) Zaria. In this work, the expected number of observations was 720.However, a sample size of 431 being the months with rainfall amounts was used for the study. Three probability distributions functions viz: normal, log normal and exponential distribution were used for rainfall generation. The data was processed to estimate the statistical parameters of the distributions for Probability Density Functions (PDFs) selected. The two types of parameters estimated using the methods of maximum likelihood via the statistical package called EasyFitXL5.6 are the scale and location parameters. The parameters' estimates were further used in various PDFs equations to generate new sets of rainfall amounts. Five statistical goodness of fit test were used in order to select the best fit probability distribution and are; Anderson-Darling, Chi-square and Kolmogorov-Smirnov, ANOVA and Student t test. From the analysis of the data, it was discovered that the normal PDF predicts correctly the monthly amount of rainfall and specific return periods rainfall values followed by the log-normal and exponential PDF, the least predictor. After carefully observing and testing the three PDFs, it was discovered that the normal PDF estimated closely the monthly amounts of rainfall compared with the log-normal and exponential PDFs when considered generally for monthly and return periods rainfall values estimation. Nevertheless, the normal PDFs have been found to predict well the daily rainfall amounts in Samaru. This work should be made available to local and immediate environments to Samaru through local and regional experts on climatic information dissemination for use in planning and management of specific rainfed crop(s) production.

TABLE OF CONTENTS

Cover Page…….. Error! Boo

Title Page……..**ii**

DECLARATION……..**iii**

CERTIFICATION……**iv**

DEDICATION…… **v**

ACKNOWLEDGEMENTS…… **vi**

ABSTRACT…… **vii**

TABLE OF CONTENTS…… **viii**

LIST OF TABLES……... **xii**

LIST OF FIGURES……..**xiv**

LIST OF APPENDICES…….. **xv**

ABBREVIATIONS, UNITS AND SYMBOLS…….. **xvi**

CHAPTER ONE: INTRODUCTION

1.1 Background of the Study…… **1**

1.2 Statement of Problems……. **4**

1.3.Justification of the Study……. **6**

1.4.Objectives of the Study…… **7**

CHAPTER TWO: LITERATURE REVIEW

2.2 Periodicity in Hydrology……..……**9**

2.3 Statistical Terms……...……**9**

2.4 Frequency Analysis of Rainfall……..**11**

2.4.1 Random Variable……**11**

2.4.2 Probability relationship…….…….**.....13**

2.5 Return Period……..…….**14**

2.6 Climate Data……...17

2.6.1 Uncertainty in model parameters.....18

2.6.2 Rainfall and climate data under climate change scenario.....18

2.7 Rainfall Amount Models 19

2.7.1 Normal (Gaussian) distribution 20

2.7.2 Exponential distributions.....21

2.7.3 Probability Log-normal distribution.....22

2.8 Goodness of Fit Tests 23

2.8.1 Kolmogorov-Smirnov Test.....24

2.8.2 Anderson-Darling Test 24

2.8.3 Chi-Squared test.....25

2.8.4 ANOVA Test 27

2.8.5 Paired t-Test 28

2.8.6 Expert Modeler 29

CHAPTER THREE: MATERIALS AND METHODS

3.1 Study Location 30

3.2 Data Assembly 30

3.2.1 Preliminary Data Analysis Method 30

3.2.2 Formulation of Hypothesis 31

3.3 Fitting of Probability Distributions 32

3.4 Rainfall Prediction Strategy **33**

3.5 Statistical Analysis Framework 33

CHAPTER FOUR: RESULTS AND DISCUSSION

4.1 Preliminary Analysis 35

4.2 Analysis of Statistical Parameters 40

4.2.1 Chi-Square Test 42

4.2.2 Kolmogorov-Smirnov and Anderson-Darling Tests Statistic **42**

4.3 Comparison of Observed Monthly Rainfall Amounts with Generated **43**

4.3.1 Testing the Normal Model with the Observed Value **44**

4.3.2 Testing the Log-normal Model with the Observed Value 46

4.3.3 Testing the Exponential Model with the Observed Value 47

4.4 Analysis of Trend 48

4.5 Monthly Rainfall Generation based on the Selected Models 49

4.5 .1 Monthly Rainfall generation for different Return Periods.....52

CHAPTER FIVE: SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1 Summary……**63**

5.2 Conclusion…….65

5.3 Recommendations……...……..66

REFERENCE……**...67**

APPENDICES……**70**

LIST OF TABLES

Table 4.1**:** Descriptive statistics of observed rainfall in millimeters across periods of five years from 1953-2012 in Samaru - - - - - 37

Table 4.2: Descriptive Statistics of observed monthly Rainfall from 1953-2012 -38

Table 4.3: Descriptive statistics for Predicted rainfall Values - - -39

Table 4.4: Log-normal Descriptive Statistics of observed Monthly Rainfall from 1953-2012 - - - - - - - -39

Table 4.5: Chi-square test of monthly Observed and Predicted rainfall values -42

Table 4.6 : Kolmogorov-Smirnov and Anderson-Darling Goodness of Fit Test - -43

Table 4.7: Model Statistics for Normal PDF - - - - - -45

Table 4.7a: Model Statistics for Log-normal PDF - - - - -46

Table 4.7b: Model Statistics for Exponential PDF - - - - - -47

Table 4.8: Generated Probabilities for 20 years - - - - -50

Table 4.9: Computed Z-values based on Generated probabilities - - -51

Table 4.10: Predicted Normal and Log-normal Mean and SD (mm) - - -52

Table 4.11: Rainfall amounts in (mm) for the Month of May - - -53

Table 4.12: ANOVA of Return Period for the month of May - - -55

Table 4.13: Rainfall amounts in (mm) for the Month of June - - -55

Table 4.14: ANOVA of Return Period for the month of May - - -57

Table 4.15: Turkey HSD Multiple Comparisons Post-HOC test for the month of June - - - - - - - -58

Table 4.16: t-Test results on rainfall values obtained at different return periods -62

LIST OF FIGURES

Figure 2.1: population and sample - - - - - - -10

Figure 2.2: Randomization of data - - - - - - - - -12

Figure 4.1: Trend of the percentage of rainfall occurrence across period of ten years from 1953-2012 in Samaru - - - - - - 36

Figure 4.2: Comparing mean monthly values of observed and predicted (mm) -41

Figure 4.3: Sequence chart for Observed Value - - - - -44

Figure 4.4: Fit Pattern of the Normal and Observed Value - -- - -45

Figure 4.5: Fit Pattern of the Log-normal and Observed Value - - -46

Figure 4.6: Fit Pattern of the Exponential and Observed Value - - -48

Figure 4.7: Time series plot of rainfall amount across periods of five years from 1953-2012 in Samaru - - - - - - -49

Figure 4.8: PDF Chart of return Periods for the Month of May - - -54

Figure 4.9: PDF Chart of return Periods for the Month of June - - -56

Figure 4.10: PDF Chart of return Periods for the Month of July - - -59

Figure 4.11: PDF Chart of return Periods for the Month of August - - -60

Figure 4.12: PDF Chart of return Periods for the Month of September - -61

LIST OF APPENDICES

Appendix A1: Generated probabilities based on random numbers for the 60 years between May and September - - - - - - 70

Appendix A2 : Computed Z-values based on Generated probabilities - - 73

Appendix A3: Predicted 60 Years rainfall values (mm) for the Months of May-September - - - - - - - - 76

Appendix B1: Monthly rainfall predicted values for Normal PDF for 20 years (mm) 79

Appendix B2: Monthly rainfall predicted values for Log-Normal PDF for 20 years (mm) 80

Appendix B3: Monthly rainfall predicted values for Exponential PDF for 20 years (mm) 81

Appendix C: Statistical Tests Results for Return Periods - - - - 82

ABBREVIATIONS, UNITS AND SYMBOLS

ANOVA	-	Analysis of variance
SPSS	-	Statistical Package for Social Science
PDF	-	Probability Density Function
RN	-	Random Number
p-value	-	probability value
SD	-	Standard deviation
Df	-	degree of freedom
Sig	-	significant
std.error	-	standard error
IAR	-	Institute for Agricultural Research
Mm	-	millimeter
%	-	Percentage
Km	-	kilometer
^{o}C	-	degree centigrade
viz-	-	namely
vs	-	versus
i.e	-	that is
>	-	greater than
<	-	less than
Fig	-	figure

CHAPTER ONE

INTRODUCTION

1.1. Background of the Study

Synthetically generated daily or monthly rainfall data are frequently necessary in data scarce environments as input into the planning and design of water resources and soil conservation projects; simulation studies of crop growth and yield; farming systems and field farm operations scheduling (Jamaludin and Jemain 2007). Rainfall is the principal phenomenon driving many hydrological extremes such as floods, droughts, landslides, debris and mud-flows; its analysis and modelling are typical problems in applied hydrometeorology. Rainfall exhibits a strong variability in time and space. Hence, its stochastic modeling is not an easy task (De Michele and Bernardara, 2005).

A good understanding of the pattern and distribution of rainfall is important for water resource management of an area. Knowledge of rainfall characteristics, its temporal and spatial distribution play a major role in the design and operation of agricultural systems, telecommunications, runoff control, erosion control, as well as water quality systems. Generated weather is needed to supplement existing weather data, provide alternative weather realizations for a particular historical record, or identify possible weather sequences for a seasonal climate forecast (Walpole and Mayers, 1989).

The amount and pattern of rainfall are among the most important weather characteristics which affect agriculture profoundly. In addition to their direct effects on water balance in soil, they are strongly related to other weather variables such as solar radiation, temperature, humidity and evaporation which are also important factors affecting the growth and development of crops, pests, diseases and weeds. However, rainfall data form an essential input into many climatologic studies for agriculture, wherein considerable research is focused on rainfall analysis and modelling (Nnaji, 2001). This work is concerned with the modeling of the distribution of monthly rainfall amounts in Samaru, Zaria Kaduna state. Evaluating a range of scenarios that accurately reflect precipitation variability is critical in hydrological and water resource applications. Inputs to these applications can be provided using location- and interval-specific probability distributions. These distributions make it possible to estimate the likelihood of rainfall within a specified range. In order to improve the ability of African decision makers to prepare for and deal with the consequences of precipitation anomalies, it is important to provide them with a more complete understanding of the range and likelihood of rainfall totals a location could possibly receive.

Models of rainfall probability distributions over various timescales are useful tools for gaining this kind of understanding. Modelling rainfall variability in Africa presents an imposing problem for many reasons, including the need to summarize rainfall data for many years at many sites and the difficulty in finding a single method to represent such a variety of rainfall regimes. Rainfall regimes across the vast African continent differ widely

in terms of total accumulations, seasonal timing, and amounts of variability. Any method that is applicable across this wide range of conditions has to be quite flexible.

The typical approach to gaining a better understanding of the spatial and temporal variability in precipitation starts with the acquisition of historical rainfall data. These historical data provide necessary information about accumulation amounts in both time and space for the region and form the basis for fitting and testing distribution models. When historical data is unavailable in a region, or available data is inaccurate or incomplete in a spatial or temporal sense, geophysical models can be used to 'fill in' the missing values. These geophysical models are based on available data at other locations and times, as well as additional variables that add information to the model. Assuming the historical values, recorded or modelled – exist and are accepted as reasonably accurate, it is possible to fit parametric statistical distribution models to rainfall histories at individual locations of interest. Using a parametric distribution model allows for a more stable and extensive analysis of the rainfall probabilities than would be available using the raw data directly (Saleh,2011).The resulting distributions describe the estimated probability of different amounts of rain at a location for a select time interval such as annual, seasonal or monthly based on the historical values for that interval at that location.

There are many probability distributions that could be successfully utilized to parameterize rainfall distributions. The critical component for these distributions is that they be flexible enough to represent a variety of rainfall regimes. From a practical point of view, there is little difference between many of the commonly used distributions when estimating parameters based on a limited number of points, as is the case in much of the developing

world. Out of these available distributions, the normal, log-normal and exponential distributions are more widely understood, making them a good choice for implementation of this work (Rafa and Khanbilvardi,2014).Once the parameters of the distribution have been estimated, they can be used to describe rainfall regimes and be used in a variety of applications.

The distribution parameters might complement or even replace such common measures as the median, variance, minimum, maximum and quartile values as descriptors of the rainfall at any location. For locating potential hazard hotspots, distribution parameters may be used to identify areas with a disposition towards certain precipitation related hazards such as drought, flood, outbreak of disease, or reliability in providing adequate water for rain-fed agriculture. Monitoring of rainfall conditions may use distribution parameters as the foundation for the standardized precipitation index.

Combining distributions with probabilistic forecasts may result in a quantitative estimation of seasonal rainfall accumulations. More specifically, probability distribution parameters are estimated from monthly model derived historical rainfall values with a spatial resolution compatible with current agro-climatic models. One of the ways by which the hydrological parameters may be analyzed is through the use of different probability distribution models to estimate rainfall amounts for data-scarce areas/environments for agricultural planning.

1.2 Statement of Problems

Lack of extensive rainfall data for effective agricultural planning and hydrological structures construction. It is a known fact in Nigeria that most local areas or environments

are without recorded rainfall data or are data-scarce areas. Proper recorded rainfalls are often scarce to come by in those areas. Therefore in such areas, it is imperative to estimate rainfall data for agricultural planning purposes such as crop water requirement determination, irrigation requirement of crops, effective rainfall determination and other hydrological analyses.

Lack of quantitative forecast over small spatial scale or areas by the agricultural and forestry sector for effective planning of field operations, management of different agronomic and horticultural practices, irrigation and pest and disease management practices to avoid the weather risks during the crop growing season. With increasing demand for more food by millions of Nigerians, it has become highly necessary that all relevant hydrological parameters that directly affect food crop and animal production, water harvesting system, rate of ground water recharge and yields of crops should be collected and analyzed in such a way to make agricultural planning easier.

Variability in extreme events' dynamics for hydrological structures construction. In most cases, the return periods of interest exceed usually the periods of available records and could not be extracted directly from the recorded data. The design and construction of certain projects, such as dams and urban drainage systems, the management of water resources, and the prevention of flood damage require an adequate knowledge of extreme events of high return periods.

1.3. Justification of the Study

The result of this study should provide and Identify the appropriate model/s for monthly rainfall amount for Samaru and its environ as almost all of the climate variables are dependent on rainfall events (Deni and Jemain, 2009). Improved accuracy in seasonal rainfall calculation will help farmers to plan well and take effective approach in their agricultural activities. Also, it will facilitate the government and other relevant stakeholders in planning and effectively implementing adaptation strategies in different sectors, particularly agriculture, water and energy to lessen the impacts of drought.

The study should be a tool and useful guides for the designers, planners and decision maker's to prepare for and deal with the consequences of precipitation anomalies since models of rainfall probability distributions over various timescales can provide useful information. Agricultural and hydrological parameters have been observed to affect the rate of runoff generation, soil erosion, and the amount of water infiltration for crop production, ground water recharge and the yield of crop during the growing season (Edoga, 2007).

The study should give information on certain aspects of rainfall such as, the expected amount for a particular period which are valuable to agriculturists. Recent advancement in statistical methods particularly in the application of generalized linear models has much improved the range of techniques available for analysis of data that are not normally distributed.

1.4. Objectives of the Study

The general objective of this study is to generate rainfall amounts based on the statistical parameters obtained from long time recorded rainfall data for Samaru and its environs (within a distance of 50 km radius away from Samaru) for agricultural planning purposes.

The specific objectives include:

1. To determine the statistical parameters of mean and standard deviation of the observed rainfall data.
2. To calculate monthly rainfall amounts using three selected PDFs and compare with the observed values.
3. To calculate monthly rainfall amounts for selected return periods and compare with those determined from the observed data.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

Over large parts of West Africa, Nigeria inclusive, the crucial problem is scarcity or incomplete precipitation (rainfall) data. This poses a major obstacle to the proper planning and development of agriculture in these regions: planning as to the types of crops to produce amount of effective rainfall, irrigation requirement of crops and calculation of crops' consumptive use amongst others. Crops are usually planted during the rainy season, implying that crop productivity is largely a function of rainfall. Irregularities in rainfall reliability and spread have therefore contributed significantly to the poor yields and high variability in production from year to year (Mortimore, 1989).

The development of a frequency model starts with the historical data acquisition. These data provide the necessary information about the accumulated rainfall in different locations and describe the time and spatial variability of these rainfall values. When the parameters of the best fitted frequency distribution are estimated and the model is verified, then the model results could be considered dependable and can be used for variety of applications in water resources systems planning and design.

2.2 Periodicity in Hydrology

Olaniran, (2002) found that there has been a progressive early retreat of rainfall over the whole country (Nigeria) spanning up to half a century now and consistent with these pattern, there have also been a significant decline of rainfall frequency. The combined effect of these declines has been found to lead to a significant decrease in annual rainy days and amounts.

The periodicity in hydrological phenomenon implies that there is a cyclical occurrence in probabilities of an event (rainy season). South Africa is in one of the regions where a periodicity of 18 to 20 years was detected; being composed of about 9 to 10 consecutive rainy years followed by similar number of dry ones (Goldreich, 1995).

2.3 Statistical Terms

Descriptive Statistics

The statistics deal with the gathering of information and organizing it, calculating summary numbers and drawing graphs.

Probability

Probability is the use of past experience to give an idea of what to expect or the process of determining the likelihoods of occurrence of a certain random variable.

Inferential Statistics

Making a decision based on which outcome is most likely to consider for analysis or drawing a conclusion based on the data.

The group of interest is called the population. The characteristic of the population that we are interested in is called the population parameter. The same characteristic in a sample is called a statistic.

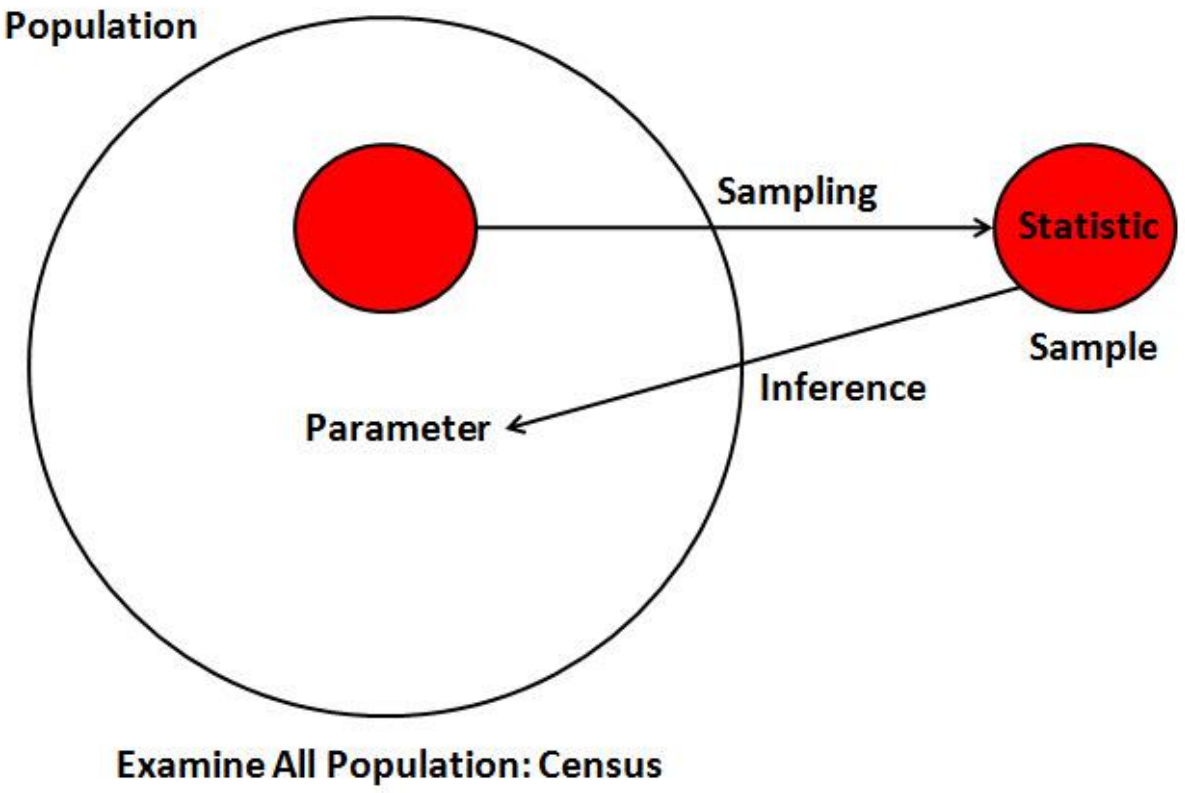


Fig 2.1 population and sample

Source: STAT301/303(Texas, A&M University)

Some samples can be biased, that is they favour one side or do not represent the entire population. In order to avoid this, statisticians use randomization. The simplest method is the simple random sample.

2.4 Frequency Analysis of Rainfall

Hydrologic phenomena are characterized by great variability, randomness and uncertainty. Precipitation, stream- flow and other quantities of importance in water resources engineering must be treated as random variables, with associated measures of frequencies that represent the likelihood percentage of time or probability. Frequency analysis methods are essential to hydrologic design and assessment.

It has been found that 80% of meteorological stations with long records receive rainfall according to the normal distribution (Kihupi *et al.,* 2002). This function is therefore most commonly used to describe variation in rainfall. One condition though is that, the period of observation is sufficiently long, that is, 30years and above (Stern and Coe, 1984).

2.4.1 Random Variable

A random variable is a quantity that depends on chance. The value or range of values can be predicted only with an associated probability, not with certainty. A random variable may be measured either as a discrete integer values (0, 1, 2 …) as in binominal distribution or over a continuous scale as in Normal, log-Normal, Gumbel distributions amongst others. A frequency relationship or probability distribution function represents the likelihood of occurrence of values of a random variable. Frequency relationships are developed based on

observed and/or generated data. In order to use random samples to help predict/determine what is going on in the population, we consider probability of occurrence; that is the proportion of times something will occur in the long run.

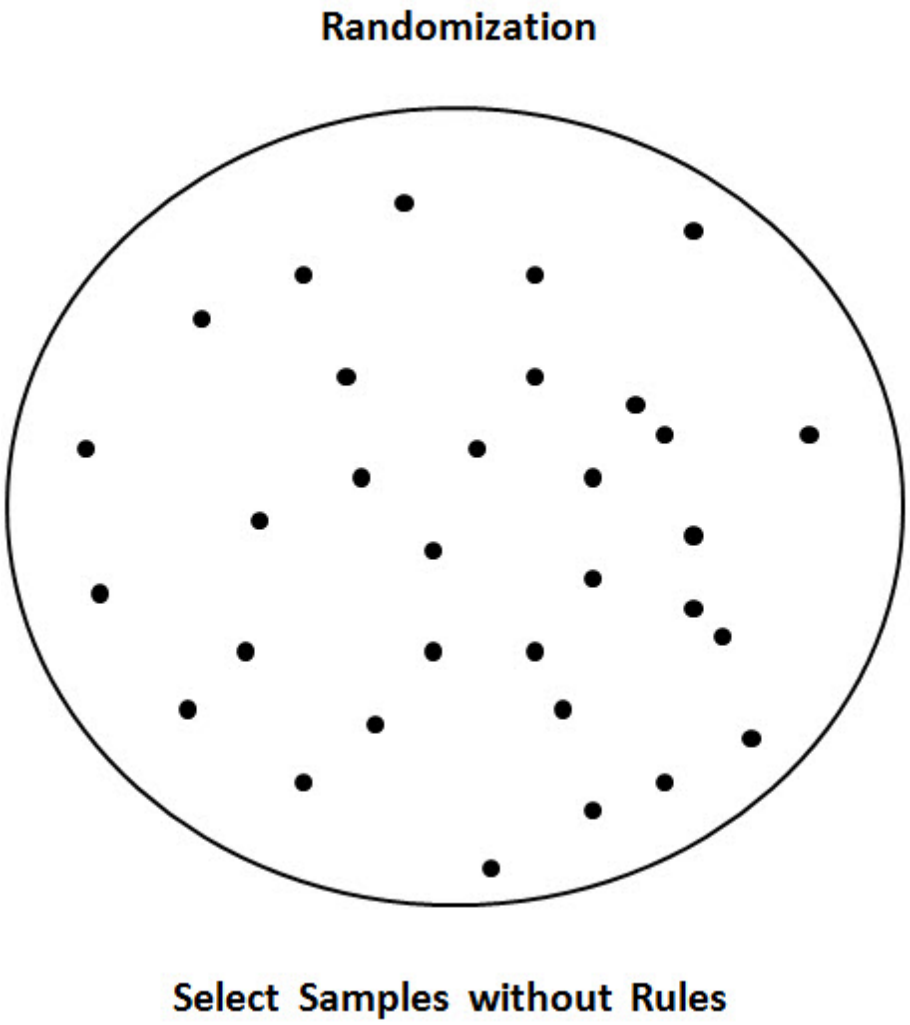


Fig 2.2: Randomization of data

Source: STAT301/303(Texas, A&M University)

Sample versus Population: Samples will vary, where the population is the 'whole truth'. Looking at multiple samples and seeing how they are different, that is, how they vary, helps

us determine how often one particular sample will occur. By studying sampling variability, we get a better idea of what the population looks like. For example, the coins toss: when the coin is tossed 10 times and the number of heads is counted, each of these 'experiments' will give a different number of heads, but most will be around 5.It will be very rare to get 0 or 10 but if the distribution of all of the experiments' outcomes is closely observed, it would be centered at 5, the true population value. So, samples can be used to determine or estimate probabilities when the true (population) probabilities are unknown. The population distribution is often unknown but can be estimated from the probabilities using sample proportions.

2.4.2 Probability relationship

The cumulative probability F(x) and probability p(x) for a random variable X are complements of each other defined as:

$$F(x) = \text{prob}\,(X < x) = 1\text{-}p(x) \tag{2.1}$$

$$P = p(x) = \text{prob}\,(X > x) = 1\text{-}F(x) \tag{2.2}$$

For continuous random variables,

$$\text{Prob}\,(X = x) = 0 \tag{2.3}$$

A distribution function provides a probabilistic model of the phenomena represented by a particular random variable. Model parameters are computed from sample observations. The parameters of mean,$\bar{X}$, standard deviation, S and skewness coefficient, G are computed from n observations X_i with the following formulae:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2.4}$$

$$S = \left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\right)^{0.5} \tag{2.5}$$

$$G = \frac{n\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^3}{(n-1)(n-2)s^3} \tag{2.6}$$

Probability distributions for continuous random variables are typically derived by statisticians and mathematicians as a probability density function that is integrated between limits to obtain the probability associated with ranges of values of the random variable.

2.5 Return period

Return period is the average time interval between occurrences of a hydrological event of a given magnitude, usually expressed in years. For example, a 100-year flood will occur on average once in every 100years.It is the inverse of probability of an event. Therefore, a 100-year flood has 1% probability to occur in each year. Upadhaya and Singh (1998) stated that it is possible to predict rainfall amount fairly accurately using various probability distributions for certain return periods although the rainfall varies with space, time and have erratic nature.

In hydrology, the term “return period” is used to define the risk of a particular design of a hydraulic structure. There are two different terms of return period reported in literature, first

arrival time and inter-arrival time or recurrence interval. These terms give different values when the events are dependent in time. On the other hand, for single and independent events, the first arrival time and recurrence interval give the same value (Fernandez and Salas 1999a). Most of the completed studies use first order Markov Chain to describe the events as serially dependent. Extensive theories and applications on the return period definitions and serial dependence are discussed in Fernandez and Salas (1999a and 1999b).

Woodyer et al. (1972), Kite (1978), Lloyd (1970), Loaiciga and Mariño (1991) and Şen (1999) define recurrence interval as the average elapsed time between the occurrences of critical events, such as earthquake of high magnitude and extreme floods or drought. The calculation of return period by these authors is for any event that has a value of equal or greater to the critical event.

Lloyd (1970) compared the return periods for river flows using two different conditions, (1) mutually independent, and (2) dependent time series, which was serially correlated using the first order Markov Chain model. His research found that the same 35 return periods were calculated for the independent and dependent time series. Therefore, the method proposed by Lloyd (1970) may not be suitable to estimate the annual stream flow data with strong persistence.

Şen (1999) compared the independent and dependent return periods for short and long records. For short records, there are significant differences in return periods but as the number of records grew, the dependence theory reduced to independent case. This study shows that the length of record may influence the calculation of return period. Another

definition of return period used by other authors (Vogel 1987; Douglas et al. 2002) is the average number of trials required for the first occurrence of a critical event. This definition may be more useful in determining the return period for a reservoir operation because the interest is knowing the first time that the reservoir is at the risk of failure, rather than the average time between failures (Douglas et al. 2002).

Many researches had been conducted to find the probability distribution function parameters of rainfall data. Husak et al. (2006) had presented the use of Gamma distribution to reflect the variability of monthly rainfall in Africa. They used the Kolmogorov-Smirnov test for the goodness of fit of Weibull and Gamma functions for different locations in Africa. They observed that for 98% of the locations, the Gamma distribution proved the best fit. Sharma and Singh (2010) had used the normal, log-normal, Gamma, Weibull, and Pearson type III distributions to fit the daily rainfalls in Pantagar, India. The tests of goodness of fit used are the Kolmogorov-Smirnov, Chi-Square and Anderson-Darlin tests. They concluded that different types of distributions could fit for different locations. Becker (2009) had examined the frequency distributions of the daily precipitation in the United States, using North America regional analysis. He observed that the daily precipitation frequencies can be fitted well by the Gamma distribution. Olofintoye et al. (2009) had studied the peak daily rainfall distribution characteristics in Nigeria, by using different frequency distributions such as Gumbel, Log-Gumbel, Normal,Log-Normal, Pearson and Log-Pearson distributions.

Mathematically, the return period, Tr, is given by:

Tr =$\frac{1}{Pr(Q1)}$ (2.7)

Where:

$Q1$= a data value.

Pr is the probability of occurrence of Q1.

The probability of non-occurrence of Q1 can be written, in term of the return period as:

Pr (not Q1) =1-Pr (Q1) =1-$\frac{1}{Tr}$ (2.8)

The probability that Q1 will not occur for n successive years is given by:

Pr {not Q1}*Pr2 {not Q1} ----------*Prn {Q1} =Pr $\{not\ Q1\}^n$= {1-$\frac{1}{Tr}$}^{n} (2.9)

Risk, R1, is defined as the probability that Q1 will occur at least once in n years, or

R=1-Pr $\{not\ Q1\}^n$=1-{1-$\frac{1}{Tr}$}^{n} (2.10)

Equation 2.10 can be used to calculate the risk of failure as a function of project design life, n, and the design return period, Tr.

2.6 Climate Data

One major use of climate data in conjunction with rainfall data is in computer simulation of hydrological and agricultural systems. Rainfall-runoff models and crop growth models require, in addition to rainfall, net radiation or evaporation as a measure of energy input. In irrigation simulation studies, both rainfall and evaporation are also required. A special characteristic that must be preserved in stochastic modeling of climate data is the cross

correlation between variables. The models for generating climate (rainfall amount) data on monthly basis are discussed in this work.

2.6.1 Uncertainty in model parameters

In most of the above models, the estimated parameter values are assumed to be the "true" values and the sampling errors are ignored. Klemes and Bulu (1979) showed that this traditional approach to generating stochastic data underestimates the range of mean and variance of the likely future sequences. Several researchers have used Bayesian methods to overcome this deficiency.

2.6.2 Rainfall and climate data under climate change scenario

Concern has increased in recent years over climate change caused by increasing concentration of Carbodioxides (CO_2) and other trace gases in the atmosphere. A major effect of climate change may be alterations in regional hydrological cycles and changes in regional water availability. The use of modified water balance models offers many advantages in evaluating the regional impacts of global climate change (Gleick, 1989). The main source of climate change projections is general circulation models (GCMs). While current GCMs perform reasonably well in simulating the present climate with respect to annual and seasonal averages over large areas, they are considerably less reliable in specific catchment scale information that are necessary for hydrological studies. As a result, climate change impact studies have had to use a spectrum of climate change scenarios. These are generally constructed using observed records of temperature and rainfall, adjusted to reflect climate changes obtained from monthly average GCM results**.**

2.7 Rainfall Amount Models

Wang and Nathan (2000) developed a daily and monthly mixed (DMM) algorithm for the generation of daily rainfall. Daily rainfall data are generated month by month using the usual two-part model with two sets of parameters for the Gamma distribution, one estimated from the daily rainfall data and the other from monthly rainfall data. The monthly total is obtained by summing the daily values generated from the monthly Gamma parameters and adjusted for serial correlation. The generated daily rainfalls from the daily Gamma parameters are linearly scaled to match the serially correlated monthly rainfalls. Results for the Lake Eppalock catchment rainfall and for six other sites around Australia showed that the DMM algorithm reproduced the mean, coefficient of variation and skewness of daily, monthly and annual rainfall. The results were examined in detail for the Lake Eppalock catchment (in southern Australia); the algorithm worked well in reproducing the mean, coefficient of variation and skewness of monthly maximum daily rainfall, but not as well for the annual maximum rainfall. For the other six sites, the algorithm worked well in reproducing the mean and coefficient of variation but not as well in reproducing the skewness of the annual maximum daily rainfall.

The amount of precipitation is formulated to fit a particular water resources engineering application. Precipitation (mm), stream discharge (m^3/s), evaporation (mm) are some example of random variables. In order to obtain an adequate model, it is necessary to choose an appropriate type of probability distribution as well as suitable values for the distribution parameters. Stern and Coe (1982) observed that the cube-roots of monthly and annual precipitation amounts are distributed normally. The estimation of the probabilities of

rainfall values needs the mathematical definition of the probability density function of the selected distribution. The mathematical description of these density functions could be found in the literature, and the following are only those used in the work reported herein.

2.7.1 Normal (Gaussian) distribution

The normal distribution is an important family of continuous probability distributions, applicable in many fields. Each member of the family may be defined by two parameters, location and scale: the mean ("average", μ) and variance (standard deviation squared), σ^2.

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\{-1/2[\frac{x-\mu}{\sigma}]\text{^}2\} \quad (2.11)$$

where:

P(x) = probability density function

μ = average or mean

σ = standard deviation

x = random variable (rainfall amount)

The distribution of a normal random variable with mean, zero and variance, one is called a standard normal distribution.

In order to find probabilities concerning X, we convert its values to z-scores using

$$Z = \frac{\times-\mu}{\sigma} \quad (2.12)$$

In which Z is a standardized normal variate.

Whenever X assumes a value of x, the corresponding value of Z is given by Eq. 2.11 which can be transformed to:

$$X = \sigma z + \mu \qquad (2.13)$$

In which z value can be obtained from normal distribution table.

The normal distribution is often used to model variation when the distribution is symmetric and has a single mode.

2.7.2 Exponential distributions

Exponential distributions are a class of continuous probability distribution. An exponential distribution is a special case of gamma distribution where $\alpha = 1$ (α is shape parameter of gamma distribution) and it arises naturally when modeling the time between independent events which occur at a constant average rate. The probability density function is given as (Hann, 1994):

$$P(x) = \lambda e^{-\lambda(x-\mu)} \qquad x \geq \mu;\ \lambda > 0 \qquad (2.14)$$

In which λ is the scale parameter which is often equals to $1/\beta$.

The case where $\mu = 0$ and $\beta = 1$ is called the standard exponential distribution. The equation for the standard exponential distribution is

$$P(X) = e^{-x}, \text{ for } x \geq 0 \qquad (2.15)$$

$\mu=1/\lambda=\beta$=mean

$\sigma^2=1/\lambda^2=\beta^2$=variance

The Cumulative density function is given by:

$P(X\leq x)=1-e^{-\lambda x}$ and

$$P(X>x)=e^{-\lambda x} \quad (2.16)$$

$\ln R_{u=-\lambda x}$ and $x=-\frac{lnR_u}{\lambda}=-\mu \ln R_u$ (2.17)

$R_{u=}$ occurrence of rainfall.

2.7.3 Probability Log-normal distribution

The log-normal distribution occurs in practice whenever we encounter a random variable which is such that its logarithm has a normal distribution, that is, if the random variable Y = ln (X) has a normal distribution with mean μ and standard deviation σ. The resulting probability density function of X is:

$$P(x)=\frac{1}{\sigma_y\sqrt{2\pi}}\exp\left[-12(\frac{\ln(x)-\mu_y}{\sigma_y})\text{^}2\right] \quad (2.18)$$

For this distribution,

$$Z=(\frac{lnx-\mu_y}{\sigma_y}) \quad (2.19)$$

This can be written as:

$$\ln x=\sigma_y Z+\mu_y \quad (2.20)$$

where:

$\mu_y = mean$ and σ_y =standard deviation of the distribution.

The mean and variance of lognormal distribution are given as:

$$\mu_y = e^{\mu+\frac{\sigma^2}{2}} \tag{2.21}$$

and

$$\sigma_y{}^2 = e^{2\mu+\sigma^2(e^{\sigma^2}-1)} \tag{2.22}$$

2.8 Goodness of Fit Tests

Various criteria can be employed to evaluate the suitability of a probability distribution for describing a set of data. Statistical goodness-of-fit tests as well as graphical display such as probability plots are effective ways to determine whether the fitted distributions are consistent with the given set of observations (Stedinger et al., 1993). In addition, the predictive ability of a model is important in applying the model for prediction of future events. Often in selecting a particular distribution, one may be tempted to select a distribution with large number of parameters. Generally, the more parameters a distribution has, the better it will fit to the data. However, difficulty in the parameter estimation arises, and the distribution may be too rigid to accurately extrapolate beyond the range of the available data.

2.8.1 Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov statistic is based on the largest vertical difference between the theoretical and the empirical cumulative distribution function. The statistics can be specified as:

$$D = \max_{1 \le i \le n} \left(F(x_i) - \frac{i-n}{n}, \frac{i}{n} - F_n(x_i) \right) \quad (2.23)$$

where:

$$F_n(x_i) = \frac{1}{n} * [Number\ of\ observations \le x] \quad (2.24)$$

Given that rainfall amount is assumed to follow a continuous random process, the theoretical cumulative distribution function $F(x_i)$ for a given probability density function (exponential, normal and lognormal) can in general be written as:

$$F(x_i) = \int_{-\infty}^{a} f(x_i) \quad (2.25)$$

The result to be obtained from the generated rainfall amount will be compared to the ones from the observed values for different return periods. This test is used to decide if a sample comes from a hypothesized continuous distribution

2.8.2 Anderson-Darling Test

The Anderson-Darling procedure is a general test to compare the fit of an observed cumulative distribution function to an expected cumulative distribution function. This test

gives more weight to the tails than the Kolmogorov-Smirnov test.The Anderson-Darling statistic (A^2) is defined as:

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1) * \left[lnF(X_i) + ln\left(1 - F(X_{n-i+1})\right)\right] \quad (2.26)$$

in which all the terms are as previously defined.

2.8.3 Chi-Squared test

The test is used to test for a significant difference between the distribution suggested by a data sample and a selected probability distribution. It is the most widely used one-sample analysis for testing a population distribution. Many statistical tests, such as the t-test for a mean, assume that the data have been drawn from a normal population, so it may be necessary to use a statistical test, such as the Chi-square test, to check the validity of the assumption for a given sample of data.

The test statistic, which is a random variable, is a function of the observed and expected frequencies, which are also random variables**.**

The Chi-Square test is used to determine if a sample comes from a population with a specific distribution. This test is applied to binned data, so the value of the test statistic depends on how the data is binned. It should be noted that this test is available for continuous sample data only.

Usually there is no optimal choice for the number of bins (k), there are several formulae which can be used to calculate this number based on the sample size (N). For example, EasyFit employs the following empirical expression:

$$k = 1 + log_2 N \tag{2.27}$$

The Chi-Squared statistic is defined as (Chow et al, 1988):

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \tag{2.28}$$

Where

$O_i =$ Observed frequency for bin i and $E_i =$ Expected frequency for bin i calculated as

$$E_i = F(x_2) - F(x_1) \tag{2.29}$$

F = the CDF of the probability distribution being tested

The degree of freedom is given by:

ν=k-h-1,

where:

k= no. of observation and

h= no.of parameters used in calculation

The test had been used by many researchers such as Zekai and Ali, (1999) and Bhakar.et al. (2008).

2.8.4 ANOVA Test

ANOVA is a statistical test which analyzes variance. It is helpful in making comparison of two or more means which enables a researcher to draw various results and predictions about two or more sets of data. ANOVA test includes one-way ANOVA, two-way ANOVA or multiple ANOVA depending upon the type and arrangement of the data. One-way ANOVA has the following test statistics:

$$F = \frac{MST}{MSE} \tag{2.30}$$

Where, F = ANOVA Coefficient

MST = Mean sum of squares due to treatment

MSE = Mean sum of squares due to error.

MST is given below as:

$$MST = \frac{SST}{p-1} \quad \text{and} \quad SST = \sum n(x - xi)^2 \tag{2.31}$$

Where, SST = Sum of squares due to treatment

p = Total number of populations

n = Total number of samples in a population.

MSE is given below as:

$$MSE = \frac{SSE}{N-p} \quad \text{and} \quad SSE = \sum (n-1)S^2 \tag{2.32}$$

Where, SSE = Sum of squares due to error

S = Standard deviation of the samples

N = Total number of observations

2.8.5 Paired t-Test

The t-Test for two related samples is given by:

$$t=\frac{\bar{D}}{S_m} \tag{2.33}$$

$$S_{m=}\sqrt{\frac{\sum d^2}{n(n-1)}} \tag{2.34}$$

$$\sum_i d^{2=} \sum_i D_i^{2-} \frac{(\sum_i D_i)^2}{n} \tag{2.35}$$

where

$S_{m=}$Sampling variation and

n is the number of pairs

D_i=difference between the ith pair of values and $\bar{D}$ is the average difference between values.

$$\nu = n-1 \tag{2.36}$$

ν is the degree of freedom.

The test described here assumes that the pairs being compared are related and that the differences values are normally distributed. The computed test statistic will be compared with table t value for level of significant α and ν.

2.8.6 Expert Modeler

The model statistics provides summary information and goodness-of-fit statistics for each estimated model. Although the Time Series Modeler offers a number of different goodness-of-fit statistics, we opted only for the stationary R-squared value. This statistic provides an estimate of the proportion of the total variation in the series that is explained by the model and is preferable to ordinary R-squared when there is a trend or seasonal pattern, as is the case here. A significance value less than 0.05 implies that there is structure in the observed series which is not accounted for by the model.

CHAPTER THREE

MATERIALS AND METHODS

3.1 Study Location

The Study was carried out in Samaru which is located at 11^{o}11' North Latitude and 7^{o}38' East Longitude with an altitude of 686m above mean sea level. The area is situated in the Northern Guinea Savannah ecological zone of Nigeria. It has three distinct seasons; namely, the hot dry season from March to May, the warm rainy season from June to October, and a cool dry season from November to February. The area has a relative humidity of 36% during the dry season and 78.5% during the wet season. The average minimum and maximum temperatures recorded in the area were 15.6^{o}C and 38.5^{o}C, respectively.

3.2 Data Assembly

Weather data was sourced from the metereological unit of the Institute for Agricultural Research (IAR), Ahmadu Bello University (ABU) Zaria. Sixty years (1953-2012) of daily data on rainfall, minimum and maximum relative humidity (%), minimum and maximum temperature (^{o}C), sunshine (hours) and wind speed (km/day) were collected. The data were recorded by an IAR trained personnel using weather instrument installations which conformed with the World Meteorological Organization (WMO) standard.

3.2.1 Preliminary Data Analysis Method

The statistical parameters estimated from the rainfall data collected, were the mean and the Standard deviation. The parameters estimation methods of these distributions are well

known and available in literature. The two mostly used methods are the moments and the maximum likelihood methods. The parameters were estimated in this work based on monthly values using maximum likelihood method through the EasyFitXL 5.6 software. EasyFitXL is a statistical package that makes use of some 'Excel' features to manipulate climatic data for better results. Random numbers were then generated to obtain probability of rainfall occurrence and the Z-scores were estimated from probability values and further used to generate the monthly rainfall amounts using the selected probability distribution model equations.

The data were analyzed statistically using SPSS (Statistical Package for Social Sciences) version 20. In the analysis of the data, a 95% confidence level was adopted and values less than 0.05 are considered to be statistically significant and values that are greater than 0.05 are considered not statistically significant.

3.2.2 Formulation of Hypothesis

The goodness of fit test measures the compatibility of random sample with the theoretical probability distribution. The goodness of fit test is applied for testing the following null (H0) and alternative (HA) hypotheses:

H0: the monthly rainfall data follow the specified distribution

HA: the monthly rainfall data does not follow the specified distribution.

3.3 Fitting of Probability Distributions

Monthly rainfall amount can be characterized by probability distribution functions known from the statistical literature. In the present study, the normal, exponential, and log-normal distribution functions were used with the aid of Eqs 2.13, 2.17 and 2.20, respectively to estimate monthly rainfall amounts. This method is detailed by Richard. (2003).

The results obtained from the generated rainfall amount were compared to those from the observed values for different return periods. Usually, a common method to analyze or evaluate model is to plot the measured values against generated values and the correlation between them is used for model evaluation (Kobayashi and Salam, 2000). Ideally, this relationship should be linear with a slope of unity and intercept of zero. This procedure is popularly known as statistical paired sample t-test to see whether the difference between measured and generated result is significant or not.

Three PDFs were equally used based on the estimated parameters to estimate rainfall amounts with return periods ranging from 2yr to 50yr in order to know which of the model(s) is/are the best for estimating return periods and rainfall amounts for agricultural and water resources planning purposes.

3.4 Rainfall Prediction Strategy

The characteristics of a population are given by the parameters. A population of hydrological variables is not always exactly known. Therefore, statisticians use samples taken from the population. Parameters estimated from a sample are called statistics, which are functions of the elements of the sample and reflect the properties of the population.The statistical parameters were obtained from observed rainfall data on monthly basis using a programme called EasyFitXL through the maximum likelihood method.Thereafter, SPSS version 20 software was used to generate random numbers based on the parameters of the Distributions.The generated random numbers were processed to obtain the probability of monthly rainfall occurrence.Subsequently both the parameters and probability were fitted to the reduced or the resulted equations of integration for all the three PDFs to obtain the predicted/generated values of monthly rainfall amounts.

3.5 Statistical Analysis Framework

The following goodness-of-fit tests; viz. Kolmogorov-Smirnov, Anderson-Darling, ANOVA analysis and expert modeller goodness of fit test were used along with the Chi-square test at $\alpha(0.05)$ level of significance for the selection of the best fit Probability distribution as indicated in Eqs 2.21,2.24,and 2.26,respectively.

The goodness of fit tests mentioned above were fitted to the monthly rainfall data. The test statistic of each test were computed and tested at (α=0.05) level of significance. Accordingly, the ranking of different probability distributions were taken from 1 to 3 based

on minimum test statistic value. The distribution holding the first rank was selected for all the three tests independently. The assessments of all the probability distributions were made on the basis of total test score obtained by combining the entire three tests. Maximum score of 3 was awarded to rank first probability distribution based on the test statistic and further less score were awarded to the distribution having rank more than 1, that is 2 to 3. Thus, the total score of the entire three tests were summarized to identify the best fit distribution on the basis of highest score obtained. The probability distribution having the maximum score was selected as the best probability distribution.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Preliminary Analysis

The length of the series was sixty years, ranging from January 1953 to December 2012. In other words, the expected number of observations was 720. However, it is important to point out that observations with unrecorded rainfall data were dropped. This was perhaps due to the season or the failure to take records by the recording officer. Practically, during the Harmattan and dry season, rain is very limited and during some months is absent. Eventually a sample size of 431 was used for the study rather than the initial sample size of 720 and the amplitude of ten years was used to categorize the pooled data of rainfall for a better understanding of the data.

Figure 4.1 reveals that the occurrence of rainfall data first decreased from the period 1953-2002 before it later increased from 2003-2012. However, the rate of occurrence of rainfall in later years was still at a low level compared to the early years.

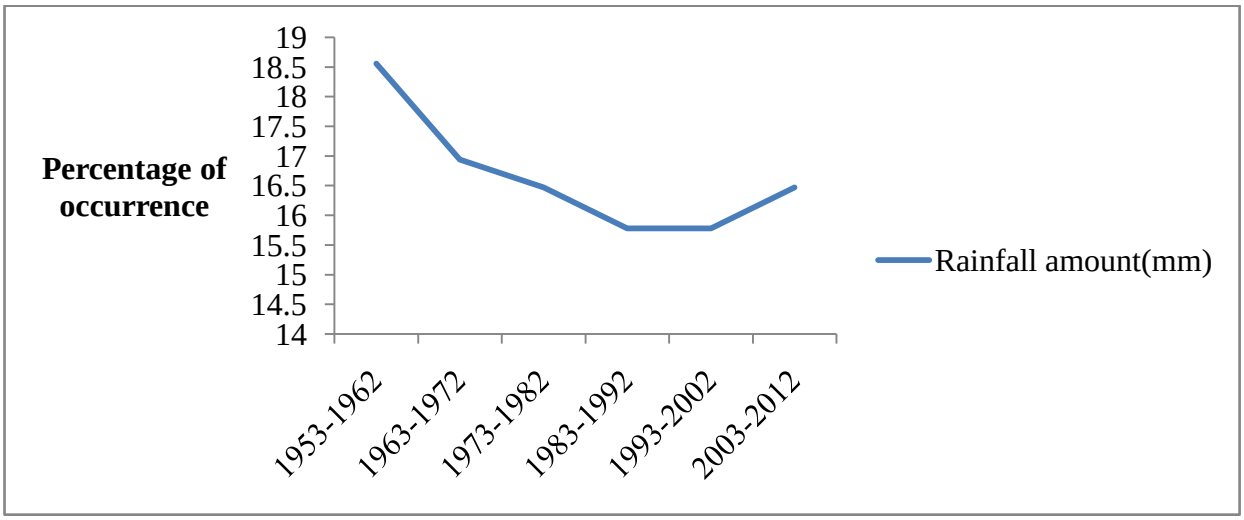


Figure 4.1: Trend of the percentage of rainfall occurrence across period of ten years from 1953-2012 in Samaru

Evidence from Fig 4.1 shows that 18.56% of the rainfall data occurred within the period 1953-1962 which represents the period with the highest level of rainfall occurrence. The period 1983-1992 and 1993-2002 had the smallest percentage of occurrence with 15.78% each.

Table 4.1 presents the descriptive statistics of the observed rainfall amount in Samaru from 1953-2012 across 12 periods of five years. The results show that the period 1953-1957 had the highest average rainfall amount which was observed to be 154 mm while the period 1978-1982 had the least average rainfall amount which was observed to be 102 mm in the study area. The minimum amount of rainfall was calculated to be 0.3 mm and was observed within the period 1968-1972 and 1973-1977 while the maximum amount of rainfall was estimated to be 474mm and was observed during the period 1998-2002. However, the

period 2008-2012 exhibited more fluctuations in the amount of rainfall than any other period based on the categorization of the years considered. By pooling the data together, it can be observed that the minimum quantity of rainfall amount from 1953-2012 in Samaru was 0.3 mm while the maximum was 473 mm. The average amount of rainfall in the study area was 126mm with a standard deviation of 104mm. This implies that a lot of variations in rainfall amount were present in the study area from one period to another.

Table 4.1: Descriptive statistics of observed rainfall in millimeters across periods of five years intervals from 1953-2012 in Samaru

Years classes	Minimum	Maximum	Mean	Standard Deviation
1953-1957	0.80	413.10	154.09	116.00
1958-1962	0.80	381.00	129.83	109.77
1963-1967	2.80	379.90	133.90	108.24
1968-1972	0.30	445.30	131.29	111.25
1973-1977	0.30	329.00	114.69	100.69
1978-1982	1.30	357.30	101.50	98.06
1983-1987	3.30	322.10	114.94	102.81
1988-1992	2.60	328.90	117.12	89.63
1993-1997	3.80	269.00	115.14	76.56
1998-2002	1.60	473.60	126.90	110.73
2003-2007	1.70	364.50	121.99	102.23
2008-2012	2.30	439.70	138.15	120.90
Pool value	0.30	473.60	125.53	104.40

Source*: adapted*

Table 4.2 shows the descriptive statistics of monthly rainfall for the observed sixty years and PDFs parameters for every month were used to estimate and predict the amount of rainfall in the past and in the future, respectively. From the table, it can also be seen that the mean monthly rainfall for January and December for the whole 60 years are 1.4mm and 5.0mm, respectively which was a clear indication that rainfall is hardly observed during the period.

Table 4.2: Descriptive Statistics of observed monthly Rainfall from 1953-2012

Year	Months	Frequency	Mean(mm)	Standard Deviation(mm)
1953-2012	January	2	1.40	.42
1953-2012	February	7	9.27	6.58
1953-2012	March	24	20.62	26.29
1953-2012	April	53	40.64	34.67
1953-2012	May	60	105.25	67.99
1953-2012	June	60	138.58	72.09
1953-2012	July	60	201.46	105.70
1953-2012	August	60	215.88	118.69
1953-2012	September	59	146.03	97.79
1953-2012	October	43	45.14	46.23
1953-2012	November	5	40.32	58.10
1953-2012	December	1	5.00	0

Tables 4.3 and 4.4 show the mean and standard deviation of the 60 years predicted values from the months of May to September.

Table 4.3: Descriptive statistics for Predicted rainfall Values

Month	Frequency (years)	Mean(mm)	Standard Deviation(mm)
May	60	104.15	67.00
June	60	141.95	69.45
July	60	196.28	98.85
August	60	222.53	122.43
September	60	156.86	94.09

Table 4.4: Log-normal Descriptive Statistics of observed Monthly Rainfall (mm) from 1953-2012

Year	Months	Frequency	Mean	Std. Deviation
1953-2012	February	7	1.75	1.41
1953-2012	March	24	1.93	1.76
1953-2012	April	53	3.12	1.30
1953-2012	May	60	4.35	.94
1953-2012	June	60	4.63	1.05
1953-2012	July	60	5.05	.88
1953-2012	August	60	5.01	1.15
1953-2012	September	59	4.63	1.04
1953-2012	October	43	3.18	1.29
1953-2012	November	5	2.93	1.42

4.2 Analysis of Statistical Parameters

This section compared the values obtained from the normal, log-normal and the exponential distribution with the observed values in order to identify which of the PDFs can best be used to predict rainfall amounts in Samaru. The tests that were conducted to achieve this are: the Chi-square test, Kolmogorov-Smirnov test, Anderson Darling test, ANOVA analysis and expert modeller goodness of fit test. The tests were conducted based on 95% confidence level that is α=0.05 level of significance. However, Values less than 0.05 will be considered to be statistically significant and values greater than 0.05 will be considered not statistically significant.

Figure 4.2 shows the mean monthly rainfall amount (mm) between the observed data and the predicted value for the month of May, June, July August and September. However, careful observation of the graph indicates that the mean monthly amount of rainfall between the observed and predicted are close to each other but sometimes inconsistent and fluctuate.

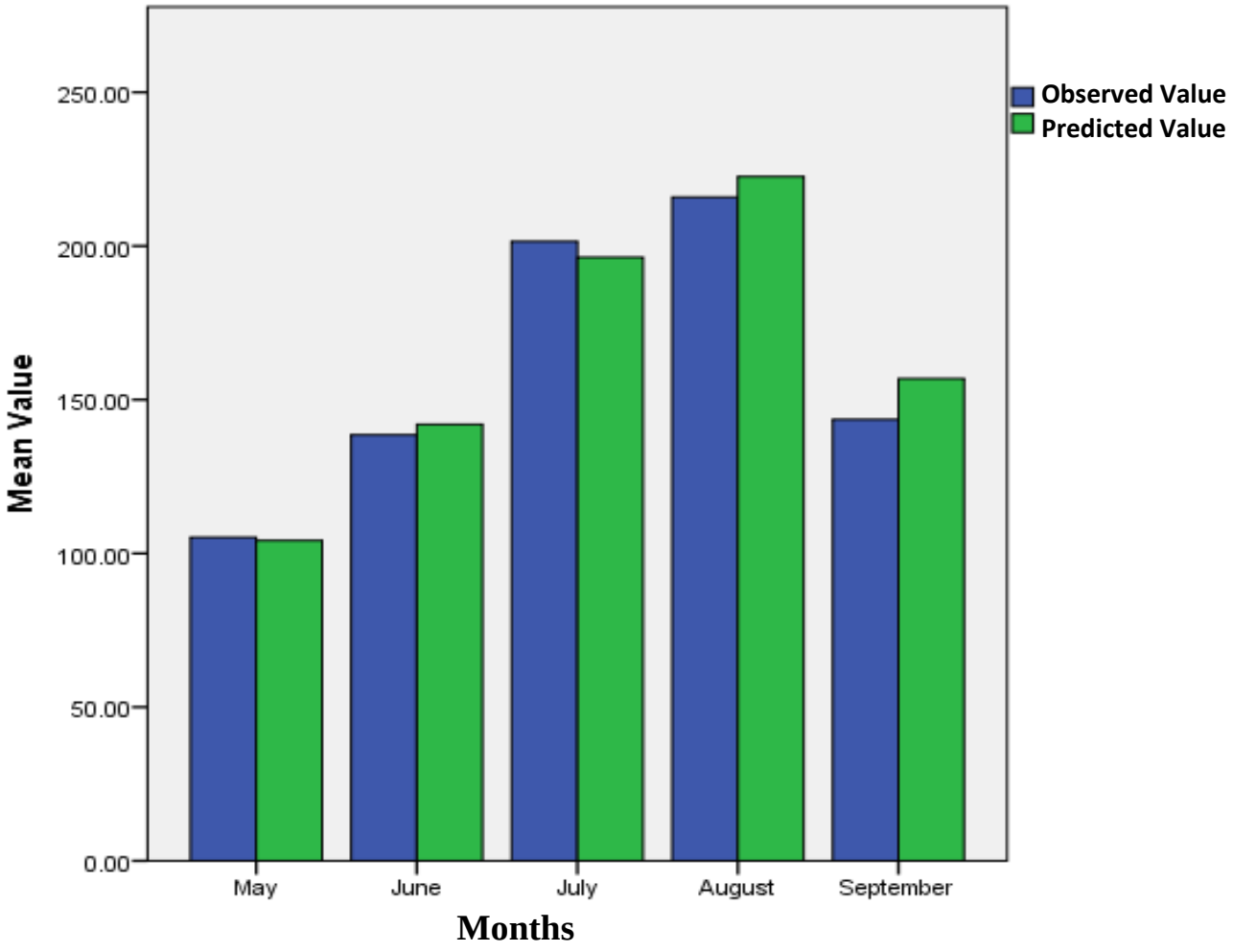


Figure 4.2: Comparing mean monthly values of observed and predicted (mm)

4.2.1 Chi-Square Test

The result of the analysis from Table 4.5 for the observed and predicted rainfall values between the months of May-September indicates that their mean monthly rainfall amounts are not statistically significant from each other with all P-values greater than 0.05.

Table 4.5: Chi-square test of monthly Observed and Predicted rainfall values

May-September	Observed and Predicted	N	Mean (mm)	Standard Deviation(mm)	Chi-square (P-value)	Comment
May	Observed	60	105.25	68.00	.356	Not Significant
	Predicted	60	104.15	67.01		
June	Observed	60	138.58	72.09	.406	Not Significant
	Predicted	60	141.95	69.45		
July	Observed	60	201.46	105.70	.431	Not Significant
	Predicted	60	196.28	98.85		
August	Observed	60	215.88	118.69	.356	Not Significant
	Predicted	60	222.53	122.43		
September	Observed	60	143.59	98.78	.381	Not Significant
	Predicted	60	156.86	94.09		

4.2.2 Kolmogorov-Smirnov and Anderson-Darling Tests Statistic

From Table 4.6, the rule of thumb is that the smaller the value of a statistics for a given probability density function, the better the distribution. Thus, based on the Kolmogorov-Smirnov and Anderson-Darling test, the normal PDF was estimated to be the best PDF in terms of estimating monthly rainfall amounts followed by the log-normal PDF. However, the exponential PDF was the least.

Table 4.6: Kolmogorov-Smirnov and Anderson-Darling Goodness of Fit Test

Distributions	Kolmogorov-Smirnov test	Rank	Anderson-Darling test	Rank
Normal	0.1	1	6.79	1
Exponential	0.15	3	15.13	3
Log-normal	0.12	2	10.12	2

4.3 Comparison of Observed Monthly Rainfall Amounts with Generated

Between the three distributions viz; normal, log-normal and the exponential, a goodness of fit tests were conducted to determine the distribution that will best predict the monthly rainfall amount such that the predicted values can be used to simulate the observed values. An expert modeller in SPSS was equally used to determine the PDF that has a statistically significant relationship with the observed values collected. The reason for the use of the expert modeller is because it is a time series model which takes into account seasonal changes that occurs.

Figure 4.3 shows the pattern in which the observed values follow on a monthly basis from 1953-2012. The monthly series exhibited numerous peaks, many of which appear not to be equally spaced, as well as the upward trend. The unequally spaced peaks suggest the presence of a periodic component to the time(monthly) series, that is to say the high peak visible in the months of July-September indicates a seasonal component in the data.

There are also peaks that do not appear to be part of the seasonal pattern and which represent significant deviations from the neighbouring data points. These points may be outliers, which must be addressed by the Expert Modeller.

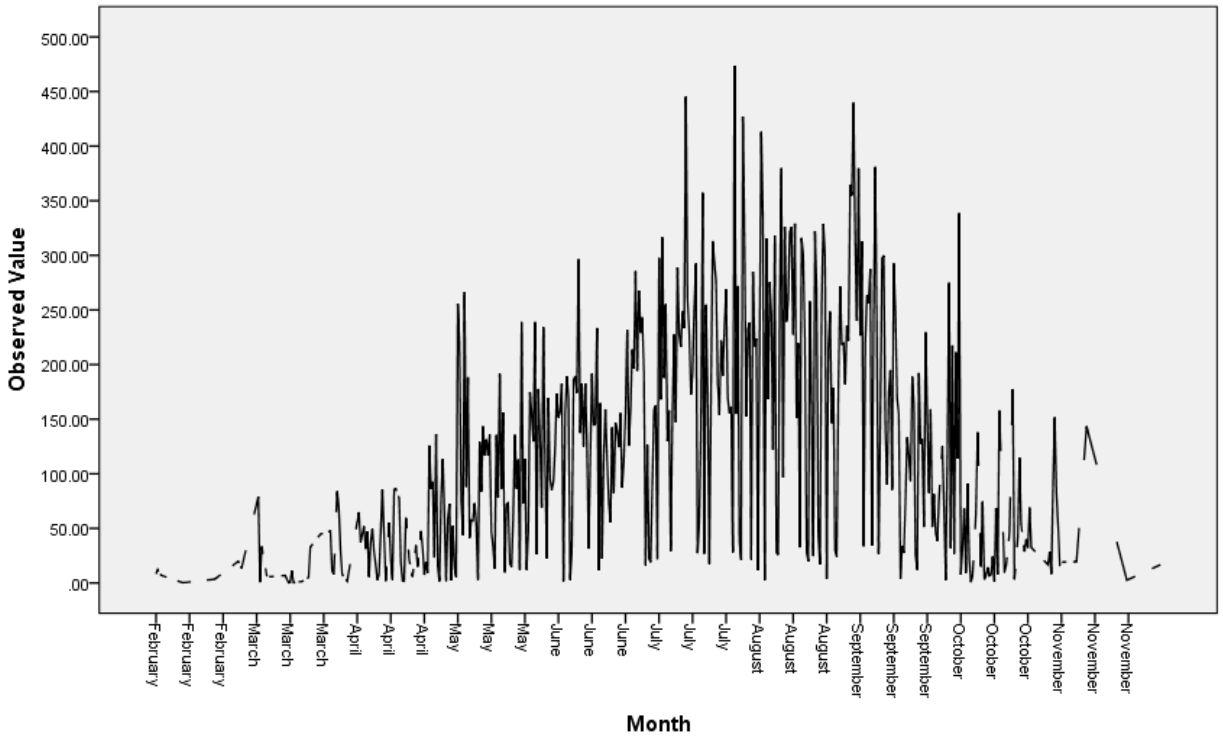


Figure 4.3: Sequence chart for Observed Value

4.3.1 Testing the Normal Model with the Observed Value

From the model fit statistics in Table 4.7, a value of 1.000 means that the model does an excellent job of explaining the observed variation in the series. The Ljung-Box statistic, also known as the modified Box-Pierce statistic, provides an indication of whether the model is correctly specified. The value of 0.200 shown here is not significant statistically, so it can be said that the model is correctly specified. The Expert Modeller detected nine points that were considered to be outliers and the results obtained without their effects.

Table 4.7: Model Statistics

Model 1	Number of Predictors	Model Fit statistics	Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Statistics	df	Sig.	
Normal Model_	1	1.000	20.47	16	.200	9

Fig.4.4 shows how well the normal model fits the observed value. The predicted values show good agreement with the observed values, indicating that the model has satisfactory predictive ability.

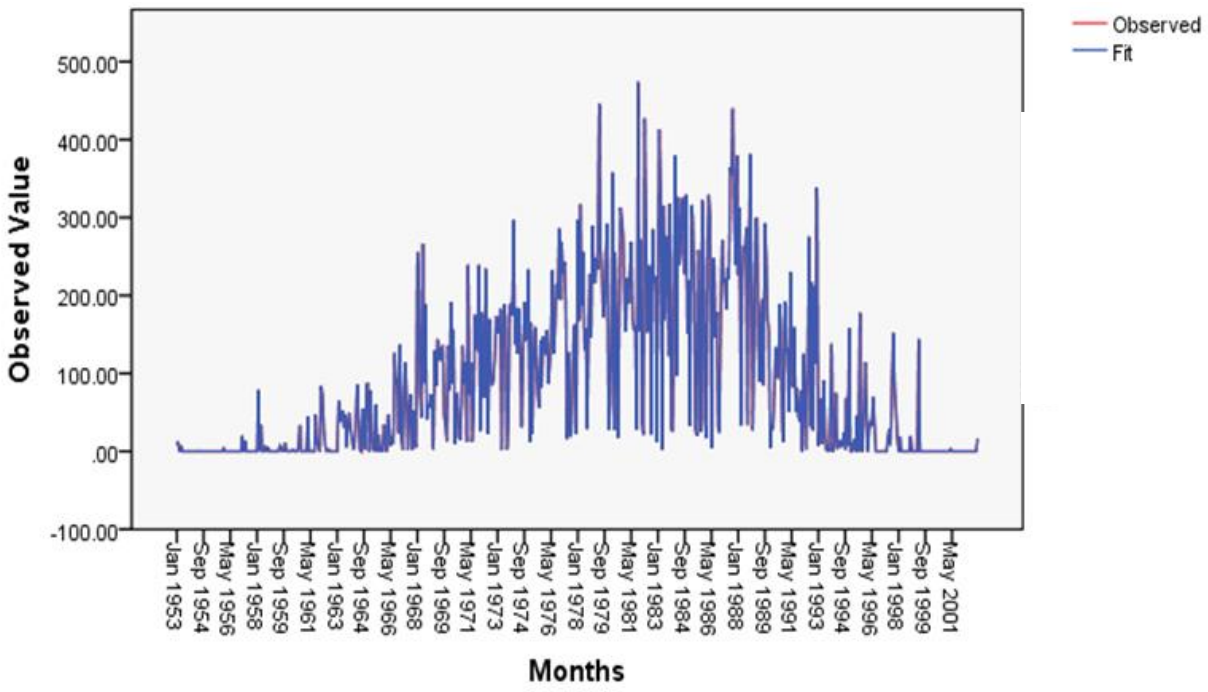


Figure 4.4: Fit Pattern of the Normal and Observed Value

4.3.2 Testing the Log-normal Model with the Observed Value

From the model fit statistics in Table 4.7a, a value of 0.934 means that the model does a good job of explaining the observed variation in the series. A significance value less than 0.05 implies that there is structure in the observed series which is not accounted for by the model. The value of 0.596 shown here is not significant statistically.

Table 4.7a: Model statistics

Model 2	Number of Predictors	Model Fit statistics	Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Statistics	df	Sig.	
Log-normal Model_	1	.934	21.77	16	.596	15

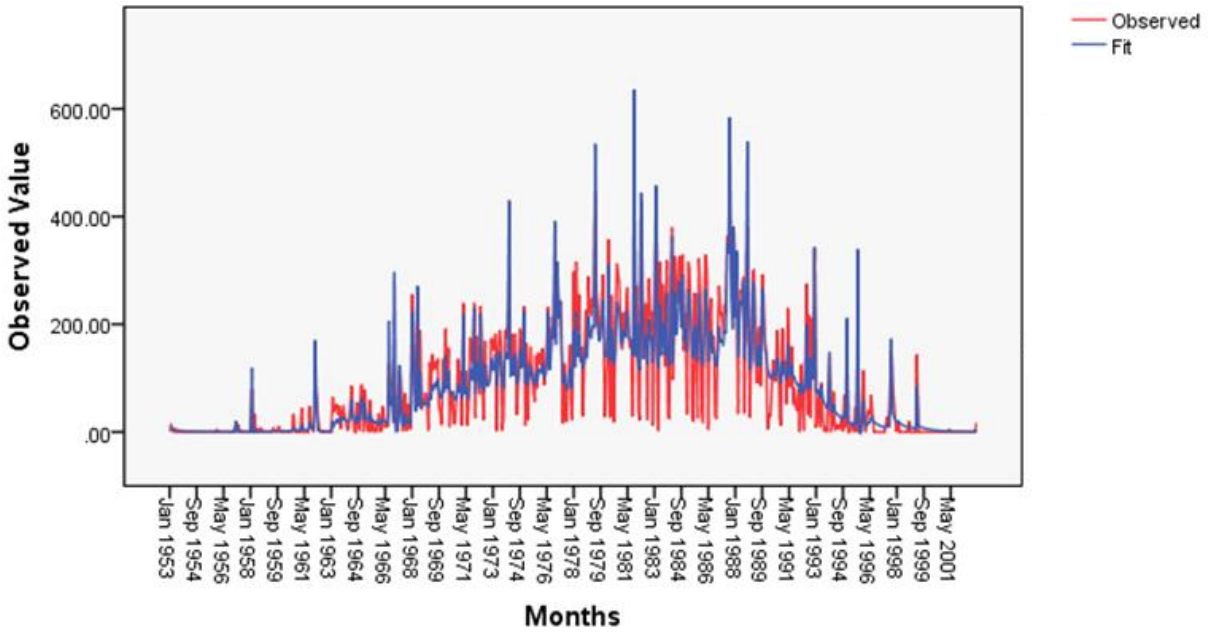


Figure 4.5: Fit Pattern of the Log-normal and Observed Value

Figure 4.5 shows how well the log-normal model fits the observed value. The predicted value shows some degree of variation with the observed value especially where there is seasonal fluctuation.

4.3.3 Testing the Exponential Model with the Observed Value

From the model fit statistics in Table 4.7b, a value of 0.911 means that the model does a good job of explaining the observed variation in the series. The value of 0.151 shown here is not significant statistically.

Table 4.7b: Model Statistics

Model 3	Number of Predictors	Model Fit statistics	Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Statistics	df	Sig.	
Exponential Model_	1	.911	14.04	16	.151	17

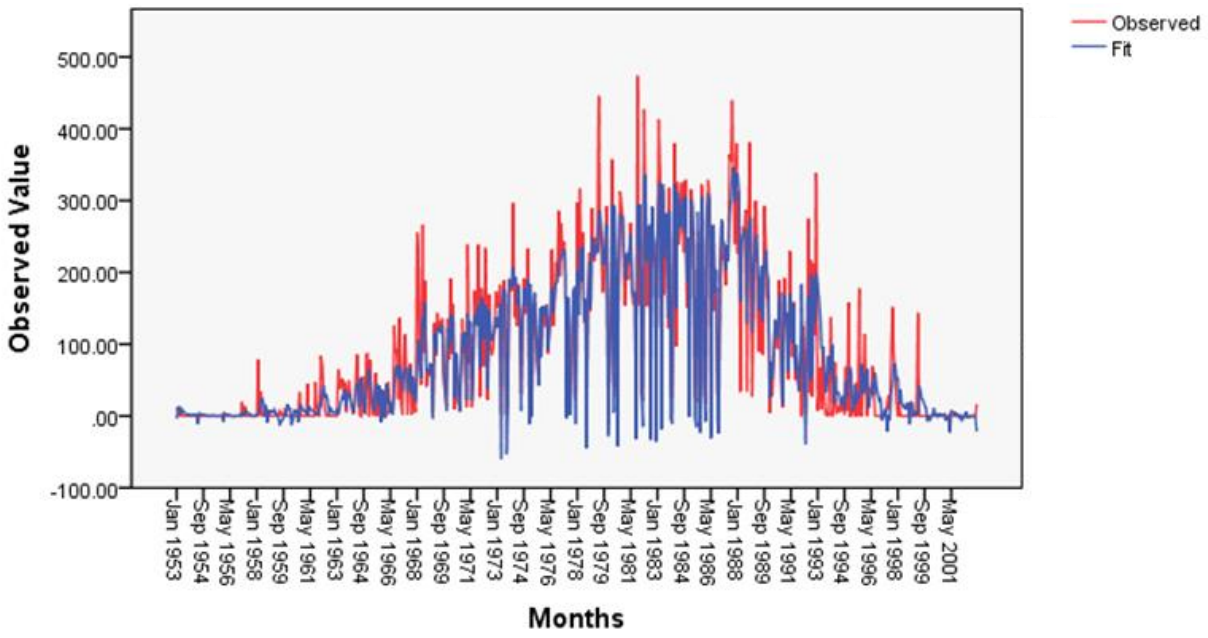


Figure 4.6: Fit Pattern of the Exponential and Observed Value

Figure 4.6 indicates that the exponential model does a good job of predicting the observed value compared with the log-normal. However, the generating power of the exponential model is not as strong enough as that of normal and log-normal models that is 0.911 which is less than 1.

4.4 Analysis of Trend

Figure 4.7 shows the time series plot of the average rainfall amount from 1953-2012 after grouping the data by intervals of five years. The graph shows an up-and-down pattern occurring, however, in an irregular manner. It can be readily seen that there was a nonlinear trend component in the form of a U-shape in the observed series. Specifically, the trend in average rainfall amount was decomposed into two phases. In the first phase which was

comprised between 1953 and 1982, the average rainfall amount decreased linearly while in the second phase which was comprised of the remaining period (between 1983 and 2012) increased linearly with a global minimum occurring around 1982 and 1983.

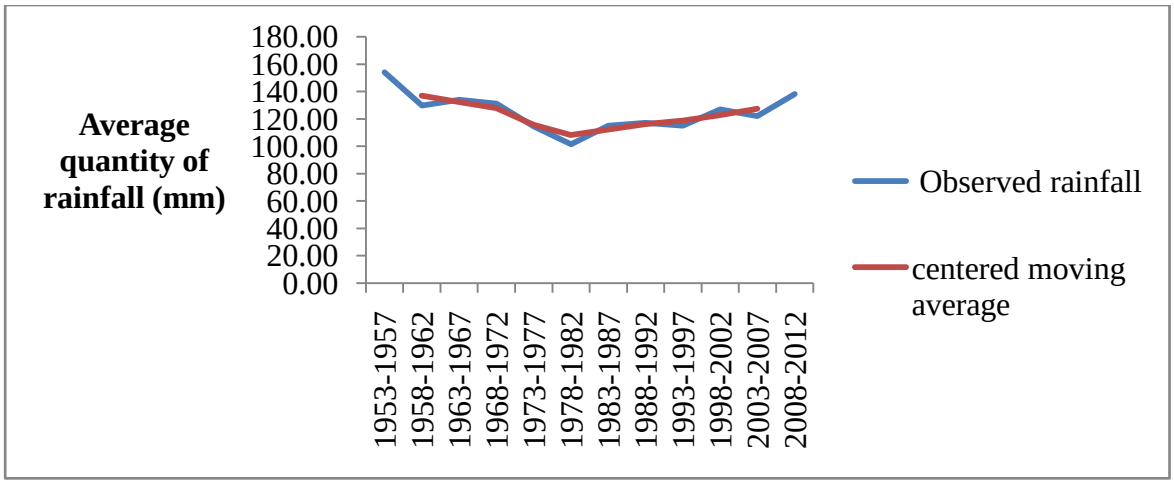


Figure 4.7: Time series plot of rainfall amount across periods of five years from 1953-2012 in Samaru

4.5 Monthly Rainfall Generation based on the Selected Models

Z-scores were estimated from the generated probabilities/generated random numbers,using the inverse distribution function in SPSS which returns the value for the normal distribution with a mean and standard deviation.

Table 4.8 shows probabilities of rainfall occurrence generated for 20years (February to November). The random numbers were generated with an in built function in SPSS using the command "RAND (mean, stddev) which generates random values from a normal

distribution with standardized mean and standard deviation. The remaining Tables of probability values are given in Appendix A1.

Table 4.8: Generated Probabilities for 20 years (random numbers method)

Year \ RN	February	March	April	May	June	July	August	September	October	November
1	0.700	0.770	0.650	0.890	0.790	0.210	0.610	0.730	0.930	0.360
2	0.330	0.840	0.600	0.810	0.290	0.490	0.110	0.510	0.730	0.340
3	0.510	0.820	0.390	0.870	0.200	0.830	0.560	0.850	0.430	0.980
4	0.850	0.680	0.320	0.760	0.590	0.770	0.980	0.740	0.540	0.370
5	0.650	0.860	0.900	0.660	0.700	0.340	0.300	0.740	0.640	0.690
6	0.690	0.510	0.760	0.870	0.310	0.890	0.510	0.680	0.430	0.020
7	0.270	0.290	0.400	0.020	0.960	0.230	0.350	0.540	0.260	0.140
8	0.430	0.180	0.110	0.490	0.060	0.110	0.490	0.320	0.880	0.940
9	0.230	0.290	0.030	0.250	0.910	0.130	0.160	0.040	0.210	0.870
10	0.780	0.700	0.430	0.880	0.080	0.380	0.930	0.360	0.150	0.390
11	0.480	0.730	0.440	0.020	0.040	0.130	0.910	0.490	0.260	0.850
12	0.500	0.890	0.930	0.100	0.020	0.820	0.010	0.040	0.990	0.730
13	0.830	0.890	0.730	0.420	0.790	0.690	0.670	0.770	0.670	0.960
14	0.060	0.740	0.380	0.600	0.230	0.520	0.430	0.920	0.960	0.440
15	0.570	0.820	0.610	0.410	0.650	0.380	0.490	0.510	0.840	0.480
16	0.020	0.630	0.270	0.140	0.060	0.870	0.530	0.420	0.310	0.530
17	0.880	0.060	0.420	0.170	0.280	0.610	0.110	0.790	0.980	0.140
18	0.830	0.250	0.540	0.030	0.450	0.570	0.130	0.800	0.990	0.400
19	0.130	0.870	0.870	0.140	0.100	0.320	0.210	0.330	0.300	0.790
20	0.450	0.750	0.410	0.340	0.340	0.320	0.900	0.850	0.700	0.290

From Table 4.9, the Z-values were obtained from the random numbers generated in Table 4.8 with an in-built function in SPSS using the command "IDF.NORMAL (randn, mean, stddev)" which give the Z-value with a mean of 0 and a standard deviation of 1. The remaining Tables of Z- values and generated rainfall amounts for various months are given in Appendix A2.

Table 4.9: Computed Z-values based on Generated probabilities

Year	Feb	March	April	May	June	July	August	Sept	Oct	Nov
1	.538	.740	.396	1.201	.793	-.804	.285	.603	1.496	-.348
2	-.435	.983	.248	.895	-.553	-.014	-1.220	.015	.606	-.425
3	.020	.926	-.271	1.140	-.854	.968	.154	1.055	-.189	2.122
4	1.048	.459	-.481	.708	.217	.725	2.053	.648	.088	-.342
5	.388	1.062	1.307	.403	.539	-.403	-.515	.641	.356	.484
6	.505	.023	.709	1.112	-.503	1.245	.020	.459	-.179	-2.002
7	-.613	-.542	-.253	-1.985	1.735	-.727	-.380	.096	-.645	-1.092
8	-.164	-.907	-1.202	-.032	-1.515	-1.210	-.013	-.468	1.151	1.583
9	-.723	-.547	-1.953	-.673	1.365	-1.117	-.996	-1.719	-.806	1.146
10	.783	.531	-.188	1.152	-1.408	-.317	1.506	-.370	-1.020	-.271
11	-.041	.621	-.156	-2.045	-1.707	-1.133	1.359	-.023	-.631	1.039
12	.003	1.244	1.473	-1.292	-2.065	.910	-2.333	-1.767	2.275	.602
13	.969	1.224	.625	-.195	.793	.505	.444	.726	.442	1.773
14	-1.588	.631	-.301	.260	-.750	.054	-.182	1.410	1.751	2.602
15	.176	.927	.285	-.222	.375	-.315	-.037	.037	.996	-.059
16	-2.131	.336	-.614	-1.063	-1.565	1.124	.085	-.198	-.495	.072
17	1.176	-1.576	-.201	-.958	-.576	.279	-1.215	.822	2.003	-1.082
18	.970	-.669	.104	-1.910	-.126	.174	-1.129	.844	2.208	-.244
19	-1.117	1.139	1.120	-1.063	-1.267	-.482	-.816	-.446	-.519	.807
20	-.113	.684	-.237	-.414	-.411	-.465	1.279	1.022	.511	-.547

4.5 .1 Monthly Rainfall generation for different Return Periods

In order to obtain the monthly return period of rainfall for months of May to September based on the three PDFs for the purpose of planning agricultural production, the mean and standard deviation for the required months were used with respect to the observed and generated rainfall amounts.

Table 4.10 shows the mean monthly values and standard deviations calculated from the generated rainfall values and were used to determine the monthly return period values for the normal and log-normal PDFs.

Table 4.10: Generated Normal and Log-normal Mean and SD (mm)

		Normal PDF		Log-normal PDF	
Month	Frequency	Mean	Standard Deviation	Mean	Standard Deviation
May	60	104.15	67.01	4.51	.73
June	60	141.95	69.45	4.87	.57
July	60	196.28	98.85	5.13	.66
August	60	222.53	122.43	5.34	.56
September	60	156.86	94.09	4.91	.83

Table 4.11 shows the values for the observed, generated normal, log-normal and exponential rainfall values of return periods of 2,5……,50 years for the month of May**.**

Table 4.11:Rainfall amounts in (mm) for the Month of May

T(year)	Observed	Normal Predicted	Log-normal predicted	Exponential predicted
2	105.25	104.15	90.92	72.19
5	162.48	160.54	168.07	167.62
10	192.40	190.02	231.72	239.81
15	207.15	204.56	271.49	282.04
20	217.10	214.36	302.10	312.00
30	230.26	227.33	347.95	354.23
40	238.53	235.48	380.23	384.19
50	244.91	241.76	407.17	407.43

Figure 4.8 shows the mean value for the various PDFs based on the return periods considered. However, the ANOVA analysis given in Table 4.12 was used to determine if the differences in their means were significant or not.

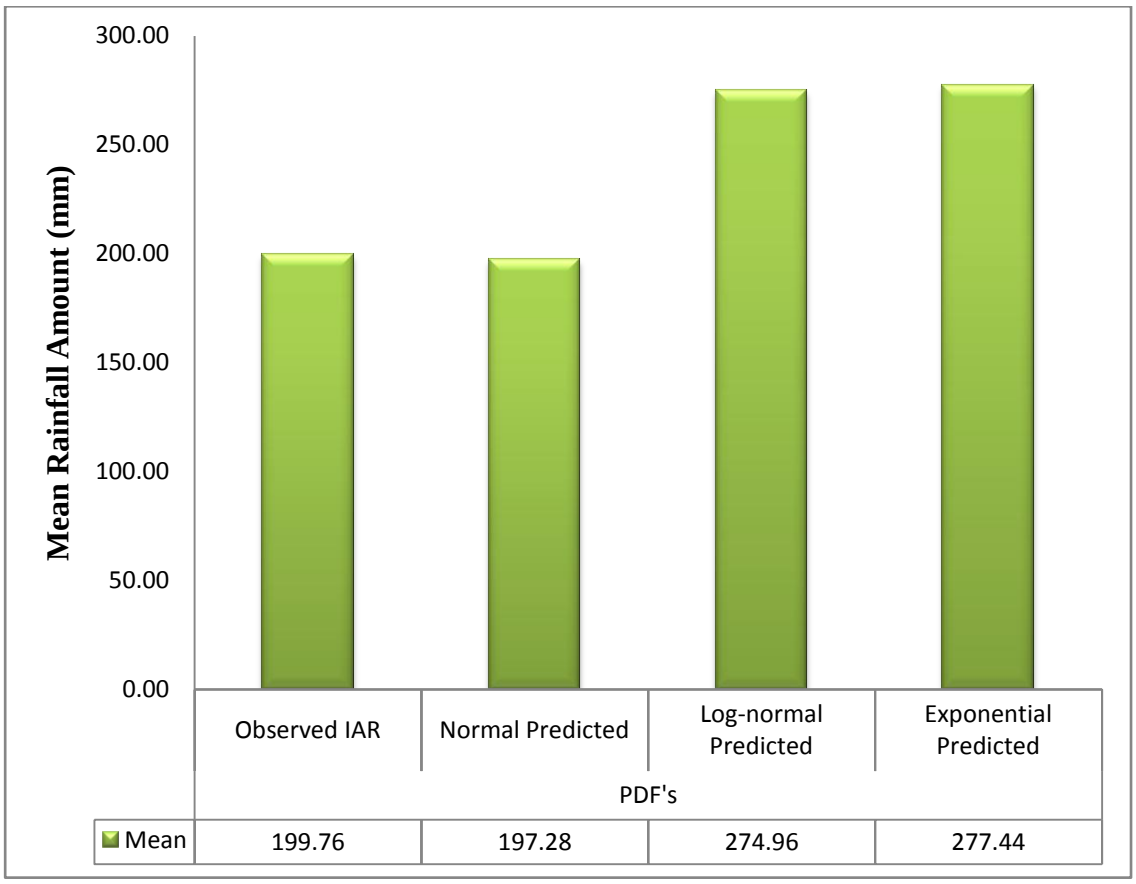


Figure 4.8: PDF Chart of Return Periods for the Month of May

The result of the analysis from Table 4.12 indicates that there is no significant difference in the mean value of the PDFs (observed, predicted normal, log-normal and exponential rainfall values) based on the return periods with a P-value of 0.107 and df of 3, that is

P-value >0.05.

Table 4.12: ANOVA of Return Period for the month of May

	Sum of Squares	df	Mean Square	F	Sig.
Between PDF's	48323.546	3	16107.849	2.224	.107
Within PDF's	202777.109	28	7242.040		
Total	251100.654	31			

Table 4.13 shows the values for the observed; predicted normal, log-normal and exponential rainfall values for return period specified.

Table 4.13: Rainfall amounts in (mm) for the Month of June

T(year)	Observed	Normal Predicted	Log-normal predicted	Exponential predicted
2	138.58	141.95	130.32	98.39
5	199.25	200.40	210.55	228.46
10	230.97	230.96	270.56	326.85
15	246.61	246.02	306.17	384.41
20	257.16	256.19	332.81	425.24
30	271.12	269.63	371.63	482.80
40	279.88	278.07	398.29	523.64
50	286.64	284.59	420.16	555.31

Figure 4.9 shows the mean value for the various PDFs based on the return periods. However, the ANOVA analysis (Table 4.14) was used to determine if the differences in their means are significant or not.

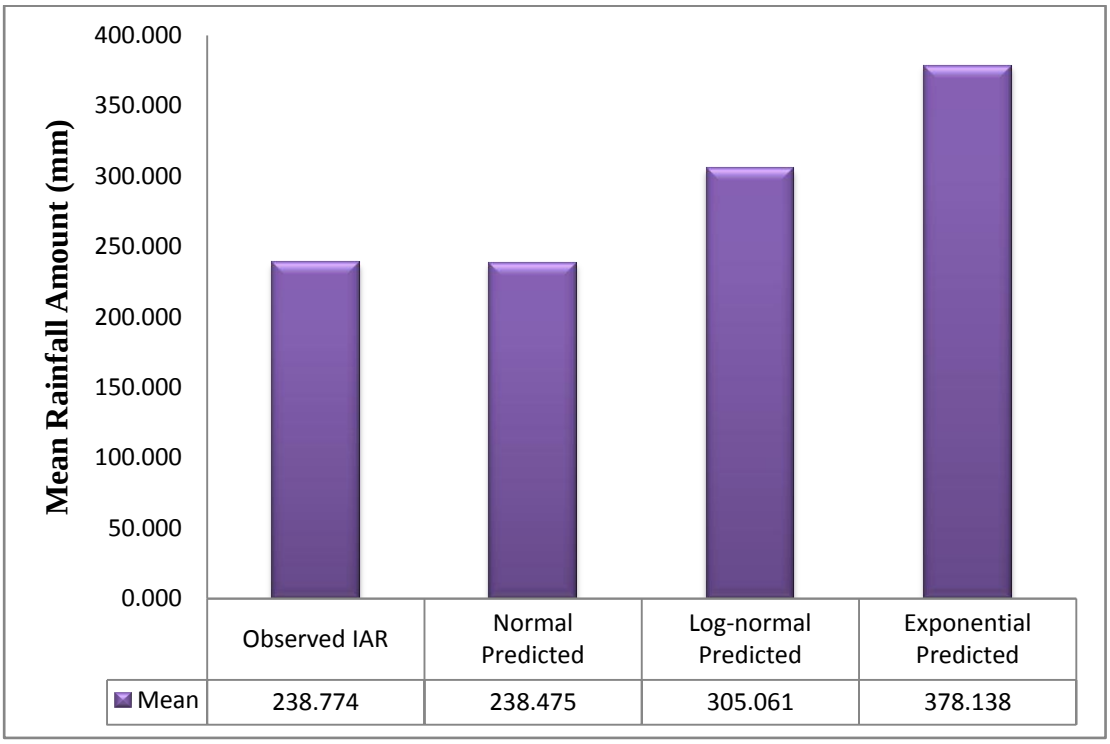


Figure 4.9: PDF Chart of Return Periods for the Month of June

The result of the analysis from Table 4.14 indicates that there is a statistical significant difference in the mean value of the PDFs between the observed and the predicted rainfall for the month of June based on the return periods with a P-value of 0.024 and df of 3, that is P-value <0.05.

Table 4.14: ANOVA of Return Period for the month of June

Sources of error	Sum of Squares	df	Mean Square	F	Sig.
Between PDFs	106191.533	3	35397.178	3.673	.024
Within PDFs	269810.649	28	9636.095		
Total	376002.182	31			

Table 4.15 is a multiple comparison test that shows where the level of significance lies based on the ANOVA test from Table 4.14 for the month of June. The result of the analysis from Table 4.15 indicates that the observed rainfall value is not statistically significant from the normal predicted value and log-normal predicted value since the

P-values are >0.05, that is 1.00 and 0.540, respectively. However, the test result indicated that the observed rainfall value is statistically significant from the exponential predicted value since P-value <0.05, that is 0.039. In general, the exponential predicted value is statistically significant from the observed rainfall values while the remaining PDFs are not statistically significant from each other. All other Tables of test results for other months are given in Appendix C.

Table 4.15: Turkey HSD Multiple Comparisons Post-HOC test for the Month of June

(I) PDFs	(J) PDFs	Mean Difference (I-J)	Std. Error	Sig.
Observed	Normal Predicted	.299	49.08	1.000
	Log-normal Predicted	-66.287	49.08	.540
	Exponential Predicted	-139.364	49.08	.039
Normal Predicted	Observed	-.299	49.08	1.000
	Log-normal Predicted	-66.586	49.08	.536
	Exponential Predicted	-139.663*	49.08	.039
Log-normal Predicted	Observed	66.287	49.08	.540
	Normal Predicted	66.586	49.08	.536
	Exponential Predicted	-73.077	49.08	.457
Exponential Predicted	Observed	139.364*	49.08	.039
	Normal Predicted	139.663*	49.08	.039
	Log-normal Predicted	73.077	49.08	.457

Figure 4.10 shows the mean value for the various PDFs based on the return periods. However, the ANOVA analysis (Appendix C) was used to determine if the differences in their means are significant or not.

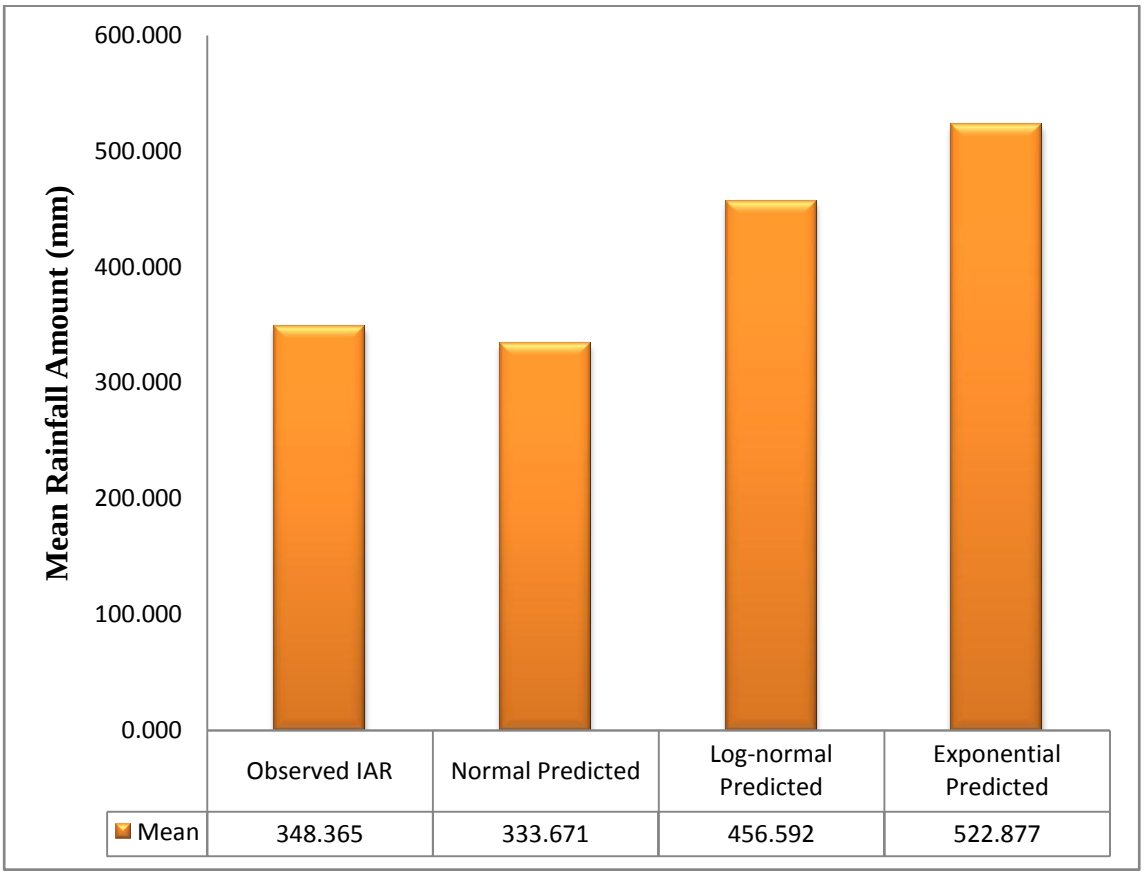


Figure 4.10: PDF Chart of Return Periods for the Month of July

Figure 4.11 shows the mean value for the various PDFs based on the return periods. However, the ANOVA analysis (Appendix C) was used to determine if the differences in their means are significant or not.

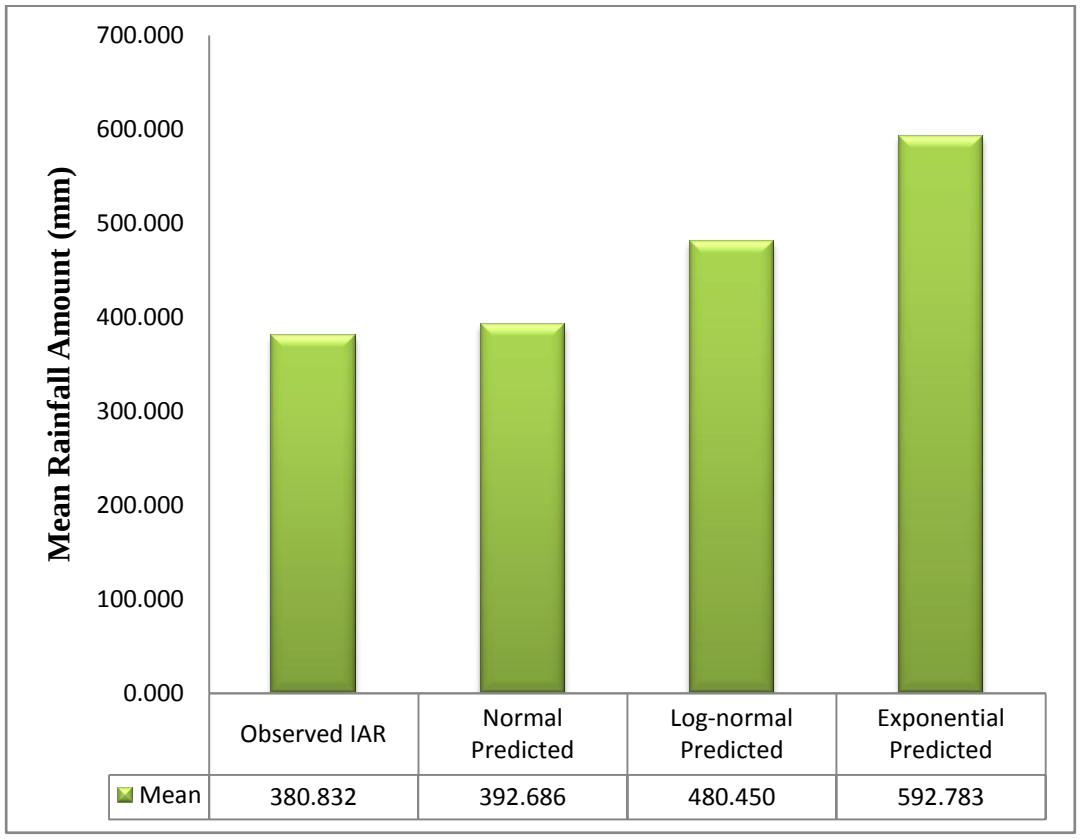


Figure 4.11: PDF Chart of Return Periods for the Month of August

Figure 4.12 shows the mean value for the various PDFs based on the return periods. However, the ANOVA analysis (Appendix C) was used to determine if the differences in their means are significant or not.

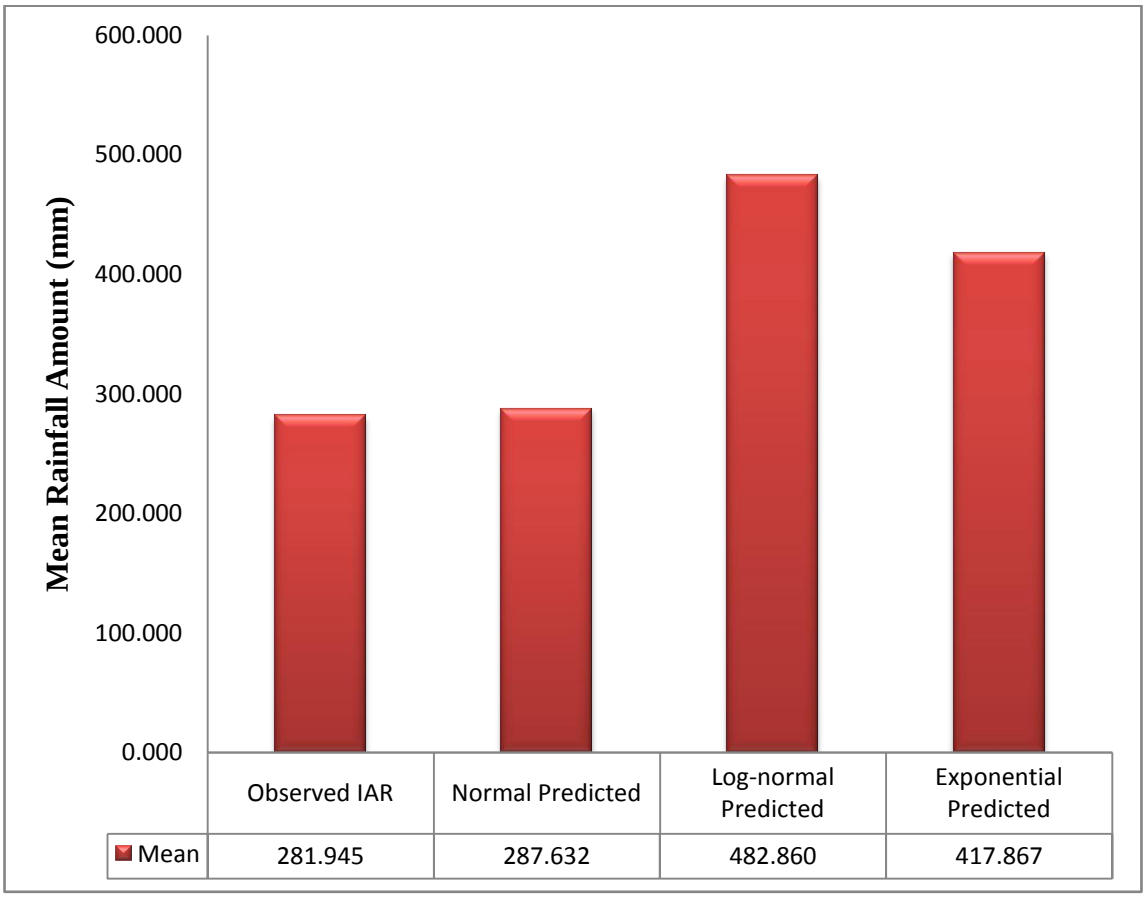


Figure 4.12: PDF Chart of Return Periods for the Month of September

From Table 4.16, it is equally clear that virtually all PDFs always show that there is no statistically significant difference between the observed and predicted rainfall amounts for various return periods except for the month of May and August for normal PDF.

Table 4.16: t-Test results on rainfall values obtained at different return periods

Month	NORMAL	LOG-NORMAL	EXPONENTIAL	t- Critical value
MAY	3.65	1.18	1.13	2.37
JUNE	0.16	1.31	1.30	2.37
JULY	0.78	1.12	1.21	2.37
AUG.	4.61	1.34	1.29	2.37
SEP.	2.24	1.38	1.28	2.37

CHAPTER FIVE

SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1 Summary

This study was set to determine the best probability density function for generating rainfall amount based on the known parameters of mean and standard deviation from a data series that covered from 1953-2012 in Samaru, Zaria, Kaduna State. Secondary data that were used for this study were collected from the Institute for Agricultural Research (IAR) Zaria, Kaduna State. The other objective of the study was to identify the most suitable probability density function (PDF) for generating monthly rainfall amount from 1953-2012 in Samaru and to further generate rainfall amounts for some selected return periods. In order to accomplish this, three PDFs were considered in the study which were: normal, log-normal and exponential probability density functions. Also, in order to ascertain the most suitable PDF to generate rainfall amounts, five major tests were conducted which were the: Kolmogorov-Smirnov test, Anderson-Darling test, Chi-square test, ANOVA test and Expert modeller goodness of fit test. Furthermore, mean test scores were computed to make decisions given the heterogeneous nature of the results provided by the various tests conducted. However, other objectives were also considered for the study which are, to examine the trend existing in the rainfall amount in the study area and to use various return periods to generate rainfall amounts based on the normal, log-normal and exponential PDFs in the study area.

The result of the analysis indicates that the mean monthly amount of rainfall for January, February, November and December were the least with frequencies of 2, 7, 5 and 1 and a mean monthly amount of 1.40mm, 9.27mm, 40.32mm and 5.00mm respectively for the period of 60 years. Also, the result of the analysis indicates that the months of June, July, August and September experience the highest amount of rainfall with frequencies of 60 each and mean monthly amount of 138.58mm, 201.46mm, 215.88mm and 146.03mm respectively.

In terms of ascertaining the most suitable PDF to adopt, the results of the analysis from the study show a range of R-squared value of between 0.911 to 1.00. The normal PDF has the highest value of 1,while the exponential PDF has the least value of 0.911. The higher the R-squared value the better the model fit and predictive power, therefore the normal PDF has the best generating power followed by log-normal PDF and exponential PDF, the least since it always give higher values of rainfall amounts.

Also, the monthly rainfall amounts based on some selected return periods and different months ranging from months of May to September were also generated and two tails t-Test and ANOVA were conducted to test for their mean difference in order to know which of the PDFs is suitable for generating rainfall amount at different return periods in Samaru and its environs. However, for the log-normal and normal PDFs, observed and generated values did not differ significantly for most of the months. The normal PDF was found to generate best followed by log-normal PDF, while the exponential PDF always give higher values of rainfall amounts for the five months considered.

5.2 Conclusion

From the study, it can be concluded that:

1. The statistical parameters of mean and standard deviation were obtained from the observed rainfall data. The monthly means for the sixty years recorded data range from 1.40mm for month of January to 215mm for month of August. The standard deviations range from 0mm for the month of November to 118.69mm for the month of August.
2. The monthly rainfall amounts were generated from the parameters and the comparison made between them and the observed values. The various tests conducted indicate that the generated rainfall values were approximately the same as the observed ones for both normal and log-normal PDFs.The values generated using the exponential PDF were found to be higher than the observed.The Kolmogorov-Smirnov and Anderson-Darling tests ranked normal PDF to be first followed by log-normal PDF and exponential PDF the least suitable for rainfall generation. The model fit statistics test equally confirmed the aforementioned.
3. The monthly rainfall amounts generated based on some return periods using the three selected PDFs show that there were no statistical significant differences between the mean monthly rainfall amounts for observed and generated values at 95% probability level. The Chi-square test, ANOVA and student t-Test were all indicated that the observed and generated values were approximately the same with normal PDF being the best model.

5.3 Recommendations

The following recommendations can be put forward based on the results obtained:

1. .Since the goodness of fit test showed that the normal and log-normal PDFs were equally good in terms of generating monthly rainfall amounts in the study area, it is therefore recommended that the two distributions be used in predicting and generating monthly rainfall amounts in Samaru, Zaria, Kaduna State.
2. This work should be done using rainfall data from several stations within a drainage basin or catchment area.
3. The work should be made available to local and immediate environments to Samaru through Extension specialists on climatic information dissemination for use in planning and management of specific rainfed crop production.
4. It is suggested that other PDF should be investigated for their appropriateness in generation of rainfall amounts from observed data.

REFERENCE

Becker, E. J.(2009) "*The Frequency Distributions of Daily Precipitation Over the United States*", A Ph. D. Dissertation, University of Maryland, College Park,

Bhakar, S.R, Bansal, A.N, Chhajed N, Purohit RC (2006). Frequency analysis of consecutive days maximum rainfall at Banswara, Rajasthan India. ARPN J. Eng. Appl. Sci. 1(3): 64-67**.**

Bhakar, S.R., Mohammed, Iqbal. Mukesh,Devanda.,(2008).Probability analysis of rainfall at Kota.Indian. *J.Agric.Res*. 42(3):201-206.

Chow, V.T., Maidment, D.R. and Mays, L.W. (1988). *Applied Hydrology*. McGraw-Hill International Editions, New York.

De Michele, C. and Bernardara, P. (2005). *Spectral analysis and modeling of space time*. Atmospheric Research, pp. 124-136

Deni, S.M. and Jemain, A. A. (2009). *Fitting the distribution of dry and wet spells with alternative probability models*. Met. and Atmosph. Physics 104: 13-27.

Douglas, E. M., Vogel, R. M., and Kroll, C. N. (2002). "Impact of streamflow persistence on hydrologic design*." Journal of Hydrologic Engineering, 7(3), 220-227 Hydrologic Engineering*, 7(3), 220-227.

Edoga, R. N. (2007). Determination of length of growing season in Samaru using different potential Evapotranspiration Models. *AU J. T*. 11 (1): 28-35.

Fernández, B., and Salas, J. D. (1999). "Return period and risk of hydrologic events. I: Mathematical formulation." *Journal of Hydrologic Engineering*, 4(4), 297-307

Fernández, B., and Salas, J. D. (1999). "Return period and risk of hydrologic events. II: Applications." *Journal of Hydrologic Engineering*, 4(4), 308-316

Gleick, P.H. (1989). *Climate change, hydrology, and water resources*. Rev. Geophys., 27, 329–344.

Goldreich, Y. (1995): *Temporal variations of rainfall in Israel*. Climate Research 5: 167-179.

Hann, C.T. (1994). *Statistical methods in hydrology*. Affiliated East-West Press Pvt. Ltd., New Delhi.

Husak G.J., Michaelson J. and Funk C. (2007). "Use of Gamma Distribution to Report Monthly Rainfall in Africa Drought Monitoring Applications", *International Journal of Climatology*, Vol. 27, Issue 7.pp 935-944.

Jamaludin, S. Jemain, A.A. (2007) Fitting daily rainfall amount in Peninsular Malaysia using several types of exponential distributions. *Journal of Applied Sciences Research*, No 3(10), pp.1027–1036.

Kihupi, N.I., Kingamkono, R., Rwamugira, W., Kingamkono, M., Mhita, M., and Brien, K.O. (2002). *Promotion and integration of indigenous knowledge in seasonal climate forecasts*. Consultancy report submitted to drought monitoring centre (Harere, Zimbabwe).

Kite, G.W. (ed.) (1978). *Frequency and risk analyses in hydrology*, 2nd Edition, Water Resources Publication, Colorado.

Klemes,V. and Bulu, A. (1979). Limited confidence in confidence limits derived by operational stochastic hydrologic models. *J. Hydrol.*, 42, 9-22**.**

Kobayashi, K. and Salam, M.U. (2000). Comparing simulated and measured values using mean squared deviation and its components. *Agron. J*., 92, 345-352.

Larsen, G.A. and Pense, R. B. (1982). Stochastic simulation of daily climatic data for agronomic models. *Agron. J*. 74: 510-514.

Lloyd, E.H. (1970). "Return periods in the presence of persistence." *Journal of Hydrology*, 10, 291-298.

Loaiciga, H.A., and Marino, M. A. (1991). "Recurrence interval of geophysical events." *Journal of Water Resources Planning and Management*, 117(3), 367-382

Mortimore, M.J. (1989). *Adopting to drought: farmers, famines and desertification in West Africa*. Cambridge University press, Cambridge, U.K.

Nnaji, A.O (2001). Forecasting seasonal rainfall for agricultural decision making in northern Nigeria. *J.oj Agril. And forest Met*. 107: 193-205

Olaniran, O.J. (2002). *Rainfall anomalies in Nigeria*: the contemporary understanding at the inaugural Lecture Department of Geography, University of Ilorin, Kwara state.

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), pp 1-12.

Rafa, H. A. and R. Khanbilvardi.(2014) " Frequency Analysis of the Monthly Rainfall Data at Sulaimania Region, Iraq". *American Journal of Engineering Research (AJER*). Volume-03, Issue-05, pp-212-222.

Richard, H.M. (2003).*Modeling hydrologic change: statistical methods*. Lewis publishers, New York, Washington, D.C.

Richardson, C.W. (1985). *Weather simulation for crop management models*. Trans. ASAE, 28: 1602-1606.

Roldan, J. and Woolhiser, D.A (1982). Stochastic daily precipitation models, A comparison of occurrence processes. Water Resource. Res. 18: 1461-1468.

Saleh, A.A., Developing an empirical formulae to estimate rainfall intensity in Riyadh region, *J. Engineering Sciences*,23(2), 81-88 (2011**)**

Şen, Z. (1999). "Simple risk calculations in dependent hydrological series." *Hydrological Science Journal*, 44(6), 871-878

Sharma N.A. and Singh J. B. (2010) "Use of Probability Distribution in Rainfall Analysis", *New York Science Journal*, 3 (9), pp40-49.

Stedinger, J.R., Vogel, R.M. and Foufoula-Georgiou, E. (1993) Frequency Analysis of Extreme Events, *Handbook of Hydrology*, McGraw-Hill, New York, NY, USA.

Stern, R.D. and Coe, R. (1982). The use of rainfall models in agricultural planning. *J. Agri.l Met*. 26:35-50.

Stern, R.D. and Coe, R.,(1984).A model fitting analysis of daily rainfall data.*J.Roy.Statist.Soc*.A,147,part 1,1-34.

Upadhaya, A., and Singh,(1998). S.R. Estimation of consecutive day's maximum rainfall by various methods and their comparison. *Indian Journal of S. Cons*. 26 (2): 193-2001.

Usman, D.I. (2011).*Performance evaluation of three probability distribution models for predicting rainfall amount in Samaru*, Nigeria. MSc. Thesis, Agric. Engineering Dept. A.B.U., Zaria.

Vogel, R.M. (1987). "Reliability indices for water supply systems." *Journal of Water Resources Planning and Management*, 113(4), 563-579

Walpole, R.E., Mayers R.H. (1989). *Probability and Statistics for engineers and Scientists.* Fourth edition. Macmillan publishing company. New York.

Wang, Q.J. and Nathan, R.J. (2000). *A daily and monthly mixed model algorithm for stochastic generation of rainfall time series*. Unpublished manuscript.

Woodyer, K.D., McGilchrist, C.A., and Chapman, T.G. (1972). "Recurrence intervals between exceedence of selected river levels 4. Seasonal streams." *Water Resources Research*, 8(2), 435-443

Zekai, S.and Ali, G.E, (1999).Rainfall distribution function for Libya and rainfall prediction. *Hydrological sciences journal*, 44:5,665-680.

APPENDICES

Appendix A1: Generated probabilities based on random numbers for the 60 year between May and September

Prb. / Year	May	June	July	August	September
1	0.829	0.430	0.039	0.380	0.393
2	0.034	0.251	0.425	0.818	0.785
3	0.219	0.664	0.265	0.920	0.946
4	0.702	0.886	0.898	0.900	0.610
5	0.650	0.005	0.513	0.370	0.354
6	0.352	0.264	0.594	0.562	0.291
7	0.617	0.516	0.846	0.326	0.095
8	0.141	0.850	0.000	0.595	0.483
9	0.513	0.786	0.382	0.814	0.963
10	0.772	0.745	0.518	0.350	0.735
11	0.090	0.646	0.996	0.883	0.968
12	0.752	0.808	0.217	0.579	0.183
13	0.681	0.945	0.392	0.048	0.079
14	0.326	0.155	0.251	0.883	0.144
15	0.113	0.584	0.547	0.080	0.380
16	0.014	0.422	0.719	0.635	0.526
17	0.853	0.761	0.542	0.950	0.080
18	0.428	0.923	0.412	0.006	0.638
19	0.372	0.687	0.114	0.256	0.561
20	0.491	0.618	0.050	0.428	0.788

21	0.474	0.685	0.173	0.525	0.083
22	0.657	0.541	0.077	0.695	0.665
23	0.967	0.594	0.150	0.326	0.404
24	0.284	0.811	0.237	0.564	0.253
25	0.668	0.017	0.412	0.985	0.806
26	0.015	0.325	0.872	0.551	0.581
27	0.398	0.617	0.943	0.533	0.358
28	0.192	0.124	0.070	0.538	0.521
29	0.467	0.862	0.643	0.276	0.047
30	0.700	0.153	0.890	0.674	0.926
31	0.772	0.156	0.578	0.240	0.487
32	0.112	0.604	0.357	0.231	0.952
33	0.603	0.275	0.413	0.856	0.599
34	0.986	0.743	0.442	0.266	0.708
35	0.807	0.837	0.798	0.730	0.423
36	0.661	0.489	0.769	0.356	0.621
37	0.225	0.863	0.775	0.229	0.840
38	0.336	0.297	0.706	0.799	0.488
39	0.261	0.067	0.142	0.146	0.604
40	0.705	0.183	0.646	0.682	0.866
41	0.826	0.413	0.740	0.030	0.346
42	0.818	0.974	0.745	0.324	0.568
43	0.111	0.837	0.325	0.997	0.803
44	0.616	0.668	0.456	0.778	0.377

45	0.792	0.619	0.507	0.966	0.808
46	0.078	0.134	0.027	0.737	0.209
47	0.176	0.975	0.229	0.488	0.856
48	0.957	0.673	0.534	0.833	0.878
49	0.545	0.758	0.589	0.152	0.854
50	0.357	0.192	0.834	0.236	0.570
51	0.948	0.674	0.152	0.575	0.292
52	0.253	0.465	0.570	0.537	0.929
53	0.156	0.070	0.071	0.302	0.348
54	0.837	0.615	0.372	0.231	0.008
55	0.804	0.137	0.737	0.328	0.023
56	0.278	0.776	0.311	0.204	0.433
57	0.227	0.346	0.721	0.933	0.895
58	0.678	0.288	0.954	0.008	0.358
59	0.167	0.661	0.794	0.517	0.900
60	0.956	0.045	0.089	0.764	0.681

.

Appendix A2 : Computed Z-values based on Generated probabilities

Year \ Z-score	(May)	(June)	(July)	(August)	(September)
1	.950	-.176	-1.762	-.305	-.272
2	-1.825	-.671	-.189	.908	.789
3	-.776	.423	-.628	1.405	1.607
4	.530	1.206	1.270	1.282	.279
5	.385	-2.576	.033	-.332	-.375
6	-.380	-.631	.238	.156	-.550
7	.298	.040	1.019	-.451	-1.311
8	-1.076	1.036	0.000	.240	-.043
9	.033	.793	-.300	.893	1.787
10	.745	.659	.045	-.385	.628
11	-1.341	.375	2.652	1.190	1.852
12	.681	.871	-.782	.199	-.904
13	.470	1.598	-.274	-1.665	-1.412
14	-.451	-1.015	-.671	1.190	-1.063
15	-1.211	.212	.118	-1.405	-.305
16	-2.197	-.197	.580	.345	.065
17	1.049	.710	.105	1.645	-1.405
18	-.181	1.426	-.222	-2.512	.353
19	-.327	.487	-1.206	-.656	.154
20	-.023	.300	-1.645	-.181	.800

21	-.065	.482	-.942	.063	-1.385
22	.404	.103	-1.426	.510	.426
23	1.838	.238	-1.036	-.451	-.243
24	-.571	.882	-.716	.161	-.665
25	.434	-2.120	-.222	2.170	.863
26	-2.170	-.454	1.136	.128	.204
27	-.259	.298	1.580	.083	-.364
28	-.871	-1.155	-1.476	.095	.053
29	-.083	1.089	.366	-.595	-1.675
30	.524	-1.024	1.227	.451	1.447
31	.745	-1.011	.197	-.706	-.033
32	-1.216	.264	-.366	-.736	1.665
33	.261	-.598	-.220	1.063	.251
34	2.197	.653	-.146	-.625	.548
35	.867	.982	.834	.613	-.194
36	.415	-.028	.736	-.369	.308
37	-.755	1.094	.755	-.742	.994
38	-.423	-.533	.542	.838	-.030
39	-.640	-1.499	-1.071	-1.054	.264
40	.539	-.904	.375	.473	1.108
41	.938	-.220	.643	-1.881	-.396
42	.908	1.943	.659	-.457	.171

43	-1.221	.982	-.454	2.748	.852
44	.295	.434	-.111	.765	-.313
45	.813	.303	.018	1.825	.871
46	-1.419	-1.108	-1.927	.634	-.810
47	-.931	1.960	-.742	-.030	1.063
48	1.717	.448	.085	.966	1.165
49	.113	.700	.225	-1.028	1.054
50	-.366	-.871	.970	-.719	.176
51	1.626	.451	-1.028	.189	-.548
52	-.665	-.088	.176	.093	1.468
53	-1.011	-1.476	-1.468	-.519	-.391
54	.982	.292	-.327	-.736	-2.409
55	.856	-1.094	.634	-.445	-1.995
56	-.589	.759	-.493	-.827	-.169
57	-.749	-.396	.586	1.499	1.254
58	.462	-.559	1.685	-2.409	-.364
59	-.966	.415	.820	.043	1.282
60	1.706	-1.695	-1.347	.719	.470

Appendix A3: Predicted 60 Years rainfall values (mm) for the Months of May September

	May	June	July	August	September
1	169.865	125.860	15.170	179.620	119.477
2	18.850	90.175	181.468	323.617	223.206
3	52.511	169.100	135.077	382.641	303.206
4	141.301	225.487	335.724	367.980	173.344
5	131.452	-47.128	204.903	176.490	109.401
6	79.415	93.079	226.599	234.397	92.197
7	125.488	141.467	309.213	162.351	17.864
8	32.093	213.296	0.000	244.412	141.860
9	107.466	195.719	169.723	321.833	320.747
10	155.941	186.074	206.229	170.144	207.443
11	14.079	165.578	481.785	357.128	327.159
12	151.544	201.337	118.761	239.536	57.625
13	137.244	253.796	172.485	18.315	7.962
14	74.583	65.383	130.496	357.128	42.122
15	22.921	153.869	213.940	49.113	116.155
16	44.165	124.388	262.752	256.839	152.407
17	176.608	189.728	212.607	411.100	8.623
18	92.910	241.349	177.950	82.282	180.561
19	83.044	173.711	74.033	138.051	161.040
20	103.716	160.220	27.596	194.339	224.214

21	100.815	173.305	101.848	223.319	10.569
22	132.742	145.998	50.777	276.416	187.703
23	230.263	155.723	91.906	162.351	122.264
24	66.422	202.133	125.778	235.000	80.989
25	134.789	14.270	177.950	473.438	230.448
26	42.316	105.861	321.524	231.091	166.023
27	87.670	160.031	368.515	225.706	110.451
28	46.053	55.290	45.466	227.199	151.179
29	99.619	217.111	240.197	145.286	17.741
30	140.909	64.776	331.104	269.403	287.499
31	155.941	65.685	222.258	132.048	142.842
32	22.565	157.588	162.720	128.576	308.811
33	123.006	95.480	178.222	341.984	170.551
34	254.665	185.626	186.036	141.703	199.575
35	164.199	209.387	289.666	288.610	127.035
36	133.483	136.587	279.208	172.061	176.160
37	53.882	217.439	281.306	127.794	243.280
38	76.459	100.145	258.720	315.343	143.087
39	61.712	30.541	88.213	90.811	171.818
40	141.891	73.402	241.048	272.051	254.352
41	169.066	122.726	269.461	7.349	107.289
42	166.978	278.664	271.098	161.691	162.779

43	22.207	209.387	153.495	542.003	229.386
44	125.309	169.893	189.777	306.727	115.384
45	160.560	160.409	203.313	432.481	231.162
46	8.782	58.718	2.210	291.139	66.827
47	41.961	279.878	123.013	212.306	249.936
48	221.998	170.889	210.478	330.539	259.962
49	112.937	189.033	225.238	93.879	249.077
50	80.329	75.813	303.998	130.514	163.277
51	215.802	171.089	92.809	238.323	92.482
52	60.025	132.242	220.101	226.900	289.626
53	36.500	32.179	46.249	154.319	107.819
54	172.040	159.654	166.940	128.576	89.546
55	163.458	59.711	268.486	163.009	49.106
56	65.212	193.277	149.346	117.673	129.527
57	54.334	110.016	263.380	393.731	268.619
58	136.674	98.257	379.558	70.030	110.451
59	39.556	168.508	288.173	220.936	271.355
60	221.261	16.347	59.085	301.240	192.040

Appendix B1: Monthly rainfall predicted values for Normal PDF for 20 years (mm)

Month / Year	Feb	March	April	May	June	July	August	Sept	October	Nov
1	12.811	40.073	54.367	186.917	195.746	116.474	249.703	204.992	114.302	20.102
2	6.409	46.462	49.235	166.109	98.707	199.978	71.079	147.490	73.156	15.628
3	9.403	44.963	31.239	182.769	77.007	303.777	234.155	249.195	36.402	163.606
4	16.167	32.685	23.958	153.393	154.219	278.091	459.541	209.393	49.208	20.450
5	11.824	48.539	85.955	132.654	177.434	158.860	154.753	208.708	61.598	68.440
6	12.594	21.222	65.220	180.865	102.312	333.056	218.251	190.910	36.865	75.994
7	5.237	6.367	31.863	29.728	263.658	124.613	170.776	155.411	15.321	23.124
8	8.192	3.230	1.042	103.074	29.353	73.560	214.334	100.256	98.352	132.291
9	4.514	6.235	27.082	59.487	236.983	83.390	97.665	-22.083	7.878	106.901
10	14.423	34.578	34.117	183.585	37.067	167.951	394.620	109.840	2.016	24.575
11	9.001	36.944	35.227	33.808	15.511	81.699	377.173	143.774	15.968	100.685
12	9.291	53.324	91.711	17.395	10.299	297.646	61.020	-26.777	150.316	75.296
13	15.647	52.798	62.307	91.990	195.746	254.837	268.574	217.021	65.574	143.330
14	1.178	37.207	30.199	122.930	84.505	207.166	194.276	283.911	126.090	191.494
15	10.429	44.990	50.518	90.154	165.610	168.162	211.486	149.641	91.186	36.892
16	4.751	29.451	19.346	32.967	25.748	320.266	225.965	126.660	22.256	44.503
17	17.009	20.819	33.667	40.107	97.049	230.949	71.672	226.409	137.741	22.543
18	15.654	3.028	44.242	24.628	129.491	219.850	81.879	228.560	147.218	26.144
19	1.921	50.564	79.471	32.967	47.232	150.510	119.028	102.407	21.146	87.206
20	8.527	38.601	32.418	77.098	108.944	152.307	367.678	245.967	68.764	8.540

Appendix B2: Monthly rainfall predicted values for Log-Normal PDF for 20 years (mm)

Month / Year	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov
1	12.277	25.524	38.290	238.731	235.859	77.144	207.955	191.122	165.234	11.471
2	3.102	39.175	31.584	179.220	57.239	154.728	36.729	103.811	52.652	10.283
3	5.902	35.429	16.078	225.468	41.704	367.531	178.825	305.543	18.956	382.696
4	25.251	15.552	12.234	150.415	128.674	296.701	1594.086	200.261	27.061	11.569
5	9.931	45.029	125.262	113.025	180.554	109.833	82.741	198.812	38.186	37.385
6	11.717	7.211	57.536	219.630	60.331	469.114	153.245	164.587	19.201	1.095
7	2.411	2.663	16.459	12.062	635.378	82.559	96.664	112.916	10.551	3.988
8	4.550	1.399	4.788	75.190	20.806	53.946	147.529	62.879	106.063	178.014
9	2.064	2.640	1.802	41.240	430.514	58.552	47.542	17.162	8.579	95.710
10	17.360	17.657	17.911	228.018	23.284	118.477	848.875	69.612	6.516	12.797
11	5.414	20.694	18.672	11.403	17.000	57.733	716.636	99.796	10.742	82.218
12	5.762	62.064	155.458	23.090	11.665	349.223	10.190	16.328	449.608	44.205
13	22.582	59.914	51.580	64.540	235.859	244.424	249.757	217.150	42.648	233.145
14	0.607	21.062	15.462	98.852	46.525	164.280	121.430	441.678	229.301	756.604
15	7.359	35.492	33.141	62.928	151.942	118.686	143.506	106.209	86.909	17.292
16	0.282	12.521	10.290	28.616	19.739	421.679	165.161	83.219	12.793	20.827
17	30.261	0.430	17.611	31.575	55.871	200.297	36.941	239.903	316.985	4.045
18	22.614	2.129	26.188	12.940	89.698	182.599	40.788	245.445	412.518	13.297
19	1.182	51.576	98.211	28.616	27.008	102.448	58.497	64.332	12.405	59.142
20	4.890	23.124	16.805	52.567	66.462	103.994	653.543	295.254	46.602	8.647

Appendix B3: Monthly rainfall predicted values for Exponential PDF for 20 years (mm)

Month / Year	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov
1	3.307	5.388	17.505	12.265	32.665	314.405	106.707	45.957	3.201	41.193
2	10.279	3.595	20.758	22.178	171.539	143.710	476.500	98.328	13.883	43.498
3	6.243	4.091	38.263	14.657	223.028	37.538	125.169	23.732	37.231	0.815
4	1.507	7.951	46.302	28.884	73.117	52.654	4.361	43.970	27.182	40.088
5	3.994	3.109	4.281	43.733	49.426	217.335	259.910	43.970	19.687	14.961
6	3.440	13.882	11.152	14.657	162.297	23.477	145.360	56.318	37.231	157.733
7	12.139	25.521	37.234	411.740	5.657	296.078	226.632	89.981	59.424	79.274
8	7.825	35.353	89.694	75.080	389.869	444.674	153.996	166.390	5.639	2.495
9	13.626	25.521	142.492	145.907	13.069	411.019	395.612	470.049	68.846	5.615
10	2.304	7.353	34.295	13.454	350.003	194.928	15.666	149.191	83.689	37.966
11	6.805	6.488	33.361	411.740	446.056	411.019	20.360	104.170	59.424	6.553
12	6.426	2.403	2.949	242.347	542.109	39.980	994.150	470.049	0.443	12.689
13	1.728	2.403	12.789	91.304	32.665	74.754	86.454	38.167	17.667	1.646
14	26.084	6.208	39.319	53.764	203.661	131.739	182.194	12.176	1.801	33.102
15	5.212	4.091	20.086	93.841	59.696	194.928	153.996	98.328	7.691	29.594
16	36.270	9.526	53.206	206.933	389.869	28.056	137.056	126.680	51.665	25.598
17	1.185	58.003	35.252	186.498	176.401	99.580	476.500	34.422	0.891	79.274
18	1.728	28.581	25.039	369.065	110.653	113.244	440.437	32.585	0.443	36.945
19	18.916	2.871	5.659	206.933	319.081	229.549	336.908	161.897	53.112	9.504
20	7.403	5.931	36.231	113.545	149.496	229.549	22.745	23.732	15.734	49.911

Appendix C: Statistical Tests Results for Return Periods

PDFs Values for the Month of July (mm)

T	Observed	Normal Predicted	Log-normal predicted	Exponential predicted
2	201.46	196.28	169.02	136.05
5	290.42	279.48	294.56	315.91
10	336.92	322.97	393.79	451.96
15	359.85	344.41	454.42	531.55
20	375.32	358.88	500.50	588.01
30	395.78	378.01	568.70	667.60
40	408.63	390.03	616.20	724.07
50	418.54	399.30	655.55	767.87

ANOVA of Return Period for the Month of July

	Sum of Squares	df	Mean Square	F	Sig.
Between PDF's	195372.320	3	65124.107	3.119	.042
Within PDF's	584579.856	28	20877.852		
Total	779952.176	31			

Turkey HSD Multiple Comparisons Post-HOC test for the Month of July

(I) PDFs	(J) PDFs	Mean Difference (I-J)	Std. Error	Sig.
Observed IAR	Normal Predicted	14.695	72.25	.840
	Log-normal Predicted	-108.227	72.25	.145
	Exponential Predicted	-174.512*	72.25	.022
Normal Predicted	Observed IAR	-14.695	72.25	.840
	Log-normal Predicted	-122.922	72.25	.100
	Exponential Predicted	-189.207*	72.25	.014
Log-normal Predicted	Observed IAR	108.227	72.25	.145
	Normal Predicted	122.922	72.25	.100
	Exponential Predicted	-66.285	72.25	.367
Exponential Predicted	Observed IAR	174.512*	72.25	.022
	Normal Predicted	189.207*	72.25	.014
	Log-normal Predicted	66.285	72.25	.367

*. The mean difference is significant at the 0.05 level.

PDF Values for the Month of August

T	Observed	Normal Predicted	Log-normal predicted	Exponential predicted
2	215.88	222.53	208.51	154.24
5	315.77	325.57	334.06	358.14
10	367.98	379.43	427.38	512.39
15	393.73	405.99	482.59	602.61
20	411.10	423.91	523.80	666.63
30	434.07	447.61	583.78	756.85
40	448.50	462.49	624.89	820.87
50	459.63	473.97	658.59	870.53

ANOVA of Return Period for the Month of August

	Sum of Squares	df	Mean Square	F	Sig.
Between PDF's	230694.272	3	76898.091	3.193	.039
Within PDF's	674425.101	28	24086.611		
Total	905119.373	31			

Turkey HSD Multiple Comparisons Post-HOC test for the Month of August

(I) PDF's	(J) PDF's	Mean Difference (I-J)	Std. Error	Sig.
Observed IAR	Normal Predicted	-11.854	77.60	.880
	Log-normal Predicted	-99.618	77.60	.210
	Exponential Predicted	-211.951*	77.60	.011
Normal Predicted	Observed IAR	11.854	77.60	.880
	Log-normal Predicted	-87.764	77.60	.268
	Exponential Predicted	-200.097*	77.60	.015
Log-normal Predicted	Observed IAR	99.618	77.60	.210
	Normal Predicted	87.764	77.60	.268
	Exponential Predicted	-112.333	77.60	.159
Exponential Predicted	Observed IAR	211.951*	77.60	.011
	Normal Predicted	200.097*	77.60	.015
	Log-normal Predicted	112.333	77.60	.159

*. The mean difference is significant at the 0.05 level.

PDF Values for the Month of September

T	Observed	Normal Predicted	Log-normal predicted	Exponential predicted
2	146.03	156.86	135.64	108.73
5	228.33	236.05	272.75	252.46
10	271.36	277.44	392.95	361.19
15	292.57	297.86	470.49	424.80
20	306.88	311.63	531.25	469.92
30	325.81	329.84	623.84	533.53
40	337.70	341.28	690.05	578.65
50	346.87	350.10	745.91	613.66

ANOVA of Return Period for the Month of September

	Sum of Squares	df	Mean Square	F	Sig.
Between PDF's	236346.798	3	78782.266	3.841	.020
Within PDF's	574294.670	28	20510.524		
Total	810641.468	31			

Turkey HSD Multiple Comparisons Post-HOC test for the Month of September

(I) PDF's	(J) PDF's	Mean Difference (I-J)	Std. Error	Sig.
Observed IAR	Normal Predicted	-5.687	71.61	.937
	Log-normal Predicted	-200.915*	71.61	.009
	Exponential Predicted	-135.922	71.61	.068
Normal Predicted	Observed IAR	5.687	71.61	.937
	Log-normal Predicted	-195.228*	71.61	.011
	Exponential Predicted	-130.235	71.61	.080
Log-normal Predicted	Observed IAR	200.915*	71.61	.009
	Normal Predicted	195.228*	71.61	.011
	Exponential Predicted	64.993	71.61	.372
Exponential Predicted	Observed IAR	135.922	71.61	.068
	Normal Predicted	130.235	71.61	.080
	Log-normal Predicted	-64.993	71.61	.372

*. The mean difference is significant at the 0.05 level.