MODELLING ANNUAL MAXIMUM FLOOD OF RIVER KADUNA

BY

ABDULLAHI ABDULLAHI

B.Eng. WREE (A.B.U)

P16EGWR8022

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

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER DEGREE IN WATER RESOURCES AND ENVIRONMENTAL ENGINEERING

DEPARTMENT OF WATER RESOURCES AND ENVIRONMENTAL ENGINEERING,
FACULTY OF ENGINEERING,
AHMADU BELLO UNIVERSITY,
ZARIA, NIGERIA

APRIL, 2021

DECLARATION

I, ABDULLAHI ABDULLAHI, hereby declare that this thesis titled "**MODELLING ANNUAL MAXIMUM FLOOD OF RIVER KADUNA**" has been carried out by me under the supervision of Prof. M. A Ajibike and Dr. B. K. Adeogun. Hence, it has not been presented for award of any degree in any institution. All sources of information are specifically acknowledged by means of reference

…… ……

ABDULLAHI, Abdullahi **Date**

CERTIFICATION

This thesis titled "**MODELLING ANNUAL MAXIMUM FLOOD OF RIVER KADUNA**" submitted by ABDULLAHI ABDULLAHI meets the requirements governing the award of the degree of Master of Science in Water Resources and Environmental Engineering Department of Ahmadu Bello University Zaria and is approved for its contribution to knowledge and literary representation.

…….. **……**

Prof. M. A. Ajibike **Date**

Chairman, Supervisory Committee

…….. **……**

Dr. B. K. Adeogun **Date**

Member, Supervisory Committee

…….. **……**

Prof. M. A. Ajibike **Date** Head of Department

…….. **……**

Prof. Sani A. Abdullahi **Date** Dean Postgraduate School

ACKNOWLEDGEMENT

In the name of ALLAH, the beneficent the most merciful

All praise is due to ALLAH, the lord of the worlds

The beneficent, the merciful

Master of the Day of Judgment

Thee do we serve and Thee do we beseech for help.

Keep us on the right path.

The path of those upon whom Thou hast bestowed favors. Not (the path) of those upon whom Thy wrath is brought down, nor of those who go astray.

My sincere gratitude goes to my humble parents whom had taken care of me since the time I was born up to this time or moment, just to make them upspring happy and give the adequate knowledge he needs from them, praying for him day and night to be among good and excellent people and to acquire the best character among others. I pray to Allah to reward them abundantly and may he make jannatul firdaus their final abode.

My brother and sister's, I really appreciate your efforts in terms of prayers, finance and advices that done to me, nothing much to say to you but to pray for you for a better life and have good family members. May almighty Allah bless you and your family.

I solemnly thank my patient supervisors, mentors, teachers, parent in person of Prof. M.A. Ajibike and Dr. B. K. Adeogun for spending their time to make sure we did the right thing. May almighty Allah increase them in knowledge and uplift them to the highest level. Hence, my able mentor, late father in person of Dr. A. W. Alayande who assisted me with secondary data of this research work. May almighty ALLAH bestow you with his doors of success.

The entire lecturers in the department, I'm sending my warm gratitude for being my support to become not only a graduate but student of knowledge that is groomed and had the capacity to compete with world professionals. You are really in my heart and will never forget you, where ever I find myself. Furthermore, I'm proud to be your student and never regret

My able Dean too “Prof. Abubakar Ismail, memory will never forget you, May almighty Allah bestow rahmah onto you. The departmental PG Coordinators also, Dr. S.B Saulawa and Dr. S. A. Argungu thank you for being our support, I say ***Jazakumullahu bi khairan wa bil Jannah*** to you.

ABSTRACT

Records of flooding are increasing due to heavy storm which tend to filled most of the water bodies, thereby overflowing onto the land, causing damages and endangering the lives people. Flood also occur due to other factors such dam break, overflowing rivers etc. Most of this event occur with inadequate capacity of reservoirs, channels, drainages etc to store or convey the water appropriately Sediments transport due to heavy rain also contributed so much in reducing the capacity of most of the water bodies. Hence, water bodies tend to experience siltation as a result of this flood. By considering effect of flood, studies must be carried out to determine its magnitude for proper flood design. The study intended to employ the concept of frequency-based flood as one of the common methods of flood design. This method aimed at estimating the annual maximum flood of adequate length within a case study. Several methods of parameter estimate for flood frequency analysis evolves and determination of frequency factor of various distributions helps a lot to fit in the probability of exceedance into the distribution model through the statistical parameter of the distribution. The probability of exceedance gives the recurrence interval of flood. The magnitude of this flood must be studied as well as knowing the reliable probability distribution and plotting position to be used. Employing the principle of flood frequency analysis will be of help for determination of such magnitude. This study went by using secondary discharge data which was gotten form National Water Resources Institute, Kaduna. The data was utilized by MATLAB software. For the purpose of this research, 30years streamflow data was used of Kaduna State Water Works Gauging Station using MATLAB Software for the statistical and programming analysis. Four probability distributions were matched using eight various number of plotting positions which result or determine predicted discharges at various probability of exceedance. The statistical measurement was based on the concept of Goodness of Fit (GOF) in which Coefficient of determination (R-Square) and Root mean square error (RMSE). In the end, it was found that the best suitable plotting position for the catchment was determined to be Modified California, having the highest Adjusted R-square as 0.9604, 0.9622, 0.9728 and 0.9584 as well as validated RMSE value as 128.79, 175.39, 107.04 and 130.56 for normal, lognormal, pearson type III and log-pearson type III respectively. Furthermore, the best distribution that suit the catchment was discovered to be Pearson type III with Adjusted R-square and Validated RMSE value of 0.9728 and 107.04 respectively. Finally, the magnitude of the flood at different recurrence interval of 25, 50 and 100 years were

determined to be 2838.60, 2936 and 2970.70 m^3/sec using the exponential model of Pearson Type III distribution.

TABLE OF CONTENT

Title Page i

Declaration ii

Certification iii

Acknowledgement iv

Abstract vi

Table of Content viii

List of Equation xv

List of Tables xviii

List of Figures xix

Appendices xxi

CHAPTER ONE 1

1.0 INTRODUCTION 1

1.1 Preamble 1

1.2 Statement of Problem 1

1.3 Justification 2

1.4 Aim and Objectives 3

1.5 Scope of Work 3

CHAPTER TWO 4

2.0 LITERATURE REVIEW 4

2.1 Background Review 4

2.2 Flooding 5

2.2.1 Common Types of Flood 6

2.2.1.1 Coastal (Surge Flood) 6

2.2.1.2 Fluvial (River Flood) 7

2.2.1.3 Pluvial (Surface Flood) 7

2.2.2 Causes of Floods. 8

2.2.3 Effects of Floods 10

2.2.4 Ways to Prevent Flood. 10

2.2.4.1 Improving Drainage *10*

2.2.4.2 Building Dikes and Levees *11*

2.2.4.3 Building Canals. *11*

2.2.4.4 Harvesting Rain Water. *11*

2.3 Methods of Parameter Estimation. **11**

2.3.1 Graphical Method. 11

2.3.2 Frequency Factor Method. 12

2.3.3 Method of Moments (MOM). 13

2.3.4 Least Square Method (LSM). 14

2.3.5 Probability Weighted Moments (PWM). 14

2.36 Maximum Likelihood Estimation (MLE) Method. 17

2.3.5 L- Moments. 17

2.3.5.1 Graphical Depictions of L-Cv and L-Skewness *19*

2.4 Probability Distribution. **20**

2.4.1 Normal Family: Normal, Lognormal and Lognormal III 20

2.4.1.1 Characteristics of Normal Distribution. *22*

2.4.1.2 Standardized Normal Distribution. *22*

2.4.2 Pearson Type III Family: Log-Pearson Type III and Pearson Type III 22

2.4.3 GEV Family: Gumbel, GEV and Weibull 25

2.4.4 Principle and Issues in Selecting a Distribution 25

2.5 Probability Plots 27

2.5.1 Plotting Positions 27

2.6 Goodness of Fit (GOF) 30

2.6.1 Goodness of Fit Test 30

2.6.1.1 Chi-square. *31*

2.6.1.2 Kolmogorov-Smirnov. *31*

2.6.1.3 Anderson-Darling *32*

2.6.1.4 Shapiro-Wilk *32*

2.6.1.5 Probability Plot Correlation Coefficient (PPCC) *32*

2.6.1.6 Coefficient of Determination (R_Squared) *33*

2.6.1.7 Root mean square error (RMSE) *34*

2.6.2 L-Moment Diagram 34

2.7 Standard Errors and Confidence Interval for Quantiles 35

2.7.1 Confidence Interval 35

2.8 Matrix Laboratory (MATLAB) 36

2.8.1 Overview 36

2.8.2 MATLAB System 38

2.8.2.1 MATLAB language. *38*

2.8.2.2 MATLAB working environment. *38*

2.8.2.3 Handle Graphics. *38*

2.8.2.4 MATLAB mathematical function library. *38*

2.8.2.5 MATLAB Application Program Interface (API). *38*

2.8.3 Features of MATLAB. 39

2.8.4 MATLAB power of Computational Mathematics. 40

2.8.5 MATLAB program files. 41

CHAPTER THREE **31**

3.0 MATERIALS AND METHODS

3.1 Overview **42**

3.1.1 Description of the Study Area. 42

3.2 Materials. **43**

3.3 Method **44**

3.3.1 Normality Test using Shapiro Wilk Approach 46

3.3.2 Probability of Exceedance 46

3.3.3 Normal Distribution 48

3.3.4 Log-Normal Distribution 49

3.3.5 Pearson Type III Distribution 50

3.3.6 Log-Pearson Type III Distribution 51

3.4 Probability Distribution Curve. **52**

3.5 Determination of Goodness of Fit (GOF) 53

3.6 Determination of Predicted Discharge. 54

CHAPTER FOUR 55

4.0 RESULTS AND DISCUSSIONS 55

4.1 Hydrograph 55

4.2 Normal Distribution 56

4.2.1 California plotting position 56

4.2.2 Weibull plotting position. 57

4.2.3 Chegodayev plotting position. 58

4.2.4 Turkey plotting position. 59

4.2.5 Blom plotting position 60

4.2.6 Cunnane plotting position 61

4.2.7 Gringorten plotting position 62

4.2.8 Hazen plotting position. 63

4.3 Log-Normal Distribution. 64

4.3.1 California plotting position. 64

4.3.2 Weibull plotting position. 65

4.3.3 Chegodayev plotting position. 66

4.3.4 Turkey plotting position. 67

4.3.5 Blom plotting position 68

4.3.6 Cunnane plotting position. 69

4.3.7 Gringorten plotting position. 70

4.3.8 Hazen plotting position. 71

4.4 Pearson Type III Distribution. 72

4.4.1 California plotting position 72

4.4.2 Weibull plotting position. 73

4.4.3 Chegodayev plotting position. 74

4.4.4 Turkey plotting position. 75

4.4.5 Blom plotting position. 76

4.4.6 Cunnane plotting position. 77

4.4.7 Gringorten plotting position. 78

4.4.8 Hazen plotting position. 79

4.5 Log-Pearson Type III Distribution 80

4.5.1 California plotting position 80

4.5.2 Weibull plotting position. 81

4.5.3 Chegodayev plotting position. 82

4.5.4 Turkey plotting position. 83

4.5.5 Blom plotting position. 84

4.5.6 Cunnane plotting position. 85

4.5.7 Gringorten plotting position. 86

4.5.8 Hazen plotting position. 87

4.6 Discussion Summary. 88

4.6.1 Normal Distribution. 88

4.6.2 Log-Normal Distribution. 88

4.6.3 Pearson Type III Distribution. 88

4.6.4 Log-Pearson Type III Distribution. 89

4.7 Flood Magnitude. 89

CHAPTER FIVE 92

5.0 CONCLUSIONS AND RECOMMENDATIONS 92

5.1 Conclusions 92

5.2 Recommendations 93

REFERENCES 94

APPENDICES 98

LIST OF EQUATIONS

Equation 2.1	General frequency equation	12
Equation 2.2	Method of Moments	13
Equation 2.3	Theoretical moments	14
Equation 2.4	Central Moment (Variance)	14
Equation 2.5	Least Square Method	14
Equation 2.6	Probability Distribution Function for Probability Weighted Moments	14
Equation 2.7a	Conventional Moment about the origin of PWM order I	15
Equation 2.7b	Conventional Moment about the origin of PWM order II	15
Equation 2.7c	Conventional Moment about the origin of PWM order III	15
Equation 2.8	Moment about the PWM origin	15
Equation 2.9	Moment about the PWM origin of the first order	15
Equation 2.10	L-moment mean	18
Equation 2.11	L-moment scale	18
Equation 2.12	L-moment Skewness	18
Equation 2.13	L-moment Kurtosis	18
Equation 2.14a	L-moment equation I	18
Equation 2.14b	L-moment equation II	18
Equation 2.14c	L-moment equation III	18
Equation 2.14d	L-moment equation IV	18
Equation 2.15a	Probability Weighted Moments equations I	18
Equation 2.15b	Probability Weighted Moments equations II	18
Equation 2.15c	Probability Weighted Moments equations III	18

Equation 2.15d	Probability Weighted Moments equations IV	18
Equation 2.16	Normal Probability Distribution Function	20
Equation 2.17	Probability of Exceedance	20
Equation 2.18	Intermediate Variable	20
Equation 2.19	Statistical Variate	20
Equation 2.20	Mean Discharge determination	20
Equation 2.21	Standard Deviation	20
Equation 2.22	Log-normal Probability Distribution Function	21
Equation 2.23	Logarithmic of the observed data	21
Equation 2.24	Mean of Logarithmic of the observed data	21
Equation 2.25	Standard Deviationwith respect to y	21
Equation 2.26	Statistical variate with respect to y	22
Equation 2.27	Predicted discharge with respect to y	22
Equation 2.28	Pearson Type III Probability Distribution Function	23
Equation 2.29	Pearson Type III Cumulative Distribution Function	23
Equation 2.30	Statistical parameter with respect to coefficient of skewness	23
Equation 2.31	Frequency factor with respect to Pearson Type III family	23
Equation 2.32	Coefficient of Skewness.	23
Equation 2.33	Log- Pearson Type III Probability Distribution Function.	24
Equation 2.34	Coefficient of Skewness with respect to Log- Pearson Type Distribution	24
Equation 2.35	General expression for plotting position.	27
Equation 2.36	Shapiro Wilk test	32

Equation 2.37 Coefficient of Shapiro Wilk data 32

Equation 2.38 Vector for Shapiro Wilk test. 32

Equation 2.39 Correlation Coefficient. 33

Equation 2.40 Statistic median for PPCC. 33

Equation 2.41 Median value I 33

Equation 2.42 Median value II 33

Equation 2.43 Median value III 33

Equation 2.44 Ordinary R-Square 34

Equation 2.45 Adjusted R-Square 34

Equation 2.46 Root Mean Square Error. 34

Equation 2.47 Standard error for confidence interval 36

Equation 2.48 Confidence Interval 36

LIST OF TABLES

Table 2.1 Frequency factor for different distributions 13

Table 2.2 Probability distribution parameters in relation to sample moments 16

Table 2.3 General Descriptions of Relative Magnitude of L-Cv 19

Table 2.4 General Descriptions of Relative Magnitude of L-Skewness. 19

Table 2.5 Plotting position Formulae. 29

Table 2.6 Line colour option in MATLAB software. 39

Table 2.7 Line style using MATLAB software command. 40

Table 3.1 Flow Regime of Kaduna River at Kaduna South 1971 to 2000 (NWRI, 2017). 45

Table. 4.1 Response of the distribution at various return period with respect to the plotting positions 90

Table. 4.2 Response of the distribution at various return period with respect to the plotting positions Cont'd 91

LIST OF FIGURES

Figure 3.1	Location of the study area on satellite imagery.	43
Figure 3.2	Flow Chart for the flood study.	44
Figure 3.3	Normality test using Shapiro Wilk Approach	46
Figure3.4	Plotting Position Determination.	47
Figure 3.5	MATLAB steps for Normal Distribution	49
Figure 3.6	MATLAB steps for Log-Normal Distribution	50
Figure 3.7	MATLAB steps Pearson Type III Distribution	51
Figure 3.8	MATLAB steps for Log-Pearson Type III.	52
Figure 4.1	Annual hydrograph of River Kaduna.	55
Figure 4.2	Normal Distribution Curves using California plotting position.	56
Figure 4.3	Normal Distribution Curves using Weibull plotting position	57
Figure 4.4	Normal Distribution Curves using Chegodayev plotting position.	58
Figure 4.5	Normal Distribution Curves using Turkey plotting position.	59
Figure 4.6	Normal distribution curves using Blom plotting position.	60
Figure 4.7	Normal distribution curves using Cunnane plotting position.	61
Figure 4.8	Normal distribution curves using Gringorten plotting position.	62
Figure 4.9	Normal distribution curve using Hazen plotting position.	63
Figure 4.10	Log-normal distribution curves using California plotting position.	64
Figure 4.11	Log-normal distribution curves using Weibull plotting position.	65
Figure 4.12	Log-normal distribution curves using Chegodayev plotting position.	66
Figure 4.13	Log-normal distribution curves using Turkey plotting position	67

Figure 4.14 Log-normal distribution curves using Blom plotting position. 68

Figure 4.15 Log-normal distribution curves using Cunnane plotting position 69

Figure 4.16 Log-normal distribution curves using Gringorten plotting position. 70

Figure 4.17 Log-normal distribution curves using Hazen plotting position. 71

Figure 4.18 Pearson type III distribution curves California plotting position. 72

Figure 4.19 Pearson type III distribution curves Weibull plotting position. 73

Figure 4.20 Pearson type III distribution curves Chegodayevs plotting position. 74

Figure 4.21 Pearson type III curves distribution Turkey plotting position. 75

Figure 4.22 Pearson type III curves distribution using Blom plotting position. 76

Figure 4.23 Pearson type III distribution curves Cunnane plotting position. 77

Figure 4.24 Pearson type III Distribution Gringorten plotting position. 78

Figure 4.25 Pearson type III Distribution Hazen plotting position 79

Figure 4.26 Log-Pearson type III distribution curve using California plotting position. 80

Figure 4.27 Log-Pearson type III distribution curves using Weibull plotting position. 81

Figure 4.28 Log-Pearson type III distribution curves using Chegodayev plotting position 82

Figure 4.29 Log-Pearson type III distribution curves using Turkey plotting position. 83

Figure 4.30 Log-Pearson type III distribution curves using Blom plotting position. 84

Figure 4.31 Log-Pearson type III distribution curves using Cunnane plotting position. 85

Figure4.32 Log-Pearson type III distribution curves using Gringorten plotting position. 86

Figure 4.33 Log-Pearson type III distribution curves using Hazen plotting position. 87

APPENDICES

Appendix I Normal Distribution using Probability of Exceedance. 98

Appendix II Normal Distribution using Probability of Exceedance Cont'd. 99

Appendix III Log- Normal Distribution using Probability of Exceedance. 100

Appendix IV Log- Normal Distribution using Probability of Exceedance Cont'd. 101

Appendix V Pearson Type III Distribution using Probability of Exceedance. 102

Appendix VI Pearson Type III Distribution using Probability of Exceedance Cont'd. 103

Appendix VII Log-Pearson Type III Distribution using Probability of Exceedance. 104

Appendix VIII Log-Pearson Type III Distribution using Probability of Exceedance Cont'd. 105

Appendix IX Exponential Model and MATLAB steps for normal distribution on how to plot curves using the respective plotting position formulae as titled to each of the plate below. 106

Plate 1. California. 106

Plate 2. Weibull. 106

Plate 3. Chegodeyev. 107

Plate 4. Turkey. 107

Plate 5. Blom. 108

Plate 6. Cunnane. 108

Plate 7. Gringorten. 109

Plate 8. Hazen. 109

Appendix X Exponential Model and MATLAB steps for Log-normal distribution on how to plot curves using the respective plotting position formulae as titled to each of the plate below. 110

Plate 9. California. 110

Plate 10. Weibull 110

Plate 11. Chegodeyev. 111

Plate 12. Turkey 111

Plate 13. Blom. 112

Plate 14. Cunnane 112

Plate 15. Gringorten. 113

Plate 16. Hazen. 113

Appendix XI. Exponential Model and MATLAB steps for Pearson Type III distribution on how to plot curves using the respective plotting position formulae as titled to each of the plate below. 114

Plate 17. California. 114

Plate 18. Weibull 114

Plate 19. Chegodeyev. 115

Plate 20. Turkey 115

Plate 21. Blom. 116

Plate 22. Cunnane 116

Plate 23. Gringorten. 117

Plate 24. Hazen. 117

Appendix XII. Exponential Model and MATLAB steps for Log-Pearson Type III distribution on how to plot curves using the respective plotting position formulae as titled to each of the plate below. 118

Plate 25. California. 118

Plate 26. Weibull 118

Plate 27. Chegodeyev. 119

Plate 28. Turkey 119

Plate 29. Blom. 120

Plate 30. Cunnane 120

Plate 31. Gringorten. 121

Plate 32. Hazen. 121

Appendix XIII. Results of R-Square and RMSE values for Normal Distribution in respect of the below plotting positions. 122

Plate 33a. California using Observed data 122

Plate 33b. California using Predicted data 122

Plate 34a. Weibull using Observed data. 122

Plate 34b. Weibull using Predicted data. 122

Plate 35a. Chegodayev using Observed data. 123

Plate 35b. Chegodayev using Predicted data. 123

Plate 36a. Turkey using Observed data. . 123

Plate 36b. Turkey using Predicted data. 123

Plate 37a. Blom using Observed data. 124

Plate 37b. Blom using Predicted data. 124

Plate 38a. Cunnane using Observed data. 124

Plate 38b. Cunnane using Predicted data. 124

Plate 39a. Gringorten using Observed data. 125

Plate 39b. Gringorten using Predicted data. 125

Plate 40a. Hazen using Observed data. 125

Plate 40b. Hazen using Predicted data. . 125

Appendix XIV. Results of R-Square and RMSE values for Log- Normal Distribution in respect of the below plotting positions. 126

Plate 41a. California using Observed data. 126

Plate 41b. California using Predicted data. 126

Plate 42a. Weibull using Observed data. 126

Plate 42b. Weibull using Predicted data. 126

Plate 43a. Chegodayev using Observed data**.** 127

Plate 43b. Chegodayev using Predicted data. 127

Plate 44a. Turkey using Observed data. 127

Plate 44b. Turkey using Predicted data. 127

Plate 45a. Blom using Observed data. 128

Plate 45b. Blom using Predicted data. 128

Plate 46a. Cunnane using Observed data. 128

Plate 46b. Cunnane using Predicted data. 128

Plate 47a. Gringorten using Observed data**.** 129

Plate 47b Gringorten using Predicted data. 129

Plate 48a. Hazen using Observed data 129

Plate 48b. Hazen using Predicted data. 129

Appendix XV Results of R-Square and RMSE values for Pearson Type III Distribution in respect of the below plotting positions. 130

Plate 49a. California using Observed data. 130

Plate 49b. California using Predicted data. 130

Plate 50a. Weibull using Observed data. 130

Plate 50b. Weibull using Predicted data. 130

Plate 51a. Chegodayev using Observed data. 131

Plate 51b. Chegodayev using Predicted data. 131

Plate 52a. Turkey using Observed data. 131

Plate 52b. Turkey using Predicted data. 131

Plate 53a. Blom using Observed data. 132

Plate 53b. Blom using Predicted data. 132

Plate 54a. Cunnane using Observed data. 132

Plate 54b. Cunnane using Predicted data. 132

Plate 55a. Gringorten using Observed data. 133

Plate 55b Gringorten using Predicted data. 133

Plate 56a. Hazen using Observed data. 133

Plate 56b. Hazen using Predicted data. 133

Appendix XVI Results of R-Square and RMSE values for Log-Pearson Type III Distribution in respect of the below plotting positions. 134

Plate 57a. California using Observed data 134

Plate 57b. California using Predicted data. 134

Plate 58a. Weibull using Observed data. 134

Plate 58b. Weibull using Predicted data. 134

Plate 59a. Chegodayev using Observed data. 135

Plate 59b. Chegodayev using Predicted data. 135

Plate 60a. Turkey using Observed data. 135

Plate 60b. Turkey using Predicted data. 135

Plate 61a. Blom using Observed data. 136

Plate 61b. Blom using Predicted data. 136

Plate 62a Cunnane using Observed data. 136

Plate 62b. Cunnane using Predicted data. 136

Plate 63a. Gringorten using Observed data. 137

Plate 63b Gringorten using Predicted data. 137

Plate 64a. Hazen using Observed data 137

Plate 64b. Hazen using Predicted data. 137

Appendix XVII Summary for Normal Distribution GOF result 138

Appendix XVIII Summary for Log-normal Distribution GOF result 138

Appendix XIX Summary for Pearson Type III Distribution GOF result 139

Appendix XX Summary for Log-Pearson Type III Distribution 139

Appendix XXI Summary of GOF for the Observed Data 140

Appendix XXII Summary of the Adjusted R_square values 141

Appendix XXIII Summary of Validated RMSE values 141

CHAPTER ONE
INTRODUCTION

1.1 Preamble.

Flooding is generally a temporary overflowing of water onto adjoiningland that is normally dry. Flooding may happen with only a few inches or centimeters of water, or it may cover a house to the rooftop. There are many possible causes of floods including heavy rain or snowmelt, coastal storms and storm surge, blocked waterway, or overflow of levees, dams, or waste water systems. Flooding usually occur slowly over many days or happen very quickly with little or no warning, called flash floods (Doswell, 2003).

Flood frequency analysis has been one of the methods to estimates floods using graphical, frequency factor, method of moments, probability weighted moments, L-moments, Maximum likelihood estimation and least square method. Determination of probability of exceedance of event depends on recurrence internal between events. To determine the probability using the frequency factor method, the concept of plotting position is use so as to determine the probability of exceedance or non-exceedance at various recurrence interval. Several formulae for plotting positions were available and need to make analytical comparison to conclude the best suitable plotting positions as well as best probability distribution model which suit the study catchment.

1.2 Statement of Problem.

The occurrence of flood continues due to heavy stormwater which filled most of the water bodies and drainages, water tend to overflow thereby causing damages to people's properties as well endangering the lives of human and other inhabitants.Furthermore, such event is the primary cause of soil erosion which result in sediments transport thereby displacingthe upper soil layer. More so, in terms of economic gain, flooding tends to cause damages to the society especially the rural dwellers where they solemnly depend on the cultivated farmlands.Looking at financial implication, People invest hugely in their businesses, millions of money or wealth were missing, all these are as a result of flooding.

Additionality, flood usually occur due to heavy rain, dam break, overflowing rivers, urban drainages, storm surges, tsunami, channels with steep sides and lack of vegetations.This event has been a threat to the society thereby needs an urgent attention in termsof determining the magnitudes and way out to control such flooding. However, several studies have been carried out for flood frequency analysis to determine the magnitude and effect of flood on the environment, estimating its magnitude and likely occurrence in near or far away.

Earlier research has focused on three (3) of the main distributions as well as three (3) plotting position for each which means that their chances or probability of getting the best plotting position as well as the distribution is very low. Hence, they usually employ single test for the determination of goodness of fit for the distributions.Soalso,the tool usually used for their frequency analysis were the useof an average performance technical computing software which have in terms of mathematical computations, algorithm development, modelling, simulations and optimizations.Furthermore, the best distribution model as well as plotting positions need to be determined for accurate predictions of flood events.

1.3 Justification.

Flood has posed serious issues to our environment, the strategies that will be of help to determine the quantity with reference to the time that it occur must be studied. Several method of flood distribution analyses had been employed since, in order to determine the rate at which the flood occurred.

In this study, its intended to experiment the use of MATLAB programming software for computational analysis so as to estimate the maximum flood that will likely occur at various recurrence interval. Chow (1964) describe frequency-based flood as one (1) of the best methods of flood design in estimating the annual flood events of adequate length. This involves but not limited to the use of frequency factor method which aimed at estimating the different peak flood series at certain recurrence interval, it may be achieved by matching eight (8) different plotting positions with four (4) fitted distribution models using MATLAB software.

The study will help the government to know the magnitude of flood to occur at various return period of 25, 50 and 100 years. Hence, knowing this by the stakeholder’s adequate measures will

be put in place so as to plan for the future properly in order to prevent or minimize such event occurring.

The study will also help researchers, hydrological site engineers, statisticians etc to learn on how to compute stream flow data using MATLAB software and also to determine maximum flood of the catchment. Hence, suitable plotting position as well as best distribution model will be known for hydrological analysis within the catchment.

1.4 Aim and Objectives.

Aim.

The aim of this study is to modelledannual maximum flood of River Kaduna using MATLAB software.

Objectives.

i. To determine the best plotting position suitable for the catchment.
ii. To determine the best probability distribution that suit the catchment
iii. To determine the magnitude of the flood at various return period.

1.5 Scope of Work.

The study centered on determining the best plotting position suitable for River Kaduna, best probability distribution model that suit the catchment as well as to determine the magnitude of flood event at various return period. Its intended to adopt the concept of statistical approach of probability distribution using frequency factor parameter estimation method. The modelling will employ the application of MATLAB programming software throughout the study. The intended distribution models for this study are Normal, Log-Normal, Pearson Type III and Log-Pearson Type III distribution, so also, the plotting positions are California, Weibull, Chegodayev, Turkey, Blom, Cunnane, Gringorten and Hazen. Hence, to determine the Goodness of fit, coefficient of determination (R-Square) and Root mean square error (RMSE) is to be employed for accurate decision making.

CHAPTER TWO

LITERATURE REVIEW

2.1 Background Review.

Ewemoje and Ewemooje (2011) employed 3 types of distribution (i.e Normal, Lognormal, and Log-Pearson type III distributions) for modelling site annual maximum flood using the Hazen, Weibull, and California as plotting positions at Ogun-Oshun river basin in Nigeria. His research shows that the three (3) probability distributions when matched with Weibull plotting position gave almost same values near the center of the distribution but varied considerably in the tails. Furthermore, he added that the Weibull plotting position when matched with Normal, Log-normal and Log Pearson Type III probability distributions gave the highest R-Square value of 0.967, 0.987, and 0.986 respectively. Likewise, Hazen plotting position gave minimal errors with the RMSE of 6.988, 6.390, and 6.011 for Normal, Log-normal, and Log-Pearson Type III probability distributions respectively. He concluded that, comparing the probability distributions, Log-Pearson Type III distribution with the least absolute differences or highest r-square for all the plotting positions is the best distribution among the three for Ona River under Ogun-osun river basin study location. Hence, there is need to explore more probability distribution as well as plotting position for reliable and acute result

According to Adeboye and Alatise(2007) in their research stated that 18-year streamflow record of River Osun at Apoje gauging station was obtained from Ogun-Osun River Basin and Rural Development Authority Abeokuta, Nigeria. He was able to draw the annual hydrograph for the station by plotting the peak discharges as the ordinate while hydrologic years as the abscissa. He added that the peak discharges were fitted into three major statistical distributions namely normal, log-normal and log-Pearson type III while seven plotting positions of Hazen, Weibull, Blom, Cunnane, California, Gringorton and Chegodajev were employed in determining their probabilities of exceedance. Results obtained showed that the annual maximum discharges at Apoje station varied from $169m^3/s$ to $400m^3/s$ within duration of 18 years. Weibull's plotting position combined with normal distribution gave the highest fit, most reliable and accurate predictions of the flood in the study area having the coefficient of determination (R-Square)and root mean square error (RMSE) of 0.9950 and 35.09 respectively. However, with

their research, conclusion can be drawn but since there are more plots as well as probability distribution, more research is needed to come up with the best.

Abdulkarim and Chowdhury (1995) eliminate lognormal, Gumbel and log-Pearson type III distributions as probability distribution for modelling at-site annual maximum discharge series in Bangladesh. They show that the generalized extreme value distribution best represented the statistical characteristics of observed data. Comparisons of distributions were based upon the root mean square deviation in fit, the probability plot correlation coefficient and L-moment ratio diagrams. The L-moment ratio diagrams method assesses whether a candidate distribution is capable of simulating the site-to-site variation in statistical characteristics that is present in an observed flood sample.Furthermore, by employing only one (1) distribution, judge cannot be made because no comparison was taking among distribution instead.

2.2 Flooding.

Flood is said to occur when there is an unusual high stage of a stream or river due to runoff from precipitations in quantities which is too large to be confined in the normal water surface elevation of the stream or river (Mustafa and Yusuf, 2012).

The following methods are used for flood design of major hydraulics structure (Mustafa and Yusuf, 2012).

i. Probable maximum flood (PMF).

   The Probable Maximum Flood (PMF) is defined as the flood that may be expected from the most severe combination of critical meteorological and hydrologic conditions. PMF also has physical meanings which provide an upper limit of the interval within which the engineer must operate and design (Lariyah*et al*, 2012). The PMF is computed by using the Probable Maximum Storm (PMS) which is an estimate of the physical upper limit to storm rainfall over the catchment. This is obtained from the studies of all the storms that have occurred over the region and maximizing them for the most critical atmospheric conditions.

ii. Standard Project Flood (SPF).

This is the flood resulting from the most severe combination of meteorological and other hydrological conditions considered characteristic of the region. The SPF is computed from the Standard Project Storm (SPS) over the watershed considered and may be taken as the largest storm observed in the region of the watershed. It is not maximized for the most critical atmospheric conditions but it may be transposed from an adjacent region to the watershed under consideration (Mustafa and Yusuf, 2012).

iii. Frequency based Flood.

This flood is estimated by frequency analysis of the annual flood values of adequate length. Sometimes when the flood data is inadequate, frequency analysis recorded storm data is made and the storm of a particular frequency applied to the unit hydrograph to derive the design flood (Mustafa and Yusuf, 2012). This flood usually has a return period greater than the storm.

2.2.1 Common Types of Flood.

Ivan (2014) shows that there are several different kinds of flood, and each one bears a different impact in terms of how it occurs, the damage it causes, and how it is forecasted. Hence, the following are the commonness types of flood.

2.2.1.1 Coastal (Surge Flood)

A coastal flood, as the name suggests, occurs in areas that lie on the coast of a sea, ocean, or other large body of open water. It is typically the result of extreme tidal conditions caused by severe weather. Storm surge is produced when high winds from hurricanes and other storms pushed water while the onshore is the leading cause of coastal flooding and often the greatest threat associated with a tropical storm. In this type of flood, water overwhelms low-lying land and often causes devastating loss of life and property.

Coastal flooding is categorized in three levels:

i. Minor: A slight amount of beach erosion will occur but no major damage is expected.
ii. Moderate: A fair amount of beach erosion will occur as well as damage to some homes and businesses.

iii. Major: Serious threat to life and property. Large-scale beach erosion will occur, numerous roads will be flooded, and many structures will be damaged. Citizens should review safety precautions and prepare to evacuate if necessary.

2.2.1.2 Fluvial (River Flood)

Fluvial, or riverine flooding, occurs when excessive rainfall over an extended period of time causes a river to exceed its capacity. It can also be caused by heavy snow melt and ice jams. The damage from a river or channel flood can be widespread as the overflow affects smaller rivers downstream, often causing dams and dikes to break and swamp nearby areas.

There are two main types of riverine flooding:

i. Overbank flooding occurs when water rises overflows over the edges of a river or stream. This is the most common and can occur in any size channelfrom small streams to huge rivers.
ii. Flash flooding is characterized by an intense, high velocity torrent of water that occurs in an existing river channel with little to no notice. Flash floods are very dangerous and destructive not only because of the force of the water, but also the hurtling debris that is often swept up in the flow.

2.2.1.3 Pluvial (Surface Flood)

A pluvial, or surface water flood, is caused when heavy rainfall creates a flood event independent of an overflowing water body. One of the most common misconceptions about flood risk is that one must be located near a body of water to be at risk. Pluvial flooding debunks that myth, as it can happen in any urban area even at higher elevation areas that lie above coastal and river floodplains.

There are two common types of pluvial flooding:

i. Intense rain saturates an urban drainage system. The system becomes overwhelmed and water flows out into streets and nearby structures.

ii. Run-off or flowing water from rain falling on hillsides that are unable to absorb the water. Hillsides with recent forest fires are notorious sources of pluvial floods.

2.2.2 Causes of Floods.

Floods are caused by many factors: heavy precipitation or downpour, severe winds over water, unusual high tides, tsunamis, or failure of dams, levees, retention ponds, or other structures that contained the water (Chow, 1964).

Periodic floods occur on many rivers, forming a surrounding region known as the flood plain. During times of rain or snow, some of the water is retained in ponds or soil, some is absorbed by grass and vegetation, some evaporates, and the rest travels over the land as surface runoff. Floods occur when ponds, lakes, riverbeds, soil, and vegetation cannot absorb all the water. Water then runs off the land in quantities that cannot be carried within stream channels or retained in natural ponds, lakes, and man-made reservoirs. River flooding is often caused by heavy rain, sometimes increased by melting snow.

Nwigweand Emberga (2014) describedthat thegenerally causes of flood could be as a result of the following.

i. Heavy Rainfall: The simplest explanation for flooding is heavy rains. No matter where someone live, people are surrounded by infrastructure and systems designed to move rainwater into appropriate basins and reservoirs. In most cases, the infrastructure does its job, and you never have to think about where the rain goes when it runs off.When it rains heavily, however, those systems are overwhelmed, and that water doesn't drain nearly as quickly as it needs to. In short, the drainage systems back up, and the water risessometimes into homes. This typically happens only in cases of sustained heavy rains over a long period.

ii. Overflowing Rivers:It's not necessarily needs to have heavy rains to experience flooding in manyareas. For example, if someone live along the downstream of a river and areas upstream experiencing heavy rains, it could lead to a serious overflow which could eventually lead to loss of lives and properties. Most larger rivers include a series of dams to help manage large amounts of rainfall, and most river systems are managed by

government authorities. Most times, those authorities have to make tough decisions about how to operate dams. They often can manage the water and prevent flooding altogether but not always.

iii. Broken Dams: When heavy rains come, and water levels rise, aging dams tend to fail and unleash torrents of water on unsuspecting households. This is part of what happened after Hurricane Katrina hit New Orleans in 2005. Levees failed and made the flooding far worse than it would have been otherwise. While ithas come to depend on 20th-century architecture, and much of it does its job well, there is always a possibility that a structure will certainly fail.

iv. Urban Drainage Basins: Drainage systems are constructed to ensure that waste water and sewage is transported neatly to disposal points, thereby keeping the environment well drained and free of waste. Examples of components that make up a good drainage system includes; closed ditches, having pipe drains, good drainage pipes, channels and conduits that may convey water without causing any damage or loss of property or even endangering life. Sustainable Drainage Systems are approaches put in place to manage the water quantity (flooding), water quality (pollution) and amenity issues in the environment. Sustainable drainage about drainage. So, lack of this can cause serious threat to the environment.

v. Storm Surges and Tsunamis: Rainwater is not always the culprit when it comes to flooding. Storm surges related to hurricanes and other storms can lead to significant flooding, as can tsunamis that are sometimes caused by underwater earthquakes. Given modern technology, its usually known about storm surges and tsunamis before they arrive, but this is not always the case. For example, in 2004, an earthquake off the coast of Indonesia created a tsunami that gave little warning before coming ashore.

vi. Channels with Steep Sides: Flooding usually occurs when there is fast runoff into lakes, rivers and other reservoirs. This is usually the case with rivers and other channels that feature steep sides. It is a similar issue to having a lack of vegetation also.

vii. A Lack of Vegetation: Vegetation help a lot to slow runoff and prevent flooding. When there is a lack of vegetation, there is little to stop water from run off. This can be a bit of a conundrum after a drought or storm. While residents' areas likely welcome the rain, the lack of vegetation after the drought or storm can cause flash flooding. This does not

always happen given that basins and reservoirs are close to empty, but it can occur in cases of extreme rains when experiencing long periods of drought.

viii. Melting Snow and Ice: A winter of heavy snow and other precipitation can lead to spring of flooding. More so, the snow and ice have reached somewhere when they melt. Most mountainous areas experience relatively consistent snowfall overall from year to year, but an unusually heavy wind.

2.2.3 Effects of Floods.

Flooding has many impacts. It damages property and endangers the lives of humans and other species. Rapid water runoff causes soil erosion and concomitant sediment deposition elsewhere (such as further downstream or down a coast). The spawning grounds for fish and other wildlife habitats can become polluted or completely destroyed. Some prolonged high floods can delay traffic in areas which lack elevated roadways. Floods can interfere with drainage and economic use of lands, such as interfering with farming. Structural damage can occur in bridge abutments, bank lines, sewer lines, and other structures within floodways. Waterway navigation and hydroelectric power are often impaired. Financial losses due to floods are typically millions of dollars each year (Chow, 1964).

2.2.4 Ways to Prevent Flood.

Floods can be prevented with proper planning and adoption of effective mitigation measures, particularly in areas prone to flooding (Alayande, 2010).Properly conveying advance warnings of impending storms and floods will not only give people the opportunity to be proactive in preventing damage to their property but will also save lives.

2.2.4.1 Improving Drainage.

Improving water drainage helps control floods by facilitating easy flow of excess water, especially in urban areas during flash floods. Drainage efficacy can be improved by rehabilitating and cleaning water drainage systems, including removing debris and solid waste from drainage systems like storm drains and French drains, de-silting trenches and underground water channels and building galleys. Effective drainage systems can ensure flood waters find a clear waterway, leading to non-risk areas.

2.2.4.2 Building Dikes and Levees.

Dikes and levees are flood-control structures built to fight river flooding and water surges. Dikes and levees restrain rivers during floods by providing artificial water channels that prevent runoffs from bursting floodplains. Floodplains are natural waterways that carry excess river water during heavy rainfall. Dike engineers/constructors alter and develop floodplains by building dikes and levees to boost flood-prevention when floodplains well up.

2.2.4.3 Building Canals.

Canals are artificial water channels that can be crucial to flood prevention. Canals facilitate control of water levels passing through, and form linear reservoirs and water locks. During flooding, excess water is channeled through canals to non-risk areas further downstream or to other areas with high demand for water, such as arid and semi-arid areas. Canals can also be used to lock excess flood waters in its linear reservoirs to prevent flooding and store water for future use.

2.2.4.4 Harvesting Rain Water.

Harvesting rainwater involves collecting and storing rainwater and can not only prevent floods, it can also curb urban water scarcity. Harvested rain water can be cleaned for human consumption and distributed to people in rural areas, as well as urban areas in times of water scarcity. Rain water harvesting is carried out in different ways, including building ponds, storm drains, water retention basins and flood-control dams.

2.3 Methods of Parameter Estimation.

Parameter estimation is important in flood frequency analysis. Any set of flood occurrences can be described using some statistical distribution. Statistical distribution may be mathematically represented through functional relationships which may contain a set of parameters(Ohja *et al*, 2008). There are four well-known parameter estimation techniques, viz.:

2.3.1 Graphical Method.

The graphical method for estimating the parameters of accelerated life data involves generating two types of plots. First, the life data at each individual stress level are plotted on a probability paper appropriate to the assumed life distribution (i.e., Weibull, exponential, or lognormal). This can be done using either Probability Plotting or Least Squares (Rank Regression) method. The

parameters of the distribution at each stress level are then be estimated from the plot. Once these parameters have been estimated at each stress level, the second plot is created on a paper that linearizes the assumed life-stress relationship (e.g., Arrhenius, inverse power law, etc.). To do this, a life characteristic must be chosen to be plotted. The life characteristic can be any percentile, the scale parameter, mean life, etc. The plotting paper used is a special type of paper that linearizes the life-stress relationship. For example, a log-log paper linearizes the inverse power law relationship, and a log-reciprocal paper linearizes the Arrhenius relationship. The parameters of the model are then estimated by solving for the slope and the intercept of the line.

2.3.2 Frequency Factor Method

The general frequency equation proposed by (Chow, 1964; Ojha e*t al*, 2008) and applicable for different distributions, is in the following form:

$$\boldsymbol{x_T = m + K_T S} \tag{2.1}$$

Where,

$\boldsymbol{x_T}$ = magnitude of the different peak flood series, or magnitude of the flood at the required return period

T=Return Period,

$\boldsymbol{K_T}$= Frequency factor corresponding to $\boldsymbol{x_T}$,

$\boldsymbol{m}$= Mean,

$\boldsymbol{S}$= Standard deviation of the population.

Hence,**Table 2.1** shows the frequency factors for different distribution.

Table 2.1 Frequency factor for different distributions

Normal Distribution	$w = [In\left(\frac{1}{p^2}\right)]^{\frac{1}{2}}, P = \frac{1}{T}$ $z = w - \frac{2.515517 + 0.802853w + 0.010328w^2}{1 + 1.432788w + 0.189269w^2 + 0.001308w^3}$ $K_T = z$
EV-I Distibution	$K_T = \frac{-\sqrt{6}}{\pi}\{0.5772 + In\left[In\left(\frac{T}{T-1}\right)\right]\}$
Log pearson Tpye III Distribution	$K_T = z + (z^2 - 1)k + \frac{1}{3}(z^3 - 6z)k^2 - (z^2 - 1)k^3 + zk^4 + \frac{1}{3}k^5$ Where $k = \frac{g}{6}$ and g is skewness

Source:(Chow, 1964; Ojha e*t al*, 2008)

2.3.3 Method of Moments (MOM)

In statistics, the method of moments is a method of estimation of population parameters such as mean, variance, median, etc. (which need not be moments), by equating sample moments with unobservable population moments and then solving those equations for the quantities to be estimated.According to Hazelton(2011), the rationale for the method of moments is that the sample moments are natural estimators of the corresponding model-based moments, and so a good estimate of (θ) will reproduce these observed moments. In practice it is usual (although not essential) to use moments of the lowest possible orders in order to obtain parameter estimates.

Furthermore, the intuitively appealing approach to estimation is the method-of -moments. It is simple to apply. Given *p(x)* a probability mass function, the ith theoretical moment is define as.

$$\mu_i = \sum_j x^i p(x_j) \qquad (2.2)$$

for i = 1, 2…., and given *f(x)* a probability density function, The theoretical moment is define as

$$\mu_i = \int x^i f(x)dx \quad (2.3)$$

for i = 1, 2..., These are the moments about zero, or the non-central moments. The first moment is simply the mean and that the variance is

$$\sigma^2 = E(x-\mu)^2 \quad (2.4)$$

$$= \int (x-\mu)^2 f(x)dx$$

$$= \int (x^2 - 2\mu x + \mu^2) \quad f(x)dx$$

$$= \int x^2 f(x)dx \quad - \mu^2 \int f(x)dx$$

$$= \mu_2 - \mu_1^2$$

The same result of Eq. 2.4 holds for discrete random variables with the integrals changed to summations. The variance is the second central moment

2.3.4 Least Square Method

The Method of Least Squares is a procedure to determine the best fit line to data; the proof uses simple calculus and linear algebra. The basic problem is to find the best fit straight line y = ax + b given that, for n ∈ {1,…N}, the pairs (x_n, y_n) are observed (Miller, 2012). The method easily generalizes to finding the best fit of the form

$$y = a_1 f_1(x) + \cdots + c_k f_k(x); \quad (2.5)$$

it is not necessary for the functions f_k to be linearly in x, all that is needed is that y is to be a linear combination of these functions.

Despite its popularity and versatility, LSM has its problems. Probably, the most important drawback of LSM is its high sensitivity to outliers (i.e., extreme observations). This is a consequence of using squares because squaring exaggerates the magnitude of differences (e.g., the difference between 20 and 10 is equal to 10 but the difference between 202 and 102 is equal to 300) and therefore gives a much stronger importance to extreme observations (Abdi, 2016).

2.3.5 Probability Weighted Moments

According to Greenwood*et al* (1979) stated that, a distribution function.

$$F=F(x) = P(X \leq x) \quad (2.6)$$

may be characterized by probability weighted moments, which are defined as.

$$M_{l,j,k} = E[X^l F^j (1-F)^k] = \int_0^1 [x(F)]^j \, F^j (1-F)^k dF \tag{2.7a}$$

where *l*, *j*, and k are real numbers. If $j = k = 0$ and *l* is a nonnegative integer, *then* $M_{l,0,0}$ represents the conventional moment about the origin of order 1. If $M_{l,0,0}$exists and *X* is a continuous function of *F*, then $M_{l,j,k}$ exists for all nonnegative real numbers *j* and *k*. If *j* and *k* are nonnegative integers, then.

$$M_{l,0,k} = \sum_{j=0}^{k} \binom{k}{j} (-1)^j M_{l,j,0} \tag{2.7b}$$

$$M_{l,j,0} = \sum_{k=0}^{j} \binom{j}{k} (-1)^k M_{l,0,k} \tag{2.7c}$$

where if $M_{l,0,k}$ exists and *X* is a continuous function of *F*, $M_{l,j,0}$ exists. In general, even if $M_{l,j,k}$exists, it may be difficult to derive its analytical form, particularly if the inverse $x = x(F)$ of the distribution $F=F(x)$ cannot be analytically defined. In the special case where *l*,*j*,and *k* are nonnegative integers, $M_{l,j,k}$is proportional to $E[X^l_{j+1,\ k+j+1}]$, the *l*th moment about the origin of the $(j + 1)$th order statistic for a sample of size $k+j + 1$. More specifically.

$$E\left[X^l_{j+1,\ k+j+1}\right] = \frac{M_{l,j,k}}{B[j+1,k+1]} \tag{2.8}$$

where *B*denotes the beta function. If $j = 0$, $(k + 1)$ $M_{l,0,k}$ represents the *l*th moment about the origin of the first-order statistic for a sample of size $k + 1$ and if $k = 0$, $(j + 1)$ $M_{l,j,0}$ represents the *l*th moment about the origin of the $(j + 1)$*th* order statistic for a sample of size $j + 1$. Hence, the expected value of the range of *X* in a sample of size$n = k+ I =j+ 1$ is

$$E\left[X_{n,n} - X_{1,n}\right] = n(M_{1,n-1,0} - M_{1,0,n-1}) \tag{2.9}$$

Probability Weighted Moments(PWM) method is much simpler than the other methods like method of moments and maximum likelihood method for the estimation of Fréchet distribution parameters (Hooda *et al*, 2018). Furthermore, this method estimates result in explicit expressions and therefore, may serve as a better alternative for parameter estimation.

Table 2.2 Probability distribution parameters in relation to sample moments.

Distribution	***Probability Density Function***	***Range***	***Equation of parameters in terms of the sample moments.***
Exponential	$f(x) = \frac{1}{\beta}\exp\left[\frac{-(x)}{\beta}\right]$	$x \geq 0$	$\beta = \bar{x}$
Extreme Value type-I	$f(x) = \frac{1}{\beta}\exp\left[-\frac{x-u}{\beta} - exp\left(-\frac{x-u}{\beta}\right)\right]$	$-\infty < x < \infty$	$u = \bar{x} - 0.5772\beta, \quad \beta = \frac{\sqrt{6s_x}}{\bar{x}}$
Gamma	$f(x) = \frac{(x)^{\gamma-1}e^{-x/\beta}}{\beta^{\gamma}\Gamma(\gamma)}$	$x \geq 0$	$\beta = \frac{(s_x)^2}{\bar{x}}$
Normal	$f(x) = \frac{1}{\sigma\sqrt{2\pi}}exp\left(-\frac{(x-u)^2}{2\sigma^2}\right)$	$-\infty \leq x \leq \infty$	$u = \bar{x}, \quad \sigma = s_x$
Lognormal	$f(x) = \frac{1}{x\sigma_L\sqrt{2\pi}}exp\left[-\frac{1}{2}\left(\frac{Inx - u_L}{\sigma_L}\right)^2\right]$	$x > 0$	$u_y = \bar{y}, \quad \sigma_y = s_y$
Pearson type-III	$f(x) = \frac{(x-u)^{\gamma-1}exp\left[-\frac{(x-u)}{\beta}\right]}{\beta^{\gamma}\Gamma(\gamma)}$	$x \geq u$	$u = \bar{x} - s_x\sqrt{\gamma}, \quad \beta = \frac{\sqrt{\gamma}}{s_x}, \gamma = \left(\frac{2}{g}\right)^2$
Log-Pearson type-III	$f(x) = \frac{(Inx-u)^{\gamma-1}\exp\left[-\frac{(Inx-u)}{\beta}\right]}{\lvert\beta\rvert \ \ \Gamma(\gamma)}$	$Inx \geq u$	$u = \bar{y} - s_y\sqrt{\gamma}, \quad \beta = \frac{\sqrt{\gamma}}{s_y}, \gamma = \left(\frac{2}{g_y}\right)^2$

Source: (Ojha*et al*, 2008)

Where,

x is defined as the sample data along horizontal axis.

y is definedas the sample data along vertical axis.

u is define as the mean.

σ is the standard deviation of the sample data.

s_x is the standard deviation of x sample data.

s_y is defined as the standard deviation of y sample data.

g is defined as the coefficient of skewness

β is the probability weighted moments.

2.3.6 Maximum Likelihood Estimation (MLE) Method.

Mauro (2017) show that the maximum likelihood estimation method was introduced by the eminent British statistician Ronald Fisher (1890–1962). It basically involves maximizing a function of the distribution parameters, known as the likelihood function. By equating this function to the condition, it reaches its maximum value, it results in six (6) Parameter and Quantile estimation 215 a system of an identical number of equations and unknowns, whose solutions yield the MLE estimators for the distribution parameters.

2.3.7 L- Moments.

L-moment statistics are used for computing sample statistics for data at individual sites; for testing for homogeneity/heterogeneity of proposed groupings of sites (regions); for conducting goodness-of-fit tests for identifying a suitable probability distribution(s); and for solving for distribution parameters for the selected probability distribution. In particular, estimation of the L-moment ratios of L-Cv and L-Skewness is a key element in determining the success of the regional frequency analysis in computing quantile estimates for selected sites. The L-moment measure of location (L_1), and L-moment ratio measures of scale (t_2), skewness (t_3) and kurtosis (t_4) are as follows.

Location, mean:

Mean= L_1 (2.10)

Scale, L-Cv (t_2): $t_2 = L_2 / L_1$ (2.11)

L-Skewness (t_3): $t_3 = L_3 / L_2$ (2.12)

L-Kurtosis (t_4): $t_4 = L_4 / L_2$ (2.13)

Where:

$L_1 = \beta_0$

(2.14a) $L_2 = 2\beta_1 - \beta_0$

(2.14b) $L_3 = 6\beta_2 - 6\beta_1 + \beta_0$

(2.14c) $L_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0$

(2.14d)

L_1, L_2, L_3 *and* L_4 are called the L-moments.

Hence, the data ($x_{1:n}$) are first ranked in ascending order from 1 to n and:

$\beta_0 = n^{-1} \sum_{j=1}^{n} x_j$ (2.15a)

$\beta_1 = n^{-1} \sum_{j=2}^{n} x_j \, [(j-1)/(n-1)]$ (2.15b)

$\beta_2 = n^{-1} \sum_{j=3}^{n} x_j \, [(j-1)(j-2)/(n-1)(n-2)]$

(2.15c) $\beta_3 = n^{-1} \sum_{j=4}^{n} x_j \, [(j-1)(j-2)(j-3)/(n-1)(n-2)(n-3)]$

(2.15d)

Where

$\beta_0, \beta_1, \beta_2$ and β_3 are called probability weighted moments (PWM).

j is the ranking order number.

x is the sample data.

n is the number of ranking.

2.3.5.1 Graphical Depictions of L-Cv and L-Skewness.

Table 2.3 General Descriptions of Relative Magnitude of L-Cv.

ASSESSMENT OF THE RELATIVE MAGNITUDE OF L-Cv	
$0.000<\|L\text{-}Cv\| \leq 0.025$	Minimal variability-often found in controlled manufacturing processes
$0.025<\|L\text{-}Cv\| \leq 0.075$	Minor variability
$0.075<\|L\text{-}Cv\| \leq 0.150$	Moderate variability
$0.150<\|L\text{-}Cv\| \leq 0.400$	Large variability- often accompanied by large skewness
$0.400<\|L\text{-}Cv\|$	Very large variability-often accompanied by very large skewness

Source. (Izinyon and Ehiorobo, 2014)

L-Coefficient of variation *(L-Cv)* is a dimensionless measure of variability. For a distribution or sample data that only has positive values, *L-Cv* is in the range $0 < L\text{-}Cv < 1$. Negative values of L-Cv are only possible if the at-site mean has a negative value. Table 2.3 lists some guidelines for describing the relative magnitude of L-Cv for a dataset or distribution (Izinyon and Ehiorobo, 2014).

Table 2.4 General Descriptions of Relative Magnitude of L-Skewness.

ASSESSMENT OF THE RELATIVE MAGNITUDE OF L-SKEWNESS	
L-skewness = 0.0	Symmetrical Distribution
$0.000<\|$ L-skewness $\| \leq 0.050$	Minor skewness
$0.050<\|$ L-skewness $\| \leq 0.150$	Moderate skewness
$0.150<\|$ L-skewness $\| \leq 0.300$	Large skewness
$0.300<\|$ L-skewness $\|$	Very large skewness -often accompanied by very large skewness

Source. (Izinyon and Ehiorobo, 2014)

L-Skewness is a dimensionless measure of asymmetry, which may take on positive or negative values. For a distribution or sample data, L-Skewness is in the range 0 < | L-Skewness | < 1. Table 2.4lists some guidelines for characterizing the relative magnitude of L-Skewness for a dataset or distribution.

2.4 Probability Distribution.

2.4.1 Normal Family: Normal, Lognormal and Lognormal III.

The normal distribution is a family of distributions that is often used to model operational processes. The normal distribution is the most widely known and used of all distributions. Because the normal distribution approximates many natural phenomena so well, it has developed into a standard of reference for many probability problems.

a. Normal Probability Density Function

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-Z^2/2}, -\infty < z < \infty \quad (2.16)$$

Where z is the standard normal variable

e is the exponential value as 2.71 approximately

Statistical parameters of this distribution are shown below

$$P = \frac{1}{T} \quad (2.17)$$

$$w = [In\left(\frac{1}{p^2}\right)]^{\frac{1}{2}} \quad (2.18)$$

$$z = w - \frac{2.515517 + 0.802853\,w + 0.010328\,w^2}{1 + 1.432788\,w + 0.189269\,w^2 + 0.001308\,w^3} \quad (2.19)$$

$$K_T = z$$

$$\bar{Q} = \frac{1}{n}\sum_{i=1}^{n} Q_{max} \quad (2.20)$$

$$S_Q = \sqrt{\sum_{i=1}^{n}(Q_{max} - \bar{Q})^2/(n-1)} \quad (2.21)$$

Where.

p is the probability of exceedance.

w is the intermediate variable.

z is the frequency factor, in other word standard normal variable.

$\bar{Q}$is the mean of the maximum annual discharge (m^3/sec).

Q_{max} is the annual maximum discharge (m^3/sec).

S_Q is the standard deviation of the annual maximum discharge (m^3/sec).

n is the number of the observed data used.

The corresponding discharges (Predicted flood) at different or various return period of the observed flood data were determined using the mathematical expression Eq. 2.1.

Where Q_T is the predicted flood with reference to the standard variable.

b. Log-normal Probability Density Function

$$\boldsymbol{f(x) = \frac{1}{x\sigma\sqrt{2\pi}} exp\left(\frac{-y-u_y}{2\sigma_y^2}\right)} \quad (2.22)$$

Statistical parameters

$$y = \log_{10} Q_{max} \quad (2.23)$$

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} \log_{10} Q_{max} \quad (2.24)$$

$$S_y = \sqrt{\frac{\sum_{i=1}^{n}(y-\overline{y})^2}{n-1}} \quad (2.25)$$

Where

y is the logarithm of the observed data at base 10.

$\overline{y}$ is the mean of y (m^3/sec).

S_y is the standard deviation with respect to y (m^3/sec).

Hence, intermediate value and frequency factor were determined using the Eq. 2.18 and Eq. 2.19 respectively.

The statistical variate and the predicted discharge were determined using the below equations respectively

$$y_T = \overline{y} + z * S_y \quad (2.26)$$

$$Q_T = 10^{y_T} = 10^{(\overline{y} + z * S_y)} \quad (2.27)$$

2.4.1.1 Characteristics of the Normal distribution

i. Symmetric, bell shaped.
ii. Continuous for all values of x between -∞ and ∞ so that each conceivable interval of real numbers has a probability other than zero.
iii. Two parameters mean (μ) and standard deviation(σ). Note that the normal distribution is actually a family of distributions, since μ and σ determine the shape of the distribution.
iv. The notation N (μ, σ^2) means normally distributed with mean μ and variance σ^2. If we say X ~ N(μ, σ^2) we mean that X is distributed N(μ, σ^2).
v. About 2/3 of all cases fall within one standard deviation of the mean, that is
$P(\mu - \sigma \leq X \leq \mu + \sigma) = .6826.$
vi. About 95% of cases lie within 2 standard deviations of the mean, that is
$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0.9544$

2.4.1.2 The standardized normal distribution.

i. General Procedure. Its suspect from the formula for the normal density function, it would be difficult and tedious to do the calculus every time. In view of that, there is a new set of parameters for μ and σ. So instead, it's usually work with the standardized normal distribution, where μ = 0 and σ=1, i.e., N (0,1). That is, rather than directly solve a problem involving a normally distributed variable x with mean μ and standard deviation σ, an indirect approach is used.
ii. Rules for using the standardized normal distribution. It is very important to understand how the standardized normal distribution works, so we will spend some time here going over it.

2.4.2 Pearson Type III Family: Log-Pearson Type III and Pearson Type III

The Pearson family of distributions is made up of seven distributions: Type I-VII. Pearson type III distributions is the generalized gamma distribution and also one of the most popular distributions for hydrologic frequency analysis. Bobee and Robitaille (1977) compared Pearson type III and log Pearson type III distribution using several long-term records of annual flood flows and found Pearson type III distribution to be much preferable, especially when the method of moments (MOM) was applied to observed sample data. Bobee (1973), Chang and Moore (1983), among others, used it for flood frequency analysis. Markovic (1963) found practically no difference in fitting of Pearson and lognormal distributions to annual precipitation and runoff data. Matalas (1963) found Pearson type III distribution to be representative of low flows.

It covers many specified average, standard deviation, skewness and kurtosis. Together they form a 4-parameter family of distributions that covers the entire skewness-kurtosis region other than the impossible region (Bobee, 1973).

The seven types are described below:

i. Type I: Beta Distribution
ii. Type II: Special case of beta distribution that is symmetrical
iii. Type III: Gamma Distribution
iv. Type IV: Region above Type V
v. Type V: 3 parameter distribution represented by curve
vi. Type VI: Region between Gamma and Type V
vii. Type VII: Special case of Type IV that is symmetrical

The special cases can be ignored (II and VII). Type I and III and are alias for distribution already covered. That leaves Type IV, V and VI as new distributions (Bobee, 1973).

a. Pearson Type III Distribution

Probability Density Function

$$\boldsymbol{f(x) = \frac{1}{aГ}(\frac{x-c}{a})^{b-1} exp[-\left(\frac{x-c}{a}\right)]} \quad (2.28)$$

Where a > 0, b > 0 and 0 < c < x are parameters

Cumulative Density Function

$$\boldsymbol{F(x) = \frac{1}{a\Gamma(b)} \int_0^{\infty} (\frac{x-c}{a})^{b-1} exp\left(-\frac{x-c}{a}\right) dx} \quad (2.29)$$

Statistical parameters are as flows but $\bar{Q}$ and s_Q were earlier stated at Eq. 2.20 and 2.21 respectively.

$$k = \frac{g}{6} \quad (2.30)$$

$$K_T = z + (z^2 - 1)k + \frac{1}{3}(z^3 - 6z)k^2 - (z^2 - 1)k^3 + zk^4 + \frac{1}{3}k^5 \quad (2.31)$$

If k =0, then K_T=z

$$g = n\sum_{i=i}^{n}(Qmax - \bar{Q})^3/(n-1)(n-2)s_Q^3 \quad (2.32)$$

Where

g is skewness coefficient

w, *z*, $\bar{Q}$ and Q_{max} were already defined in chapter 2.4.1.

The statistical variate is being determined using Eq. 2.1

b. Log Pearson Type III Distribution

Probability Density Function

$$\boldsymbol{f(x) = \frac{(Inx-u)^{\gamma-1} exp[-(Inx-u)/\beta}{|\beta|\Gamma(\gamma)}} \quad (2.33)$$

$\boldsymbol{u = \bar{y} - s_y\sqrt{\gamma}}$, $\boldsymbol{\beta = \frac{\sqrt{\gamma}}{s_y}}$, $\boldsymbol{\gamma = (\frac{2}{g_y})^2}$

Statistical parameters y, $\bar{y}$, k, S_y and K_T already mention earlier in chapter. 2.4.1

$$Cs = g = n\sum_{i=i}^{n}(y - \bar{y})^3/(n-1)(n-2)s_y^3 \quad (2.34)$$

Hence, intermediate value (*w*) and frequency factor (*z*) also were determined using the equation earlier mention under Normal Distribution

The statistical variate and the predicted discharge also were determined using Eq. 2.26 and 2.27 respectively.

2.4.3 GEV Family: Gumbel, GEV and Weibull

Generalized Extreme Value (GEV) distribution: The GEV distribution is a family of continuous probability distributions developed within extreme value theory. Extreme value theory provides the statistical framework to make inferences about the probability of very rare or extreme events. The GEV distribution unites the Gumbel, Fréchet and Weibull distributions into a single family to allow a continuous range of possible shapes. These three distributions are also known as type I, II and III extreme value distributions. The GEV distribution is parameterized with a shape parameter, location parameter and scale parameter. The GEV Is equivalent to the type I, II and III, respectively. Based on the extreme value theorem the GEV distribution is the limit distribution of properly normalized maxima of a sequence of independent and identically distributed random variables. Thus, the GEV distribution is used as an approximation to model the maxima of long (finite) sequences of random variables.According to (Ripunjai *et al*, 2010) said the extreme value distribution family models are used to capture the uncertainty of data and finally Generalized Extreme Value (GEV) distribution model is foundthe best fitted distribution model.

Adeyemi and Adebanji (2006) stated that the Generalized Extreme-Value (GEV) distribution is often used to model extremes events of natural phenomena such as river heights, sea levels, stream flows, rainfall and air pollutions in order to obtain the distribution of daily or annual maxima. Additionally, in reliability context, analogous analyses are performed where the interest is in sample minima strengths and failure times. The general equation for (GEV) is shown Table. 2.2.

2.4.4 Principle and issues in selecting distribution.

The selection of the most suitable probability distribution and associated parameter estimation procedure, as a fundamental step in flood frequency analysis, has remained the most difficult task for many researchers and water practitioners.

Langat *et al*, (2019) described that the current approaches that are used to identify the probability distribution functions that are best suited for the estimation of maximum, minimum, and mean streamflow are by comparing the performance of probability distributions, and illustrates fitting

tests along with evaluation procedures, and selection procedures through using a river basin as a case study. The assemblage ofcomputer statistical packages in an integrated development environment helps a lot in determination or selecting the best distributions. Goodness-of-fit (GOF) tests-based analysis, and information criteria-based selection procedures were used to identify the most suitable distribution models.

Statistics and statistical reasoning is based on probability and probabilistic reasoning. Statistical conclusions or proofs are never absolutely certain, but are always qualified, in that, a certain probability or likelihood is attached to the conclusion or proof. In statistics, a conclusion may be reported as being true, but attached to this conclusion is a probability that the conclusion is false. Hopefully this latter probability is very small. If a particular problem can be analyzed using probabilistic reasoning, then probabilities or degrees of belief can be attached to particular conclusions. That is, principles of probability allow uncertainty to be quantified, and limits can be placed on this uncertainty. Where this can be done, it is often possible to devise methods which can be used to increase the probability of, or degree of belief in, the conclusion.

In examining probability, there are three possible interpretations of probability. Each approach is a legitimate one and each approach has its own particular uses. These interpretations are the classical approach, the frequency interpretation and subjective probabilities. These are outlined as follows.

i. Classical Approach to Probability The classical or theoretical approach to probability depends only on theoretical considerations. In this interpretation, various assumptions or postulates are used and these form the basis for the theory of probability. Building on these assumptions, probability distributions can be constructed. Games of chance are used to illustrate this approach. However, these need not be real games of chance that are actually conducted. Rather, it is only necessary to imagine the possibility of the game of chance.
ii. Frequency Interpretation of Probability An alternative view of probability begins by examining the frequency of occurrence of particular outcomes in a large series of trials of a probability experiment. The number of times that a particular outcome occurs in these trials forms the basis for defining probabilities in this manner.

iii. Subjective Interpretation of Probability A third possible way of defining probabilities is to take an event and subjectively attach a number to this event. This number then becomes an estimate of the probability of the occurrence of the event. This type of probability has also been called judgment or personal probability, because a personal judgment concerning the likelihood or probability of a particular event occurring has been made. In doing this, the number representing the subjective probability of the event should be between 0 and 1. Numbers in this range are used in order to obtain probabilities which are consistent with the other interpretations of probability, where probabilities are always between 0 and 1.

2.5 Probability Plots.

2.5.1 Plotting Positions.

Plotting order-ranked data is a standard technique that is used in estimating the probability of extreme weather events. Typically, observations, say, annual extremes of a period of N years, are ranked in order of magnitude and plotted on probability paper. Some statistical model is then fitted to the order-ranked data by which the return periods of specific extreme events are estimated (Lesse, 2005).

The exceedance probability for a particular sample value is a function of sample size and the rank of the particular sample. For exceedance probabilities, the sample values are ranked from largest to smallest. The general expression in common use for plotting position is

$$p = \frac{r-b}{n+1-2b} \tag{2.35}$$

where

r is the ordered rank of a sample value, *n* is the sample size, *b* is a constant between 0 and 1,

Given a plot produced in this way, one might simply draw a smooth curve through the points, and usually extend it to the recurrence intervals of interest. This constitutes an entirely empirical approach and is sometimes made easier by constructing the plot using a transformed scale for the cumulative frequency. The simplest such transformation is to plot the logarithm of the cumulative frequency, which flattens the curve and makes extrapolation easier (Alam and Matin, 2004).

A second approach would be to choose a distribution type, and adjust its free parameters, so a plot of the distribution matches the plot of the sample. This is commonly done by least squares fitting. Fitting by eye is also possible if an appropriate *probability paper* is adopted, on which the transformed axis is not logarithmic, but is transformed in such a way that the corresponding distribution plots as a straight line; however, this cannot be done for all distributions (Alam and Matin, 2004).

These simple methods based on plotting positions, although widely used, are not recommended. Two fundamental problems with the methods are seldom addressed. First, it is inherent in the methods that each of N quantile bins of the distribution is occupied by one and only one sample point, an extremely unlikely outcome. Second, when a least squares fit is made for an analytical distribution form, the error being minimized is taken as the difference between the sample value and the distribution value, whereas the true error is not in the value but in its frequency position (Alam and Matin, 2004).

Table2.5 Plotting position Formulae.

S/N	Formular Name	Plotting position formulae P(x ≥ X)
1	Blom	$\frac{m-0.375}{n+0.25}$
2	California(modified)	$\frac{m-1}{n}$
3	Chegodayev	$\frac{m-0.3}{n+0.4}$
4	Cunnane	$\frac{m-0.4}{n+0.2}$
5	Gringorten	$\frac{m-0.44}{n+0.12}$
6	Hazen	$\frac{m-0.5}{n}$
7	Turkey	$\frac{3m-1}{3n+1}$
8	Weibull	$\frac{m}{n+1}$

Where m is the ranking order number. n is the totalnumber of ranking order.

Chow (1964) has demonstrated theoretically that California method is suitable for plotting annual exceedance series or partial duration series. However, this simple formula, plots the data at the edge of group intervals and also produces a probability of 100 percent for m=N, which may not be easily plotted on a probability scale. Hazen (1930) proposed his formulae, which plots data at the centers of group intervals. As the extreme distribution was later introduced to frequency analysis, the Weibull (1939) formula was soon found to be a compromise with more statistical justification. Chow (1964) has shown that this formula is theoretically suitable for plotting the annual maximum series; the US Water Resources Council adopted this formula as the standard plotting position method. Another compromise is the Chegodajew (1955) formula widely used in Russia and the Eastern European countries. Cunnane (1978) studied the various available plotting position methods using criteria of unbiasedness and minimum variance. For normally distributed data, he found that the Bloom (1958) plotting position is close to being unbiased, while for data distributed according to the EV-1 distribution, the Gringorton (1963) formula is the best. Adamwaski (1981) advocates for his formula, for EV-1 distribution, for high values of probability of exceedance based on mean square criteria. Similarly, other plotting position formulae have been recommended for one reason or the other. According to Alam and Matin (2004), all the methods of determining plotting positions give practically the same results in the middle of a distribution but produces different positions near the tails of the distribution. Thus, the choice of a plotting position formula becomes important for fitting the extreme value flood data.

2.6 Goodness of Fit (GOF).

2.6.1 Goodness of Fit Test

Olivares and Forero (2010) elaborate that the goodness of fit (GOF) of a statistical model describes how well it fits into a set of observations. GOF indices summarize the discrepancy between the observed values and the values expected under a statistical model. GOF statistics are GOF indices with known sampling distributions, usually obtained using asymptotic methods, that are used in statistical hypothesis testing. As large sample approximations may behave poorly in small samples, a great deal of research using simulation studies has been devoted to investigate under which conditions the asymptotic p-values of GOF statistics are accurate (i.e.,

how large the sample size must be for models of different sizes). Goodness of fit tests commonly used in statistics are:

i. The chi-square.
ii. Kolmogorov-Smirnov.
iii. Anderson-Darling.
iv. Shipiro-Wilk.
v. Probability Plot Correlation Coefficient (PPCC)
vi. Coefficient of determination (R_Square)
vii. Root Mean Square Error (RMSE)

2.6.1.1 Chi-square.

A chi square statistic is a measurement of how expectations compare with results. The data used in calculating a chi square statistic must be random, raw, mutually exclusive, drawn from independent variables, and drawn from a large enough sample. For example, the results of tossing a coin 100 times meets these criteria (Will, 2018).

Two potential disadvantages of chi square are:

i. The chi square test can only be used for data put into classes (bins). A non-binned dataneed to make a frequency table or histogram before performing the test.
ii. Another disadvantage of the chi-square test is that it requires a sufficient sample size in order for the chi-square approximation to be valid.

There is another type of chi-square test, called the chi-square test for independence. The two are sometimes confused but they are quite different.

i. The chi-square test for independence compares **two** sets of data to see if there is a relationship.
ii. The chi-square Goodness of fit is to fit **one** categorical variable to a distribution.

Typically, this test is run using software. The null hypothesis for the chi-square goodness of fit test is that the data comes from a specified distribution. The alternate hypothesis is that the data does not come from a specified distribution.

To interpret the test, its needed to choose an alpha level (1%, 5% and 10% are common). The chi-square test will return a p-value. If the p-value is small (less than the significance level), the null hypothesis is rejected that the data comes from the specified distribution.

2.6.1.2 Kolmogorov-Smirnov.

Although this is called a test for normality, it actually doesn't tell you whether a particular sample likely came from a normal population. Instead, it will tell you when it is unlikely that you have a normal distribution. One advantage to this test is that it doesn't make any assumptions about the distribution of data. A sample can be compared to a distribution using a one-sample K–S test or two-sample K–S test. The test is usually performed using software (like SPSS), because critical values have to be calculated for each distribution and finding the tables of critical values isn't an easy task. The test is usually recommended for large samples over 2000. Usually Shapiro-Wilk is more prepared.

2.6.1.3 Anderson-Darling

This test is a modification of Kolmogorov-Smirnov. It is more sensitive to deviations in a distribution's tails. Like the Kolmogorov-Smirnov, this test will tell when it is unlikely to have a normal distribution and is normally run using statistical software.

2.6.1.4 Shapiro-Wilk

This test calculates a W value that will tell if a random sample came from a normally distributed population. The test is recommended for samples up to n=2000. The equation for the test is as follows

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_i\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad (2.36)$$

Where x_i is defined as sample data. $\bar{x}$ is defined as the mean of the sample data.

The coefficients given a_i are given by

$$(a_i, \dots\dots a_n)=\frac{m^T V^{-1}}{C} \quad (2.37)$$

$$C = \left||V^{-1}m|\right| = (m^T V^{-1} V^{-1} m)^{1/2}$$

V is the covariance matrix and the vector m,

$$m = (m_1 \dots\dots m_n)^T \quad (2.38)$$

2.6.1.5 Probability Plot Correlation Coefficient (PPCC)

Filliben PPCC test is a formalization of a technique used by statistical hydrologists for many decades; that is , it determine the linearity of a probability plot. Furthermore, Filliben PPCC test statistic is define as the product moment correlation coefficient between the ordered observation y_i and the order statistic median M_i for a standardized normal distribution (Vogel, 1986).

The correlation coefficient CC is expressed as follows,

$$PPCC = \sum_{i=1}^{n}(X_i - \bar{X})(M_i - \bar{M}) \Big/ \sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(M_i - \bar{M})^2} \quad (2.39)$$

where $\bar{X}$ and $\bar{M}$ represent the mean values of the observation X_i and the fitted quantiles M_i , respectively, and n is the sample size. If correlation coefficient CC is close to 1.0, the observations can be defined by the fitted probability distribution. The order statistic median for $M_i$by (Filliben,1975; Kim *et al*, 2010) is explained as follow.

$$M_i = \varphi^{-1} m_i \quad (2.40)$$

Where φ^{-1} means the inverse of cumulative distribution function for the standard normal distribution and the median value m_i is given as shown in the three equations below:

$$m_i = 1 - (0.5)^{\frac{1}{n}}, \text{ for i=1} \quad (2.41)$$

$$m_i = \frac{i-0.3175}{n+0.365}, \quad for\ i = 2,3, \dots, n-1 \quad (2.42)$$

$$m_i = (0.5)^{\frac{1}{n}}, for\ i = n \quad (2.43)$$

If the following condition is satisfied, the null hypothesis cannot be rejected at the q significant level$r > r_q(n)$.

As stated above, the PPCC test statistics are influenced by the characteristics of various significance levels, sample sizes, plotting position formulas, and shape parameters for the fitted probability distribution. Therefore, the application of a proper plotting position formula for the fitted probability distribution is important to estimate the PPCC test statistics.

2.6.1.6 Coefficient of Determination (R-Squared)

The coefficient of determination is well-defined for linear regression models, and is popularly used in practice as a measure of goodness-of-fit of the underlying models. It indicates the proportionate amount of variation in the response variable *y* explained by the independent variables *X* in the linear regression model. The larger the R-squared is, the more variability is explained by the linear regression model (Zhang, 2016).

R-squared is the proportion of the total sum of squares explained by the model. R-squared, a property of the fitted model, is a structure with two fields:

i. Ordinary — Ordinary (unadjusted) R-squared

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \tag{2.44}$$

ii. Adjusted — R-squared adjusted for the number of coefficients

$$R^2_{adj} = 1 - \left(\frac{n-1}{n-p}\right)\frac{SSE}{SST} \tag{2.45}$$

Where

SSE is the sum of squared error,

SSR is the sum of squared regression,

SST is the sum of squared total, *n* is the number of observations,

p is the number of regression coefficient.

Because R-squared increases with added predictor variables in the regression model, the adjusted R-squared, adjusts for the number of predictor variables in the model. This makes it more useful for comparing models with a different number of predictors.

2.6.1.7 Root mean square error (RMSE)

The root mean square error of a distribution fitted to the observed discharge data at a site is the square root of the sum of the squares of the differences between the observed and predicted values (Izinyon and Ehiorobo, 2014). It is computed using the equation:

$$\sqrt{\frac{\sum(x_i-y_i)^2}{n-m}} \tag{2.46}$$

where x_i , i=1,..,n are the observed values and y_i, i=1,….,n are the values computed from the assumed probability distributions, the number of parameters estimated for the distribution is denoted by m.

2.6.2 L-Moment Diagram.

Hosking and Wallis (1995) describe goodness-of-fit as a measure for identifying the probability distribution(s) that most closely matches the weighted-average regional L-Skewness and L-Kurtosis values for the grouping of sites for the proposed region. The test can best be visualized using the L-Moment Ratio Diagram shown in Figure 2.1. The test can be viewed as plotting the regional L-Skewness and L-Kurtosis values obtained from the sites for the proposed region and identifying the probability distribution(s) whose L-Kurtosis value most closely matches the regional L-Kurtosis value. Figure 2.1 gives an example of L-Moment Ratio Diagram for Selecting Best-Fit Probability Distribution

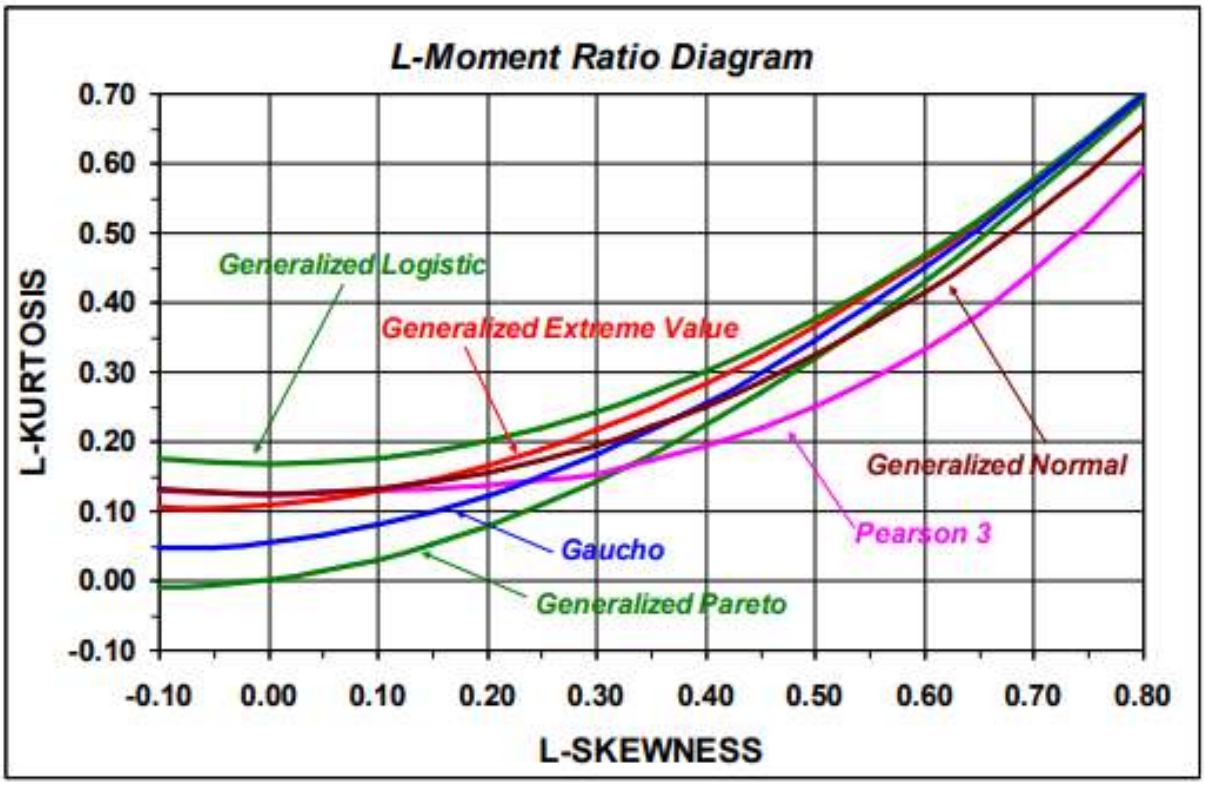


Figure 2.1 Example of L-Moment Ratio Diagram for Selecting Best-Fit Probability DistributionSource: (Peel *et al*, 2001)

2.7 Standard Errors and Confidence Interval for Quantiles.

2.7.1 Confidence Interval for Quantiles.

After having estimated the parameters of a probability distribution FX(x), another important objective of Statistical Hydrology is to estimate the quantile XF, corresponding to the non-exceedance probability F, or, equivalently, the quantile XT, corresponding to the return period T. The quantile XF can be estimated by the inverse function F(x), which is denoted here by φ (F), with the formal meaning. It is clear that a point estimator, such as X^ T, contains errors that are inherent to the uncertainties resulting from the estimation of population characteristics and parameters from finite samples of size N (Naghettini, 2017). A measure frequently used to quantify the uncertainties of X^ T and the reliability of quantile estimators is the standard error of estimate, denoted by ST and formally defined as:

$$S_T = \sqrt{E[\{X_T - E[X_T]\}^2]} \quad (2.47)$$

Confidence interval (CI). $CI =$

$$\bar{x} \pm z\frac{s}{\sqrt{n}} \quad (2.48)$$

Where $\bar{x}$ is define as sample mean.

z is defined as confidence level value

s is defined as sample standard deviation.

n is defined as sample size

2.8 Matrix Laboratory (MATLAB).

2.8.1 Overview.

MATLAB is an interactive system whose basic data element is an array that does not require dimensioning. This allows solving to solve many technical computing problems, especially those with matrix and vector formulations, in a fraction of the time it would take to write a program in a scalar non interactive language such as C or FORTRAN. The name MATLAB stands for matrix laboratory. MATLAB was originally written to provide easy access to matrix software developed by the LINPACK and EISPACK projects, which together represent the state-of-the-art in software for matrix computation(Hunt *et al*, 2001).

MATLAB has evolved over a period of years with input from many users. In university environments, it is the standard instructional tool for introductory and advanced courses in mathematics, engineering, and science. In industry, MATLAB is the tool of choice for high-productivity research, development, and analysis.MATLAB features a family of application-specific solutions called toolboxes. Very important to most users of MATLAB, toolboxes allow individual to learn and apply specialized technology. Toolboxes are comprehensive collections of MATLAB functions (M-files) that extend the MATLAB environment to solve particular classes of problems. Areas in which toolboxes are available include signal processing, control systems, neural networks, fuzzy logic, wavelets, simulation, and many others.

Mbele, and Kusakana, (2015) in their research said, it's important to invest in technologies that aim to harvest excess energy/pressure considering the unstable use of water of the world level. Ithighlighted that, three models for energy recovery system using MATLAB/Simulink software. The models comprise of the components of energy recovery turbine, Drive Train and Permanent Magnet Synchronous Generator (PMSG). The models can be integrated to form a complete model for energy recovery system. Developers of conduit hydropower can use the model to simulate the technology during the planning stage. This shows that Matlab software is a strong tool for writing programme intended to solve computational problems.

Bello *et al*, (2015) shows that Matlab software can solve many technical computational problems, especially those with matrix and vector formulations, in a fraction of the time it would take to write a program in a scalar non-interactive language as compared to C or FORTRAN.

This concluded that the use of Matlab software can be of use in computational analysis of various water related problems.

MATLAB is a high-performance language for technical computing. It integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation. Typical uses include:

i. Math and computation.
ii. Algorithm development.
iii. Modeling, simulation, and prototyping.
iv. Data analysis, exploration, and visualization.
v. Scientific and engineering graphics.
vi. Application development, including Graphical User Interface building.

2.8.2MATLAB System.

The MATLAB system consists of five main parts:

2.8.2.1 MATLAB language.

This is a high-level matrix/array language with control flow statements, functions, data structures, input/output, and object-oriented programming features. It allows both "programming in the small" to rapidly create quick and dirty throw-away programs, and "programming in the large" to create complete large and complex application programs.

2.8.2.2 MATLAB working environment.

This is the set of tools and facilities that you work with as the MATLAB user or programmer. It includes facilities for managing the variables in your workspace and importing and exporting data. It also includes tools for developing, managing, debugging, and profiling M-files, MATLAB's applications.

2.8.2.3 Handle Graphics.

This is the MATLAB graphics system. It includes high-level commands for two-dimensional and three-dimensional data visualization, image processing, animation, and presentation graphics. It

also includes low-level commands that allow you to fully customize the appearance of graphics as well as to build complete Graphical User Interfaces on your MATLAB applications.

2.8.2.4 MATLAB mathematical function library.

This is a vast collection of computational algorithms ranging from elementary functions like sum, sine, cosine, and complex arithmetic, to more sophisticated functions like matrix inverse, matrix eigenvalues, Bessel functions, and fast Fourier transforms.

2.8.2.5 MATLAB Application Program Interface (API).

This is a library that allows you to write C and Fortran programs that interact with MATLAB. It includes facilities for calling routines from MATLAB (dynamic linking), calling MATLAB as a computational engine, and for reading and writing MAT-files.

2.8.3 Features of MATLAB

Following are the basic features of MATLAB:

i. It is a high-level language for numerical computation, visualization and application development.
ii. It also provides an interactive environment for iterative exploration, design and problem solving.
iii. It provides vast library of mathematical functions for linear algebra, statistics, Fourier analysis, filtering, optimization, numerical integration and solving ordinary differential equations.
iv. It provides built-in graphics for visualizing data and tools for creating custom plots.
v. MATLAB's programming interface gives development tools for improving code quality, maintainability, and maximizing performance.
vi. It provides tools for building applications with custom graphical interfaces.
vii. It provides functions for integrating MATLAB based algorithms with external applications and languages such as C, Java, .NET and Microsoft Excel.
viii. It provides line color and style commands for plotting of graphs as shownat Table 2.6 and 2.7 respectively.

Table 2.6 Line colour option in MATLAB software.

S/N	Option	Description
1	'red' or 'r'	Red
2	'green' or 'g'	Green
3	'blue' or 'b'	Blue
4	'yellow' or 'y'	Yellow
5	'magenta' or 'm'	Magenta
6	'cyan' or 'c'	Cyan
7	'white' or 'w'	White
8	'black' or 'k'	Black
9	'none'	No color

Source: (Hunt *et al*, 2001)

Table 2.7 Line style using MATLAB software command.

S/N	Value	Description
1	'o'	Circle
2	'+'	Plus sign
3	'*'	Asterisk
4	'.'	Point
5	'x'	Cross
6	'square' or 's'	Square
7	'diamond' or 'd'	Diamond
8	'^'	Upward-pointing triangle
9	'v'	Downward-pointing triangle
10	'>'	Right-pointing triangle
11	'<'	Left-pointing triangle
12	'pentagram' or 'p'	Five-pointed star (pentagram)
13	'hexagram' or 'h'	Six-pointed star (hexagram)
14	'none	No markers

Source: (Hunt *et al*, 2001)

2.8.4 MATLAB's Power of Computational Mathematics.

MATLAB is used in every facet of computational mathematics. Following are some commonly used mathematical calculations where it is used most commonly:

i. Dealing with Matrices and Arrays.
ii. 2-D and 3-D Plotting and graphics.
iii. Linear Algebra.
iv. Algebraic Equations.
v. Non-linear Functions.
vi. Statistics.
vii. Data Analysis.
viii. Calculus and Differential Equations.
ix. Numerical Calculations.
x. Integration.
xi. Transforms.
xii. Curve Fitting.
xiii. Various other special functions.

2.8.5 MATLAB program files.

MATLAB allows writing two kinds of program files:

i. Scripts: script files are program files with .m extension. In these files, you write series of commands, which you want to execute together. Scripts do not accept inputs and do not return any outputs. They operate on data in the workspace.
ii. Functions: functions files are also program files with .m extension. Functions can accept inputs and return outputs. Internal variables are local to the function. You can use the MATLAB editor or any other text editor to create your .m files. In this section, we will discuss the script files. A script file contains multiple sequential lines of MATLAB commands and function calls. You can run a script by typing its name at the command line.

CHAPTER THREE

MATERIALS AND METHODS

3.1 Overview

The flood frequency analysis within the study relied on secondary discharge data which served as input data in Matlab software for the computation of predicted discharge using the selected distribution models. The data was obtained from National Water Resources Institute, Mando, Kaduna, Kaduna State. The data was subjected to normality as shown in Fig. 3.3, so also Fig. 3.2 described the entirely methodology for this study, right from flood frequency analysis using secondary data through testing the various plotting positions when fitted with selected distribution models which were further subjected to test for GOF, adopting the use of coefficient of determination and root mean square error. Field visit was carried out to Kaduna State Water Works for checkmating the stream gauge works

3.1.1 Description of the Study Area.

According to Alayande (2010), the study area covers the reaches of the Kaduna River extending between the confluences with the Kangimi and Rigasa Rivers located upstream and downstream Kaduna city respectively as shown in plate 3.1, where A, B, C, D, E, F and G are points along the river. For the purpose of this study, the reach extends between the Eastern Bye Pass bridge at Malali to the Kaduna South Waterworks. The reach is the most developed of the three reaches in terms of physical developments in the floodplain and host to the Kaduna South waterworks. The gauging station situated at a coordinate of latitude and longitude (N10^{o}30’07.63” and E07^{o}25’06.61”) respectively



Fig 3.1 Location of the study area on satellite imagery.

3.2 Materials.

i. HP 250 dual core Computer

ii. Microsoft word

iii. Microsoft Excel

iv. Secondary Stream flow data (see Table 3.1)

v. MATLAB software R2017a (9.2.0.538062) and R2018a (9.4.0.813654)

vi. Google Earth software

3.3 Methods

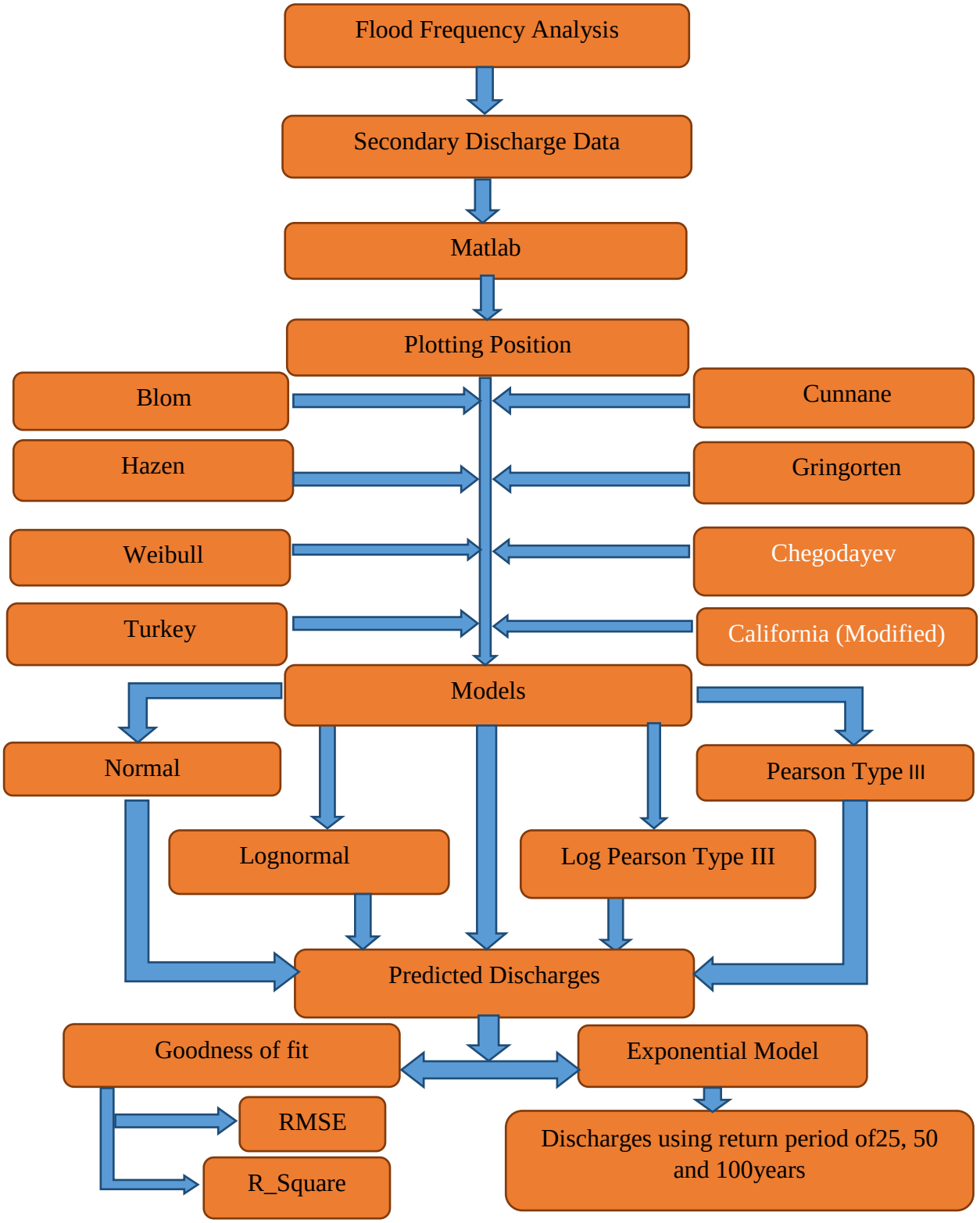


Fig 3.2 Flow Chart for the flood study.

Table 3.1 Flow Regime of Kaduna River at Kaduna South 1971 to 2000 (NWRI, 2017).

Year	Avg Daily Flow (m^3/sec)	Max 7-days Flow (m^3/sec)	Max Daily Flow Parameters		
			Discharge (m^3/sec)	Water stage (m)	Date of occurrences
1971	821.81	1,095.09	1,469.64	6.12	10th Sept 1971
1972	650.93	893.97	2270.70	6.89	21st June 1972
1973	760.01	1372.96	2296.52	6.92	11th July 1973
1974	725.97	1112.16	1441.33	6.07	1st Sept 1974
1975	941.38	1301.85	1795.29	6.45	19th Sept 1975
1976	1015.94	1591.03	2362.50	6.97	21st Sept 1976
1977	475.33	629.67	814.50	4.95	26th Sept 1977
1978	500.73	815.82	1005.40	5.39	9th Sept 1978
1979	699.12	835.19	1068.11	5.57	25th Sept 1979
1980	541.40	713.42	1066.41	5.57	2nd Sept 1980
1981	750.15	1125.74	1192.14	5.16	8th Sept 1981
1982	492.32	953.63	1251.04	5.79	7th Aug 1982
1983	395.69	561.28	700.56	4.68	1st Aug 1983
1984	233.77	302.26	422.49	3.79	2nd Oct 1984
1985	522.75	716.52	1100.82	5.58	15th Aug 1985
1986	501.82	1169.03	2871.75	7.96	22nd Jul 1986
1987	623.50	1298.34	2370.38	6.98	21st Jun 1987
1988	536.00	1112.96	2072.57	6.78	25th and 26th Jun 1988
1989	67.02	93.28	126.20	2.38	15th Aug 1989
1990	658.06	1124.76	1440.94	5.67	19th Sep 1990
1991	853.71	1163.84	1460.45	5.70	28th Aug 1991
1992	1093.47	1686.62	2579.50	7.14	1st Sep 1992
1993	828.40	1523.63	2926.31	7.39	18th Sep 1994
1994	550.09	727.30	862.01	5.08	10th Aug 1995
1995	643.89	838.28	1151.79	5.67	12th Sep 1996
1996	668.08	867.41	1013.98	5.41	20th Aug 1997
1997	960.22	1675.79	2395.85	7.00	10th Sept 1998
1998	923.92	1269.23	2488.34	7.07	4th Sept 1999
1999	831.71	958.29	1083.83	5.55	25th Aug 2000
2000	NA	869.56	1055.51	5.50	26th Aug 2001

3.3.1 Normality Test Using Shapiro Wilk (*W Test*) Approach

The above test has been found to be the most powerful test in most situation (Brzezinski, 2012). The probability values for *W* are valid for sample sizes greater than three (3). Hence, was developed by Shapiro and Wilk for sample sizes up to 20. Below is the algorithms in computing normality test using W test.

i. ***Mupad*** was typed in the MatLAB command window

ii. An icon titled ***Choose Symbolic Editor*** display or pop up, given option to choose.

iii. ***New MuPAD Notebook*** was selected and then the notebook icon displays.

iv. The below command was followed to determine the goodness of fit by using shapiro wilk test.

x:= (x1,x2,x3…..xn):

stats::swGOFT(x)

v. The results of the peak value (p_value) displays as shown in fig. 3.3.

vi. The null hypothesis of the data set is accepted since the p_value is greater than 0.05 as the chosen alpha level. Hence, the data set came from a normally distributed dischage.

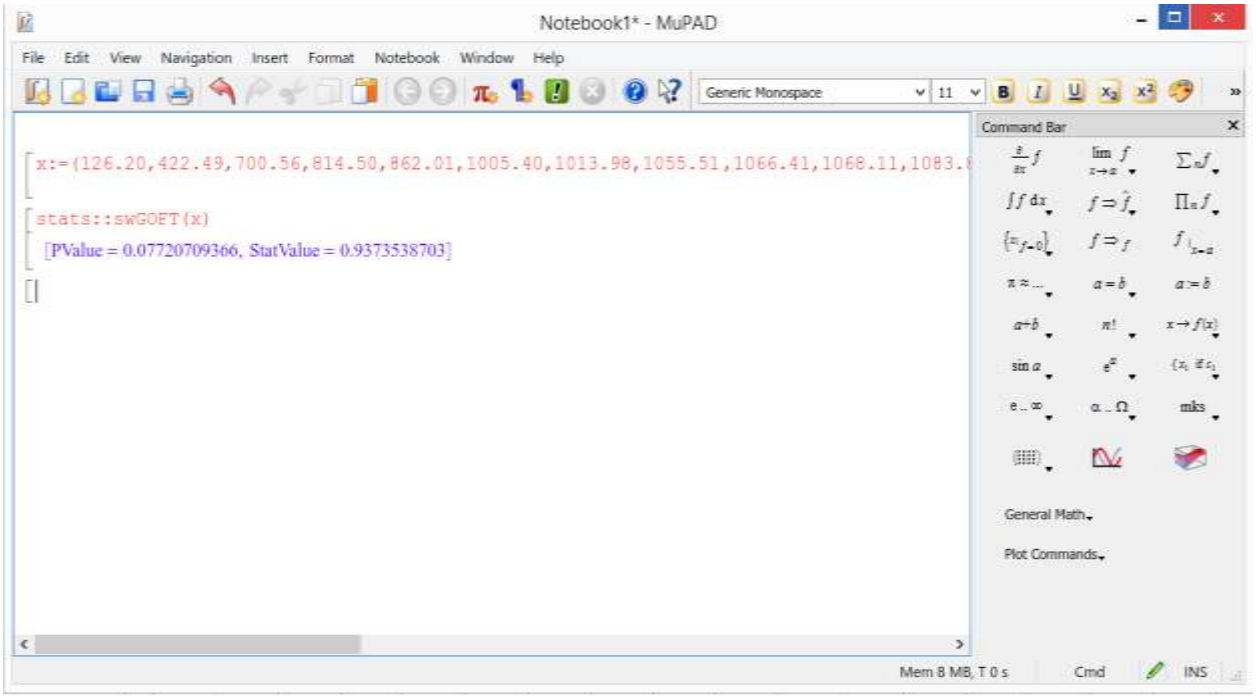


Fig 3.3 Normality test using Shapiro Wilk Approach

3.3.2 Probability of Exceedance

i. Stream flow data of 30years as shown **Table 3.1** was used for the computation of the plotting position values of various formulae

ii. The data was arranged in descending order i.e (higher to lower values) as shown in (**Appendix. I-VIII**) for the determination of the probability of exceedance of various formulae. Hence, they were ranked 1-30 from highest to lowest value respectively.

iii. The ranking numberswere used in the determination of probability of exceedance.

iv. The various ranking order number was defined as ***m*** while the totalranking number is defined as ***n***; hence, they were inputted into the MATLAB software for the execution of the various plotting position formulae

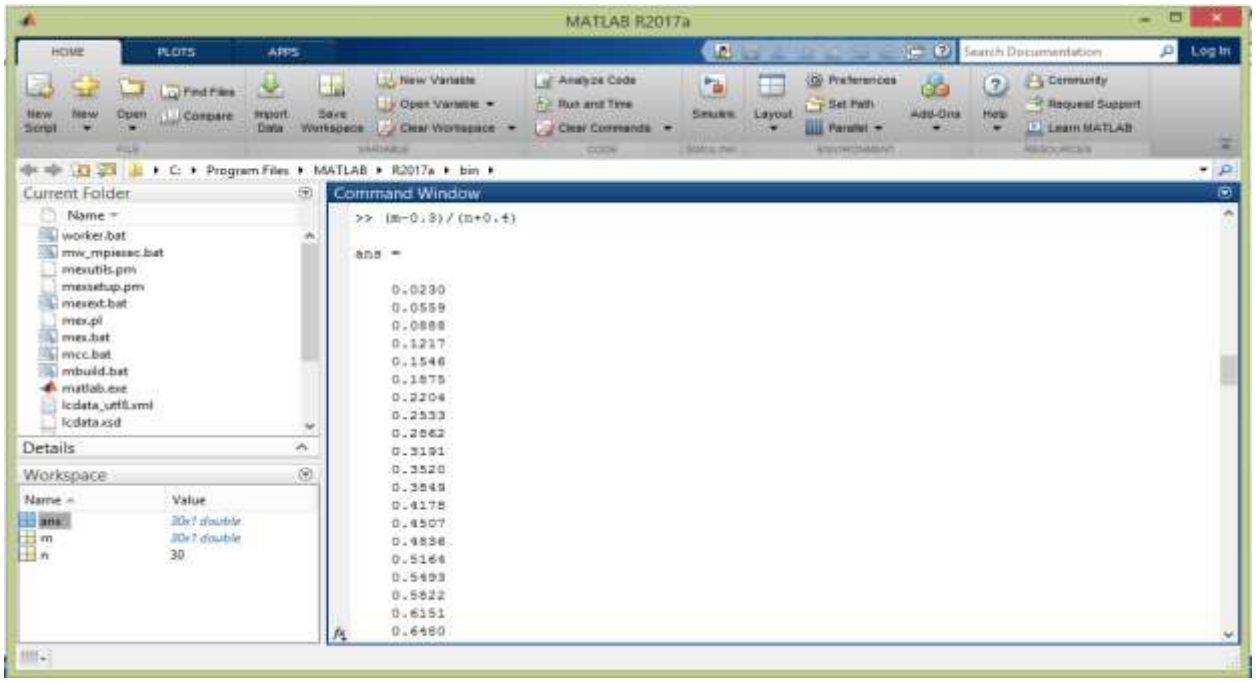
v. The eight (8) plotting position formulaeas presented in Table2.5 were adopted in this research andprogrammed one at a time within the Matlab command window background for the determination of probability of exceedance with reference to the corresponding ranking order number.

vi. Each of the formulae was used and the probability of exceedance was determined using the command window in the software as shown in Fig 3.4.

vii. The values were exported from the software to Microsoft Excel by right clicking on ***ans*** in the workspace and then clicking on ***open selection*.**

viii. The values appearin spreadsheet format in ***editor command window*** in the Matlab.

ix. The values were highlighted and copied which were taken to Microsoft-Excel for compilation or data arrangement.



3.4 Plotting Position Determination.

The probability of exceedance values for the corresponding observed discharge data and the return period was used to estimate the corresponding predicted discharge value.

i. The probability plot or plotting position were defined; hence, inputted into the command window of the Matlab software.

ii. Taking power of a number was separated by a dot which indicated a scalar quantity Hence, the Matlab language only understand it but anything other than this will not compute the result using the equations.

iii. The intermediate variable ***w*** and the frequency factor ***z*** were also defined using their respective equation as presented in Eq. 2.18 and 2.19,of each of the plot models. Hence, were inputted into the command window of the Matlab software.

iv. The discharges were inputted also in same command window which were arranged or ranked in descending order.

3.3.3 Normal Distribution.

The forecasting formulae or normal distribution model equation was used to predict the corresponding discharge using MATLAB software and the following algorithms were followed.

i. The MATLAB software application was installed and a shortcut icon was created on the desktop background.

ii. The MATLABshortcut icon on the desktop was opened by a double click on it.

iii. The MATLAB command window or workspace displayed.

iv. The ***Home*** tab at the menu bar was then clicked and ***new*** was selected as one of the components of the ***Home*** tab.

v. ***Live function*** was selected which falls under the ***new*** icon.

vi. A text editor icon displayed where the algorithms for the computation of normal distribution was then typed or written as shown in Fig 3.4. Hence, the empirical formulae used for normal distribution are presented in Eq. 2.1, 2.17 - 2.21, which helps in determination of predicted and corresponded discharges.

vii. The ***save*** ribbon under ***live editor*** was then selected and the file was named as ***norms.***

viii. ***norms*** file was subsequently recalled for discharge determination by replacing the various probability of exceedance values at different ranking order number ***m*** up to***n.***

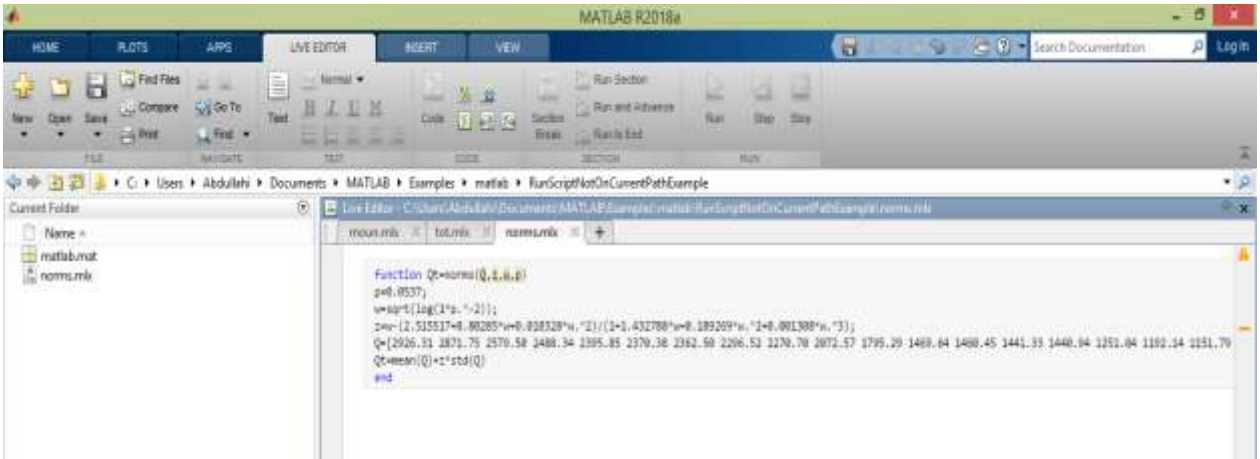


Fig 3.5 Matlab steps for Normal Distribution

3.3.4 Log-Normal Distribution

i. Same algorithms were followed as in normal distribution, from roman number I to V.

ii. A text editor icon displayed where the algorithms for the computation of log-normal distribution was then typed or written as shown in Fig 3.6. Hence, the empirical formulae for log-normal distribution are presented in Eq. 2.1, 2.18-2.19 and 2.23-2.27 which helps in determination of predicted and corresponded discharges.

iii. Log base 10 of the discharge were determined

iv. The lognormal distribution model equation was used to predict the corresponding discharge.

v. The ***save*** ribbon under ***live editor*** was then selected and the file was named as ***lgnorms.***

vi. ***lgnorms*** file was subsequently recalled for discharge determination by replacing the various probability of values at different ranking order number ***m*** up to ***n.***



Fig 3.6 Matlab steps for Log-Normal Distribution

3.3.5 Pearson Type III Distribution

i. Same algorithms were followed as in normal distribution, from roman number I to V

ii. A text editor icon displayed where the algorithms for the computation of normal distribution was then typed or written as shown in Fig 3.7. Hence, the empirical formulae for normal distribution are presented in Eq. 2.1, 2.18-2.19 and 2.30-2.32 which helps in determination of predicted and corresponded discharges.

iii. The ***save*** ribbon under ***live editor*** was then selected and the file was named as ***pears.***

iv. ***pears***file was subsequently recalled for discharge determination by replacing the various probability of values at different ranking order number ***m*** up to ***n.***

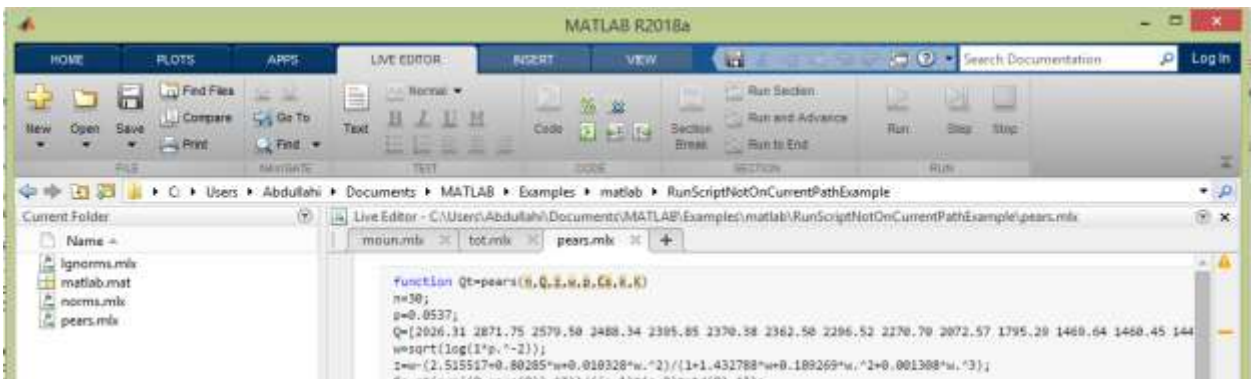


Fig 3.7 Matlab steps Pearson Type III Distribution

3.3.6 Log-Pearson Type III Distribution

i. Same algorithms were followed as in normal distribution, from roman number I to V

ii. A text editor icon displayed where the algorithms for the computation of logpearson type III distribution was then typed or written as shown in Fig 3.8. Hence, the empirical formulae for log-pearson type III distribution are presented in Eng. 2.1, 2.18-2.19, 2.30-32 and 2.34 which helps in determination of predicted and corresponded discharges.

iii. The ***save*** ribbon under ***live editor*** was then selected and the file was named as lg***pears.***

iv. ***lgpears*** file was subsequently recalled for discharge determination by replacing the various probability of values at different ranking order number ***m*** up to ***n.***

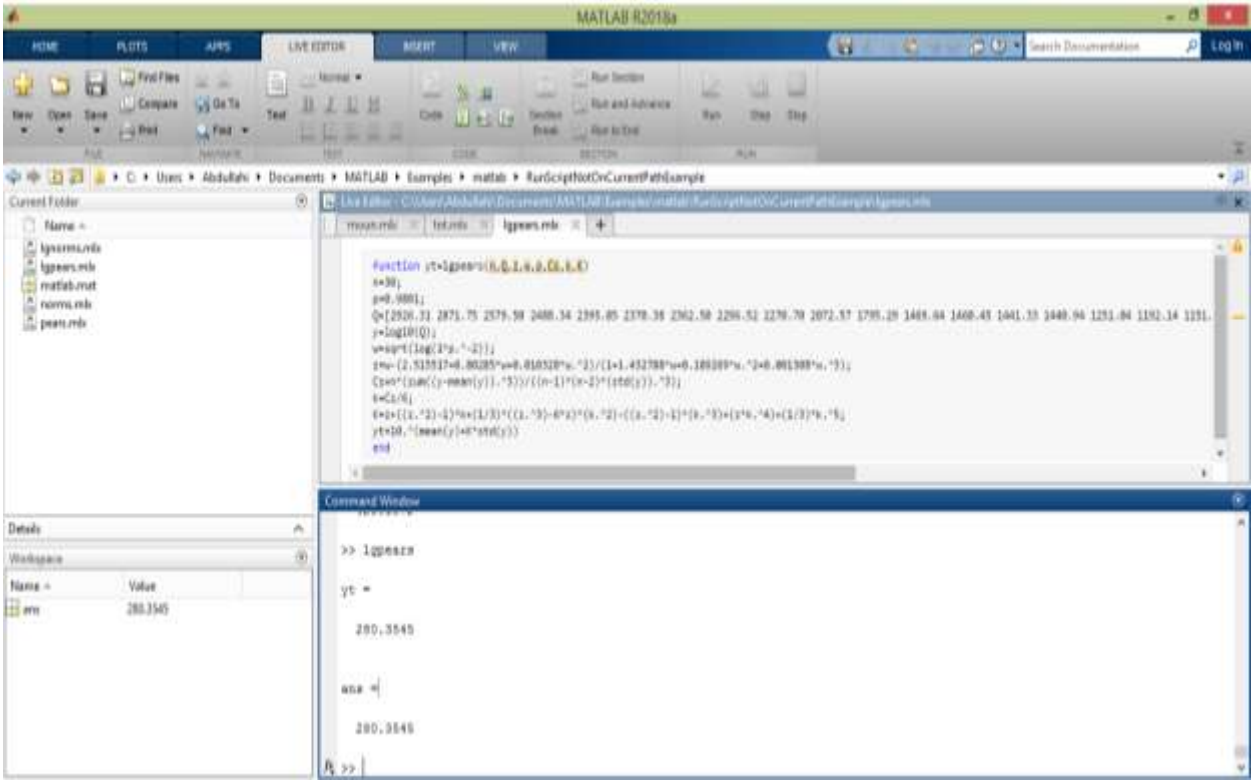


Fig 3.8MATLAB steps for Log-Pearson Type III.

3.4 Probability Distribution Curve.

Step on how to plot observed and predicted data curve using MATLAB.

The following terms were defined and typed into the MATLAB command window for plotting fitted curves.

i. The probability of non-exceedance was determined using the below equation

$$P(x < X) = 1 - \frac{1}{T}$$

$$= 1 - p$$

ii. The value for x, y1 and y2 were define within the command window (see. Appendix IX-XII) and the steps were followed accordingly as shown below.

$$x = [\quad];$$

$$y1 = [\quad];$$

$$y2 = [\quad];$$

$$f1 = fit(x, y1,' exp1');$$

$$f2 = fit(x, y2,' exp1');$$

$$plot(f1,' blue', x, y1, 'o')$$

$$hold\ on$$

$$plot(f2,' red', x, y2, 'p')$$

Where

p is the probability of exceedance

'x' is defined as the probability of **non-exceedance.**

'y1' is define as the observed discharge data.

'y2' is define as the predicted discharge data.

'f1' and 'f2' are define as functions for exponential equation 1 and 2 respectively.

'blue' and 'red' define the colour of the curve generated.

'o' and 'p' are referred to as circle and pentagram nature of line respectively, hence they define the nature of the fitted curve.

exp1 is referred to as single term exponential model

3.5 Determination of Goodness of Fit (GOF)

i. These terms $x1, y1\ and\ y2$were defined and typed into the Matlab software (i.e ***command window***) for each of the selected plotting positions when fitted with the four (4) distribution models.

ii. ***APPS***at the menu bar was clicked and then ***curve fitting***was selected**.**

iii. An icon displays which request for **X** and **Y data.**

iv. The value of $x1\ and\ y1$were imported through **X** and **Y data** respectively for observed data.

v. Different curves displayed and were tested with a click on each in which exponential model best fit the curve.

vi. ***Fit***was clicked at themenu bar and **specify validation data was s**elected**.**

vii. Same data as mention in step was selected again for R-square and RMSE values.

viii. Result of the R-squared and RMSEvalue then displayed.

ix. Step (*ii*) to (*vii*)were repeated while replacing $y1$with$y2$for predicted discharge data.

Where.

$x1 = X\ data$= Probability of non-exceedanceof the various plotting positions namely (Blom, California, Chegodayev, Cunnane, Gringorten, Hazen, Turkey and Weibull).

$y1 =$ Observed Discharge data.

$y2 =$ Predicted Discharge data of the four (4) distributions, namely (Normal, Lognormal, Pearson type III and Log-Pearson Type III).

Y data= This may be $y1$ or $y2$, depend on whether interested in determining R-Square and RMSE for Observed or Predicted.

3.6 Determination of Predicted Discharge.

Algorithms on how determine predicted discharges at various return period using MATLAB software.

i. The probability of exceedance was determined at return periods of 25, 50 and 100yearson the command window of Matlab software.

ii. Above step (i) was used in determining the probability of non-exceedance so that it can be fitted with the individual distribution models.

iii. The result of (ii) above was used in determining the magnitudes of flood using the exponential models (Appendix. XIII-XVI).

iv. Above algorithms (i-iii) wererepeated for both observed and predicted discharge data.

v. Result of (iii) above displayed and were recorded accordingly and the absolute difference was determined.

CHAPTER FOUR

RESULTSAND DISCUSSIONS

4.1 Hydrograph

The results of the entire studyare presented below with their respective values, tables and chart.

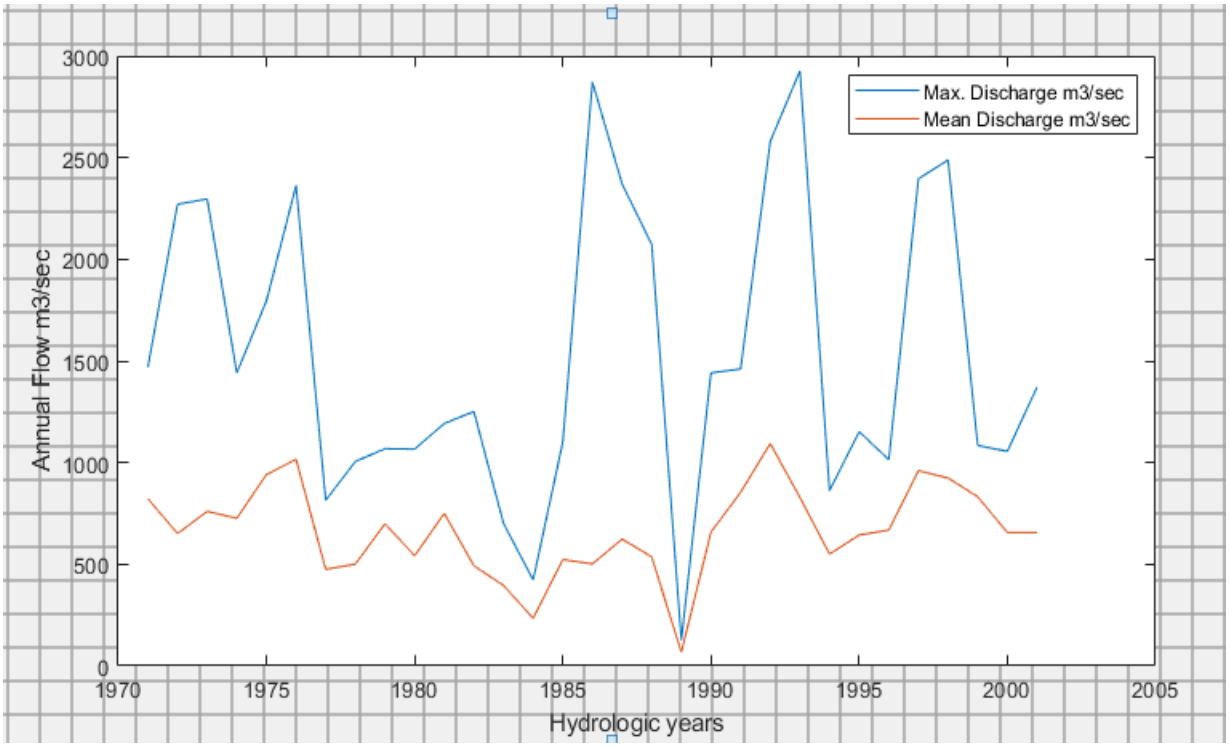


Fig 4.1 Annual hydrograph of River Kaduna.

The above (Fig. 4.1) shows the year which experience the highest maximum storm which occurred in1993 and had experienced the lowest storm in the year 1989 with magnitudes of 2926.31m^3/sec and 126.20m^3/sec respectively. Considering an annual mean discharge, it was discovered that, River Kaduna hada mean maximum and minimum annual flow of 1093.47m^3/sec and 67.02 m^3/sec respectively.Between 1986 and 1989, it discovered that there is a high decline of flow most especially in 1989. The difference in magnitudes can be attributed to diversion of water for small scale irrigation activities occurring along the course of the River. Emphasis in time series analyses should be laid on the mechanisms that generate the data but not importantly on the future sequence of events over a period of time (Kottegoda, 1980).

4.2 Normal Distribution

4.2.1 California plotting position.

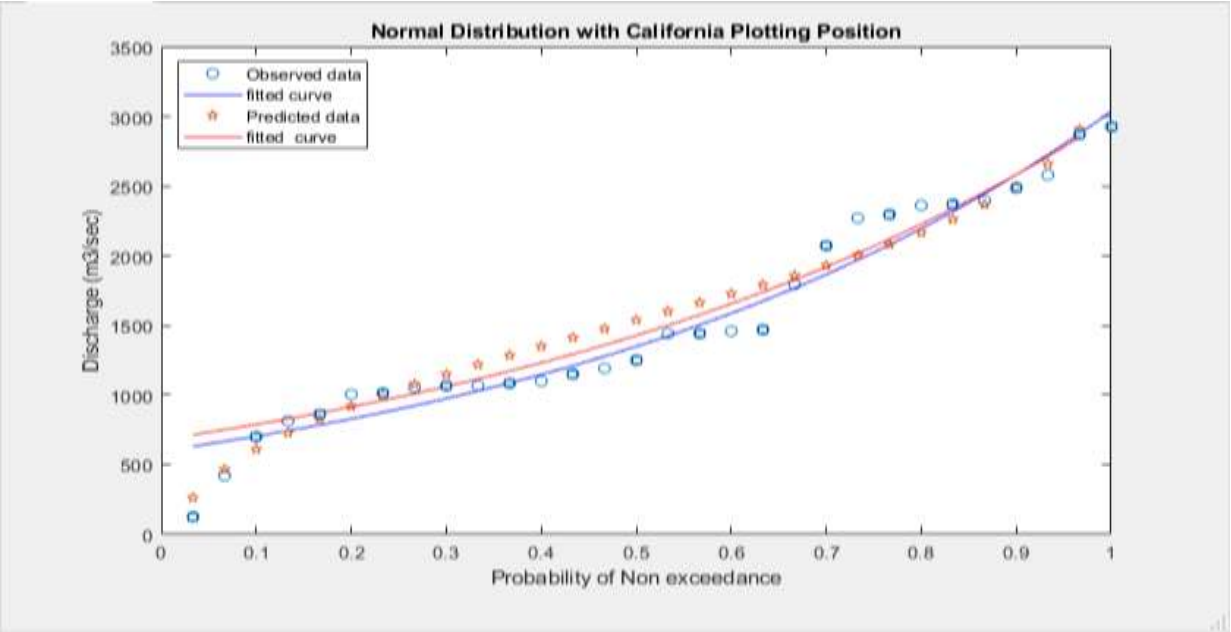


Fig 4.2 Normal Distribution Curves using California plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix I) using normal distribution model when matched with Blom plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.2. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIII) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9604 and 128.79 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 33a and 33b respectively (Appendix XIII). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 2909.49 m^3/s against a probability of exceedance of 0.0333, invariably within a recurrence interval of approximately 30 years in which the flood may occur.

4.2.2Weibull plotting position.

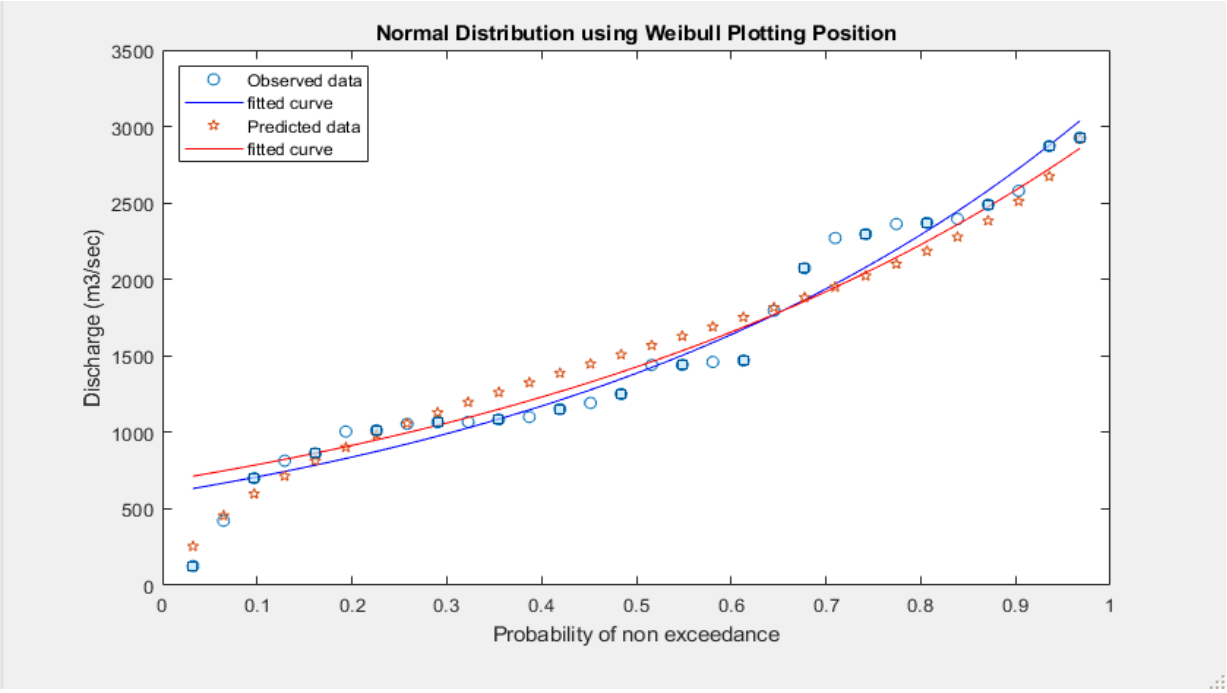


Fig 4.3 Normal Distribution Curves using Weibull plotting position

The observed discharge data and the result of predicted discharges as presented in (Appendix II) using normal distribution model when matched with Weibull plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig. 4.3. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIII) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9602 and 129.63 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 34a and 34b respectively (Appendix XIII). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 2920.46 m^3/s against a probability of exceedance of 0.0323, invariably within a recurrence interval of approximately 31 years in which the flood may occur.

4.2.3 Chegodayev plotting position.

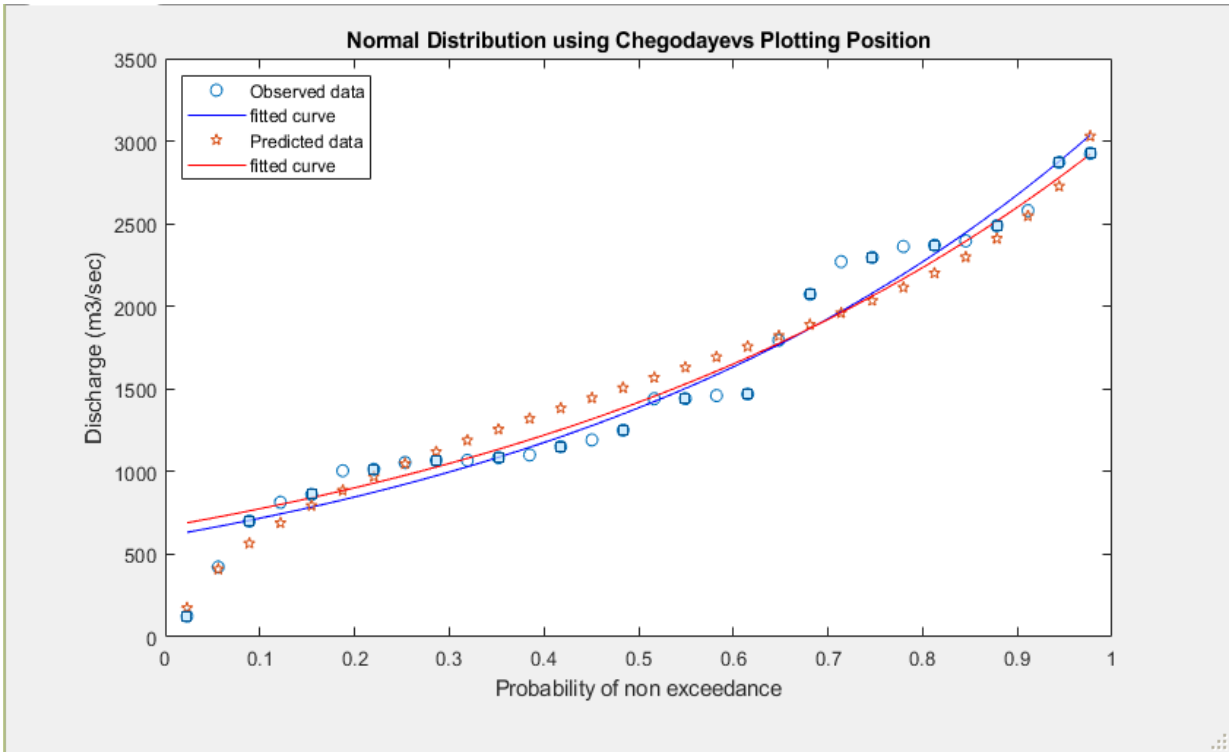


Fig 4.4 Normal Distribution Curvesusing Chegodayev plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix I) using normal distribution model when matched with Chegodayev plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig. 4.4. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIII) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9552 and 143.35 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 35a and 35b respectively (Appendix XIII). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 3029.83 m^3/s against a probability of exceedance of 0.0230, invariably within a recurrence interval of approximately 43 years in which the flood may occur.

4.2.4 Turkey plotting position.

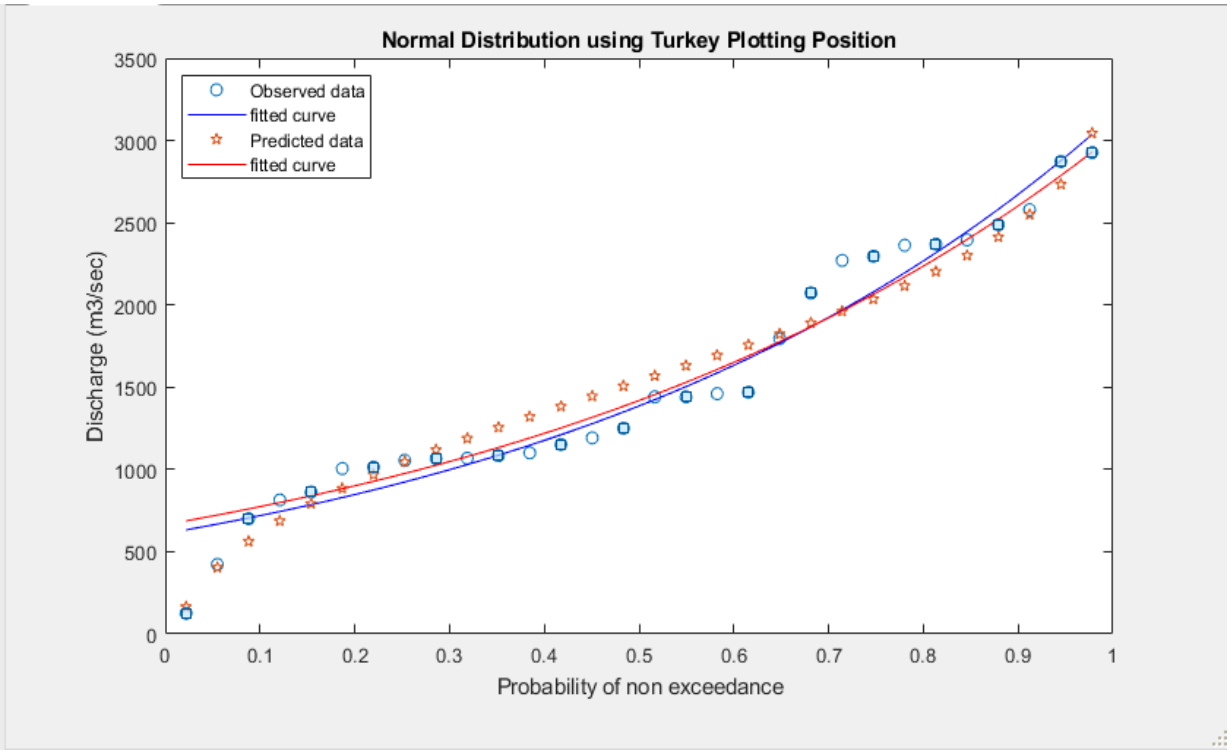


Fig 4.5 Normal Distribution Curves using Turkey plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix II) using normal distribution model when matched with Turkey plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig. 4.5. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIII) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9546 and 145.17 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 36a and 36b respectively (Appendix XIII). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 3044.48 m^3/s against a probability of exceedance of 0.0220, invariably within a recurrence interval of approximately 45 years in which the flood may occur.

4.2.5 Blom plotting position.

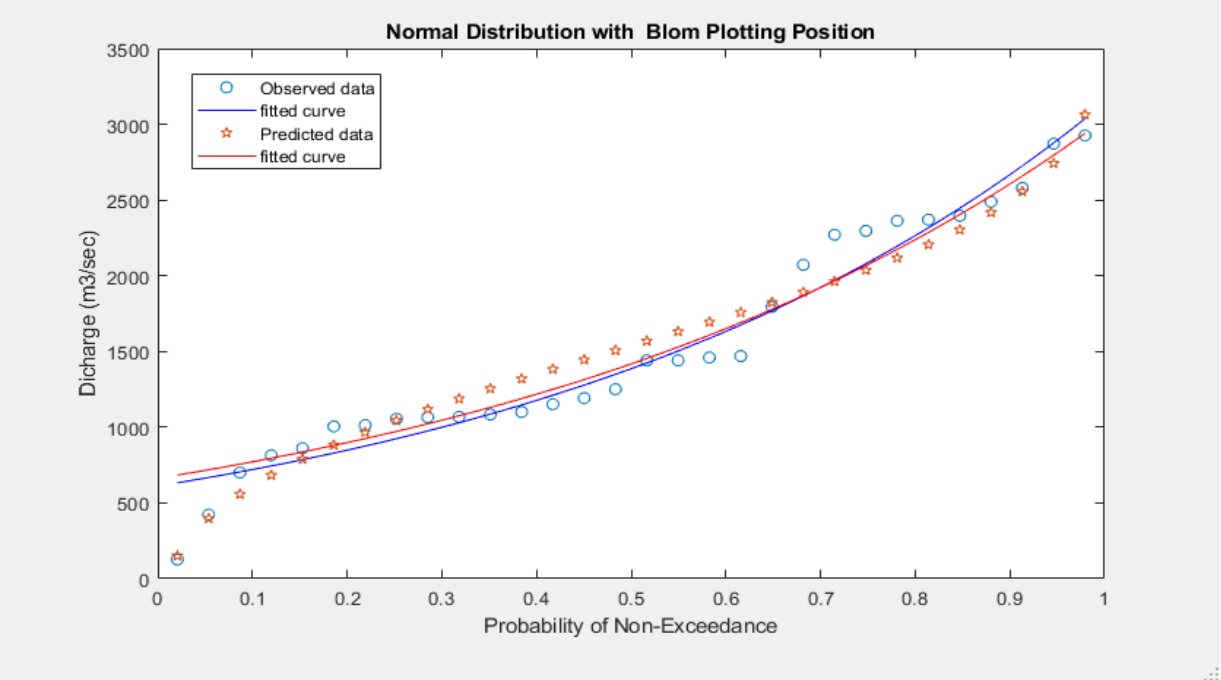


Fig 4.6 Normal distribution curves using Blom plotting position. The observed discharge data and the result of predicted discharges as presented in (Appendix I) using normal distribution model when matched with Blom plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.6. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIII) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9537 and 147.54 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 37a and 37b respectively (Appendix XIII). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 3063.70 m^3/s against a probability of exceedance of 0.0207, invariably within a recurrence interval of approximately 48 years in which the flood may occur.

4.2.6 Cunnane plotting position.

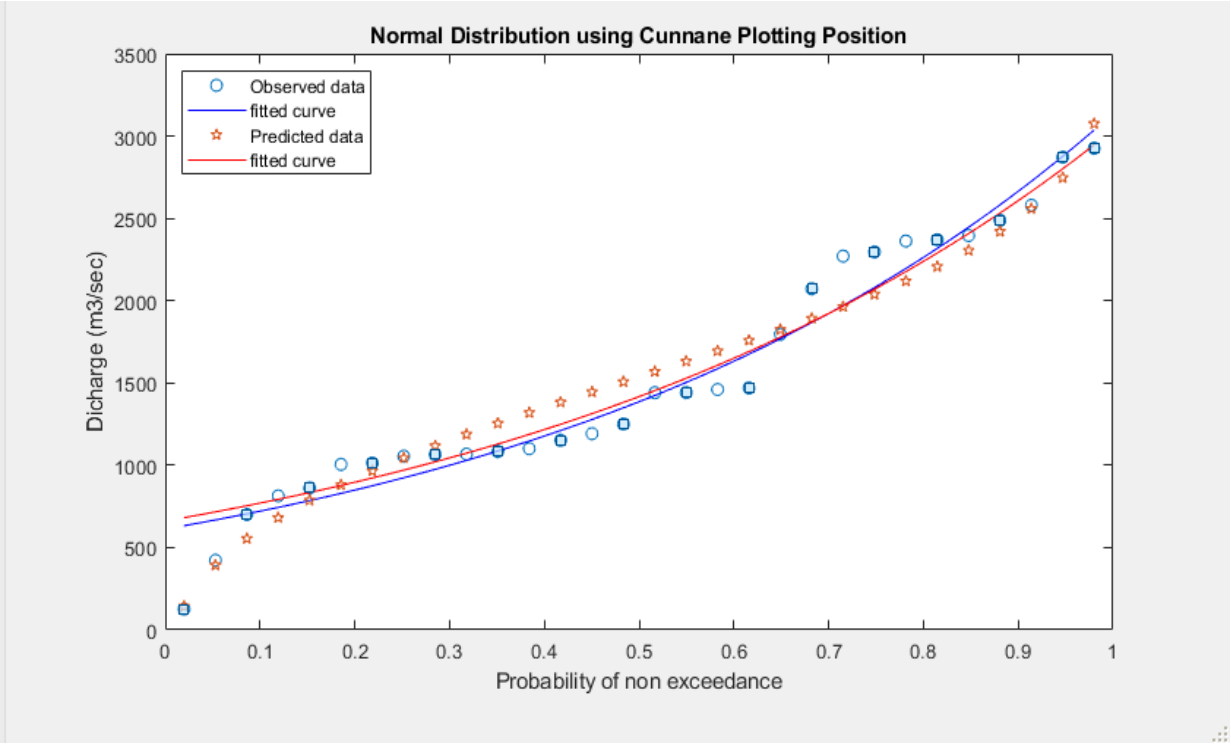


Fig 4.7 Normal distribution curvesusing Cunnane plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix I) using normal distribution model when matched with Cunnane plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.7. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIII) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9531 and 149.04 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 38a and 38b respectively (Appendix XIII). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 3075.85 m^3/s against a probability of exceedance of 0.0199, invariably within a recurrence interval of approximately 50 years in which the flood may occur.

4.2.7 Gringorten plotting position.

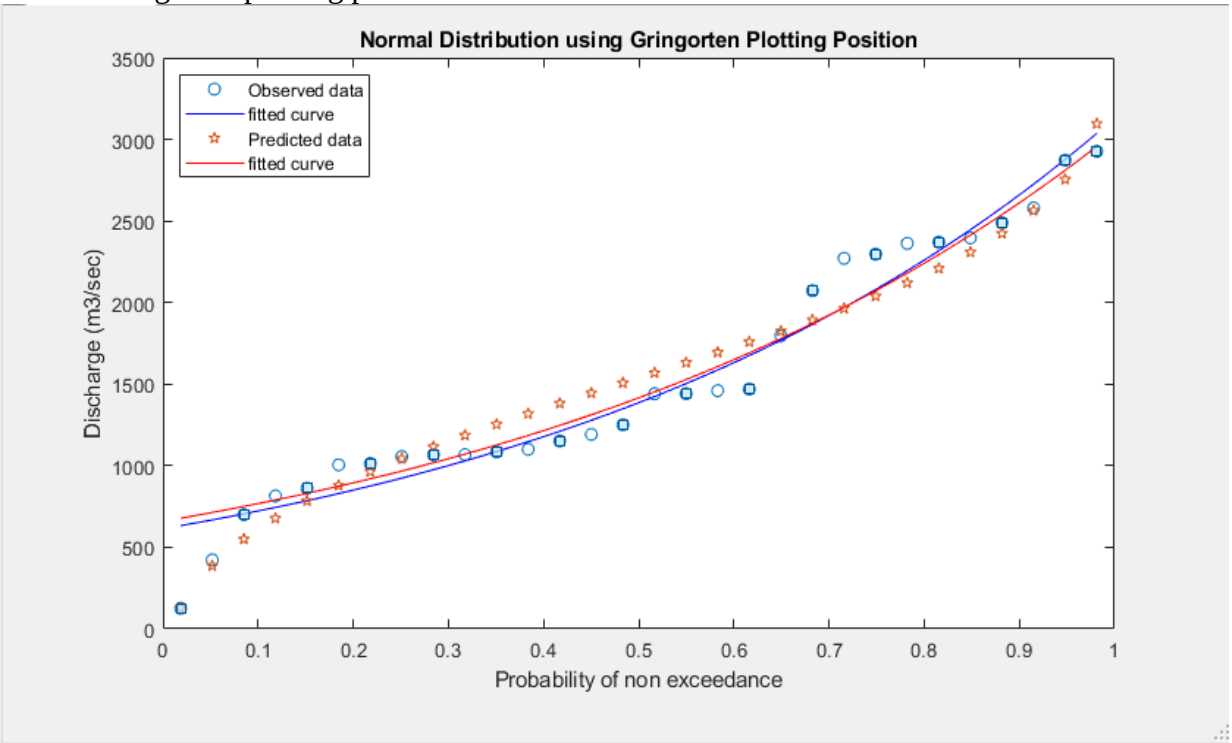


Fig 4.8 Normal distribution curves using Gringorten plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix II) using normal distribution model when matched with Gringorten plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.8. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIII) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9522 and 151.54 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 39a and 39b respectively (Appendix XIII). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 3096.22 m^3/s against a probability of exceedance of 0.0186, invariably within a recurrence interval of approximately 54 years in which the flood may occur.

4.2.8 Hazen plotting position.

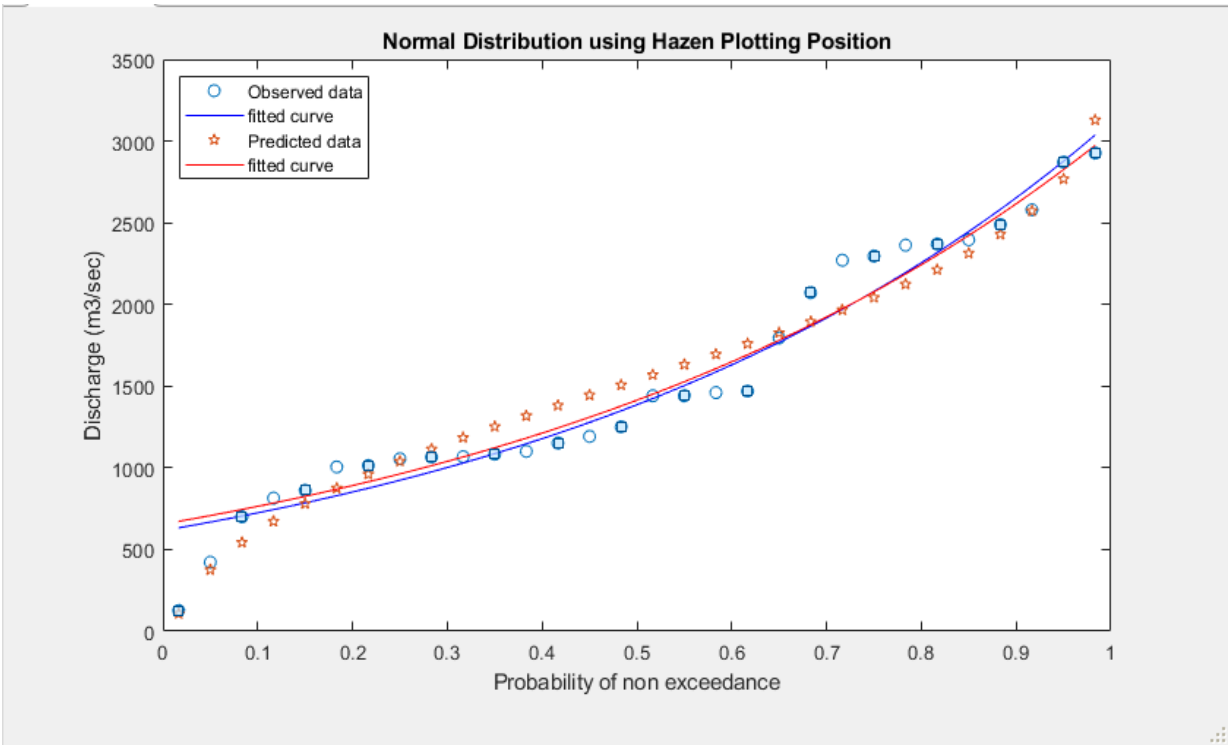


Fig 4.9 Normal distribution curve using Hazen plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix II) using normal distribution model when matched with Gringorten plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.9. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIII) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9506 and 155.58 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 40a and 40b respectively (Appendix XIII). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 3129.33 m^3/s against a probability of exceedance of 0.0167, invariably within a recurrence interval of approximately 60 years in which the flood may occur.

4.3 Log-normal Distribution.

4.3.1 California plotting position.

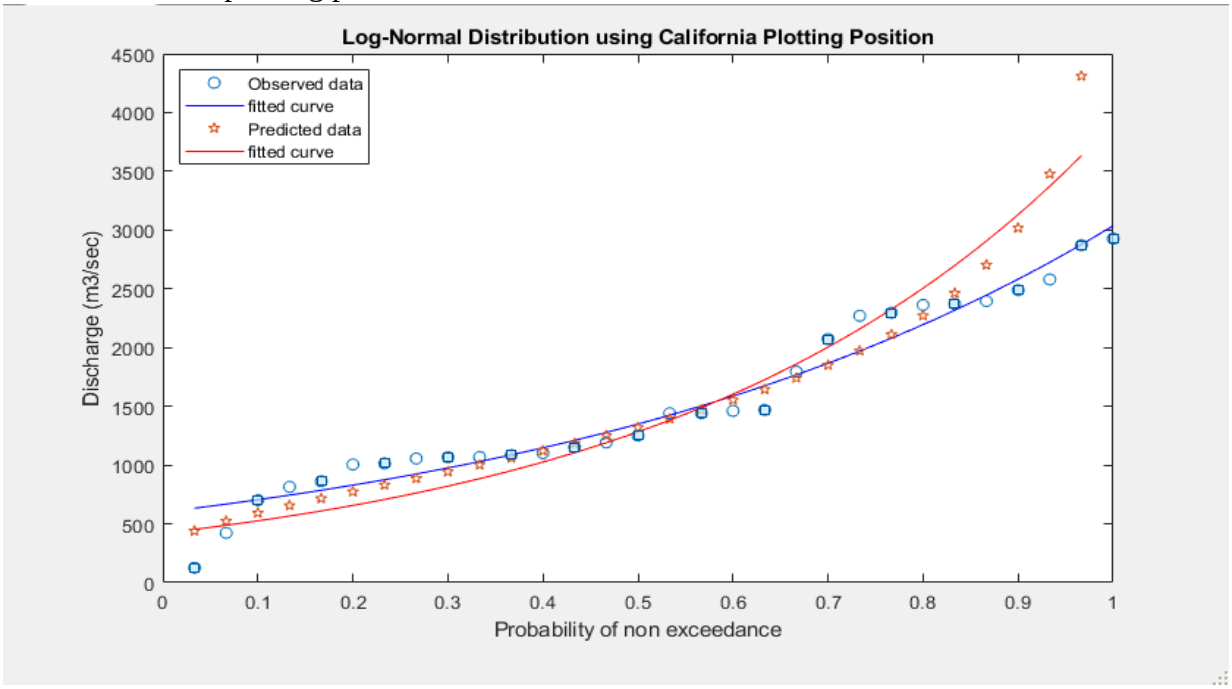


Fig 4.10 Log-normal distribution curves using California plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix III) using log-normal distribution model when matched with California plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.10. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9622 and 175.39 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 41a and 41b respectively (Appendix XIV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 4310.67 m^3/s against a probability of exceedance of 0.0333, invariably within a recurrence interval of approximately 30 years in which the flood may occur.

4.3.2 Weibull plotting position.

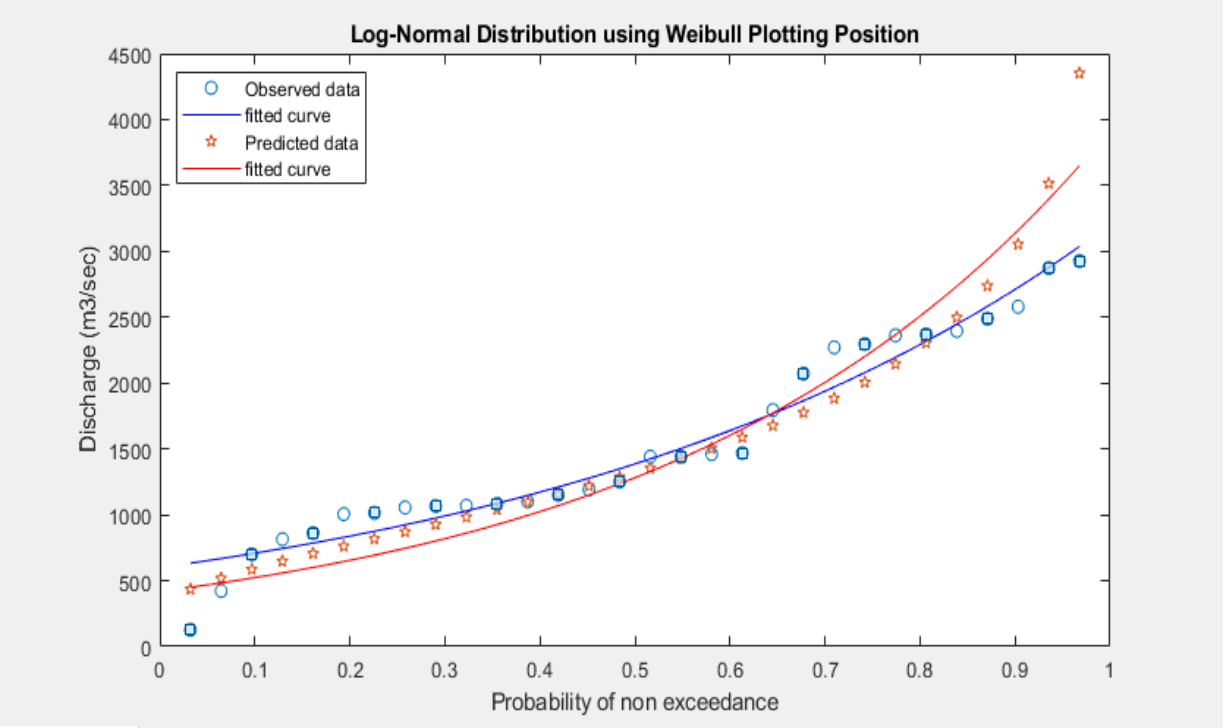


Fig 4.11 Log-normal distribution curves using Weibull plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix IV) using log-normal distribution model when matched with Weibull plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.11. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9612 and 178.47 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 42a and 42b respectively (Appendix XIV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 4351.70 m^3/s against a probability of exceedance of 0.0323, invariably within a recurrence interval of approximately 31 years in which the flood may occur.

4.3.3 Chegodayev plotting position.

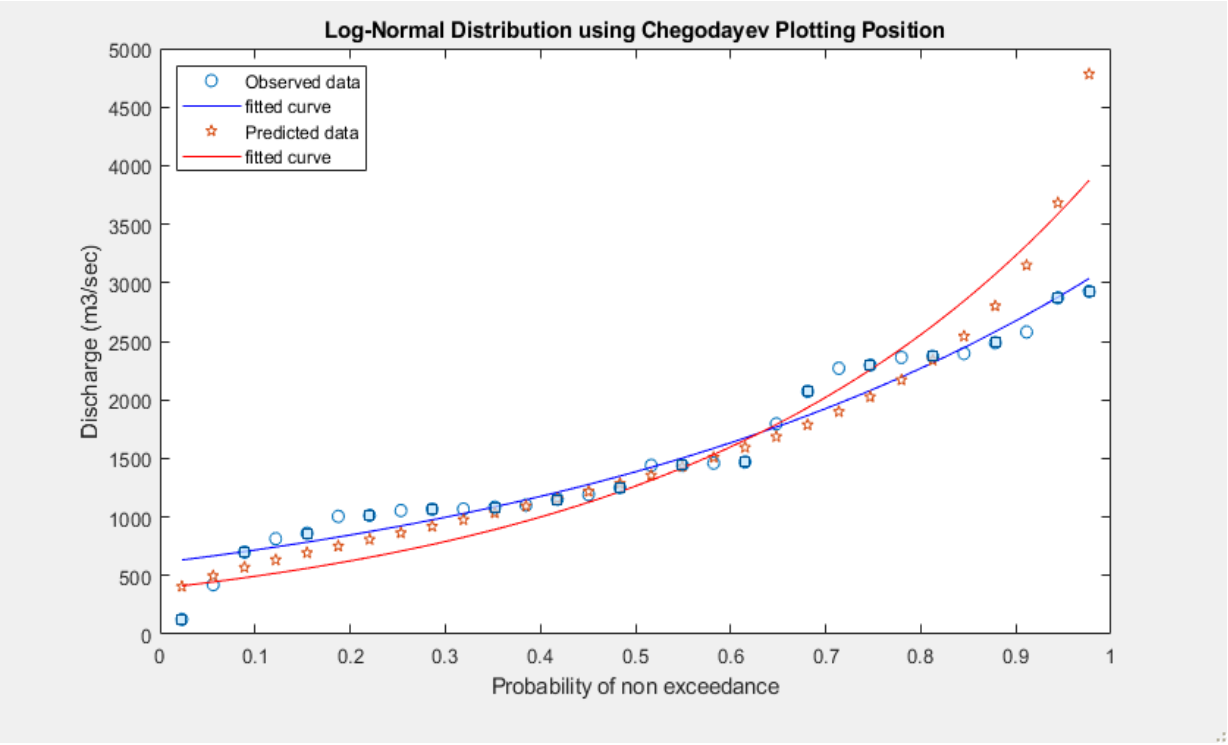


Fig 4.12 Log-normal distribution curves using Chegodayev plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix III) using log-normal distribution model when matched with Chegodayev plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.12. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9466 and 225.93 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 43a and 43b respectively (Appendix XIV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 4782.64 m^3/s against a probability of exceedance of 0.0230, invariably within a recurrence interval of approximately 43 years in which the flood may occur.

4.3.4 Turkey plotting position.

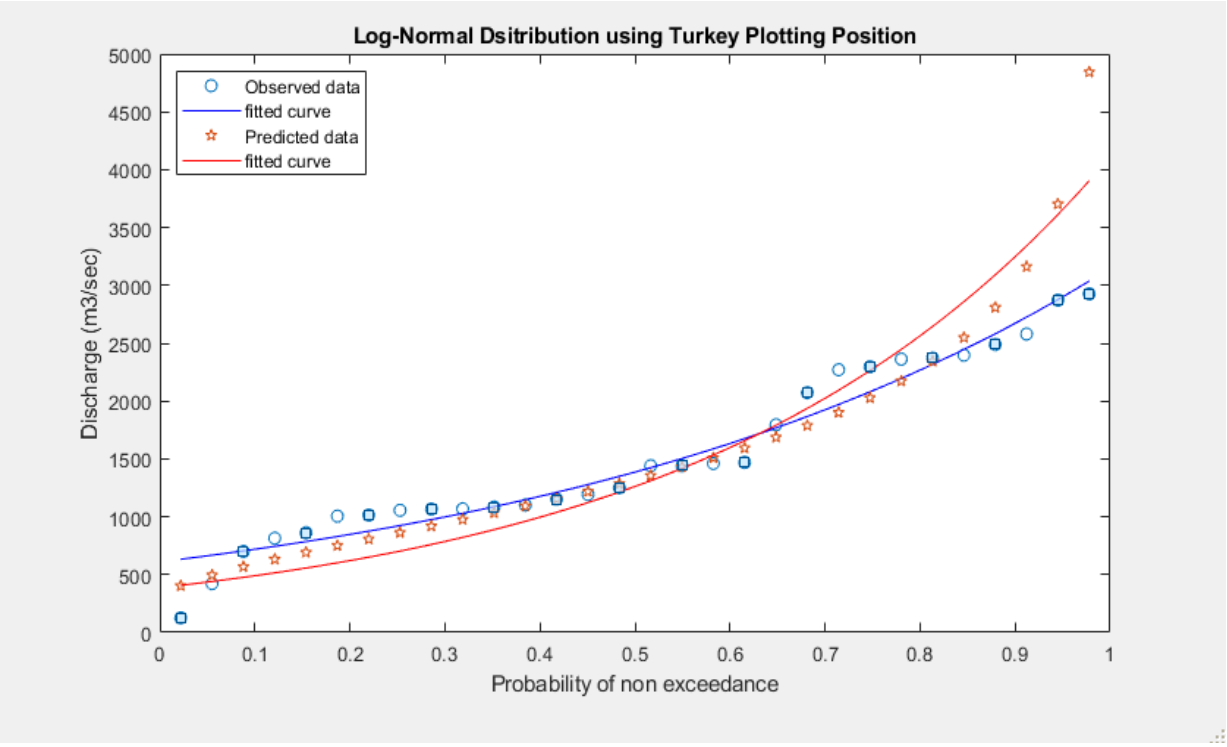


Fig 4.13 Log-normal distribution curves using Turkey plotting position

The observed discharge data and the result of predicted discharges as presented in (Appendix IV) using log-normal distribution model when matched with Turkey plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.13. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9444 and 232.86 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 44a and 44b respectively (Appendix XIV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 4843.50 m^3/s against a probability of exceedance of 0.0220, invariably within a recurrence interval of approximately 45 years in which the flood may occur.

4.3.5 Blom plotting position.

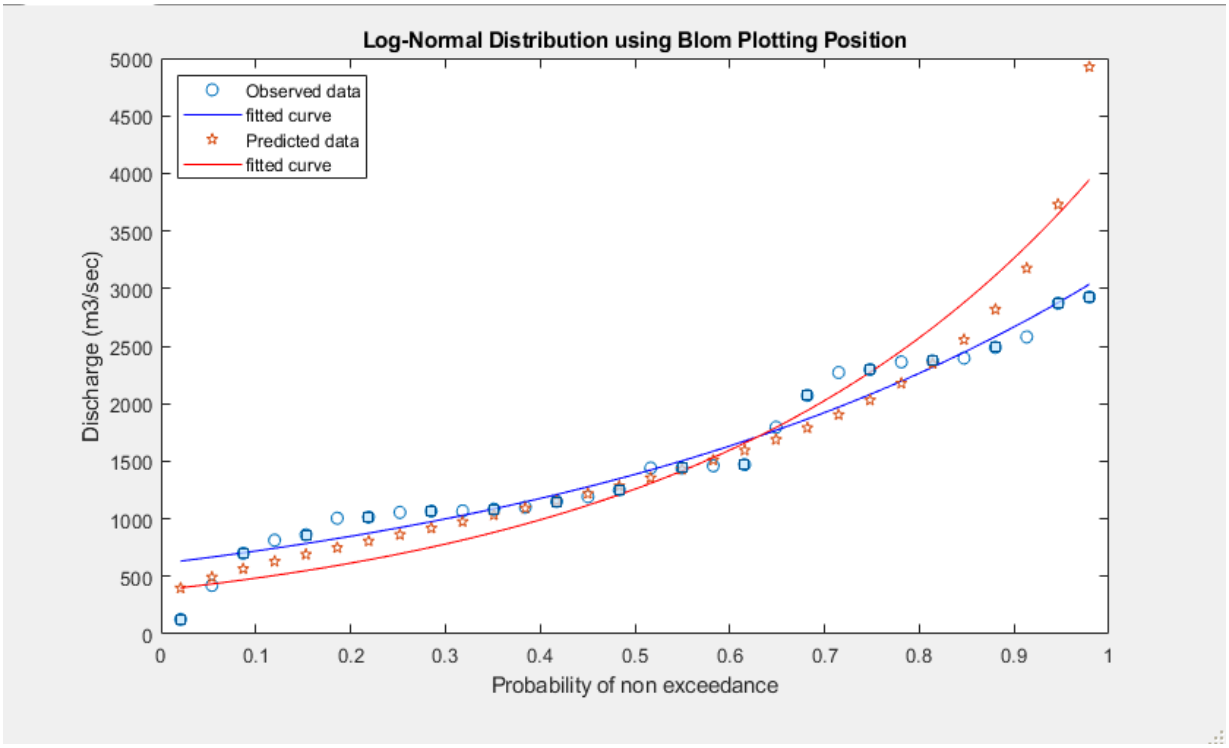


Fig 4.14 Log-normal distribution curves using Blom plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix III) using log-normal distribution model when matched with Blom plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.14. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9413 and 242.21 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 45a and 45b respectively (Appendix XIV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 4924.73 m^3/s against a probability of exceedance of 0.0207, invariably within a recurrence interval of approximately 48 years in which the flood may occur.

4.3.6 Cunnane plotting position.

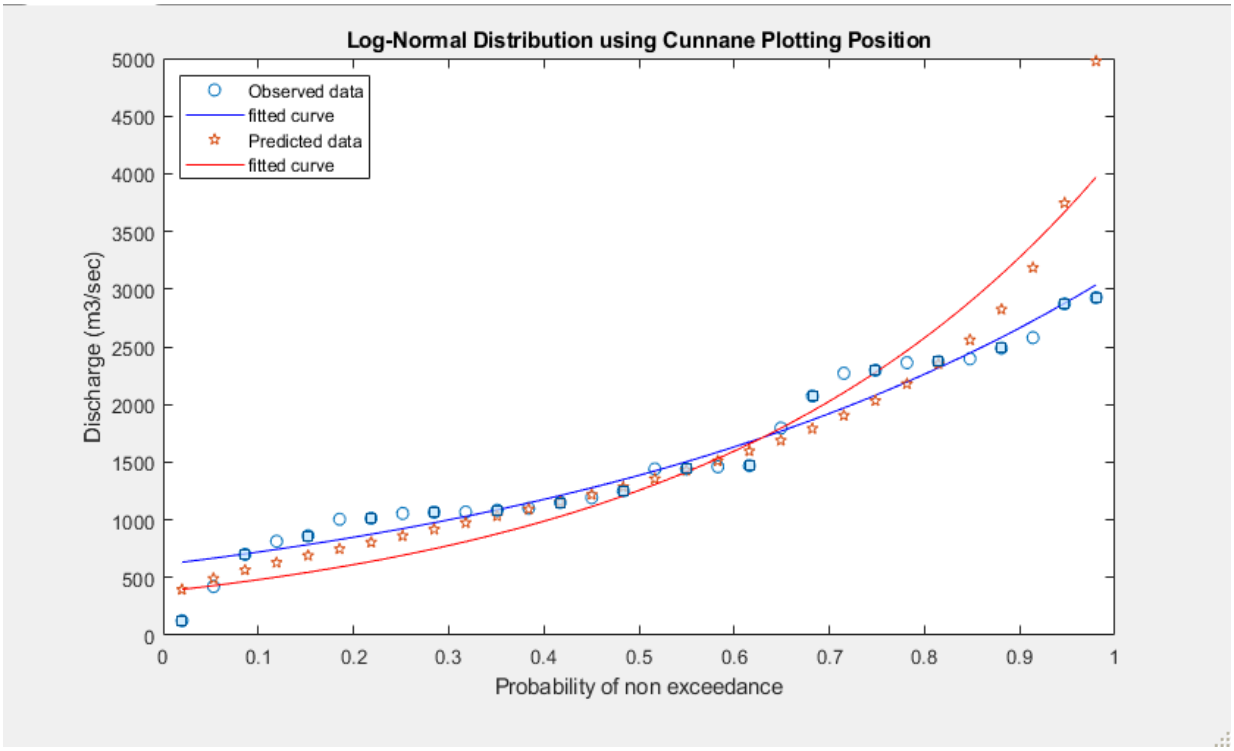


Fig 4.15 Log-normal distribution curves using Cunnane plotting position

The observed discharge data and the result of predicted discharges as presented in (Appendix III) using log-normal distribution model when matched with Cunnane plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.15. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9394 and 248.41 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 46a and 46b respectively (Appendix XIV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 4976.48 m^3/s against a probability of exceedance of 0.0199, invariably within a recurrence interval of approximately 50 years in which the flood may occur.

4.3.7 Gringorten plotting position.

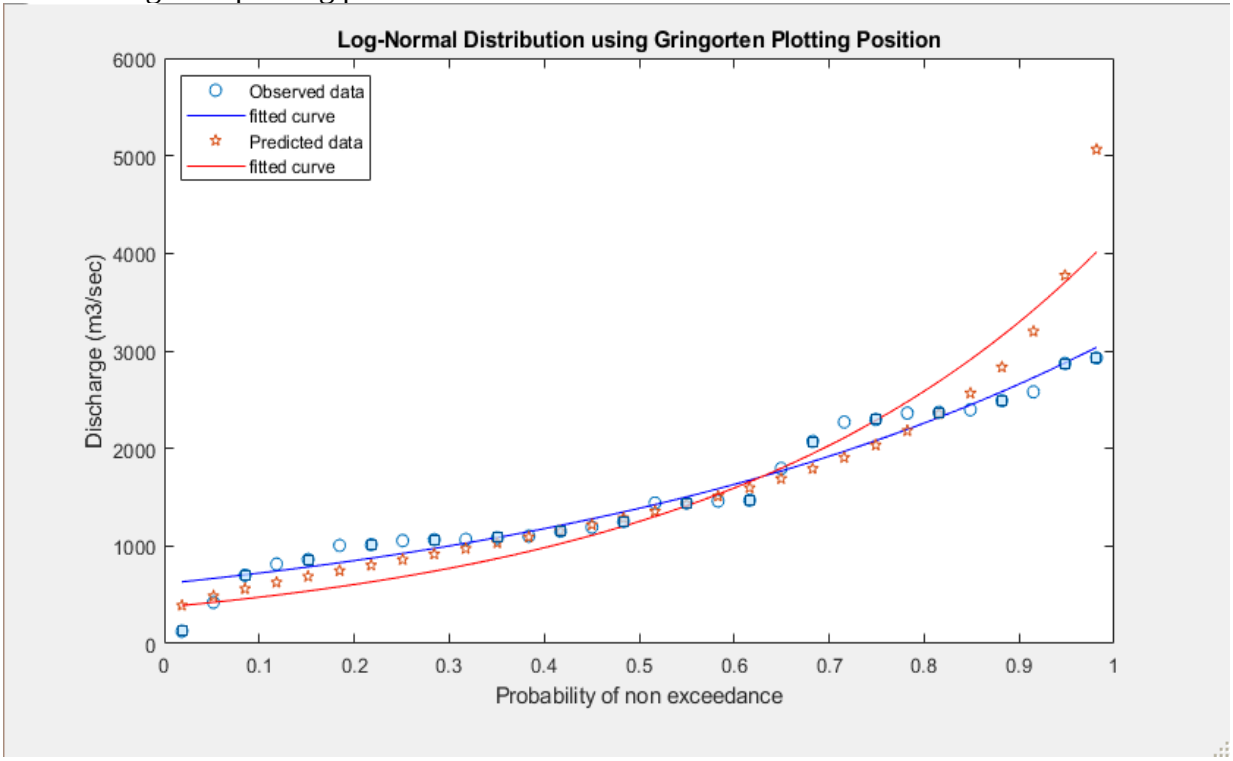


Fig 4.16 Log-normal distribution curves using Gringorten plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix IV) using log-normal distribution model when matched with Gringorten plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.16. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9360 and 258.52 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 47a and 47b respectively (Appendix XIV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 5064.78 m^3/s against a probability of exceedance of 0.0186, invariably within a recurrence interval of approximately 54 years in which the flood may occur.

4.3.8 Hazen plotting position.

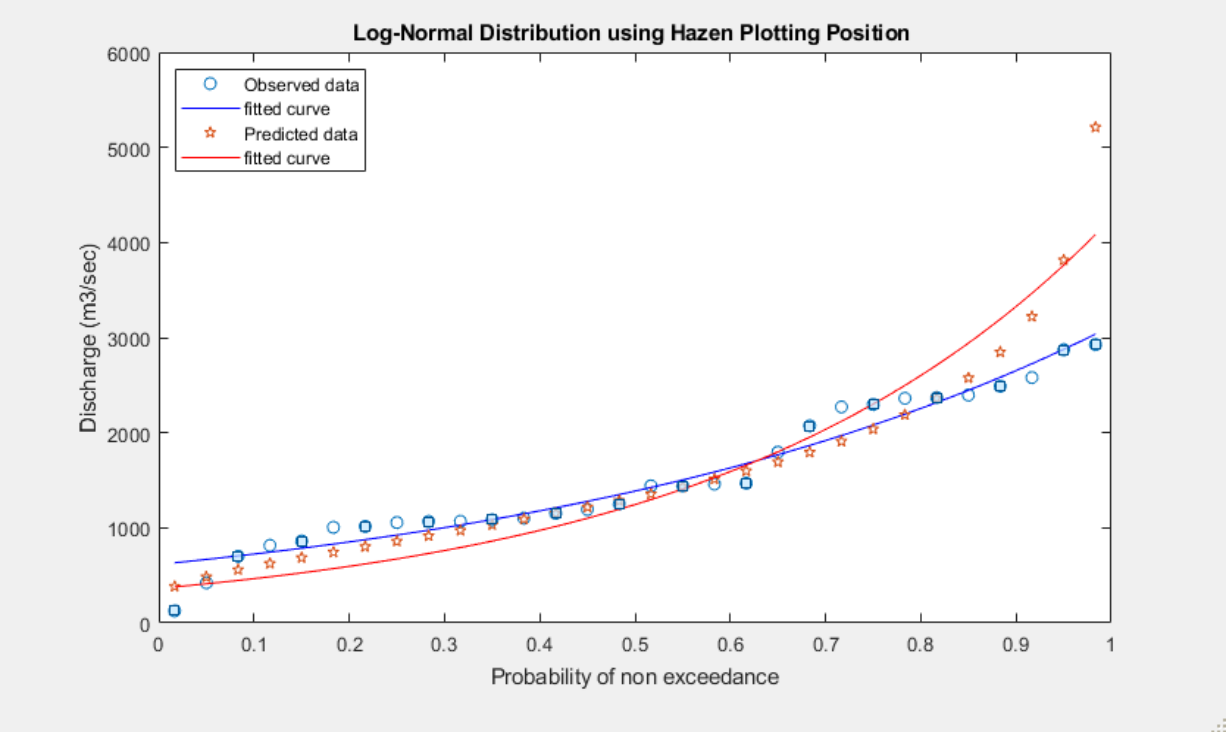


Fig 4.17 Log-normal distribution curves using Hazen plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix IV) using log-normal distribution model when matched with Hazen plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.17. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XIV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9301 and 275.91 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 48a and 48b respectively (Appendix XIV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 5211.64 m^3/s against a probability of exceedance of 0.0167, invariably within a recurrence interval of approximately 60 years in which the flood may occur.

4.4 Pearson Type III Distribution.

4.4.1 California plotting position.

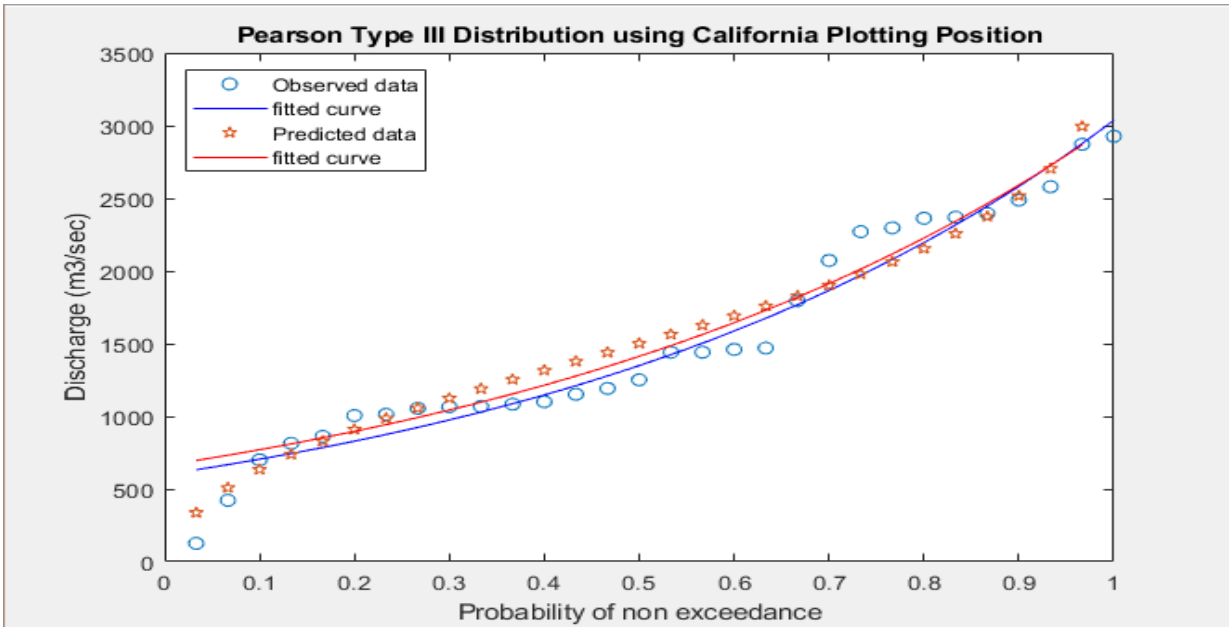


Fig 4.18 Pearson type III distribution curves California plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix V) using Pearson type III distribution model when matched with California plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.18. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9728 and 107.04 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 49a and 49b respectively (Appendix XV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 2994.44 m^3/s against a probability of exceedance of0.0333, invariably within a recurrence interval of approximately 30 years in which the flood may occur.

4.4.2 Weibull plotting position.

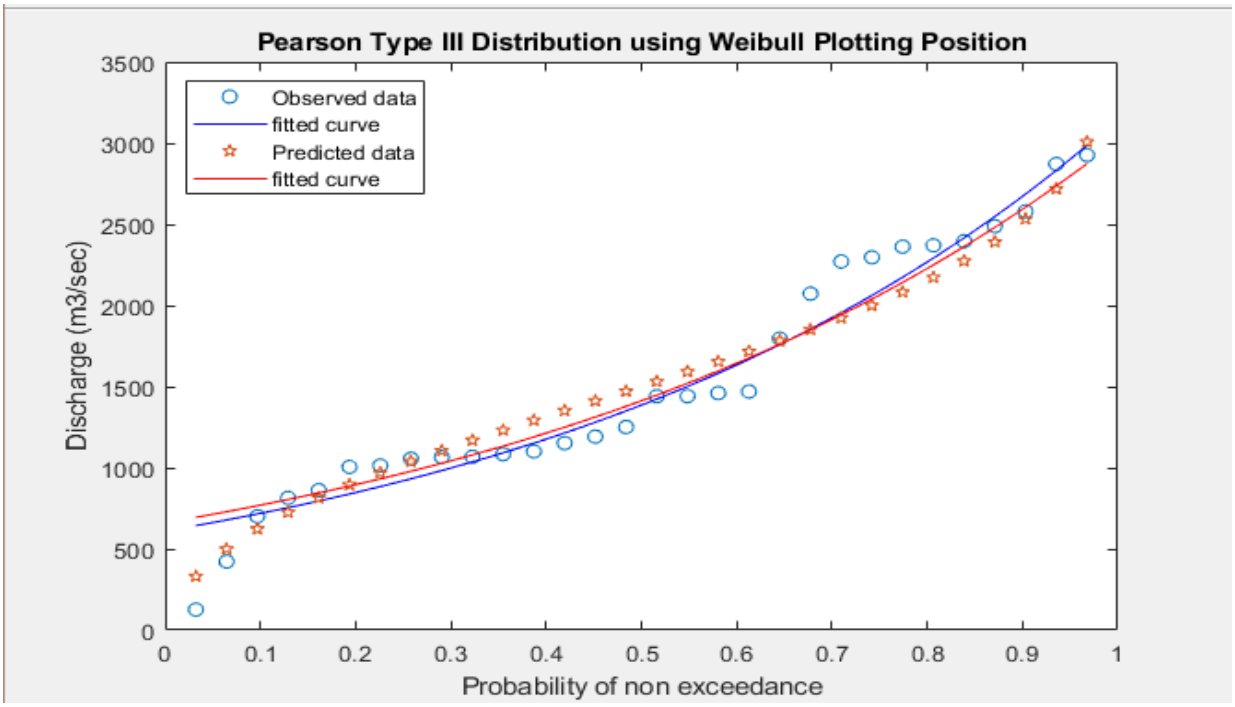


Fig 4.19 Pearson type III distribution curves Weibull plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix VI) using Pearson type III distribution model when matched with Weibull plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.19. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9725 and 107.86 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 50a and 50b respectively (Appendix XV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 3007.50 m^3/s against a probability of exceedance of 0.0323, invariably within a recurrence interval of approximately 31 years in which the flood may occur.

4.4.3Chegodayev plotting position.

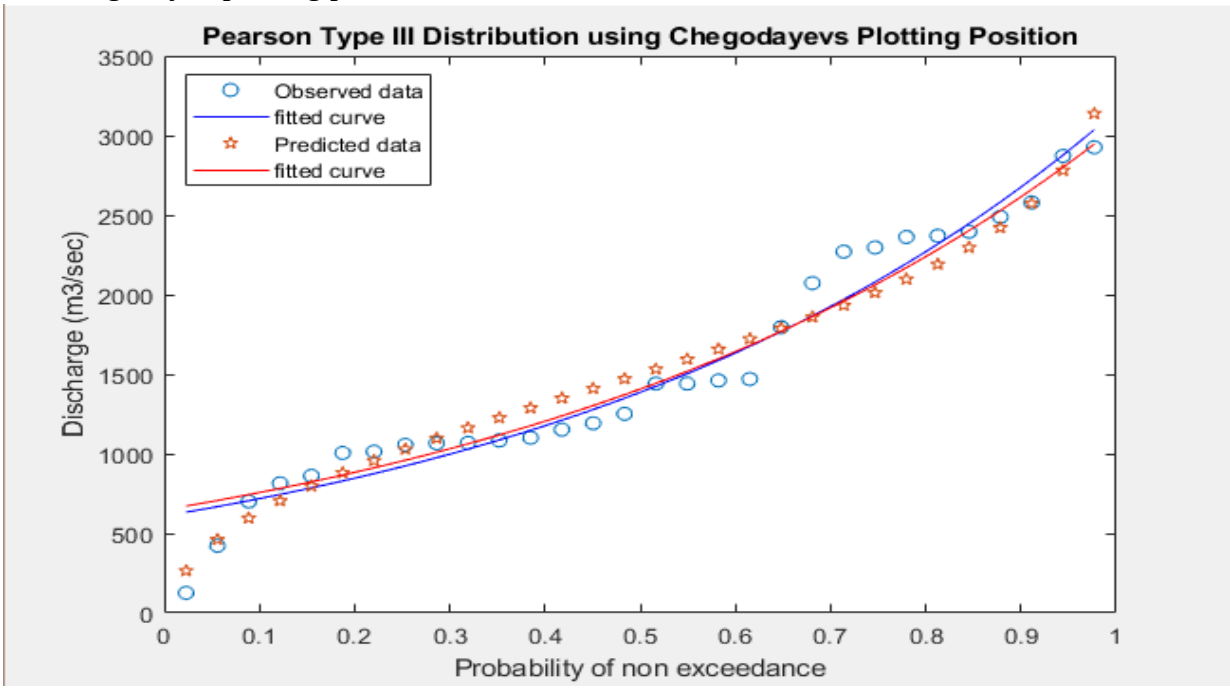


Fig 4.20 Pearson type III distributioncurves Chegodayevs plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix V) using Pearson type III distribution model when matched with Chegodayev plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.20. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9682 and 121.26 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 51a and 51b respectively (Appendix XV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 3138.22 m^3/s against a probability of exceedance of 0.0230, invariably within a recurrence interval of approximately 43 years in which the flood may occur.

.4.4.4 Turkey plotting position.

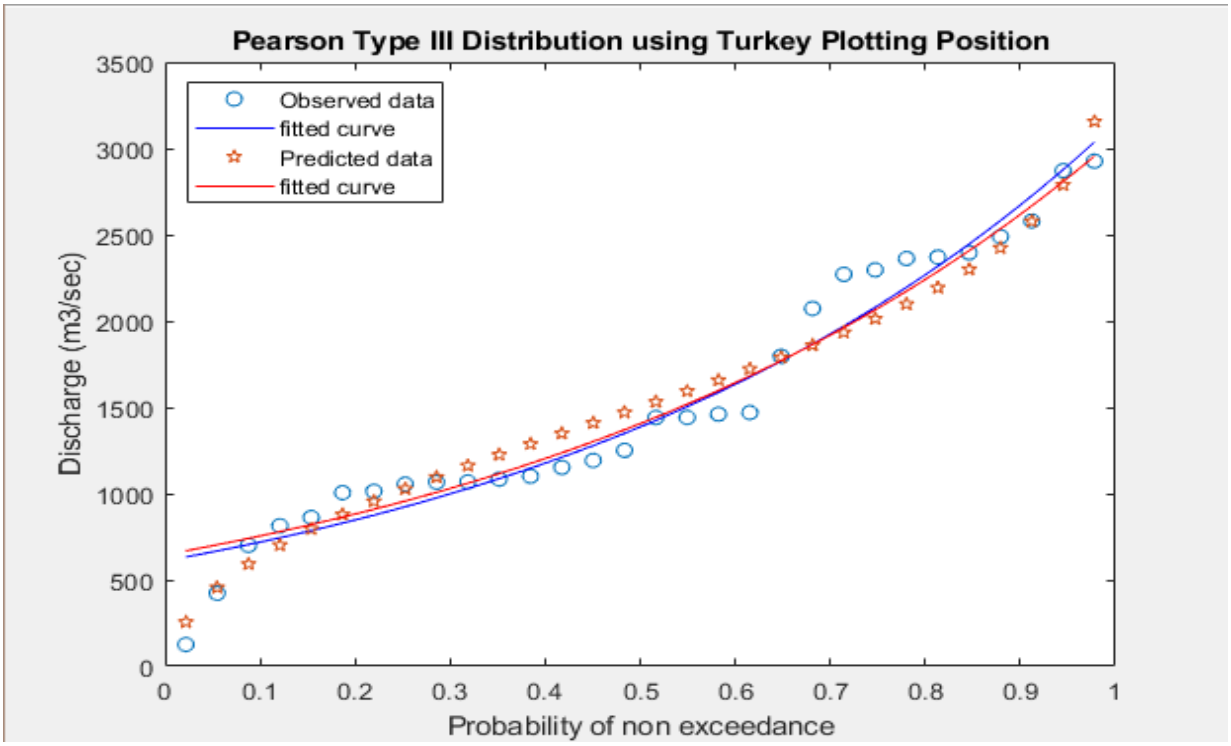


Fig 4.21 Pearson type III curves distribution Turkey plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix VI) using Pearson type III distribution model when matched with Turkey plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.21. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9676 and 123.09 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 52a and 52b respectively (Appendix XV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 3155.90 m^3/s against a probability of exceedance of 0.0220, invariably within a recurrence interval of approximately 45 years in which the flood may occur.

4.4.5 Blom plotting position.

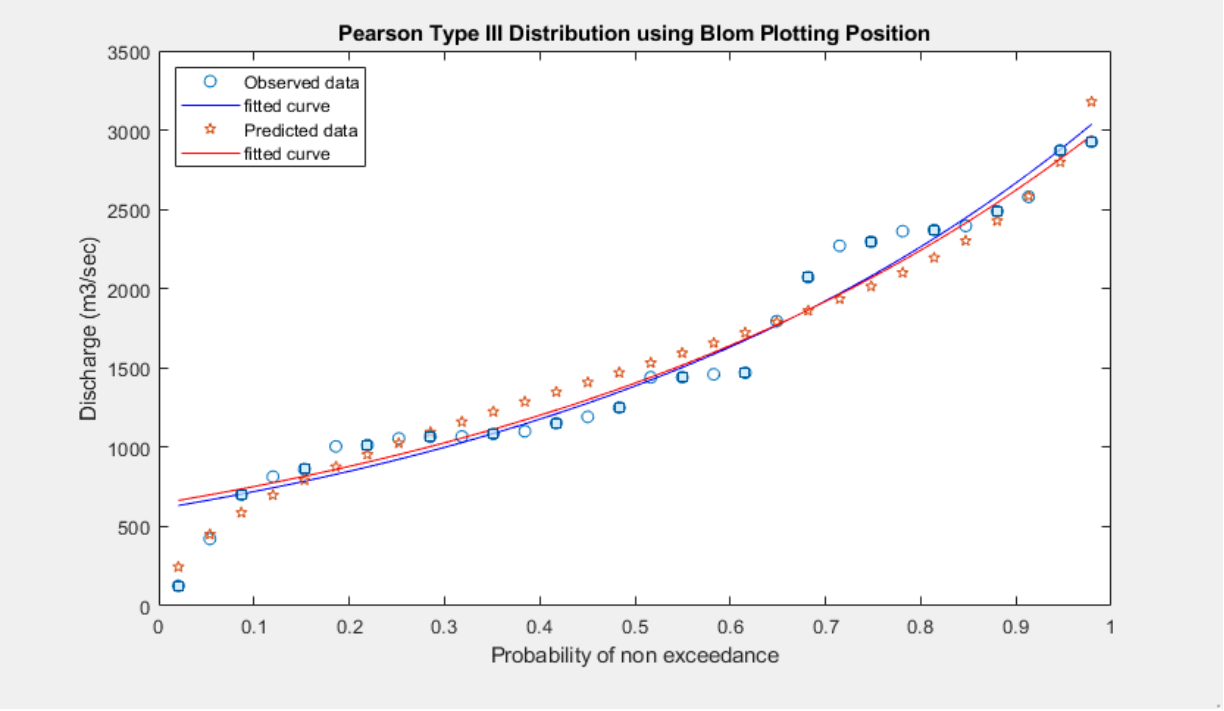


Fig 4.22 Pearson type III curves distribution using Blom plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix V) using Pearson type III distribution model when matched with Blom plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.22. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9667 and 125.52 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 53a and 53b respectively (Appendix XV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 3179.12 m^3/s against a probability of exceedance of 0.0207, invariably within a recurrence interval of approximately 48 years in which the flood may occur.

4.4.6 Cunnane plotting position.

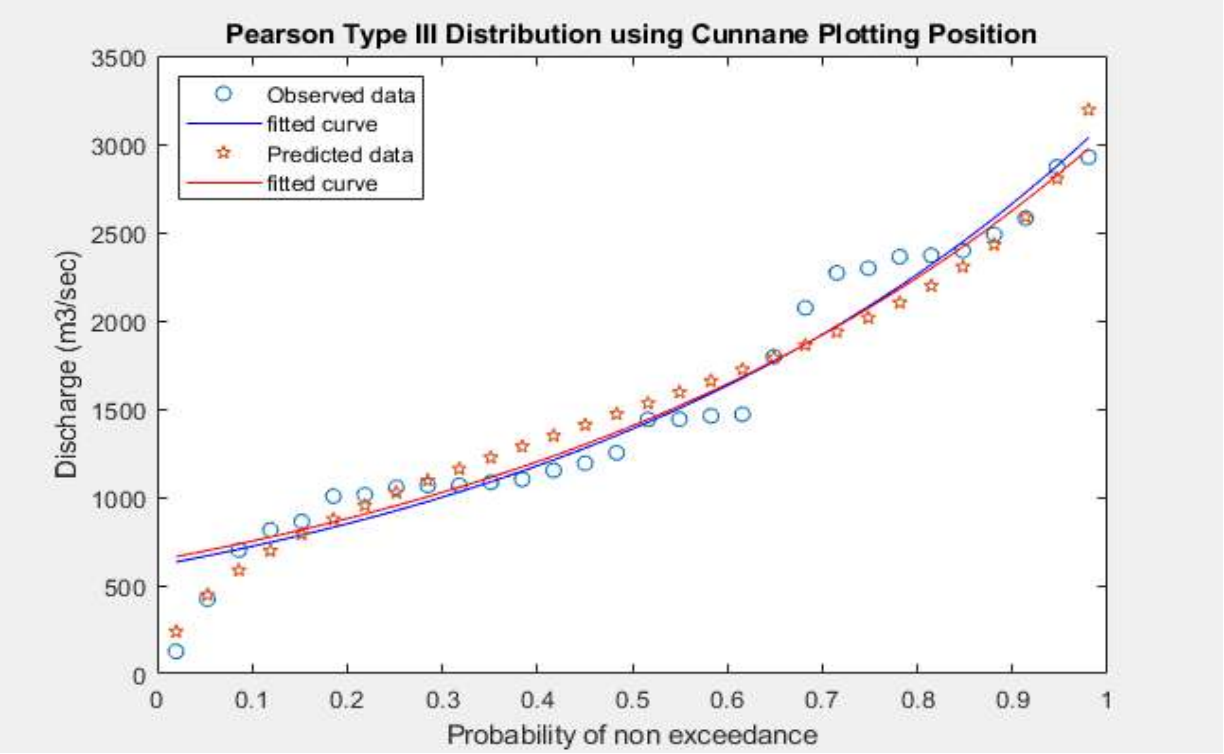


Fig 4.23 Pearson type III distribution curves Cunnane plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix V) using Pearson type III distribution model when matched with Cunnane plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.23. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9662 and 127.05 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 54a and 54b respectively (Appendix XV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 3193.76 m^3/s against a probability of exceedance of 0.0199, invariably within a recurrence interval of approximately 50 years in which the flood may occur.

4.4.7 Gringorten plotting position.

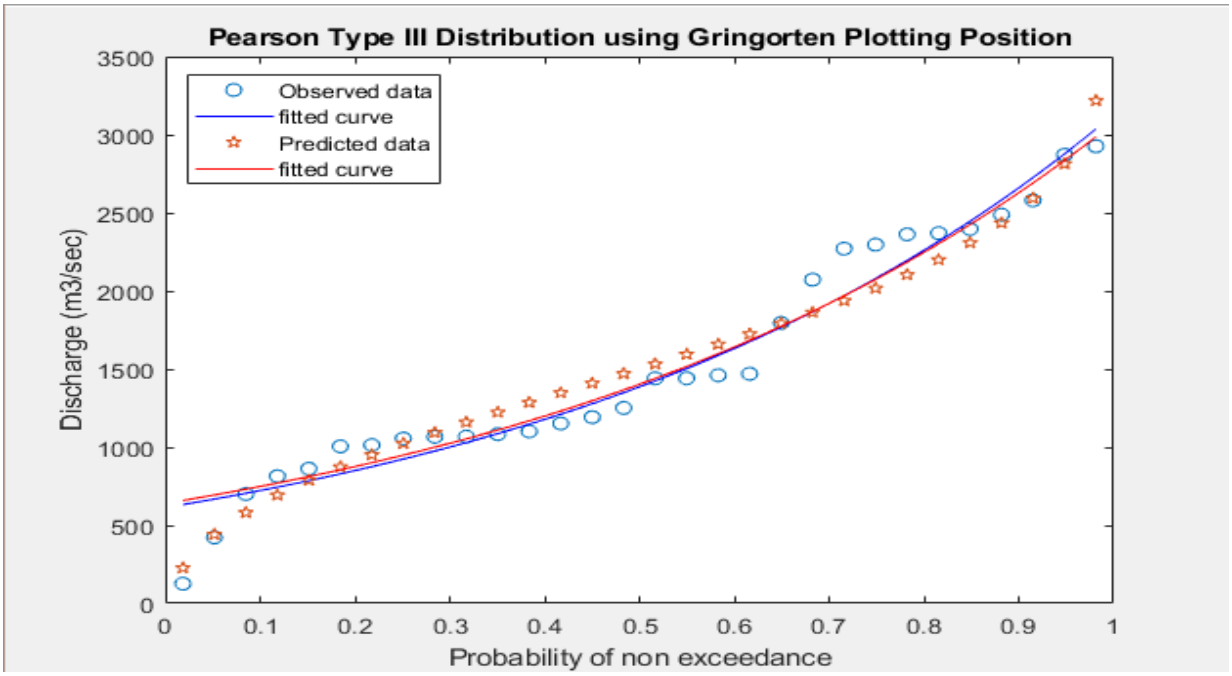


Fig 4.24 Pearson type III Distribution Gringorten plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix VI) using Pearson type III distribution model when matched with Gringorten plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.24. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9653 and 129.64 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 55a and 55b respectively (Appendix XV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 3218.40 m^3/s against a probability of exceedance of 0.0186, invariably within a recurrence interval of approximately 54 years in which the flood may occur.

4.4.8 Hazen plotting position.

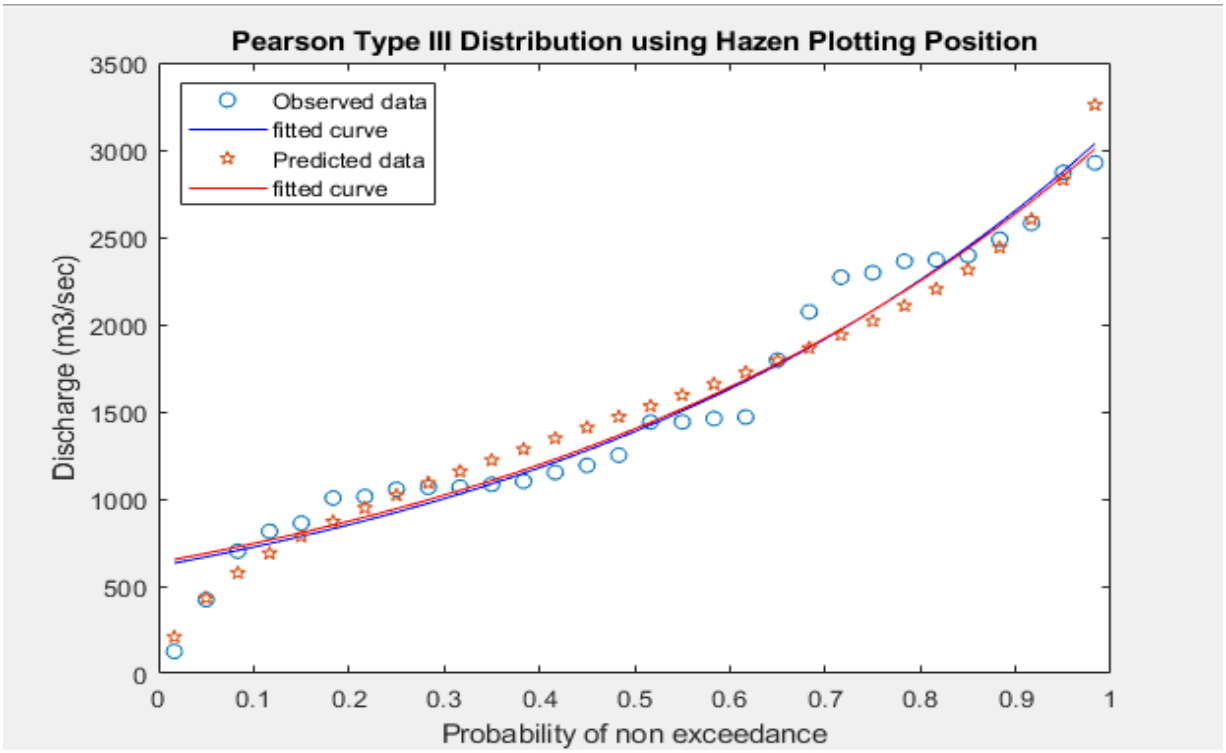


Fig 4.25 Pearson type III Distribution Hazen plotting position

The observed discharge data and the result of predicted discharges as presented in (Appendix VI) using Pearson type III distribution model when matched with Hazen plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.25. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XV) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9638 and 133.90 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 56a and 56b respectively (Appendix XV). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 3258.70 m^3/s against a probability of exceedance of 0.0167, invariably within a recurrence interval of approximately 60 years in which the flood may occur.

4.5 Log-Pearson Type III Distribution.

4.5.1 California plotting position.

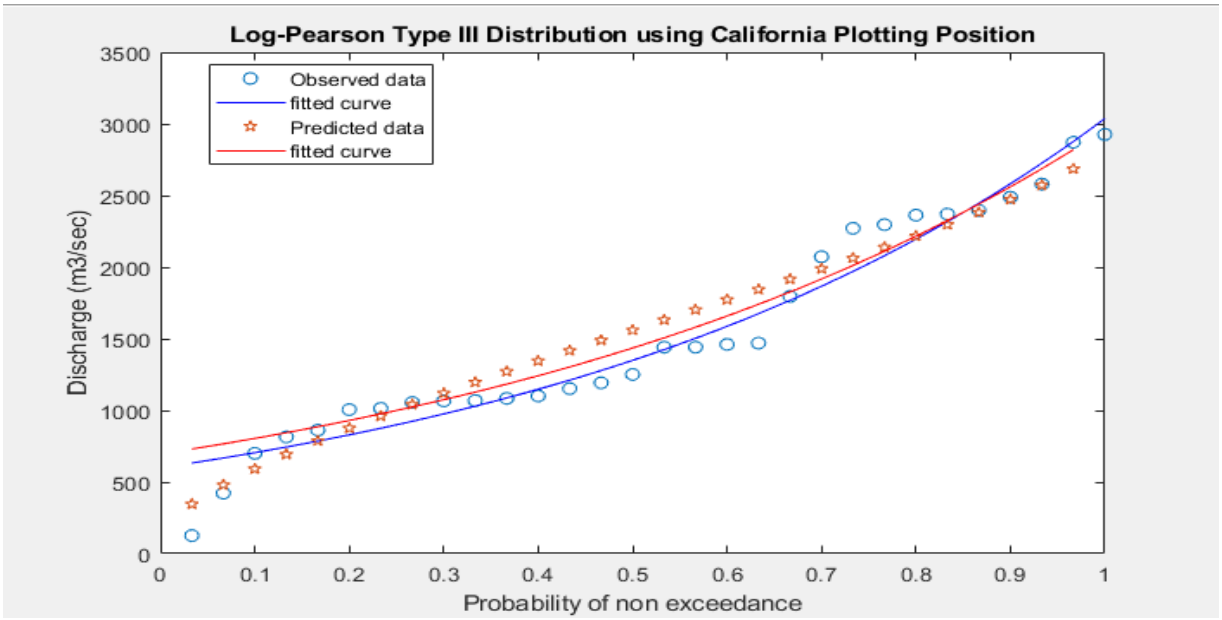


Fig 4.26 Log-Pearson type III distribution curve using California plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix VII) using Log-Pearson type III distribution model when matched with California plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.26. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XVI) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9584 and 130.57 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 57a and 57b respectively (Appendix XVI). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 2685.40 m^3/s against a probability of exceedance of 0.0333, invariably within a recurrence interval of approximately 30 years in which the flood may occur.

4.5.2 Weibull plotting position.

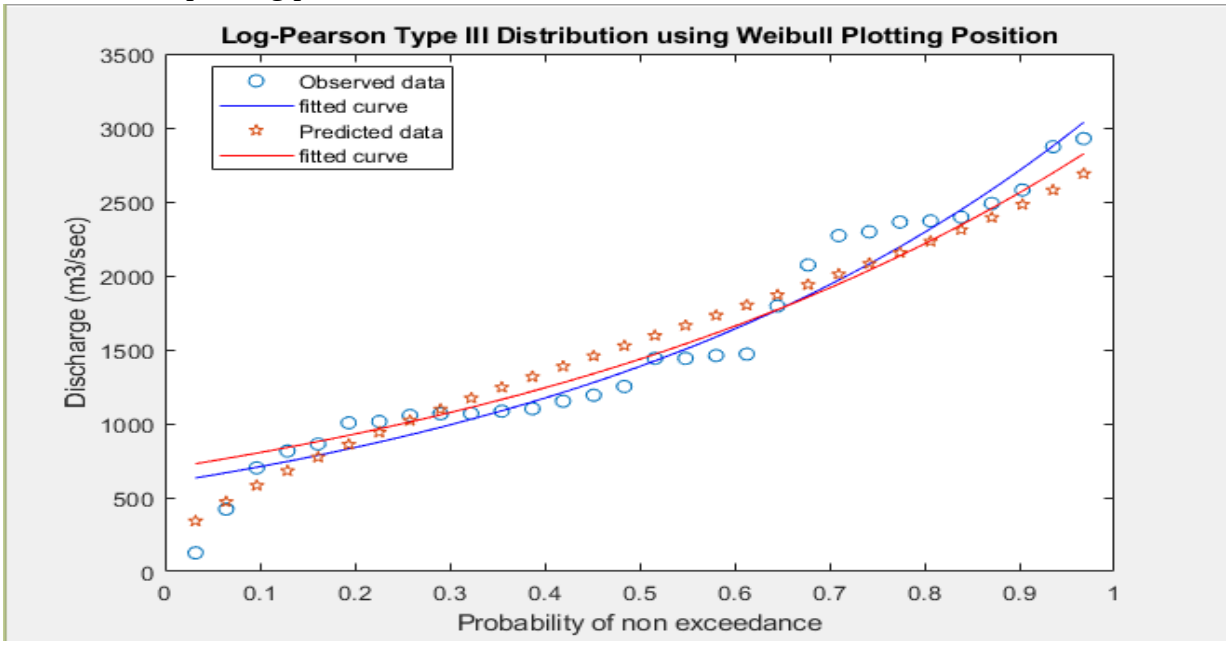


Fig 4.27 Log-Pearson type III curves distribution curves using Weibull plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix VIII) using Log-Pearson type III distribution model when matched with Weibull plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.27. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XVI) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9583 and 130.99 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 58a and 58b respectively (Appendix XVI). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 2689.20 m^3/s against a probability of exceedance of

0.0323, invariably within a recurrence interval of approximately 31 years in which the flood may occur.

4.5.3 Chegodayev plotting position.

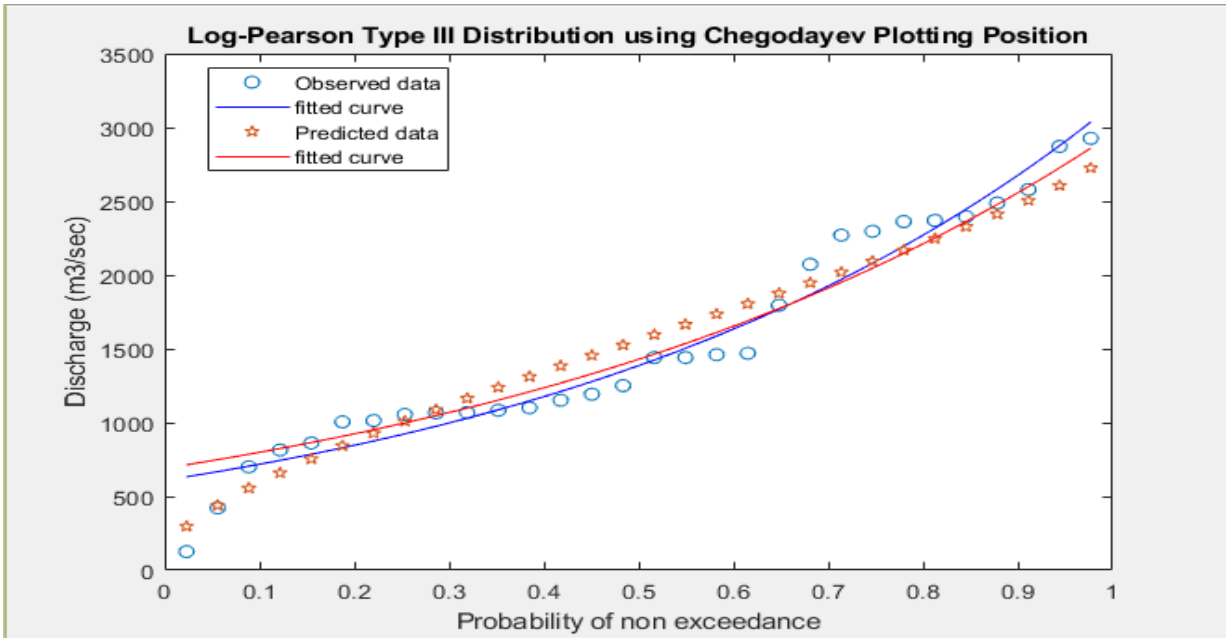


Fig. 4.28 Log-Pearson type III distribution curves using Chegodayev plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix VII) using Log-Pearson type III distribution model when matched with Chegodayev plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.28. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XVI) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9559 and 138.25 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 59a and 59b respectively (Appendix XVI). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an

estimated maximum annual discharge of 2725.60 m^3/s against a probability of exceedance of 0.0230, invariably within a recurrence interval of approximately 43 years in which the flood may occur.

4.5.4 Turkey plotting position.

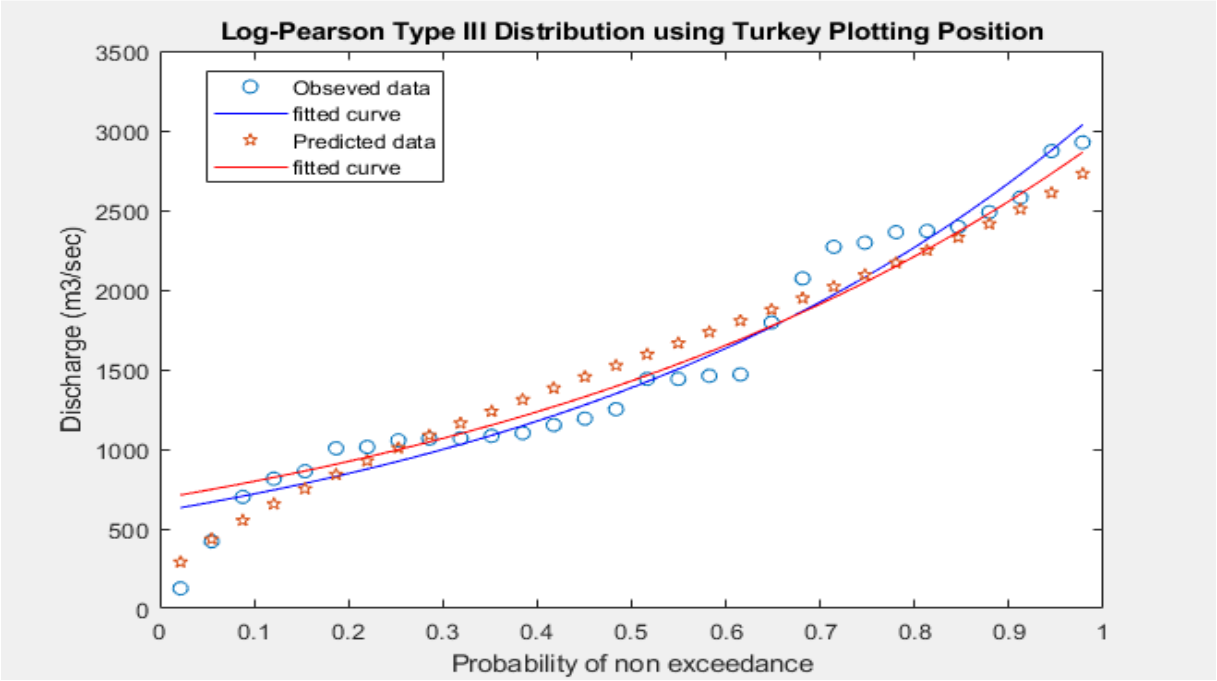


Fig 4.29 Log-Pearson type III distribution curves using Turkey plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix VIII) using Log-Pearson type III distribution model when matched with Turkey plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.29. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XVI) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9556 and 139.11 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 60a and 60b respectively (Appendix XVI). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend

to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 2729.70 m^3/s against a probability of exceedance of 0.0220, invariably within a recurrence interval of approximately 45 years in which the flood may occur.

4.5.5 Blom plotting position.

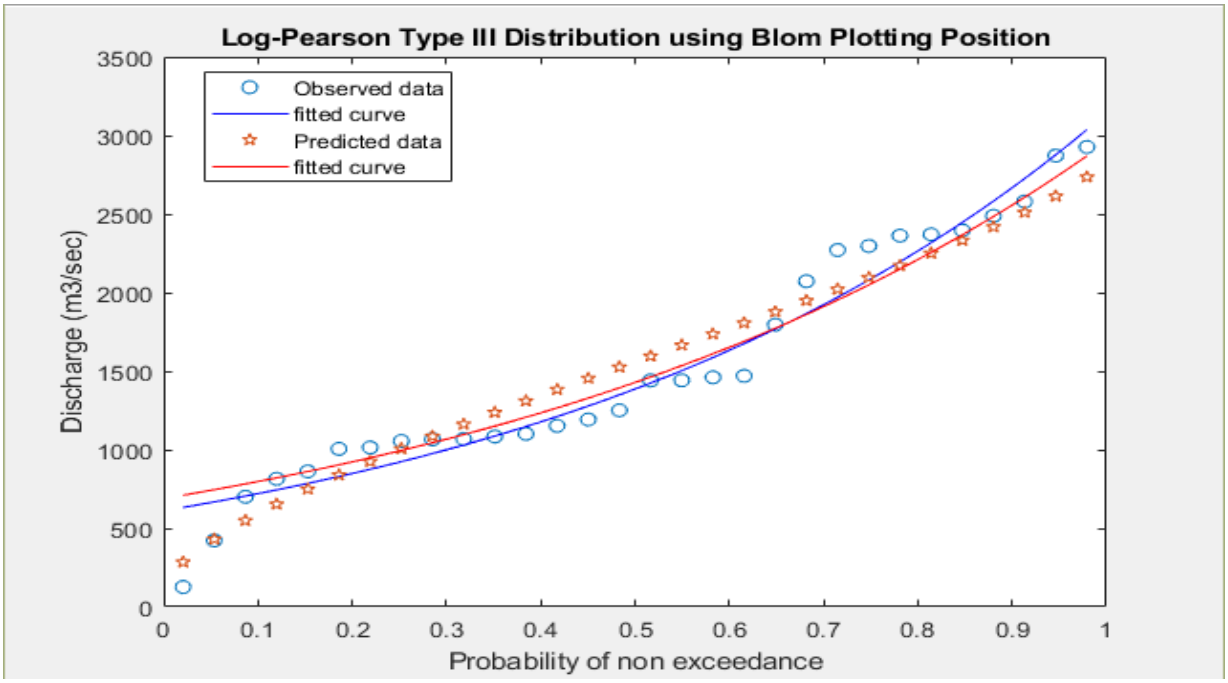


Fig 4.30 Log-Pearson type III distribution curves using Blom plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix VII) using Log-Pearson type III distribution model when matched with Blom plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.30. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XVI) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9553 and 140.21 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 61a and 61b respectively (Appendix XVI). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend

to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 2735.10 m^3/s against a probability of exceedance of 0.0207, invariably within a recurrence interval of approximately 48 years in which the flood may occur.

4.5.6 Cunnane plotting position.

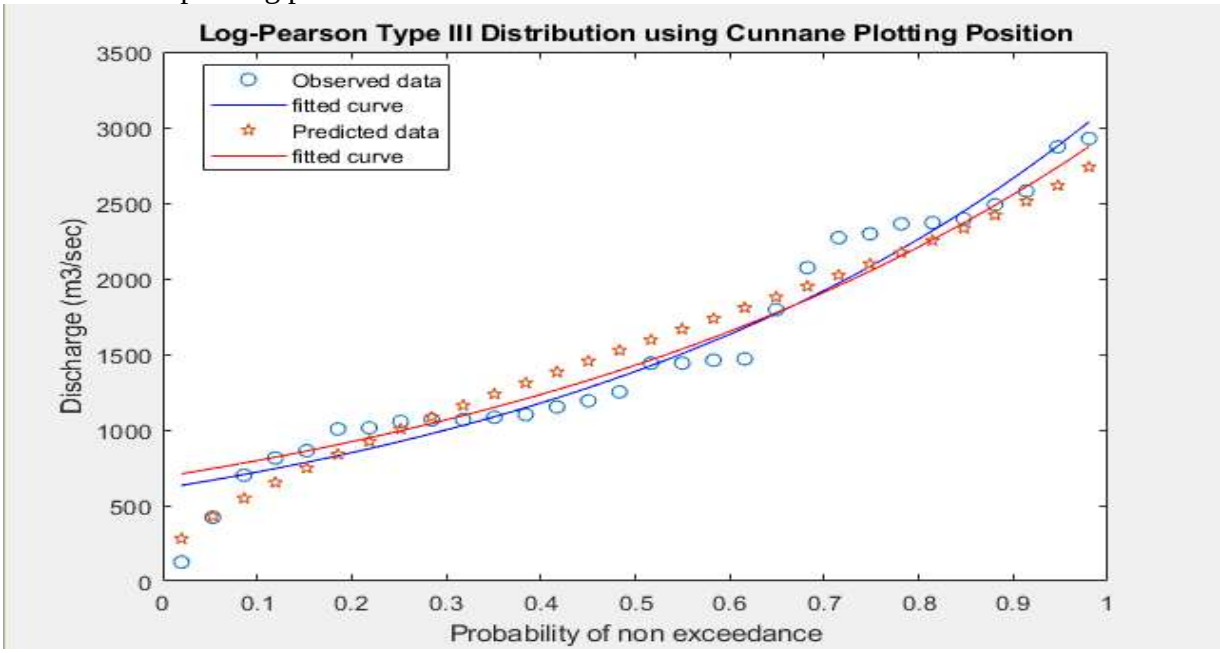


Fig 4.31 Log-Pearson type III distribution curves using Cunnane plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix VII) using Log-Pearson type III distribution model when matched with Cunnane plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.31. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XVI) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9550 and 140.88 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 62a and 62b respectively (Appendix XVI). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher

degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 2738.40 m^3/s against a probability of exceedance of 0.0199, invariably within a recurrence interval of approximately 50 years in which the flood may occur.

4.5.7 Gringorten plotting position.

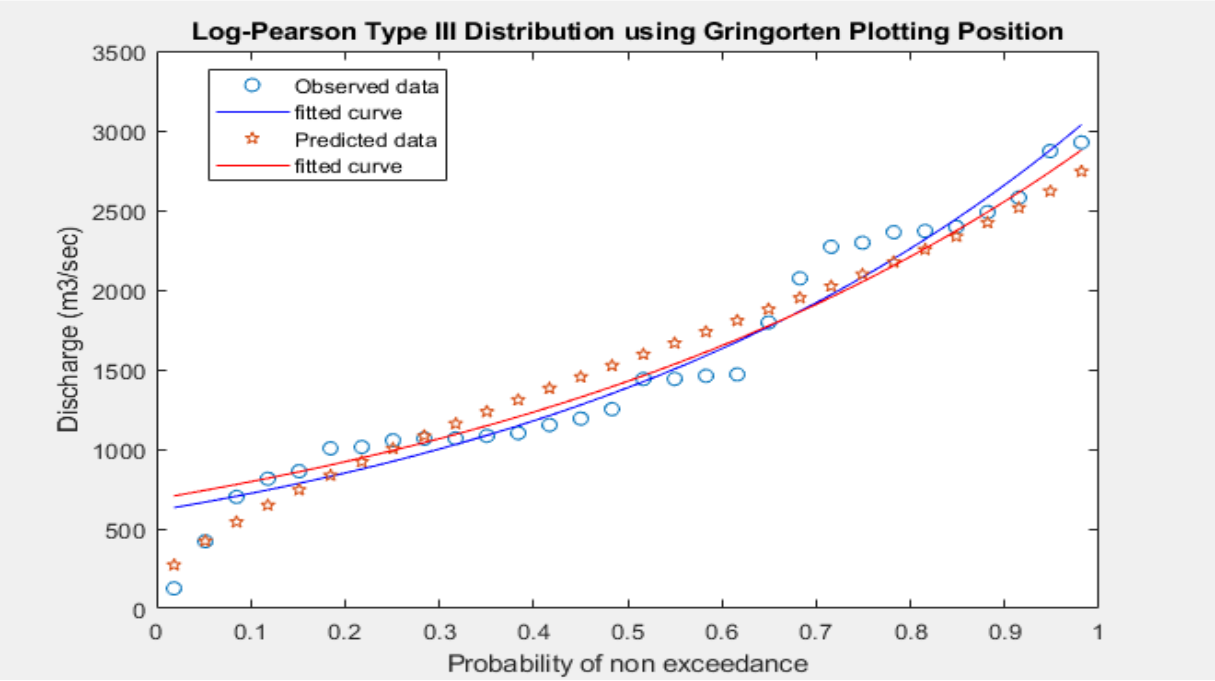


Fig. 4.32 Log-Pearson type III distribution curves using Gringorten plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix VIII) using Log-Pearson type III distribution model when matched with Gringorten plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.32. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XVI) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9547 and 141.99 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are also presented in plate 63a and 63b respectively (Appendix XVI). By comparison

between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 2743.90 m^3/s against a probability of exceedance of 0.0186, invariably within a recurrence interval of approximately 54 years in which the flood may occur.

4.5.1 Hazen plotting position.

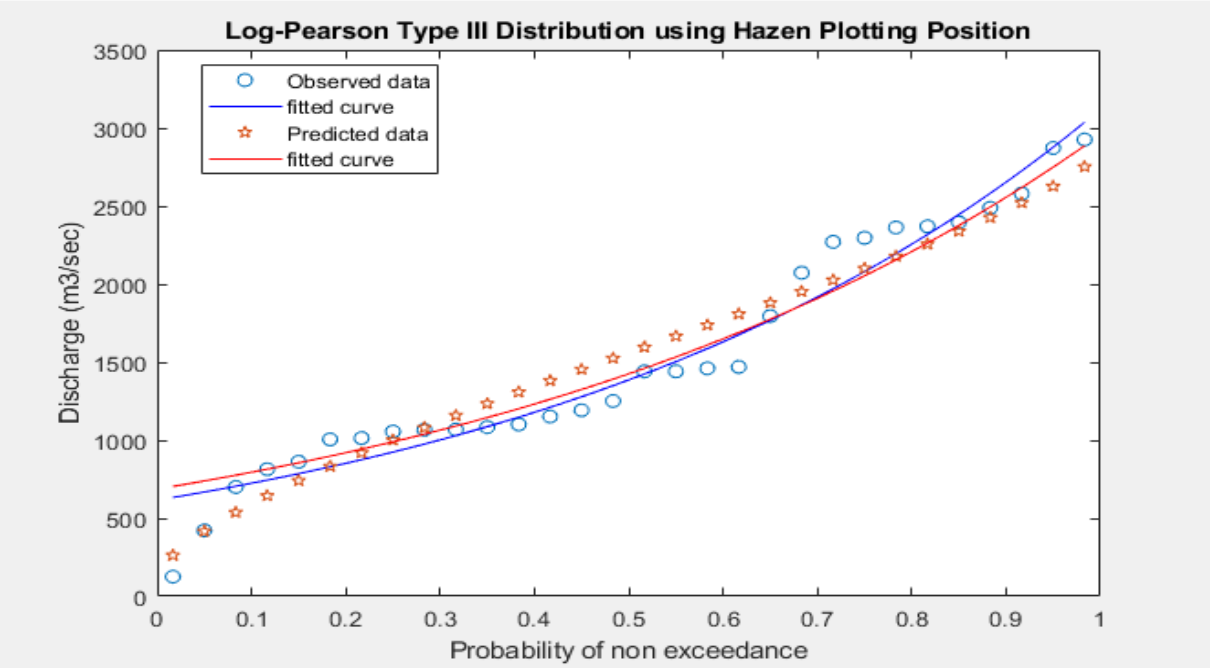


Fig. 4.33 Log-Pearson type III distribution curves using Hazen plotting position.

The observed discharge data and the result of predicted discharges as presented in (Appendix VIII) using Log-Pearson type III distribution model when matched with Hazen plotting position were used in plotting curves of Observed and predicted discharge data against probability of non-exceedance respectively, as shown in Fig 4.33. Statistical hypothesis test was applied using the concept of Goodness of fit in order to determine how the sample data fitted the distribution. Coefficient of determination (R_square) and Root Mean Square Error (RMSE) were determined (Appendix XVI) while the algorithms followed were elaborated in chapter 3.4. The Observed discharge curve gave R-square and RMSE value as 0.9523 and 157.71 respectively, so also the predicted discharge curve gave R-square and RMSE value as 0.9541 and 143.69 respectively. Hence, the result of the R-square and RMSE for both the observed and predicted discharge are

also presented in plate 64a and 64b respectively (Appendix XVI). By comparison between Observed and predicted discharge curve, the result shows that the predicted discharge data tend to fit in more using exponential curve with much reliability and accuracy due to higher degree of R-square and least RMSE value. Furthermore, the predicted discharge curve gave an estimated maximum annual discharge of 2752.00 m^3/s against a probability of exceedance of 0.0167, invariably within a recurrence interval of approximately 60 years in which the flood may occur.

4.6 Discussion Summary.

4.6.1 Normal Distribution.

Figs. 4.2-4.9, show the graphical representation/illustration of normal probability distribution, in which the observed and the predicted data were plotted against different number of plotting position values of eight formulae (i.eBlom, California, Chegodayev, Cunnane, Gringorten, Hazen, Turkey and Weibull). Appendices I and II show the result of the predicted data using the various plotting positions formulae with respect to observed data. Furthermore, by determining the goodness of fit and goodness of validation of these formulae, it was discovered that the California plotting position fit more for this distribution because it recorded the highest coefficient of determination (Adjusted R-sqaure) as 0.9604 as well as low RMSE of 128.79for the predicted data and an Adjusted R-sqaure of 0.9523 as well as RMSE 157.71 being the observed data which is the flat figure for all of the distribution (Appendices XVII and XXI).

4.6.2 Log-Normal Distribution.

Figs. 4.10-4.17, show the graphical representation/illustration of log-normal probability distribution, in which the observed and the predicted data were plotted against different number of plotting position values of eight formulae (i.e Blom, California, Chegodayev, Cunnane, Gringorten, Hazen, Turkey and Weibull). Appendices III and IV show the result of the predicted data using the various plotting positions formulae with respect to observed data. Furthermore, by determining the goodness of fit and goodness of validation of these formulae, it was discovered that the California plotting position fit more for this distribution because it recorded the highest coefficient of determination (Adjusted r_sqaure) as 0.9622 as well as low RMSE of 175.39for the predicted data and an Adjusted r_sqaure of 0.9523 as well as RMSE 157.71being the observed data which is the flat figure for all of the distribution (Appendices XVIII and XXI).

4.6.3 Pearson Type III Distribution.

Figs. 4.18-4.25, show the graphical representation/illustration of Pearson type III probability distribution, in which the observed and the predicted data were plotted against different number of plotting position values of eight formulae (i.e Blom, California, Chegodayev, Cunnane, Gringorten, Hazen, Turkey and Weibull). Appendices V and VI show the result of the predicted data using the various plotting positions formulae with respect to observed data. Furthermore, by determining the goodness of fit and goodness of validation of these formulae, it was discovered that the California plotting position fit more for this distribution because it recorded the highest coefficient of determination (Adjusted R_sqaure) as 0.9728 as well as low RMSE of 107.04for the predicted data and an Adjusted R_sqaure of 0.9523 as well as RMSE 157.71being the observed data which is the flat figure for all of the distribution (Appendices XIX and XXI).

4.6.4 Log-Pearson Type III Distribution.

Figs. 4.26-4.33, show the graphical representation/illustration of Log-Pearson type III probability distribution, in which the observed and the predicted data were plotted against different number of plotting position values of eight formulae (i.e Blom, California, Chegodayev, Cunnane, Gringorten, Hazen, Turkey and Weibull). Appendices VII and VIII show the result of the predicted data using the various plotting positions formulae with respect to observed data. Furthermore, by determining the goodness of fit and goodness of validation of these formulae, it was discovered that the California plotting position fit more for this distribution because it recorded the highest coefficient of determination (Adjusted R_sqaure) as 0.9584 as well as low RMSE of 130.57 for the predicted data and an Adjusted R_sqaure of 0.9523 as well as RMSE 157.71 being the observed data which is the flat figure for all of the distribution (Appendices XX and XXI).

4.7 Flood Magnitude.

The response of the distributions at various recurrence interval of 25, 50 and 100 years with respect to plotting positions were determined using their individual exponential models (Appendices XIII-XVI). The magnitude of the flood at recurrence interval of 25, 50 and 100 years were determined to be 2838.60, 2936 and 2970.70 m^3/sec using the Pearson Type IIIdistribution (Tables 4.1 and 4.2). This shows that the performance of the distributions along

with various plotting positions estimate maximum discharges, correlate and in comparison, with both the observed and predicted data. So also, by comparison between the predicted discharge and the observed discharge data, it was discovered that Modified California plotting position became the best, having the least absolute difference of 6.60, 13.00 and 16.40 m^3/sec when matched with Pearson Type III.

Table. 4.1 Response of the distribution at various return period with respect to the plotting positions.

Plotting Positions	Normal Distribution (Predicted Flow m^3/s)			Log-normal Distribution (Predicted Flow m^3/s)			Observed Flow m^3/s		
	Return Period			Return Period			Return Period		
	25	50	100	25	50	100	25	50	100
California	2822.30	2907.20	2950.60	3582.00	3745.60	3830.10	2845.20	2939.00	2987.10
Weibull	2824.40	2909.50	2952.90	3588.20	3752.40	3837.30	2999.40	3101.80	3154.20
Chegodayev	2847.00	2934.40	2979.10	3724.90	3904.00	3996.70	2953.40	3052.10	3102.70
Turkey	2850.60	2938.40	2983.30	3741.20	3922.30	4016.00	2947.40	3045.70	3096.00
Blom	2854.40	2942.60	2987.70	3767.10	3951.10	4046.40	2942.40	3040.20	3090.40
Cunnane	2856.60	2945.00	2990.30	3782.50	3968.20	4064.50	2938.20	3035.80	3085.70
Gringorten	2860.80	2949.70	2995.10	3810.80	3999.70	4097.70	2933.20	3030.30	3080.10
Hazen	2869.50	2959.10	3005.00	4335.10	4564.70	4684.00	2923.00	3019.40	3068.80
	Absolute Difference								
California	22.90	31.80	36.50	736.80	806.60	843.00			
Weibull	175.00	192.30	201.30	588.80	650.60	683.10			
Chegodayev	106.40	117.70	123.60	771.50	851.90	894.00			
Turkey	96.80	107.30	112.70	793.80	876.60	920.00			
Blom	88.00	97.60	102.70	824.70	910.90	956.00			
Cunnane	81.60	90.80	95.40	844.30	932.40	978.80			
Gringorten	72.40	80.60	85.00	877.60	969.40	1017.60			
Hazen	53.50	60.30	63.80	1412.10	1545.30	1615.20			

Table. 4.2 Response of the distribution at various return period with respect to the plotting positions

Plotting Positions	Pearson type III Distribution (Predicted Flow m³/s)			Log-Pearson type III Distribution (Predicted Flow m³/s)			Observed Flow m³/s		
	Return Period			Return Period			Return Period		
	25	50	100	25	50	100	25	50	100
California	2838.60	2926.00	2970.70	2793.08	2875.09	2917.00	2845.20	2939.00	2987.10
Weibull	2840.70	2928.20	2973.00	2791.88	2873.86	2915.74	2999.40	3101.80	3154.20
Chegodayev	2869.50	2959.80	3006.10	2790.11	2872.43	2914.50	2953.40	3052.10	3102.70
Turkey	2872.50	2963.20	3009.60	2790.36	2872.75	2914.86	2947.40	3045.70	3096.00
Blom	2870.50	2970.50	3017.20	2790.21	2872.65	2914.79	2942.40	3040.20	3090.40
Cunnane	2881.30	2972.80	3019.60	2790.86	2873.39	2915.56	2938.20	3035.80	3085.70
Gringorten	2887.10	2979.20	3026.30	2790.30	2872.86	2915.06	2933.20	3030.30	3080.10
Hazen	2897.70	2990.70	3038.30	2790.79	2873.48	2915.74	2923.00	3019.40	3068.80
				Absolute Difference					
California	6.60	13.00	16.40	52.12	63.91	70.10			
Weibull	158.70	173.60	181.20	207.52	227.94	238.46			
Chegodayev	83.90	92.30	96.60	163.29	179.67	188.20			
Turkey	74.90	82.50	86.40	157.04	172.95	181.14			
Blom	71.90	69.70	73.20	152.19	167.55	175.61			
Cunnane	56.90	63.00	66.10	147.34	162.41	170.14			
Gringorten	46.10	51.10	53.80	142.90	157.44	165.04			
Hazen	25.30	28.70	30.50	132.21	145.92	153.06			

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions.

i. The determination of suitable plotting position for River Kaduna catchment was carried out by employing eight (8) plotting positions (i.e California, Weibull, Chegodayev, Turkey, Blom, Cunnane, Gringorten, Hazenwhenmatched with four (4) selected distribution models. The study discovered that California (Modified) emerges the best plotting position suitable for the catchment. Comparison was made between the R-sqaure and RMSE values, in which California (modified) recorded the highest Adjusted R-square as 0.9604 0.9622, 0.9728 and 0.9584 as well as validated least RMSE value of 128.79, 175.39, 107.04 and 130.57for all the distribution namely Normal, Lognormal, Pearson type III and log-Pearson type III respectively(Appendix XXII and XIII).

ii. The determination of probability distribution that best suit the catchment of River Kaduna was carried out by employing four (4) distribution models when matched with four (4) selected distribution namely Normal, Lognormal, Pearson type III and log-Pearson type III respectively. The study discovered that Pearson Type III emerges the best probability distribution suitable for the catchment. Comparison was made between the R-square and RMSE values, in which Pearson Type III recorded the highest Adjusted R-square as 0.9728 as well as validated least RMSE value of 107.04(Appendix XXII and XIII).

iii. Probability of non-exceedance for recurrence interval of 25, 50 and 100 years were determined which were fitted into the exponential models of the four (4) selected distributions (Appendices XIII-XVI). This is because, the exponential model was initially derived based on probability of non-exceedance as shown in (Figs. 4.2-4.33). Hence, the magnitude of the flood at different recurrence interval of 25, 50 and 100 years were determined to be 2838.60, 2936 and 2970.70 m^3/sec using the exponential model of Pearson Type III distribution(Tables 4.1 and 4.2).

5.2 Recommendations.

i. It is recommended that Modified California plotting position should be used for any hydrological engineering within the catchment.

ii. Pearson type III distribution should be adopted also for any flood discharge estimation analysis within the catchment.

iii. Further research should focus on evaluating the performance of other distribution models and determining their degree of accurate discharge predictions.

iv. Derivation of more plotting positionsformulae should be carried outin determination of accurate predictions of flood. The formulae should be able to have a coefficient of determination value of relatively one (1), if possible.

v. Governmentshould support the agencies responsible for providing adequate drainages, canals etc so as to prevent or minimize flood events.

REFERENCES.

Abdi, H (2016). *The Method of Least Squares*. Retrieved from https://personal.utdallas.edu/ ~herve/Abdi-LeastSquares06-pretty.pdf on February 24, 2021.

Abdulkarim M.D and Chowdhury J.U (1995). A comparison of four distributions used in flood frequency analysis in Bangladesh.*HydrologicalSciences-Journal-desSciences Hydrologiques,* Vol. 40, l, February 1995.

Adamowski, K. (1981). Plotting Formulae for Flood Frequency.*Water Resources Bulletin*, 17(2), pp. 197.

Adeboye, O. B and Alatise, M. O (2007). Performance of Probability Distributions and PlottingPositions in Estimating the Flood of River Osun at Apoje Sub-basin, Nigeria. *AgriculturalEngineering International*: the CIGR Ejournal. Manuscript LW 07 007. Vol. IX. July, 2007

Adeyemi, S and Adebanji, A.O. (2006). The exponentiated generalized extreme value distribution. *Journal of Applicable Functional Differential Equations (JAFDE)*, Volume 1, No. 1, pp. 41-47.

Alam, M. J. B. and Matin, A. (2004).Study of plotting position formulae for Surma basin in Bangladesh.*Journal of Civil Engineering (IEB)*, 33 (1) (2005) 9-17.

Alayande, A.W (2010).*Impact of urbanization on Kaduna River flooding and streamflow modeling*. Department of civil engineering, Faculty of Engineering, University of Nigeria, Nsukka, Nigeria; pp. 3-5.

Bello, A. D., Alayande2, W A., Otun., J. A., Ismail, A and Lawan, U. F (2015). Optimization of the Designed Water Distribution System Using MATLAB. *International Journal of Hydraulic Engineering 2015*, 4(2): 37-44 DOI: 10.5923/j.ijhe.20150402.03

Blom, G. (1958). *Statistical Estimates and Transformed Beta Variables*. Wiley, New York, pp.197.

Bobee, B., (1973). Sample error of T-year events computed by fitting a Pearson type III distribution.*Water Resources Research,* Vol. 9, No. 5, pp. 1264-1270.

Bobee, B. and Robitaille, R., (1977). Correction of bias in estimation of the coefficient of skewness. *Water Resources Research*, Vol. 11, No. 6, pp. 851-854.

Brzezinski, M. (2012). The Chen-Shapiro test for normality. *The stata journal*, No. 3, pp. 308-374

Chang, S.K. and Moore, S.M., (1983). Flood frequency analysis for small watersheds in southern Illinois.*Water Resources Research*, Vol. 19, No. 2, pp. 277-282.

Chegodajev, N. N. 1955. Formulas for Calculation of the Confidence of Hydrologic Quantities. *Handbook of Applied Hydrology*, McGraw Hill, N.Y.

Chow, V.T. (1964). *Hand Book of Applied Hydrology*. McGraw-Hill, New York.

Cunnane, C. (1978). Unbiased plotting positions-A review. *Journal of Hydrology*, 37(3/4): 205-222.

Doswell, C. A. (2003). Chapter 8: Flooding. University of Oklahoma, Norman, OK, USA Copyright 2003 Elsevier Science Ltd. All Rights Reserved. Pp. 769.

Ewemoje T.A and Ewemooje O.S (2011). *Best Distribution and Plotting Positions of Daily Maximum Flood Estimation at OnaRiver in Ogun-Oshun River Basin, Nigeria*. Agricultural Engineering International: CIGR Journal. Vol.13, No. 3, 2011. Manuscript No.1380.

Filliben, J.J. (1975).*The Probability Plot Correlation Coefficient Test for Normality*. Technometrics, 17(1): pp.111-117.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C., And Wallis, J. R (1979). *Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressible in Inverse Form*. Water Resources Research. Vol. 15, No. 5.

Gringorton, I.I. (1963). *A plotting Rule for Extreme Probability Paper*, Journal of Geophysical Res., 68(3), pp.813.

Hazelton, M. L (2011). *Methods of Moments Estimation*. International Encyclopedia of Statistical Science (pp.816-817). DOI: 10.1007/978-3-642-04898-2_364.

Hazen, A. (1930). *Flood Flows-A Studies of Frequencies and Magnitudes*, John Wiley and Sons. Inc., New York.

Hooda, E., Hooda, B. K and Tanwar, N (2018). *Probability weighted moments (PWMs) and partial probability weighted moments (PPWMs) of type-II extreme value distribution*. Department of Mathematics, Statistics and Physics, CCS HAU, Hisar, Haryana- 125004, India. Retrieved from https://www.researchgate.net/publication/328252711.

Hosking, J. R. M. & Wallis, J. R. (1995); *A comparison of unbiased and plotting-position estimators of L-Moments*. Wat. Resoures. Pp.78-85

Hunt, B. R., Lipsman, R. L., Rosenberg, J. M., Coombes, K. R., Osborn, J. E., and Stuck, G. J (2001). *A guide to MATLAB Beginners and Experienced Users*. Cambridge University Press.The Edinburgh Building, Cambridge $_{CB2\ 2RU}$ United Kingdom. Pp. 36-40.

Ivan, M. (2014).*The Risks of Hazard*. retrieved from http://www.intermap.com/risks-of-hazard-blog/three-common-types-of-flood-explained on March 15, 2018.

Izinyon, O.C and Ehiorobo. J.O. (2014). L-Moments Approach For Flood Frequency Analysis Of River Owena River Basin Okhuwan In Benin-Owena River Basin In Nigeria.*Nigerian Journal of Technology (NIJOTECH),* Vol. 33, No. 1, pp. 15.

Kim. S, Shin. H, Kim.T and Heo. J (2010). Derivation of the probability plot correlation coefficient test statistics for the generalized logistic distribution. *International Workshop Advances in Statistical Hydrology*; Taormina, Italy; pp.4

Kottegoda, N. T. 1980. Stochastic Water Resources Technology, 1st edition, Macmillan Press Ltd, London, 208-215

Lariyah, M.S, Faizah C.R, Mohd Noh, A., Rahsidi S.M., Azwin Zailti, A.R., Mohd Nor, M.D., Hidayah & Intan Shafilah A.A (2012).*Development of probable Maximum Flood (PMF) for sultan Abu Bakar Dam*. EnCon 2012, 5th Engineering Conference, "Engineering Towards Change - Empowering Green Solutions" 10-12th July 2012, Kuching Sarawak.

Langat, P.K., Kumar, L and Koech, R. (2019). *Identification of the Most Suitable Probability Distribution Models for Maximum, Minimum, and Mean Streamflow*. Water 2019, 11, 734; doi:10.3390/w11040734.

Lesse, M. (2005). Plotting Positions in Extreme Value Analysis. *Journal of applied meteorology and climatology*, Vol.35; Published by VTT Technical Research Centre of Finland, Espoo, Finland p.334

Markovic, R.D., (1963). Probability functions of best fir to distributions of annual precipitation and runoff.*Hydrology paper* No. 8, pp. 33, Colorado State University, Fort Collins, Colorado.

Matalas, N. C., (1963). *Probability distribution of low flows*. U.S. Geological Survey, Professional Paper 434-A, pp. 1-27.

Mauro, N (2017).*Fundamental of statistical hydrology*. Published by springer International Publishing Switzerland, p.215-216.

Miller, S.J (2012). *The Method of Least Squares*. Mathematics Department Brown University Providence, RI 02912

Mustafa, S and Yusuf, I.M (2012). *A textbook of hydrology and water resources*. Topsmerit page publishing company. Abuja, Nigeria. Revised Edition.

Naghettini, M (2017). *Fundamentals of Statistical Hydrology*. The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland, Springer International Publishing Switzerland 2017, ISBN 978-3-319-43560-2, pp.227

Nwigwe, C and Emberga, T.T (2014). An Assessment of causes and effects of flood in Nigeria. *Standard Scientific Research and Essays* Vol2 (7): 307-315, July 2014 (ISSN: 2310-7502). pp. 309.

Ojha, C.S.P., Berndtsson, R and Bhunya, P. (2008).Flood Frequency Analysis.*Engineering Hydrology*, Oxford University Press ISBN: 978- 0- 19- 569461- 1, pp 250, 257.

Olivares, A. M. and Forero, C. G (2010). *Goodness-of-Fit Testing*. International Encyclopedia of Education, vol. 7, pp. 190-196. Elsevier Ltd. All rights reserved.

Peel, M. C., Wang, Q. J., Vogel, R. M. and Mcmahon, T. A. (2001). The utility of L-moment ratio diagrams for selecting a regional probability distribution.*Hydrological Sciences Journal des Sciences Hydrologiques*, 46(1) February. Pp. 147-155.

Ripunjai, K. S., Trivedi, M and Kumar, M (2010). On the proficient use of GEV distribution: a case study of subtropical monsoon region in India. Anale, Seria Informatica. Vol. VIII fasc. I - 2010 Annals. Computer Science Series 8th Tome 1st Fasc.

Stedinger, J.R, Vogel, R.M and Georgiou, E.F (1993). *Frequency analysis of extreme events, Hand book hydrology*.pp.18-26.

Weibull, W. (1939). *A Statistical Theory of Strength of Materials*. Stockholm, pp.151.

Will Kenton (2018), "Financial Analysis", Retrieved From https://www.investopedia. com/terms/c/chi-square-statistic.asp on January 22, 2019

Vogel, R. M. (1986). *The Probability Plot Correlation Coefficient Test for the Normal, Lognormal, and Gumbel Distributional Hypotheses*. Water Resources Research, Yolo 22, No.4, Pages 587-590, April.

Zhang, D (2016). *A Coefficient of Determination for Generalized Linear Models*. Department of Statistics, Purdue University.

Appendix I Normal Distribution using Probability of Exceedance

YEAR	Max. Discharge Q (m^3/sec)	S/N	Blom	Predicted Q (m^3/sec)	California	Predicted Q (m^3/sec)	Chegodayevs	Predicted Q (m^3/sec)	Cunnane	Predicted Q (m^3/sec)
1993	2926.31	1	0.0207	3063.70	0.0000	NaN	0.0230	3029.83	0.0199	3075.85
1986	2871.75	2	0.0537	2742.00	0.0333	2909.49	0.0559	2727.11	0.0530	2747.04
1992	2579.50	3	0.0868	2555.81	0.0667	2660.65	0.0888	2546.24	0.0861	2559.06
1998	2488.34	4	0.1198	2417.41	0.1000	2496.49	0.1217	2410.43	0.1192	2419.77
1997	2395.85	5	0.1529	2303.94	0.1333	2368.77	0.1546	2298.54	0.1523	2305.76
1987	2370.38	6	0.1860	2205.82	0.1667	2261.54	0.1875	2201.50	0.1854	2207.27
1976	2362.50	7	0.2190	2118.02	0.2000	2167.44	0.2204	2114.51	0.2185	2119.20
1973	2296.52	8	0.2521	2037.58	0.2333	2082.39	0.2533	2034.72	0.2517	2038.54
1972	2270.70	9	0.2851	1962.56	0.2667	2003.87	0.2862	1960.22	0.2848	1963.34
1988	2072.57	10	0.3182	1891.60	0.3000	1930.19	0.3191	1889.72	0.3179	1892.23
1975	1795.29	11	0.3512	1823.71	0.3333	1860.15	0.3520	1822.23	0.3510	1824.21
1971	1469.64	12	0.3843	1758.13	0.3667	1792.86	0.3849	1757.01	0.3841	1758.50
1991	1460.45	13	0.4174	1694.23	0.4000	1727.60	0.4178	1693.45	0.4172	1694.49
1974	1441.33	14	0.4504	1631.49	0.4333	1663.79	0.4507	1631.03	0.4503	1631.65
1990	1440.94	15	0.4835	1569.45	0.4667	1600.93	0.4836	1569.30	0.4834	1569.50
1982	1251.04	16	0.5165	1507.69	0.5000	1538.56	0.5164	1507.84	0.5166	1507.64
1981	1192.14	17	0.5496	1445.79	0.5333	1476.27	0.5493	1446.25	0.5497	1445.64
1995	1151.79	18	0.5826	1383.34	0.5667	1413.62	0.5822	1384.11	0.5828	1383.08
1985	1100.82	19	0.6157	1319.89	0.6000	1350.18	0.6151	1321.00	0.6159	1319.52
1999	1083.83	20	0.6488	1254.96	0.6333	1285.48	0.6480	1256.42	0.6490	1254.47
1979	1068.11	21	0.6818	1187.98	0.6667	1218.97	0.6809	1189.83	0.6821	1187.36
1980	1066.41	22	0.7149	1118.26	0.7000	1150.02	0.7138	1120.55	0.7152	1117.49
2000	1055.51	23	0.7479	1044.93	0.7333	1077.83	0.7467	1047.72	0.7483	1044.00
1996	1013.98	24	0.7810	966.84	0.7667	1001.36	0.7796	970.24	0.7815	965.71
1978	1005.40	25	0.8140	882.39	0.8000	919.19	0.8125	886.52	0.8146	881.00
1994	862.01	26	0.8471	789.16	0.8333	829.27	0.8454	794.26	0.8477	787.45
1977	814.50	27	0.8802	683.31	0.8667	728.40	0.8783	689.76	0.8808	681.13
1983	700.56	28	0.9132	557.83	0.9000	611.06	0.9112	566.35	0.9139	554.94
1984	422.49	29	0.9463	397.37	0.9333	466.11	0.9441	409.75	0.9470	393.14
1989	126.20	30	0.9793	151.36	0.9667	263.56	0.9770	175.03	0.9801	143.05

Appendix II Normal Distribution using Probability of Exceedance

YEAR	Max. Discharge Q (m^3/sec)	S/N	Gringorten	Predicted Q (m^3/sec)	Hazen	Predicted Q (m^3/sec)	Turkey	Predicted Q (m^3/sec)	Weibull	Predicted Q (m^3/sec)
1993	2926.31	1	0.0186	3096.22	0.0167	3129.33	0.0220	3044.48	0.0323	2920.46
1986	2871.75	2	0.0518	2755.33	0.0500	2768.15	0.0549	2733.63	0.0645	2673.25
1992	2579.50	3	0.0850	2564.32	0.0833	2572.35	0.0879	2550.46	0.0968	2510.40
1998	2488.34	4	0.1182	2423.58	0.1167	2429.38	0.1209	2413.51	0.1290	2383.88
1997	2395.85	5	0.1514	2308.70	0.1500	2313.16	0.1538	2300.93	0.1613	2277.81
1987	2370.38	6	0.1846	2209.61	0.1833	2213.16	0.1868	2203.41	0.1935	2184.85
1976	2362.50	7	0.2178	2121.10	0.2167	2123.97	0.2198	2116.07	0.2258	2100.95
1973	2296.52	8	0.2510	2040.09	0.2500	2042.43	0.2527	2035.99	0.2581	2023.61
1972	2270.70	9	0.2842	1964.60	0.2833	1966.51	0.2857	1961.26	0.2903	1951.16
1988	2072.57	10	0.3174	1893.24	0.3167	1894.77	0.3187	1890.55	0.3226	1882.40
1975	1795.29	11	0.3506	1825.00	0.3500	1826.21	0.3516	1822.89	0.3548	1816.47
1971	1469.64	12	0.3838	1759.10	0.3833	1760.01	0.3846	1757.51	0.3871	1752.66
1991	1460.45	13	0.4170	1694.91	0.4167	1695.54	0.4176	1693.79	0.4194	1690.40
1974	1441.33	14	0.4502	1631.89	0.4500	1632.27	0.4505	1631.23	0.4516	1629.23
1990	1440.94	15	0.4834	1569.59	0.4833	1569.71	0.4835	1569.37	0.4839	1568.70
1982	1251.04	16	0.5166	1507.56	0.5167	1507.43	0.5165	1507.77	0.5161	1508.44
1981	1192.14	17	0.5498	1445.39	0.5500	1445.01	0.5495	1446.05	0.5484	1448.04
1995	1151.79	18	0.5830	1382.66	0.5833	1382.03	0.5824	1383.77	0.5806	1387.14
1985	1100.82	19	0.6162	1318.92	0.6167	1318.02	0.6154	1320.51	0.6129	1325.31
1999	1083.83	20	0.6494	1253.68	0.6500	1252.49	0.6484	1255.77	0.6452	1262.12
1979	1068.11	21	0.6826	1186.36	0.6833	1184.85	0.6813	1189.01	0.6774	1197.03
1980	1066.41	22	0.7158	1116.26	0.7167	1114.39	0.7143	1119.53	0.7097	1129.44
2000	1055.51	23	0.7490	1042.49	0.7500	1040.21	0.7473	1046.49	0.7419	1058.55
1996	1013.98	24	0.7822	963.87	0.7833	961.10	0.7802	968.73	0.7742	983.36
1978	1005.40	25	0.8154	878.76	0.8167	875.36	0.8132	884.69	0.8065	902.47
1994	862.01	26	0.8486	784.68	0.8500	780.48	0.8462	792.01	0.8387	813.85
1977	814.50	27	0.8818	677.62	0.8833	672.29	0.8791	686.91	0.8710	714.35
1983	700.56	28	0.9150	550.27	0.9167	543.15	0.9121	562.59	0.9032	598.50
1984	422.49	29	0.9482	386.28	0.9500	375.72	0.9451	404.30	0.9355	455.31
1989	126.20	30	0.9814	129.24	0.9833	107.27	0.9780	164.74	0.9677	255.21

Appendix III Log-Normal Distribution using Probability of Exceedance.

YEAR	Max. Discharge Q (m^3/sec)	S/N	Blom	Predicted Q (m^3/sec)	California	Predicted Q (m^3/sec)	Chegodayevs	Predicted Q (m^3/sec)	Cunnane	Predicted Q (m^3/sec)
1993	2926.31	1	0.0207	4924.73	0.0000	NaN	0.0230	4782.64	0.0199	4976.48
1986	2871.75	2	0.0537	3730.15	0.0333	4310.67	0.0559	3682.66	0.0530	3746.57
1992	2579.50	3	0.0868	3176.37	0.0667	3477.31	0.0888	3150.25	0.0861	3185.28
1998	2488.34	4	0.1198	2818.63	0.1000	3017.79	0.1217	2801.69	0.1192	2824.38
1997	2395.85	5	0.1529	2555.59	0.1333	2702.70	0.1546	2543.70	0.1523	2559.62
1987	2370.38	6	0.1860	2348.01	0.1667	2463.74	0.1875	2339.27	0.1854	2350.96
1976	2362.50	7	0.2190	2176.61	0.2000	2271.49	0.2204	2170.03	0.2185	2178.82
1973	2296.52	8	0.2521	2030.57	0.2333	2110.67	0.2533	2025.56	0.2517	2032.26
1972	2270.70	9	0.2851	1903.22	0.2667	1972.32	0.2862	1899.39	0.2848	1904.50
1988	2072.57	10	0.3182	1790.12	0.3000	1850.76	0.3191	1787.22	0.3179	1791.09
1975	1795.29	11	0.3512	1688.21	0.3333	1742.17	0.3520	1686.06	0.3510	1688.94
1971	1469.64	12	0.3843	1595.28	0.3667	1643.84	0.3849	1593.74	0.3841	1595.79
1991	1460.45	13	0.4174	1509.65	0.4000	1553.78	0.4178	1508.64	0.4172	1509.99
1974	1441.33	14	0.4504	1430.06	0.4333	1470.50	0.4507	1429.49	0.4503	1430.25
1990	1440.94	15	0.4835	1355.48	0.4667	1392.82	0.4836	1355.30	0.4834	1355.54
1982	1251.04	16	0.5165	1285.09	0.5000	1319.81	0.5164	1285.26	0.5166	1285.03
1981	1192.14	17	0.5496	1218.22	0.5333	1250.70	0.5493	1218.70	0.5497	1218.05
1995	1151.79	18	0.5826	1154.27	0.5667	1184.84	0.5822	1155.04	0.5828	1154.01
1985	1100.82	19	0.6157	1092.74	0.6000	1121.70	0.6151	1093.79	0.6159	1092.39
1999	1083.83	20	0.6488	1033.17	0.6333	1060.76	0.6480	1034.48	0.6490	1032.73
1979	1068.11	21	0.6818	975.12	0.6667	1001.56	0.6809	976.68	0.6821	974.59
1980	1066.41	22	0.7149	918.15	0.7000	943.68	0.7138	919.97	0.7152	917.54
2000	1055.51	23	0.7479	861.83	0.7333	886.66	0.7467	863.91	0.7483	861.13
1996	1013.98	24	0.7810	805.64	0.7667	830.01	0.7796	808.00	0.7815	804.85
1978	1005.40	25	0.8140	748.99	0.8000	773.17	0.8125	751.66	0.8146	748.09
1994	862.01	26	0.8471	691.07	0.8333	715.42	0.8454	694.11	0.8477	690.04
1977	814.50	27	0.8802	630.71	0.8667	655.75	0.8783	634.23	0.8808	629.53
1983	700.56	28	0.9132	565.95	0.9000	592.57	0.9112	570.13	0.9139	564.54
1984	422.49	29	0.9463	492.74	0.9333	522.86	0.9441	498.03	0.9470	490.94
1989	126.20	30	0.9793	398.45	0.9667	438.98	0.9770	406.67	0.9801	395.60

Appendix IV Log-Normal Distribution using Probability of Exceedance.

YEAR	Max. Discharge Q (m^3/sec)	S/N	Gringorten	Predicted Q (m^3/sec)	Hazen	Predicted Q (m^3/sec)	Turkey	Predicted Q (m^3/sec)	Weibull	Predicted Q (m^3/sec)
1993	2926.31	1	0.0186	5064.78	0.0167	5211.64	0.0220	4843.50	0.0323	4351.70
1986	2871.75	2	0.0518	3773.49	0.0500	3815.47	0.0549	3703.45	0.0645	3515.33
1992	2579.50	3	0.0850	3199.78	0.0833	3222.06	0.0879	3161.75	0.0968	3054.26
1998	2488.34	4	0.1182	2833.68	0.1167	2847.90	0.1209	2809.16	0.1290	2738.20
1997	2395.85	5	0.1514	2566.13	0.1500	2576.02	0.1538	2548.96	0.1613	2498.58
1987	2370.38	6	0.1846	2355.71	0.1833	2362.94	0.1868	2343.13	0.1935	2305.89
1976	2362.50	7	0.2178	2182.40	0.2167	2187.81	0.2198	2172.94	0.2258	2144.77
1973	2296.52	8	0.2510	2034.98	0.2500	2039.10	0.2527	2027.78	0.2581	2006.23
1972	2270.70	9	0.2842	1906.58	0.2833	1909.72	0.2857	1901.08	0.2903	1884.57
1988	2072.57	10	0.3174	1792.66	0.3167	1795.04	0.3187	1788.50	0.3226	1775.97
1975	1795.29	11	0.3506	1690.10	0.3500	1691.86	0.3516	1687.01	0.3548	1677.69
1971	1469.64	12	0.3838	1596.62	0.3833	1597.88	0.3846	1594.42	0.3871	1587.76
1991	1460.45	13	0.4170	1510.54	0.4167	1511.37	0.4176	1509.09	0.4194	1504.68
1974	1441.33	14	0.4502	1430.56	0.4500	1431.02	0.4505	1429.74	0.4516	1427.27
1990	1440.94	15	0.4834	1355.63	0.4833	1355.78	0.4835	1355.38	0.4839	1354.60
1982	1251.04	16	0.5166	1284.94	0.5167	1284.80	0.5165	1285.18	0.5161	1285.92
1981	1192.14	17	0.5498	1217.79	0.5500	1217.40	0.5495	1218.48	0.5484	1220.59
1995	1151.79	18	0.5830	1153.59	0.5833	1152.97	0.5824	1154.70	0.5806	1158.06
1985	1100.82	19	0.6162	1091.83	0.6167	1090.98	0.6154	1093.32	0.6129	1097.87
1999	1083.83	20	0.6494	1032.03	0.6500	1030.97	0.6484	1033.90	0.6452	1039.57
1979	1068.11	21	0.6826	973.75	0.6833	972.49	0.6813	975.98	0.6774	982.77
19800	1066.41	22	0.7158	916.57	0.7167	915.09	0.7143	919.16	0.7097	927.06
2000	1055.51	23	0.7490	860.01	0.7500	858.32	0.7473	862.99	0.7419	872.02
1996	1013.98	24	0.7822	803.57	0.7833	801.65	0.7802	806.96	0.7742	817.21
1978	1005.40	25	0.8154	746.64	0.8167	744.46	0.8132	750.48	0.8065	762.09
1994	862.01	26	0.8486	688.39	0.8500	685.90	0.8462	692.76	0.8387	705.95
1977	814.50	27	0.8818	627.62	0.8833	624.74	0.8791	632.67	0.8710	647.84
1983	700.56	28	0.9150	562.27	0.9167	558.82	0.9121	568.28	0.9032	586.17
1984	422.49	29	0.9482	488.04	0.9500	483.61	0.9451	495.69	0.9355	518.01
1989	126.20	30	0.9814	390.91	0.9833	383.57	0.9780	403.08	0.9677	435.82

Appendix V Pearson Type III Distribution using Probability of Exceedance.

YEAR	Max. Discharge Q (m^3/sec)	S/N	Blom	Predicted Q (m^3/sec)	California	Predicted Q (m^3/sec)	Chegodayevs	Predicted Q (m^3/sec)	Cunnane	Predicted Q (m^3/sec)
1993	2926.31	1	0.0207	3179.12	0.0000	NaN	0.0230	3138.22	0.0199	3193.76
1986	2871.75	2	0.0537	2797.79	0.0333	2994.44	0.0559	2780.56	0.0530	2803.70
1992	2579.50	3	0.0868	2584.02	0.0667	2703.82	0.0888	2573.17	0.0861	2587.71
1998	2488.34	4	0.1198	2428.28	0.1000	2516.94	0.1217	2420.50	0.1192	2430.92
1997	2395.85	5	0.1529	2302.62	0.1333	2374.19	0.1546	2296.68	0.1523	2304.62
1987	2370.38	6	0.1860	2195.39	0.1667	2256.12	0.1875	2190.71	0.1854	2196.97
1976	2362.50	7	0.2190	2100.60	0.2000	2153.83	0.2204	2096.83	0.2185	2101.86
1973	2296.52	8	0.2521	2014.68	0.2333	2062.43	0.2533	2011.63	0.2517	2015.70
1972	2270.70	9	0.2851	1935.34	0.2667	1978.93	0.2862	1932.89	0.2848	1936.17
1988	2072.57	10	0.3182	1861.02	0.3000	1901.35	0.3191	1859.06	0.3179	1861.67
1975	1795.29	11	0.3512	1790.55	0.3333	1828.30	0.3520	1789.02	0.3510	1791.06
1971	1469.64	12	0.3843	1723.07	0.3667	1758.73	0.3849	1721.93	0.3841	1723.45
1991	1460.45	13	0.4174	1657.88	0.4000	1691.86	0.4178	1657.09	0.4172	1658.15
1974	1441.33	14	0.4504	1594.42	0.4333	1627.03	0.4507	1593.95	0.4503	1594.57
1990	1440.94	15	0.4835	1532.17	0.4667	1563.69	0.4836	1532.02	0.4834	1532.22
1982	1251.04	16	0.5165	1470.72	0.5000	1501.37	0.5164	1470.87	0.5166	1470.67
1981	1192.14	17	0.5496	1409.63	0.5333	1439.65	0.5493	1410.09	0.5497	1409.48
1995	1151.79	18	0.5826	1348.52	0.5667	1378.09	0.5822	1349.27	0.5828	1348.27
1985	1100.82	19	0.6157	1286.96	0.6000	1316.28	0.6151	1288.03	0.6159	1286.60
1999	1083.83	20	0.6488	1224.50	0.6333	1253.79	0.6480	1225.90	0.6490	1224.03
1979	1068.11	21	0.6818	1160.65	0.6667	1190.12	0.6809	1162.41	0.6821	1160.10
1980	1066.41	22	0.7149	1094.81	0.7000	1124.73	0.7138	1096.96	0.7152	1094.09
2000	1055.51	23	0.7479	1026.24	0.7333	1056.92	0.7467	1028.84	0.7483	1025.37
1996	1013.98	24	0.7810	953.98	0.7667	985.82	0.7796	957.10	0.7815	952.93
1978	1005.40	25	0.8140	876.71	0.8000	910.27	0.8125	880.47	0.8146	875.44
1994	862.01	26	0.8471	792.47	0.8333	828.58	0.8454	797.04	0.8477	790.92
1977	814.50	27	0.8802	698.14	0.8667	738.15	0.8783	703.85	0.8808	696.22
1983	700.56	28	0.9132	588.15	0.9000	634.57	0.9112	595.56	0.9139	585.65
1984	422.49	29	0.9463	450.35	0.9333	508.99	0.9441	460.87	0.9470	446.77
1989	126.20	30	0.9793	245.22	0.9667	337.86	0.9770	264.64	0.9801	238.42

Appendix VI Pearson Type III Distribution using Probability of Exceedance.

YEAR	Max. Discharge Q (m^3/sec)	S/N	Gringorten	Predicted Q (m^3/sec)	Hazen	Predicted Q (m^3/sec)	Turkey	Predicted Q (m^3/sec)	Weibull	Predicted Q (m^3/sec)
1993	2926.31	1	0.0186	3218.40	0.0167	3258.70	0.0220	3155.90	0.0323	3007.50
1986	2871.75	2	0.0518	2813.30	0.0500	2828.30	0.0549	2788.10	0.0645	2718.30
1992	2579.50	3	0.0850	2593.70	0.0833	2602.80	0.0879	2578.00	0.0968	2532.60
1998	2488.34	4	0.1182	2435.20	0.1167	2441.60	0.1209	2423.90	0.1290	2391.00
1997	2395.85	5	0.1514	2307.90	0.1500	2312.80	0.1538	2299.30	0.1613	2273.90
1987	2370.38	6	0.1846	2199.50	0.1833	2203.40	0.1868	2192.80	0.1935	2172.70
1976	2362.50	7	0.2178	2103.90	0.2167	2107.00	0.2198	2098.50	0.2258	2082.30
1973	2296.52	8	0.2510	2017.30	0.2500	2019.80	0.2527	2013.00	0.2581	1999.80
1972	2270.70	9	0.2842	1937.50	0.2833	1939.50	0.2857	1934.00	0.2903	1923.40
1988	2072.57	10	0.3174	1862.70	0.3167	1864.30	0.3187	1859.90	0.3226	1851.40
1975	1795.29	11	0.3506	1791.90	0.3500	1793.10	0.3516	1789.70	0.3548	1783.10
1971	1469.64	12	0.3838	1724.10	0.3833	1725.00	0.3846	1722.40	0.3871	1717.50
1991	1460.45	13	0.4170	1658.60	0.4167	1659.20	0.4176	1657.40	0.4194	1654.00
1974	1441.33	14	0.4502	1594.80	0.4500	1595.20	0.4505	1594.20	0.4516	1592.10
1990	1440.94	15	0.4834	1532.30	0.4833	1532.40	0.4835	1532.10	0.4839	1531.40
1982	1251.04	16	0.5166	1470.60	0.5167	1470.50	0.5165	1470.80	0.5161	1471.50
1981	1192.14	17	0.5498	1409.20	0.5500	1408.90	0.5495	1409.90	0.5484	1411.80
1995	1151.79	18	0.5830	1347.90	0.5833	1347.20	0.5824	1348.90	0.5806	1352.20
1985	1100.82	19	0.6162	1286.00	0.6167	1285.10	0.6154	1287.60	0.6129	1292.20
1999	1083.83	20	0.6494	1223.30	0.6500	1222.10	0.6484	1225.30	0.6452	1231.40
1979	1068.11	21	0.6826	1159.10	0.6833	1157.70	0.6813	1161.60	0.6774	1169.20
1980	1066.41	22	0.7158	1092.90	0.7167	1091.20	0.7143	1096.00	0.7097	1105.30
2000	1055.51	23	0.7490	1024.00	0.7500	1021.80	0.7473	1027.70	0.7419	1038.90
1996	1013.98	24	0.7822	951.25	0.7833	948.69	0.7802	955.72	0.7742	969.20
1978	1005.40	25	0.8154	873.41	0.8167	870.32	0.8132	878.80	0.8065	895.00
1994	862.01	26	0.8486	788.44	0.8500	784.67	0.8462	795.02	0.8387	814.67
1977	814.50	27	0.8818	693.12	0.8833	688.40	0.8791	701.32	0.8710	725.65
1983	700.56	28	0.9150	581.59	0.9167	575.41	0.9121	592.29	0.9032	623.58
1984	422.49	29	0.9482	440.95	0.9500	432.00	0.9451	456.24	0.9355	499.75
1989	126.20	30	0.9814	227.14	0.9833	209.23	0.9780	256.19	0.9677	330.91

Appendix VII Log-Pearson Type III Distribution using Probability of Exceedance.

YEAR	Max. Discharge Q (m^3/sec)	S/N	Blom	Predicted Q (m^3/sec)	Califonia	Predicted Q (m^3/sec)	Chegodayevs	Predicted Q (m^3/sec)	Cunnane	Predicted Q (m^3/sec)
1993	2926.31	1	0.0207	2735.10	0.0000		0.0230	2725.60	0.0199	2738.40
1986	2871.75	2	0.0537	2613.50	0.0333	2685.40	0.0559	2606.30	0.0530	2615.90
1992	2579.50	3	0.0868	2510.50	0.0667	2571.50	0.0888	2504.70	0.0861	2512.60
1998	2488.34	4	0.1198	2418.00	0.1000	2472.60	0.1217	2412.90	0.1192	2419.60
1997	2395.85	5	0.1529	2331.70	0.1333	2382.10	0.1546	2327.40	0.1523	2333.20
1987	2370.38	6	0.1860	2249.90	0.1667	2297.10	0.1875	2246.30	0.1854	2251.40
1976	2362.50	7	0.2190	2171.70	0.2000	2216.40	0.2204	2168.40	0.2185	2172.80
1973	2296.52	8	0.2521	2095.70	0.2333	2138.60	0.2533	2092.90	0.2517	2096.60
1972	2270.70	9	0.2851	2021.70	0.2667	2062.80	0.2862	2019.30	0.2848	2022.40
1988	2072.57	10	0.3182	1949.00	0.3000	1988.80	0.3191	1947.00	0.3179	1949.70
1975	1795.29	11	0.3512	1877.50	0.3333	1916.20	0.3520	1875.80	0.3510	1878.00
1971	1469.64	12	0.3843	1806.60	0.3667	1844.20	0.3849	1805.30	0.3841	1807.00
1991	1460.45	13	0.4174	1736.10	0.4000	1773.10	0.4178	1735.20	0.4172	1736.50
1974	1441.33	14	0.4504	1666.00	0.4333	1702.30	0.4507	1665.40	0.4503	1666.20
1990	1440.94	15	0.4835	1595.70	0.4667	1631.40	0.4836	1595.50	0.4834	1595.90
1982	1251.04	16	0.5165	1525.40	0.5000	1560.60	0.5164	1525.60	0.5166	1525.20
1981	1192.14	17	0.5496	1454.50	0.5333	1489.50	0.5493	1455.10	0.5497	1454.30
1995	1151.79	18	0.5826	1383.10	0.5667	1417.60	0.5822	1384.00	0.5828	1382.70
1985	1100.82	19	0.6157	1310.60	0.6000	1345.10	0.6151	1311.90	0.6159	1310.20
1999	1083.83	20	0.6488	1237.00	0.6333	1271.60	0.6480	1238.80	0.6490	1236.50
1979	1068.11	21	0.6818	1162.10	0.6667	1196.60	0.6809	1164.20	0.6821	1161.40
1980	1066.41	22	0.7149	1085.20	0.7000	1120.10	0.7138	1087.80	0.7152	1084.50
2000	1055.51	23	0.7479	1006.30	0.7333	1041.50	0.7467	1009.20	0.7483	1005.30
1996	1013.98	24	0.7810	924.35	0.7667	960.13	0.7796	927.88	0.7815	923.09
1978	1005.40	25	0.8140	839.19	0.8000	875.80	0.8125	843.14	0.8146	837.60
1994	862.01	26	0.8471	749.25	0.8333	787.38	0.8454	754.00	0.8477	747.57
1977	814.50	27	0.8802	653.29	0.8667	693.29	0.8783	659.00	0.8808	651.48
1983	700.56	28	0.9132	549.08	0.9000	592.02	0.9112	555.71	0.9139	546.75
1984	422.49	29	0.9463	430.91	0.9333	479.50	0.9441	439.38	0.9470	428.20
1989	126.20	30	0.9793	284.63	0.9667	346.00	0.9770	296.61	0.9801	280.35

Appendix VIIILog-Pearson Type III Distribution using Probability of Exceedance.

YEAR	Max. Discharge Q (m^3/sec)	S/N	Gringorten	Predicted Q (m^3/sec)	Hazen	Predicted Q (m^3/sec)	Turkey	Predicted Q (m^3/sec)	Weibull	Predicted Q (m^3/sec)
1993	2926.31	1	0.0186	2743.90	0.0167	2752.00	0.0220	2729.70	0.0323	2689.20
1986	2871.75	2	0.0518	2619.90	0.0500	2626.00	0.0549	2609.60	0.0645	2578.50
1992	2579.50	3	0.0850	2515.80	0.0833	2520.90	0.0879	2507.30	0.0968	2481.60
1998	2488.34	4	0.1182	2422.30	0.1167	2426.40	0.1209	2415.00	0.1290	2393.50
1997	2395.85	5	0.1514	2335.50	0.1500	2339.00	0.1538	2329.40	0.1613	2310.60
1987	2370.38	6	0.1846	2253.30	0.1833	2256.40	0.1868	2248.00	0.1935	2231.90
1976	2362.50	7	0.2178	2174.50	0.2167	2177.00	0.2198	2169.80	0.2258	2155.90
1973	2296.52	8	0.2510	2098.10	0.2500	2100.40	0.2527	2094.30	0.2581	2082.10
1972	2270.70	9	0.2842	2023.70	0.2833	2025.70	0.2857	2020.40	0.2903	2010.20
1988	2072.57	10	0.3174	1950.70	0.3167	1952.30	0.3187	1947.90	0.3226	1939.40
1975	1795.29	11	0.3506	1878.80	0.3500	1880.10	0.3516	1876.70	0.3548	1869.80
1971	1469.64	12	0.3838	1807.60	0.3833	1808.70	0.3846	1805.90	0.3871	1800.60
1991	1460.45	13	0.4170	1736.90	0.4167	1737.50	0.4176	1735.60	0.4194	1731.80
1974	1441.33	14	0.4502	1666.40	0.4500	1666.80	0.4505	1665.80	0.4516	1663.40
1990	1440.94	15	0.4834	1595.90	0.4833	1596.10	0.4835	1595.70	0.4839	1594.90
1982	1251.04	16	0.5166	1525.20	0.5167	1525.00	0.5165	1525.40	0.5161	1526.30
1981	1192.14	17	0.5498	1454.10	0.5500	1453.60	0.5495	1454.70	0.5484	1457.10
1995	1151.79	18	0.5830	1382.20	0.5833	1381.60	0.5824	1383.50	0.5806	1387.50
1985	1100.82	19	0.6162	1309.50	0.6167	1308.40	0.6154	1311.30	0.6129	1316.80
1999	1083.83	20	0.6494	1235.60	0.6500	1234.30	0.6484	1237.90	0.6452	1245.10
1979	1068.11	21	0.6826	1160.30	0.6833	1158.70	0.6813	1163.20	0.6774	1172.20
1980	1066.41	22	0.7158	1083.10	0.7167	1080.90	0.7143	1086.60	0.7097	1097.40
2000	1055.51	23	0.7490	1003.60	0.7500	1001.20	0.7473	1007.70	0.7419	1020.80
1996	1013.98	24	0.7822	921.32	0.7833	918.54	0.7802	926.37	0.7742	941.44
1978	1005.40	25	0.8154	835.48	0.8167	832.03	0.8132	841.30	0.8065	858.89
1994	862.01	26	0.8486	745.05	0.8500	741.11	0.8462	751.77	0.8387	772.58
1977	814.50	27	0.8818	648.46	0.8833	643.92	0.8791	656.60	0.8710	680.69
1983	700.56	28	0.9150	543.07	0.9167	537.36	0.9121	552.74	0.9032	581.78
1984	422.49	29	0.9482	423.51	0.9500	416.42	0.9451	435.54	0.9355	471.50
1989	126.20	30	0.9814	273.28	0.9833	262.63	0.9780	291.45	0.9677	341.46

Appendix IX. Exponential Model and Matlab steps for normal distribution on how to plot curves using the respective plotting position formulae as titled to each of the plate below.

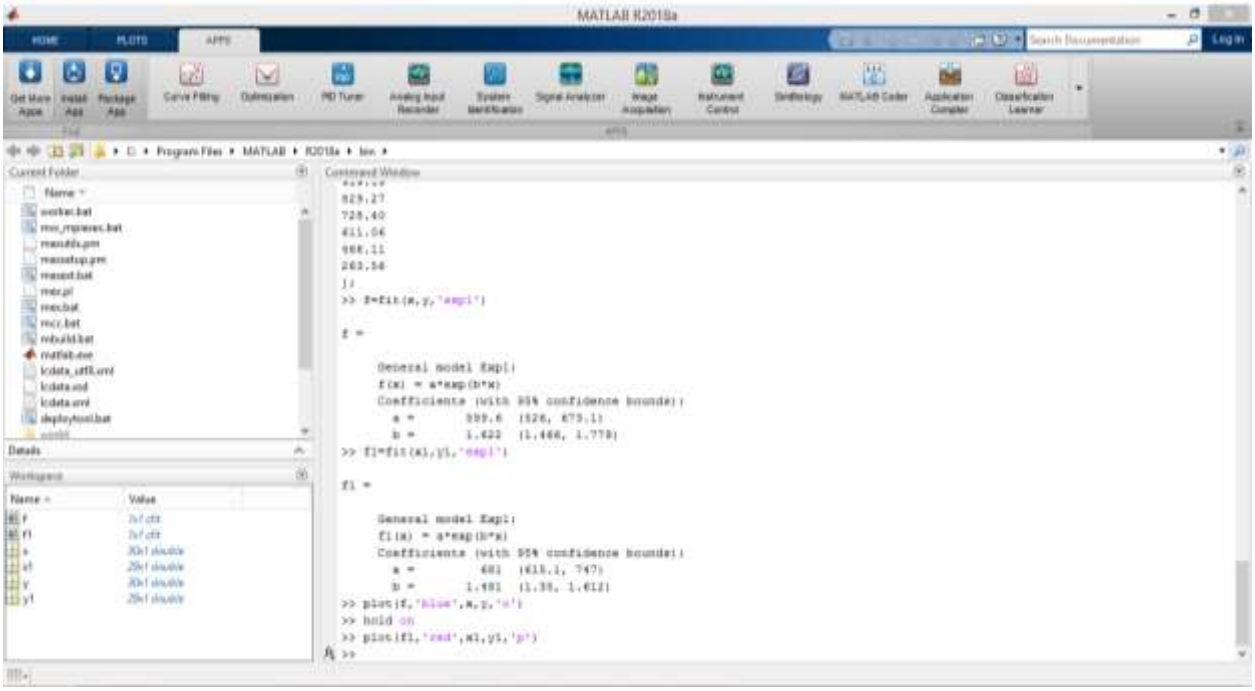


Plate 1. California.

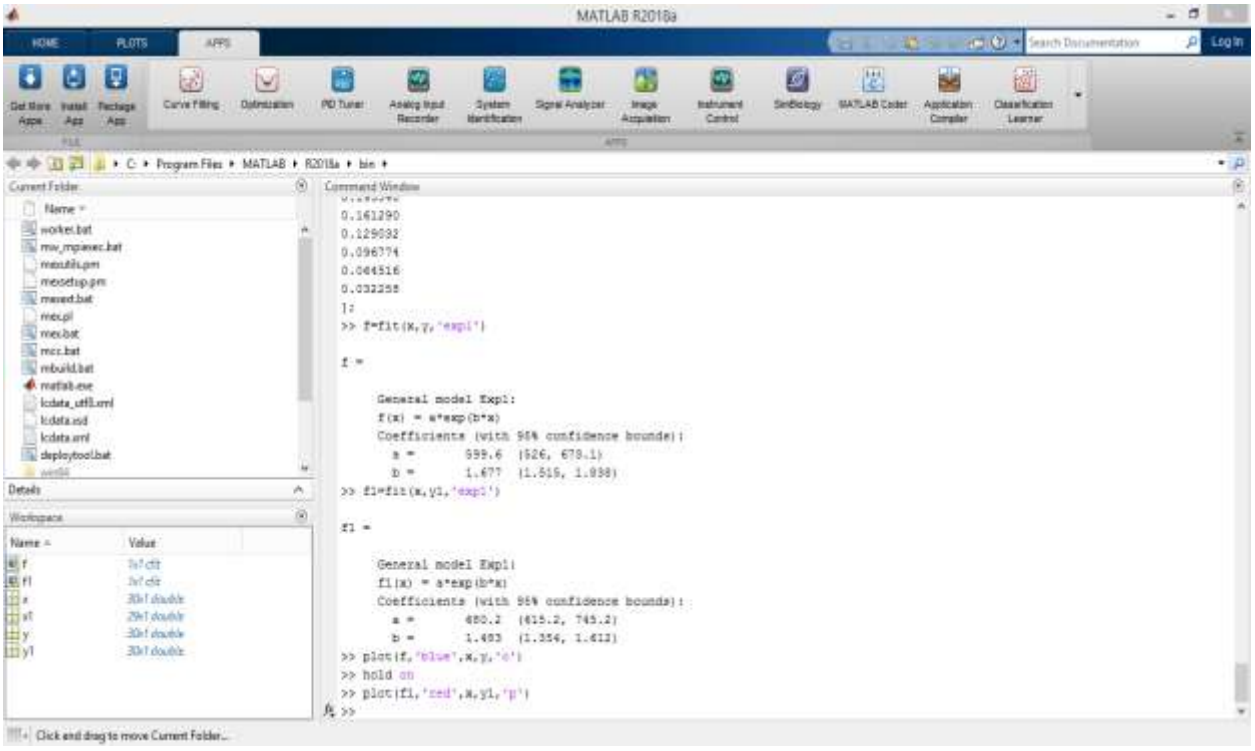


Plate 2. Weibull

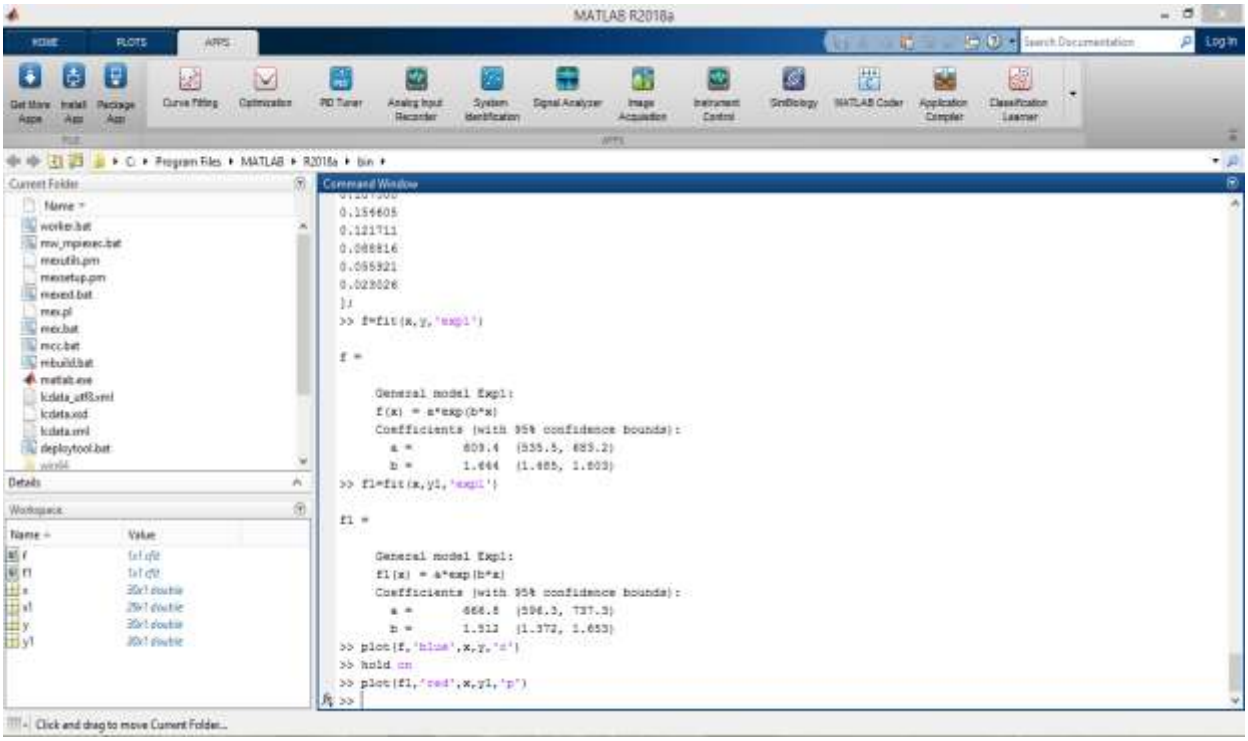


Plate 3. Chegodeyev.

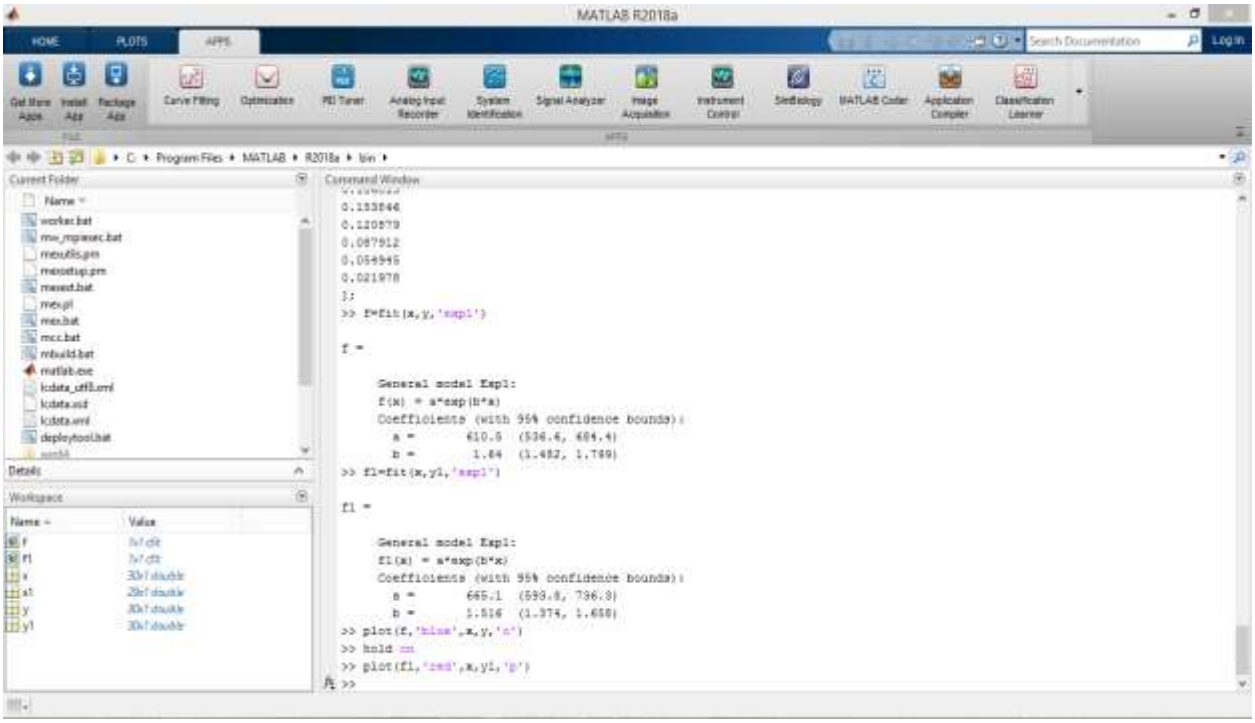


Plate 4. Turkey

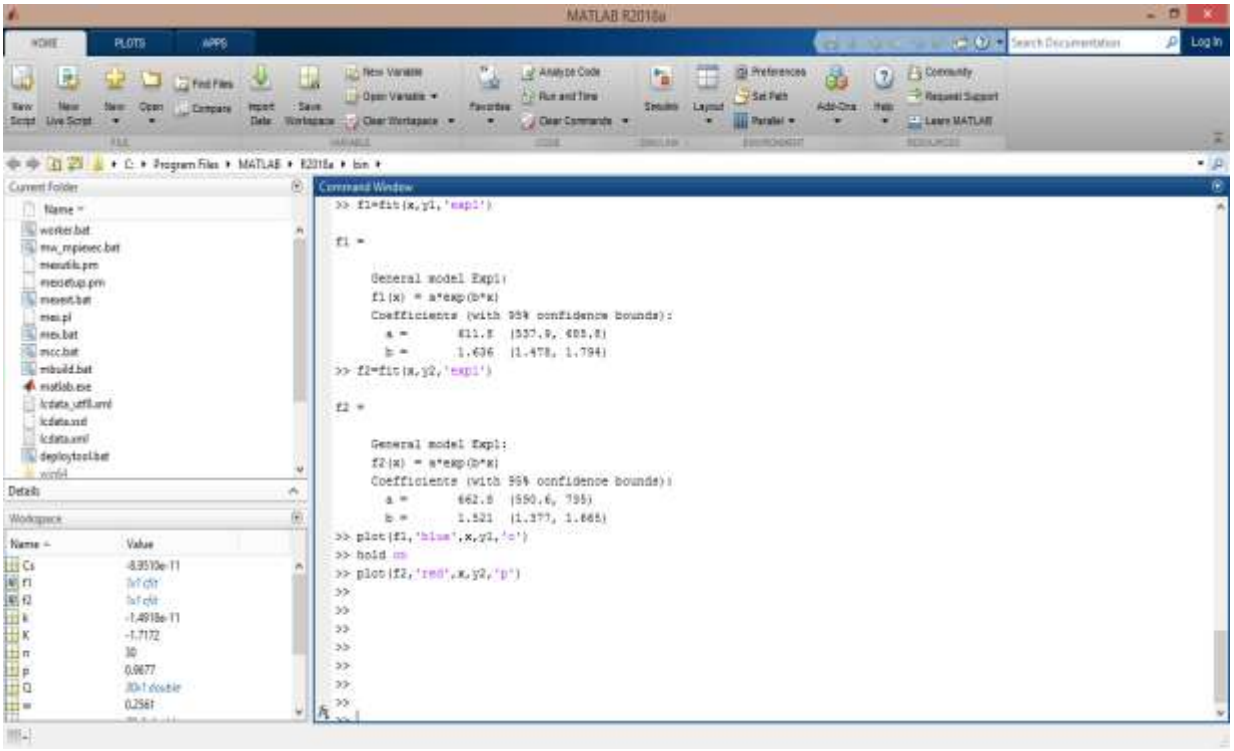


Plate 5. Blom.

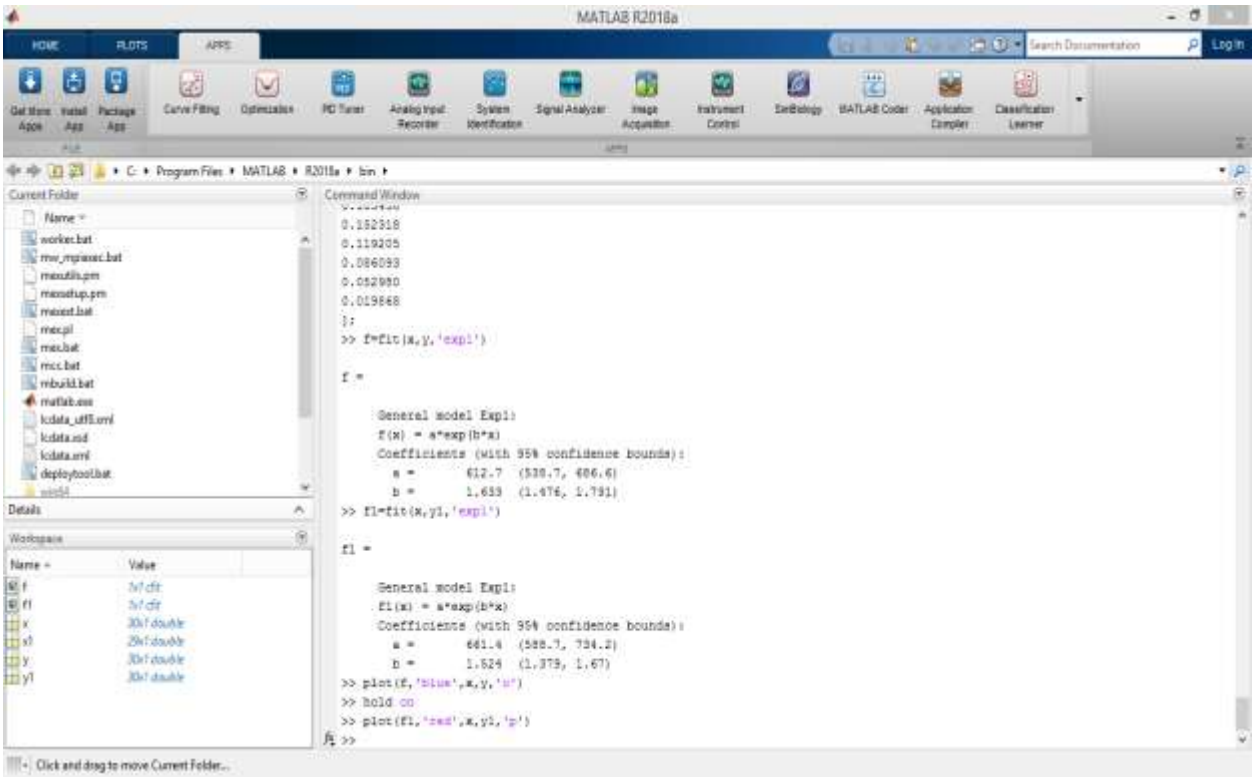


Plate 6. Cunnane

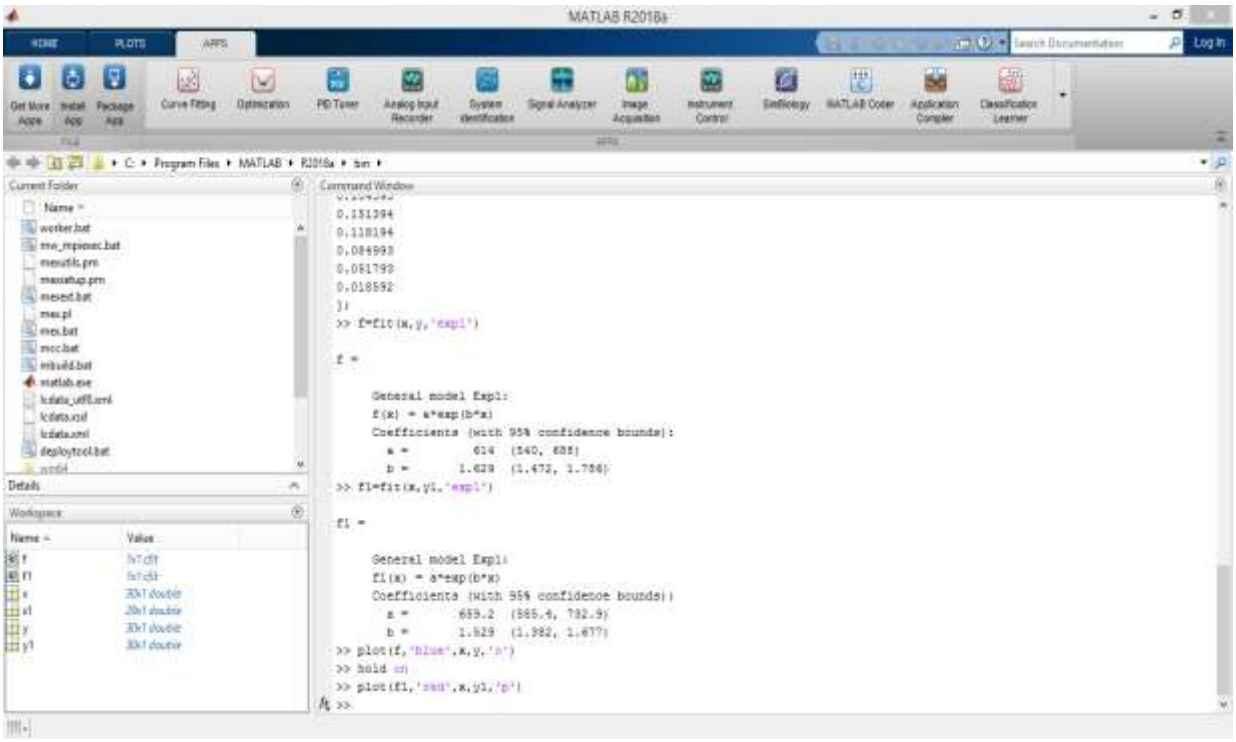


Plate 7. Gringorten.

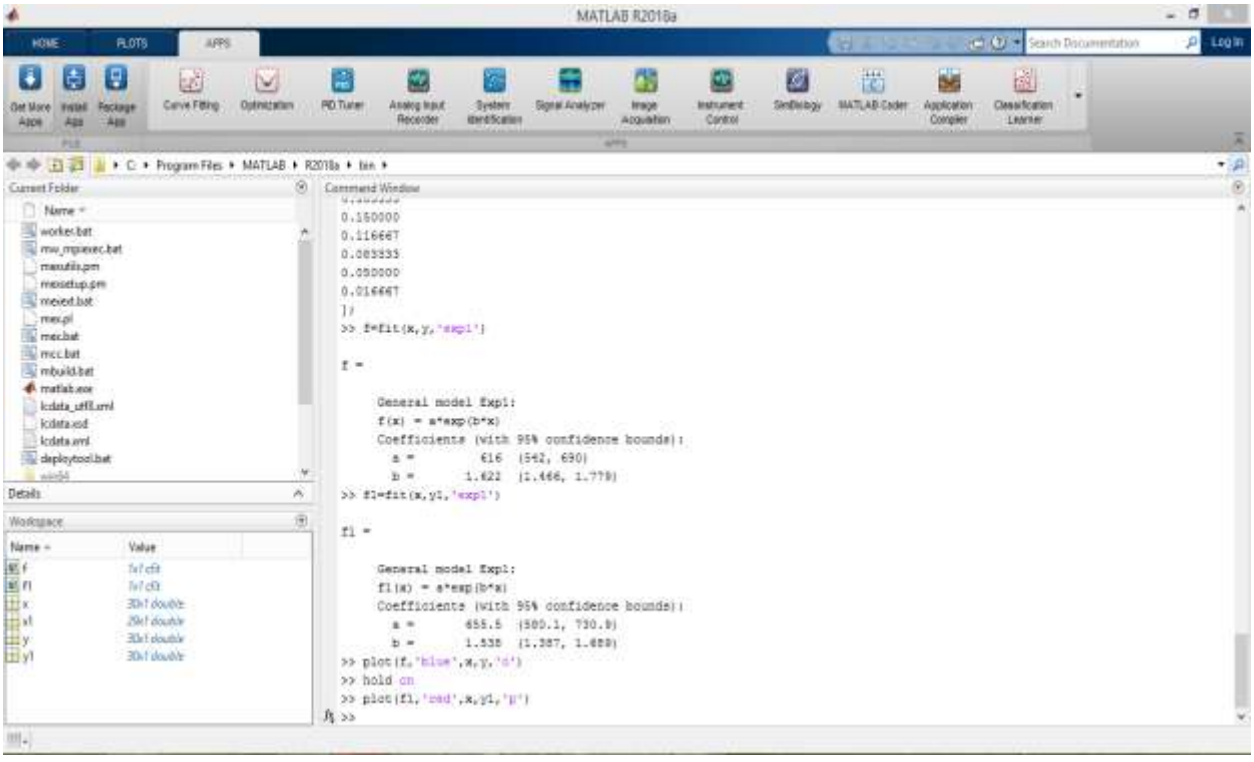


Plate 8. Hazen.

Appendix X. Exponential Model and Matlab steps for **log-normal distribution** on how to plot curves using the respective plotting position formulae as titled to each of the plate below.

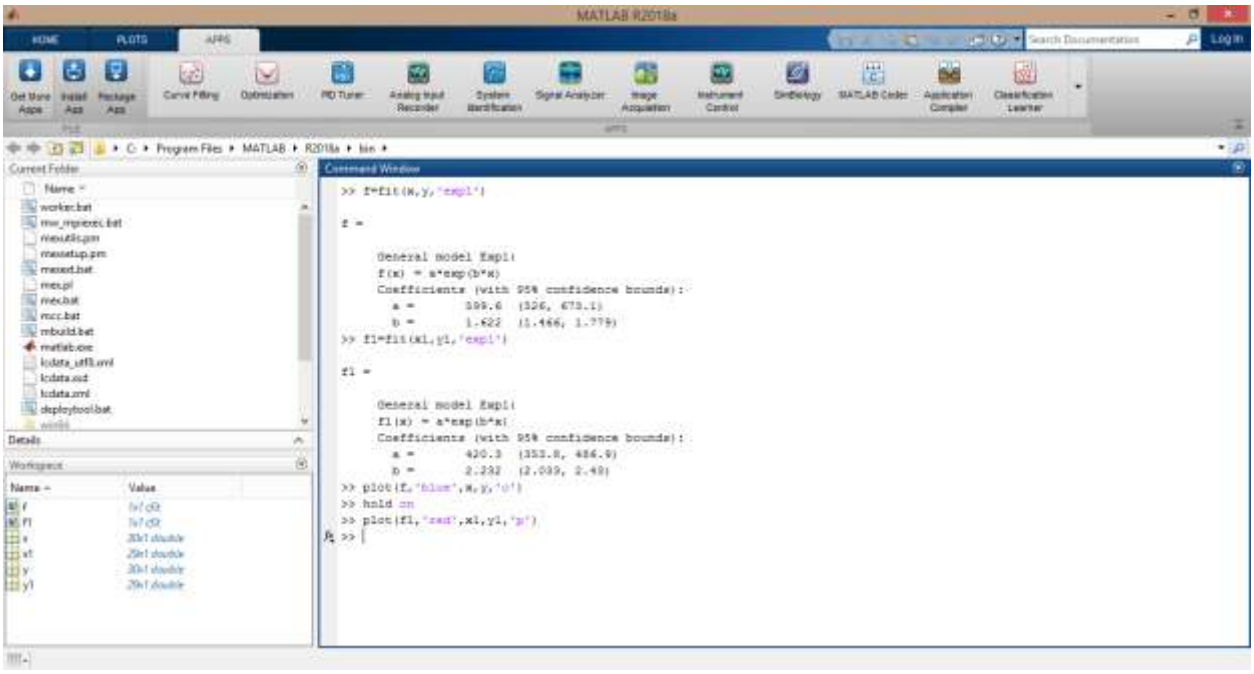


Plate 9. California.

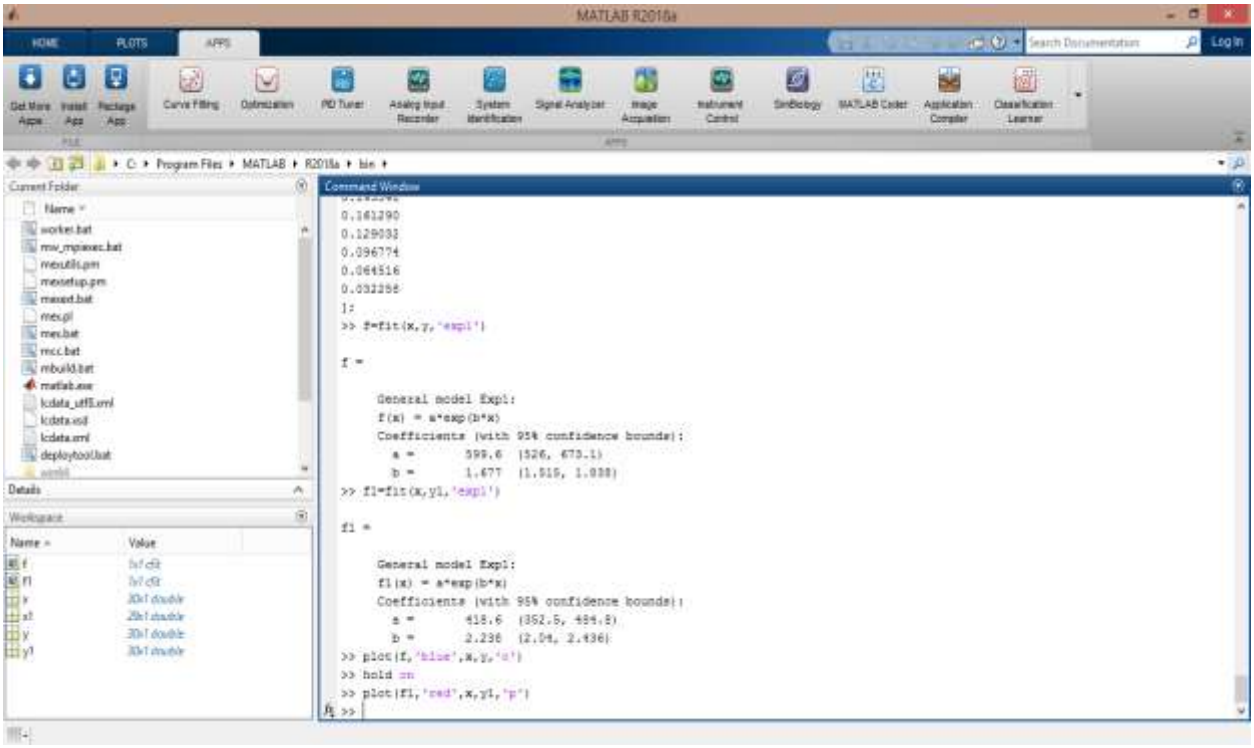


Plate 10. Weibull.

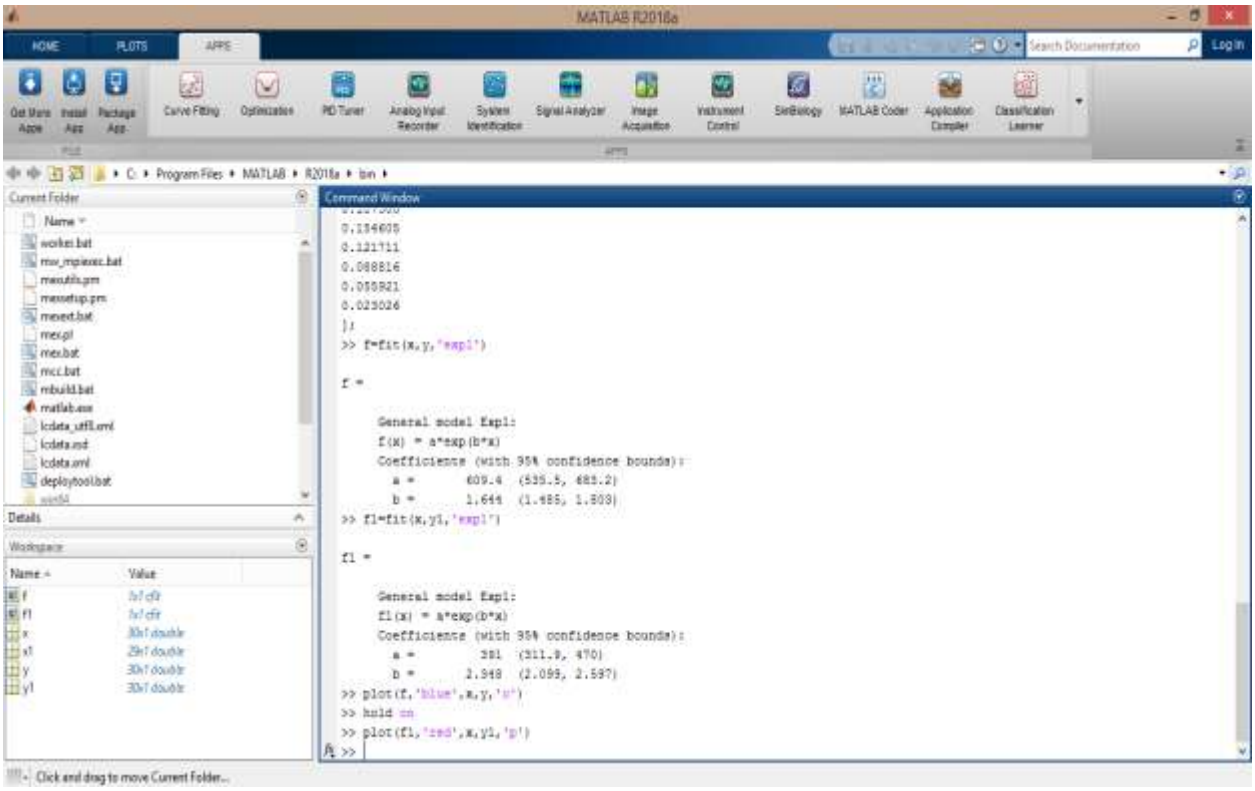


Plate 11. Chegodayev.

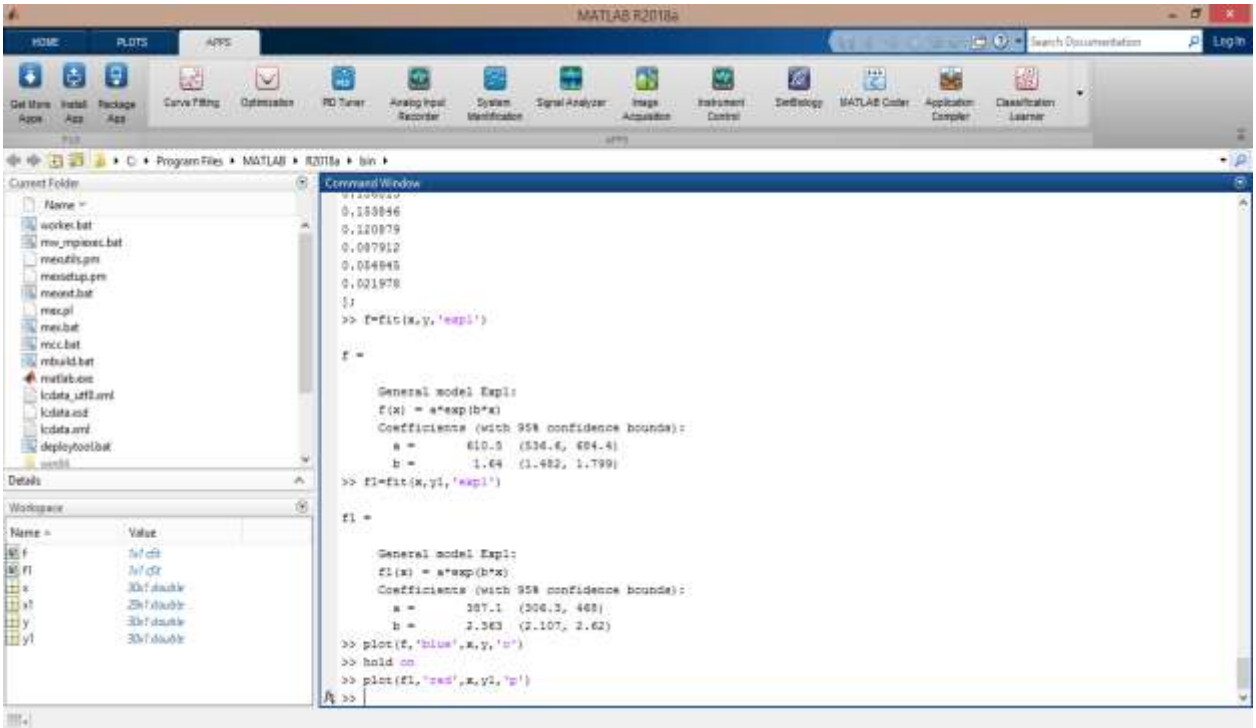


Plate 12. Turkey.

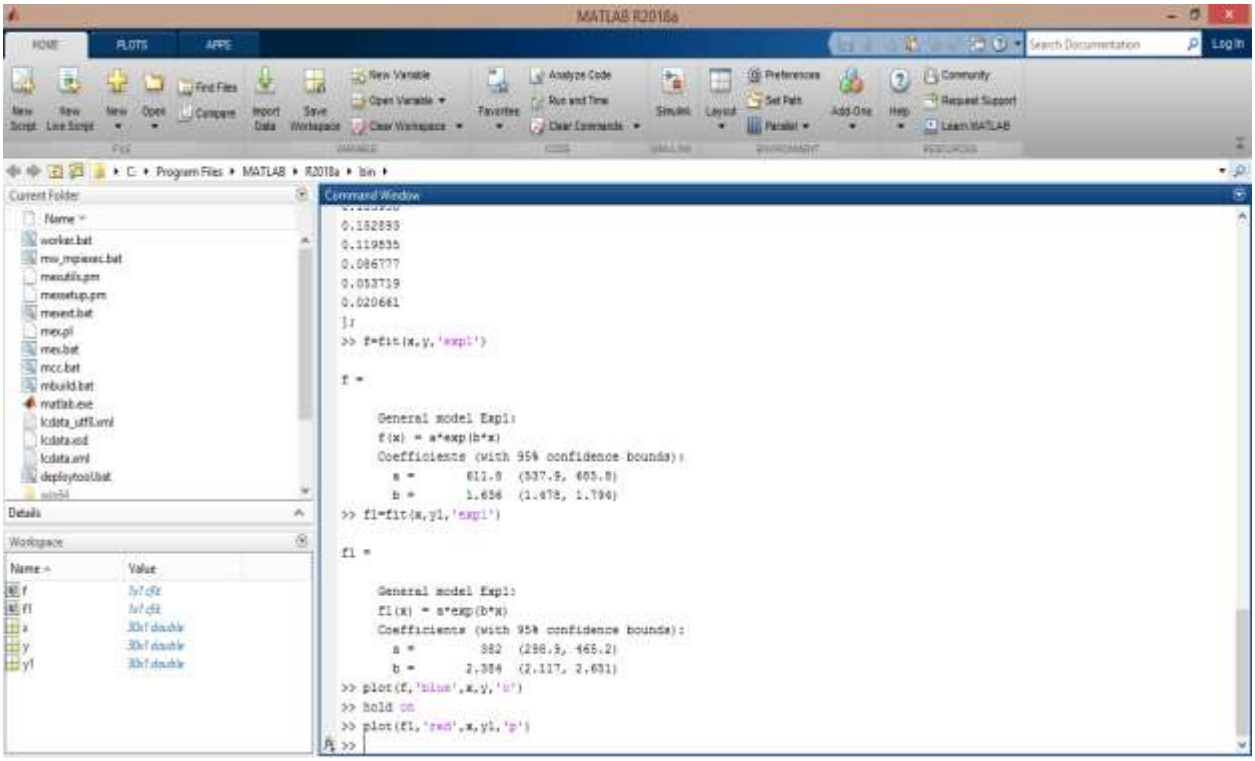


Plate 13. Blom.

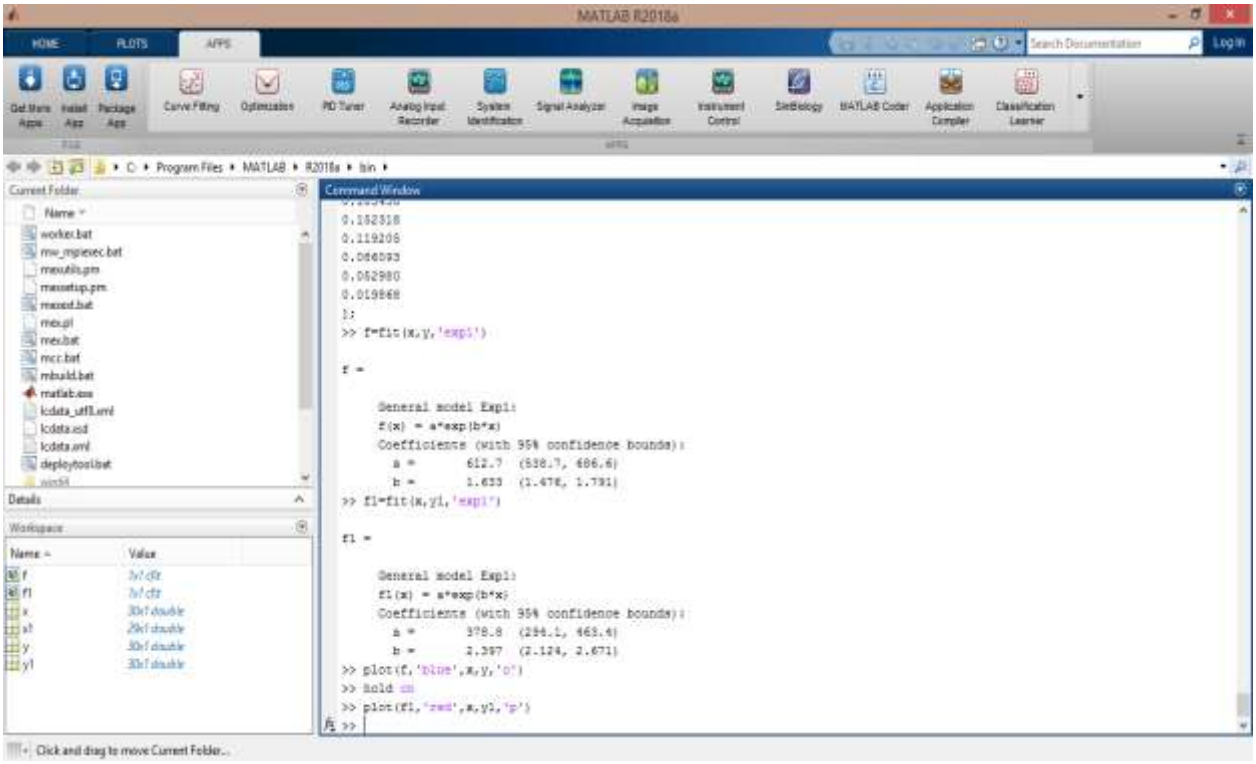


Plate 14. Cunnane

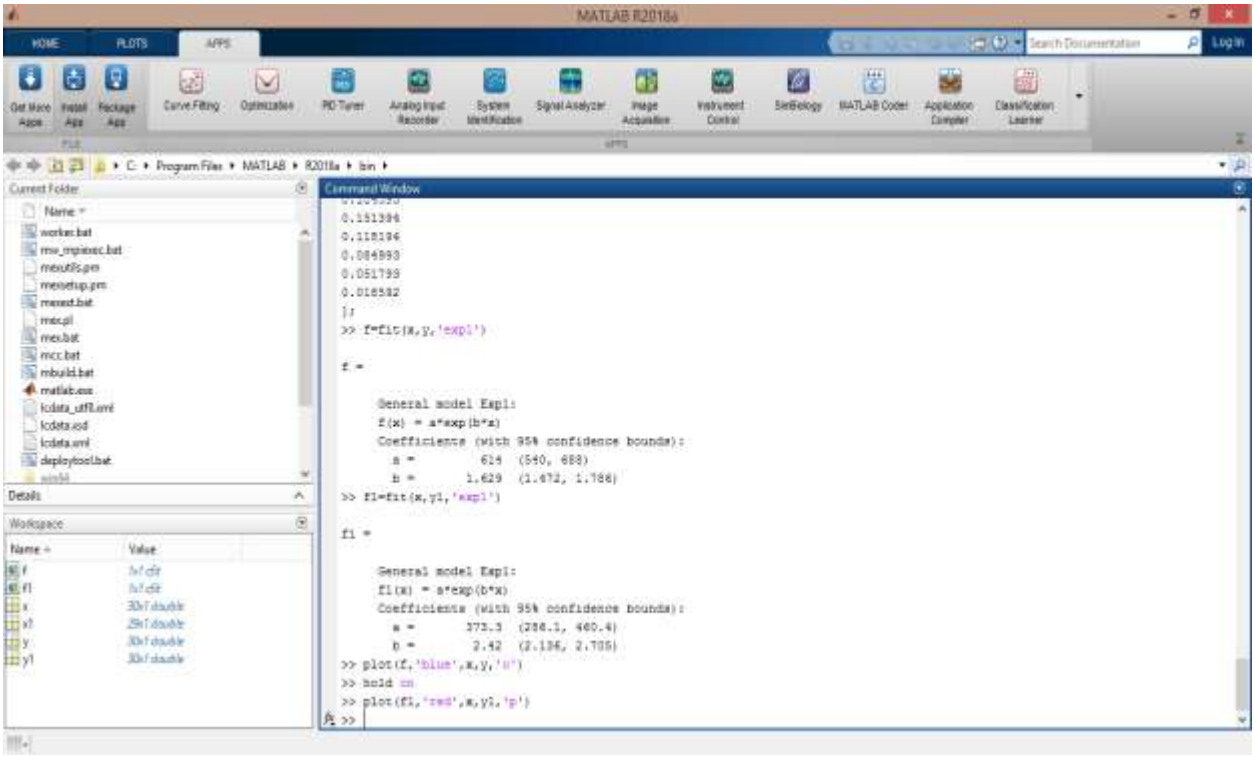


Plate 15. Gringorten.

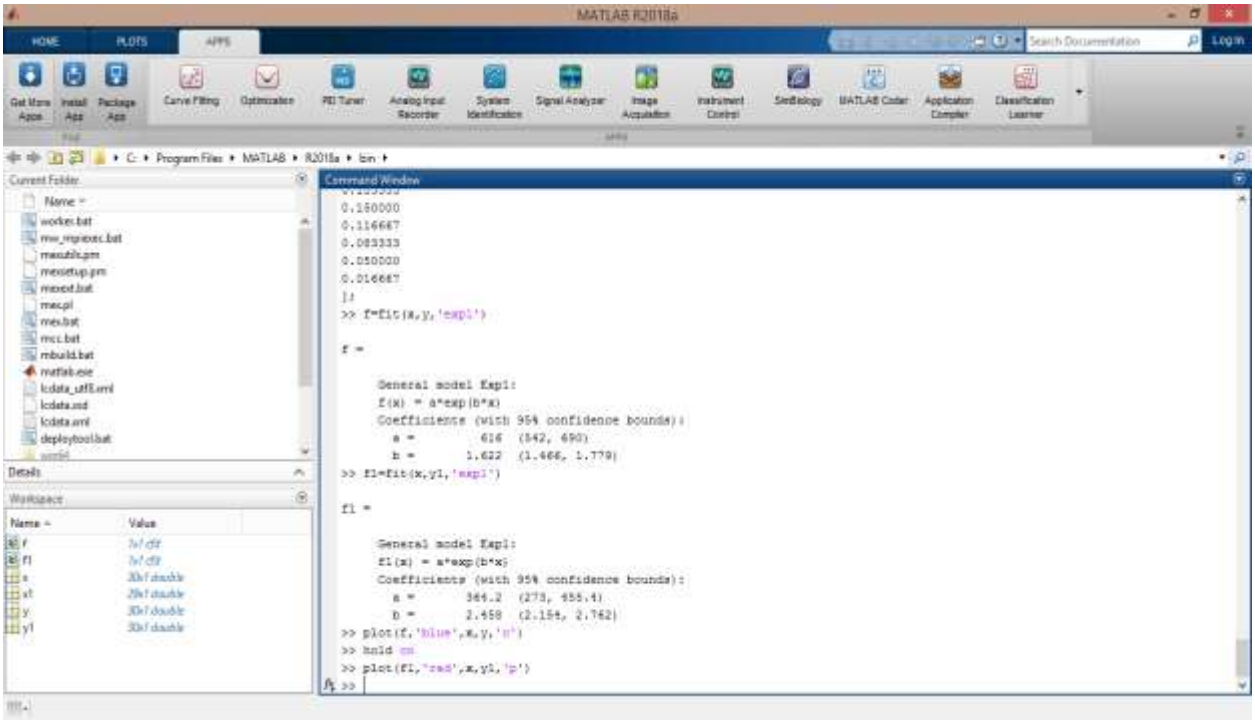


Plate 16. Hazen

Appendix XI. Exponential Model and Matlab steps for **Pearson type III distribution** on how to plot curves using the respective plotting position formulae as titled to each of the plate below.

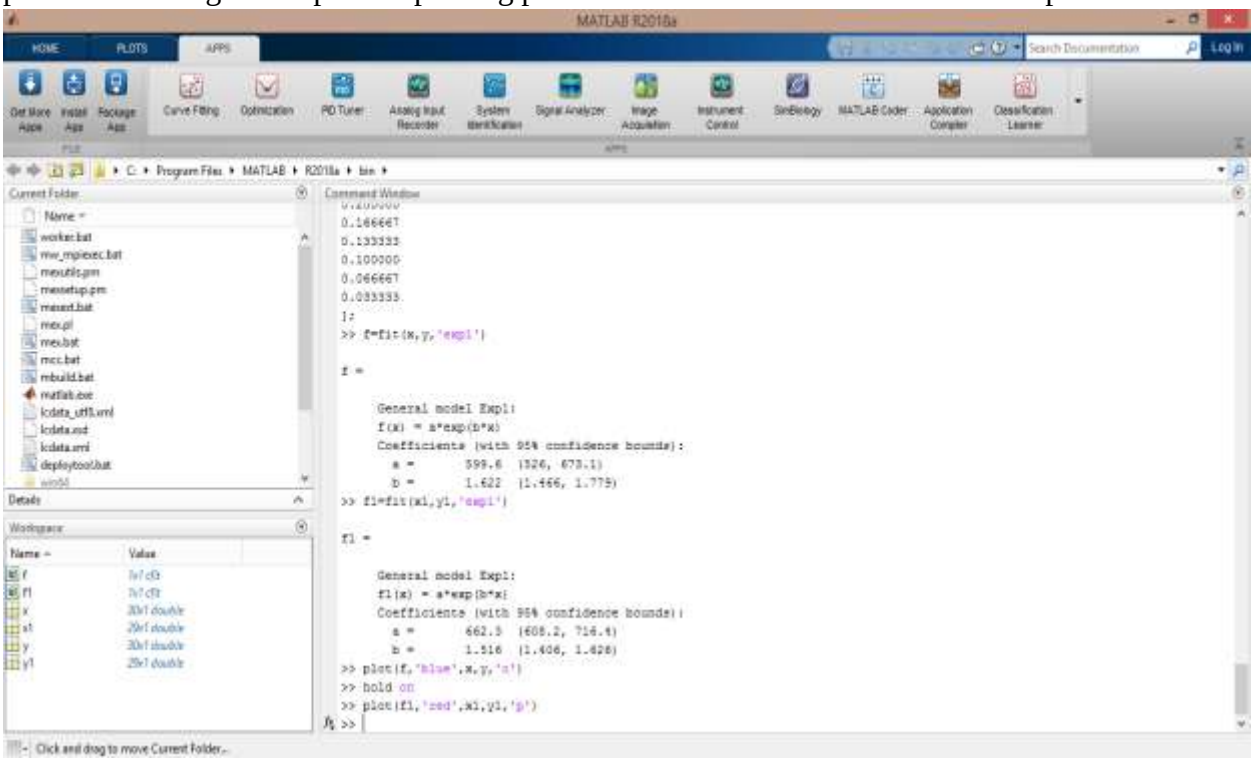


Plate 17. California

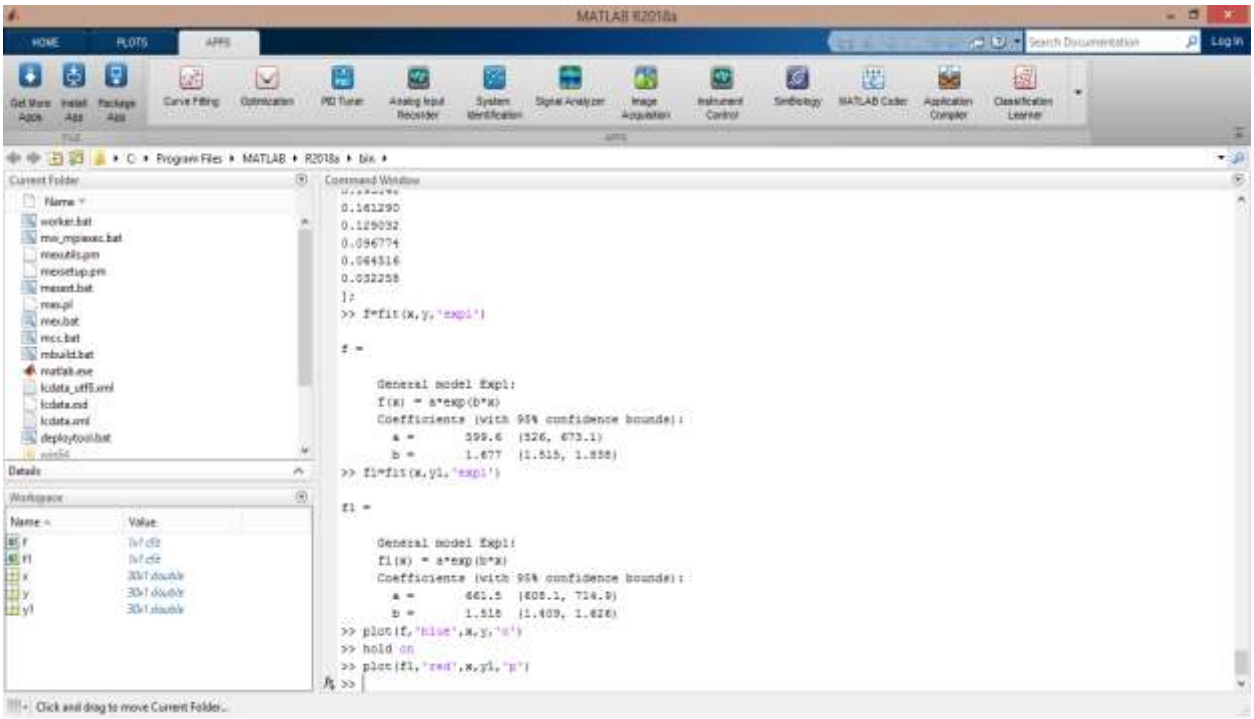


Plate 18. Weibull.

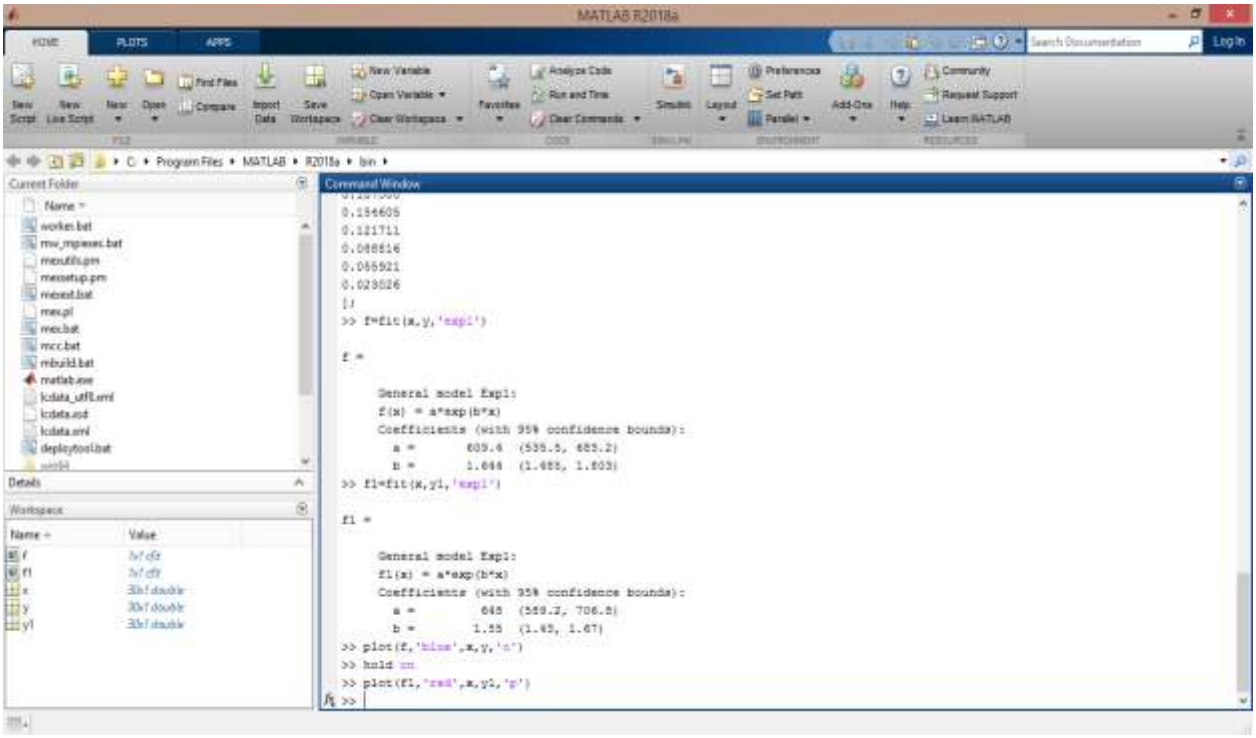


Plate 19. Chegodayev.

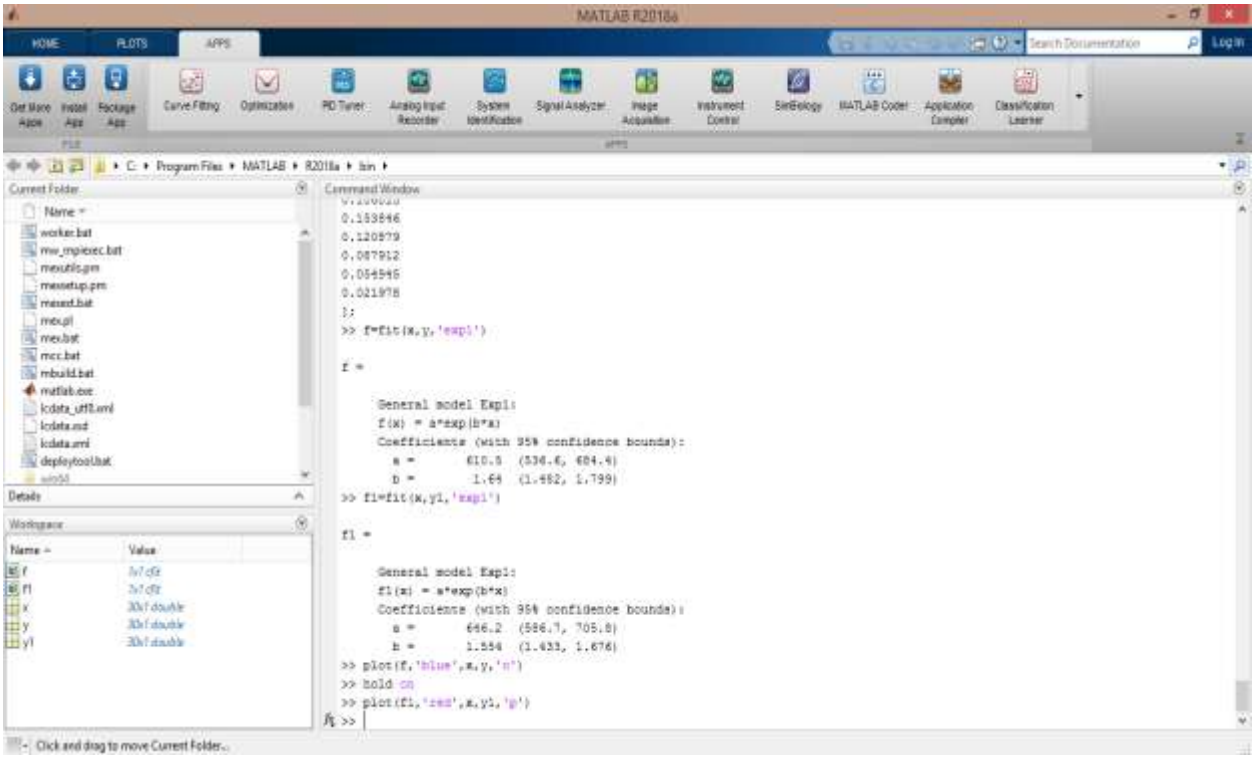


Plate 20. Turkey

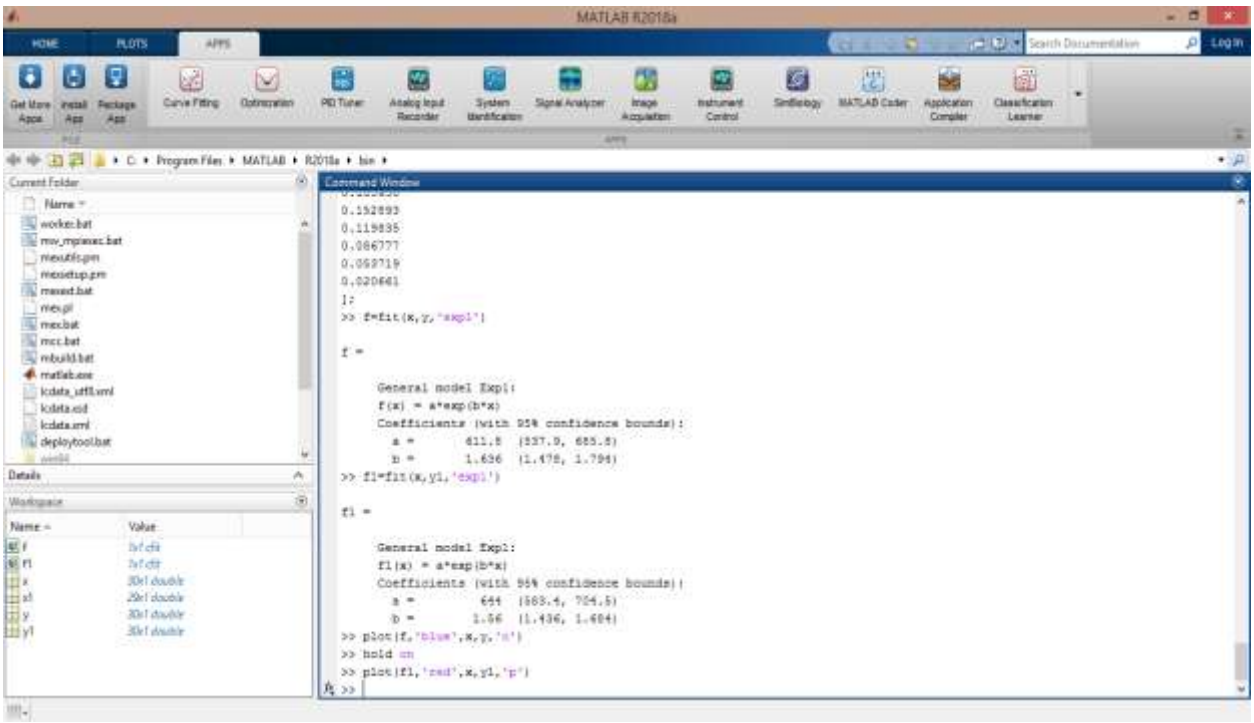


Plate 21. Blom.

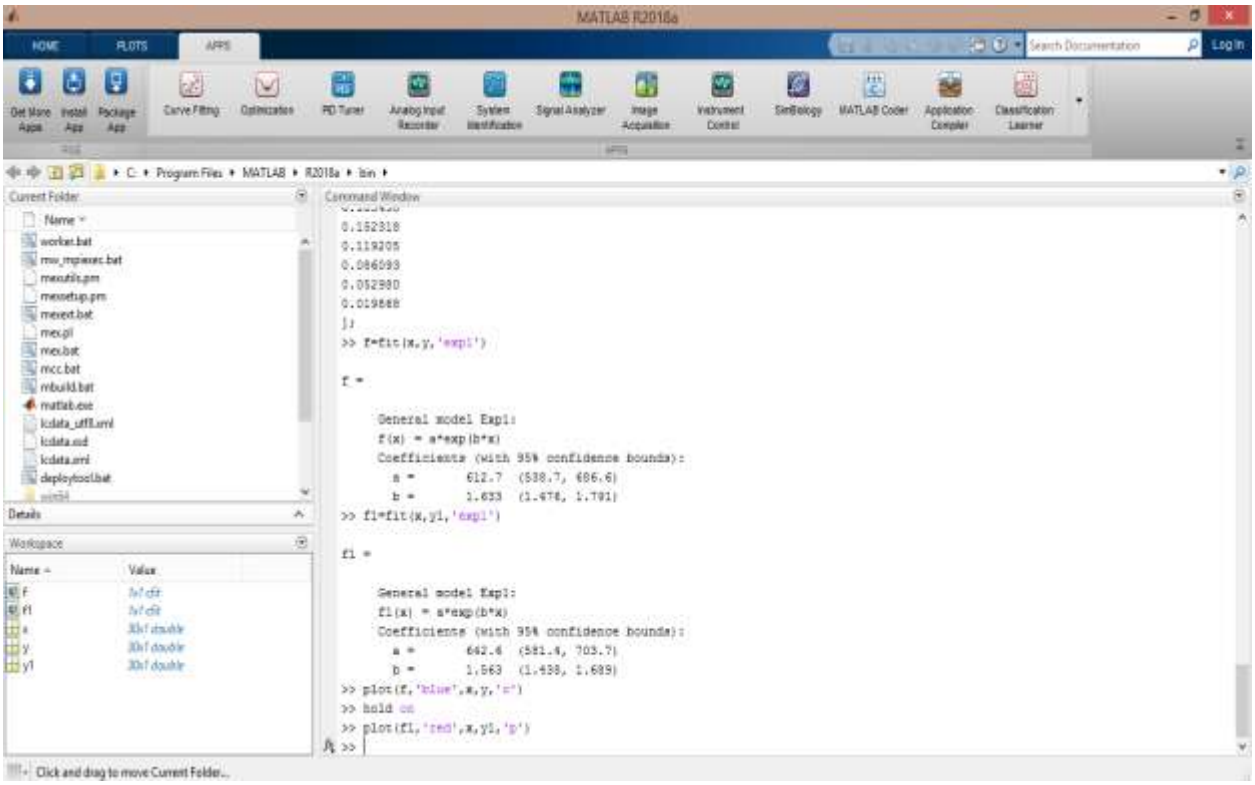


Plate 22. Cunnane

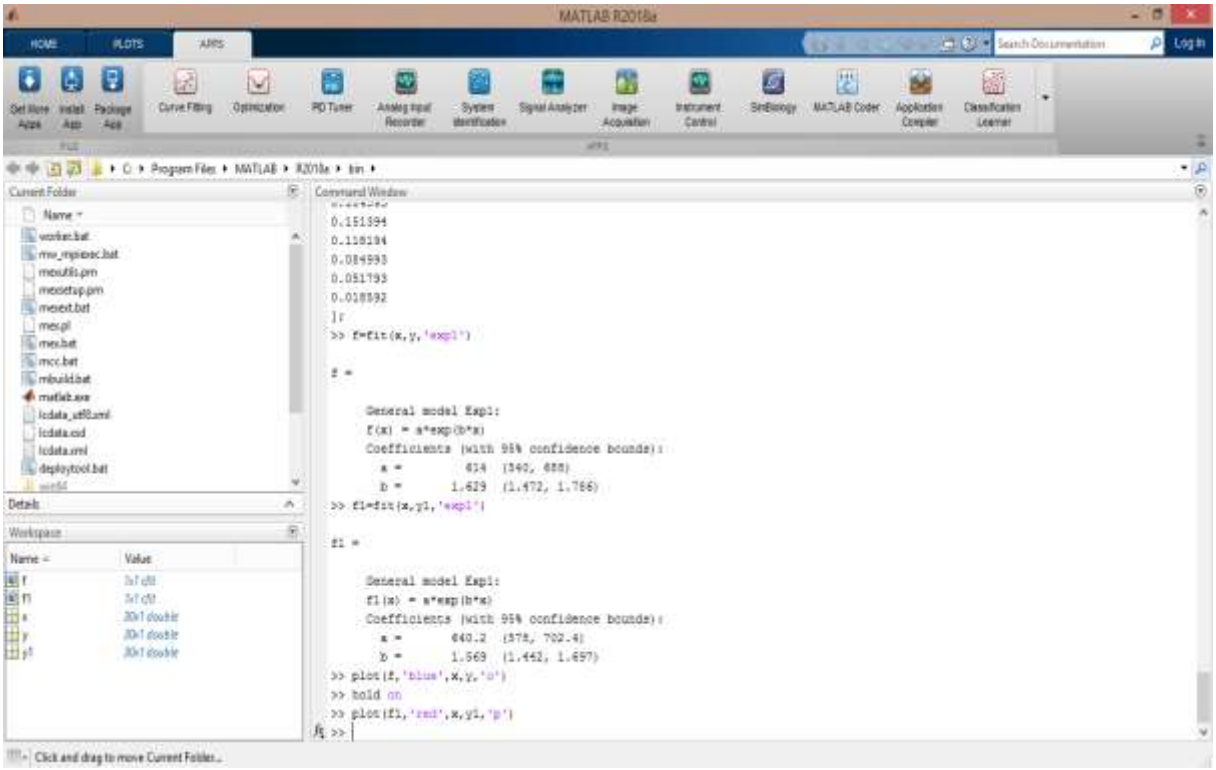


Plate 23. Gringorten.

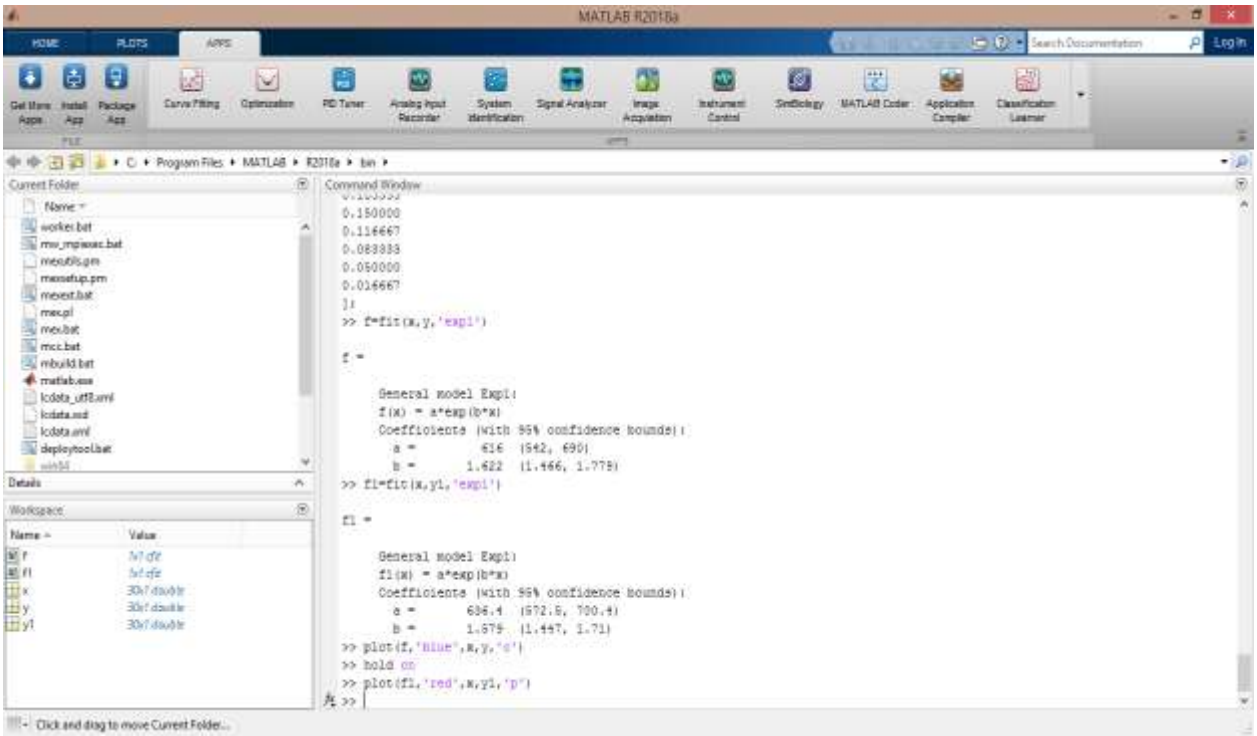


Plate 24. Hazen.

Appendix XII. Exponential Model and Matlab steps for **Log-Pearson Type III** distribution on how to plot curves using the respective plotting position formulae as titled to each of the plate below.

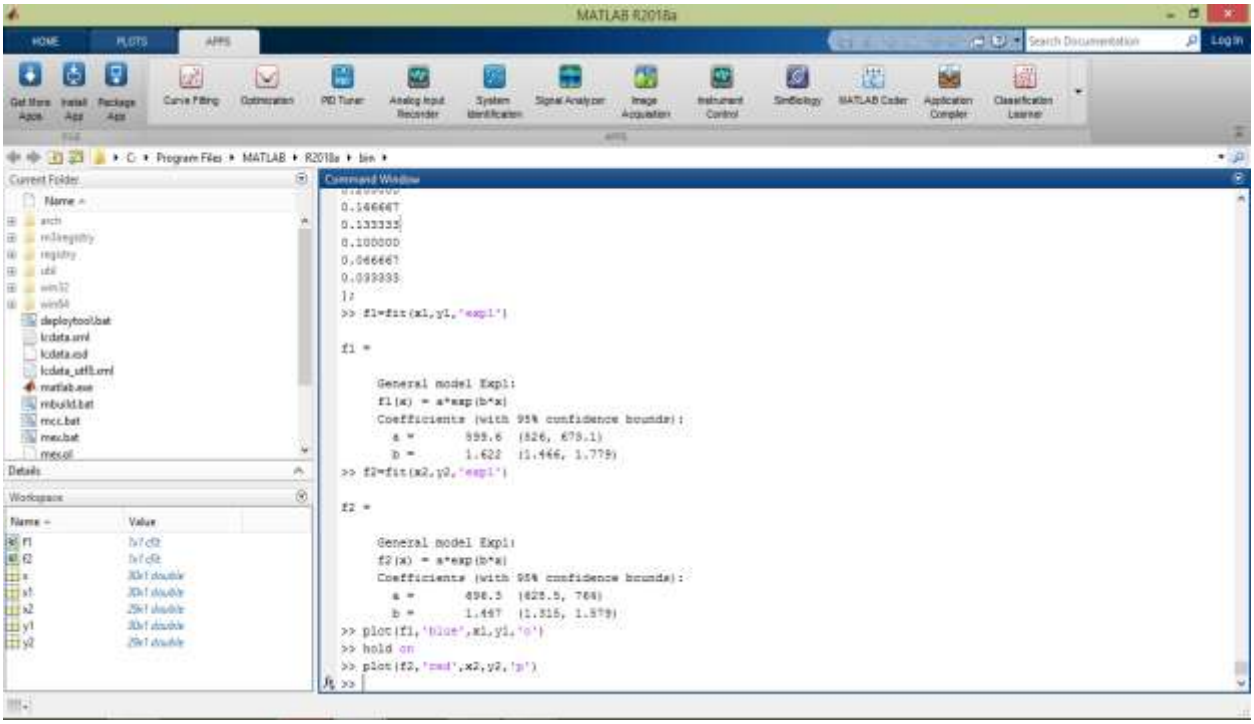


Plate 25. California

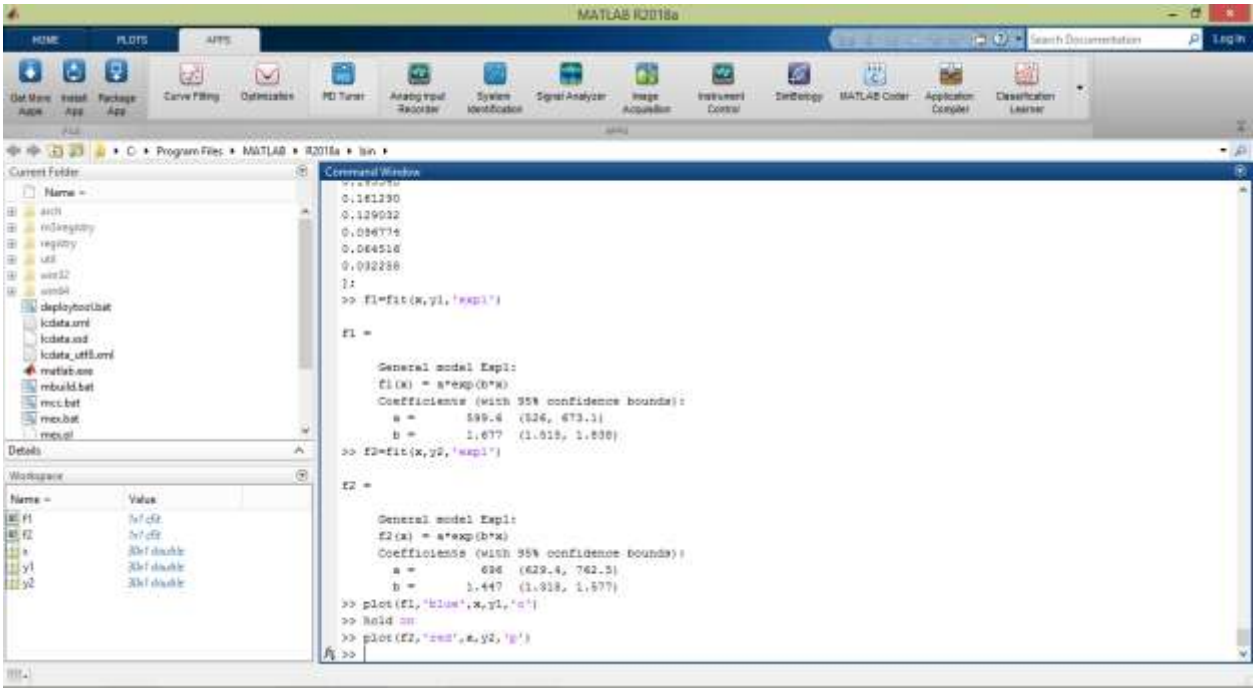


Plate 26. Weibull.

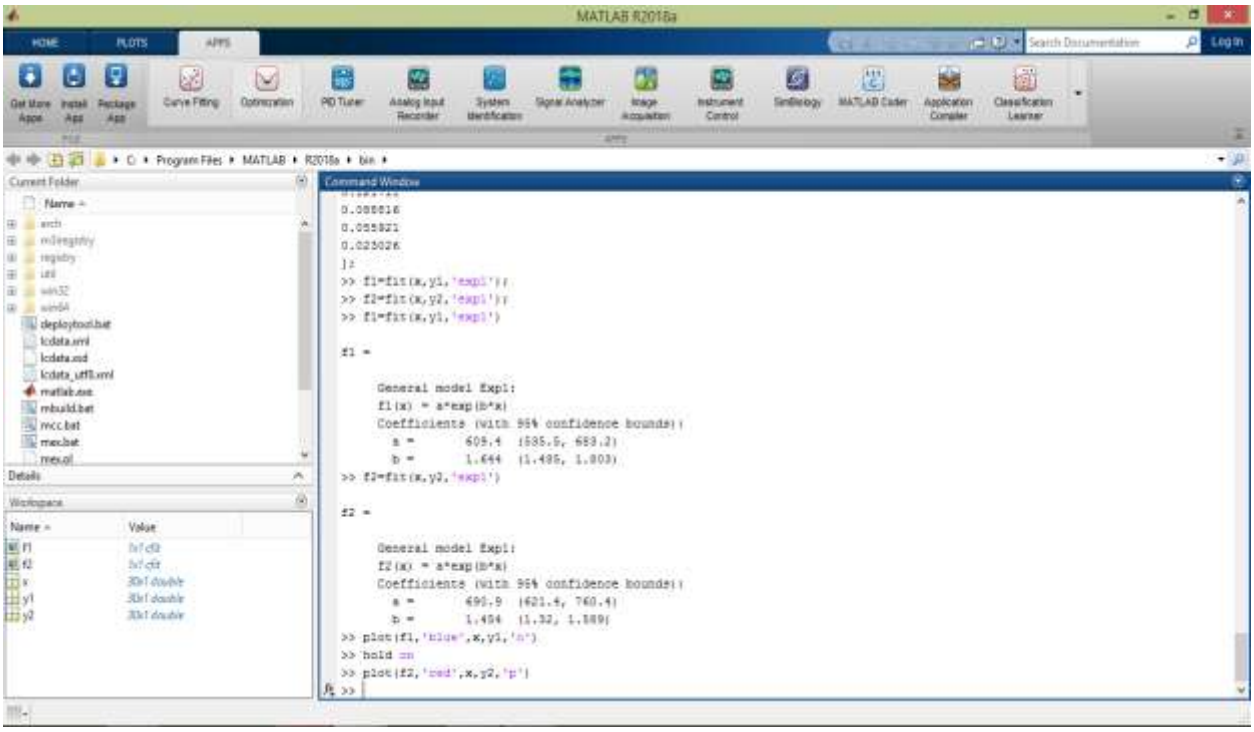


Plate 27. Chegodayev.

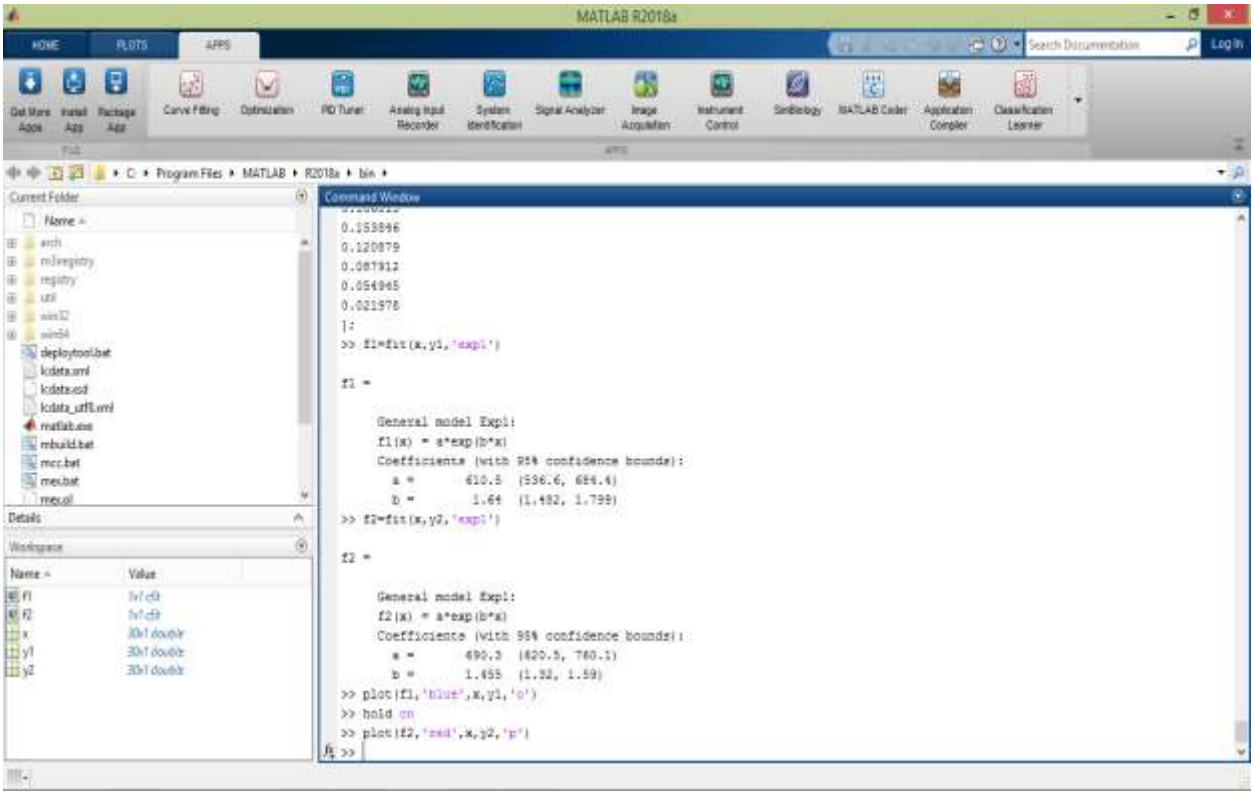


Plate 28. Turkey.

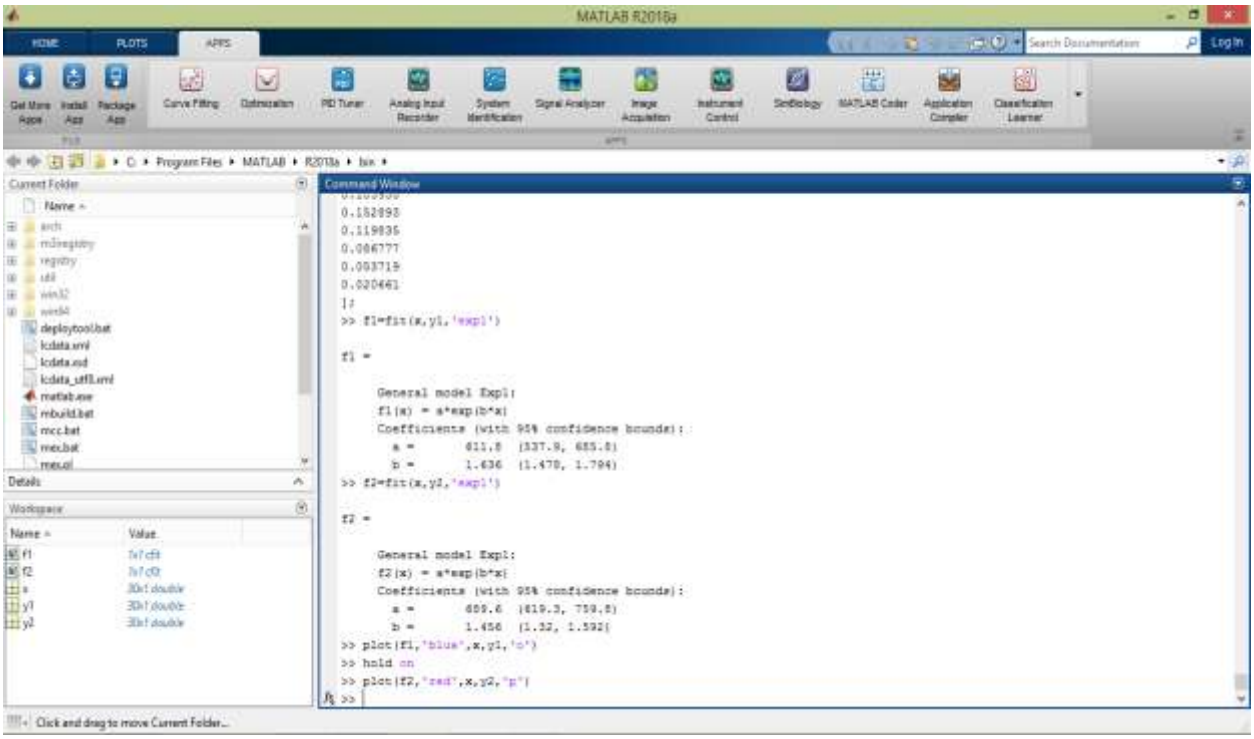


Plate 29. Blom.

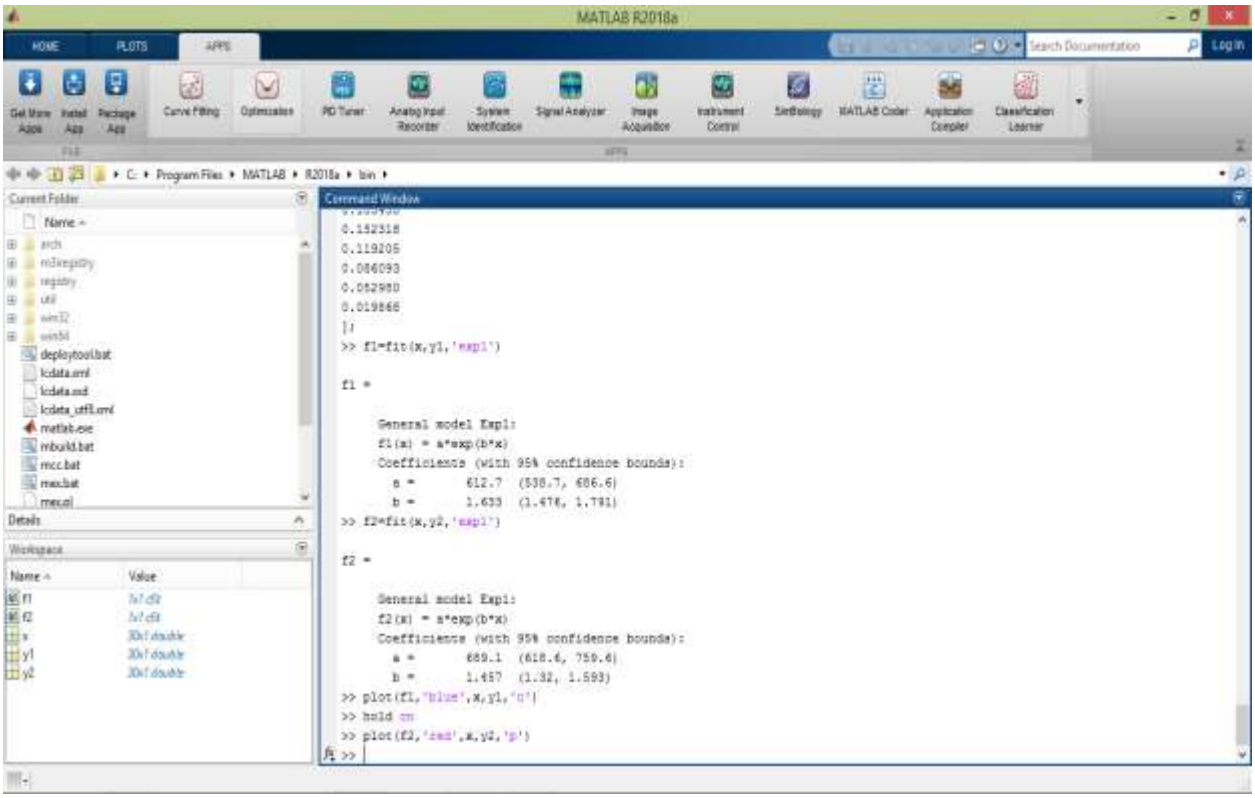


Plate 30. Cunnane

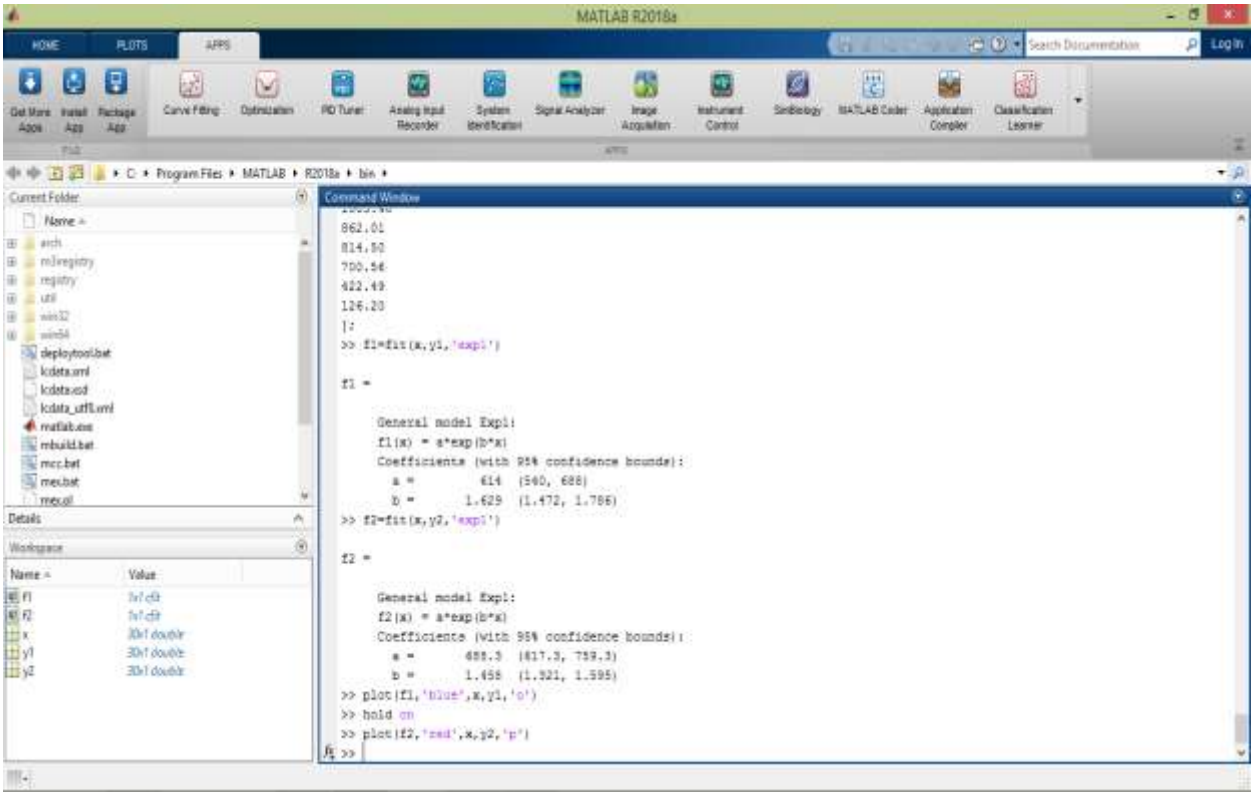


Plate 31. Gringorten

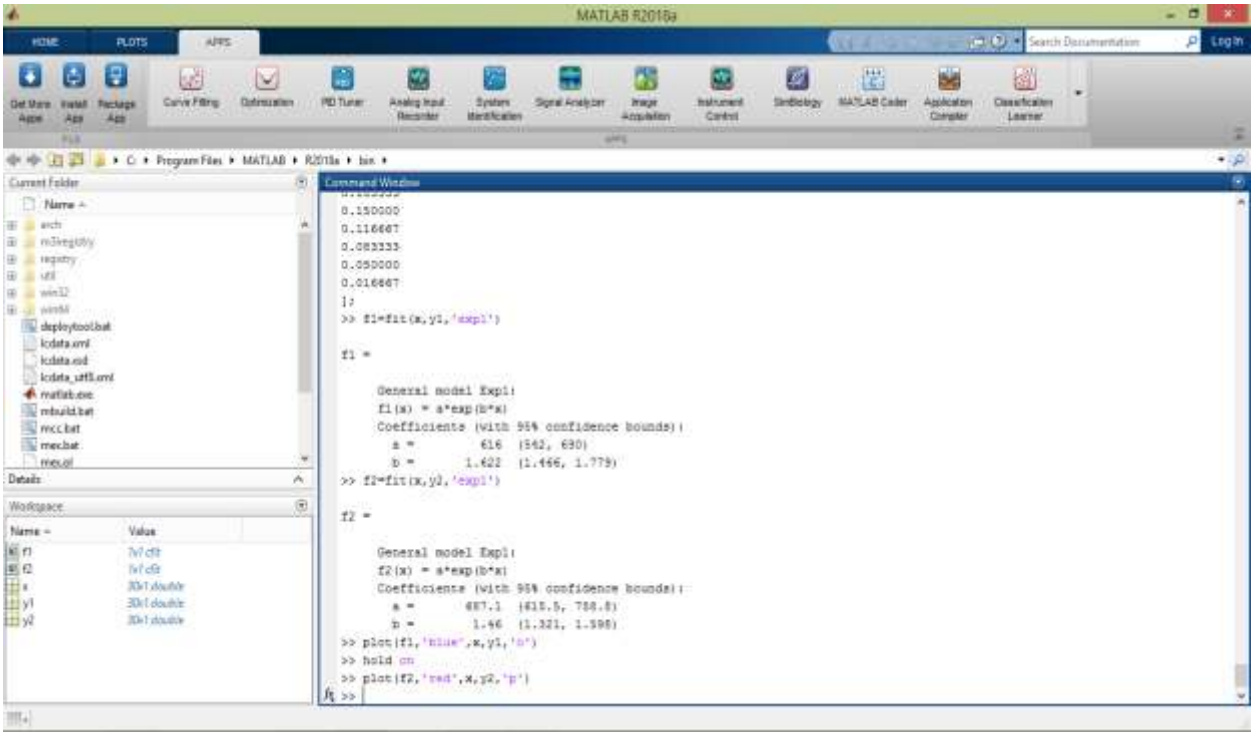


Plate 32. Hazen.

Appendix XIII. Results of R-Square and RMSE values for Normal Distribution in respect of the below plotting positions.

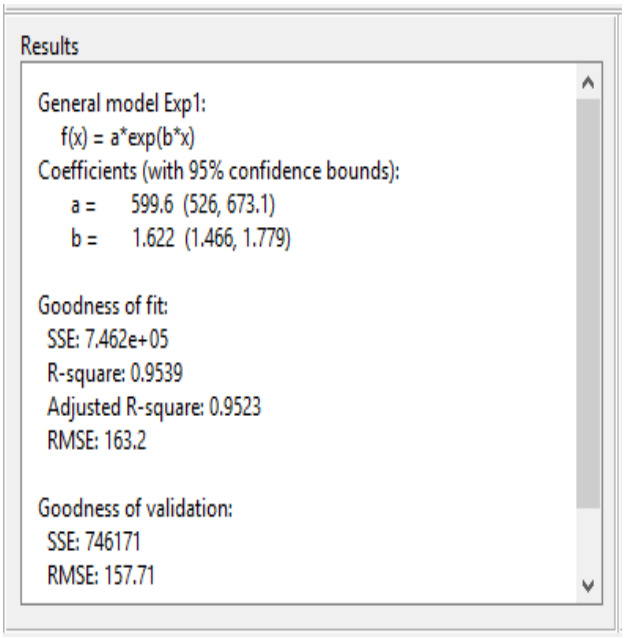


Plate 33a. California using Observed data

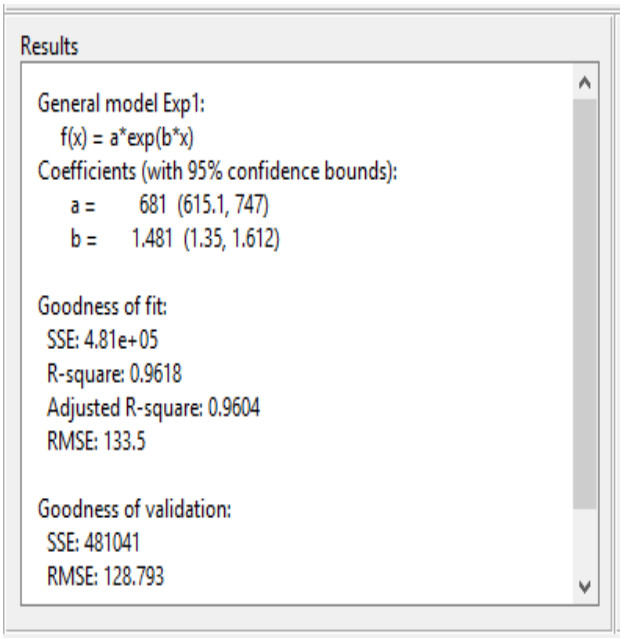


Plate 33b. California using Predicted data

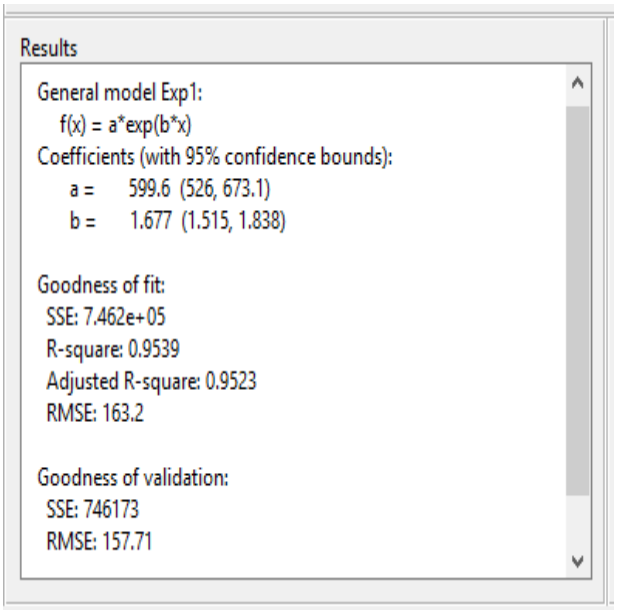


Plate 34a. Weibull using Observed data.

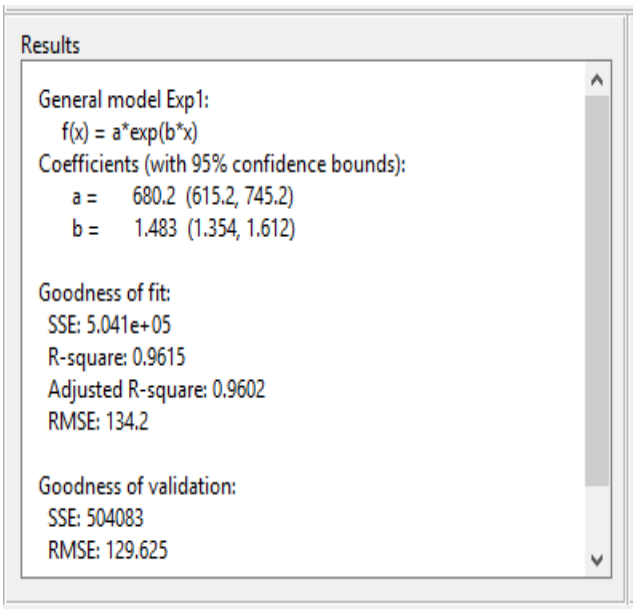


Plate 34b. Weibull using Predicted data.

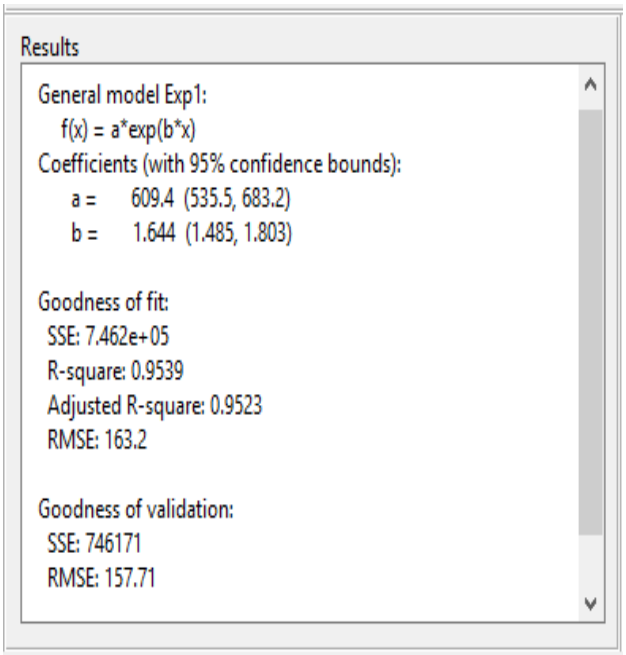


Plate 35a. Chegodayev using Observed data.

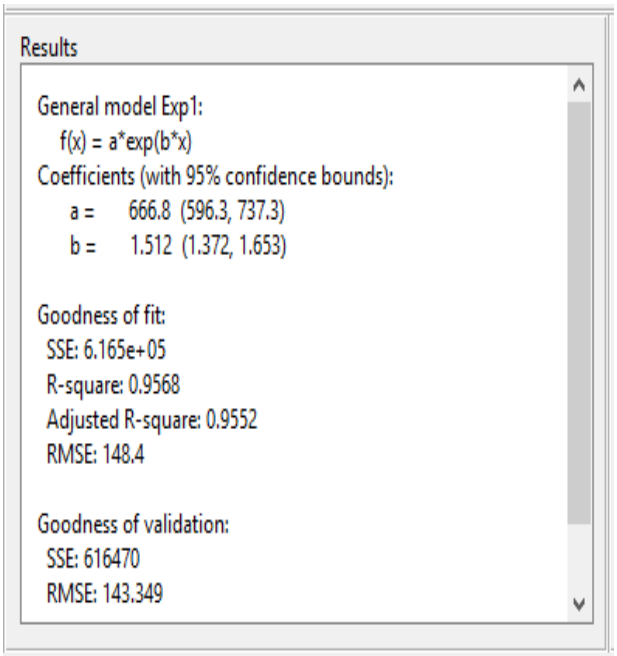


Plate 35b. Chegodayev using Predicted data.

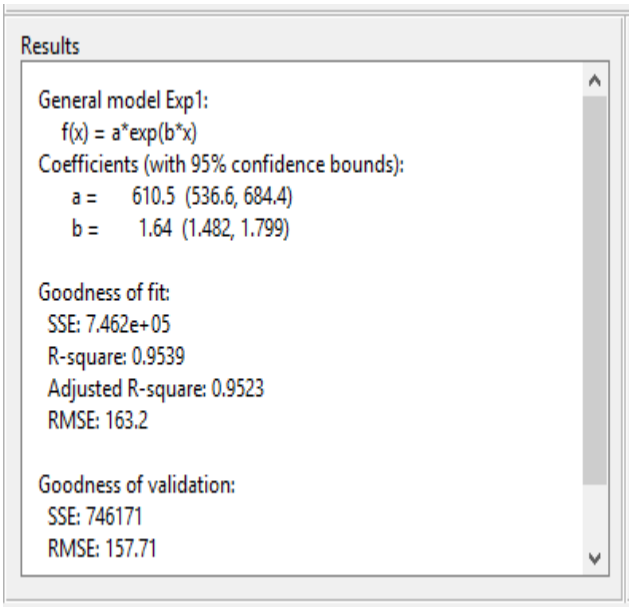


Plate 36a. Turkey using Observed data.

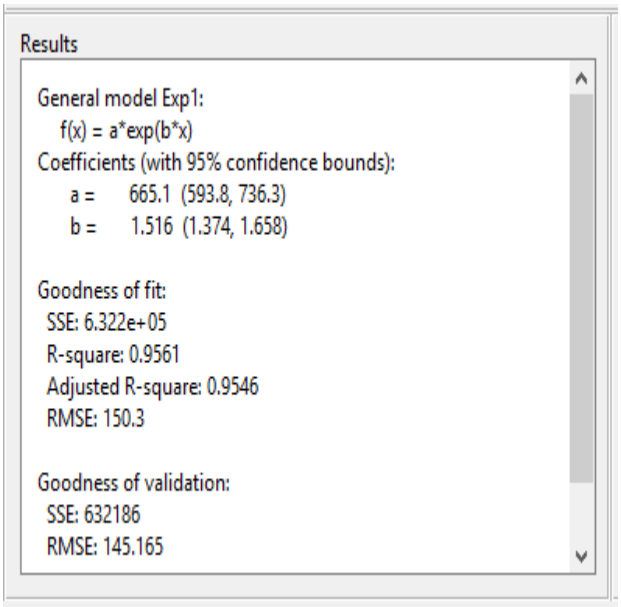


Plate 36b. Turkey using Predicted data.

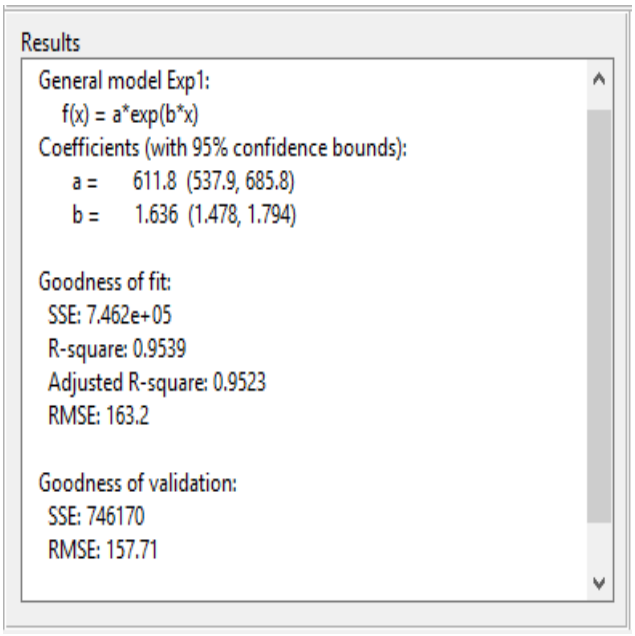


Plate 37a. Blom using Observed data.

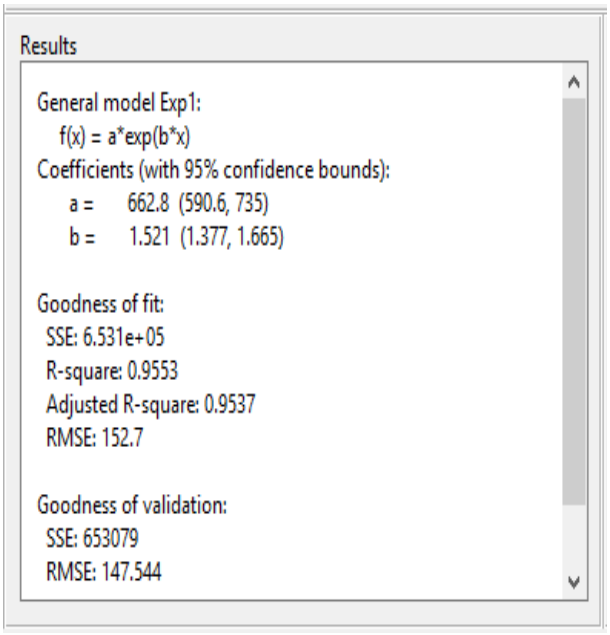


Plate 37b. Blom using Predicted data.

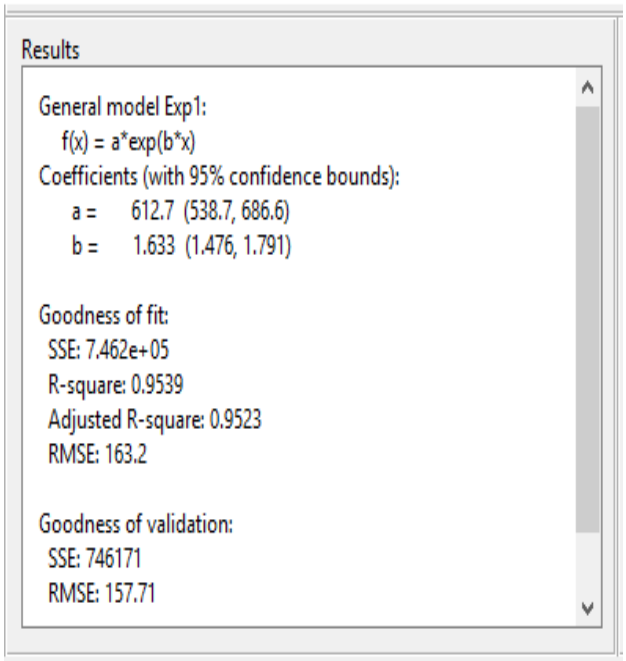


Plate 38a. Cunnane using Observed data.

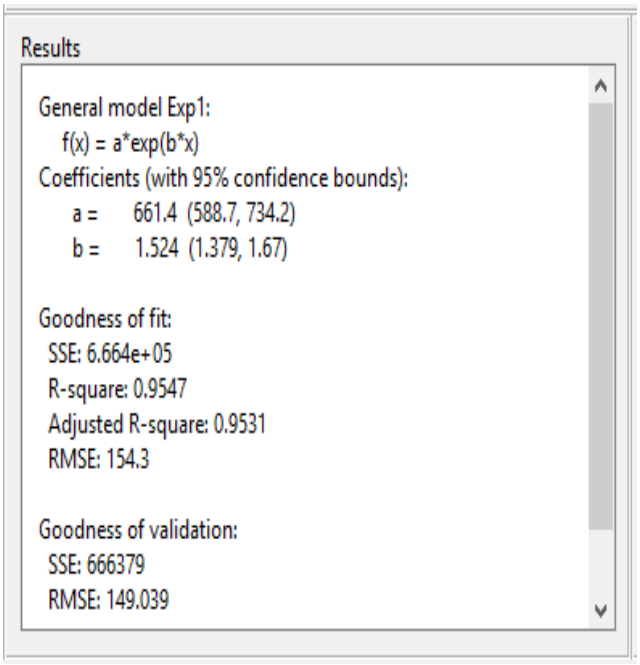


Plate 38b. Cunnane using Predicted data.

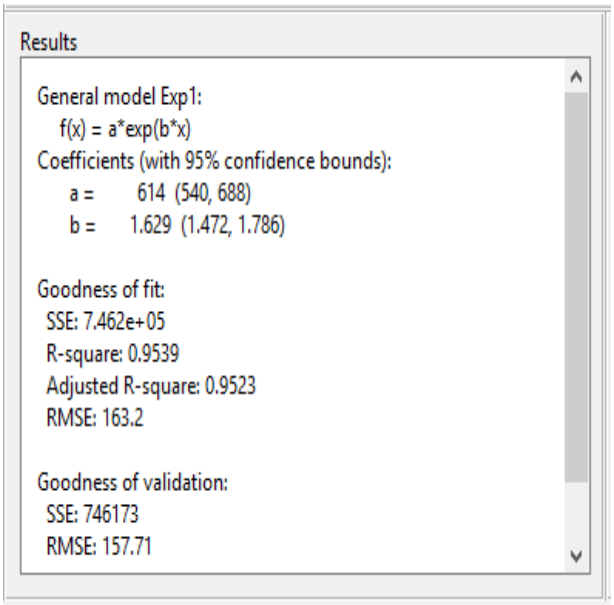


Plate 39a. Gringorten using Observed data.

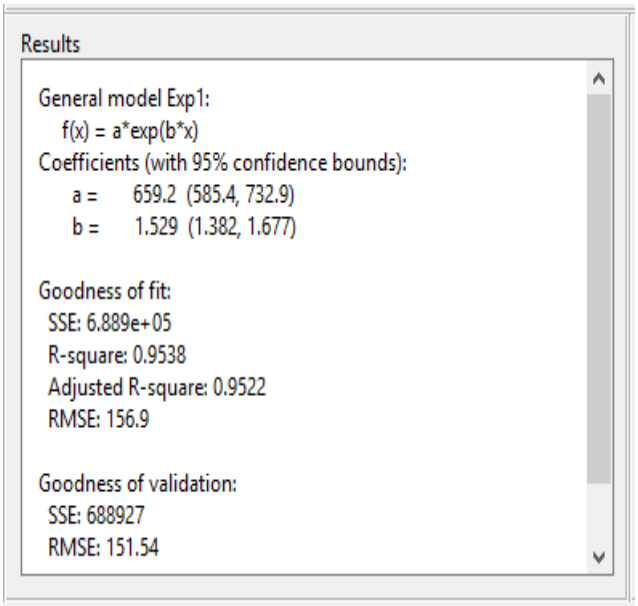


Plate 39b. Gringorten using Predicted data.

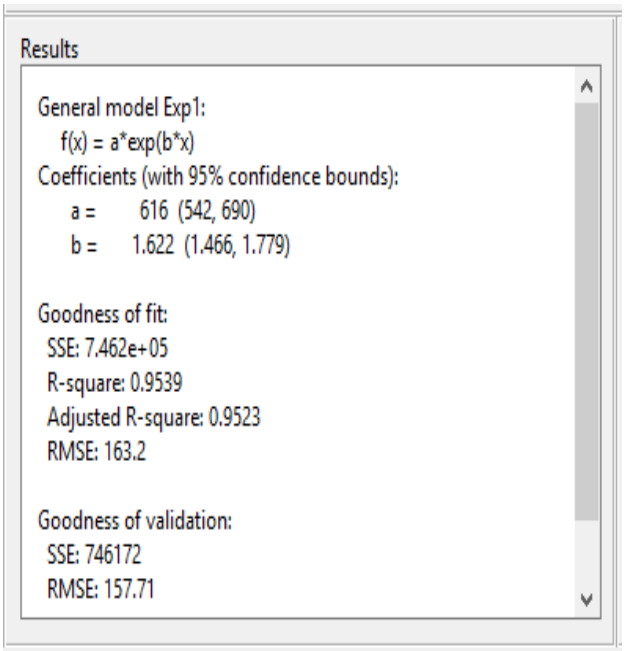


Plate 40a. Hazen using Observed data

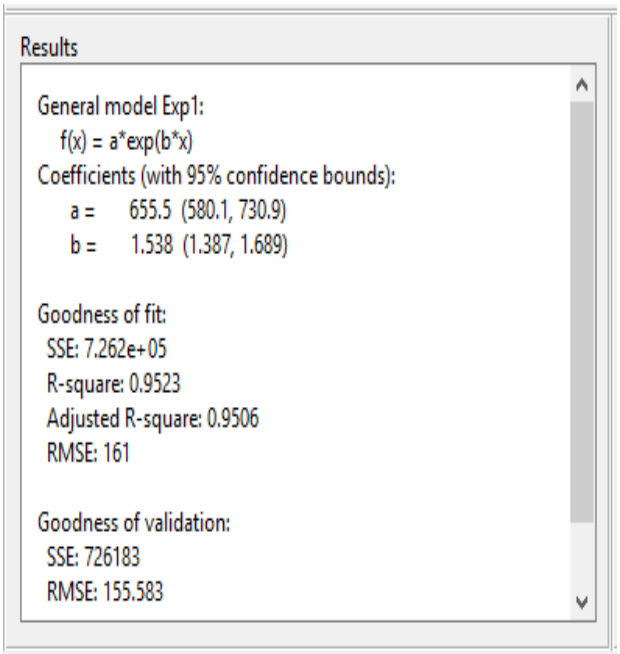


Plate 40b. Hazen using Predicted data.

Appendix XIV. Results of R-Square and RMSE values for Log- Normal Distribution in respect of the below plotting positions.

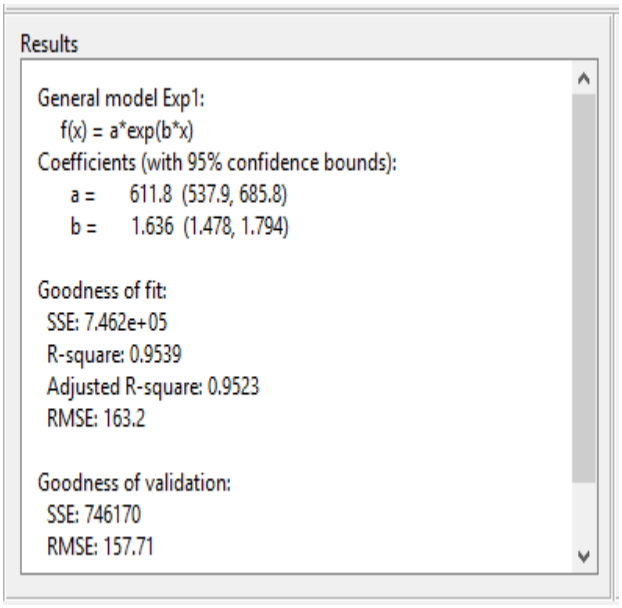


Plate 41a. California using Observed data.

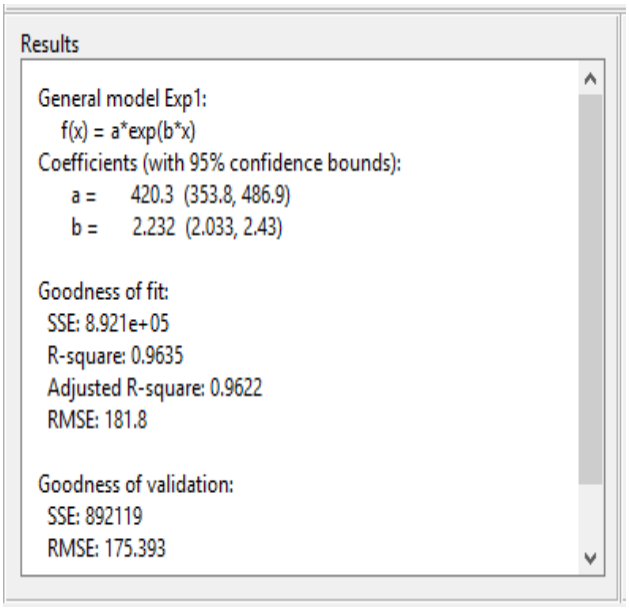


Plate 41b. California using Predicted data.

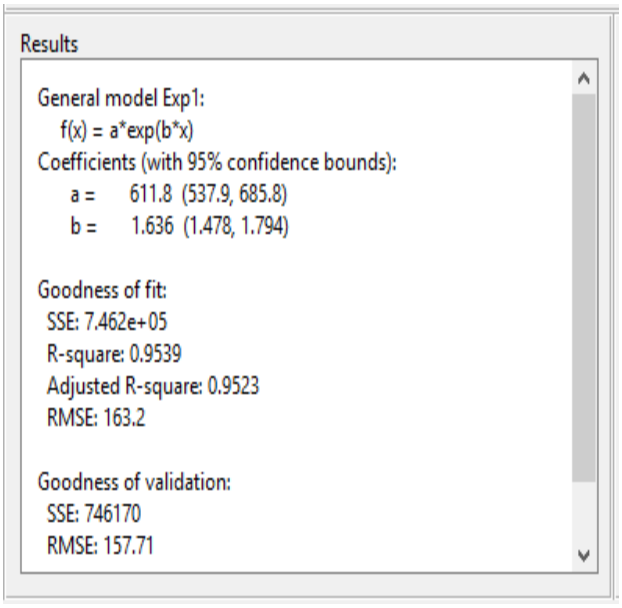


Plate 42a. Weibull using Observed data.

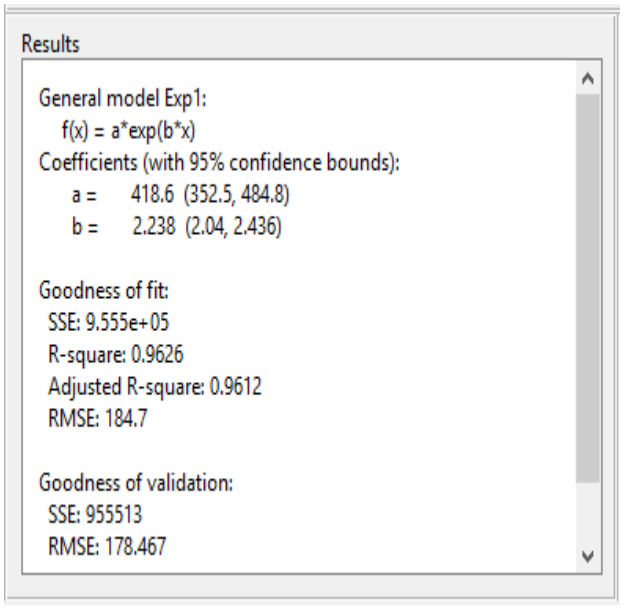


Plate 42b. Weibull using Predicted data.

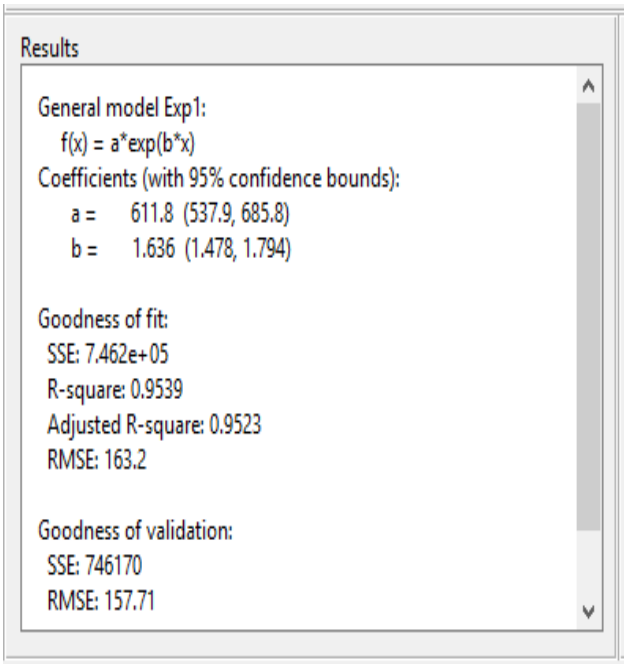


Plate 43a. Chegodayev using Observed data**.**

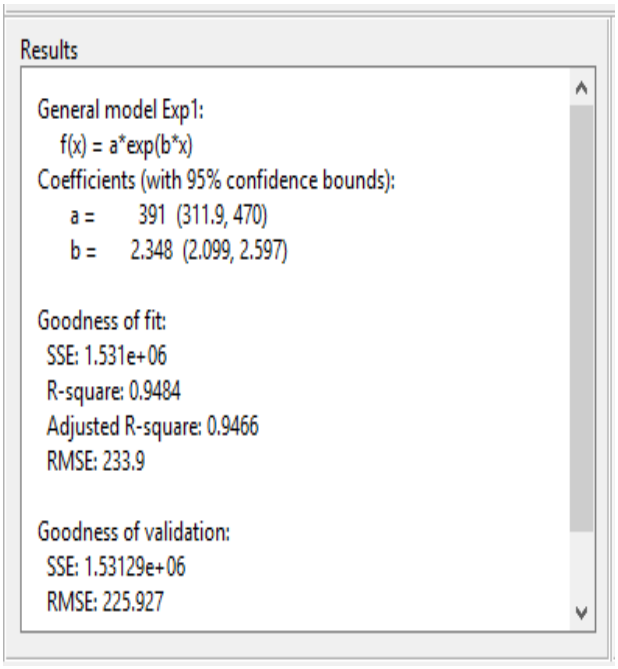


Plate 43b. Chegodayev using Predicted data.

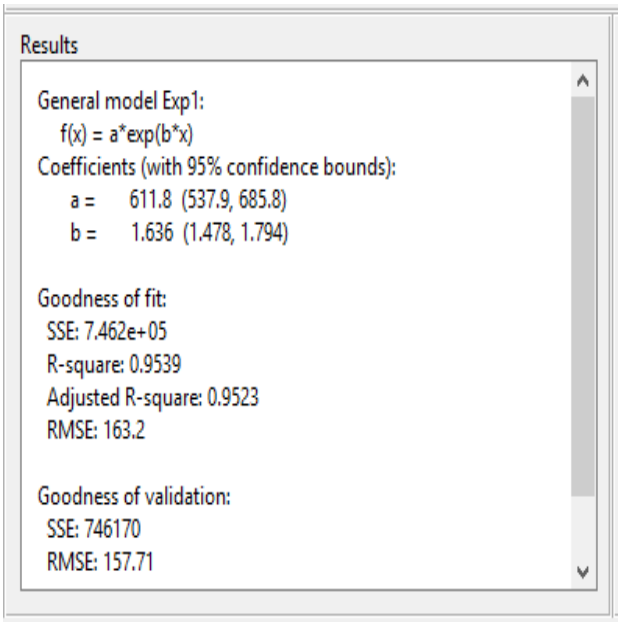


Plate 44a. Turkey using Observed data.

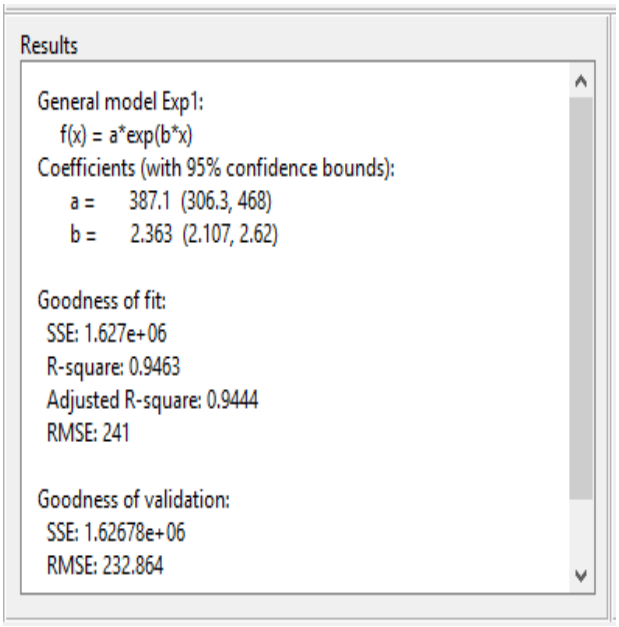


Plate 44b. Turkey using Predicted data.

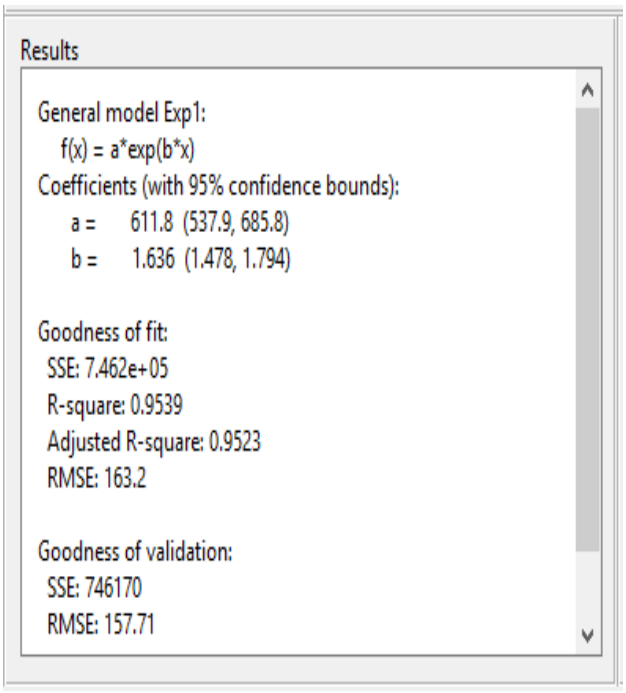


Plate 45a. Blom using Observed data.

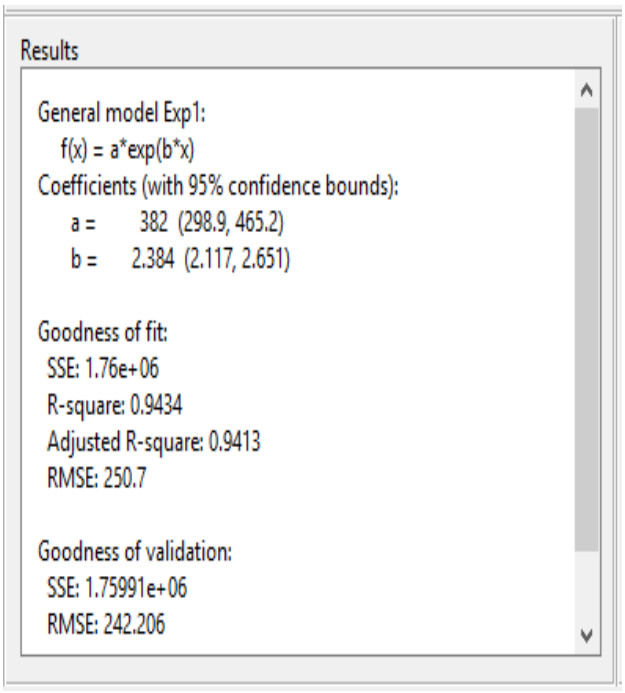


Plate 45b. Blom using Predicted data.

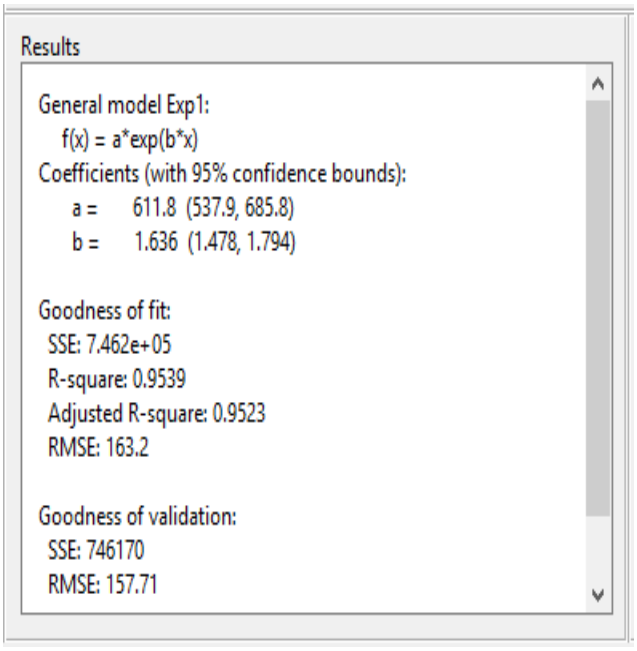


Plate 46a.Cunnane using Observed data

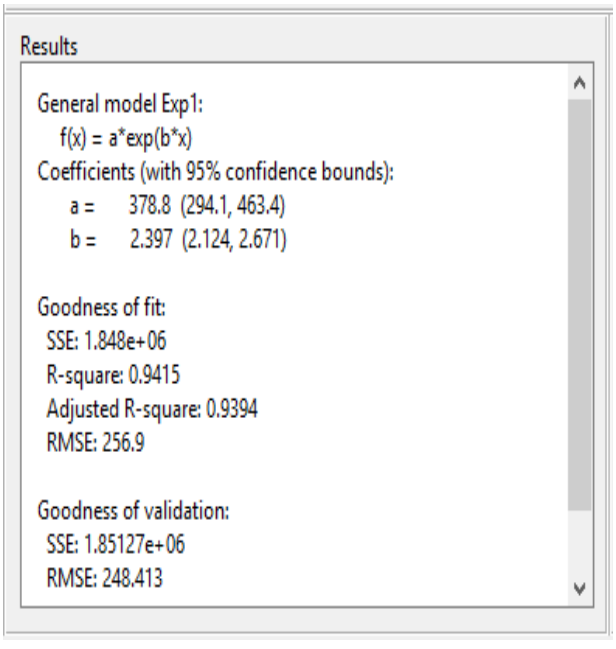
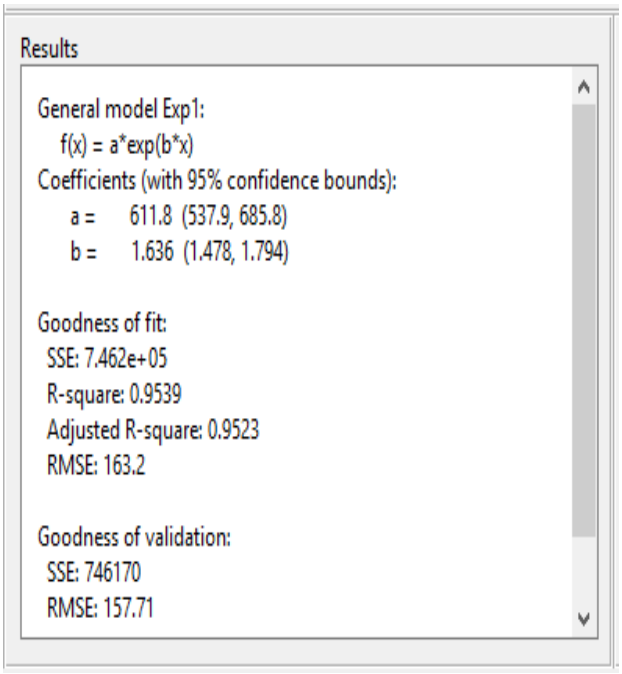


Plate 46b. Cunnane using Predicted data



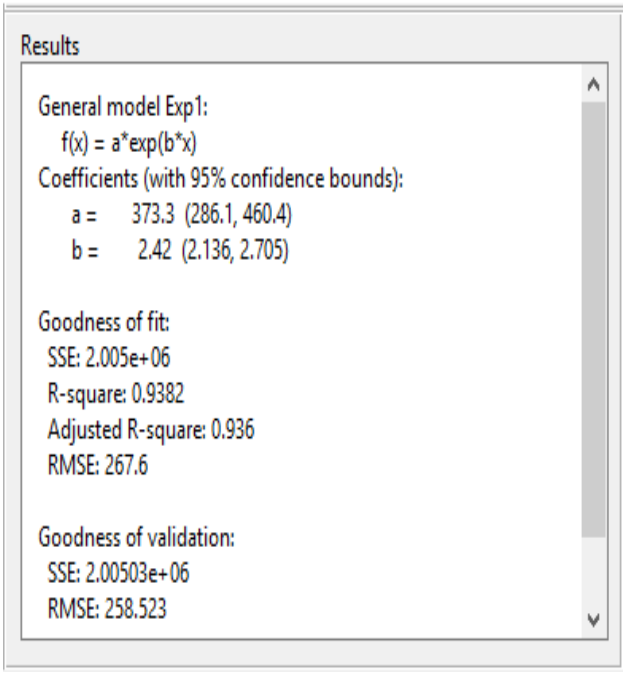
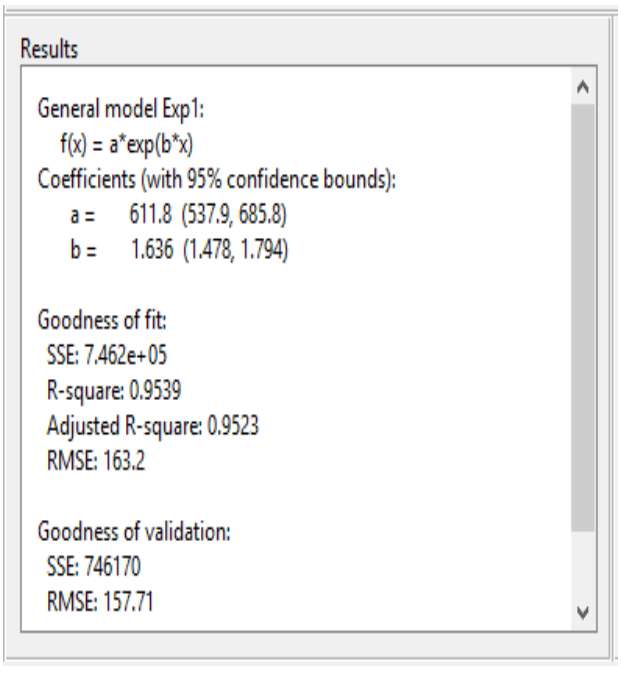


Plate 47a. Gringorten using Observed data**.** **Plate 47b** Gringorten using Predicted data.



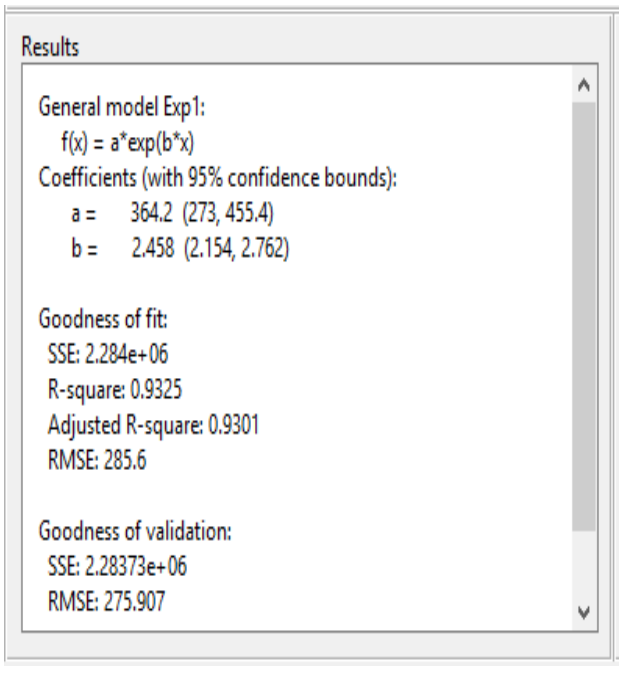


Plate 48a. Hazen using Observed data **Plate 48b**. Hazen using Predicted data.

Appendix XV. Results of R-Square and RMSE values for Pearson Type III Distribution in respect of the below plotting positions.

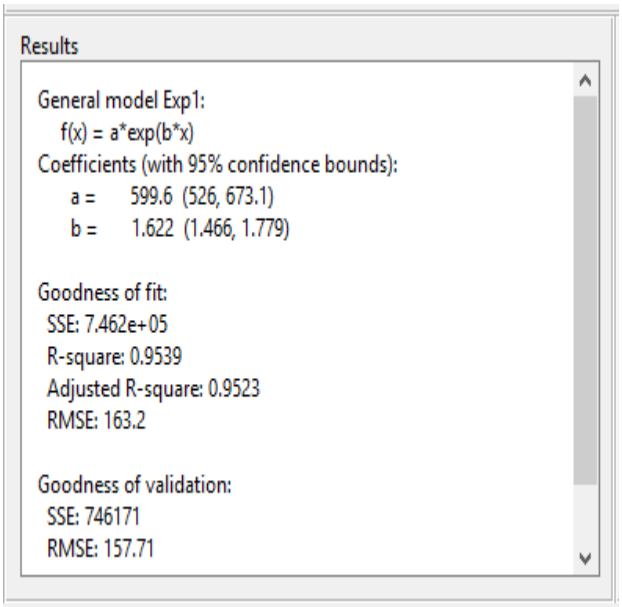


Plate 49a. California using Observed data.

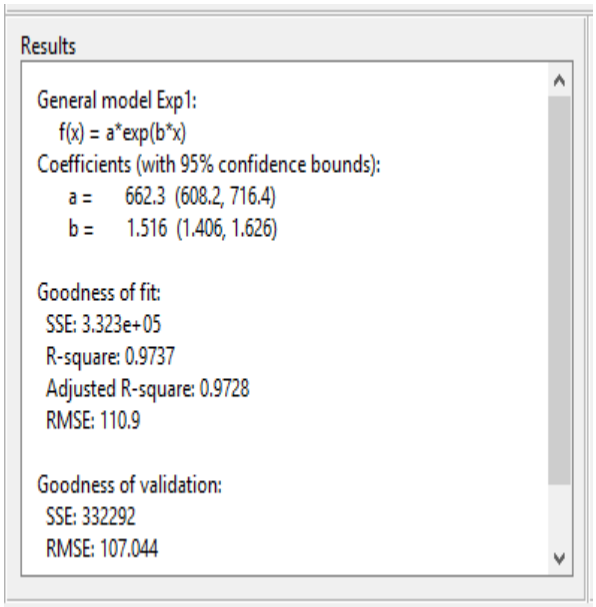


Plate 49b. California using Predicted data.

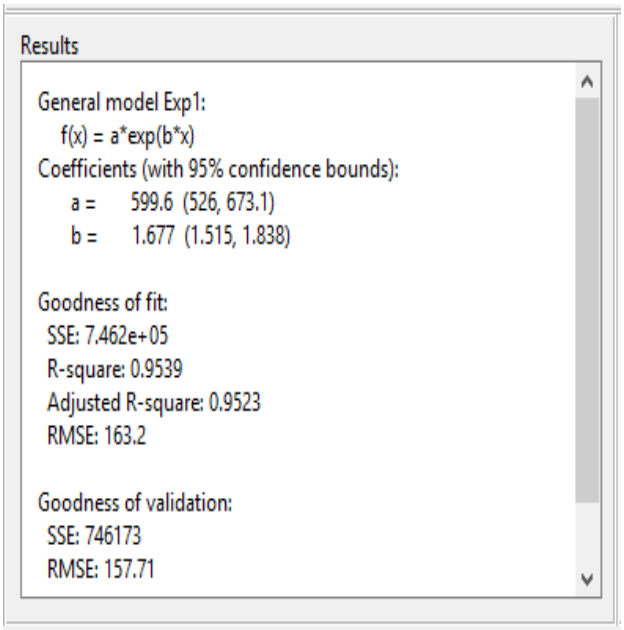


Plate 50a. Weibull using Observed data.

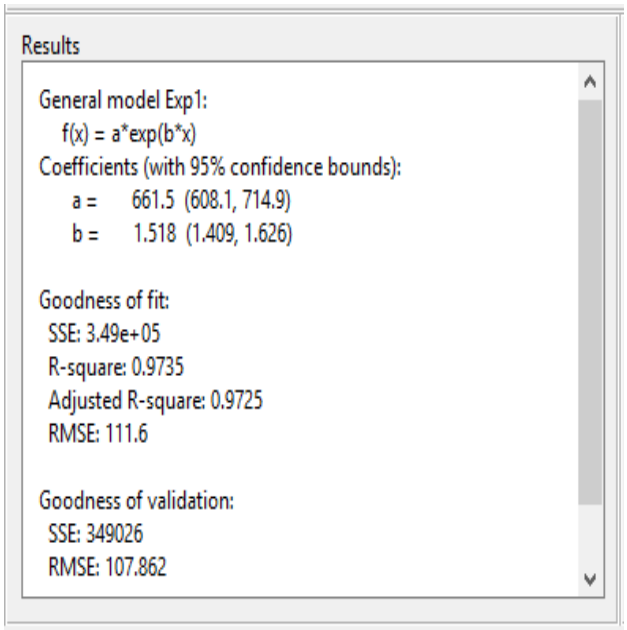


Plate 50b. Weibull using Predicted data.

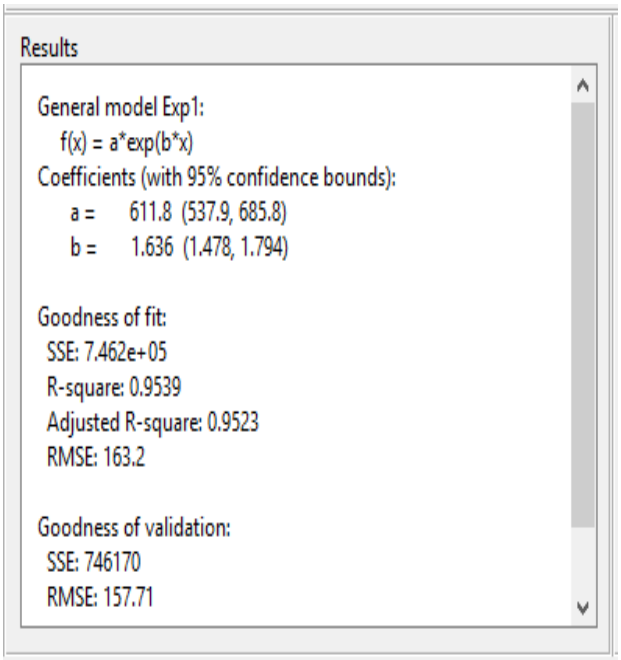


Plate 51a. Chegodayev using Observed data.

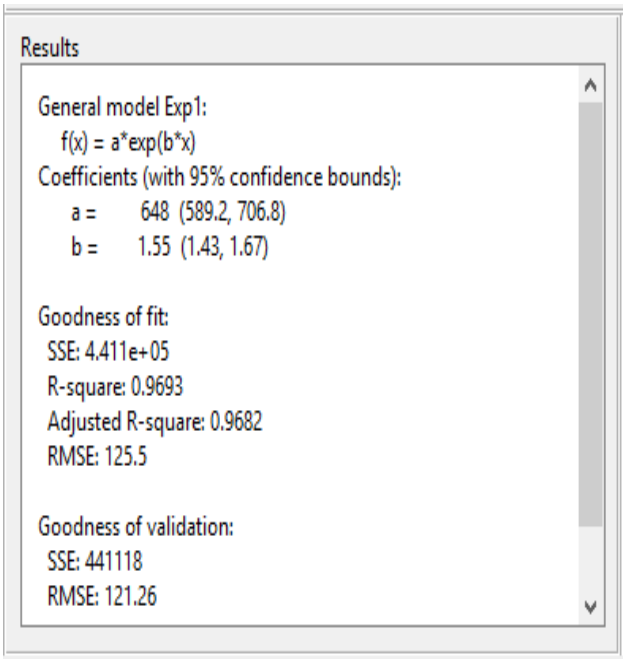


Plate 51b. Chegodayev using Predicted data.

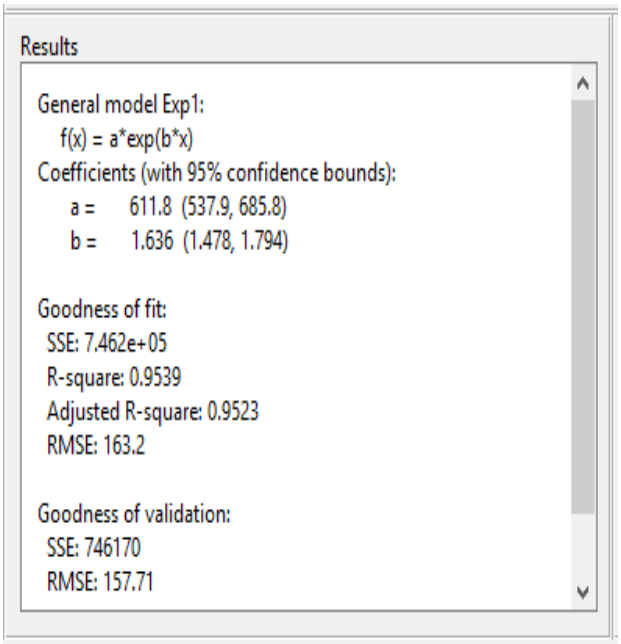


Plate 52a. Turkey using Observed data.

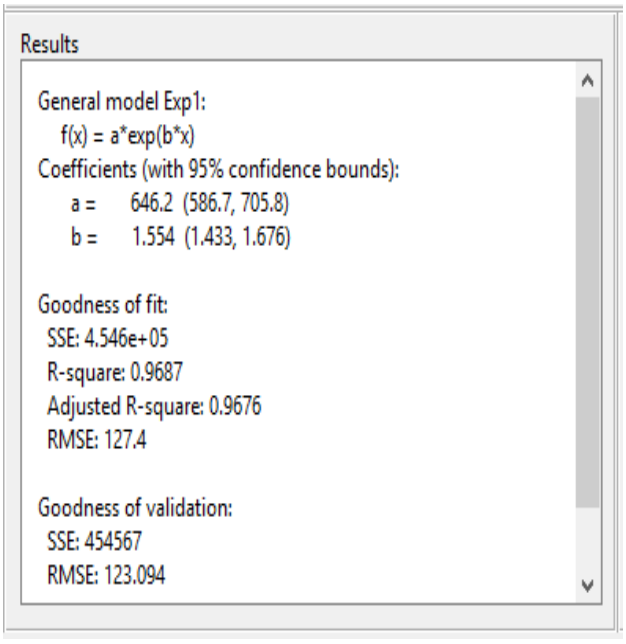


Plate 52b. Turkey using Predicted data.

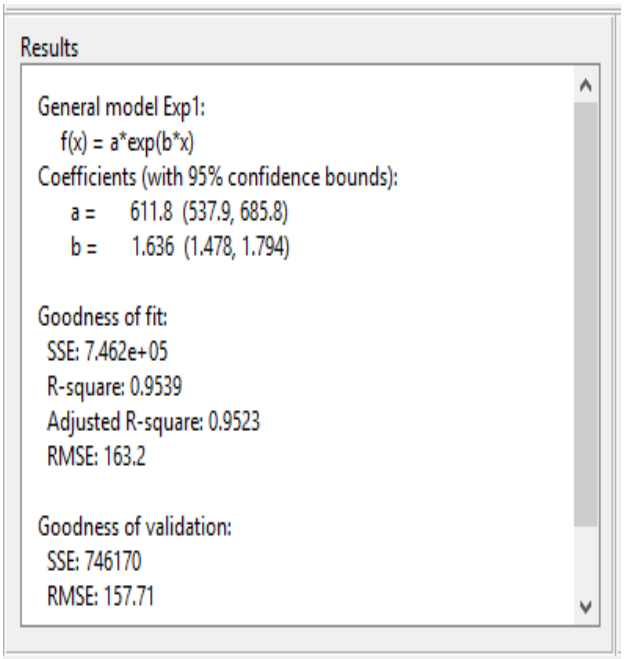


Plate 53a. Blom using Observed data.

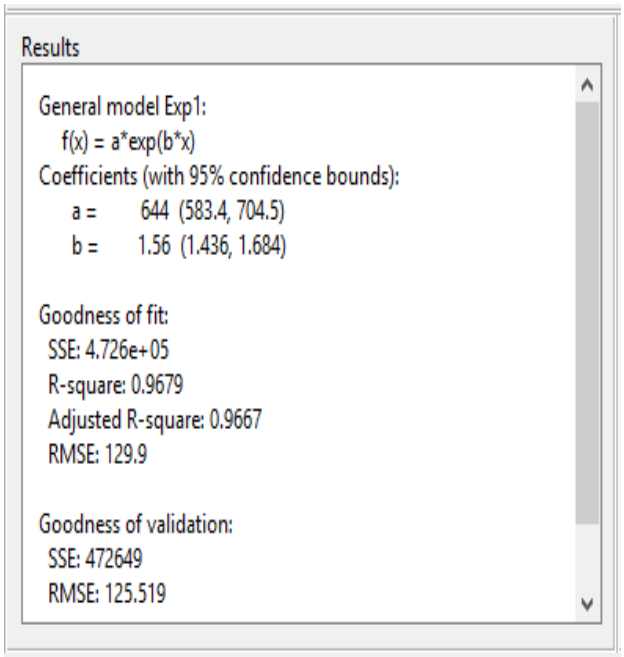


Plate 53b. Blom using Predicted data.

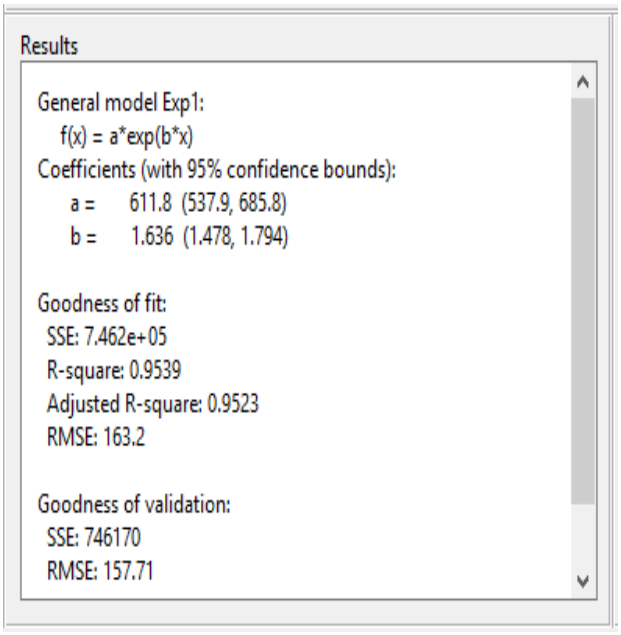


Plate 54a.Cunnane using Observed data.

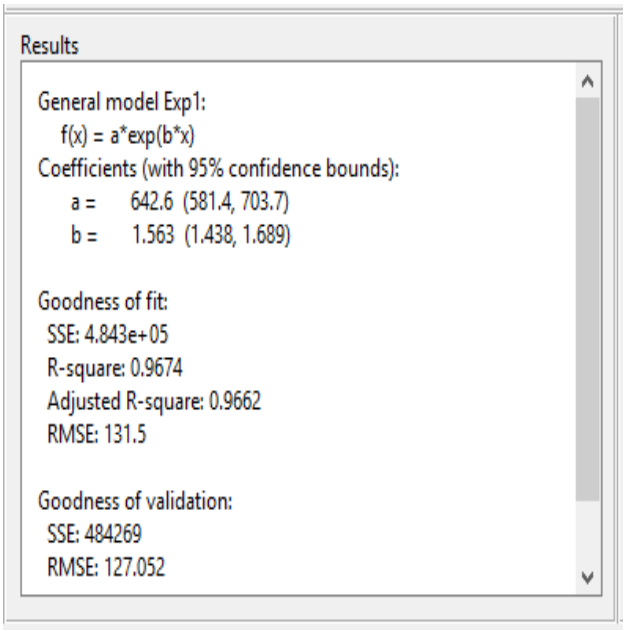


Plate 54b. Cunnane using Predicted data.

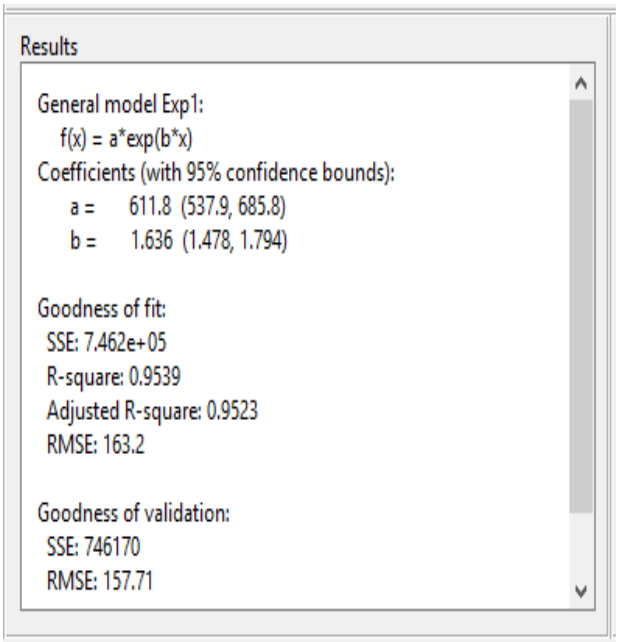


Plate 55a. Gringorten using Observed data.

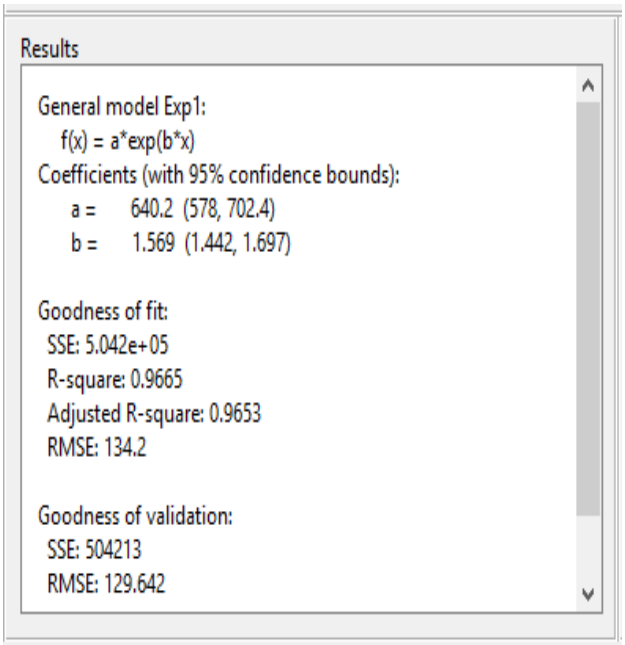


Plate 55b Gringorten using Predicted data.

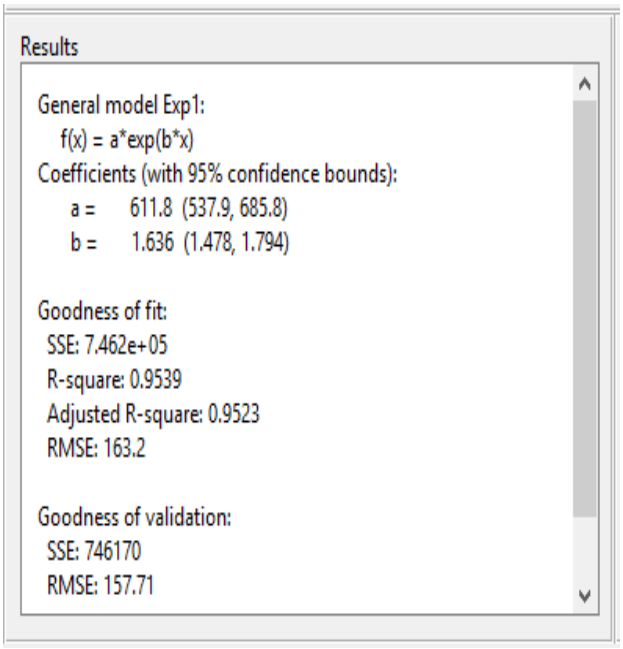


Plate 56a. Hazen using Observed data.

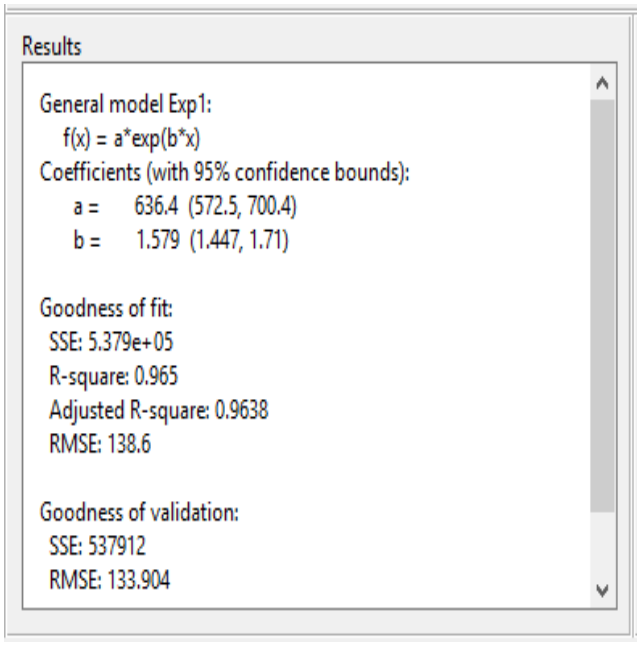


Plate 56b. Hazen using Predicted data.

Appendix XVI. Results of R-Square and RMSE values for Log-Pearson Type III Distribution in respect of the below plotting positions.

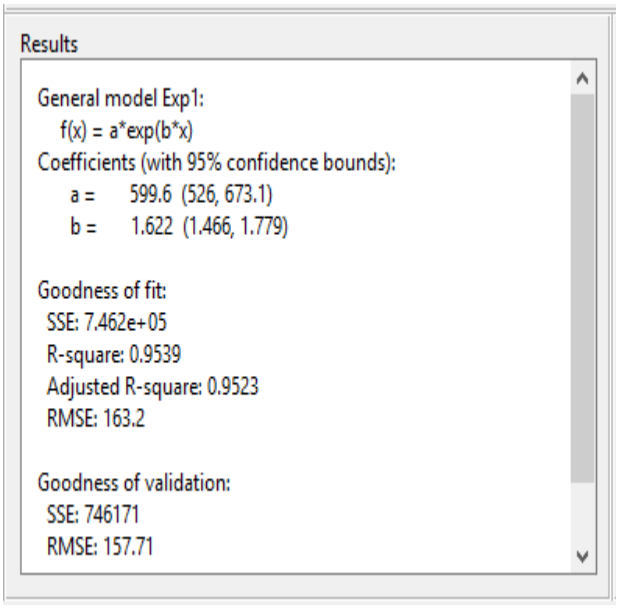


Plate 57a. California using Observed data

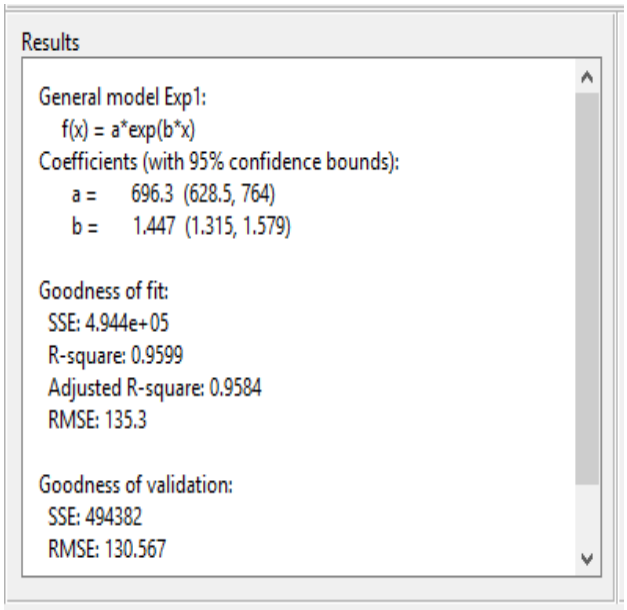


Plate 57b. California using Predicted data.

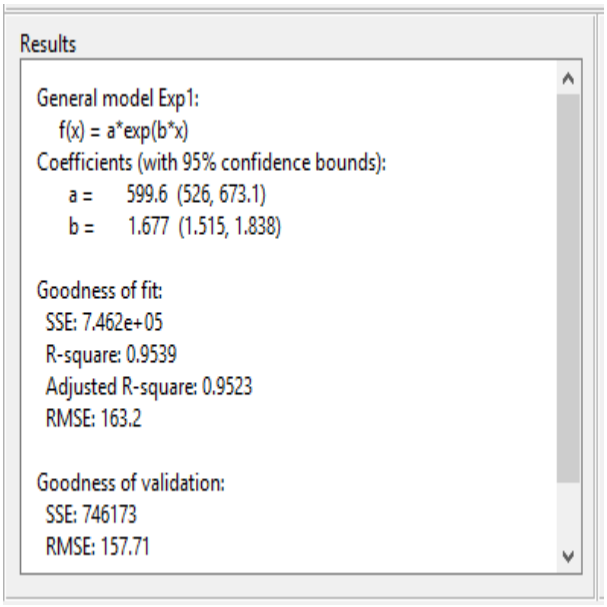


Plate 58a. Weibull using Observed data.

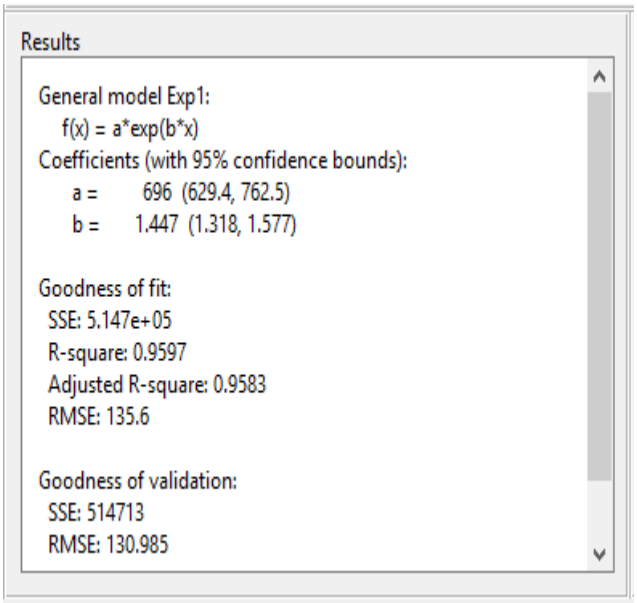


Plate 58b. Weibull using Predicted data.

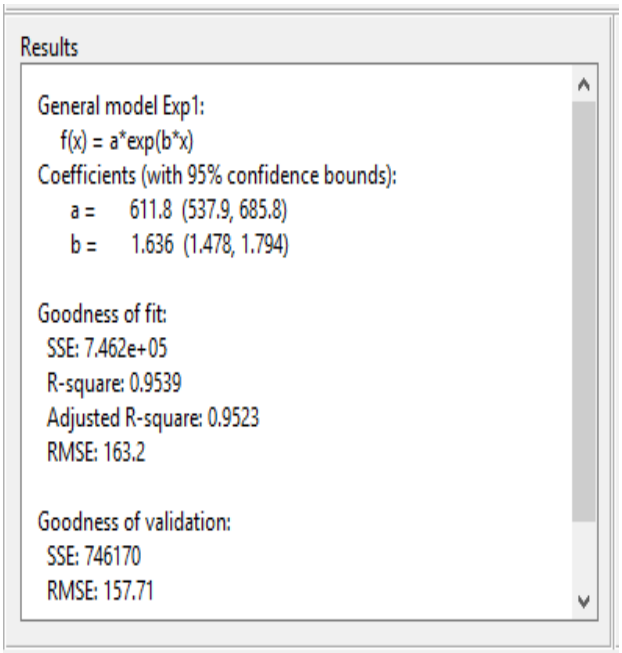


Plate 59a. Chegodayev using Observed data.

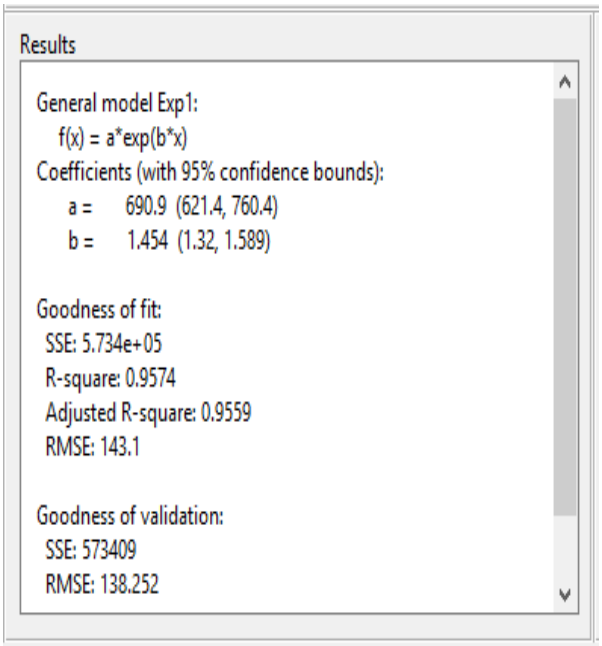


Plate 59b. Chegodayev using Predicted data.

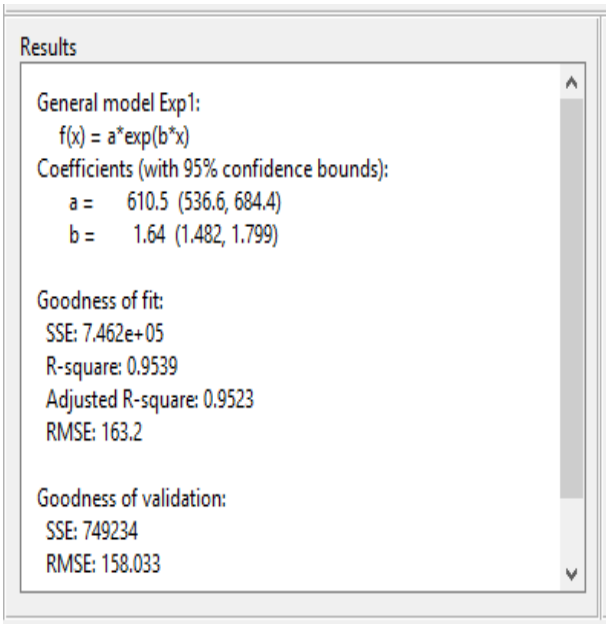


Plate 60a. Turkey using Observed data.

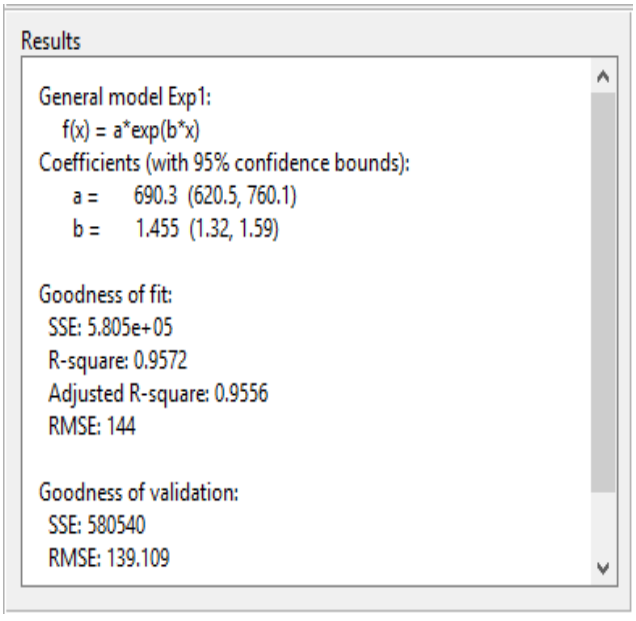


Plate 60b. Turkey using Predicted data.

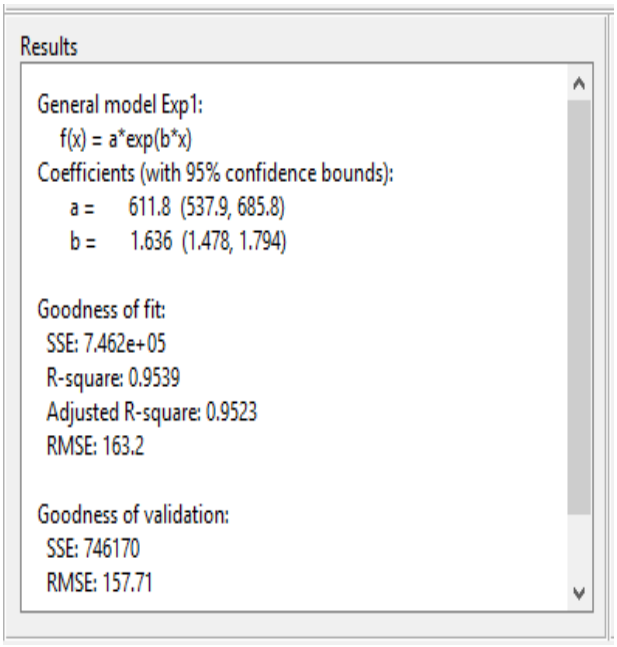


Plate 61a. Blom using Observed data.

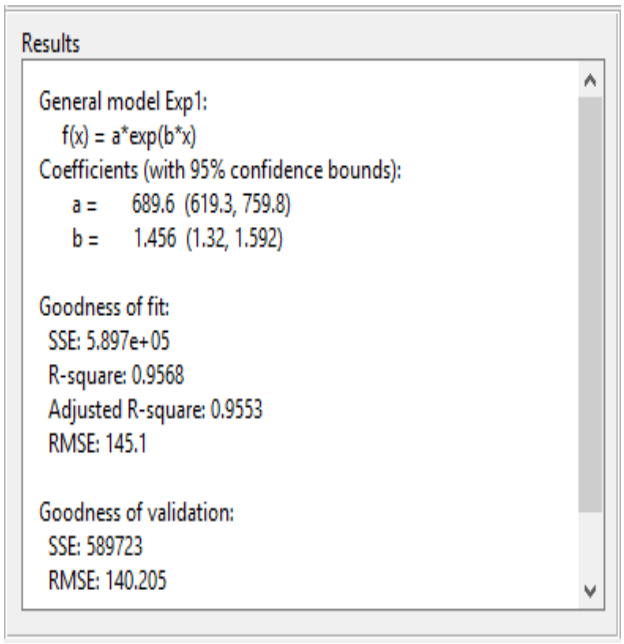


Plate 61b. Blom using Predicted data.

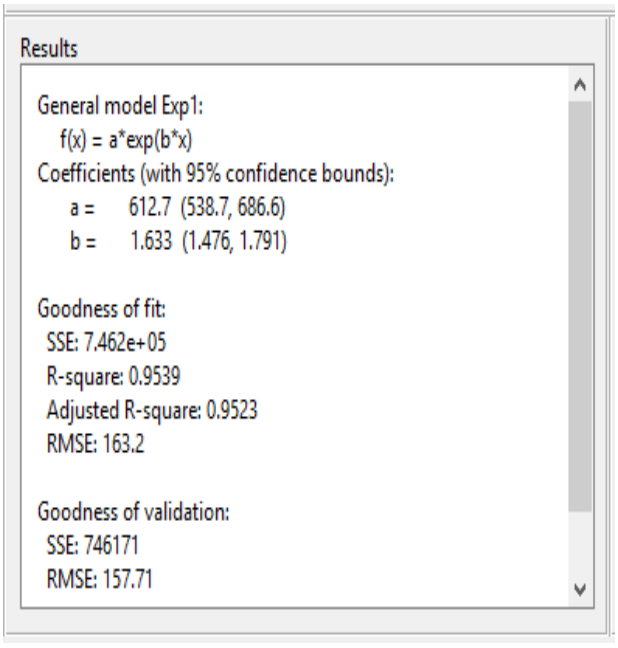


Plate 62a.Cunnane using Observed data

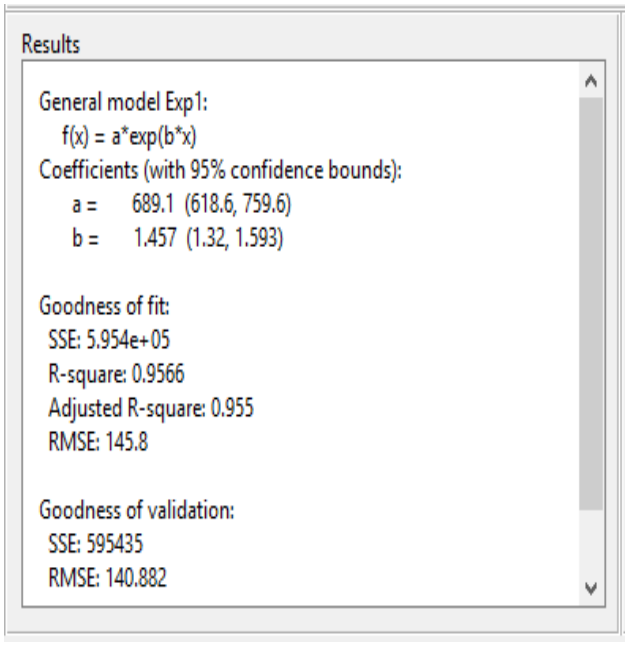


Plate 62b. Cunnane using Predicted data

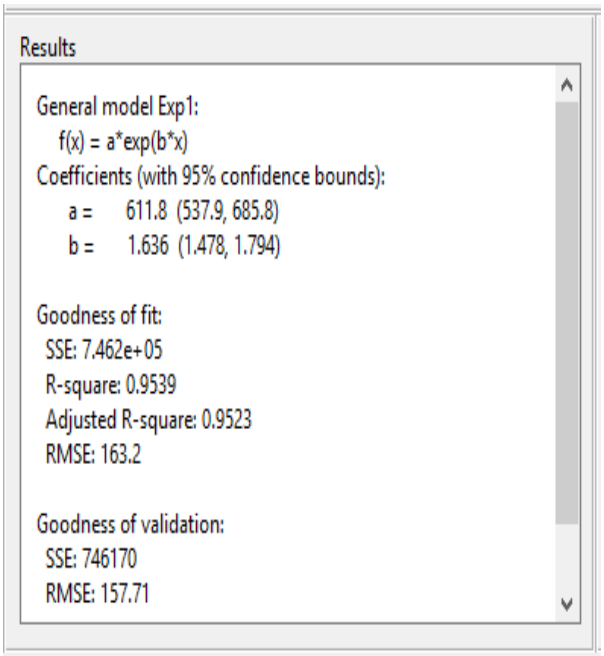


Plate 63a. Gringorten using Observed data.

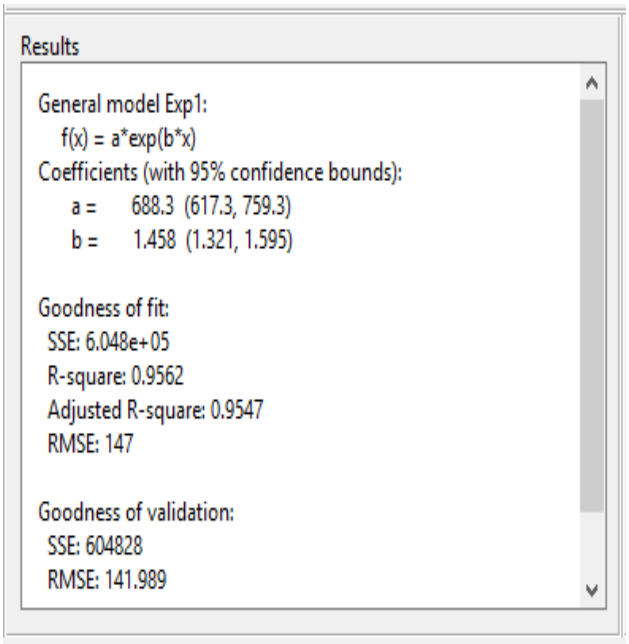


Plate 63b Gringorten using Predicted data.

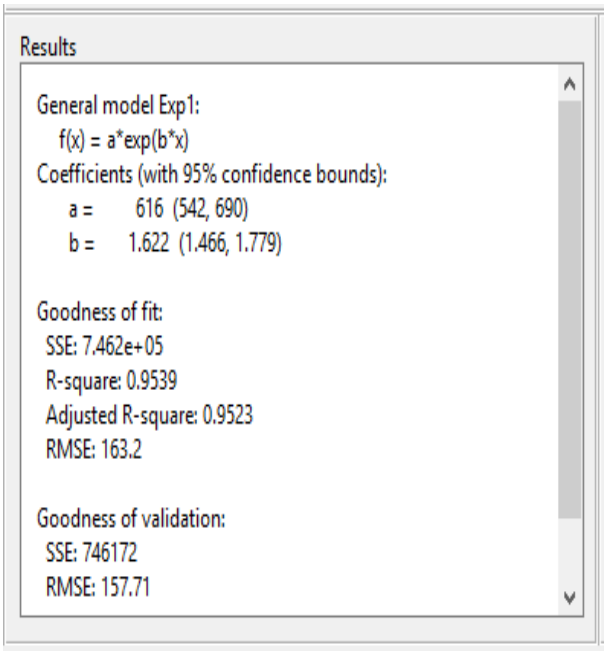


Plate 64a. Hazen using Observed data

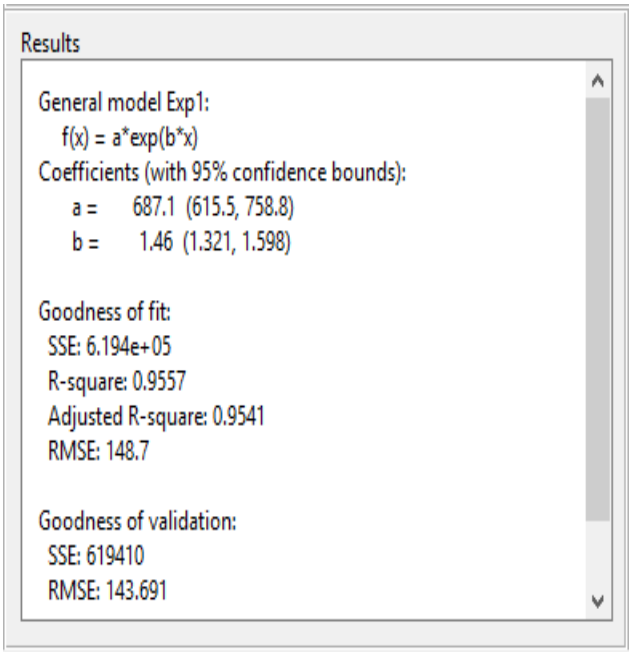


Plate 64b. Hazen using Predicted data.

Appendix XVII Summary for Normal Distribution GOF result

Plotting Position Formulae	Goodness of Fit		Goodness of Validation	
	Adjusted R_Square	RMSE	SSE	RMSE
California	0.9604	133.50	4.81E+05	128.79
Weibull	0.9602	134.20	5.04E+05	129.63
Chegodayev	0.9552	148.40	6.17E+05	143.35
Turkey	0.9546	150.30	6.32E+05	145.17
Blom	0.9537	152.70	6.53E+05	147.54
Cunnane	0.9531	154.30	6.66E+05	149.04
Gringorten	0.9522	156.90	6.89E+05	151.54
Hazen	0.9506	161.00	7.26E+05	155.58

Appendix XVIII Summary for Log-normal Distribution GOF result

Plotting Position Formulae	Goodness of Fit		Goodness of Validation	
	Adjusted R_Square	RMSE	SSE	RMSE
California	0.9622	181.80	8.92E+05	175.39
Weibull	0.9612	184.70	9.56E+05	178.47
Chegodayev	0.9466	233.90	1.53E+06	225.93
Turkey	0.9444	241.00	1.63E+06	232.86
Blom	0.9413	250.70	1.76E+06	242.21
Cunnane	0.9394	256.90	1.85E+06	248.41
Gringorten	0.9360	267.60	2.01E+06	258.52
Hazen	0.9301	285.60	2.28E+06	275.91

Appendix XIX Summary for Pearson Type III Distribution GOF result

Plotting Position Formulae	Goodness of Fit		Goodness of Validation	
	Adjusted R_Square	**RMSE**	**SSE**	**RMSE m^3/s**
California	0.9728	110.90	3.32E+05	107.04
Weibull	0.9725	111.60	3.49E+05	107.86
Chegodayev	0.9682	125.50	4.41E+05	121.26
Turkey	0.9676	127.40	4.55E+05	123.09
Blom	0.9667	129.90	4.73E+05	125.52
Cunnane	0.9662	131.50	4.84E+05	127.05
Gringorten	0.9653	134.20	5.04E+05	129.64
Hazen	0.9638	138.60	5.38E+05	133.90

Appendix XX Summary for Log-Pearson Type III Distribution

Plotting Position Formulae	Goodness of Fit		Goodness of Validation	
	Adjusted R_Square	**RMSE**	**SSE**	**RMSE**
California	0.9584	135.30	4.94E+05	130.57
Weibull	0.9583	135.60	5.15E+05	130.99
Chegodayev	0.9559	143.10	5.73E+05	138.25
Turkey	0.9556	144.00	5.81E+05	139.11
Blom	0.9553	145.10	5.90E+05	140.21
Cunnane	0.9550	145.80	5.95E+05	140.88
Gringorten	0.9547	147.00	6.05E+05	141.99
Hazen	0.9541	148.70	6.19E+05	143.69

Appendix XXI Summary of GOF for the Observed Data

Plotting Position Formulae	Goodness of Fit		Goodness of Validation	
	Adjusted R_Square	**RMSE**	**SSE**	**RMSE**
California	0.9523	163.20	7.46E+05	157.71
Weibull	0.9523	163.20	7.46E+05	157.71
Chegodayev	0.9523	163.20	7.46E+05	157.71
Turkey	0.9523	163.20	7.46E+05	157.71
Blom	0.9523	163.20	7.46E+05	157.71
Cunnane	0.9523	163.20	7.46E+05	157.71
Gringorten	0.9523	163.20	7.46E+05	157.71
Hazen	0.9523	163.20	7.46E+05	157.71

Appendix XXII Summary of the Adjusted R_square values.

Type of Distribution	California (Modified)	Weibull	Chegodayev	Turkey	Blom	Cunnane	Gringorten	Hazen	observed
Pearson type III	0.9728	0.9725	0.9682	0.9676	0.9667	0.9662	0.9653	0.9638	0.9523
Log-Pearson type III	0.9584	0.9583	0.9559	0.9556	0.9553	0.9550	0.9547	0.9541	0.9523
Normal	0.9604	0.9602	0.9552	0.9546	0.9537	0.9531	0.9522	0.9506	0.9523
Lognormal	0.9622	0.9612	0.9466	0.9444	0.9413	0.9394	0.9360	0.9301	0.9523

Appendix XXIII Summary of Validated RMSE values.

Type of Distribution	California (Modified)	Weibull	Chegodayev	Turkey	Blom	Cunnane	Gringorten	Hazen	observed
Pearson type III	107.04	107.86	121.26	123.09	125.52	127.05	129.64	133.9	157.71
Log-Pearson type III	130.57	130.99	138.25	139.11	140.21	140.88	141.99	143.69	157.71
Normal	128.79	129.63	143.35	145.17	147.54	149.04	151.54	155.58	157.71
Lognormal	175.39	178.47	225.93	232.86	242.21	248.41	258.52	275.91	157.71