Analyzing L-moments, Maximum Likelihood, and Maximum Product of Spacings Estimation Methods for Modeling Climate Extremes

By

Fokia Aamer

(Registration No: 00000-359911)

Department of Sciences

School of Interdisciplinary Engineering & Sciences

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

(August 2024)

Analyzing L-moments, Maximum Likelihood, and Maximum Product of Spacings Estimation Methods for Modeling Climate Extremes

By

Fokia Aamer

(Registration No: 00000-359911)

A thesis submitted to the National University of Sciences and Technology, Islamabad,

in partial fulfillment of the requirements for the degree of

Masters of Science in

Computational Science

and Engineering

Supervisor: Dr. Zamir Hussain

School of Interdisciplinary Engineering & Sciences

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

(August 2024)

Annex A to NUST Letter No.
0972/102/Exams/Thesis-Cert
dated 23 Dec 16.

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by Ms. Fokia Aamer Registration No. 00000359911 of **SINES** has been vetted by undersigned, found complete in all aspects as per NUST Statutes/Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature with stamp: Associate Professor SINES - NUST, Sector H-12 Islamabad

Name of Supervisor: Dr. ZAMIR HUSSAIN

Date: NOV 05 2024

Signature of HoD with stamp: Mian Ilyas Ahmad ... Engineering ...fessor ...NES - NUST, Sector H-12 ...slamabad

Date: 05 Nov 2024

Countersign by

Signature (Dean/Principal): Dr. SYED IRTIZA ALI SHAH Principal & Dean SINES - NUST, Sector H-12 Islamabad

Date: 6/11/24

FORM TH-4

MASTER THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by: Fokia Aamer (00000359911)

Titled: Analyzing L-Moments, Maximum Likelihood and Maximum Product of Spacings Estimation Methods for Modeling Climate Extremes be accepted in partial fulfillment of the requirements for the award of Masters of Science in Computational Science and Engineering degree with (_______Grade).

Examination Committee Members

1. **Name:** Dr. Salma Sherbaz **Signature:** ______________

2. **Name:** Dr. Absaar-ul-Jabbar **Signature:** _____ -------------

Supervisor's name: Dr. Zamir Hussain

Signature: _______________

Date: ___________________

Head of Department **Date**

COUNTERSINGED

Date: __________ **Dean/Principal**

CERTIFICATE OF APPROVAL

This is to certify that the research work presented in this thesis, entitled "Analyzing L-Moments, Maximum Likelihood and Maximum Product of Spacings for Modeling Climate Extremes" was conducted by Ms. Fokia Aamer under the supervision of Dr. Zamir Hussain.

No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted to the SINES in partial fulfillment of the requirements for the degree of Master of Science in Field of CSE.

Department of Science, SINES, National University of Sciences and Technology, Islamabad.

Student Name: Fokia Aamer **Signature:** _________

Examination Committee:

a) **Examiner 1:** Dr. Salma Sherbaz **Signature:** _________
Assistant Professor, SINES

b) **Examiner 2:** Dr. Absaar-ul-Jabbar **Signature:**
Assistant Professor, SINES

c) **External Examiner:** **Signature:** _________

Supervisor: Dr. Zamir Hussain **Signature:** _________

Dean: **Signature:** _________

AUTHOR'S DECLARATION

I Fokia Aamer hereby state that my MS thesis titled "Analyzing L-Moments, Maximum Likelihood and Maximum Product of Spacings for Modeling Climate Extremes" is my own work and has not been submitted previously by me for taking any degree from National University of Sciences and Technology, Islamabad or anywhere else in the country/ world.

At any time if my statement is found to be incorrect even after I graduate, the university has the right to withdraw my MS degree.

Name of Student: Fokia Aamer

PLAGIARISM UNDERTAKING

I solemnly declare that research work presented in the thesis titled "Analyzing L-Moments, Maximum Likelihood and Maximum Product of Spacings for Modeling Climate Extremes" is solely my research work with no significant contribution from any other person. Small contribution/ help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero-tolerance policy of the HEC and National University of Sciences and Technology (NUST), Islamabad towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS degree, the University reserves the rights to withdraw/revoke my MS degree and that HEC and NUST, Islamabad has the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized thesis.

Student Signature: Fokia

Name: Fokia Aamer

DEDICATION

With heartfelt gratitude, this research is a tribute

to my parents, the architects of my dreams,

and my teachers, the guiding stars on this academic journey.

ACKNOWLEDGEMENTS

My profound gratitude is directed towards Almighty Allah for the divine blessings that sustained me throughout the demanding yet fulfilling pursuit of this research. His unwavering support empowered me to navigate challenges and ultimately achieve my desired goals, resulting in the successful completion of this project.

I extend my deepest gratefulness to my esteemed supervisor, Dr. Zamir Hussain, whose support, insightful guidance, and profound knowledge were helpful in the successful completion of this research. His constant motivation, enthusiasm, and patience were indispensable throughout the entire process. Undoubtedly, his dedication to research serves as a perpetual source of inspiration for me.

Furthermore, I acknowledge the pivotal role played by the School of Interdisciplinary Engineering and Sciences (SINES), NUST, in providing the necessary research infrastructure and enabling environment for this project. I am deeply indebted to the members of my Guidance and Examination Committee (GEC), Dr. Salma Sherbaz and Dr. Absaar-ul-Jabbar whose expert guidance, critiques, and support proved invaluable throughout the research process.

TABLE OF CONTENTS

LIST OF TABLES

Page No.

LIST OF FIGURES

Page No.

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS

PE3	Pearson Type-3
LM	L-Moments
MLE	Maximum Likelihood Estimation
MPS	Maximum Product of Spacings
RMSE	Root Mean Square Error
AMRS	Annual Maximum Rainfall Series
Mu (μ)	Location
Sigma (σ)	Scale
Gamma (γ)	Shape
°F	Fahrenheit
°C	Celsius
°N	North
°E	East

ABSTRACT

Choice of model and estimation method plays a vital role in dealing extremes, especially when the interest is in the extreme upper quantiles. Pearson Type-3 (PE3) probability distribution is frequently used due to its proficiency in effectively accommodating asymmetrical and non-normal data distributions. The location (μ), shape (σ), and scale (γ) parameters of PE3 play a crucial role in determining and controlling skewness, spread, and position of the distribution respectively. The objective of this study is to assess and contrast the impacts of three parameter estimation methods L-moments (LM), Maximum Likelihood Estimation (MLE), and Maximum Product of Spacing (MPS) for the Pearson Type-3 (PE3) distribution. Therefore, empirical analyses are conducted using real world data with diverse degrees of skewness and moderate sample size to compare the estimated parameters of each method against the true parameters. The data being considered is the series of annual maximum rainfall, obtained from 16 different meteorological observatories in zone D and E of Pakistan. This study underscores the importance of selecting an appropriate estimation method tailored to the data's characteristics, highlighting that no single method excels in all scenarios. In conclusion, MPS method is particularly effective for datasets exhibiting severe skewness and kurtosis, allowing it to handle extreme distributions well. In contrast, the LM method performs effectively with datasets that show mild skewness and kurtosis. However, the MLE method struggles when applied to skewed datasets, making it less suitable for such distributions. These insights are crucial for improving the reliability of extreme event modeling and providing a stronger foundation for accurate method selection.

Keywords: Climate Extremes; Annual Maximum Rainfall; Maximum Product of Spacings method; Pearson Type-3 Distribution; Root Mean Square Error; Bias.

1. INTRODUCTION

Climate extremes are becoming more frequent and intense due to climate change. A climate extreme refers to an exceptionally uncommon and severe weather event that significantly deviates from the typical or anticipated conditions in a specific geographical area. It is characterized by its rarity and intensity, as it represents an extreme departure from the average weather patterns. Climate extremes can manifest as extreme temperatures (such as heatwaves or cold snaps), intense storms (such as hurricanes or blizzards), heavy rainfall leading to floods, prolonged droughts, or other extraordinary weather phenomena. These events are notable for their severity and often require special attention in terms of preparation, adaptation, and response strategies.

1.1 Extreme Values

Extreme values frequently appear within datasets of extreme events during rare and intense climatic phenomena. These values typically manifest as outliers within datasets, demonstrating a pronounced departure from the statistical norms of a given variable. Moreover, extreme values frequently display asymmetry in their distribution, with longer tails towards the higher end of the scale, indicating a higher likelihood of extreme values exceeding a certain threshold. Furthermore, extreme events data often reveal dependencies and correlations between different variables, highlighting the interconnected nature of climatic systems and the potential for cascading impacts across multiple sectors. Identifying the characteristics of extreme values is crucial for assessing the risks associated with future extreme events.

Understanding and describing the patterns, trends, and characteristics of a single variable using statistical techniques like descriptive statistics and historical climate data is fundamental in climate research and decision-making processes. Through this analysis, researchers can discern central tendencies, variations, and distributions of the variable over time. These insights help detect long-term trends like gradual temperature increases or shifts in precipitation patterns, offering valuable understanding of climate evolution. Statistical techniques also uncover temporal and spatial patterns, revealing seasonal fluctuations, periodic oscillations, and spatial gradients. Additionally, characterizing statistical properties enhances predictive capabilities, facilitating accurate climate modeling and scenario projections.

1.2 Distribution Fitting

Probabilistic distribution fitting, particularly the application of the Pearson Type-3 distribution, plays a pivotal role in analyzing extreme events due to its ability to effectively model skewed data. In the realm of extreme events analysis, where occurrences often exhibit skewed distributions with tails extending towards higher values, the flexibility of the Pearson Type-3 distribution becomes highly advantageous. The Pearson Type-3 distribution's versatility in accommodating different degrees of skewness and its applicability across various fields, including hydrology, climatology, and engineering, further underscore its significance. Moreover, the availability of robust parameter estimation techniques enhances its practical utility, enabling reliable analysis even with limited data. As extreme events continue to pose significant challenges exacerbated by climate change, leveraging appropriate probabilistic models like the Pearson Type-3 distribution remains essential for understanding, predicting, and mitigating associated risks.

1.3 Estimation Methods

Parameter estimation methods play a critical role in fitting probability distributions to empirical data, particularly in hydrology, climatology, and other fields dealing with extreme events. This thesis seeks to evaluate following three widely used parameter estimation methods across a spectrum of settings and varying data conditions:

a) L-moments (LM),
b) Maximum Likelihood Estimation (MLE)
c) Maximum Product of Spacings (MPS)

The selection of a method hinges on various factors including data characteristics, analysis objectives, computational complexity, sample size requirements, and robustness considerations. Each method offers unique advantages and drawbacks, highlighting the importance of thoughtful selection based on specific modeling needs in extreme event analysis

1.4 Empirical Validation

Empirical analysis serves as a cornerstone in validating and refining the suitability of chosen probability distribution alongwith estimation methods by leveraging historical data on variable of interest, such as precipitation, to evaluate its fit and predictive capacity. This typically involves statistical techniques such as goodness-of-fit tests, where the observed precipitation data is compared to the modeled distribution to determine how well they align. Additionally, measures of predictive performance are employed to evaluate how well the model predicts precipitation levels for unseen data.

Analyzing various data types, including historical records of extreme weather events, observational data from weather stations, and climate model projections, is essential for understanding the complex interplay between geography, rainfall patterns, and extreme events in Pakistan. Utilizing suitable estimation methodologies, such as statistical analysis and climate modeling simulations, facilitates the optimal selection of models tailored to forthcoming precipitation data. Moreover, this multidisciplinary approach allows for the identification of vulnerable areas, the prediction of future trends in extreme weather, and the development of targeted adaptation strategies to mitigate the impacts of climate change on Pakistan's environment and society.

1.5 Climate of Pakistan

Pakistan, nestled between latitudes 24° and 37° N and longitudes 61° and 77° E, boasts a varied topography, from the towering peaks of the Himalayas, Karakoram, and Hindu Kush ranges in the north to the vast plains and coastal areas in the south. The nation covers a vast geographical expanse, totaling 803,943 square kilometers in area. Subjected to a myriad of climate extremes, the region experiences significant variability in rainfall patterns due to its diverse geography, which exposes it to a wide range of climatic influences.

The country's monsoon climate is the dominant factor influencing its rainfall patterns. The southwest monsoon, originating from the Arabian Sea, brings heavy rainfall to the southern and southeastern parts of the country during the summer months (July to September). Coastal areas, such as Karachi, receive substantial rainfall during this period, contributing to their annual precipitation totals.

In contrast, the northern and western regions of Pakistan, including the mountainous areas such as the Himalayas, Karakoram, and Hindu Kush ranges, experience comparatively lower rainfall during the monsoon season. These regions often rely on winter precipitation, primarily in the form of snowfall, for their water supply.

1.6 Study Area

The intricate tapestry of Pakistan's climatic diversity unfolds across its varied landscapes. The climatic zones of the country have been delineated based on their diversity, with a total of five distinct zones identified in reference to existing research [1]. From this literature, two specific zones denoted as D and E, along with their corresponding latitudinal extents, were chosen for analysis, as depicted in Figure In this study, stations were carefully selected considering factors such as their latitude, elevation above sea level, length of data record, and data reliability to ensure a comprehensive and reliable synthesis of information.

In addition to the existing stations, several more were incorporated into the designated zones for comprehensive analysis. These additional stations were selected based on similar criteria. This augmentation allowed for a more robust representation of climatic conditions across the selected zones, facilitating a thorough examination of rainfall trends and patterns.

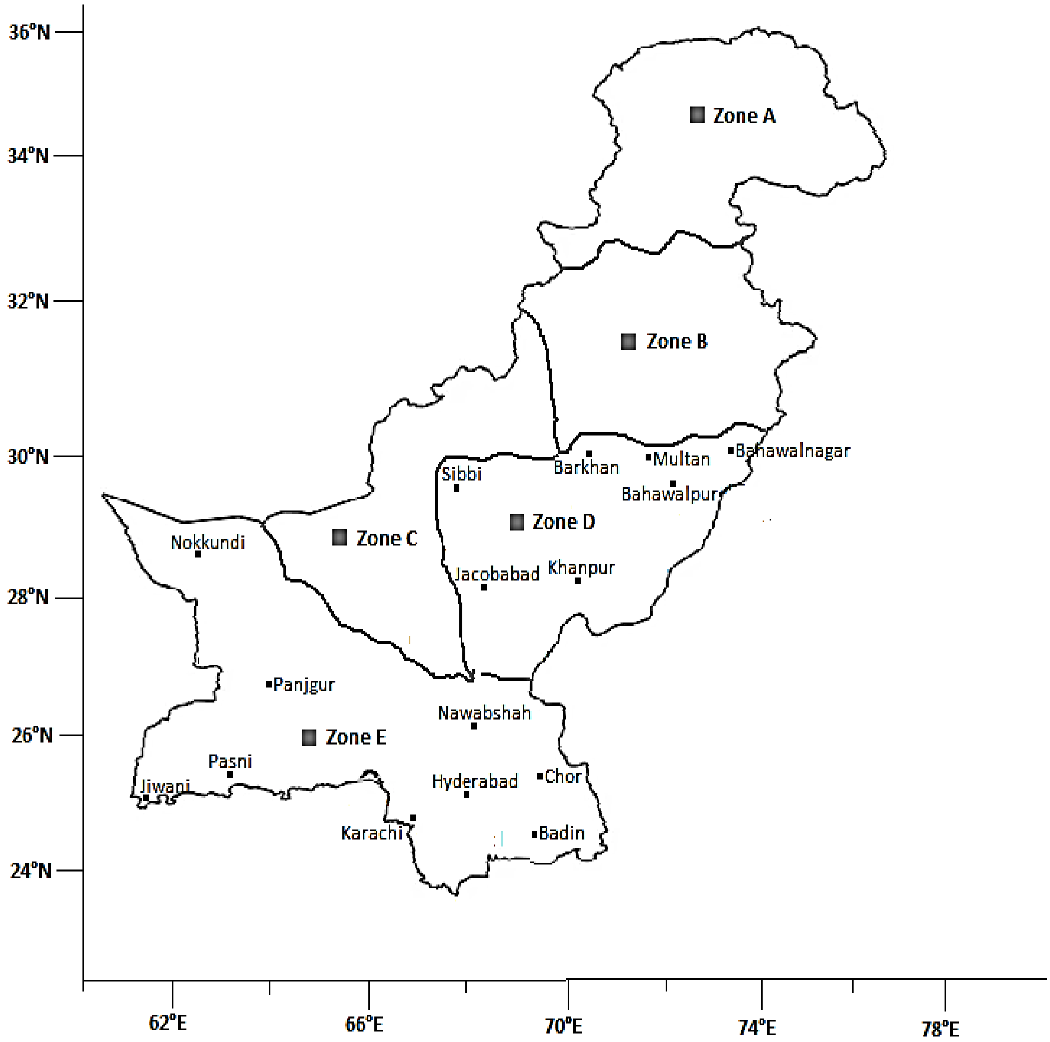

Figure 1.1: A map showing all zones including zone D and E with selected stations for study area.

1.6.1 Zone D

This is the most arid and sweltering zone of Pakistan, characterized by scorching temperatures and scant rainfall, encompasses vast expanses of land predominantly in the southwestern regions of the country. Jacobabad, Nawabshah, and Sukkur, temperatures ascended to 50.2°C (122.4°F) in May 2018 and 50°C in June 2019 [2]. Due to the considerable distance from the sea, temperatures persist at elevated levels, with infrequent but occasionally intense rainfall events leading to flooding. Stations in this zone include Sibbi, Jacobabad, Bahawalpur, Khanpur, Multan, Bahawalnagar and Barkhan.

1.6.2 Zone E

Zone E is a sizable region featuring numerous stations and coastal cities situated in proximity to the Arabian Sea. The region's flat topography and extensive network of rivers, including the Indus, make it susceptible to inundation during heavy rainfall or when river levels rise significantly. The stations identified within this zone are Hyderabad, Karachi, Nawabshah, Jewani, Badin, Chor, Pasni, Panjgur and Nokkundi.

1.7 Problem Identification

With ongoing climate change, there is increasing evidence that certain types of extreme events are becoming more frequent, intense, or prolonged in many regions of the world. Extreme precipitation events are a matter of worldwide apprehension because they have the capacity to trigger significant flooding, landslides, and other catastrophic natural occurrences. These events have been linked to climate change and are a focus of research worldwide. In Pakistan, modeling extreme events is challenging due to insufficient availability of long-term, site-specific data, hindering an accurate representation of extreme precipitation behaviors. Extreme values often deviate from the norm or display asymmetrical behavior, necessitating specialized probabilistic modeling methods to precisely depict them. Therefore, probabilistic modeling, through appropriate estimation methods and assessments, is essential for understanding and preparing for extreme events, aiding in Trend Analysis, Risk Assessment and Climate Impact improvements in the country.

1.8 Aims and Objectives

- To analyse the performance of different estimation methods (L-moments and maximum product of spacing) for modelling climate extremes.

- To evaluate the fit and predictive power of the chosen distribution and estimation methods through empirical analysis of precipitation patterns in selected zones of Pakistan.

1.9 Relevance to National Needs

Probabilistic models are crucial for assessing and managing risks associated with climate extremes. By combining probabilistic projections of extreme events with vulnerability assessments of exposed systems (e.g., infrastructure, ecosystems, human populations), decision-makers can prioritize adaptation measures, allocate resources, and develop resilience strategies to minimize the impacts of extreme events within Pakistan.

Climate change is expected to have significant impacts on the frequency, intensity, and duration of extreme events in Pakistan. The research's findings on behavior of extreme events and their application to native datasets have the potential to inform future research and practice in a range of fields including risk management, climate modeling and disaster preparedness. As such, the research can be considered a scientific contribution that advances our knowledge and understanding of extreme value estimation and its application to national problems.

1.10 Thesis Organization

The thesis comprises 5 chapters and a References section. Chapter 1 provides a comprehensive overview of the study's background, objectives, and flow. Chapter 2 outlines existing research and identifies gaps. Chapter 3 details the methodology employed, including Probability distribution fitting, Empirical Analysis, Parameter Estimation methods selection, visual representation, and predictive analysis. Chapter 4 presents the results of various estimation methods and forecasts future extreme events. Chapter 5 offers a brief conclusion and suggests potential future research directions. The References section lists all cited works along with their publication details.

2. Literature Review

2.1 Existing Literature

This paper provides insights into a wide range of applications, including the analysis of precipitation and streamflow extremes, as well as the assessment of economic damage linked to these extreme events [3].

Topics that are addressed include: trends in hydrologic extremes and statistical downscaling of hydrologic extremes. The authors specifically focus on methodological developments involving maximum likelihood estimation in the presence of covariates, combined with approaches like block maxima or peaks over threshold.

Two examples were considered:

I. The maximum of daily precipitation amount for the month of January at Chico, CA, USA, for 78 years is modeled with the covariate being the mean sea level pressure. the fitted GEV appears reasonably satisfactory for this example.

II. When applying maximum likelihood estimation to fit the GEV distribution to annual peak flow data from the Salt River near Roosevelt, AZ, USA, for the years 1924-1999, it resulted in a very poor fit for the highest observations. Instead, a Generalized Pareto distribution, also fitted using maximum likelihood, provided a more satisfactory fit compared to the GEV distribution.

The central theme of this review revolves around the efficiency of MPS estimators, especially in cases where sample sizes are small, offering a more reliable alternative to MLE [4]. Simulations were conducted to assess the benefit of using the MPS method compared to MLE for GEV and GPD, across various sample sizes and parameter choices.

In this study, the MPS method was also used to analyze four real datasets: The annual oldest ages at death in Sweden from 1905 to 1958. The wave dataset contains the yearly maximum heights, in feet.

Then, in the wind data, the yearly maximum wind speed in miles per hour is considered. The last example is the flood data which consists of the yearly maximum flow discharge, in cubic meters.

The results demonstrate that MLE exhibits instability at small sample sizes, which aligns with the simulation results, while the MPS method maintains its effectiveness.

The objective of this study was to evaluate widely-used methods for estimating the parameters of the GEV distribution [5]. This assessment involved both simulated data and actual wind measurements gathered from four buoys situated in the Atlantic and Pacific Ocean regions.

The analysis suggests that methods like MPS, EP, ordinary entropy, and ML are generally better at estimating GEV distribution parameters, with lower bias and error. Surprisingly, lesser-known methods like MPS and EP also perform well in describing extreme wind speed quantiles in real data.

This study evaluates three parameter estimation methods for the Pearson Type-3 (PE3) distribution: L-moments (LM), maximum likelihood estimation (MLE), and maximum product of spacing (MPS) [6]. The assessment involves simulation experiments and empirical analyses considering different sample characteristics. The study also assessed these methods using real data from four sites in KPK, Pakistan, measured in Annual Maximum river discharge AMRD (cubic feet per second).

The findings suggest that LM is best for small samples with low skewness and kurtosis, MPS works well for highly skewed and kurtotic data with moderate sample size, while MLE is suitable for very large sample sizes with low shape characteristics. These results offer valuable guidance for fitting the PE3 distribution, particularly for extreme values.

The paper explores two primary approaches for (maximum flood discharge) MFD estimation: deterministic models that rely on extreme storm events and probabilistic frequency analysis [7]. In this study, the authors employ the method of moments and L-moments (LMO) to determine parameters for six different probability distributions. They further evaluate the adequacy of these distributions through goodness-of-fit tests. Moreover, the diagnostic test is utilized to select the most appropriate distribution for MFD estimation.

The findings of the study underscore the superior performance of the Extreme Value Type-1 distribution, particularly when using L-moments, among the six distributions considered for estimating MFD.

The study conducted regional flood frequency analysis (RFFA) to estimate peak discharges at the regional level over Kerala State, India, along with at-site flood frequency analysis [8]. The researchers used annual peak discharges from 43 gauging stations with data lengths ranging from 14 to 47 years.

To identify the best-fit distribution for both at-site and regional analyses, five distributions were considered: Generalized Extreme Value (GEV), Generalized Pareto Distribution (GPA), Generalized Logistic (GLO), Generalized Normal (LN3), and Pearson Type III (PE3). Chi-square test, ranking method using statistical indicators, and L-moment ratio diagram were employed for at-site analysis.

The results showed that GPA was the best fit for 27 stations, GLO for 14, LN3 for 1, and GEV for 2 stations. After discordancy and heterogeneity tests, five homogeneous regions were identified for RFFA. The best-fit distributions for each zone were used to derive flood growth curves incorporating catchment characteristics. Interestingly, the best-fit distribution for a gauging site in at-site analysis differed from the RFFA results. The RFFA growth curves provide flood magnitudes for various return periods, which can be used to estimate flood magnitude and frequency at ungauged sites in each region of Kerala State.

This study aims to compare the probabilities of extreme still water levels estimated using the block maxima method and the peaks over threshold method [9]. The study uses a wide range of strategies to create extreme value datasets and considers different model setups. The focus is on testing the influence of detrending techniques, sample sizes, and record lengths on the estimates of extreme value statistics.

The study finds that using different techniques can significantly bias the results from extreme value statistics. Therefore, it recommends using a 1-year moving average of high waters (or hourly records if available) to correct for seasonal and long-term sea level changes. Additionally, the study finds that the peaks over threshold method yields more reliable and stable estimates of probabilities of extreme still water levels than the block maxima method. The study also recommends using the 99.7th percentile water level as the threshold for the peaks over threshold method.

This study describes methods to calculate extreme wind speeds, focusing on both classical statistical approaches and newer techniques designed for short data sets [10]. Traditional methods include the Generalized Extreme Value (GEV) distribution and the Generalized Pareto Distribution (GPD). The GEV is often used with annual maxima, fitting the highest wind speeds observed each year to predict extremes, while the GPD is applied in peak-over-threshold (POT) approaches, modeling exceedances over a set threshold. These classical methods rely on robust statistical foundations and techniques like maximum likelihood estimation for parameter calculation. Ensuring data independence and minimizing standard errors are crucial to enhancing the reliability of these methods.

In scenarios with short data sets, which are common in extreme wind speed studies, this study describes alternative techniques to address the limitations. Simulation modeling, such as Monte Carlo simulations, is used to extend limited data sets, thereby improving accuracy. Additionally, parametric methods based on the parent distribution, including Bayesian inference, offer effective solutions by estimating underlying distribution parameters and updating estimates with new data. These approaches are tailored for short-term data and help maintain prediction accuracy despite limited observations, meeting the needs of users requiring precise extreme wind speed calculations.

The study elucidates the significance of L-moments as a powerful statistical tool for analyzing probability distributions and empirical data samples [11]. Through their definition as expectations of linear combinations of order statistics, L-moments offer a robust framework that encompasses various statistical procedures, including summarization, parameter estimation, and hypothesis testing. By leveraging established techniques like order statistics and Gini's mean difference statistic, L-moments introduce innovative measures for skewness, kurtosis, and parameter estimation for multiple distributions.

One notable advantage highlighted in the study is the robustness of L-moments to sampling variability, particularly their resilience against outliers in data. This robustness makes L-moments more reliable for making inferences from small sample sizes about underlying probability distributions. Furthermore, the study suggests that L-moments often yield more efficient parameter estimates compared to maximum likelihood estimates, further enhancing their practical utility in statistical analysis.

In summary, the study underscores the versatility and efficiency of L-moments in statistical analysis, emphasizing their superiority over conventional moments in handling data variability and facilitating more accurate inferences from limited data samples. These findings establish L-moments as a valuable tool for researchers and practitioners across diverse fields, enabling more robust and reliable statistical analyses.

This study employs the TL-moments approach to analyze annual maximum streamflow data from seven stations in Johor, Malaysia, aiming to identify the best-fitting probability distributions [12]. TL-moments, with varying trimming values, are utilized to estimate parameters for selected distributions, specifically the Three-parameter lognormal (LN3) and Pearson Type III (P3) distributions. The primary objective is to derive TL-moments (t 1,0), t 1 = 1,2,3,4 methods tailored for LN3 and

P3 distributions. Through Monte Carlo simulation and analysis of streamflow data, the performance of TL-moments (t 1,0), t 1 = 1,2,3,4 is compared against L-moments, with the absolute error serving as a metric to evaluate the influence of TL-moments methods on estimated probability distribution functions.

The study reveals that TL-moments with four trimmed smallest values from the conceptual sample (TL-moments [4, 0]) of LN3 distribution exhibit superior performance across most stations in Johor, Malaysia, for modeling annual maximum streamflow series. This finding suggests the efficacy of TL-moments in capturing the underlying characteristics of the data and selecting appropriate probability distributions. By offering a comparative analysis with L-moments and utilizing Monte Carlo simulation, the study provides valuable insights into the application of TL-moments in hydrological analysis, highlighting their potential for improving the accuracy of probability distribution estimation in water resource management and related fields.

This paper addresses the challenges associated with fitting the Pearson Type-3 (P3) distribution to untransformed data, a task often hindered by traditional methods [13]. The study introduces an adaptive estimation procedure for the P3 family, leveraging fractional moments of exponentially transformed data and the mean of the original dataset. This approach offers simplicity in implementation, particularly advantageous for small sample sizes, and remains valid across the entire parameter space. The paper also provides explicit formulae for the variances and covariances of parameter estimators, as well as for the variance of the T-year event, enhancing the understanding and applicability of the proposed method.

Furthermore, the study conducts a comparative analysis by pitting two variants of the new procedure against two versions of the method of moments and a version of the method of conditional moments through Monte Carlo simulation. Results from samples generated from P3 populations indicate that one variant of the new procedure outperforms others in estimating 100-year flood events, while the other variant excels in estimating median and 10-year low-flow events. Remarkably, the robust performance of these variants extends to samples generated from distributions other than P3, underscoring their versatility and reliability in various scenarios.

Moreover, the paper introduces and investigates a modification of the procedure tailored for cases where a prior assumption of positive skewness is adopted, further demonstrating the adaptability and effectiveness of the proposed approach. Overall, this study contributes valuable insights into improving the estimation of extreme hydrological events, offering a practical and robust solution for fitting the P3 distribution to untransformed data in statistical hydrology.

This study addresses the challenges encountered in curve-fitting methods due to discrepancies in accuracy and positions of experience points, particularly in hydrological sequence data with limited sample sizes [14]. Recognizing the need to assign varying importance to data points and mitigate sampling errors in parameter estimation, the study introduces a weighted approach within the optimum curve-fitting method. Through an analysis of existing weighted methods, the study focuses on the Fuzzy Weighted Optimum Curve-fitting Method (FWOCM), which overcomes limitations such as nominal length and membership degree function determination without relying on large samples. To enhance this method, a new membership degree function is derived, assuming the hydrologic sequence is sufficiently large. Additionally, the study employs Monte Carlo statistical tests to extend the nomograph across the entire frequency range, aiming to evaluate the effectiveness of the improved FWOCMs using both ideal and real data.

By introducing enhancements such as the extension of the nomograph and the derivation of a new membership degree function, the study aims to mitigate the impact of shorter hydrologic sequences on curve-fitting accuracy. Through comparative analysis using selected benchmark methods and improved percentage methods, the study evaluates the performance of the enhanced FWOCMs, demonstrating promising results that suggest their suitability for engineering applications. This research contributes to the advancement of curve-fitting methodologies in hydrology by addressing inherent challenges and offering practical solutions to improve accuracy and reliability in curve-fitting procedures, particularly in scenarios with limited sample sizes and varying data point importance.

This paper represents a significant advancement in the Bayesian Forecasting System (BFS) for hydrological forecasts by addressing the limitations of traditional distribution assumptions in the Hydrological Forecast Processor (HUP) [15]. Traditionally, HUP assumes runoff distributions follow Logweibull or Normal distributions, which may not be accurate across different regions. To improve accuracy, this study introduces Nonparametric Kernel Density Estimation, Pearson III, and Empirical distributions as prior distributions to mitigate parameter uncertainty.

By comparing these five distributions using data from 52 floods in the ZheXi basin (2004-2014), the study demonstrates that while Logweibull and Empirical

Bayesian models show the best average performance, the new distributions also perform well regarding interval width and probability forecasting. These findings suggest that incorporating a broader range of distribution types can enhance the robustness of BFS, highlighting the need for further exploration of diverse distributions to improve hydrological forecast reliability.

This paper illustrates the results of a regional frequency analysis of Annual Maximum Monthly Rainfall Totals (AMMRT) at seven sites in Sindh, Pakistan [16]. The analysis includes run tests, lag-1 correlation coefficients, and Mann-Whitney tests, all of which indicate that the data series are random, uncorrelated, and identically distributed. Discordancy measures show no discordant sites among the seven, and the L-moments based heterogeneity measure (H) confirms the region's homogeneity.

The study identifies three suitable regional distributions—GNO, PE3, and GPA—based on L-moment ratio diagrams and Z DIST statistics. To estimate rainfall quantiles at ungauged sites, a linear regression model is developed using the mean AMMRT of gauged sites and their respective elevations. This model satisfies the assumptions of Classical Linear Regression Modeling (CLRM) as verified by formal tests.

Consequently, the estimates from this study are useful for calculating rainfall quantiles for various return periods. These findings have practical applications in flood disaster prevention, agricultural water management, and the improvement projects for rehabilitating and modernizing major barrages of the Indus River in Sindh Province.

This paper presents an analysis of the Upper Vistula River basin, dividing it into pooling groups with similar dimensionless frequency distributions of annual maximum river discharge [17]. Using cluster analysis and the Hosking and Wallis (HW) L-moment-based method, the study divides 52 mid-sized catchments into disjoint clusters based on morphometric, land use, and rainfall variables, testing the homogeneity within these clusters. The study identified three and four pooling groups alternatively.

Two methods were employed to identify the regional distribution function: the HW method and the Kjeldsen and Prosdocimi method, which uses a bivariate extension of the HW measure. Flood quantile estimates were then calculated using the index flood method. The study compared ordinary least squares (OLS) and generalized least squares (GLS) regression techniques to relate the index flood to catchment characteristics, finding that GLS improved predictive performance in the southern part of the Upper Vistula River basin.

The study's results are recommended for estimating flood quantiles at ungauged sites, useful for flood risk mapping, and beneficial in engineering hydrology for designing flood protection structures.

2.2 Research Gap

Despite significant advancements in probabilistic modeling and hydrological forecasting notable gaps remain, particularly in the framework of Pakistan. While various studies have explored Bayesian forecasting systems and Regional frequency analysis in various contexts globally, there is a scarcity of comparative model analyses specifically tailored to Pakistan's diverse hydrological regimes. Existing studies often focus on specific regions or limited datasets, lacking comprehensive comparisons across different modeling approaches that could provide insights into the best practices for forecasting.

There is a clear opportunity for valuable research contributions in addressing these gaps. The majority of existing studies tend to rely on more conventional distributions such as Logweibull or Normal distributions, potentially overlooking the benefits that Pearson Type-3 distribution might offer in capturing the characteristics of local rainfall and runoff patterns more accurately. Comparative analysis of parameter estimation methods for Pearson Type-3 distribution further elucidate the strengths and weaknesses of each approach in capturing the complexities of local hydrological processes. Additionally, investigating the applicability and advantages of this distribution in hydrological modeling could expand the toolkit available to researchers and practitioners in the region, thereby enhancing the accuracy and reliability of flood risk assessments and management strategies.

3. METHODOLOGY

This chapter presents the methodological framework employed to analyze precipitation patterns in two major zones of Pakistan. The methodology integrates various statistical and empirical techniques to ensure a comprehensive examination of historical precipitation data. By outlining the data preliminary analyses, and sophisticated modeling approaches, this chapter aims to provide a clear and structured guide to understanding the complex dynamics of rainfall distribution and its implications for future flood risk assessments.

3.1 Research Design

The research design comprises a multi-step approach aimed at accurately modeling and analyzing precipitation data. Initially, historical precipitation data is gathered and curated from the Pakistan Meteorological Department (PMD). This data undergoes a preliminary screening to ensure quality and reliability. The analysis proceeds with fitting the Pearson Type-3 (PE-3) probability distribution to the data, followed by parameter estimation using various statistical methods. The performance and accuracy of these models are evaluated through a series of diagnostic tests, ensuring robust and reliable findings.

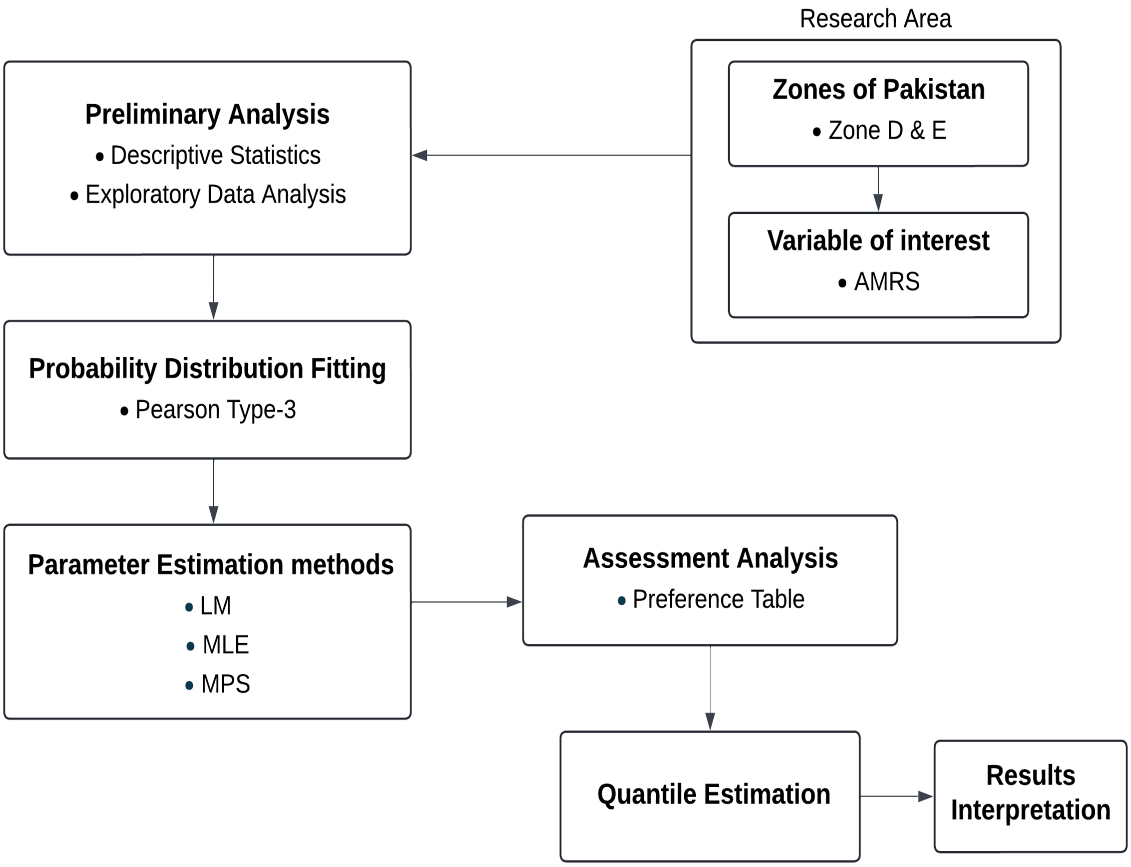

Figure 3.1: Flow of Work

3.2 Data under Consideration

To describe characteristics of the data, demographics including age and city of residence of the study population are analyzed. Prevalence of psychosocial dysfunction, attention problems, internalizing problems, and externalizing problems according to the PSC are also determined using the scoring instructions of Massachusetts General Hospital. Measures of central tendency including mean and mode are calculated for age of the subjects, total score of the PSC, and scores of the three subscales of the PSC for each age group.

3.3 Preliminary Analysis of Data

Extreme rainfall events are identified and analyzed to distinguish genuine extremes from potential errors. This involves checking for missing values, outliers, and ensuring consistency in formatting. This step ensures the dataset's integrity and the accuracy of subsequent analyses.

The dataset comprises geographical coordinates (latitude and longitude) of meteorological stations where precipitation data were recorded. These coordinates are essential for spatial analysis and mapping the distribution of extreme rainfall events across the study area.

3.3.1 Descriptive Analysis

Descriptive statistics provide concise summaries of datasets, helping to understand the overall characteristics, central tendency, and variability of the data. Measures such as minimum and maximum values, mean, standard deviation, skewness and kurtosis offer a snapshot of the dataset's distribution and spread.

3.3.2 Exploratory Data Analysis

Exploratory Data Analysis (EDA) involves examining datasets to summarize their main characteristics, often using visual methods. When dealing with extremes, one of the primary tools for EDA is the time series plot.

Time series plots are graphical representations of data points in a time-ordered sequence. They are essential for understanding the temporal dynamics of the data and identifying patterns, trends, and anomalies.

3.4 Probability Distribution Fitting

Probability distribution fitting is a fundamental tool in climate science for understanding and preparing for extreme weather events, such as heavy rainfall, heatwaves, and strong winds. These extremes often follow non-normal, skewed distributions, necessitating specialized approaches to accurately represent their behavior and impacts. By selecting appropriate distributions and estimation methods, statistical properties of these events can be better characterized, aiding in trend analysis and climate modeling.

3.4.1 Pearson Type-3

The Pearson Type-3 distribution, with its three parameters delineating shape, scale, and location, stands as a cornerstone in statistical modeling, bridging theoretical elegance with empirical application. Its versatility in accommodating skewed and heavy-tailed distributions makes it indispensable in disciplines where accurate representation of data variability is paramount.

Named after Karl Pearson 1895 [18], this distribution is part of the Pearson system of distributions. Pearson developed this system to model skewed distributions more effectively than the normal distribution, which assumes symmetry. The Pearson system classifies distributions based on their skewness and kurtosis, offering a family of curves that includes the normal, beta, and gamma distributions as special cases.

In the realm of statistical literature, the Type III distribution is regarded as a member of the gamma family, specifically the three-parameter gamma family (when $\mu \neq 0$), where the shape parameter governs the skewness of the distribution. This association highlights its capability to describe a wide range of data distributions with differing skewness characteristics. When the location parameter of the PE3 distribution is set to zero (i.e: $\mu = 0$), it corresponds to what is termed as the two-parameter gamma distribution, as discussed by Johnson et al. (1995) [19]. Chow, V.T. (1954) applied the Pearson Type-3 distribution to flood frequency analysis, establishing its importance in hydrological studies [20]. Johnson, N.L., Kotz, S. (1970) provides an in-depth exploration of continuous univariate distributions, including the Pearson Type-3 distribution, detailing their properties, applications, and methods of parameter estimation [21]. Bobée, B. and Ashkar, F. (1991) discussed the gamma distribution and its applications in hydrology, equating it to the Pearson Type

III distribution [22].

Pearson Type-3 distribution is characterized by three parameters: location (μ\mu), scale (σ\sigma), and shape (γ\gamma), which collectively define its central tendency, spread, and skewness. It can be expressed in terms of its probability density function (PDF) as:

$$f(x;\, \gamma, \sigma, \mu) = \frac{1}{\sigma\tau(\gamma)}\left(\frac{x-\mu}{\sigma}\right)^{\gamma-1} \exp\left(-\left(\frac{x-\mu}{\sigma}\right)\right), \qquad x \geq \mu$$

where:

- x is the random variable representing the observation.
- μ is the location parameter, shifting the distribution along the x-axis. It can be any real number.
- σ is the scale parameter, determining the spread or variability of the distribution. It must be positive $(\sigma > 0)$.
- γ is the shape parameter, influencing the skewness of the distribution.
- $\tau(\gamma)$ denotes the gamma function.

3.4.2 Characteristics and Parameters

1- **Location Parameter (μ):** The location parameter shifts the distribution along the x-axis. It represents the mean of the distribution and determines its position relative to the origin.

2- **Scale Parameter (σ):** The scale parameter determines the spread or variability of the distribution. Larger values of σ indicate greater variability in the data.

3- **Shape Parameter (γ):** This parameter dictates the skewness of the distribution. For $\gamma > 0$, the distribution is right-skewed (positively skewed), while $\gamma < 0$ indicates left-skewed (negatively skewed) distributions. A special case occurs when $\gamma = 0$, where the distribution simplifies to a normal distribution.

Precipitation extremes often exhibit skewed distributions due to the rare occurrence of very high values. The Pearson Type-3 distribution is well-suited for modeling these extremes because of its flexibility in handling skewed data. It can accurately represent the tail behavior of precipitation events, which is critical for risk assessment and management in hydrology.

3.5 Parameter Estimation Methods

3.5.1 L-Moments Method

The L-moments method, pioneered by J.R.M. Hosking in the 1980s [23], represents an advancement beyond traditional moments by offering robust estimators for distribution parameters. Specifically designed to handle skewed and heavy-tailed distributions, L-moments mitigate the sensitivity to outliers and departures from normality that can affect conventional moments. L-Moments, analogous to moments of a distribution, leverage weighted averages from order statistics rather than relying solely on raw data moments. Their versatility and interpretability make them a valuable tool for modern statistical analysis and modeling.

- **First L-Moment (λ1):**

The first L-moment represents the mean or the location parameter of the distribution. It is simply the average of data points.

$$\lambda_1 = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{X_i} \tag{i}$$

- **Second L-Moment (λ2)**

The second L-moment represents the dispersion or scale of the distribution. It is related to the spread of the data and is comparable to the standard deviation in traditional moments.

$$\lambda_2 = \frac{1}{2}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right) \tag{ii}$$

- **Third L-Moment (λ3)**

The third L-moment measures the skewness of the distribution. It captures the asymmetry in the data and indicates whether the distribution is skewed to the left or right.

$$\lambda_3 = \frac{1}{3}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^3\right) - \frac{1}{2}\lambda_2 \tag{iii}$$

- **Fourth L-Moment (λ4)**

The fourth L-moment measures the kurtosis of the distribution. It provides information about the peakedness or flatness of the distribution compared to a normal distribution.

$$\lambda_4 = \frac{1}{4}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^4\right) - \lambda_3 \qquad \text{(iv)}$$

For all (i), (ii), (iii) and (iv):

- X_i are the observations.
- $\bar{X}$ is the sample mean.
- n is the sample size.

- **L-moment Ratios:**
  - L-skewness (τ_3):

$$\tau_3 = \frac{\lambda_3}{\lambda_4}$$

  - L-kurtosis (τ_4):

$$\tau_4 = \frac{\lambda_4}{\lambda_2}$$

Sample L-moments along with L-moment ratios are required to calculate location, scale and shape parameters for PE3.

In summary, while a distribution can be both skewed and heavy-tailed, L-moments describe different aspects of its shape and behavior. Skewness relates to symmetry around the mean, while heavy-tailedness pertains to the behavior of the distribution's tails and the likelihood of extreme values. Moreover, L-moments are advantageous for Pearson Type-3 distribution due to their robustness against outliers and their ability to describe shape and tail behavior of skewed distributions.

3.5.2 Maximum Likelihood Estimation Method

Maximum Likelihood Estimation (MLE) is a widely used statistical method for parameter estimation. It aims to find the parameters of a statistical model that maximize the likelihood function, which measures how likely the observed data are given the model parameters. The likelihood function represents the probability of observing the given sample data as a function of the distribution parameters. It has become a cornerstone of statistical inference due to its asymptotic properties and computational feasibility.

The MLE method was developed by the British statistician Ronald A. Fisher in the early 20th century. Fisher introduced the concept of likelihood and the method of maximum likelihood in a series of papers starting from 1912 and more formally in 1922 [24]. This method has become one of the most widely used techniques for parameter estimation due to its desirable properties, such as consistency, efficiency, and asymptotic normality under certain conditions.

For the Pearson Type-3 distribution:

$$\hat{\mu}_{MLE}\ ,\hat{\sigma}_{MLE}\ ,\hat{\gamma}_{MLE} = \arg max_{\mu,\sigma,\gamma}\ \left[\prod_{i=1}^{n} f(X_i\,;\gamma,\sigma,\mu)\right]$$

where:

$f(X_i;\gamma,\sigma,\mu)$ is the probability density function (pdf) of PE-3 distribution.

3.5.3 Maximum Product of Spacings Method

Introduced by *Cheng and Amin and independently by Ranneby* [25], the Maximum Product of Spacings (MPS) method has gained recognition as a robust technique for estimating shape parameter of extreme value distributions. MPS method involves maximizing the product of spacings between successive order statistics. This approach is based on the idea that the distribution's parameters can be estimated by ensuring that the gaps (spacings) between ordered data points match those expected under hypothesized distribution.

MPS is advantageous in precipitation extremes due to its robustness against outliers and model misspecification. It is particularly useful when data quality is variable or when the distributional assumptions of traditional methods may not hold perfectly such as distributions that exhibit heavy tails or contain outliers.

For the Pearson Type-3 distribution:

$$\hat{\gamma}_{MPS} = \arg max_{\gamma}\ \left[\prod_{i=1}^{n} f(X_i\,;\gamma,\sigma,\mu)\right]^{\frac{1}{n}}$$

where:

$f(X_i;\gamma,\sigma,\mu)$ is the probability density function (pdf) of PE-3 distribution.

3.6 Assessment Analysis

RMSE and bias play crucial roles in evaluating the accuracy, performance, and reliability of parameter estimation methods in statistical analysis. They provide quantitative insights into how well these methods perform in capturing the true characteristics of the data-generating process or underlying population parameters. These metrics together serve as diagnostic tools to identify potential shortcomings or strengths in estimation methods.

3.6.1 Root Mean Square Error (RMSE)

RMSE measures the average magnitude of the differences between estimated values and true values. It provides a measure of the overall accuracy or precision of the estimation method.

RMSE is calculated as the square root of the average of the squared differences between the estimated and true values:

$$R = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_i - \theta\right)^2} \qquad \text{(i)}$$

3.6.2 Bias

Bias measures the systematic error in the estimation of a parameter. It indicates whether the estimation method tends to consistently overestimate or underestimate the true parameter value.

Bias is calculated as the difference between the expected value of the estimated parameter and its true value:

$$B = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_i - \theta\right) \qquad \text{(ii)}$$

For (i) and (ii):

- $\hat{\boldsymbol{\theta}}_{\boldsymbol{i}}$ is the estimated parameter for the i-th sample,
- $\boldsymbol{\theta}$ is the true (population) value of the parameter,
- N is the number of samples.

3.7 Comparative Framework for Estimation Approaches (PT)

A preference table (PT) is generated to determine the most suitable parameter estimation method based on tail behavior and skewness levels. This evaluation is crucial as it helps identify which method effectively captures the characteristics of the data distribution. The table highlights how skewness and kurtosis levels influence the RMSE (Root Mean Squared Error) and bias values of each estimation method. Lower RMSE values indicate greater precision in parameter estimation, while bias measures systematic errors in estimation. By comparing these metrics across different methods, one can discern which approach best aligns with the distributional properties of the data, ensuring more accurate and reliable statistical inference.

3.8 Quantile Estimation

. Estimating flood quantiles using the quantile function of the Pearson Type-3 (PE3) distribution involves using statistical methods to fit the PE3 distribution to historical flood data. Once the distribution is fitted and its parameters (location, scale, and shape) are estimated, the quantile function of the PE3 distribution is utilized. This function allows for the direct calculation of flood quantiles corresponding to specific return period, such as 100-year event. These quantiles represent the flood magnitudes that have a given probability of being exceeded in any given year over the next 100 years. The process relies on the assumption that the historical data adequately represents future flood behavior under similar climatic and environmental conditions, ensuring that the estimated quantiles provide meaningful insights for risk assessment and planning purposes in hydrological studies.

3.9 Results Interpretation

The findings are analyzed at each level. The results from tables, calculated using Bias and RMSE, are interpreted to determine the most suitable parameter estimation method for each site and each parameter. Goodness-of-fit tests identify the best fitting method of parameter estimation. The estimated quantile results offer insights into potential future flood occurrences.

4. RESULTS AND DISCUSSIONS

In this chapter, the results generated by projected methodology are detailed and analyzed in relation to comparable studies. The chapter also underscores the study's strengths and weaknesses, providing suggestions for future research directions..

4.1 Data Overview

Pakistan is divided into five distinct zones, labeled A through E, each representing distinct geographical and climatic characteristics. The study area comprises 16 meteorological stations distributed across two main zones of Pakistan: Zone D and Zone E. All the concerned stations are located between Longitude 61.74573°E to 75.63372°E and Latitude 24.65572°N to 35.974986°N.

The dataset utilized in this study is Annual Maximum Rainfall Series (AMRS), which includes records of extreme rainfall events. This data has been meticulously gathered by the Pakistan Meteorological Department (PMD). The study covers a substantial time frame of 36 years, from 1980 to 2015, providing a robust dataset with 36 annual observations.

Table 4.1 and 4.2 detailing the meteorological stations, along with their respective longitude and latitude coordinates, are provided to give context to the locations of these stations within the study area. This information is essential for understanding the spatial distribution and regional variations in extreme rainfall events across studied zones.

Table 4.1: Geographical Coordinates of Seven Stations in Zone D

Sr.no	Stations	Longitude	Latitude
1	MULTAN	71.492157°	30.181459°
2	KHANPUR	70.6569°	28.64739°
3	JACOBABAD	68.4376°	28.28187°
4	SIBBI	67.87726°	29.54299°
5	BAHAWALPUR	71.67068°	29.41806°
6	BAHAWALNAGAR	73.25884°	29.99918°
7	BARKHAN	69.6994°	29.98482°

Table 4.2: Geographical Coordinates of Nine Stations in Zone E

Sr.no	Stations	Longitude	Latitude
1	JIWANI	61.74573°	25.04852°
2	KARACHI	67.0056°	24.94621°
3	NAWABSHAH	68.41003°	26.244221°
4	HYDERABAD	68.3693°	25.3817°
5	BADIN	68.83724°	24.65572°
6	CHOR	69.7666°	25.51667°
7	PASNI	63.4667°	25.2667°
8	PANJGUR	64.2500°	26.6667°
9	NOKKUNDI	62.75002°	28.8258°

4.2 Data Analysis

4.2.1 Descriptive Statistics

The total number of observations for each station is 36, each value representing annual maximum rainfall in millimeters. Min represents the minimum annual rainfall value and max represents the maximum annual rainfall value for the observed data series. Mean represents average rainfall in 36 years. Standard deviation is indicating how consistent or variable the data is relative to its average. The positive Skewness ranging from 0.56 to 3.93 shows the shape and Kurtosis ranging from -0.35 to 31.52 with both leptokurtic and platykurtic behaviors indicate the spread of data values (table 4.3 and 4.4).

The rainfall analysis across stations for zone D reveals diverse patterns: Multan has moderate rainfall with low extreme event likelihood (leptokurtic) and moderate right skewness, indicating a tendency towards more frequent lower rainfall values. Khanpur exhibits a high likelihood of extreme events (leptokurtic) and very high skewness, suggesting a significant presence of very high rainfall values. Jacobabad and Sibbi, with high extreme event likelihood (leptokurtic), show right skewness, reflecting a tendency towards more frequent high extremes. Bahawalpur, indicating frequent extremes (leptokurtic), also has high skewness, pointing to a notable presence of very high rainfall events. Bahawalnagar, with the highest average rainfall and variability, shows a high probability of extremes (leptokurtic) and significant right skewness, indicating a concentration of extreme high rainfall events. Barkhan, with moderate rainfall and variability exhibits right skewness, signifying a tendency towards more frequent extremes.

Key Findings

- Multan and Barkhan show moderate rainfall with consistent variability and fewer extreme events.
- Khanpur, Sibbi, Bahawalpur, and Bahawalnagar exhibit high variability and a strong

tendency towards extreme events, with Khanpur having the highest extreme likelihood.

- Jacobabad has lower average rainfall but a higher chance of extreme events compared to Multan and Barkhan

Table 4.3: Descriptive statistics of stations in Zone D.

Zone D							
Sr.no	**Site Name**	**Min**	**Max**	**Mean**	**Standard Deviation**	**Skewness**	**Kurtosis**
1	Multan	4.02	23.49	12.62	3.78	0.56	1.38
2	Khanpur	3.94	31.52	12.08	5.59	3.93	31.52
3	Jacobabad	2.10	33.61	10.04	6.77	1.72	3.32
4	Sibbi	1.93	56.70	13.59	9.60	2.86	11.22
5	Bahawalpur	5.70	38.80	13.79	5.36	3.05	13.43
6	Bahawalnagar	7.68	57.20	17.85	10.66	2.56	6.96
7	Barkhan	7.07	28.12	12.75	4.476	1.44	2.64

The rainfall analysis across stations for zone E reveals varied patterns: Jiwani has higher average rainfall with very high variability, right skewness (1.43), and leptokurtic distribution (kurtosis: 1.08), indicating extreme events. Karachi exhibits a high likelihood of extreme events (leptokurtic) with very high skewness (3.92). Nawabshah and Hyderabad, both with right skewness and leptokurtic distributions, suggest a higher probability of extremes. Badin shows frequent extreme events with high skewness (3.29) and high kurtosis (14.02). Chhor and Pasni also exhibit low to moderate skewness with leptokurtic distributions. Panjgur indicates a higher likelihood of extremes with significant right skewness (2.20) and leptokurtic distribution (kurtosis: 7.77). Nokkundi display fewer extremes with moderate to right skewness and platykurtic distribution.

Key Findings

- Karachi, Badin, Panjgur, and Hyderabad exhibit high skewness and high kurtosis, indicating a high likelihood of extreme rainfall events.
- Nawabshah, Jiwani, and Pasni show moderate to high probability of extreme events with right-skewed distributions and varying levels of kurtosis mostly Leptokurtic.
- Nokkundi has platykurtic distribution, suggesting fewer extreme events and a flatter distribution compared to all other locations.

Table 4.4: Descriptive statistics of stations in Zone E.

Zone E							
Sr.no	**Site Name**	**Min**	**Max**	**Mean**	**Standard Deviation**	**Skewness**	**Kurtosis**
1	Jiwani	1.43	45.79	14.66	11.61	1.43	1.08
2	Karachi	0.88	75.96	12.24	12.57	3.92	19.36
3	Nawabshah	0.79	36.55	11.05	7.24	1.32	3.32
4	Hyderabad	2.08	51.12	13.50	10.44	2.12	4.98
5	Badin	0.64	67.26	13.93	11.38	3.29	14.02
6	Chhor	1.19	27.72	12.17	5.66	0.58	1.02
7	Pasni	1.28	37.80	14.32	10.24	1.16	0.41
8	Panjgur	0.78	60.63	13.92	11.13	2.20	7.77
9	Nokkundi	1.07	33.48	12.59	8.80	0.633	-0.35

4.2.2 Time Series Plots

To comprehensively analyze extreme rainfall events, it is essential to visualize the temporal variations in rainfall across different monitoring stations. This section presents time series plots for all selected rainfall stations in zone D and E. These plots offer a detailed view of how rainfall patterns evolve over time, highlighting both seasonal and anomalous events. By examining these visual representations, we can identify trends, detect outliers, and better understand the spatial and temporal distribution of extreme rainfall events. The following figures illustrate the rainfall data from each station, providing a basis for further statistical analysis and comparison. These graphs illustrate the presence of random variabilities within the data series at each site.

Zone D

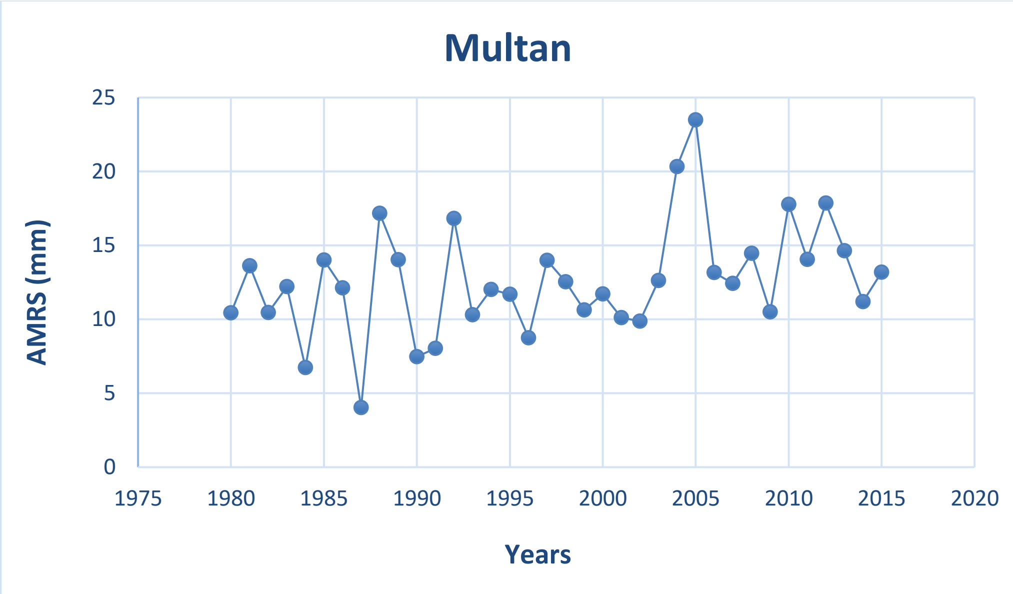

Figure 4.1: Time Series Plot of AMRS for the site Multan

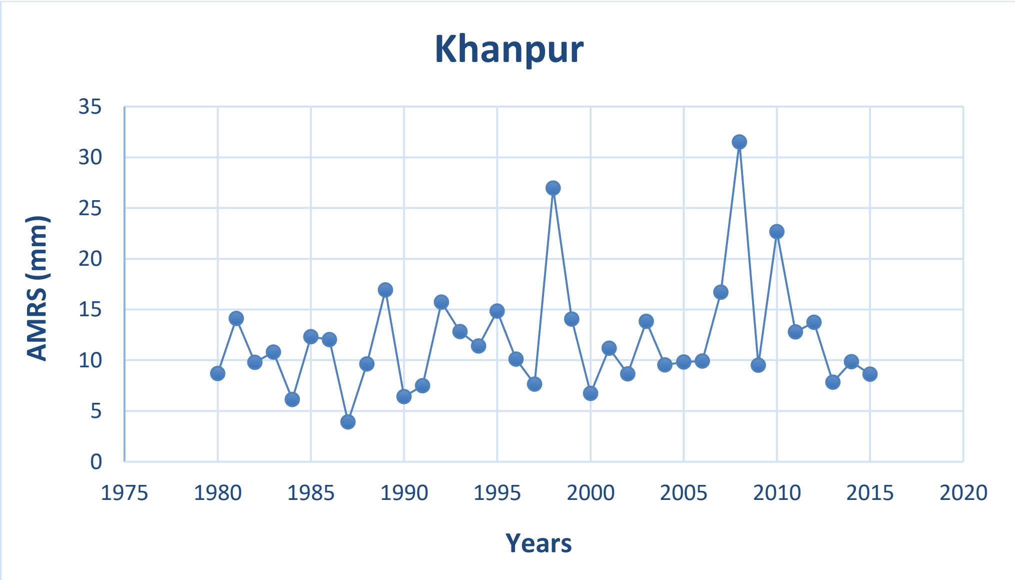

Figure 4.2: Time Series Plot of AMRS for the site Khanpur

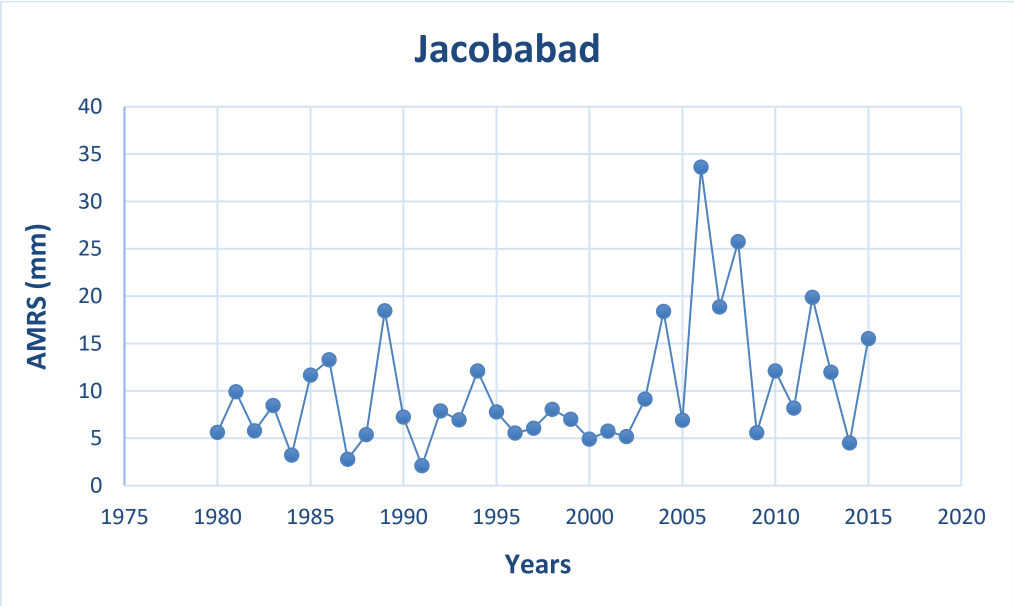

Figure 4.3: Time Series Plot of AMRS for the site Jacobabad

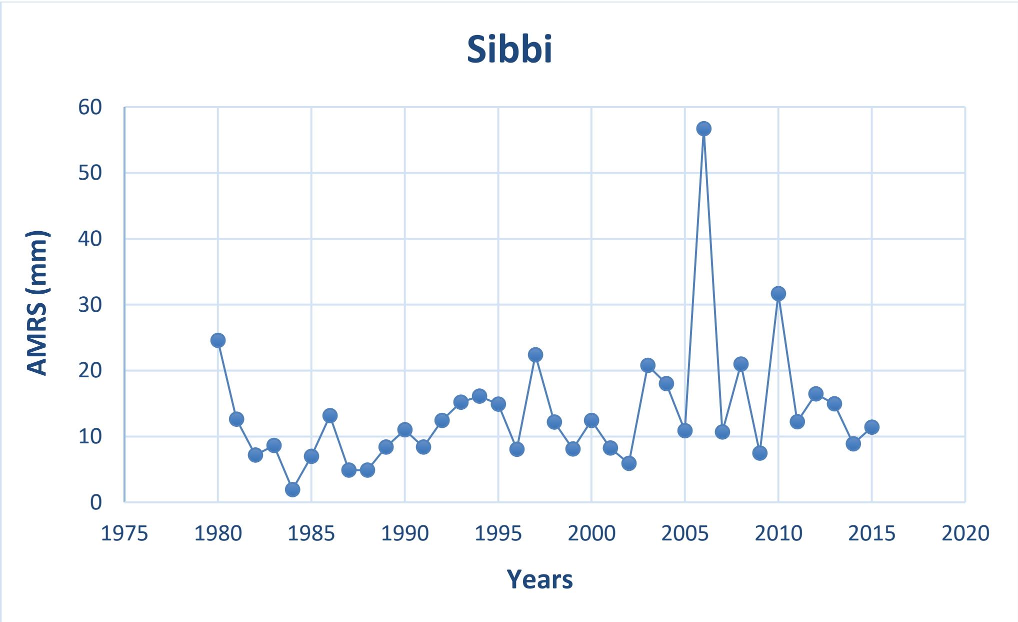

Figure 4.4: Time Series Plot of AMRS for the site Sibbi

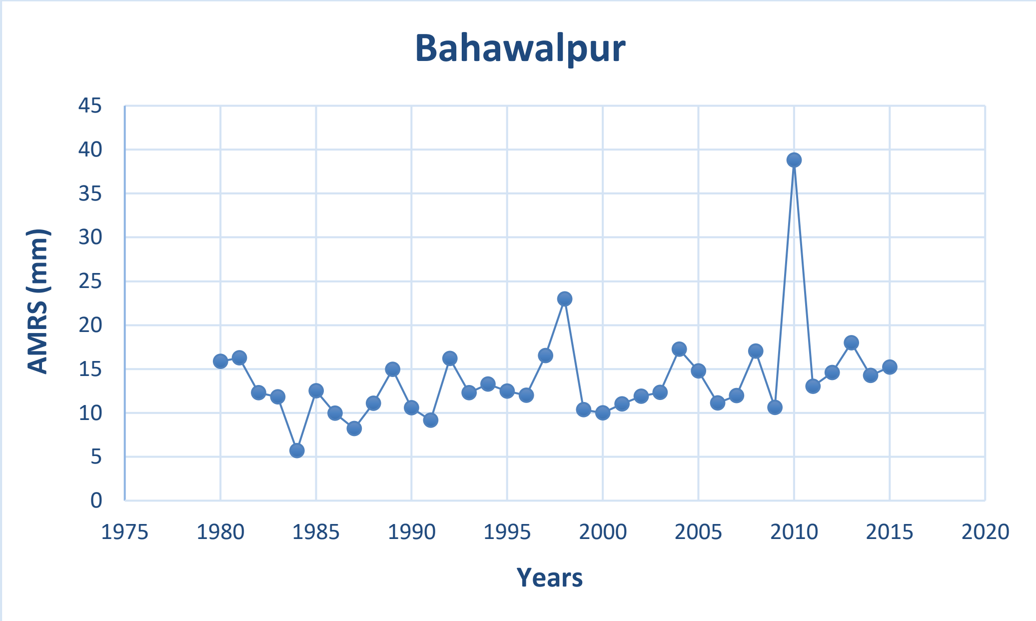

Figure 4.5: Time Series Plot of AMRS for the site Bahawalpur

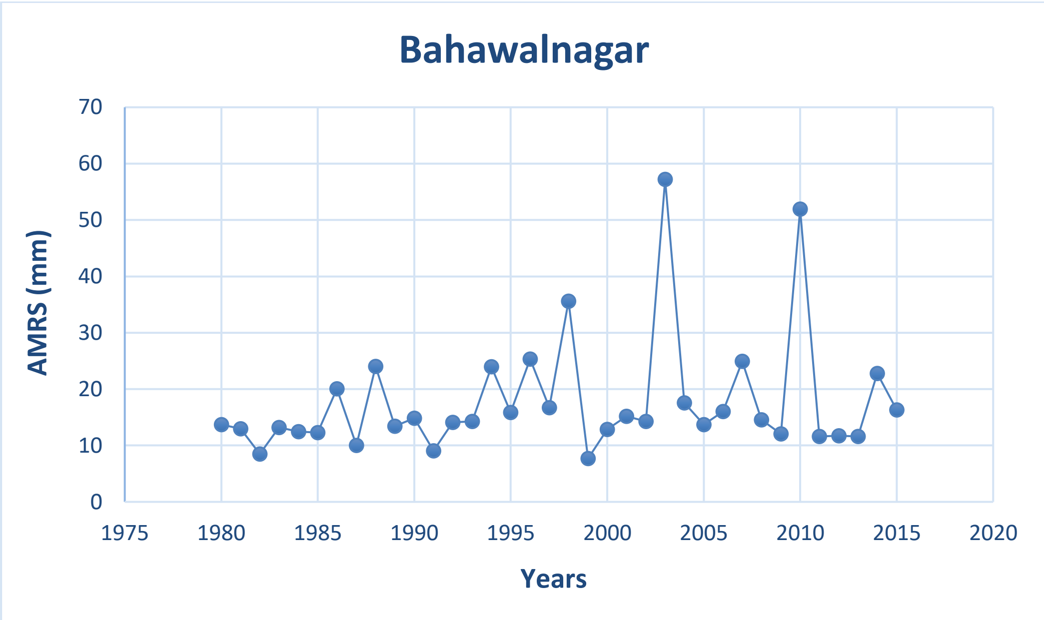

Figure 4.6: Time Series Plot of AMRS for the site Bahawalnagar

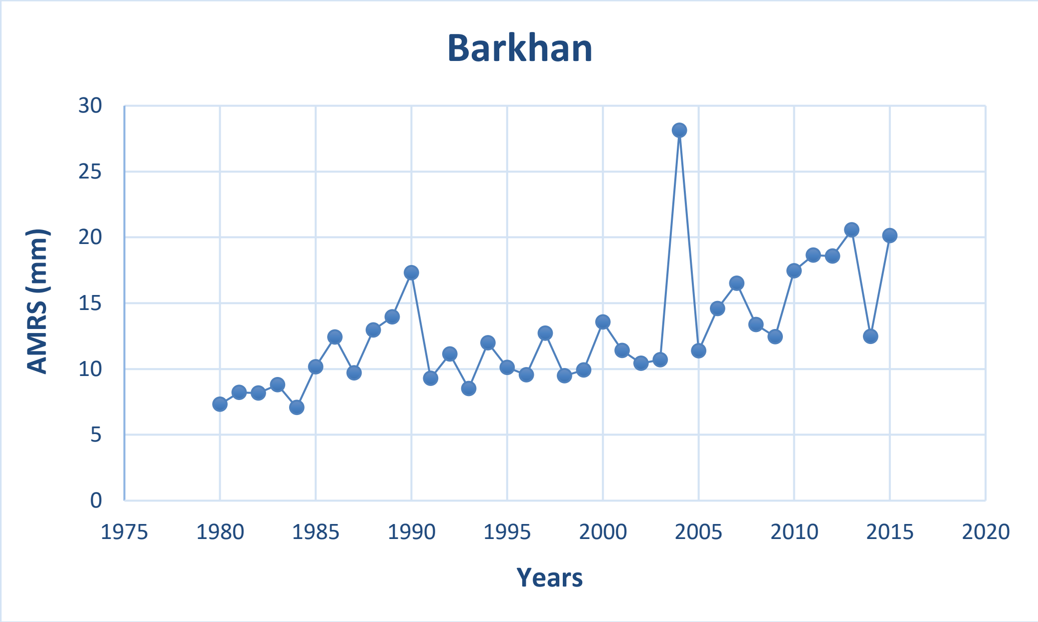

Figure 4.7: Time Series Plot of AMRS for the site Barkhan

Zone E

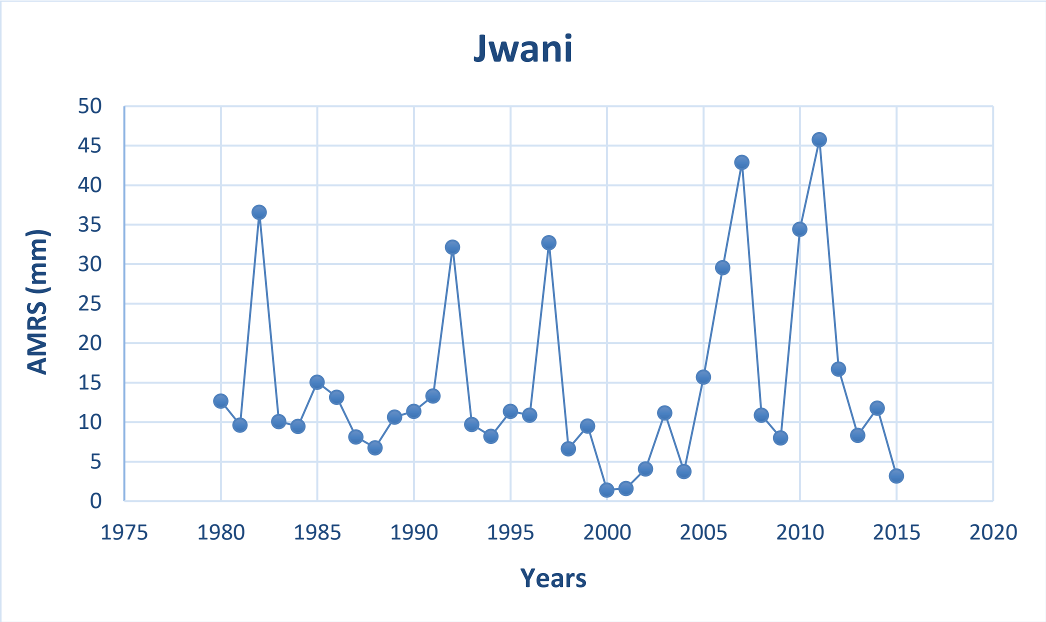

Figure 4.8: Time Series Plot of AMRS for the site Jiwani

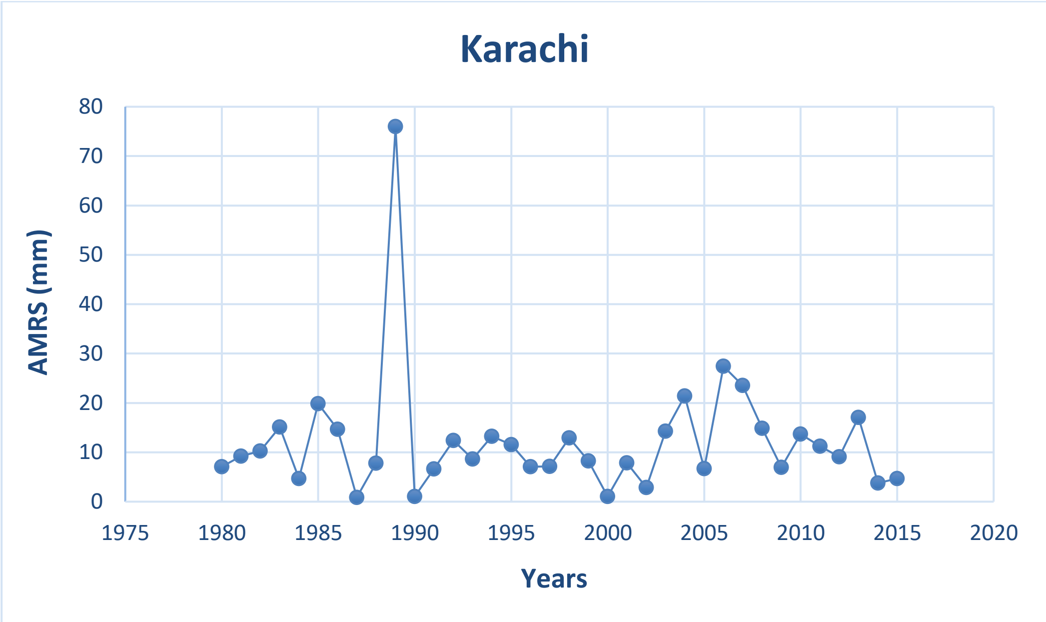

Figure 4.9: Time Series Plot of AMRS for the site Karachi

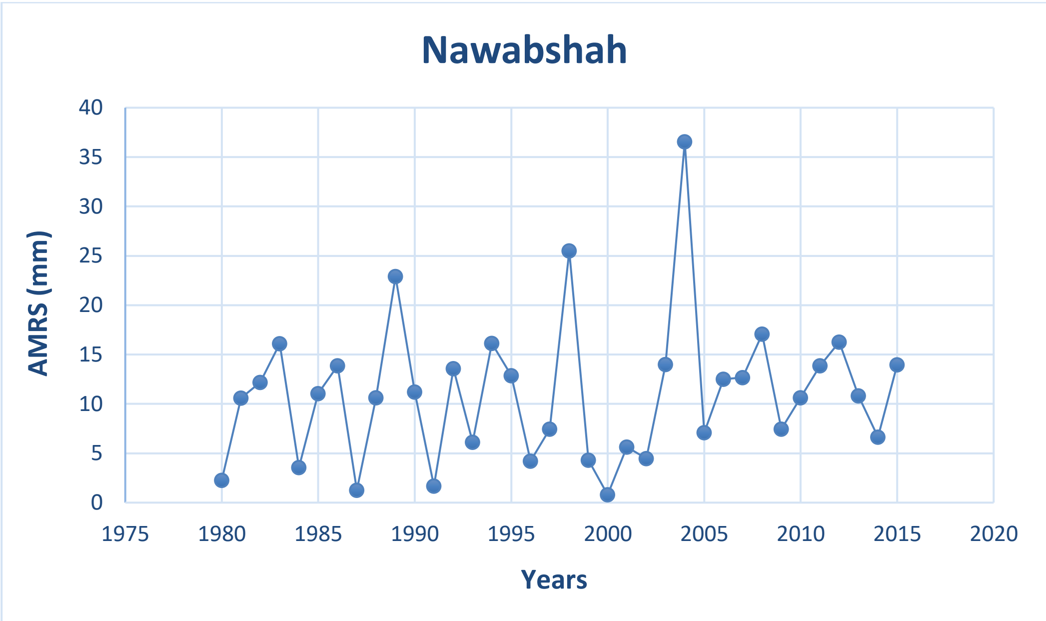

Figure 4.10: Time Series Plot of AMRS for the site Nawabshah

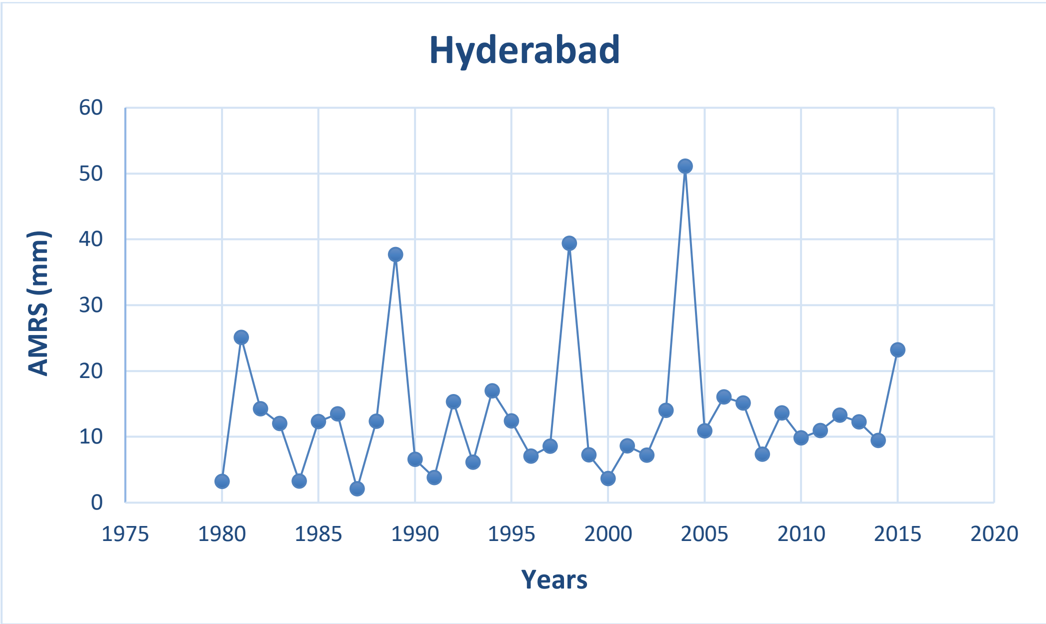

Figure 4.11: Time Series Plot of AMRS for the site Hyderabad

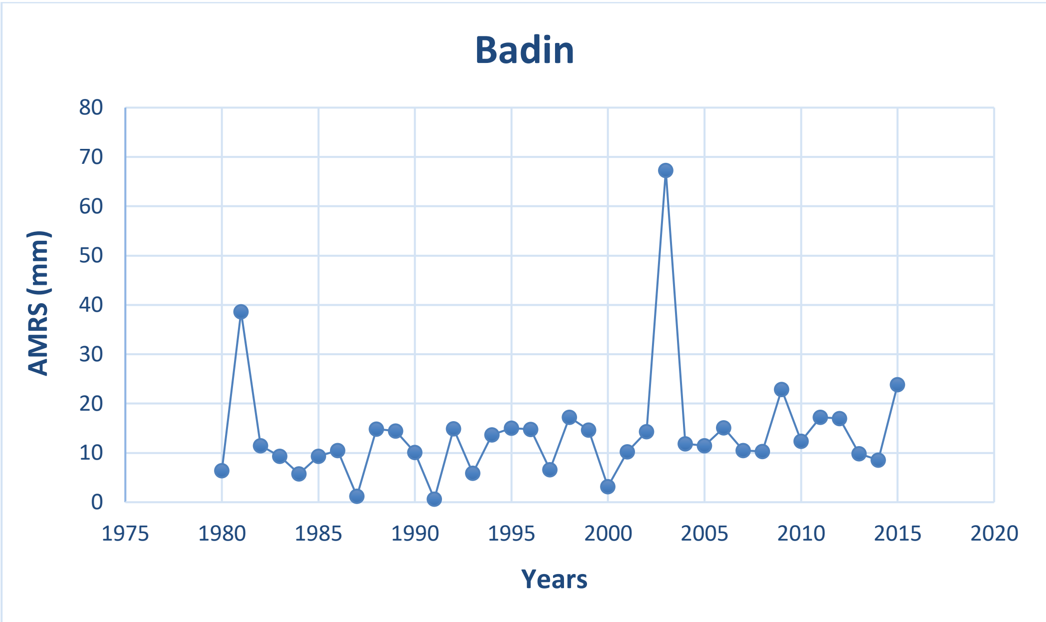

Figure 4.12: Time Series Plot of AMRS for the site Badin

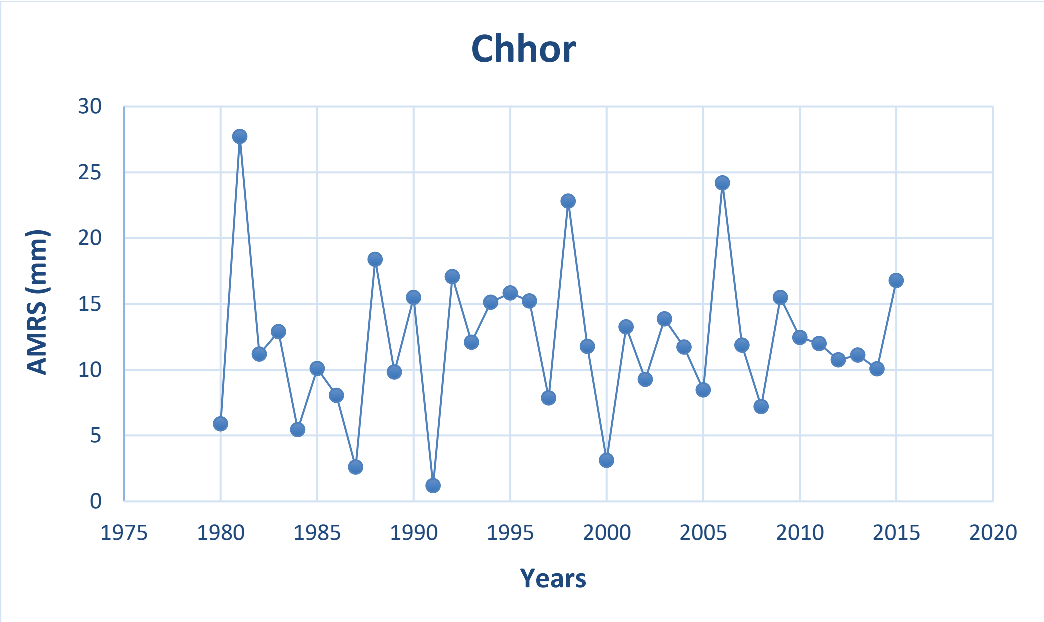

Figure 4.13: Time Series Plot of AMRS for the site Chhor

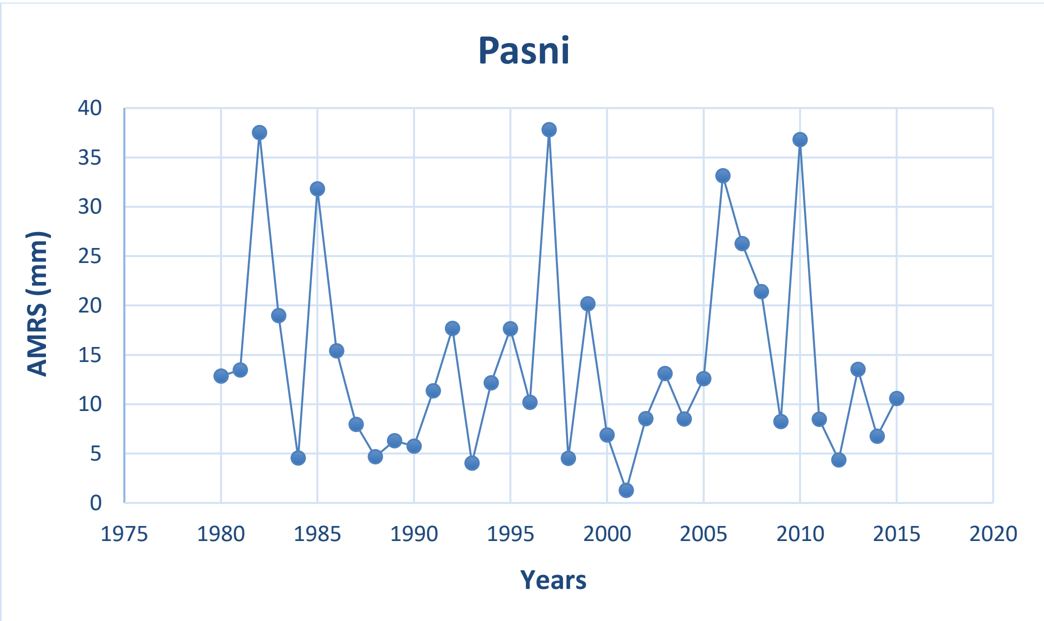

Figure 4.14: Time Series Plot of AMRS for the site Pasni

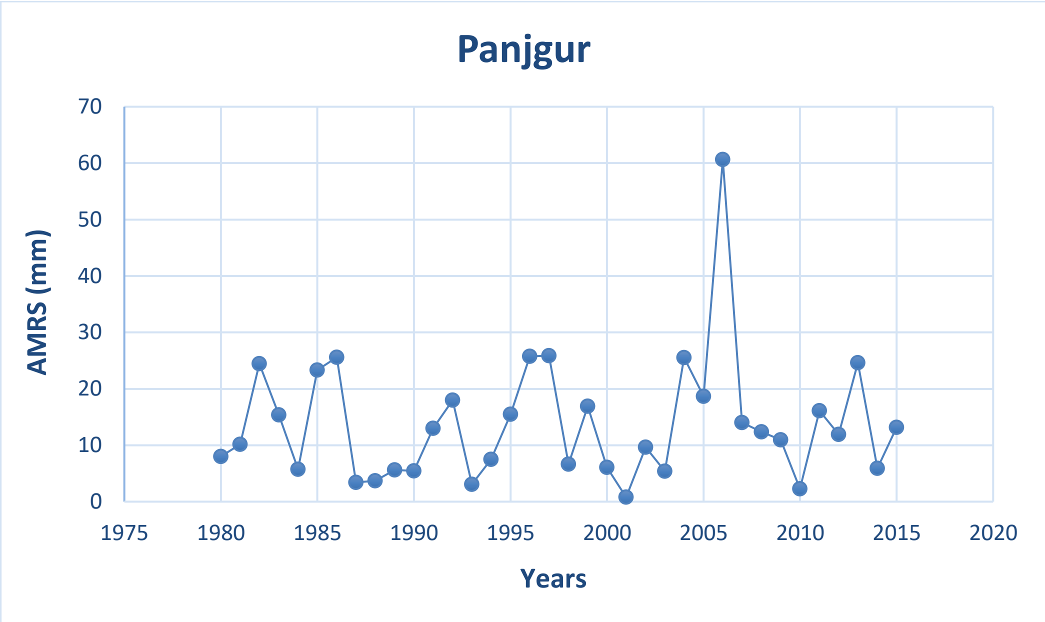

Figure 4.15: Time Series Plot of AMRS for the site Panjgur

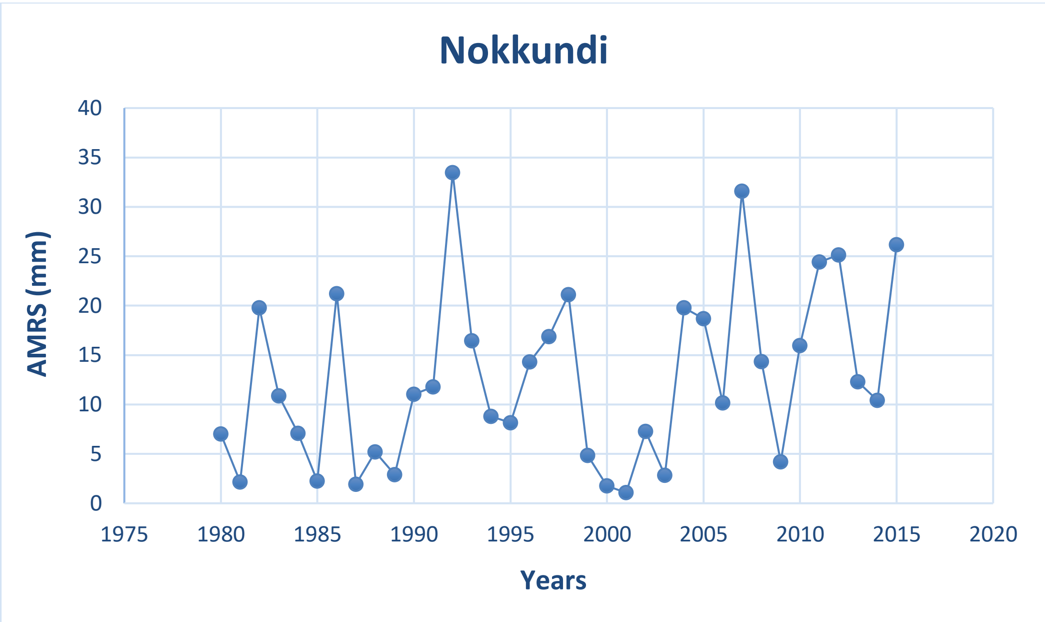

Figure 4.16: Time Series Plot of AMRS for the site Nokkundi

4.3 Model Fitting and Methodological Comparison

This detailed analysis highlights the differences in estimation methods across the stations providing insight into which method might yield the most accurate representation of the rainfall data. Table 4.5 and 4.6 (refer to Appendix A) displays the estimated parameters along with their RMSE and bias values. To determine these accuracy metrics, simulation experiments were performed. For each site, 1,000 random samples were drawn from the PE3 distribution, corresponding to the size of the observed data. The PE3 distribution was then fitted to each sample using LM, MLE, and MPS methods. The resulting RMSE and bias were computed from these simulated parameters.

Here's a summarized discussion of the key parameters (location, scale, shape) and the corresponding RMSE and Bias for each estimation method (LM, MLE, MPS) across different stations (with moderate sample size):

Zone D

Sibbi

For the location and scale parameters, the MLE method consistently yields lower RMSE values, suggesting better fit precision, while the LM method maintains a lower bias. The estimates for the shape parameter show that the MPS method produces the lowest RMSE and bias.

Jacobabad

For the location parameters, the MLE method results in both lower RMSE and bias values. For the scale and shape parameters, LM method exhibits lower bias values, whereas the LME and MPS methods achieve lower RMSE.

Bahawalpur

For location and shape parameters, LM and MPS are the most appropriate methods, respectively, as they show the lowest RMSE and bias values. For the scale parameter, LM method yields a significantly lower RMSE, while MLE method results in lower bias.

Khanpur

For location and shape parameters, the MLE and MPS methods exhibit the lowest RMSE and bias values, respectively. For the scale parameter, MLE demonstrates the lowest RMSE, while LM shows the lowest bias.

Multan

For the location and scale parameters, MLE exhibits lower RMSE values, whereas LM demonstrates lower bias values. Conversely, for the shape parameter, LM shows lower bias, while MLE has lower RMSE.

Bahawalnagar

For the three parameters location, scale and shape, MLE significantly reveals low RMSE values, while LM presents low bias values.

Barkhan

For the location parameter, LM shows low RMSE, while MLE has low bias. For the scale parameter, LM unveils low values for both bias and RMSE. For the shape parameter, MPS shows low RMSE, whereas LM has low bias. Overall, LM is the most appropriate method for this station.

Zone E

Hyderabad

For the three parameters location, scale, and shape, MLE yields the lowest RMSE values, while LM produces the smallest Bias values, indicating that these are the most appropriate methods for this station.

Karachi

For the three parameters, LM reveals the lowest Bias values. However, when considering location, scale, and shape parameters, LM, MPS, and MLE respectively result in the lowest RMSE values.

Nawabshah

At this station, LM consistently produces the lowest RMSE and Bias values for all three parameters, except for the RMSE of the shape parameter, where MPS yields the lowest value.

Jiwani

The estimates for the location parameter show that the LM method produces the lowest RMSE, while the MLE method shows slightly higher RMSE but lower bias, indicating a trade-off between accuracy and consistency. For the scale and shape parameters, the LM method consistently maintains lower Bias values, suggesting better fit precision, while the MPS and MLE methods produces lower RMSE values.

Badin

For location, scale and shape parameters LM method produces lowest Bias values while MLE produces lower RMSE values with MPS revealing slightly low RMSE particularly for shape parameter.

Chor

This station shows a similar pattern where the LM method has the lowest RMSE and bias for all three parameters. However, for the shape parameter, the MPS method provides slightly lower bias.

Pasni

MPS shows lower RMSE values across all three parameters, while LM displays lower Bias values.

Panjgur

For the location, scale, and shape parameters, LM method produces significantly lower Bias and RMSE value, except for the shape parameter, MPS produces a slightly lower Bias value.

Nokkundi

For the location parameter, the LM method produces the lowest RMSE, while the MLE method shows the lowest Bias. For the scale parameter, MLE achieves both the lowest RMSE and Bias values. For the shape parameter, MPS produces the lowest RMSE, and LM shows the lowest Bias.

Overall Results

The above analysis for zone D shows that MLE is generally the best method for achieving lower RMSE, making it most suitable for stations like Sibbi (location and scale), Jacobabad (location), Khanpur (location and scale), Multan (location and scale), and Bahawalnagar (all parameters). LM is preferred for its lower bias, especially in Jacobabad (scale and shape), Bahawalpur (location), Khanpur (scale), and Barkhan (location and scale). MPS is particularly effective for the shape parameter in Sibbi, Bahawalpur, Khanpur, and Barkhan.

The results for zone E indicate that MLE generally provides the lowest RMSE values, making it most suitable for stations like Hyderabad (all parameters), Badin (all parameters), and Nokkundi (scale). LM is preferred for its lower bias, particularly in Hyderabad (all parameters), Karachi (all parameters), Nawabshah (all parameters), Jiwani (location, scale, and shape bias), Badin (all parameters), Chor (all parameters), Pasni (all parameters), Panjgur (all parameters), and Nokkundi (location and shape bias). MPS stands out for shape parameter accuracy, showing lower RMSE in Karachi, Nawabshah, Jiwani, Badin, Chor, Pasni, Panjgur, and Nokkundi.

Overall, MLE emerges as the most suitable method across most stations with moderate sample size for achieving lower RMSE, while LM is favored for bias reduction. MPS is particularly effective and ideal for the shape parameter estimation.

4.4 Preference Table

The preference table is designed to provide a clear overview of how different stations behave in terms of skewness and tail characteristics. For each station, it records the level of skewness and the behavior of the data's tails (whether they are heavy or light). Based on these characteristics, the table also identifies the most appropriate parameter estimation methods for each station, specifically when working with a moderate sample size and variations in shape parameter. This approach helps in selecting the best statistical methods tailored to the unique data distribution of each station, ensuring more accurate and reliable analysis.

The analysis of skewness and kurtosis for the stations in table 4.7 suggests that the most appropriate method for data fitting varies depending on the specific characteristics of the data distribution. For stations with very high skewness and leptokurtic distributions, the Maximum Product Spacing (MPS) method consistently emerges as the best fit, as seen in Khanpur. Conversely, stations with moderate skewness and leptokurtic distributions, like Multan, typically favor the Linear Moments (LM) method, while those with moderate skewness and platykurtic kurtosis, such as Nokkundi, also tend to favor LM. In cases of high skewness combined with leptokurtic distributions, the combination of MPS and LM is frequently preferred, as observed in Jacobabad and Nawabshah. Therefore, the choice of method is highly dependent on the nature of skewness and kurtosis in the data, with MPS being particularly effective for extreme conditions, while LM and Maximum Likelihood Estimation (MLE) methods are more suited for moderate conditions.

Table 4.7: Tail behavior and shape-focused approach

Sr no.	Stations	Skewness	Interpretation of Skewness	Kurtosis	Interpretation of Kurtosis	Preferred Method	
						RMSE	Bias
1	Multan	0.56	Moderate Skewness	1.38	Leptokurtic	LM	MLE
2	Khanpur	3.93	Very High Skewness	31.52	Leptokurtic	MPS	MPS
3	Jacobabad	1.72	High Skewness	3.32	Leptokurtic	MPS	LM
4	Sibbi	2.86	Very High Skewness	11.22	Leptokurtic	MPS	MPS
5	Bahawalpur	3.05	Very High Skewness	13.43	Leptokurtic	MPS	MPS
6	Bahawalnagar	2.56	Very High Skewness	6.96	Leptokurtic	MLE	LM

7	Barkhan	1.44	High Skewness	2.64	Leptokurtic	MPS	LM
8	Jiwani	1.43	High Skewness	1.08	Leptokurtic	MLE	LM
9	Karachi	3.92	Very High Skewness	19.36	Leptokurtic	MLE	LM
10	Nawabshah	1.32	High Skewness	3.32	Leptokurtic	MPS	LM
11	Hyderabad	2.12	Very High Skewness	4.98	Leptokurtic	MLE	LM
12	Badin	3.29	Very High Skewness	14.02	Leptokurtic	MPS	LM
13	Chhor	0.58	Moderate Skewness	1.02	Leptokurtic	LM	MPS
14	Pasni	1.16	High Skewness	0.41	Leptokurtic	MPS	LM
15	Panjgur	2.20	Very High Skewness	7.77	Leptokurtic	MPS	LM
16	Nokkundi	0.63	Moderate Skewness	-0.35	Platykurtic	MPS	LM

4.5 Skewness and Kurtosis Levels

For interpreting the skewness and kurtosis values in Table 4.3, Table 4.4 and Table 4.7, the following thresholds have been applied.

Skewness [26]**:**

- **0 to ± 0.5:** Approximately symmetric
- **± 0.5 to ± 1:** Moderately skewed
- **± 1 to ± 2:** Highly skewed
- **± 2:** Very highly skewed

Kurtosis [27]**:**

- **Platykurtic (Kurtosis < 0):** Flat, light tails.
- **Mesokurtic (Kurtosis ≈ 0):** Normal distribution, moderate tails.
- **Leptokurtic (Kurtosis > 0):** Sharp peak, heavy tails.

4.6 Quantile Estimation

The table 4.8 presents a comparison of three estimation methods LM (L-moments), MLE (Maximum Likelihood Estimation), and MPS (Maximum Product of Spacings) across seven sites in Zone D for a 100-year return period. Quantile values show that the methods produce relatively consistent estimates, with some variability. In Zone D, LM emerges as the best method based on both Bias and RMSE, offering more accurate estimates for most stations. For example, in stations like Multan and Jacobabad, LM shows the smallest Bias, making it the most accurate method. Even when RMSE values are slightly higher than MLE or MPS in a few cases, the much lower Bias of LM ensures better reliability. LM also maintains relatively low RMSE values across the sites, with stations like Barkhan and Multan favoring LM in terms of both Bias and RMSE, highlighting its balanced performance. However, in stations like Sibbi and Bahawalnagar, while LM provides lower Bias, MLE shows lower RMSE, making it slightly better in terms of error reduction in these specific sites.

In table 4.9, the quantile estimates reflect a similar trend where the LM method generally produces lower bias compared to MLE and MPS. Based purely on RMSE, MLE outperforms LM in some stations, such as Badin and Hyderabad, where it exhibits slightly lower RMSE values. In these cases, MLE might be preferable for reducing errors, although it tends to have higher Bias compared to LM. This means that while MLE can reduce RMSE, the higher Bias suggests a trade-off in accuracy. In Nokkundi and Nawabshah clearly LM shows significant low Bias and RMSE. Overall, LM remains the best method across Zone E for both Bias and RMSE combined, but MLE offers advantages in reducing RMSE at selected stations.

Table 4.8: Quantile estimation for return period 100

Zone D					
Sr. no	**Site Name**	**Method**	**LM**	**MLE**	**MPS**
		Return Period (years)	100	100	100
1-	Multan	Quantile	23	22	23
		Bias	0.02	0.21	0.83
		RMSE	2.31	2.39	2.69
2-	Khanpur	Quantile	31	28	30

		Bias	0.12	1.07	1.72
		RMSE	5.67	4.61	5.04
3-	Jacobabad	Quantile	36	32	34
		Bias	0.04	1.72	3.08
		RMSE	7.73	6.32	7.40
4-	Sibbi	Quantile	47	39	43
		Bias	0.26	1.75	3.20
		RMSE	10.13	7.42	8.72
5-	Bahawalpur	Quantile	30	27	29
		Bias	0.09	0.64	1.34
		RMSE	4.69	3.79	4.14
6-	Bahawalnagar	Quantile	60	49	52
		Bias	0.72	2.80	4.20
		RMSE	15.65	9.07	10.96
7-	Barkhan	Quantile	28	28	29
		Bias	0.35	1.31	1.91
		RMSE	4.56	4.78	4.84

Table 4.9: Quantile estimation for return period 100

Zone E					
Sr. no	**Site Name**	**Method**	**LM**	**MLE**	**MPS**
		Return Period (years)	100	100	100
1-	Jiwani	Quantile	59	54	56
		Bias	0.39	4.37	4.89
		RMSE	14.10	12.46	12.77

2-	Karachi	Quantile	53	54	52
		Bias	0.10	2.79	5.59
		RMSE	12.79	13.04	12.79
3-	Nawabshah	Quantile	32	34	36
		Bias	0.06	1.42	2.63
		RMSE	4.89	6.59	7.37
4-	Hyderabad	Quantile	51	47	50
		Bias	0.17	2.98	4.26
		RMSE	11.20	10.06	10.92
5-	Badin	Quantile	49	44	48
		Bias	0.40	1.85	3.47
		RMSE	11.02	8.52	10.14
6-	Chhor	Quantile	27	27	29
		Bias	0.25	0.10	1.42
		RMSE	3.48	3.74	4.30
7-	Pasni	Quantile	52	48	50
		Bias	0.24	2.04	4.25
		RMSE	10.98	9.56	11.24
8-	Panjgur	Quantile	50	50	52
		Bias	0.32	3.08	4.28
		RMSE	10.42	10.84	11.46
9-	Nokkundi	Quantile	40	51	50
		Bias	0.32	3.20	4.32
		RMSE	6.68	11.76	11.59

5. CONCLUSIONS AND FUTURE RECOMMENDATIONS

5.1 Key Findings and Strengths

The key findings from the analysis illuminate the nature of extreme rainfall events in Pakistan and guide the selection of statistical methods for accurate analysis. The data analysis reveals that these events are characterized by significant skewness and kurtosis, indicating a highly asymmetric rainfall distribution with heavy tails. This suggests that extreme rainfall events are frequent and exhibit considerable variability, complicating prediction and management efforts. The study also identifies a trend of increasing intensity in these events over time, which is crucial for understanding the potential impacts of climate change on future rainfall patterns in the region.

In terms of statistical method selection, the study finds that the Maximum Likelihood Estimation (MLE) method emerges as the most suitable approach across most stations with a moderate sample size for achieving lower Root Mean Square Error (RMSE). This makes MLE the preferred method when the goal is to minimize overall prediction error. However, for bias reduction, the Linear Moments (LM) method is favored, offering better performance in reducing estimation bias. Additionally, the Maximum Product of Spacings method for the Shape parameter (MPS) proves particularly effective for accurately estimating the shape parameter in several locations. These findings are essential as they provide clear guidance on the best statistical method to use based on the sample size and shape characteristics of the data, thereby enhancing the accuracy and reliability of extreme rainfall analysis.

The strength of this study lies in its detailed and methodical statistical analysis. By focusing on ungauged sites and using advanced measures like skewness and kurtosis, coupled with an evaluation of different estimation methods, the research offers a nuanced understanding of rainfall distribution that goes beyond simple averages. This approach allows for a more accurate assessment of the risks posed by extreme weather events, which is essential for effective disaster preparedness and response. Additionally, the study's findings contribute to the broader field of climate science by offering insights that could inform future modeling efforts and policy decisions aimed at mitigating the impacts of extreme weather.

5.2 Limitations and Future Recommendations

However, the study is constrained by several limitations. The analysis was conducted on a fixed sample size of 36, with data drawn from only a few stations in Pakistan. This limited scope may restrict the generalizability of the findings to other regions within the country. Furthermore, the focus on ungauged sites means that the study may not fully capture the complete variability of rainfall across Pakistan, potentially overlooking important patterns present in gauged locations. Another limitation is the exclusive use of historical rainfall data, which may not be adequately available.

To overcome these limitations, future research should aim to expand the spatial and temporal scope of the analysis. Increasing the number of stations included in the study and extending the time span of the data would provide a more comprehensive picture of rainfall patterns across Pakistan. Additionally, integrating climate model projections could help predict how extreme rainfall events might change in frequency and intensity under different climate scenarios. Expanding the analysis to include gauged sites and regions with diverse climatic conditions would also enhance the robustness of the findings, making them more applicable to a wider range of settings. These steps would not only strengthen the current study but also contribute to a deeper understanding of extreme rainfall events on a global scale.

REFERENCES

[1] Salma, S., Rehman, S., & Shah, M. A. (2012). Rainfall trends in different climate zones of Pakistan. Pakistan Journal of Meteorology, 9(17).

[2] Hassan, A. The Socioeconomic Impacts of Heatwave and its Alternatives for Pakistan.

[3] Katz, R. W., Parlange, M. B., & Naveau, P. (2002). Statistics of extremes in hydrology. Water Resources Research, 25(8-12), 1287–1304. DOI: 10.1016/s0309-1708(02)00056-8

[4] Wong, T. S. T., & Li, W. K. (2006). A note on the estimation of extreme value distributions using maximum product of spacings. Lecture Notes-Monograph Series, 52, 272–283. DOI: 10.2307/20461444

[5] Soukissian, T., & Tsalis, C. (2015). The effect of the generalized extreme value distribution parameter estimation methods in extreme wind speed prediction. Natural Hazards, 78. DOI: 10.1007/s11069-015-1800-0

[6] Khan, M.S.u.R., Hussain, Z. & Ahmad, I. Effects of L-Moments, Maximum Likelihood and Maximum Product of Spacing Estimation Methods in Using Pearson Type-3 Distribution for Modeling Extreme Values. Water Resour Manage 35, 1415–1431 (2021). DOI: 10.1007/s11269-021-02767-w

[7] Vivekanandan, N., & Shukla, S. (2015). Flood frequency analysis using method of moments and L-moments of probability distributions. Cogent Engineering, 2(1). DOI: 10.1080/23311916.2015.1018704

[8] Drissia, T.K., Jothiprakash, V. & Anitha, A.B. Flood Frequency Analysis Using L Moments: a Comparison between At-Site and Regional Approach. Water Resour Manage 33, 1013–1037 (2019). DOI: 10.1007/s11269-018-2162-7

[9] Arns, T. Wahl, I.D. Haigh, J. Jensen, C. Pattiaratchi, Estimating extreme water level probabilities: A comparison of the direct methods and recommendations for best practise, Coastal Engineering, Volume 81, 2013, Pages 51-66, ISSN 0378- 3839. DOI: 10.1016/j.coastaleng.2013.07.003.

[10] Palutikof, J. P., Brabson, B. B., Lister, D. H., & Adcock, S. T. (1999). A review of methods to calculate extreme wind speeds. Meteorological Applications, 6(2), 119–132. DOI: 10.1017/S1350482799001103

[11] J. R. M. Hosking, L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics, Journal of the Royal Statistical Society: Series B (Methodological), Volume 52, Issue 1, September 1990, Pages 105–124. DOI: 10.1111/j.2517-6161.1990.tb01775.x

[12] Mat Jan, N.A., Shabri, A. Estimating distribution parameters of annual maximum streamflows in Johor, Malaysia using TL-moments approach. Theor Appl Climatol 127, 213–227 (2017). DOI: 10.1007/s00704-015-1623-7

[13] Koutrouvelis, I. A., and G. C. Canavos (1999), Estimation in the Pearson type 3 distribution, Water Resour. Res., 35(9), 2693–2704. DOI: 10.1029/1999WR900174

[14] Lei, GJ, Yin, JX., Wang, WC. et al. The Analysis and Improvement of the Fuzzy Weighted Optimum Curve-Fitting Method of Pearson – Type III Distribution. Water Resour Manage 32, 4511–4526 (2018). DOI: 10.1007/s11269-018-2055-9

[15] Li, W., Zhou, J., Sun, H. et al. Impact of Distribution Type in Bayes Probability Flood Forecasting. Water Resour Manage 31, 961–977 (2017). DOI: 10.1007/s11269-016-1557-6

[16] Hussain, Z., Shahzad, M.N. & Abbas, K. Application of regional rainfall frequency analysis on seven sites of Sindh, Pakistan. KSCE J Civ Eng 21, 1812–1819 (2017). DOI: 10.1007/s12205-016-0946-y

[17] Rutkowska, A., Żelazny, M., Kohnová, S., Łyp, M., Banasik, K. (2018). Regional L-Moment-Based Flood Frequency Analysis in the Upper Vistula River Basin, Poland. In: Niedzielski, T., Migała, K. (eds) Geoinformatics and Atmospheric Science. Pageoph Topical Volumes. Birkhäuser, Cham. DOI: 10.1007/978-3-319-66092-9_13

[18] Pearson Karl 1895X. Contributions to the mathematical theory of evolution.—II. Skew variation in homogeneous material Philosophical Transactions of the Royal Society of London. (A.)186343–414 DOI: 10.1098/rsta.1895.0010

[19] Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). Continuous univariate distributions, volume 2 (Vol. 289). John wiley & sons.

[20] Chow VT (1954) The log-probability law and its engineering applications. Proc Am Soc Civil Eng 80:1–25

[21] Kotz, S. (1970). Distributions in statistics: Continuous univariate distributions. John Wiley & Sons.

[22] Bobee, B., & Ashkar, F. (1991). The Gamma Family and Derived Distributions Applied In hydrology.

[23] Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. Journal of the Royal Statistical Society: Series B (Methodological), 52(1), 105-124. DOI: 10.1111/j.2517-6161.1990.tb01775.x

[24] Fisher, R. A. "On the mathematical foundations of theoretical statistics." Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character 222 (1922): 309-368.

[25] Ranneby, B. (1984). The Maximum Spacing Method. An Estimation Method Related to the Maximum Likelihood Method. Scandinavian Journal of Statistics, 11(2), 93–112. DOI: 10.2307/4615946

[26] Bulmer, M. G. (1979). Principles of Statistics. Dover Publications.

[27] Joanes, D. N., & Gill, C. A. (1998). Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(1), 183-189. DOI:10.1111/1467-9884.00122

Appendix A

Table 4.5: Estimated parameters alongwith their RMSE and bias

Zone D											
	ESTIMATION METHODS		**LM**			**MLE**			**MPS**		
Sr. no	SITES	PARAMETER	LOCATION	SCALE	SHAPE	LOCATION	SCALE	SHAPE	LOCATION	SCALE	SHAPE
1	SIBBI	ESTIMATE	13.59	9.04	2.16	13.59	8.05	1.28	14.01	8.97	1.30
		RMSE	1.55	2.31	0.55	1.27	1.56	0.42	1.52	1.73	0.38
		BIAS	2.0202	0.0844	0.0188	0.0361	0.2370	0.1379	0.3072	0.6496	0.0150
2	JACOBABAD	ESTIMATE	10.04	7.02	2.10	10.04	6.42	1.57	10.36	6.96	1.54
		RMSE	1.23	1.67	0.53	1.12	1.29	0.34	1.22	1.40	0.33
		BIAS	0.0235	0.0531	0.0313	0.0123	0.2779	0.1634	0.3213	0.6173	0.0660
3	BAHAWALPUR	ESTIMATE	13.79	4.73	1.81	13.79	4.51	0.93	13.96	5.05	0.96
		RMSE	0.78	1.01	0.50	0.78	0.73	0.47	0.86	0.84	0.43
		BIAS	0.0041	0.0529	0.0094	0.0140	0.0334	0.0878	0.1261	0.3332	0.0008
4	KHANPUR	ESTIMATE	12.08	5.52	1.78	12.08	5.09	1.10	12.29	5.64	1.10
		RMSE	0.96	1.20	0.51	0.89	0.95	0.43	0.96	0.98	0.41
		BIAS	0.0041	0.0007	0.0644	0.0020	0.1247	0.1262	0.2070	0.4476	0.0328
5	MULTAN	ESTIMATE	12.62	3.70	0.55	12.63	3.75	0.33	12.67	4.12	0.32
		RMSE	0.67	0.48	0.46	0.60	0.47	0.53	0.71	0.53	0.49
		BIAS	0.0097	0.0188	0.0342	0.0106	0.0424	0.0104	0.0098	0.2797	0.0142
6	BAHAWALNAGAR	ESTIMATE	17.85	10.66	2.86	17.85	8.92	1.73	18.38	9.75	1.73
		RMSE	1.77	3.24	0.62	1.54	1.94	0.31	1.68	2.18	0.32
		BIAS	0.0702	0.3076	0.0228	0.0831	0.4512	0.1844	0.5094	0.9852	0.0855
7	BARKHAN	ESTIMATE	12.75	4.58	1.61	12.76	4.48	1.50	12.95	4.80	1.45
		RMSE	0.74	0.88	0.46	0.77	1.00	0.36	0.83	0.96	0.34
		BIAS	0.0181	0.0497	0.0133	0.0002	0.2450	0.1844	0.2042	0.4479	0.0815

Table 4.6: Estimated parameters alongwith their RMSE and Bias

Zone E											
	ESTIMATION METHODS		**LM**			**MLE**			**MPS**		
Sr. no	SITES	PARAMETER	LOCATION	SCALE	SHAPE	LOCATION	SCALE	SHAPE	LOCATION	SCALE	SHAPE
1	HYDERABAD	ESTIMATE	13.50	10.19	2.11	13.50	9.72	1.67	14.01	10.46	1.63
		RMSE	1.77	2.43	0.53	1.71	2.09	0.31	1.76	2.17	0.34
		BIAS	0.0140	0.0780	0.0267	0.0872	0.5318	0.1981	0.4742	0.9073	0.0639
2	KARACHI	ESTIMATE	12.24	10.74	2.34	12.13	11.59	2.06	12.82	11.14	1.76
		RMSE	1.78	2.74	0.58	2.04	2.76	0.26	2.01	2.50	0.32
		BIAS	0.0236	0.1746	0.0041	0.0287	0.4928	0.1339	0.5580	1.1174	0.0862
3	NAWABSHAH	ESTIMATE	11.05	7.01	0.89	11.05	7.33	1.23	11.34	7.92	1.18
		RMSE	1.11	1.04	0.45	1.21	1.40	0.43	1.33	1.44	0.40
		BIAS	0.0202	0.0597	0.0113	0.0902	0.1428	0.13894	0.3393	0.6483	0.0217
4	JIWANI	ESTIMATE	14.66	11.99	2.16	14.67	11.50	1.70	15.13	12.01	1.62
		RMSE	1.96	2.76	0.52	1.98	2.48	0.30	2.03	2.46	0.33
		BIAS	0.0842	0.1279	0.0375	0.0080	0.7320	0.1908	0.5289	1.0987	0.0739
5	BADIN	ESTIMATE	13.93	9.78	2.05	13.93	9.32	1.27	14.42	10.42	1.30
		RMSE	1.59	2.32	0.53	1.52	1.74	0.41	1.73	1.89	0.39
		BIAS	0.0010	0.0933	0.0310	0.1415	0.1954	0.1339	0.2835	0.7713	0.0401
6	CHOR	ESTIMATE	12.1675	5.5838	0.4922	12.1675	5.5584	0.4396	12.26	6.14	0.44
		RMSE	0.95	0.72	0.47	0.95	0.74	0.53	1.00	0.84	0.49
		BIAS	0.0166	0.0266	0.0190	0.0589	0.0324	0.0438	0.0852	0.4746	0.0135
7	PASNI	ESTIMATE	14.32	10.75	1.76	14.32	10.09	1.49	14.66	10.70	1.43
		RMSE	1.86	2.32	0.50	1.80	2.19	0.37	1.79	2.09	0.36
		BIAS	0.0211	0.0092	0.0498	0.0723	0.5097	0.1876	0.4312	0.9670	0.0708
8	PANJGUR	ESTIMATE	13.92	10.69	1.58	13.92	10.58	1.56	14.39	11.35	1.52
		RMSE	1.77	2.06	0.48	1.83	2.31	0.37	1.99	2.39	0.34
		BIAS	0.0319	0.0028	0.0593	0.1244	0.4808	0.1821	0.5090	1.1071	0.0737
9	NOKKUNDI	ESTIMATE	12.59	9.16	0.94	12.59	10.89	1.88	12.95	10.69	1.66
		RMSE	1.52	1.39	0.48	1.89	2.45	2.45	1.80	2.26	0.31
		BIAS	0.0721	0.1426	0.0037	0.0182	0.6799	0.1773	0.4180	0.9578	0.0873