Quantification of Risk Associated with Climate Change by Analyzing Extremes of Rainfall for Various Met Stations of Pakistan

By

Muhammad Hassnain

00000205008

Supervisor

Dr. Zamir Hussain

Research Center for Modelling and Simulation (RCMS)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

August 2021

Quantification of Risk Associated with Climate Change by Analyzing Extremes of Rainfall for Various Met Stations of Pakistan

By

Muhammad Hassnain

00000205008

Supervisor

Dr. Zamir Hussain

A thesis submitted in conformity with the requirements for the degree of *Master of Science* in Computational Science and Engineering

Research Center for Modelling and Simulation (RCMS)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

August 2021

Declaration

I, *Muhammad Hassnain* declare that this thesis titled "Quantification of Risk Associated withClimate Change by Analyzing Extremes of Rainfall for Various Met Stations of Pakistan" and the work presented in it are my own and has been generated by me as a result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a Master of Science degree at NUST
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at NUST or any other institution, this has been clearly stated
3. Where I have consulted the published work of others, this is always clearly attributed
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work
5. I have acknowledged all main sources of help
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself

Muhammad Hassnain,

00000205008

Copyright Notice

This thesis is dedicated to *my beloved parents*

Abstract

Climate is a vast phenomenon that includes the study of weather, atmospheric variables and external factors. Some of the variables that are commonly used are precipitation, atmospheric pressure, humidity, and temperature. To study the climate of any region, usually decades of data are considered by researchers. Regarding variables that impact the environment, extreme variables play a vital role in assessing and making predictions about the climate. To be more specific, these extreme variables have proved to be very effective for studying and predicting climate hazards. Pakistan has suffered dramatically due to repeated floods in the last two decades. An immediate need to initiate research efforts to investigate and precisely predict climate hazards like floods thoroughly. Any research effort in this domain would help save life losses and stabilize the country's economy. As extreme rainfall has proved to be highly correlated with floods, in this research, we have presented the statistical analysis of extreme variables, i.e., intense rainfalls, to predict the floods. A thorough study of extreme rainfall is presented in this research. The current study provides an in-depth analysis of the extremes of rainfall of 14 met stations/observatories of the four provinces and Gilgit-Baltistan of Pakistan.

Acknowledgments

I would like to express sincere appreciation for my supervisor, Dr. Zamir Hussain, who has the confidence and the build of a great academic. He persistently conveyed a sense of undertaking in research and a passion in regards to personal triumph. Without his guidance and tenacious help, this thesis would not have been conceivable.

I would like to appreciate the efforts of my committee members, Dr. Adnan Maqsood and Ammar Mushtaq, whose direction and criticism established the importance of literature in modern technology for a proficient and well written thesis.

In addition, a thank you to The Principal RCMS, Dr. Rizwan Riaz for providing reading material and research facilities at RCMS, NUST for the ease of students. I also appreciate the resolute camaraderie of my fellow scholars and their endless motivation and debate on current research and research methodology.

Most importantly, I thank my parents for their love and support. Without their efforts to provide a comfortable environment for me, I would not have reached this pinnacle.

Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

Abbreviations

GEV	Generalized Extreme Value
GLO	Generalized Logistic
IG	Inverse Gausian
LP III	Log Pearson III
KS	Kolmogrov Smirnov
AD	Anderson Darling
KS	Khi Square

Chapter 1

Introduction

1.1 Climate Change

Climate change refers to the long-term changes in the climate that occur over time. Environmental change alludes to the long-haul modifications in the atmosphere over decades, hundreds of years, or more [3]. In this regard, the climate system is maintained by balancing energy from the sun and the energy emitted by the system to outer space. The energy added to the climate system of the earth is vital to the earth's weather. In turn, the average of weather (along with other factors such as time and geography) makes the planet's climate. Climate change can be decided through internal or external factors[4]. The external factors include variations in solar radiation, change in the earth's orbit, etc. Internal factors include the greenhouse effect, inertia of glaciers and large water bodies, etc. These factors are called climate forcing factors. It is a historical fact that the earth's climate system did not remain constant [5]. From the perspective of climate, the change is stochastic [5]. When weather pattern follows the pattern of heavy rainfall, it leads to flooding.

Climate variability and change, its impacts and vulnerabilities are growing concern worldwide[6]. These changes occur by air pollution as human activities cause ozone depletion. These heat-trapping gases are warming the earth. Consequently, we have rising ocean levels, frequent hurricanes, changes in precipitation, melting of glaciers, increasingly outrageous heat occasions, heatwave, and dry season. These changes directly or indirectly affect humans and other living beings, foundations, greenery, freshwater supplies, coastlines, and marine ecosystem.

1.2 Climate Change in Pakistan

Pakistan is a country vulnerable to natural disasters.There are indications that Pakistan has had its share of the large climatic variations[7]. Therefore, there is a need to understand the trends and tendencies of these disasters and adequately forecast the magnitude and frequency of these natural disasters to minimize the risks associated with them.Pakistan has suffered a devastating flood disaster in 2010 [8]. Floods are a major disastrous event in Pakistan and are significantly correlated to rainfall extremes [3].Pakistan is a land where rainfall not only falls in summer but also in winter.[9] Therefore, it is essential to analyze and predict the behavior of rainfall extremes in fourteen stations of the country. Accurate rainfall estimates are also necessary for water resource management, designing, constructing reservoirs and culverts, etc.

Pakistan being a developing country, with approximately 70% of its economy, is dependent on agriculture [5]. Most of the water for agriculture is acquired through rainfall. About 60 percent of rainfall in Pakistan is in summer [5]. Frequency analysis estimates how often a specified event will happen. Heavy rainfall is a significant cause of floods in Pakistan. Therefore, it is essential to analyze the patterns of extreme rainfall.

1.3 Problem Statement

Pakistan is a country exposed to natural disasters. Therefore, there is a need to understand the trends and tendencies of these adversities and adequately forecast the magnitude and frequency of these natural disasters to minimize the risks associated with them. Floods are major disastrous events in Pakistan 1.1, and these are significantly correlated with the extremes of rainfall. Therefore, there is a need to assess the magnitude and frequency associated with rains. It is also essential to analyze and predict weather extremes for water resource management, designing and construction of reservoirs and culverts, etc. The current study provides an in-depth analysis of the extremes of rainfall of fourteen stations of Pakistan's four provinces and Gilgit-Baltistan.

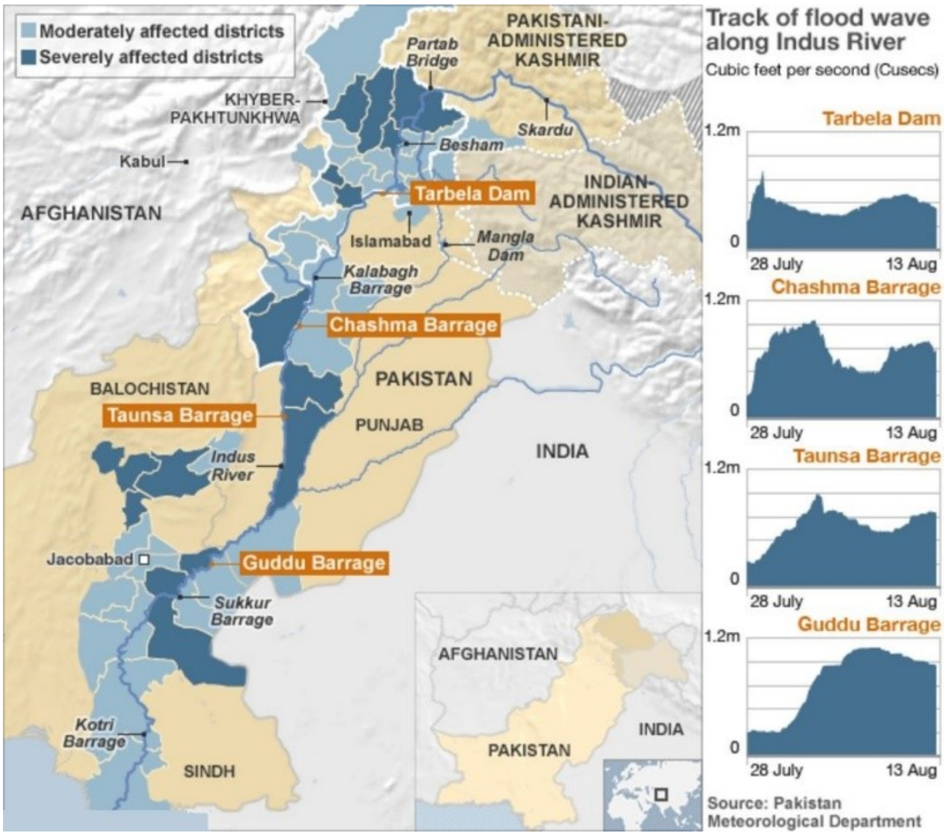


Figure 1.1: Track of flood wave across Indus river [1]

1.4 Objectives

Keeping in view the details mentioned above, the following are the objectives of the current study.

- To provide an in-depth analysis of rainfall extremes by providing analysis of complete series to compare pre and post-climate change era (For instance, the year 2000 may be used as a benchmark).

- To compute accurate and reliable estimates of rainfall extremes for various return periods (10, 20, 50, and 100 years return period) using the probability distributions approach. The proposed methods include both at-site (Rao and Hameed, 2000) and regional methods of analysis (Hosking and Wallis, 1997). The methodology will be adopted as per the tendencies of the recorded data series at various sites. Robust methods will be used for the estimation of parameters of the proposed model(s) (probability distribution(s)).

1.5 Scope

This analysis will provide helpful information regarding the trends and tendencies of the extremes of rainfall in Pakistan. The comparison will give a better understanding of whether there is a shift in the pattern of extremes of rains after the year 2000 or not. This analysis also compares the trends and tendencies concerning the spatial characteristics, as there is diversity in the climate of various provinces of Pakistan. This information will be helpful for the policymakers, scientists, and researchers dealing with water resource management, irrigation, feasibility studies of new hydraulic structures, etc. Scientists can use this analysis to formulate climate change procedures. This analysis is also very important for water resource management. This analysis will be helpful for feasibility studies of up-gradation of existing and planning/construction of new hydraulic structures.

1.6 Motivation

Agriculture contributes to about 21% of Pakistan's Gross domestic product and utilizes about 45% of the workforce[10] . In Pakistan, Punjab being the farming region known for its wheat and cotton production[11]. For the most part, Mango plantations are found in the Sindh and Punjab regions, making Pakistan the world's fourth most prominent provider of mangoes. Changes in precipitation straightforwardly influence water level, cultivation, and devastation. As indicated by the report of Team on Environmental Change (2010) in Pakistan, the nation is susceptible to various catastrophic events, including twisters, floods, dry season, extreme precipitation, and seismic tremors. Therefore, it is essential to study the variations in extreme events,[12] behaviors, especially rainfall.

Chapter 2

Literature Review

2.1 Regional Frequency Analysis Of Annual Total Rainfall in Pakistan

The atmosphere of any country is assessed based on its climate conditions. Pakistan is situated southwest to Northwest at 23–37-degree north scope and 61-76 degrees east. Pakistan has been affected by cataclysmic events, such as dry spells, floods, and storms. These monstrous distributions are caused commonly and can't be moderated even though the misfortunes could be limited to a few or more degrees by appropriate arranging. The fundamental wellsprings of precipitation in Pakistan are the late spring rainstorm, the western depression, and the thunderstorms. From July to September, there is a lot of rainfall. There is an intermission and dynamic system of waterways and water system distributaries in Pakistan, but there is a poor seepage framework. Floods bring about colossal demolition in the nation. Provincial Recurrence examination is profoundly significant for the development of various hydrological structures in the country.

Discordancy measure dependent on LMs was utilized to screen the ATR. As an issue of reality, exceptionally raised regions of Pakistan get more precipitation; the examination zone was separated into four unique areas. Further to legitimize the homogeneity of these regions, LM based heterogeneity measure (H) was determined for every station utilizing four-parameter Kappa distribution. For every station, best dissemination was found among Pearson type III (PE3), Generalized normal (GNO), Generalized extreme value (GEV), Generalized logistic (GLO), and Generalized Pareto (GPA) appropriation utilizing ZM and LM proportion outline. Provincial quantiles based on best-fit circulation for every area were resolved, and further for robustness, the exactness measures for the evaluated territorial quantiles were determined utilizing Monte Carlo simulations. It was discovered that PE3 was the most reasonable decision for a more extended return period for the initial three stations while for a short return period GNO and GEV. So also, for region IV, GEV was announced as the best fit for lower return periods as long as 20 years, while for a time of 50, 100, 500, and 1000 years, GNO was the best one. A regional frequency investigation has been completed for 30 meteorological observatories over Pakistan. After the satisfaction of initial suppositions, information saw as reasonable for Regional Frequency Investigation. Discordancy measure showed that no station is harsh, and all can be utilized to develop homogeneous districts. The district of 30 stations didn't fulfill the heterogeneity measurement characterized by [23]. Based on mean yearly precipitation and height, four homogeneous areas were framed. The LM proportion outline and ZM PE3 are the most appropriate conveyance decisions for the initial two districts, while GLO is for the third and fourth stations. Based on a reenactment study utilizing RR, BR and AR, it is discovered that PE3 the most appropriate decision for an enormous return period for the initial three districts while GNO and GEV for the low return period. Correspondingly, for the fourth area, GEV is announced as best for period quieter return periods as long as 20 years while GNO for more immense return periods. Assessment of recurrence and extent of ATR might be of extraordinary noteworthiness for making strategies suggestion as to the measures for calamity anticipation and moderation. Provincial quantiles appraisals could likewise be utilized in affordable planning and practical activity of various hydrological structures [13].

2.2 Frequency Analysis of Annual Maximum Rainfall in Monsoon Region

Extraordinary occasions like excessive rainfalls, floods, and tsunamis can take many lives and cause billions of property harm.Estimation of magnitude and frequency of extreme rainfall has immense importance to make decisions about hydraulic structures like spillways, dikes and dams etc[14]. Hydrologists are constantly inspired by estimating the probability of such occasions to develop water-powered structures like dams, repositories, spillways and storm surge barriers, etc.Frequency analysis (FA) provides information for estimating how often a specified event will happen [5]. And so the water can be put away, and lives and cash can be spared. Physical laws are wrong to manage the acquire arbitrariness and vulnerability of outrageous occasions. Statistics techniques recognize the presence of exposure and empower its belongings to be measured. Modeling precipitation has immense significance for developing water-driven structures, just as for rural arranging and power purposes in Pakistan. In light of topographic conditions, there is an extraordinary variety in the atmosphere of Pakistan. Pakistan gets precipitation in summer because of the storm and in winter because of western disturbances. Northern regions get more rainfall, and precipitation diminishes in the southern sites of Pakistan. Storm causes 60% of rain in Pakistan and 40% on account of different causes. Regional frequency analysis has been directed for yearly most extreme precipitation in the Monsoon region of Pakistan utilizing Hosking's LMs strategy [15]. The Monsoon region was partitioned into three stations dependent on at-site qualities. It was seen summed up Generalized regular as best vigorous for quantile estimation of more extended return periods and summed up generalized extreme value for quantiles of low return periods for each of the three stations [14].

2.3 Rainfall Variability for Pakistan

Pakistan is a farming nation primarily dependent on the water system through the Indus water framework. The country has built up the world's most extensive contagious canal network. Monsoon precipitation is the lifesaver of Pakistan that is not crucial for the national power supply. Standing yields water demand yet helps the stores meet the necessities of low stream period in the following 4-5 months. Future projection of climate demonstrates towards the dry land agribusiness for Pakistan. The motivation behind the present examination was to research and break down the changeability in precipitation that happened on past occasions. The precipitation fluctuation coefficient is used to dissect the past time circumstance, and the investigation covers the period from 1960-2009. The consequences of the decadal examination indicated that the high estimation of the changeability coefficient had been seen in Baluchistan (251%), Sindh (247%), and Punjab (208%) stations. Anyway, the yearly investigation demonstrated the expanding pattern of changeability coefficient from North to South in Pakistan. As indicated by the inter-seasonal analysis of fifty years' information (1960-09), variety in the coefficient of changeability was most noteworthy in post-storm and pre rainstorm seasons when contrasted with the winter and rainstorm seasons. These examinations demonstrated that forecasting is a complex activity for the forecasters where the fluctuation is unmistakably high. It uncovered that the more significant part of the northern territories had been protected aside from the post-storm period. At the same time, the southern half has been endured during the time regarding precipitation changeability. Past investigation of precipitation information has indicated a marginally diminishing pattern of precipitation changeability for the northern pieces of the nation. Then again, the circumstance for southern sites of the country is getting better regarding precipitation and temperature [16]. The motivation behind the present examination was to explore and break down the fluctuation in rainfall that happened on past occasions. The precipitation inconstancy coefficient is used to investigate the past information circumstance, and the examination covers the period from 1960-2009. It was uncovered in the investigation that the vast majority of the northern zones had been protected aside from in the post-storm period while the southern half has endured during the time of precipitation changeability. Particularly in south Balochistan, it has been seen that precipitation stayed fluctuating all through, independent of the seasons. The decadal investigation of precipitation changeability indicated that the most elevated estimation of fluctuation coefficient had been seen during the fourth decade of the examination, for example, 1990-99, most likely due to the solid ENSO scenes. Because of high fluctuation in lower some portion of the nation occasional just as extraordinary occasions forecast is troublesome. The outcomes got in this examination may help to research the future situation of precipitation in the district, yet increasingly point by point and zone detailed investigation is as yet required [6].

2.4 Classification of Rainfall Regions

Pakistan is situated from southwest to Northwest at 23-37degree north scope and 61-76-degree east scope. Pakistan has been confronting cataclysmic events, such as dry spells, floods, and storms. This investigation is expected to group precipitation areas in Pakistan. Characterization of precipitation stations is fundamental to comprehend precipitation designs in Pakistan. Precipitation designs have been researched utilizing a factor and bunch investigation procedure by 10-days precipitation parameter. The information used here has been gotten from 32 explicit climate stations of PMD (Pakistan Meteorological Office) for January 1980 to December 2006. The outcomes acquired from factor investigation give three elements, and these three components represent 94.60% of the all-out change. For a superior comprehension of precipitation regions, a bunch examination technique has been applied. The bunching system depends on the Wards method algorithm. Generally, these precipitation areas have been isolated into six gatherings. The limit of the region is dictated by the topology, for example, the Baluchistan level, Indus plain, Hindu Kush, and Himalaya ranges. In this investigation, the identical precipitation stations are recognized by utilizing the factor and group multivariate system over Pakistan. The examinations have demonstrated that the factor and bunch method is beneficial for the regionalization of precipitation systems in Pakistan. This precipitation system is connected to territorial variables, and the examination of this exploration affirms the likenesses in precipitation districts. The progressive bunch technique has accomplished this grouping for precipitation 10-days informational collection of 32 climate stations recently diminished by factor investigation. For this examination, three elements have been obtained with 94.60% that clarifies the all-out change for three components. The main factor is related to the precipitation of early harvest time, winter, and pre-summer. The next element is connected with the rain of summer storms, and the third factor demonstrated a profound association with winter precipitation. The investigation regions are separated into six precipitation districts by bunching various leveled techniques with detailed precipitation designs. This investigation gives a nitty-gritty viewpoint upon the idea of precipitation districts crosswise over Pakistan. The 10-days precipitation length gives us an excellent distributional precipitation design, for example, early, mid, and late 10-days precipitation conduct of all months [1].

2.5 Probability Distributions for Annual Daily Maximum Rainfall in Pakistan

Pakistan being a developing nation, 70% of its economy is reliant on horticulture. Water is a special requirement for the horticultural part, just like human life and other living beings. Pakistan gets precipitation in summer as well as in winter. In summer, rain happens for the most part during the monsoon season (early July to September)[17]. July and August are the pinnacles for a very long time for rainstorm precipitation in Pakistan. The most significant wellspring of water for horticulture overall is precipitation. Practically 60% of the all-out yearly rainfall in Pakistan is because of rain in summer. All Kharif crops are generally subject to the example of monsoon precipitation[18]. Winter precipitation is significant for Rabi crops in the nation. In Pakistan, water is utilized for the most part underway of agribusiness and hydroelectricity[19]. In this examination, at-site frequency investigation (AFA) of a yearly everyday most extreme precipitation (ADMR) arrangement was done utilizing the strategy for linear moments (LMs) and their variations, for example, trimmed linear moments (TLMs) and higher-order linear moments (LH-minutes). The ADMR arrangement we examined covers 28 meteorological observatories across Pakistan as collected from the Pakistan Meteorological Department (PMD).

The essential point of the investigation was to find best-fit (i.e., the most appropriate) likelihood dissemination among the class of different probability distribution. At first extraordinary goodness-of fit (GOF) measures, for example, the Kolmogorov-Smirnov test (KST), Anderson-Sweetheart test (ADT), root mean square error (RMSE), and LM proportion outline (LRD) were applied to decide the best-fit dissemination as well as the best linear estimation technique for AFA. We saw that no single likelihood dissemination could be proclaimed as the best-fit distribution for every station. Five distribution was seen as the most fitting: summed up generalized extreme value (GEV), three parameters log-normal (LN3), Pearson type III (P3), generalized logistic (GLO), and generalized Pareto (GPA). The TLMs technique was additionally applied for parameter estimation to moderate the impact of exceptions on final gauges. LH-moments were utilized for assessing the upper piece of likelihood circulations and more significant occasions in the information tests. LH moments mitigate the undesirable effects because of short example esteems that might be clear when estimating events identified with more extensive return periods. Utilizing various GOF tests, we saw that the LMs technique was best for eight stations, TLMs with cutting (1, 0), and LH-moments with level =2, 3, 4 were best for six and 14 stations, individually. A hypothetical connection between TLMs and LH-moments was additionally returned, which uncovered that LH-moments are exceptional instances of TLMs when we are propelled to make trimming just from the lower side. Awareness about precipitation displaying can be beneficial to take some careful steps to adapt to such issues as loss of harvests, human lives, and framework because of the substantial rainfall. For any practical application later on, for example, getting ready for water-related crises, maintainable water assets the executives, and development of various hydraulic structures, it is recommended that at any rate, these five distributions should be considered and looked at for final choice of best fit likelihood appropriation for the ADMR arrangement in Pakistan [5].

The atmosphere of any nation is assessed based on its climate conditions. Pakistan is situated from southwest to Northwest at 23-37degree north scope and 61-76-degree east scope. Pakistan has been confronting catastrophic events, such as dry seasons, floods, and storms. These monstrous circulations are caused commonly and can't be alleviated even though the misfortunes could be limited to a few or more degrees by legitimate arranging. The fundamental wellsprings of precipitation in Pakistan are the mid-year rainstorm, the western discouragement, and the tempests. From July to September, there is a lot of rainfall. This investigation introduces local quantiles of Annual Total Rainfall (ATR) for 30 meteorological stations utilizing index flood procedures dependent on LMs. Discordancy measure dependent on LMs was used to screen the ATR. As an issue of reality, exceptionally raised zones of Pakistan get more precipitation; the investigation zone was isolated into four unique areas.

Further to legitimize the homogeneity of these areas, LM based heterogeneity measure (H) was determined for every station utilizing four-parameter Kappa distribution. For every region, best appropriation was found among Pearson type III (PE3), Generalized Normal (GNO), Generalized Extreme Value (GEV), Generalized Logistic (GLO), and Generalized Pareto (GPA) distribution dispersion utilizing ZM and LM proportion outline. Provincial quantiles based on best-fit dispersion for every area were resolved, and further for vigor, the exactness measures for the assessed territorial quantiles were determined utilizing Monte Carlo simulations. It was discovered that PE3 was the most appropriate decision for an enormous return period for the initial three stations while for a short return period GNO and GEV. So also, for region IV, GEV was announced as the best fit for lower return periods as long as 20 years, while for a time of 50, 100, 500, and 1000 years GNO was the best one [13].

This examination introduces a systematic investigation for recognizing and attributing patterns in the yearly frequency of precipitation occasions over the conterminous US to environmental change and atmosphere fluctuation modes. A Bayesian multilevel model is created for 1244 precipitation stations all the while to test the invalid speculation of no pattern and check two interchange theories: pattern can be ascribed to changes in worldwide surface temperature peculiarities or to a mix of indeed understood repeating atmosphere modes with fluctuating periodicity and worldwide surface temperature inconsistencies. The Bayesian staggered model gives the chance to pool data over stations and diminish the parameter estimation vulnerability, henceforth distinguishing the patterns better. The decision of the best exchange speculation is made dependent on the Watanabe–Akaike data foundation, a Bayesian point-wise proactive precision measure. Measurably significant time patterns are seen in 742 of the 1244 stations. Patterns in 409 of these stations can be ascribed to changes in worldwide surface temperature oddities. These stations are dominatingly found in the U.S. Southeast and Upper east atmosphere areas. The patterns in 274 of these stations can be ascribed to El Niño–Southern Oscillation, the North Atlantic Oscillation, the Pacific decadal Oscillation, and the Atlantic multidecadal wavering alongside changes in worldwide surface temperature peculiarities. These stations are principally found in the U.S. Northwest, West, and Southwest atmosphere districts .

2.6 Climate Change Profile of Pakistan

Pakistan possibly faces a significant environmental change challenge. A deliberate exertion by the administration and everyday society at all levels is required to moderate these dangers. Over the most recent 50 years, the yearly mean temperature in Pakistan has expanded by generally 0.5°C. The quantity of warmth wave days out of each year has grown about fivefold over the past 30 years. Yearly precipitation has typically demonstrated high changeability yet has somewhat expanded over the most recent 50 years. Ocean level along the Karachi coast has risen around 10 centimeters in the century. Before this current century's over, the yearly mean temperature in Pakistan is required to ascend by 3°C to 5°C for a focal worldwide outflows situation. In comparison, higher worldwide discharges may yield an ascent of 4°C to 6°C. Average yearly precipitations are not expected to have a significant long-term trend; however, they are relied upon to display sizeable inter-annual variability. The ocean level is relied upon to ascend by a further 60 centimeters before the century is over and will influence the low-lying beachfront regions south of Karachi toward Keti Bander and the Indus Stream delta. Under future environmental change situations, Pakistan is relied upon to encounter expanded changeability of waterway streams because of the expanded fluctuation of precipitation and the liquefying of ice sheets. Interest in water system water may increment because of higher dissipation rates[20]. Yields of wheat and basmati rice are relied upon to decay and may drive generation northward, subject to water accessibility. Water accessibility for the hydropower age may decay. More sweltering temperatures are probably going to expand vitality requests because of expanded cooling necessities. Hotter air and water temperatures may diminish the proficiency of atomic and warm power plant age. Mortality because of outrageous warmth waves may increment. Urban waste frameworks might be additionally worried by high precipitation and glimmer floods. Ocean level ascent and tempest floods may antagonistically influence beach front foundation and vocations.

Adjusting to these effects may include advancement or utilization of harvest assortments with more extraordinary warmth and dry season resilience, modernizing water system foundation and utilizing water-sparing advances, coordinated watershed the executives, reforestation of catchment regions, and development of extra water stockpiling expansion of vitality blend remembering venture for inexhaustible and little hydropower ventures, improved climate anticipating and cautioning frameworks, retrofitting of basic vitality foundation, and development of dams or ocean dividers.

The National Environmental Change Approach of 2012 is Pakistan's managing record on environmental change, defining the objective of accomplishing atmosphere flexible advancement for the nation through mainstreaming ecological change in the country's monetarily and socially vulnerable parts. Roughly 6% of Pakistan's government spending plan during 2010–2014 included environmental change-related consumption, prevalently in vitality and transport. As portrayed in its Broadly Decided Commitment to the Paris Understanding under the Assembled Countries System Show on Environmental Change, Pakistan expects to decrease up to 20% of its 2030 anticipated ozone-depleting substance discharges, subject to accessibility of global awards to meet the total reduction costs adding up to around $40 billion. The nation's adjustment needs have been recognized to go between $7 billion to $14 billion every year.[21]

Pakistan being a developing nation, 70% of its economy is subject to agribusiness. Water is an incredible requirement for the rural division, just like human life and other living life forms. Pakistan gets precipitation in summer as well as in winter. In summer, precipitation generally happens during the storm season (early July to September). July and August are the pinnacles for a very long time for rainstorm precipitation in Pakistan. The most significant wellspring of water for farming overall is precipitation. Practically 60% of the all-out yearly precipitation in Pakistan is because of precipitation in summer. All Kharif crops are, for the most part, reliant on the example of storm precipitation. Winter precipitation is significant for Rabi crops in the nation. In Pakistan, water is generally utilized underway of agribusiness and hydroelectricity. Farming is the foundation of Pakistan's economy, yet tragically extraordinary rainfalls in the national cause loss of yields, lives, and framework. Extraordinary natural occasions can leave overall effects on society and the economy. In this investigation, at-site frequency analysis (AFA) of an annual daily maximum rainfall (ADMR) arrangement was done utilizing the strategy for linear moments (LMs) and their variations, for example, trimmed linear moments (TLMs) and higher-order linear moments (LH-minutes). The ADMR arrangement we examined was seen at 28 meteorological observatories crosswise over Pakistan as recovered from the Pakistan Meteorological Office (PMD). The essential point of the examination was to find best-fit (i.e., the most appropriate) likelihood circulation among the class of different likelihood dispersions. At first extraordinary goodness-of fit (GOF) measures, for example, the Kolmogorov-Smirnov test (KST), Anderson-Darling test (ADT), root mean square error (RMSE), and LMs ratio diagram (LRD) were applied to decide the best-fit conveyances as well as the best linear estimation technique for AFA. We saw that no single likelihood circulation could be pronounced as the best-fit appropriation for every station. Five dispersions were seen as the truest: generalized extreme value (GEV), three parameters log-normal (LN3), Pearson type III (P3), generalized logistic (GLO), and generalized Pareto (GPA). The TLMs strategy was likewise applied for parameter estimation to relieve the impact of anomalies on final gauges. LH-minutes were utilized for assessing the upper piece of likelihood appropriations and bigger occasions in the information tests. LH moments lighten the undesirable effects because of little example esteems that might be clear during estimation of occasions identified with bigger return periods. Utilizing specific GOF tests, we saw that the LMs strategy was best for eight stations, TLMs with cutting (1, 0), and LH-minutes with level =2, 3, 4 were best for six and 14 stations, individually. A hypothetical connection between TLMs and LH-minutes was additionally returned, which uncovered that LH-minutes are unique instances of TLMs when we are persuaded to make cutting just from the lower side. In this examination, AFA was done for ADMR arrangement in Pakistan utilizing information from 28 meteorological stations. The arrangement of most reasonable appropriations is chosen based on various GOF estimations, for example, KST, ADT, RMSE, and LRD. These GOF estimations distinguish the best likelihood conveyances and the superior estimation technique among LMs, TLMs, and LH-Moments. Practically speaking, the reason for dissecting ADMR arrangement is to anticipate outrageous precipitation for longer return periods and to dodge the disturbance impact of littler example esteems in evaluating upper quantiles. TLMs and LH-moments ought to likewise be viewed as other than straightforward LMs. Just five circulations (for example, GEV, GLO, GPA, LN3, and P3) are announced the most appropriate for various stations. LMs, TLMs with cutting (1, 0), LH-minutes with level (=2), (=3), (= 4) are reasonable for 8, 6, and 14 stations, separately. A hypothetical connection between TLMs and LH moments is additionally returned to, which convinces us that LH moments are extraordinary instances of TLMs when cutting is performed distinctly from the lower side. Mindfulness about precipitation demonstrating can be beneficial to take some careful steps to adapt to such issues as loss of harvests, human lives, and foundation because of the overwhelming precipitation [5].

2.7 Disaster Risk Reduction

Climate-related disasters in the South Asian station are getting progressively dangerous and costlier regarding economic and social effects. From 1900 to 2015, there has been a nonstop increment in the quantity of atmosphere-related disasters, cresting from 2000 to 2015. Strikingly, the pace of mortality related to such disasters has been impressively diminished in the present period, even though the number of individuals allegedly influenced has expanded tremendously. While South Asia is as often as possible presented to severe atmosphere-related disasters, a circumstance has been additionally muddled by exceptional populace development, arriving at a gigantic 1.8 billion in 2018 and anticipated to increase by another 800 million individuals by 2050. The whole populace is powerless against change in atmosphere, and outrageous occasions gave a low degree of institutional capacities, financial defenselessness, and significant reliance on climate-sensitive assets[22]. Atmosphere-related vulnerabilities are exacerbated by low populace strength; many people survive on less than USD 1 per day. The degree of this pressure is related to the high variability of climate-related disasters. South Asia is vulnerable to an assortment of hydro-meteorological risks, which are regularly cross-boundary in nature. Environmental change is relied upon to influence a significant number of these risks. In this way, atmosphere-related dangers over South Asia make disaster risk reduction (DRR) and environmental change adjustment (CCA) critical strategic objectives. As of late, there is an expanding accord that DRR including CCA ought to be inserted being developed arranging. Catastrophe hazard decrease, including CCA, has dynamically picked up significance in worldwide administration. Crosswise over South Asia, however, such a combination is just in a starter organize. This survey evaluated the current status and extent of DRR incorporating CCA being developed tasks crosswise over South Asia, with the goal that a successful and attainable consultation might be made to local policymakers. A sum of 371 ventures pertinent to CCA and DRR were looked into. The venture stock was various concerning area, scale, sectoral center, and vital significance. Bangladesh, India, and Bhutan were seen to be proactive in actualizing DRR-and CCA-related activities. Meta-investigation of the venture stock proposes a critical requirement for an individual and communitarian assembly of procedures for DRR and CCA through approaches, plans, systems, and projects. The main conclusion is to address the DRR, including CCA, over South Asia. The top needs should be to set up institutional linkages at both the national and local levels by expelling fundamental hindrances. Emphasis should build up a mix among different foundations both on a level plane and vertically. A system for the union of policies, planning, and programs ought to likewise be organized. The SAARC part states should set up a structure for guaranteeing total coordination of strategies, plans, and projects embraced for DRR, including CCA inside individual nations and the whole district. There is additionally a need to set up a stage for information sharing, similar to an environmental change information the executives focus parallel to the SAARC Calamity. The executives Center (SDMC) or a division inside the SDMC to encourage sharing information and data, directing provincial research, and reproducing excellent practices among SAARC part states [23]

Hydrologists are in every case are short of data for making decisions about water assets structures like spillways, dikes, storm flood hindrances, and dams, and so forth. The physical laws are lacking to deal with the wrong short information and critical changes in arbitrary procedures. The hydro metrological factors like extraordinary precipitation are hard to depict due to irregular changes in climate and examining blunder created by constrained information, just like a bit of example of a boundless populace. In this application, the recorded data at various destinations of a broad characterized homogeneous area is utilized to evaluate the outrageous occasions expected to happen in 100 or 1000 years to decrease the vulnerabilities of uncommon events.

Provincial precipitation recurrence investigation is of paramount significance in standard structure plans, just as it assumes a significant job in a differing scope of nonstructural issues, including characteristic dangers related to extraordinary precipitation occasions, because the examinations give the data about the event precipitation sums inside a specified repeat interim. Estimation of the extent and recurrence of extreme precipitation has vast significance in settling on water-powered structures like spillways, dikes and dams, and so forth. This examination includes estimating provincial precipitation quantiles of 23 destinations utilizing LM based list flood territorial recurrence investigation. At first, various tests are applied to check the presumptions of autonomy, stationarity, and indistinguishable dispersion. An LM-based discordancy measure is utilized to distinguish conflicting stations. Since in Pakistan, the profoundly raised region gets more precipitation. Based on this trademark, the investigation area is separated into three stations that fulfill the LM-based heterogeneity measurements utilizing Monte Carlo simulations from Kappa dissemination. The territorial quantile gauges are acquired from GEV, GNO, and GLO conveyances seen as best decisions for each of the three districts dependent on LM proportion chart, Z-Statistics, and normal weighted contrast esteems. For robust territorial gauges, some precision measures are determined utilizing a reenactment investigation of local LM calculation. Based on relative predisposition, relative total inclination, and relative RMSE, GNO is found to be best hearty for local quantile estimation at ale return times of 50, 100, 500, and 1000 and GEV at return times of 1, 2, 5, 10 and 20 for every one of the three districts. There is an extraordinary variety in the atmosphere of Pakistan, dependent on geography. A significant piece of Pakistan encounters a dry atmosphere. Humid conditions beat a short region in the north. In Pakistan, precipitation is brought about by Storm and Western Sadness. Storm happens from July to September and Western Despondency from December to spring. The average rainfall of the nation is around 300mm, of which 140mm precipitation occurs during the storm. A provincial recurrence examination was performed on a district of

Pakistan having 23 stations which are pretty much influenced by a rainstorm.

At first, the suppositions of local recurrence examination were tried by time arrangement plots, Mann-Whitney test, Kendall's tautest, and L Jung-Box-Q-Measurements. These tests fulfilled all destinations. Site 17, Multan indicated autocorrelations on account of some dry season years. Non-uniformity of climatic conditions didn't permit expelling this site from the examination.

2.8 Forecasting of Rainfall in Pakistan

Site 20, Khanpur, has had all the earmarks of being harsh. In light of having no gross mistakes for this site, it was considered in testing the heterogeneity of the area. An area of 23 stations didn't fulfill Hosking and Wallis' heterogeneity insights. Since in Pakistan, destinations having high height get more precipitation while the stations on plain regions get less rainfall. The three areas were shaped based on mean yearly precipitation and size, which are acceptably homogeneous. Station 1 were having 11 destinations, districts 2 and 3 each having five stations. Site 2 and 20 were disposed of as being creating the high estimations of heterogeneity measurements. Further investigation was performed on three homogeneous districts. [14]

Precipitation is significant for the production plan, water assets, the executives, and all movement designs. The delayed dry period or substantial downpours at the primary phases of the harvest development and improvement may prompt a critical decrease in crop yield. Pakistan is an agricultural nation, and its economy is to a great extent dependent on crop profitability. Therefore, precipitation forecast turns into a massive factor in farming countries like Pakistan. Precipitation prediction has been one of the most challenging issues in the world. Precipitation prediction includes a blend of probabilistic models, perception, and information on patterns and examples. Utilizing these strategies, sensibly precise figures can be made up. A few late research ponders created precipitation expectations using different climate and atmosphere estimating techniques. Various examinations have been directed about precipitation anticipating in Pakistan and other nations. The industriousness is commonly perceived in hydrological and climatological time arrangements over a different scope of time scales. Among them, a significant climatological variable is the measure of precipitation. Unmistakably, the precipitation examination has a substantial job in the fruitful arranging, advancement, and usage of water assets of the executives to assess building ventures and ecological issues. They incorporate hydropower age, repository activity, flood control, and control of water quality. Subsequently, a proficient investigation of worldly precipitation conduct is viewed as basically significant in hydrology. In this paper, a study has led the nation to display the precipitation pattern in Pakistan in recent decades.

For this reason, an optional dataset of average precipitation containing 65 years from 1951 to 2015 is obtained from the World Bank site. In Pakistan, unfriendly results of rain have just been watched. They are dry spells and super floods, which have seriously influenced human settlements, water the executives, and horticulture. The information is breaking down through a cut practical time arrangement model in this examination, a generally new technique for determining. The outcomes show a diminishing pattern in average precipitation over the nation. The month-to-month figures for the following ten years (2016-2025) are gotten alongside 80% forecast interims. These gauges are additionally contrasted, and the statistics got from ARIMA and exponential smoothing state space (ETS) models. This paper broke down month-to-month precipitation information through the sliced functional time series (SFTS) model. A moderately new technique for gauging was presented, and the month-to-month conjectures for the following ten years (2016-2025) were acquired alongside an 80% expectation interims. These conjectures were likewise contrasted, and the gauges received from Autoregressive Integrated Moving Average (ARIMA) and exponential smoothing state space (ETS) models. It was discovered that the SFTS model performed superior to standard ARIMA and ETS ones, and the figures acquired from SFTS models are progressively precise and solid. Yet, they additionally give restricted forecast interims when contrasted with different models [24]

2.9 Research Gaps

Various published studies have addressed the issue of climate change or frequency analysis of rainfall separately. None of the published studies so far (to the best of our knowledge) has investigated the effect of climate change on the extremes of rainfall in various regions/zones/provinces of Pakistan

So there is need to understand trend and tendencies of magnitude and frequency of natural disaster due to floods in Pakistan. as floods are majorly correlated to extremes of rainfall.further there is need for in-depth analysis and prediction of extreme of rainfall in Pakistan.

Chapter 3

Methodology

In this section, we are going to discuss the methodology, we followed in this research. This section contains the main procedures followed for data collection. This chapter also covers the comprehensive review of statistical data analysis.

3.1 Descriptive Statistics

Descriptive statistics include mean, standard deviation, minimum and maximum, skewness, kurtosis of each station.

3.1.1 Mean

Mean is the average value of the station. To calculate the mean, add all values and divide them by a total number of values.

3.1.2 Standard deviation

Standard deviation is the measure of the amount of variation from the mean. The low value of standard deviation indicates the values are close to the mean.

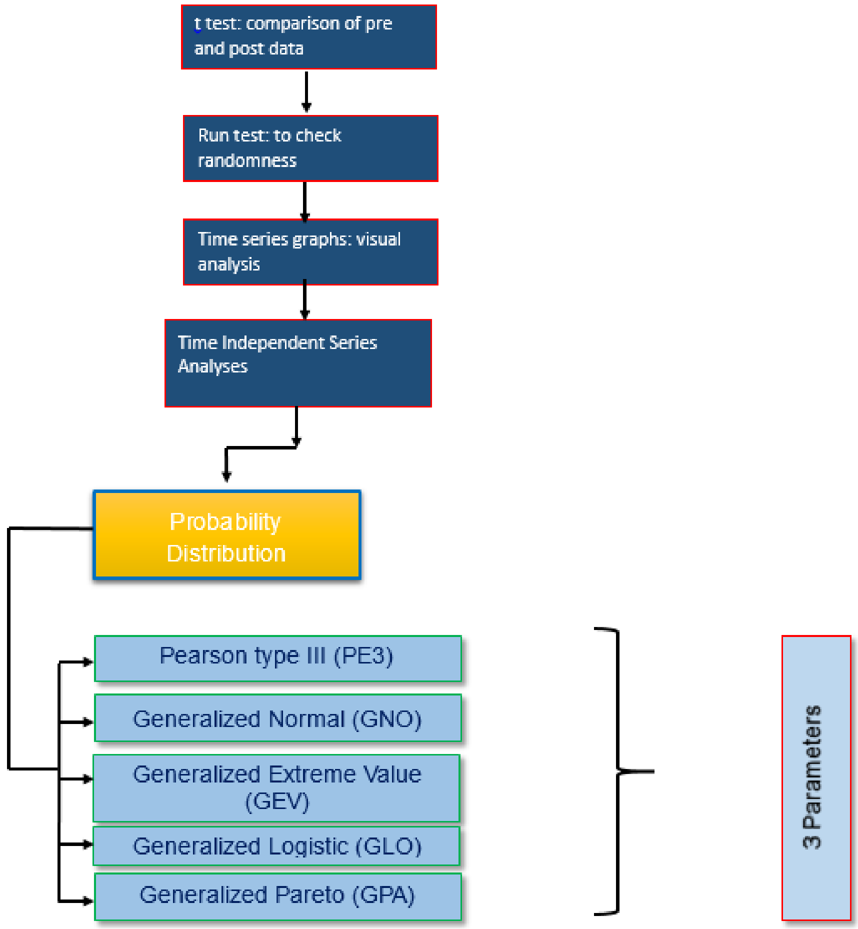


Figure 3.1: Flow chart of methodology

3.1.3 Minimum and Maximum

Minimum and maximum values are the minimum value of rainfall and the maximum value of precipitation for the corresponding station.

3.1.4 Skewness

Skewness is the measure of symmetry or lack of balance. There is an average skewness, positive skewness, and negative skewness. Normal means the data is symmetric. At the same time, it is positive or negative when information is non-symmetric.

3.2 Kurtosis

Kurtosis refers to data concerning normal distribution as data is heavy-tailed or light-tailed compared to normal distribution. The value for kurtosis of the normal distribution is 3. There are other types of kurtosis, Mesokurtic is a normal distribution of data, and Leptokurtic shows positive kurtosis and has heavy tails. In contrast, Platykurtic shows negative kurtosis and has flat tails.

3.2.1 Time-series graphs:

A time-series graph shows different values against different times. These are similar to x-y plots. Time is plotted on the x-axis, and other values of rainfall are plotted against every year. The primary purpose of a time-series graph is to see any trend in data and using that trend for prediction. Also, the mean line is drawn in the time-series graph to see the pattern of the chart. Time-series graphs make trends visible.

To construct a time series graph starts with a standard x-y plane. The horizontal x-axis is used to plot time given in years from 1980-2015 total of 35 years. The vertical y-axis is used to plot different values of extreme rainfall. So each point on the graph corresponds to a year and a value of excessive rain. The mean value line is drawn to see trends fluctuating across the mean.

Time-series graphs of fourteen stations are drawn to see the trends and variations. Also, peak values are shown for every station in the time-series graph. Most of the stations have the same trend in their peak value, which lies between the years 1990 to the year 2000.

3.3 Zone Wise Test Results

T-Tests and F Tests are performed zone-wise to test the validity of the null hypothesis.

3.3.1 t-test:

A T-test is a statistical method that tests the validity of the null hypothesis, An independent-group t test can be carried out for a comparison of means between two independent groups[25],which states a commonly accepted claim about a population.

The smaller the p-value, the more substantial the evidence that the null hypothesis should be rejected and that the alternate hypothesis might be more credible. The P-test statistic typically follows a standard normal distribution when large sample sizes are used. The t-test is used to compare the mean of pre and past benchmark years as the F test is a prerequisite to the t-test used to compare the variance of pre and post benchmark years. We assumed the null hypothesis that the variance of pre and post-bench mark years are equal. So alternative hypothesis states that variances of pre and post-bench mark years are not equal. The significance of the null hypothesis is 0.05 or 5 percent. If the p-value is less than 0.05, then the null hypothesis is rejected, and if the p-value is more significant than 0.05, then the null hypothesis is accepted

3.3.2 F-test:

An F-test is any statistical test in which the test statistic has an F-distribution under the null hypothesis. It is most often used when comparing statistical models that have been fitted to a data set to identify the model that best fits the population from which the data were sampled.[26]

3.3.3 Run test and Lag 1 auto co-relation

The run test is implemented to check whether the data is random or not. It mostly depends on no of runs present in the data[27]. A p-value of the run test determines while the information is random or not. We take the null hypothesis that data is random. The significance of this test is at 5 percent. If the P values are less than 0.05, then the null hypothesis is rejected, and if the p-value is more significant than 0.05, then the null hypothesis is accepted.

Lag 1 auto co-relation is also another test used to check whether the data is random[28].

3.3.4 Modeling of extreme events

There are two ways to model extreme events. One is time-independent series analysis, and the other one is time series analysis. A time-independent series analysis is implemented in this research. Further, there are three models in time-independent series analysis. Probability distribution has been selected for modeling extreme events.

3.4 Probability distributions

In this section, we will discuss the inferential analysis of all the stations. In this regard, the following distributions are included:

3.4.1 Generalized Extreme Value Distribution

The generalized extreme-value (GEV) distribution is widely used for modeling and characterizing extremes[29].In probability theory and statistics, the generalized extreme value (GEV) distribution is a family of continuous probability distributions developed within extreme value theory to combine the Gumbel, Fréchet, and Weibull families, also known as type I, II, and III extreme value distributions. The GEV distribution is the only possible limit distribution of properly normalized maxima of a sequence of independent and identically distributed random variables by the extreme value theorem. Note that a limit distribution needs to exist, which requires regularity conditions on the tail of the distribution. Despite this, the GEV distribution is often used as an approximation to model the maxima of long (finite) sequences of random variables. It has three parameters, i.e., location, scale, and shape.

$$F(x) = exp-[1 + \xi(x - \mu\sigma)] - 1/\xi \tag{3.4.1}$$

Define, for

$$: 1 + \xi(x - \mu)/\sigma > 0, -\infty < \mu < \infty, \sigma > 0 and - \infty < \xi < \infty \tag{3.4.2}$$

$$\mu : LocationPerimeter \xi : ShapePerimeter \sigma : ScalePerimeter \tag{3.4.3}$$

3.4.2 Generalized Pareto

In statistics, the generalized Pareto distribution (GPD) is a family of continuous probability distributions. It is often used to model the tails of another distribution.It has since been used by many authors to model data in several fields. [30]

$$F(x) = 1/\sigma(1 + \xi * z)power(-1/\xi + 1)wherez = (x - \mu)/\sigma \tag{3.4.4}$$

$$\mu : LocationPerimeter \xi : ShapePerimeter \sigma : ScalePerimeter \tag{3.4.5}$$

3.4.3 Inverse Gaussian

In probability theory, the inverse Gaussian distribution (also known as the Wald distribution) is a two-parameter family of continuous probability distributions. The inverse Gaussian distribution is related closely to the Gaussian distribution, as is suggested by its name. It is used in mathematical statistics and in various fields such as engineering to describe multiple phenomena and make quantitative analyses.

F(x) = $\sqrt{\lambda/2\pi x^3} exp(-(\lambda(x-\mu)^2)/2\mu^2 x))$
$for x > 0, \mu : Mean\lambda : ShapeParameter.$(3.4.6)

3.4.4 Log Pearson III

The log-Pearson type 3 (LP3) distribution has been one of the most frequently used distributions for hydrologic frequency analyses. It is also recommended that this distribution be fitted to sample data using mean, standard deviation, and coefficient of skewness of the logarithms of flow data [i.e., the method of moments (MOM)]. A large volume of literature on the LP3 distribution has since been published about its accuracy and fitting or parameter estimation methods.

$$F(x) = 1/a\gamma b \int ((x-c)/a)^{(} b-1) exp(-(x-c)/a)) \qquad (3.4.7)$$

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

3.5 Best Fit Distribution

To check which distribution is the best fit. Three tests are performed. Based on three tests, the best-fit probability distribution is fitted.

3.5.1 Chi-square

These statistics tell about how much difference exists between observed count and expected count.

3.5.2 Kolmogorov Smirnov

A non-parametric test is used to tell the difference between the two cumulative distributions and calculates a p-value.

3.5.3 Anderson darling

This test measures how to fit a distribution is for any data set. The smallest the value of the test, the better the distribution fits the data.

Chapter 4

Result And Discussion

In this section, we will discuss the statistical results of this research. The results of 14 stations are represented below. The result section contains two main portions. Firstly, we will discuss the graphs of data collection from stations. Secondly, we will mean p-test and t-tests on the respective stations.

4.1 Station Bibliography

In this section, a detailed review of data collection stations is presented in table 4.1.

4.2 Description of Stations.

4.2.1 Zone A

Zone A comprises those stations having a cold climate and high mountains, situated in northern Pakistan; Chitral, Gilgit, Muzaffarabad, Said-u- Sharif, Skardu, Astor, Dir, Chilas Parachinar, and Kakul4.1. These are mostly hill stations located between 34 N to 38 N in the Himalaya, Hindukush, and Koh-e- Sufaid mountain ranges.

4.2.2 Zone B

This zone has mild cold climate and Sub Mountains, located between 31N to 34 N. The stations are Sialkot, D.I.Khan, Islamabad, Peshawar, Cherat, and Lahore.

Table 4.1: Geographic details of stations

Station	Latitude	Longitude	Elevation
Bahawalpur	29.20	71.47	110.00 meter
Chhor	29.53	69.43	05 meter
Chitral	35.51	71.50	1497.8 meter
Gupis	36.10	73.24	2156.0 meter
Islamabad	33.44	73.5	540 meter
Jivani	25.04	61.48	56 meter
Karachi	24.54	66.56	22 meter
Multan	30.12	71.26	121.95 meter
Peshawar	34.02	71.56	327 meter
] Quetta	30.05	66.58	1719m
Badin	24.38	68.54	09 meter
Jacobabad	28.18	68.28	55 meter
Sibbi	29.33	67.59	133m
Gilgit	35.55	74.20	14600 meter

4.2.3 Zone C

Climate is cold in winters and hot in summers. Most of them are mountainous stations with high elevations from mean sea level and cover an area between 27 N to 32N and 64 E to 70 E. Stations included in this zone are Quetta, Zhob, Kalat, and Khuzdar.

4.2.4 Zone D

This is the hottest and dry zone of the country, where the highest maximum temperatures are recorded in Sibbi and Jacobabad. The area is almost plain, with some sites included in the Thar Desert. Stations had are Sibbi, Jacobabad, Bahawalpur, Khanpur, Multan, and Rohri.

4.2.5 Zone E

Zone E is a big zone having many stations and coastal cities near the Arabian Sea. The coastal part comprises only a tiny amount of this region, and the climate above coastal elements in Balochistan and Sindh province is mainly arid to hyper-arid. The selected stations from this zone are Hyderabad, Karachi, Nawabshah, and Jewani.

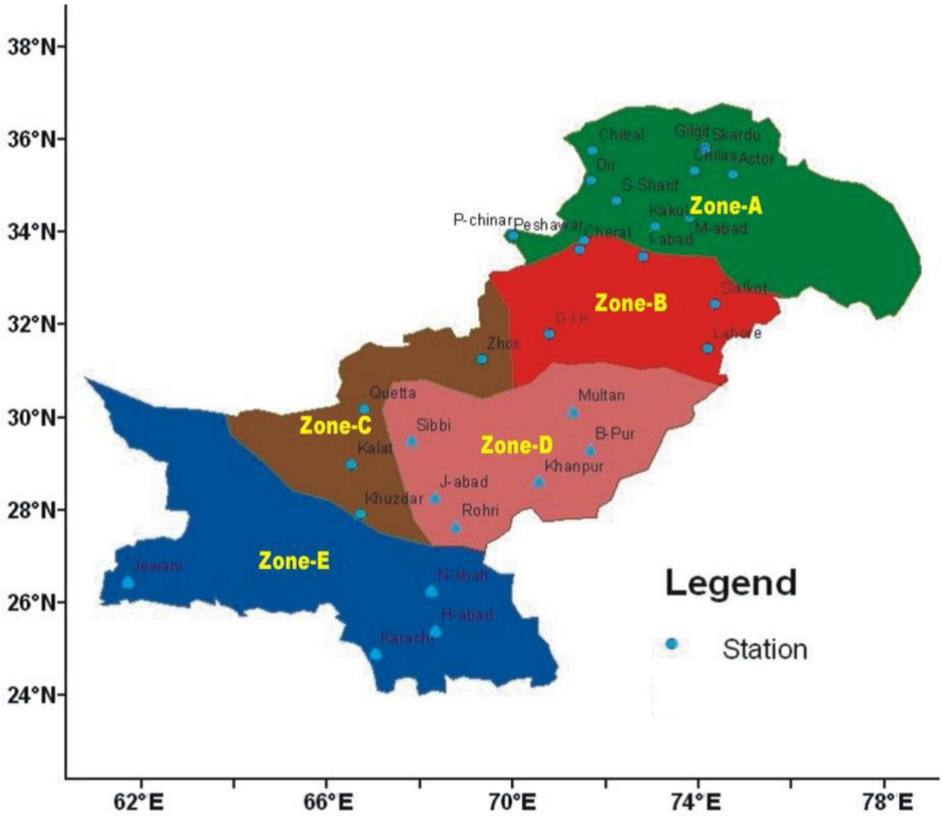


Figure 4.1: Data stations on the map of Pakistan [2]

4.3 Descriptive statistics of stations

This subsection contains the descriptive statistics of 14 stations. Following three tables table 4.2 table4.3 table4.4 shows the different essential variables of descriptive statistics of 14 stations.

4.4 Comparison of Pre and Post data

The following table shows the comparison of pre and post-data. Data is divided into two parts. The year 2000 is taken as the center point, and information is separated into two parts, i.e., pre-2000 and post 2000(1985-2000, 2000-2015). The mean of pre-2000 and

Table 4.2: Averaged descriptive statistics of 14 stations (1980-2015)

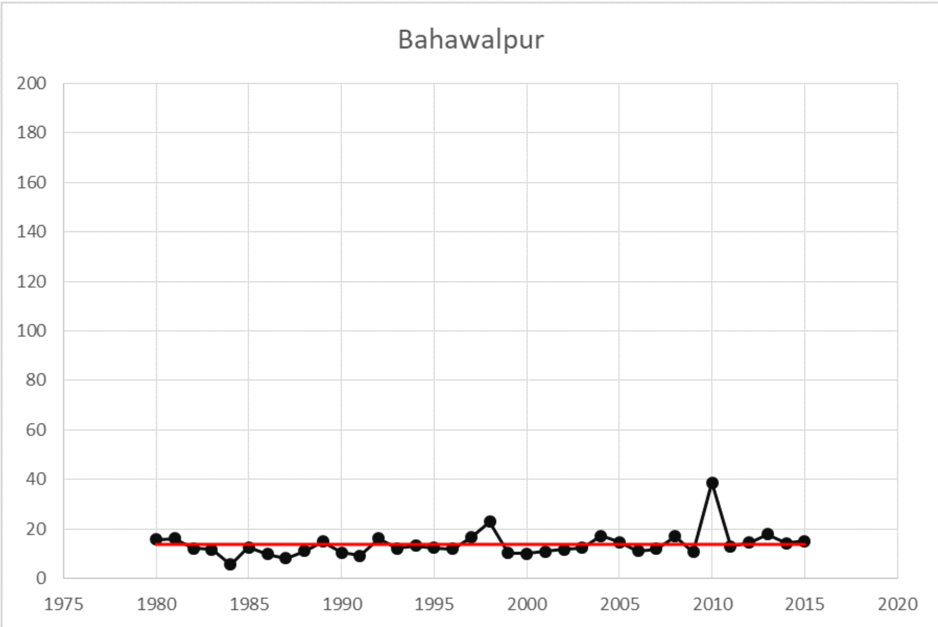
Sr no	City	Mean	Stand Deviation	Minimum	Maximum	Skewness	Kurtosis
1.	Badin	13.92	11.22	0.64	67.25	3.28	14.01
2.	Bahawalpur	13.78	5.28	5.70	38.79	3.05	13.43
3.	Chhor	12.16	5.58	1.19	27.71	0.57	1.02
4.	Chitral	29.67	7.49	18.60	51.67	0.82	0.52
5.	Gilgit	21.37	5.11	13.08	37.15	0.85	0.94
6.	Gupis	27.65	6.23	16.55	44.84	0.36	0.18
7.	Islamabad	60.18	35.02	22.16	157.28	1.42	1.57
8.	Jacobabad	10.03	6.67	2.09	33.60	1.71	3.32
9.	Jivani	14.69	11.79	1.57	45.78	1.43	1.076622
10.	Karachi	12.24	12.96	0.85	75.95	3.92	19.35
11.	Multan	12.60	3.79	4.02	23.40	0.55	1.37
12.	Peshawar	70.64	25.68	17.21	123.42	0.19	-0.31
13.	Quetta	15.07	6.06	4.74	28.51	0.80	-0.09
14.	Sibbi	13.59	9.46	1.92	56.70	2.86	11.22

post-2000 years is calculated and stated in the table. Likewise, minimum, maximum, standard deviation, skewness, and kurtosis are also calculated for pre and post-2000 year. as shown in table below table 4.3 and table 4.4. That makes a brief comparison of data for 14 stations.

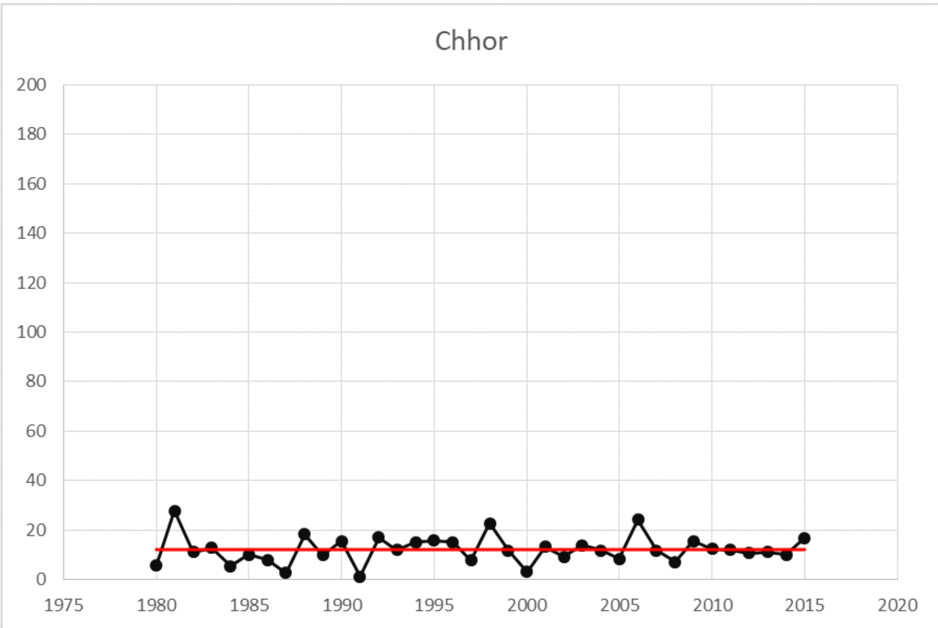
4.5 Graphs

Time-series graphs are drawn for descriptive analysis. There are 14 stations, and a graph of every station is drawn to show trends. The following are the graphs of 14 stations from graph 4.2a to graph 4.8b. Y-axis represents years, and the x-axis represents rain in millimeters. The Red line shows the mean or average value. The span of data is 36 years (1980-2015)

In the Bahawalpur graph, the peak value is at 2010, and the peak value is 40 millimeters. The average line is at 13.78. Karachi shows a significant peak value at 1989 year. Sargodha shows a significant peak at 1994 and 2013. Badin graph has a significant peak

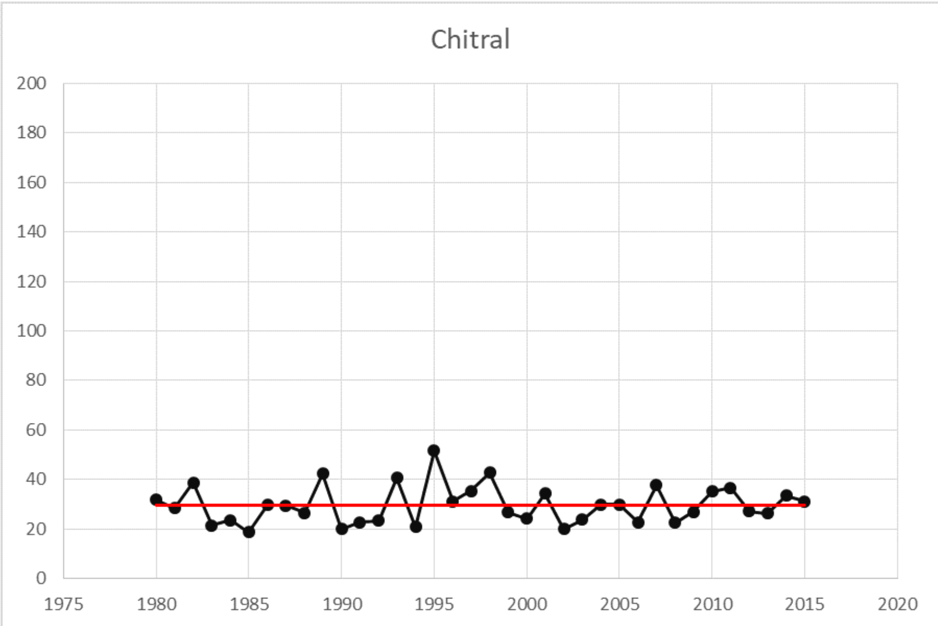


(a) Time series graph for descriptive analysis of Bahawalpur

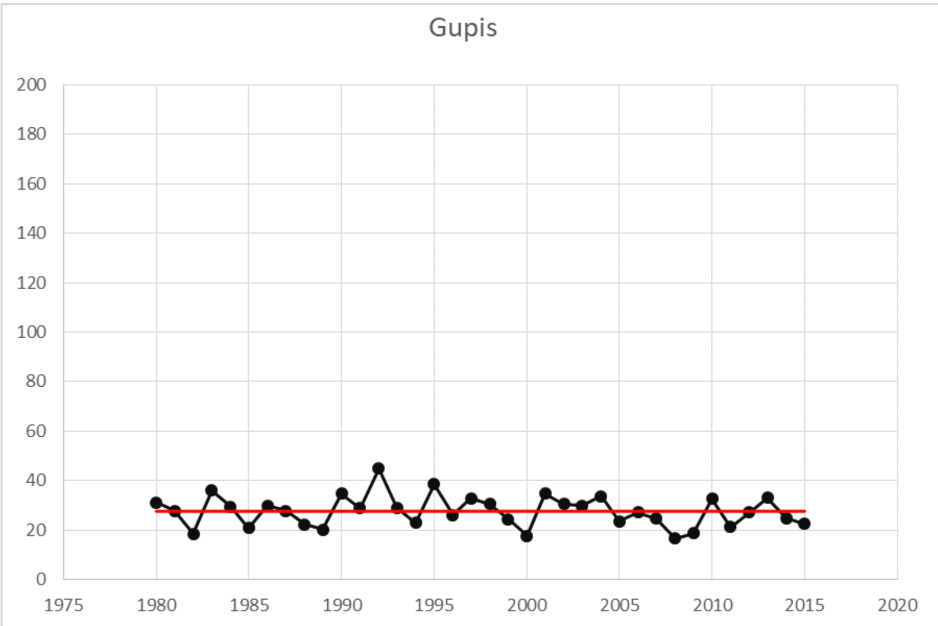


(b) Time series graph for descriptive analysis of Chhor

Figure 4.2: Time Series Graphs of Bahawlapur and Chhor

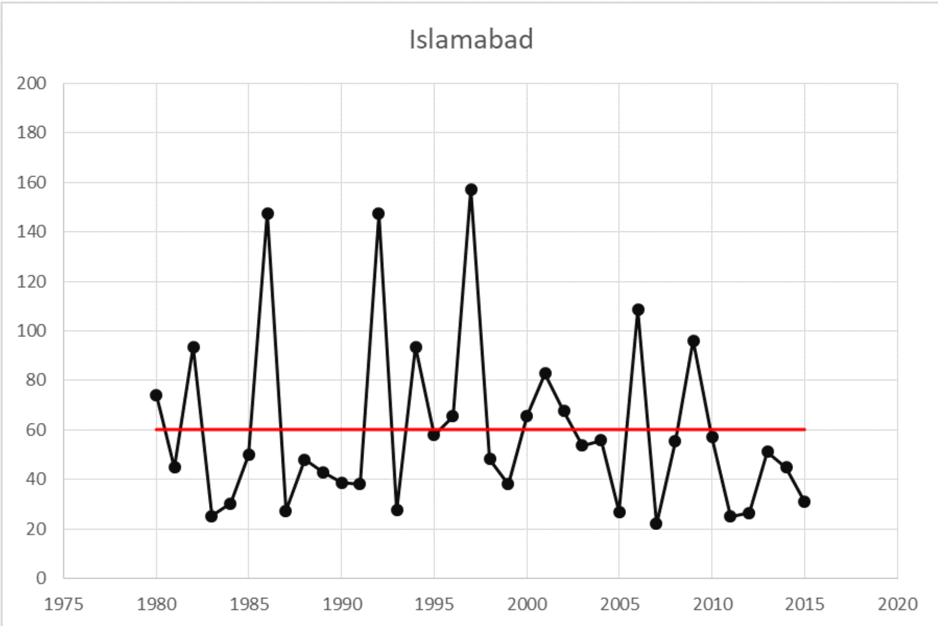


(a) Time series graph for descriptive analysis of Chitral

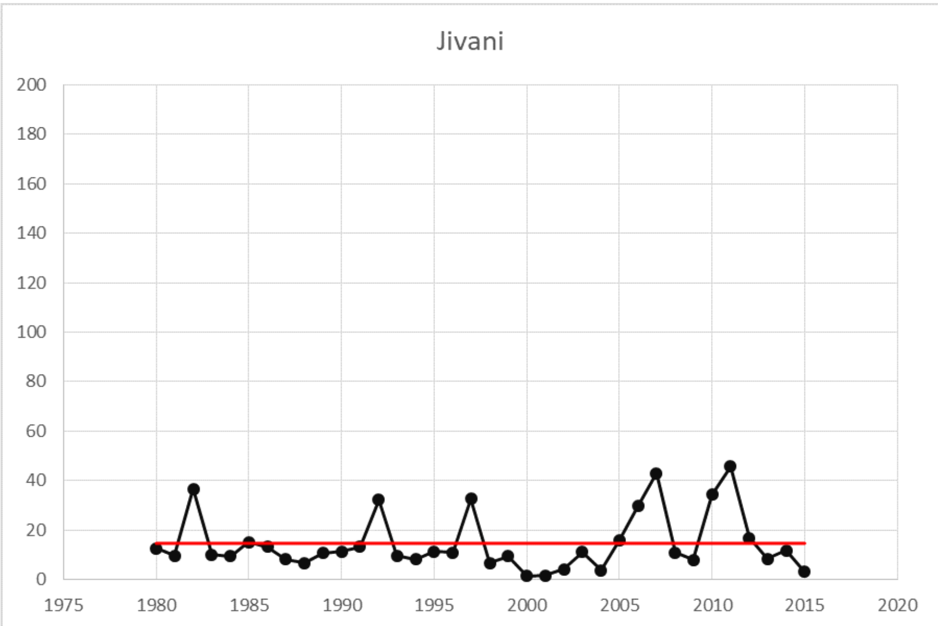


(b) Time series graph for descriptive analysis of Gupis

Figure 4.3: Time Series Graphs of Chitral and Gupis

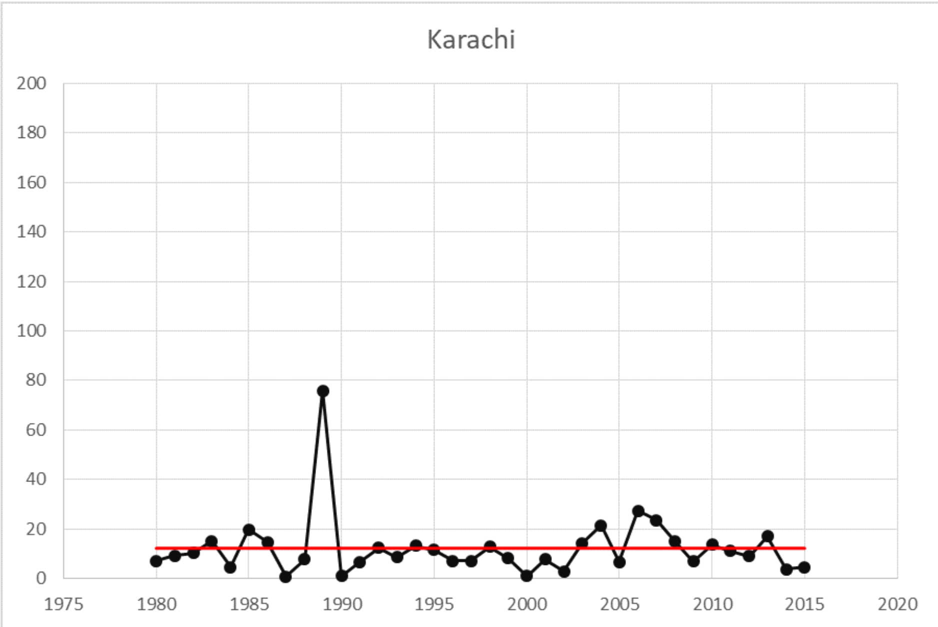


(a) Time series graph for descriptive analysis of Islamabad

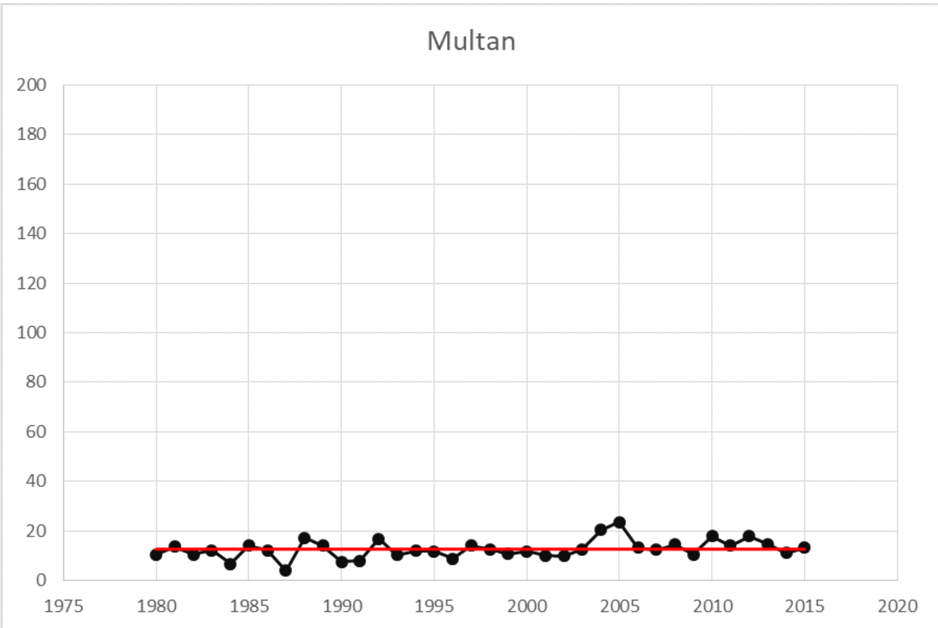


(b) Time series graph for descriptive analysis of Jivani

Figure 4.4: Time Series Graphs of Islamabad and Jiwani

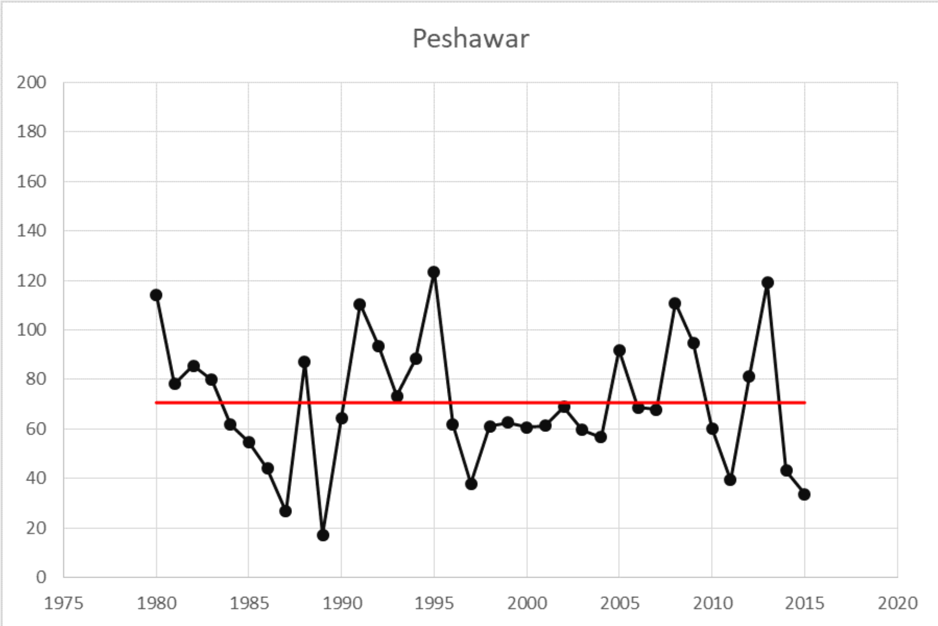


(a) Time series graph for descriptive analysis of Karachi

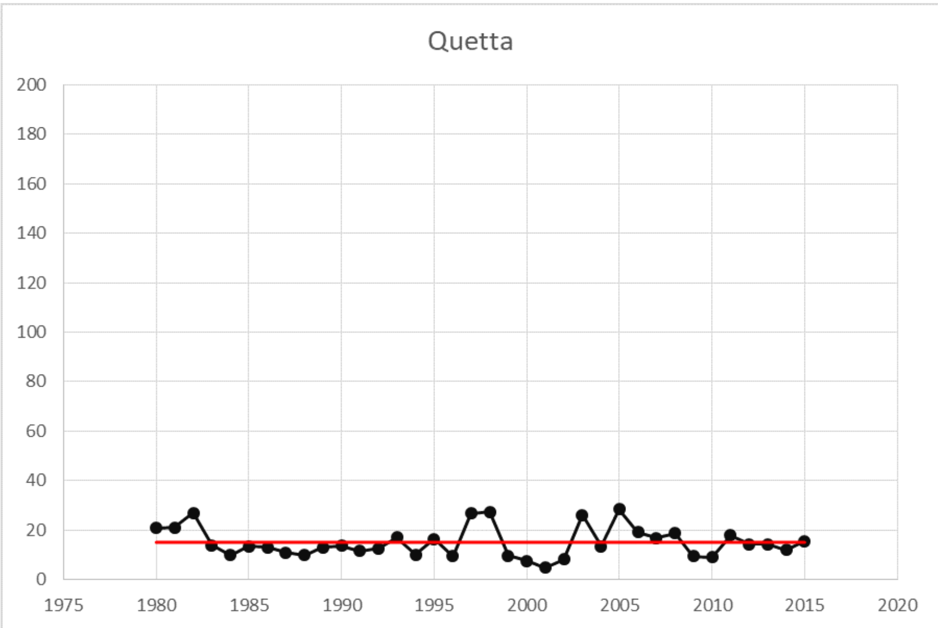


(b) Time series graph for descriptive analysis of Multan

Figure 4.5: Time Series Graphs of Multan and Karachi

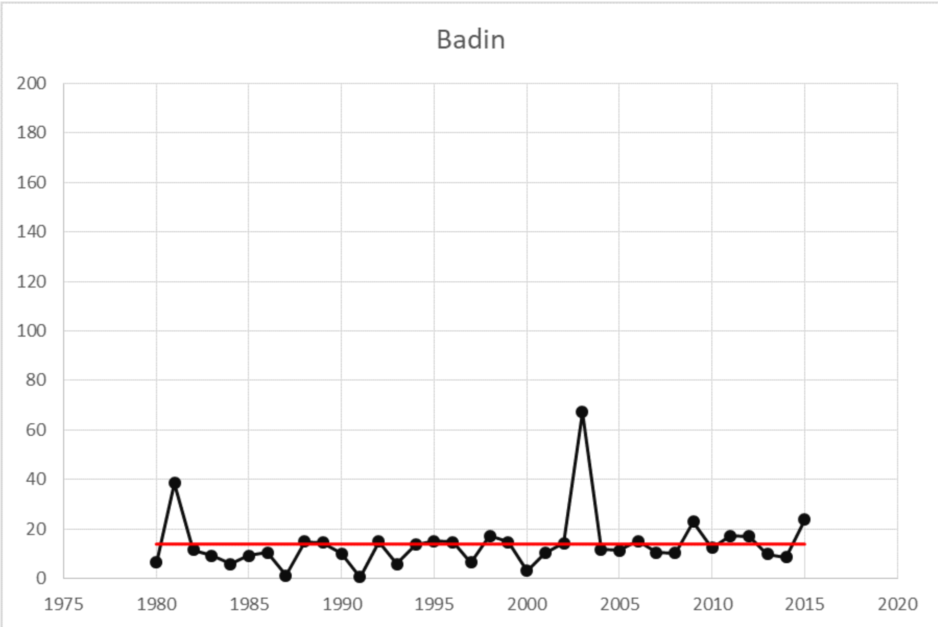


(a) Time series graph for descriptive analysis of Peshawar

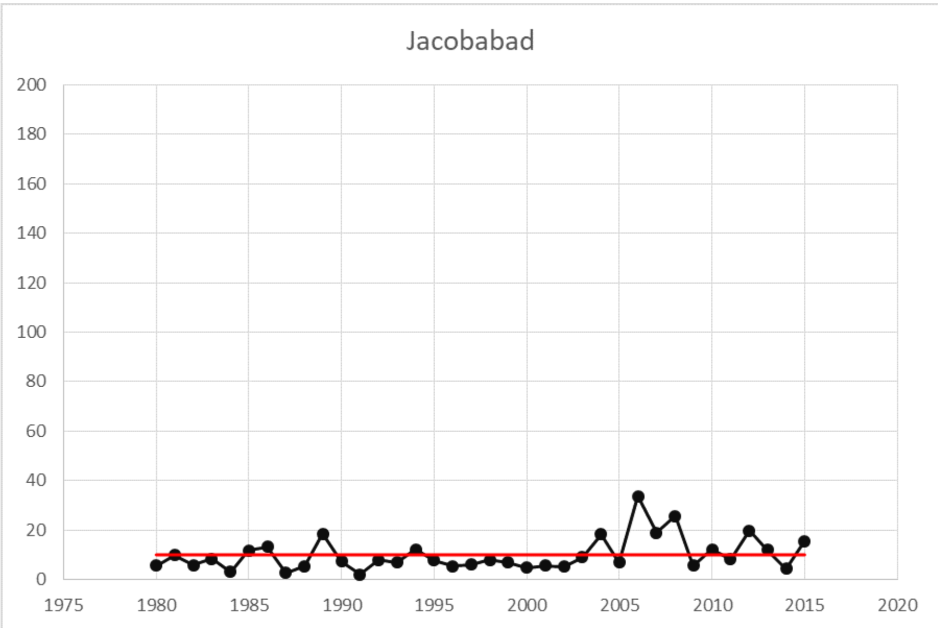


(b) Time series graph for descriptive analysis of Quetta

Figure 4.6: Time Series Graphs of Peshawar and Quetta

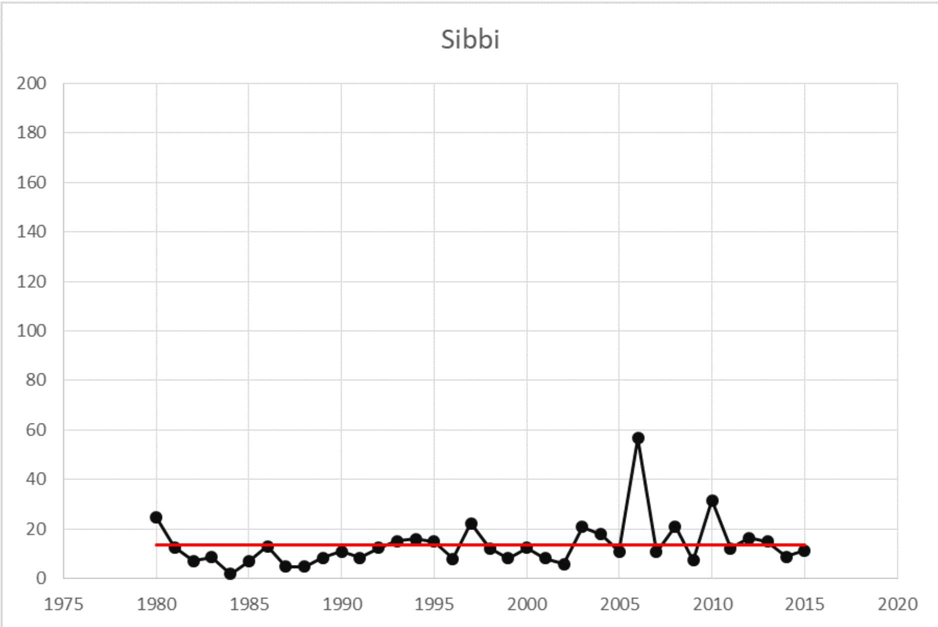


(a) Time series graph for descriptive analysis of Badin

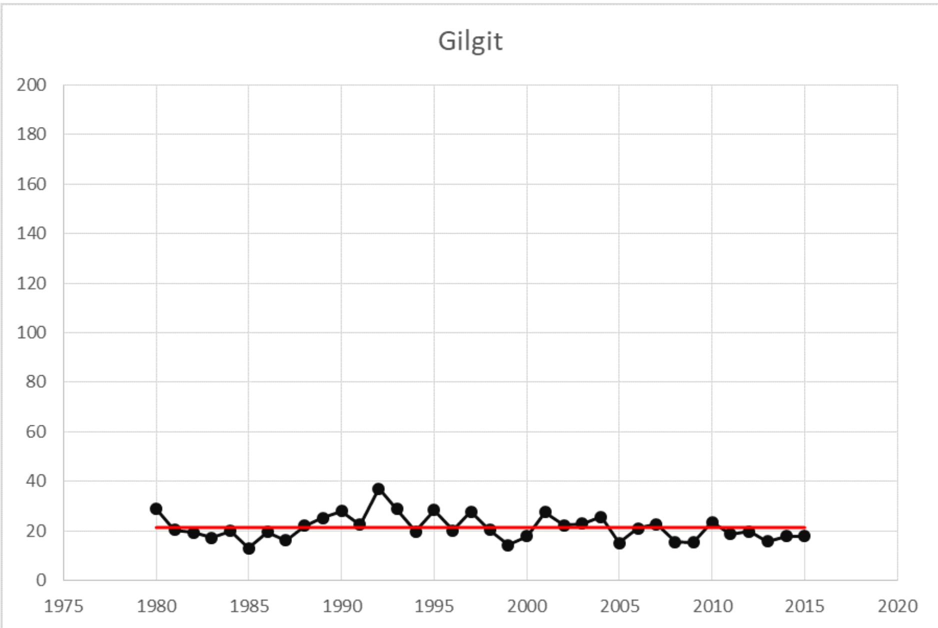


(b) Time series graph for descriptive analysis of Jacobabad

Figure 4.7: Time Series Graphs of Badin and Jocobabad



(a) Time series graph for descriptive analysis of Sibbi



(b) Time series graph for descriptive analysis of Gilgit

Figure 4.8: Time Series Graphs of Sibbi and Gilgit

Table 4.3: Descriptive statistics Mean, Minimum, Maximum (1980-2015)

Sr no	City	Mean	Mean	Minimum	Minimum	Maximum	Maximum
		1980-2000	2001-2015	1980-2000	2001-2015	1980-2000	2001-2015
1.	Badin	11.36	17.52	0.64	8.57	38.61	67.25
2.	Bahawalpur	12.59	15.46	5.71	10.63	22.97	38.79
3.	Chhor	11.88	12.56	1.13	7.19	27.71	24.18
4.	Chitral	30.01	29.19	18.69	20.07	51.67	37.80
5.	Gilgit	22.25	20.13	13.08	15.24	37.15	27.55
6.	Gupis	28.31	26.72	17.70	16.55	44.84	34.65
7.	Islamabad	64.85	53.65	25.29	22.16	157.26	108.53
8.	Jacobabad	7.62	13.41	2.09	4.48	18.44	33.60
9.	Jivani	13.35	16.53	1.43	1.63	36.59	45.78
10.	Karachi	12.16	12.36	0.88	2.85	75.95	27.40
11.	Multan	11.37	14.37	4.023	9.86	17.17	23.44
12.	Peshawar	70.77	70.47	17.21	33.61	123.42	119.13
13.	Quetta	14.99	15.18	7.57	4.74	27.29	28.51
14.	Sibbi	11.15	17.00	1.92	5.87	24.58	56.70

in 2003

Balakot, Murree, Kakul, Peshawar, Skardu, Gupis, Chitral, and Gilgit are high altitude sites of Pakistan. These sites show approximately the same trends. Usually, these sites get more precipitation as compared to other areas of Pakistan. These sites have a peak value between 1992-2000. Gilgit, Gupis, and Skardu have their peaks in 1992, while Peshawar and Chitral had peak values in 1995. Islamabad, Jhelum, Lahore, and Sargodha are neighboring sites. They all show approximately the same trends. Islamabad had a peak value in 1997.

Bahawalpur, Multan, Jacobabad, and Sibbi are neighboring sites and show approximately the same trends. These are low altitude sites of Pakistan, and they usually get low precipitation. Their peak values lie between 2004-2010. Jacobabad and Sibbi have the same peak value in the 2006 year, Multan's most precipitation year is 2005, and Bahawalpur's peak value is 2010.

Table 4.4: Descriptive statistics S.D, Skewness, Kurtosis (1980-2015)

Sr no	City	S.D	S.D	Skewness	Skewness	Kurtosis	Kurtosis
		1980-2000	2001-2015	1980-2000	2001-2015	1980-2000	2001-2015
1.	Badin	7.73	14.01	1.91	3.25	6.50	11.46
2.	Bahawalpur	3.59	6.64	0.90	3.12	2.10	10.91
3.	Chhor	6.48	3.94	0.47	1.68	0.29	4.12
4.	Chitral	8.68	5.35	0.86	-0.01	0.09	-1.16
5.	Gilgit	5.71	3.81	0.74	0.33	0.54	-0.93
6.	Gupis	6.64	5.47	0.51	-0.20	0.34	-0.99
7.	Islamabad	39.81	25.48	1.35	0.77	0.81	-0.12
8.	Jacobabad	3.73	8.24	1.17	1.05	1.97	0.65
9.	Jivani	8.85	14.10	1.76	1.06	2.35	-0.16
10.	Karachi	15.02	7.22	3.85	0.64	16.40	-0.48
11.	Multan	3.12	3.79	-0.31	1.05	0.34	0.63
12.	Peshawar	26.70	24.14	0.03	0.52	-0.14	-0.34
13.	Quetta	5.95	6.21	1.04	0.56	-0.03	0.28
14.	Sibbi	5.38	12.43	0.84	2.38	0.91	6.50

Jivani, Badin, Karachi, and Chhor are geographically neighboring sites, but their trends for rainfall are not the same. These sites also get low precipitation. Jivani site had a peak value in 2011, Karachi site had a peak value in 1989, Badin had peak value in 2003, and Chhor had peak value in 1981.

4.6 Tests

4.6.1 Run test, Lag 1 auto co-relation

This section will present the results of some basic tests, including Run Test and Lag 1 Auto Correlation. The run test determines the randomness in the occurrence of the strings of data. It determines the similarity of different events. The significance of this test is to determine whether the data is generated randomly or depends on underlying variables.

The p-value of the run test determines whether the data is random or not. If the p-value is less than 0.05, then the data is not arbitrary, and if it is more significant than 0.05, then the data is random. In 13 stations p-value is more critical than 0.05, so all these stations have arbitrary data. Only Peshawar shows some deviation as its P-value is less than 0.05 in table 4.5.

Table 4.5: Results of Run test and Auto co-relation

Sr no	Site name	No of observations	Lag 1	Run test	Run test	Group
1.	Bahawalpur	36	0.91	15	0.26	18
2.	Chhor	36	0.05	24	0.05	18
3.	Chitral	36	0.26	21	0.48	18
4.	Gupis	36	0.71	18	0.74	18
5.	Islamabad	36	0.17	18	0.70	18
6.	Jivani	36	0.17	13	0.23	18
7.	Karachi	36	0.43	21	0.30	18
8.	Multan	36	0.51	18	0.79	18
9.	Peshawar	36	0.32	12	0.02	18
10.	Quetta	36	0.21	15	0.26	18
11.	Badin	36	0.76	20	0.60	18
12.	Jacobabad	36	0.43	16	0.70	18
13.	Sibbi	36	0.79	16	0.70	18
14.	Gilgit	36	0.16	14	0.11	18

4.6.2 t-test and F-test

F test is a prerequisite to t-test. F test is used to compare variance, and the t-test is used to compare means at a certain point. In the given data, the year 2000 is selected as a benchmark. F test and t-test are used to compare variance and mean, respectively. It is a statistical hypothesis test in which two hypotheses are drawn. One is the null hypothesis, and the second is the alternative hypothesis. The null hypothesis states that the mean before 2000 and after 2000 is the same, and the alternative states that they are not the same. The level of significance for acceptance of the null hypothesis is 0.05 0r 5 percent. Based on this level of importance, we will decide on accepting or rejecting the null hypothesis. If the p-value of the F test is less than 0.05, then the null hypothesis is rejected, and if the p-value of the F test is equal to or greater than 0.05, then a null hypothesis is accepted. The same is in the case of the t-test. We can reject or accept the null hypothesis based on that p-value significance.

F test is used to compare variance, and the t-test is applied to compare the mean of the given data. As the year 2000 is selected as a benchmark, data is spliced into two parts (before 2000 and after the year 2000). Based on the F-test, we implement a t-test. So firstly, the F-test is applied. Sixteen stations showed that variance before and after are the same as they accepted the null hypothesis, and six stations rejected the null hypothesis in table 4.5. Their trends are showing a significant change in variance after comparison of pre and post-year 2000 data.

The t-test is applied based on the F test. the t-test is used to compare the mean of data. Data is again spliced into two parts (before the year 2000 and after the year 2000). .

Bahawalpur, Karachi, Badin, Jacobabad, and Sibbi are the five stations that reject the null hypothesis. Their p-value for the t-test is less than 0.05, and the significance for the null hypothesis is 0.05. So these stations rejected the null hypothesis. And other nine stations have a p-value of t-test that is more than 0.05, so the null hypothesis is accepted.

Bahawalpur, Karachi, Badin, Jacobabad, and Sibbi have different pre and post variances, while the other nine stations have the same pre and post variances on the T-test in table 4.6.

4.7 Inferential techniques

Inferential techniques are used to predict future trends. Previously given data is used to predict future trends. There are a lot of distributions to predict future trends as previous data include extreme events. So for modeling extreme events, two types of series analysis techniques are given. One is "Time Independent Series Analysis," and the second is "Time-Dependent Series Analysis." Given data is time-independent, so based on previous data, we use "Time Independent Series Analysis." Further, there are three more techniques in "Time Independent Series Analysis." Those are Probability Distribution, Linear Models, and Nonlinear Models/Machine Learning techniques.

Table 4.6: Results of t-test F-test

Sr no	Sites	F Test	P-value	Null Hypothesis
1.	Bahawalpur	0.01	0.29	Reject
2.	Chhor	0.06	2.65	Do not
3.	Chitral	0.07	2.58	Do not
4.	Gupis	0.48	1.44	Do not
5.	Islamabad	0.09	2.39	Do not
6.	Jivani	0.05	0.39	Do not
7.	Karachi	0.01	4.24	Reject
8.	Multan	0.39	0.66	Do not
9.	Peshawar	0.74	1.20	Do not
10.	Quetta	0.81	0.90	Do not
11.	Badin	0.01	0.30	Reject
12.	Jacobabad	0.01	0.20	Reject
13.	Sibbi	0.01	0.18	Reject
14.	Gilgit	0.13	2.20	Do not

According to “Hosking, J.R.M, and Wallis," J.R. (1997), the Probability distribution is the best fit model for analyzing the trends of extreme events with three parameters (location, scale, and shape). Further, four models are selected to check the best fit model for given data and to predict future trends. According to "Rao, A.R., and Hamed," K.H. (2000), the best fit four models for modeling extreme events include Generalized Extreme Value in table 4.7, Generalized Pareto in table 4.8, Inverse Gaussian in table 4.9, and Log Pearson 3 in table 4.10.

To check the best distribution, three tests are drawn for every station. Kolmogorov Smirnov, Kolmogorov Smirnov, Chi-Squared are the three test. Based on these three tests, the best fit distribution is selected.

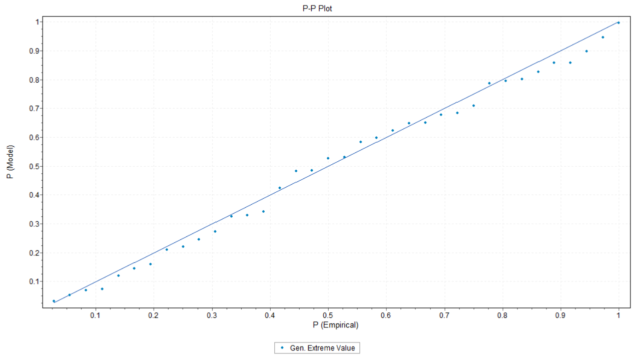
Kolmogorov Smirnov is a non-parametric test. It is to check the goodness of fit. It has the advantage of not assume distribution. To check the normality of distribution, data is standardized and then compared with a standard normal distribution.

Anderson darling test is also a non-parametric test, and it also has the advantage of making no assumptions. It is also applied to check the goodness of fit. This test is used to check whether data come from a specific distribution or not. It is a modified form of the Kolmogorov Smirnov test. There are two hypotheses in this test. The null hypothesis is that the data has a specific distribution, and the alternate hypothesis is that the data has not specific distribution. Based on 0.05 significance, any hypothesis is accepted or rejected. The Chi-square test is a parametric test, and parametric tests are superior to non-parametric tests, so it is superior to the Kolmogorov Smirnov test and Anderson darling test.

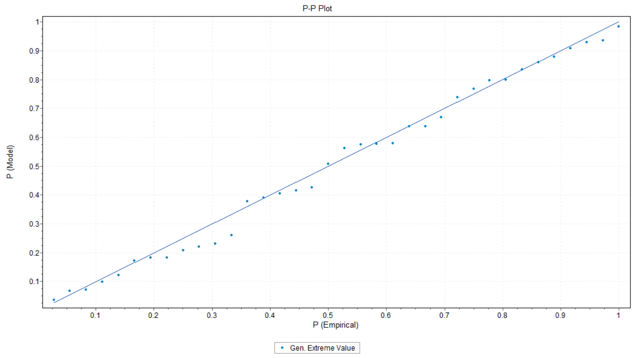
Based on these three tests, the best fit distribution has been chosen. In all fourteen stations, the Generalized Extreme value is the best fit distribution. So, for finding quantiles, Generalized Extreme Value distribution is used.

Table 4.7: Extreme value analysis using GEV Ranks

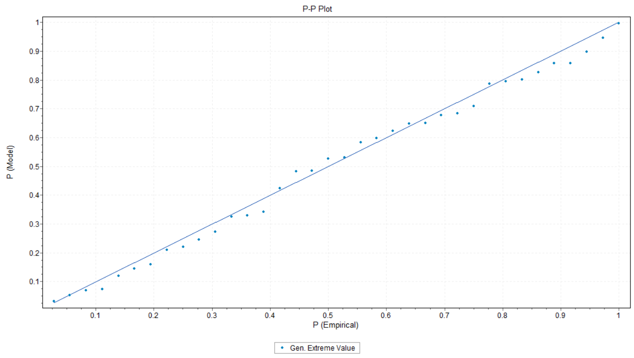
Sites	obs	KS	KS	AD	CS	CS	GEV RANKS
		Stat value	P-value	Stat value	Stat value	P-value	
Bahawalpur	36	0.102	0.80655	0.6826	1.8464	0.39726	k/1,a/1,c/1
Chhor	36	0.085	0.9358	0.36504	2.1138	0.71483	k/1,a/1,c/1
Chitral	36	0.074	0.97838	0.17644	0.68438	0.87687	k/1,a/1,c/3
Gupis	36	0.065	0.99515	0.16874	0.87382	0.92829	k/1,a/1,c/2
Islamabad	36	0.081	0.95659	0.31133	0.54106	0.96938	k/1,a/3,c/1
Jivani	36	0.125	0.58336	0.85767	2.9353	0.56871	k/1,a/1,c/2
Karachi	36	0.138	0.45752	0.58034	0.66638	0.88108	K/1,a/1,c/1
Multan	36	0.104	0.78592	0.40675	0.35777	0.99641	k/1,a/1,c/1
Peshawar	36	0.10	0.76473	0.31431	0.27192	0.96522	K/2,a/1,c/1
Quetta	36	0.07	0.96881	0.28236	0.432	0.97977	k/1,a/1,c/1
Badin	36	0.15	0.33206	1.1625	0.5272	0.91288	k/1,a/2,c/1
Jacobabad	36	0.08	0.96014	0.30922	0.36381	0.98533	k/1,a/1,c/1
Sibbi	36	0.089	0.91367	0.31268	0.17577	0.9814	k/1,a/1,c/1
Gilgit	36	0.0739	0.98107	0.22172	1.1644	0.88393	k/1,a/1,c/2



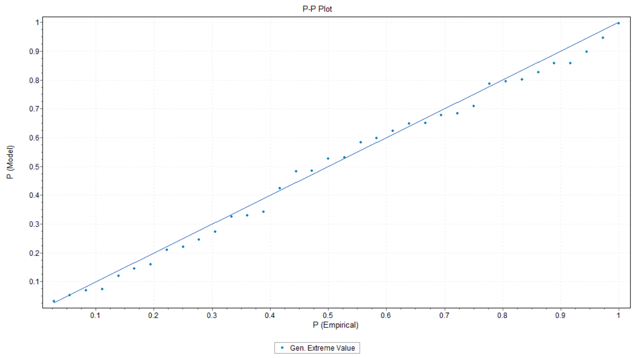
(a) a



(b) b

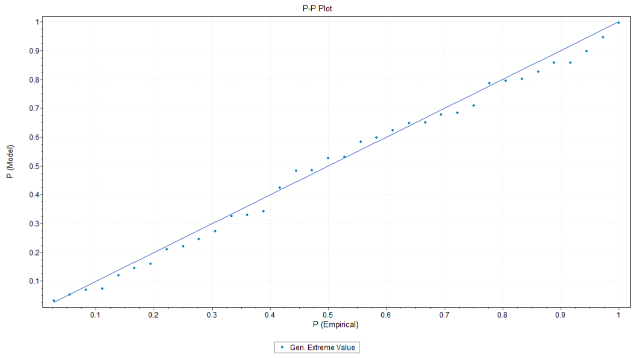


(c) c

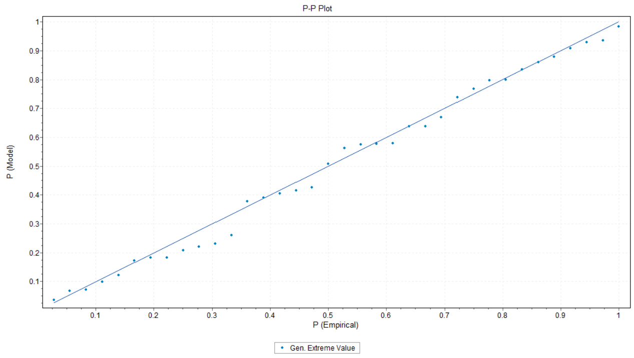


(d) d

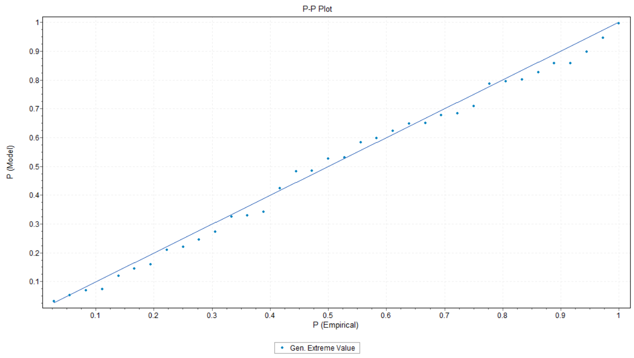
Figure 4.9: PP plots of (a) Bahawalpur, (b) Chhor, (c) Chitral, (d) Gupis



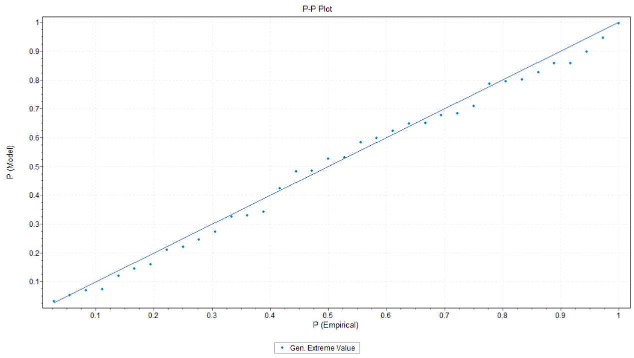
(a) e



(b) f

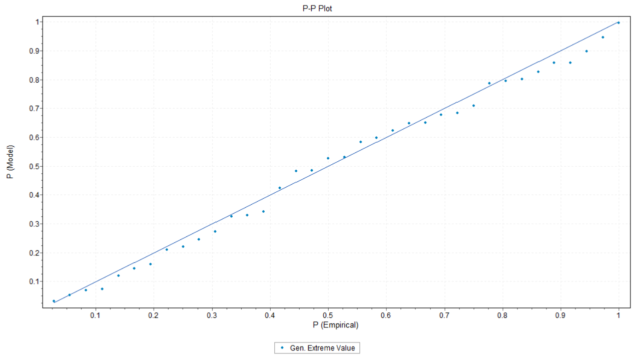


(c) g

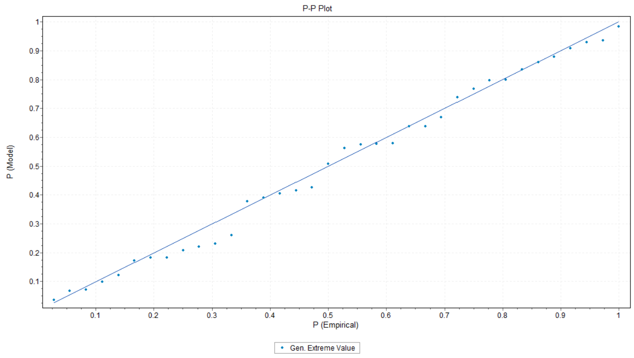


(d) h

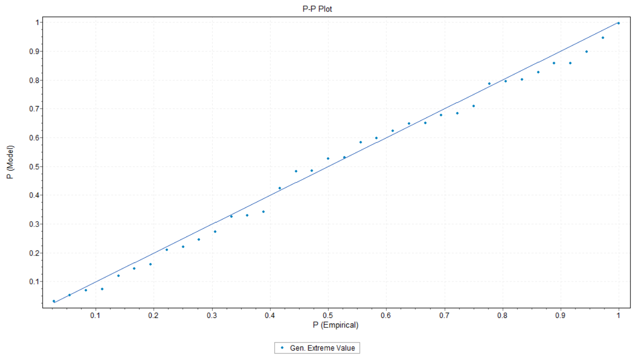
Figure 4.10: PP plots of (e) Islamabad, (f) Jivani, (g) Karachi, (h) Multan



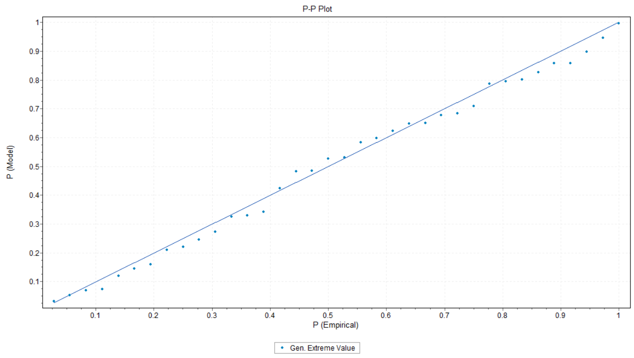
(a) i



(b) j

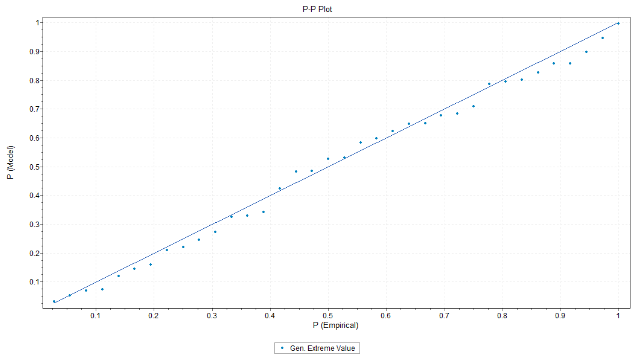


(c) k

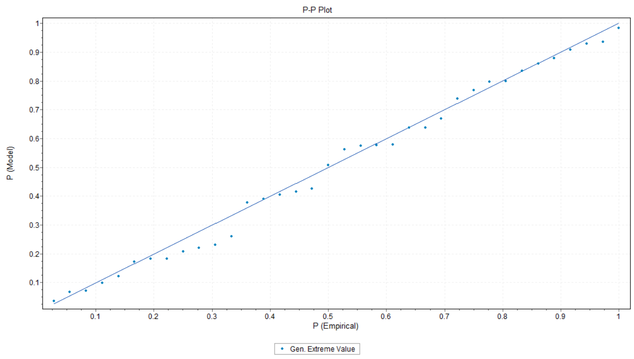


(d) l

Figure 4.11: PP plots of (i) Peshawar, (j) Quetta, (k) Badin, (l) Jacobabad



(a) m



(b) n

Figure 4.12: PP plots of (m) Sibbi, (n) Gilgit

Table 4.8: Extreme value analysis using Gen. Pareto

Sites	obs	KS	KS	AD	CS	CS	GEN PARETO
		Stat value	P-value	Stat value	Stat value	P-value	
Bahawalpur	36	0.13253	0.50982	8.3056			K/4, A/5
Chhor	36	0.11768	0.65788	11.671			k/2,a/4
Chitral	36	0.07561	0.97643	3.9712			k/2,a/5
Gupis	36	0.09039	0.9047	4.1669			k/5,a/5
Islamabad	36	0.08688	0.92695	0.29876	1.7488	0.78184	k/3,a/1,c/4
Jivani	36	0.16066	0.27956	12.053			k/5,a/5
Karachi	36	0.17642	0.18852	11.664			k/4,a/5
Multan	36	0.14566	0.14566	8.2739			k/5,a/5
Peshawar	36	0.13967	0.44365	7.9769			k/5,a/5
Quetta	36	0.09361	0.88164	4.1698			k/5,a/5
Badin	36	0.18777	0.13862	12.185			k/3,a/5
Jacobabad	36	0.11556	0.67949	8.6614			k/5,a/5
Sibbi	36	0.11722	0.66261	11.345			k/4,a/5
Gilgit	36	0.09036	0.90488	4.2816			k/4,a/5

4.8 Quantiles

Based on best fit distribution, quantiles of 10, 20, 30, 50, and 100 years for the fourteen stations are built. The following table 4.11 shows quantiles for each station.

Table 4.9: Extreme value analysis using Inverse Gaussian

Sites	obs	KS	KS	AD	CS	CS	INV GAUSSIAN
		Stat value	P-value	Stat value	Stat value	P-value	
Bahawalpur	36	0.11178	0.71759	0.8287	3.2715	0.35164	K/2, A/2, C/2
Chhor	36	1	0				k/5
Chitral	36	0.08209	0.95207	0.19929	0.54276	0.99047	k/3,a/3,c/1
Gupis	36	0.07107	0.98732	0.27399	0.98306	0.91235	k/2,a/3,c/3
Islamabad	36	0.09414	0.87761	0.38683	4.1222	0.24857	k/5,a/5,c/5
Jivani	36	0.14799	0.37277	0.98712	3.8566	0.27737	k/3,a/2,c/3
Karachi	36	0.13826	0.45631	0.62311	0.83231	0.84173	k/2,a/2,c/2
Multan	36	0.13282	0.507	0.47752	3.1796	0.52823	K3,a/2,c/3
Peshawar	36	0.11174	0.718	0.36077	1.2971	0.72982	k/3,a/2,c/4
Skardu	36	0.10283	0.80402	0.28995	0.11049	0.990055	k/5,a/1,c/1
Badin	36	0.17327	0.20459	1.1615	10.506	0.01472	k/2,a/1,c/3
Jacobabad	36	0.09599	0.86312	0.35823	2.1288	0.54611	k/4,a/3,c/4
Sibbi	36	0.09007	0.90683	0.43564	0.6406	0.72593	k/2,a/3,c/3
Gilgit	36	0.07603	0.97517	0.22993	0.44799	0.93016	k/2,a/3,c/1

Table 4.10: Extreme value analysis using Log Pearson III

Sites	obs	KS	KS	AD	CS	CS	LOG PEARSON 3
		Stat value	P-value	Stat value	Stat value	P-value	
Bahawalpur	36	0.11413	0.69396	4.493			K/3, A/4
Chhor	36	0.12718	0.56198	11.612			k/4,a/3
Chitral	36	0.08266	0.94942	0.18811	0.58382	0.96484	k/4,a/2,c/2
Gupis	36	0.0777	0.96975	0.19261	0.06869	0.99942	k/4,a/2,c/1
Islamabad	36	0.08729	0.92447	0.30634	1.0577	0.90093	k/4,a/2,c/2
Jivani	36	0.15405	0.32591	1.0418	4.3763	0.2236	k/4,a/3,c/4
Karachi	36	0.14413	0.40473	0.87974	1.2342	0.74481	K/3,a/3,c/3
Multan	36	0.13988	0.44176	0.91959	3.4171	0.4906	K/4,A/4,C/4
Peshawar	36	0.10291	0.80332	0.37357	1.1654	0.76132	k/1,a/3,c/3
Quetta	36	0.08925	0.91225	0.32126	1.4304	0.69842	k/4,a/3,c/4
Badin	36	0.21912	0.05352	9.1558			k/5,a/4
Jacobabad	36	0.09062	0.9031	0.33158	0.98871	0.80398	k/2,a/2,c/2
Sibbi	36	0.09312	0.8853	0.41843	0.63877	0.63877	k/3,a/2,c/2
Gilgit	36	0.07653	0.97361	0.22521	1.1787	0.8816	k/3,a/2,c/3

Table 4.11: Quantile for 14 Stations based on Best-Fit Distribution

Site	Distrib	Location	Scale	Shape	Q10	Q20	Q30	Q50	Q100
Bahawalpur	GEV	11.686	3.3142	-0.059	19.67	22.45	24.11	26.24	29.23
Chhor	GEV	9.9007	5.1053	0.14	19.61	22.20	23.53	25.084	27.002
Chitral	GEV	26.111	5.7770	-0.034	39.61	44.19	46.86	50.26	54.9
Gupis	GEV	25.133	5.7697	0.164	35.93	38.69	40.10	41.74	73.75
Islamabad	GEV	41.138	18.692	-0.37	107.25	143.1	168.64	206.6	271.08
Jivani	GEV	8.9109	6.2763	-0.27	28.36	37.54	43.69	52.42	66.30
Karachi	GEV	7.3704	5.4719	-0.21	23.30	30.30	34.86	41.17	50.95
Multan	GEV	11.139	3.4828	0.19	17.70	19.35	20.21	21.21	22.40
Peshawar	GEV	61.256	24.862	0.25	104.01	113.3	117.97	123.13	129.13
Quetta	GEV	12.224	4.7129	-0.01	23.03	26.55	28.59	31.16	34.09
Badin	GEV	9.4730	6.0701	-0.125	25.25	31.32	35.06	40.03	47.26
Jacobabad	GEV	6.7125	3.6726	-0.26	18.05	23.37	26.52	31.97	39.97
Sibbi	GEV	9.5428	5.0802	-0.17	23.49	29.21	32.82	37.74	45.11
Gilgit	GEV	19.053	4.1205	0.019	28.12	30.94	32.51	34.53	37.11

Chapter 5

Discussion

5.1 Discussion

In this section above, we have already shown the tabular and graphical results. Now we will elaborate on notable findings of our results. Firstly, we have to confirm and validate the randomness of the data of all the cities in the dataset. For this purpose, we have to opt to run a test. Among all the cities, data of Peshawar shows that it's not random. This fact is advocated by the run test results and graphical depiction of data, which shows patterns. On the contrary, data of all other cities were random as confirmed by run test and visual representation.

To compare the pre and post-data of year- 2000, a t-test was applied to all the cities. The test results show that Bahawalpur, Karachi, Quetta, Jacobabad, and Sibbi have equal pre and post year-2000 variance. All other sites have unequal variance. year 2000 is selected as benchmark. F test is pre requisite to t test. and used to check variance. t test is used to check mean of pre and post year 2000. as per t test 5 cities have equal pre and post means. other 9 cities have different pre and post means.

To test the closeness and agreement of data of sites, pp plots were drawn. These plots show the skewness of the data.

To forecast the data, we had to choose the best fit probability distribution. In this regard, we have used four different tests, i.e., Chi-Square, Anderson darling and Kolmogorov Smirnov. The test results show that Generalized Extreme Value was most suitable for our data. This selection of distribution was made on rankings of all four distributions. The ranking itself was decides based on a p-value. Using GEV, we have computed quantiles, i.e., Q10, Q20, Q30, Q50, and Q100. These quantiles predict weather events. GEV is the best fit model on the basis of three tests i.e Chi-Square, Anderson darling and Kolmogorov Smirnov.

Q10, Q20, Q30, Q50, and Q100, Islamabad, and Peshawar have repeatedly shown extreme weather behaviour in table 4.11. All other sites at all the quantiles offer regular weather events. This research aims to predict extreme weather events, so this research recommends the necessary actions to avoid the calamity of natural disasters.

Chapter 6

Conclusion

In this section, we will discuss the overall finding of the research and conclude this document.

It's a fact that Pakistan has repeatedly faced floods in the last two decades which existing models have failed to predict precisely and timely. Lack of water reservoirs has left this country at mercy of nature. This research work aimed to perform the statistical analysis of extreme rainfall at various met stations in Pakistan. This research domain was selected by keeping in regard the devastating effects of extreme weather conditions in a country badly hit by climate change.

In this research, standardized data validation policies were followed. For example, in order to test the consensus, the closeness of data through skewness, pp-plots were used. Not only graphical confirmation but, statistical t-test was also performed on data to test the randomness. As far as methodology is concerned, the methodological pipeline was established after an extensive literature review. In this research, only those statistical tests and variables have opted which most suited the data and were also recommend by the researchers in the literature review. As directed by literature, Chi-Square, Anderson Darling, Generalized Extreme Value, and Kolmogrovo Smirnov were selected. Still, there is a lot of research gap in this domain specifically in dataset and scope expansion

Considering our findings in the form of statistical evidence, (out of 14 stations) Islamabad and Peshawar are expected to face extreme weather in the future. In this regard, policymakers and concerned authorities are required to take measures to establish a national consensus against climate change, specifically for the more threatened cities.

6.1 Future work

Scope of this research can be extended by adding more stations. as number of stations from all over the Pakistan is increased for future research. further scope of research can be extended to region wise. Machine learning models can be used to predict extremes of rainfall. further updated data is used to predict more accurate and reliable estimates. in future, variable is changed to daily maximum rainfall data or average yearly rainfall data Moreover, it is expected that obtained results of rainfall regions can lead to further investigations of similar nature and scope with emphasis on rainfall changes in future. The rainfall changes is a hot issue, particularly for the agriculture sector, food security and one of the major concerns for A Classification of Rainfall Regions in Pakistan scientists working in the field of climate change, for there still many more needs to do research work to understand the changing behavior of rainfall. Consequently, next study will present the rainfall changes in identified rainfall regions in Pakistan.

References

[1] Mian Sabir Hussain and Seungho Lee. A classification of rainfall regions in pakistan. , 44(5):605–623, 2009.

[2] Muhammad Qasim, Salma Khalid, and Dilawar Shams. Spatiotemporal variations and trends in minimum and maximum temperatures of pakistan. 4:85–93, 01 2014.

[3] Zamir Hussain and GR Pasha. Regional flood frequency analysis of the seven sites of punjab, pakistan, using l-moments. *Water resources management*, 23(10):1917–1933, 2009.

[4] John Theodore Houghton, Geoffrey J Jenkins, and Jim J Ephraums. Climate change. 1990.

[5] Ishfaq Ahmad, Aamar Abbas, Aamir Saghir, and Muhammad Fawad. Finding probability distributions for annual daily maximum rainfall in pakistan using linear moments and variants. *Pol. J. Environ. Stud*, 25(3):925–937, 2016.

[6] Ghazala Naheed and Ghulam Rasul. Investigation of rainfall variability for pakistan. *Pakistan Journal of Meteorology*, 7(14):25–32, 2011.

[7] S Salma, S Rehman, and MA Shah. Rainfall trends in different climate zones of pakistan. *Pakistan Journal of Meteorology*, 9(17), 2012.

[8] Takahiro Sayama, Go Ozawa, Takahiro Kawakami, Seishi Nabesaka, and Kazuhiko Fukami. Rainfall–runoff–inundation analysis of the 2010 pakistan flood in the kabul river basin. *Hydrological Sciences Journal*, 57(2):298–312, 2012.

[9] Shahzada Adnan and Azmat Hayat Khan. Effective rainfall for irrigated agriculture plains of pakistan. *Pakistan Journal of Meteorology*, 6(11):61–72, 2009.

[10] Syed Ali Raza, Yasir Ali, and Farhan Mehboob. Role of agriculture in economic growth of pakistan. 2012.

[11] Anam Azam and Muhammad Shafique. Agriculture in pakistan and its impact on economy. *A Review. Inter. J. Adv. Sci. Technol*, 103:47–60, 2017.

[12] Sergio Albeverio, Volker Jentsch, and Holger Kantz. *Extreme events in nature and society*. Springer Science & Business Media, 2006.

[13] Ishfaq Ahmad, Aamir Abbas, Muhammad Fawad, and Aamir Saghir. Regional frequency analysis of annual total rainfall in pakistan using l-moments. *NUST Journal of Engineering Sciences*, 10(1):19–29, 2017.

[14] Amina Shahzadi, Ahmad Saeed Akhter, and Betul Saf. Regional frequency analysis of annual maximum rainfall in monsoon region of pakistan using l-moments. *Pakistan Journal of Statistics and Operation Research*, pages 111–136, 2013.

[15] M Latif, Abdelwaheb Hannachi, and FS Syed. Analysis of rainfall trends over indo-pakistan summer monsoon and related dynamics based on cmip5 climate model simulations. *International Journal of Climatology*, 38:e577–e595, 2018.

[16] Zarrien Ayub. Effect of temperature and rainfall as a component of climate change on fish and shrimp catch in pakistan. *The Journal of Transdisciplinary Environmental Studies*, 9(1):1–9, 2010.

[17] Muhammad Athar Haroon and Ghulam Rasul. Principal component analysis of summer rainfall and outgoing long-wave radiation over pakistan. *Pakistan Journal of Meteorology*, 5(10):109–114, 2009.

[18] H Athar et al. Water supply and effective rainfall impacts on major crops across irrigated areas of punjab, pakistan. *Theoretical and Applied Climatology*, 142(3): 1097–1116, 2020.

[19] Umar K Mirza, Nasir Ahmad, Tariq Majeed, and Khanji Harijan. Hydropower use in pakistan: past, present and future. *Renewable and Sustainable Energy Reviews*, 12(6):1641–1651, 2008.

[20] Muhammad Akram Kahlown and Abdul Majeed. Water-resources situation in pakistan: challenges and future strategies. *Water Resources in the South: present scenario and future prospects*, 20:33–45, 2003.

[21] Qamar Uz Zaman Chaudhry. *Climate change profile of Pakistan.* Asian Development Bank, 2017.

[22] Waqas Shabbir, Juergen Pilz, and Amna Naeem. A spatial-temporal study for the spread of dengue depending on climate factors in pakistan (2006–2017). *BMC Public Health*, 20(1):1–10, 2020.

[23] Rajesh K Mall, Ravindra K Srivastava, Tirthankar Banerjee, Om Prakash Mishra, Diva Bhatt, and Geetika Sonkar. Disaster risk reduction including climate change adaptation over south asia: challenges and ways forward. *International Journal of Disaster Risk Science*, 10(1):14–27, 2019.

[24] Farah Yasmeen and Shaheen Hameed. Forecasting of rainfall in pakistan via sliced functional times series (sfts). *World Environment 2018*, 8(1).

[25] Tae Kyun Kim. T test as a parametric statistic. *Korean journal of anesthesiology*, 68(6):540, 2015.

[26] Will Thalheimer and Samantha Cook. How to calculate effect sizes from published research: A simplified methodology. *Work-Learning Research*, 1:1–9, 2002.

[27] Ayman Baklizi. Testing symmetry using a trimmed longest run statistic. *Australian & New Zealand Journal of Statistics*, 49(4):339–347, 2007.

[28] Pierre Legendre. Spatial autocorrelation: trouble or new paradigm? *Ecology*, 74 (6):1659–1673, 1993.

[29] Pierre Ailliot, Craig Thompson, and Peter Thomson. Mixed methods for fitting the gev distribution. *Water Resources Research*, 47(5), 2011.

[30] Enrique Castillo and Ali S Hadi. Fitting the generalized pareto distribution to data. *Journal of the American Statistical Association*, 92(440):1609–1620, 1997.