

**Ensemble Projections on Climate Change using CMIP-6 Data
over Baluchistan, Pakistan**



By

Beenish Javed

Regn. #0000365090

Department of Mathematics and Statistics

School of Natural Sciences

National University of Sciences and Technology

H-12, Islamabad, Pakistan

August 2023

**Ensemble Projections on Climate Change using CMIP-6 Data
over Baluchistan, Pakistan**



By

Beenish Javed

Regn. #0000365090

A thesis submitted in partial fulfillment of the requirements
for the degree of **Master of Science**
in
Statistics

Supervised by: Dr. Firdos Khan


Department of Mathematics and Statistics


School of Natural Sciences
National University of Sciences and Technology
H-12, Islamabad, Pakistan


August 2023

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS thesis written by **Beenish Javed** (Registration No. **00000365090**), of **School of Natural Sciences** has been vetted by undersigned, found complete in all respects as per NUST statutes/regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS/M.Phil degree. It is further certified that necessary amendments as pointed out by GEC members and external examiner of the scholar have also been incorporated in the said thesis.


Signature: 
Name of Supervisor: Dr. Firdos Khan
Date: 08-09-2023

Signature (HoD): 
Date: 8/9/2023

Signature (Dean/Principal): 
Date: 12.9.2023

National University of Sciences & Technology**MS THESIS WORK**

We hereby recommend that the dissertation prepared under our supervision by: "**Beenish Javed**" Regn No. **00000365090** Titled: "**Ensemble Projections on Climate change using CMIP-6 data over Baluchistan, Pakistan**" accepted in partial fulfillment of the requirements for the award of **MS** degree.

Examination Committee Members1. Name: PROF. TAHIR MAHMOODSignature: 2. Name: DR. SHAKEEL AHMEDSignature: Supervisor's Name: DR. FIRDOS KHANSignature: Head of Department8/9/2023

Date

COUNTERSIGNEDDate: 12.9.2023
Dean/Principal

Dedication

With profound appreciation to my parents, for their unending encouragement, and to my teachers, for their patient guidance and unyielding faith in my abilities.

Acknowledgements

*Completing this thesis is not merely an academic milestone, but a profound journey guided by Allah's majestic benevolence. His blessings have instilled a sense of purpose, fostering resilience in the face of complexity and uncertainty. With each word of gratitude, I acknowledge the perpetual source of strength that has illuminated my path, shaping this scholarly expedition into a tapestry of faith, diligence, and devotion. With immense gratitude, I want to acknowledge my parents for their endless encouragement, which has been my anchor throughout my academic journey. Your unshakable faith in me has propelled me to achieve more than I could have imagined. I also want to express my appreciation to my teachers, specially **Dr. Firdos Khan**, whose passion for teaching and commitment to fostering a love of learning have left an indelible mark on my education. Your dedication has enriched my intellectual growth and has significantly contributed to the completion of this thesis.*

Abstract

Climate change's effect have sparked several problems that go beyond Pakistan's border. Extreme occurrences could be more severe because of raising temperature, raising concerns, about their potential impacts. The purpose of this study was to investigate the impact of climate change on distinct climatic zones in the province Baluchistan, Pakistan. The study analyzed precipitation and temperature data from 12 meteorological sites using the Reconnaissance Drought Index (RDI). A fusion of two statistical approaches was used to generate homogeneous climatic regions (HCR's). In the beginning, cluster analysis was used to classify climatic regions (CRs) by considering the distinctive characteristics of the station's sites. Through this process, three distinct climatic areas were identified. The second phase comprised the use of heterogeneity and discordancy metrics, as well as the L-moment techniques to assure the consistency of the established HCRs. These HCRs were then evaluated using GCM's specifically for the reference period of 1985 to 2014. We established these HCR's to make predictions about the climate using two alternative scenarios Shared Socioeconomic Pathways (SSPs 4.5 and 8.5) for the twenty-first century. Despite changes in temperature and precipitation patterns, the results from the HCR's show that all the stations preserve their distinct regions in both future scenarios. The maximum and minimum temperatures are expected to rise significantly in the future, with a significant rise predicted in 2075-2100 for SSP 4.5 and 8.5. This research will provide vital support for both immediate and long-term climate change policies, as well as projects linked to irrigation, water management, hydropower, and other relevant areas around the country.

Keywords: Bias Correction; Cluster analysis; Ensemble Projection; L-moments; CMIP-6; Shared Socioeconomic Pathways; Taylor Diagram; K-folds cross validation

Contents

List of figures	vi
List of Tables	vii
List of Abbreviations	viii
1 Introduction	1
1.1 Objectives	5
2 Study Area and Data set	6
3 Methodology	9
3.1 Droughts and RDI	11
3.1.1 Calculation of RDI	11
3.2 L-moment	12
3.3 Cluster Analysis	14
3.4 Data screening and homogeneity test	15
3.5 Weighted Average	16
3.6 Bias Correction	17
3.7 K-folds cross validation	18
3.7.1 Model Training and Testing	19
3.7.2 Performance Evaluation	19
3.7.3 Performance Aggregation	19
3.7.4 Model Choice	19

4 Results	21
4.1 Construction of HCRs	21
4.1.1 Region 1	23
4.1.2 Region 2	25
4.1.3 Region 3	26
4.2 Evaluation of HCRs using observed data and model baseline data	27
4.3 Assessment of combine model’s output and output of GCMs	28
4.3.1 Region1	28
4.3.2 Region 2	30
4.3.3 Region 3	32
4.4 Evaluation of HCRs using Ensemble Models output	35
4.5 Impact assessment of climate change on HCRs using SSP 4.5 and SSP 8.5 Scenarios	36
4.6 Evaluation of ensemble model output	39
4.7 Ensemble Climate Projections	45
4.7.1 Region 1	46
4.7.2 Region 2	48
4.7.3 Region 3	50
5 Conclusions and Recommendations	54
5.1 Conclusion	54
5.2 Recommendation	55
Bibliography	58
Appendix	61

List of Figures

2.1	Study area of province Baluchistan, Pakistan	7
3.1	Methodology Flowchart	10
4.1	A generated dendrogram displaying the HCRs produced using observed data from the Baluchistan region of Pakistan from 1987-2014	23
4.2	Construction of HCRs using observed data and model baseline data	27
4.3	Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of maximum temperature of Region 1	29
4.4	Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of minimum temperature of Region 1	29
4.5	Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of precipitation of Region 1	30
4.6	Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of maximum temperature of Region 2	31
4.7	Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of minimum temperature of Region 2	31

4.8	Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of precipitation of Region 2	32
4.9	Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of maximum temperature of Region 3	33
4.10	Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of minimum temperature of Region 3	34
4.11	Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of precipitation of Region 3	34
4.12	Construction of HCRs using SSP 4.5 and SSP 8.5 from 2017-2044	37
4.13	Construction of HCRs using SSP 4.5 and SSP 8.5 from 2045-2072	38
4.14	Construction of HCRs using SSP 4.5 and SSP 8.5 from 2073-2100	38
4.15	This graphs shows maximum temperature of zone 1 using model data, baseline data and observed data.	41
4.16	This graphs shows minimum temperature of zone 1 using model data, baseline data and observed data.	41
4.17	This graphs shows rainfall pattern of zone 1 using model data, baseline data and observed data.	42
4.18	This graphs shows maximum temperature of zone 2 using model data, baseline data and observed data.	42
4.19	This graphs shows minimum temperature of zone 2 using model data, baseline data and observed data.	43
4.20	This graphs shows rainfall pattern of zone 2 using model data, baseline data and observed data.	43
4.21	This graphs shows maximum temperature of zone 3 using model data, baseline data and observed data.	44

4.22	This graphs shows minimum temperature of zone 3 using model data, baseline data and observed data.	44
4.23	This graphs shows rainfall pattern of zone 3 using model data, baseline data and observed data.	45
4.24	Future Maximum Temperture trend since (2017-2100) of region 1 using SSP 4.5 and SSP 8.5	47
4.25	Future Minimum Temperture trend since (2017-2100) of region 1 using SSP 4.5 and SSP 8.5	47
4.26	Future Rainfall trend since (2017-2100) of region 1 using SSP 4.5 and SSP 8.5	48
4.27	Future Maximum Temperture trend since (2017-2100) of region 2 using SSP 4.5 and SSP 8.5	49
4.28	Future Minimum Temperture trend since (2017-2100) of region 2 using SSP 4.5 and SSP 8.5	50
4.29	Future Rainfall trend since (2017-2100) of region 2 using SSP 4.5 and SSP 8.5	50
4.30	Future Maximum Temperture trend since (2017-2100) of region 3 using SSP 4.5 and SSP 8.5	52
4.31	Future Minimum Temperture trend since (2017-2100) of region 3 using SSP 4.5 and SSP 8.5	52
4.32	Future Rainfall trend since (2017-2100) of region 3 using SSP 4.5 and SSP 8.5	53

List of Tables

2.1	A brief information including latitude, longitude, elevation, maximum and minimum temperature, and mean annual precipitation at each meteorological station of Baluchistan, Pakistan	8
3.1	Droughts classification thresholds	13
4.1	Discordancy values (D_i) and Heterogeneity measures (H_1)	22
4.2	Global Climate Models	40

List of Abbreviations

GCMs	Global Climate Models
SSPs	Shared Socioeconomics Pathways
HCRs	Homogeneous Climatic Regions
PET	Potential Evapotranspiration
RDI	Reconnaissance Drought Index
CDFs	Commulative Distribution Functions
MAP	Mean Annual Precipitation
PWM	Probability Weighted Moments
DRINC	Drought Indices Calculator
CV	Coefficients of Variation
RMSE	Root Mean Square Error
CMIP-6	Coupled Model Intercomparison Phase 6

Chapter 1

Introduction

Long-term changes in the Earth's climate, such as those in global temperature, precipitation, and sea level, are referred to as "climate change." Human activities, which have escalated since the Industrial Revolution, are the main source of these changes. Significant volumes of greenhouse gases are released into the atmosphere through the burning of fossil fuels used for energy production, transportation, and industrial activities. These emissions, which include carbon dioxide (CO₂), methane (CH₄), and nitrous oxide (N₂O), have been identified as the main causes of global warming and climate change by the Intergovernmental Panel on Climate Change. Naomi Oreskes in 2004, [1] asserts that the buildup of greenhouse gases creates a thermal blanket that traps emitted infrared light and prevents it from escaping into space. The greenhouse effect, a phenomenon, causes an increase in the average surface temperature of the Earth. The effects of this global warming trend are extensive and have an effect on several facets of the climate system on Earth. The increase in global temperature is one of the most notable consequences of climate change. The warmest years on record have occurred in recent decades, and the average surface temperature of the Earth has dramatically increased during the last century. Sea levels have risen as a result of glaciers and polar ice caps melting as a result of rising temperatures. Due to increasing floods, erosion, and loss of habitat for many species, this poses serious dangers to coastal towns, low-lying islands, and sensitive ecosystems. Climate change has significant impacts on the environment and human societies. It is causing melting

of polar ice caps and glaciers, leading to rising sea levels that threaten coastal communities and island nations. It is also leading to more frequent and severe weather events such as hurricanes, floods, droughts, and wildfires, which can cause loss of life, damage to infrastructure, and economic losses [2]. There are significant global consequences for agriculture and food security due to climate change. A considerable challenge to agricultural productivity is posed by the changing climate patterns and the frequency and severity of extreme weather events, which might disrupt food supply systems and jeopardise global food security.

Crop yields and production may be negatively impacted by changing weather patterns, including changed rainfall patterns and rising temperatures. The growth, development, and yields of crops can all be severely impacted by extreme heat, droughts, and water scarcity. On the other hand, excessive rain and flooding can cause agricultural damage, waterlogging, and soil erosion. Farmers may experience decreased crop yields and financial losses as a result of their inability to foresee and manage their agricultural activities properly due to these weather changes [3]. The unexpected nature of climate change only adds to its complexity. Climate change's effects can be clearly linked to specific occurrences or phenomena, but they can also differ significantly throughout time and space. A dynamic process, climate change is impacted by a variety of elements, such as greenhouse gas emissions, land use patterns, natural temperature variability, and feedback processes within the Earth's system. It is difficult to anticipate and completely comprehend the scope and timing of its repercussions due to its intricacy. Understanding how our global system is interrelated is essential. Climate changes can have ripple effects that cut over geographic borders and have an impact on many different facets of Earthly existence. For instance, melting ice caps and rising sea levels have an effect on coastal communities as well as marine ecosystems and ocean currents, which in turn influence fisheries and biodiversity. Temperature and rainfall variations can affect agricultural output, causing food shortages and affecting the world's food security. Ecosystems can be harmed by the loss of natural habitats brought on by climate change, which can result in the extinction of some species and a reduction in biodiversity as a whole [4]. The examination of temperature and rain-

fall patterns has captured the attention of scientists all across the world in the 20th century. For the purpose of creating efficient adaptation and mitigation measures, it is essential to comprehend the patterns and changes in these climatic parameters. For the purpose of identifying changes in precipitation patterns and assessing the effects on agricultural systems, scientists have looked at historical rainfall data. Similar to how temperature data are analysed, global warming is quantified and its consequences on ecosystems, human activities, and natural resources are identified. Karl's study looks at the global temperature data to determine whether there is a discernible rising trend. The results of this study show that the average global temperature has risen over the past century by 0.6 degrees Celsius. The observed warming trend offers important proof of the long-term modifications to the Earth's climate system. The study emphasises the need of looking at temperature data to understand the scope and effects of global warming [5], warming occurring in past few decades [6]. In the United States of America, there has been an increase in precipitation from September to December during the years 1941-1988 [7]. In order to determine the climatic variances and changes in patterns of various locations of Baluchistan, we constructed a model to analyse the climatic change of Pakistan in just the province of Baluchistan using CMIP-6 baseline data of 30 years (1987-2014). The province of Baluchistan makes up around 44% of the nation's total land area. With a population of 12.3 million, Baluchistan is the least populated province in Pakistan. Its population is diversified, with a number of ethnic groups including Baloch, Pashtuns, Brahuis, and Hazaras residing in distinct places. The province features a diverse geography, including plateaus, mountains, deserts, and coastal regions. The province's topography includes the Sulaiman range, Kirthar Mountains, and Makran Coastal Range. Baluchistan has an extremely hot and dry climate, with summertime highs of 50°C (122°F). The mountainous regions, however, are colder, and some of them get snowfall in the winter. Baluchistan's economy is mostly focused on agriculture, mining, and fishing. Gas, coal, copper, gold, and other natural resources are abundant in the province. It is still one of Pakistan's poorest and least developed regions, nonetheless. Due to climate change, the already existing problems of water shortages, food insecurity, and desertification are projected to get

worse. Though one of the most significant economic sectors in Baluchistan, agriculture is heavily dependent on rainfall. Water variability for agriculture is predicted to decrease due to climate change, which might result in crop failure and a scarcity of food. The length and timing of the growing season may also alter in the province, which might have an even greater effect on the agricultural industry. Another significant sector in Baluchistan is the fishing industry, although it too is expected to be damaged by climate change. Changes in ocean currents, sea level rise, and rising sea temperatures might all have an impact on fish populations and their habitats. This could result in dwindling fish populations and a drop in the number of people who depend on fishing for their livelihood. Baluchistan also has a significant livestock industry, although it is susceptible to the effects of climate change. The availability and quality of grazing grounds may change due to changes in temperature and rainfall patterns, which may have an effect on the welfare and productivity of cattle. Baluchistan's tourism industry is expanding, but it is also vulnerable to climate change, while sea level rise might endanger coastal infrastructure and tourism amenities, rising temperatures and changes in the weather could affect how appealing some locations are to travellers. The research study's methodology and approach are described in the methodology section. In this work, we estimate the climate using an ensemble technique. By using a range of model outputs, Buontempo, C., et al.(2015) [8] gives more reliable and trustworthy estimations of significant climatic variables like temperature and precipitation pattern. To comprehend and foresee the impacts of climate change, global climate models (GCMs) are crucial. These integrated numerical models comprise the sea surface, land surface, atmosphere, and sea ice. They provide considerable promise for researching climate change and its variations, according to Fowler, H. J., et al.(2015) [9] when making an ensemble projection, a number of different GCMs are merged to simulate a variety of plausible future climate scenarios. When making an ensemble projection, a number of different GCMs are merged to simulate a variety of plausible future climate scenarios. Researchers may examine the uncertainty brought on by model differences and the inherent natural variability of the climate system utilizing methodologies like ensemble projections that enhance our understanding of future climate dynamics [10]. Our

goal is to create uniformly climate zones throughout Baluchistan in order to examine temperature and precipitation variance and to forecast the future. The second goal is to ensemble all GCMs rather than utilising a single model to reduce model uncertainty. Baluchistan's economy is extremely susceptible to the effects of climate change. We applied some machine learning techniques as well as statistical methods for the investigation. The province will need to take action to adapt these changes, such as improving water management, diversifying the economy, and investigating in climate resilient infrastructure.

1.1 Objectives

The major objectives of this study are given below:

- To construct homogeneous climatic regions over the province of Baluchistan, Pakistan
- To develop ensemble climate projections over the developed homogeneous climatic regions

Chapter 2

Study Area and Data set

The Baluchistan region is famous for its varied topography, rich agricultural history, and numerous social issues. It is spread across several South Asian countries, including Pakistan, Iran, and Afghanistan. Baluchistan's climate is predominantly distinguished by dry and semi-arid conditions. It is a region with minimal precipitation because the majority of locations receive yearly rainfall that is typically less than 250mm. The problem of water scarcity is made worse by the low yearly rainfall, especially during the dry months when freshwater sources are limited. Baluchistan endures scorching summers with certain locations seeing temperatures rise beyond 45°C, which adds to the region's dry and hot environment. For Baluchistan's population, agriculture, and overall socioeconomic growth, the lack of water supplies poses serious problems. The region's socioeconomic challenges are further exacerbated by the absence of adequate water for agricultural activities, which has an impact on crop production and livelihoods. The issue of water scarcity and its effects on the prosperity of the region is complicated and necessitates extensive solutions. Table (2.1) gives brief informations of all Baluchistan stations. The Pakistan Meteorological Department provided baseline data for 28 years, from 1987 to 2014. The study used eight Global Climate Models (GCMs) encompassing the years 2017 to 2100 to comprehend the various climate situations in the future. These models provide information about possible temperature and precipitation changes in the area under several scenarios for the future. To evaluate how climate conditions might change depending on various socio-economic trajectories, two

specific scenarios, namely SSP 4.5 and SSP 8.5, were taken into consideration. Twelve meteorological stations in Baluchistan were the subject of the investigation. These stations offered useful weather related data, such as details on temperature and precipitation. The presence of missing values in the temperature and precipitation time series was one of the difficulties faced during the investigation. Statistical techniques were used to estimate the missing values in order to address this problem and guarantee the validity of the results. These statistical techniques assisted in bridging the dataset's gaps, allowing for a more thorough examination and evaluation of the Baluchistan climatic patterns. Figure (2.1) indicate our study area which include all meterological stations of Baluchistan, Pakistan.

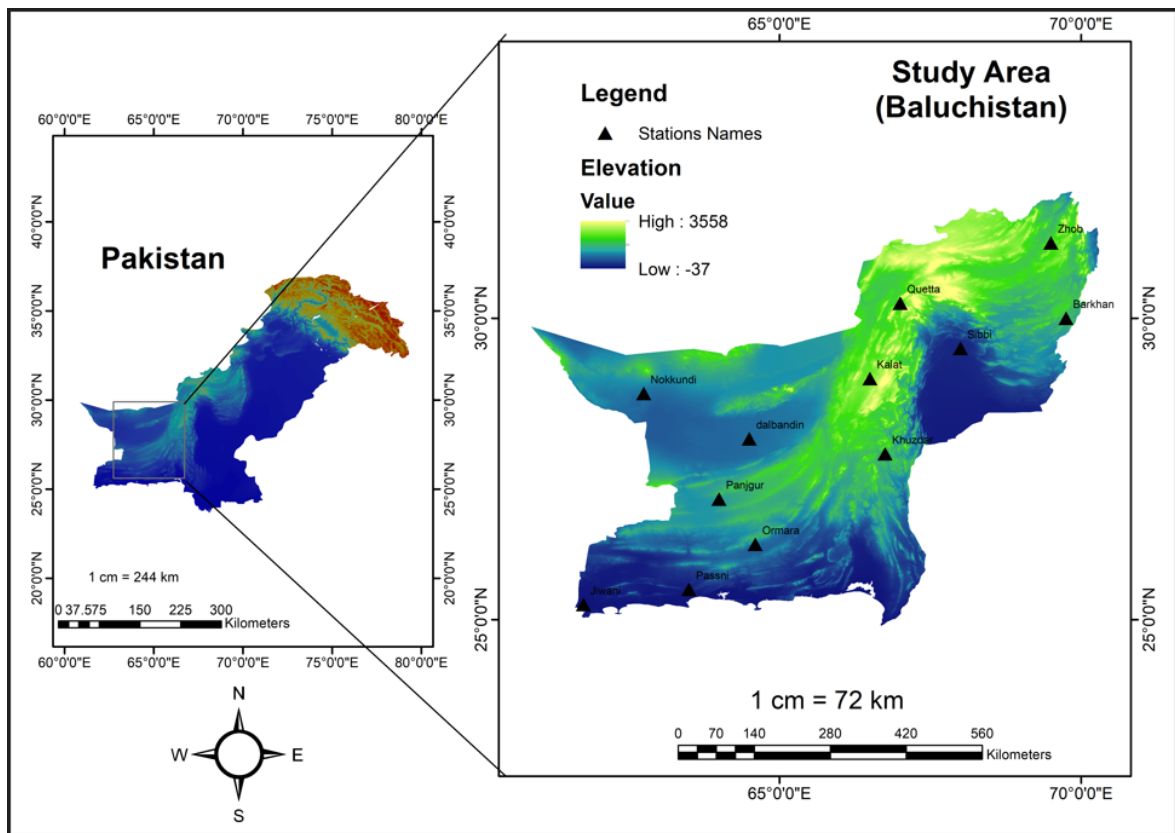


Figure 2.1: Study area of province Baluchistan, Pakistan

Stations	Longitude	Latitude	Elevation	Min. Temp	Max. Temp	MAP
Nokkundi	62.75	28.75	68.2	17.543	32.932	2799.179
Dalbandin	64.5	28	848	14.299	32.4721	2686.68
Panjgur	64	27	980	15.259	30.4482	4232.55
Sibbi	68	29.5	133	19.565	35.1040	5434.3
Passni	63.5	25.5	4	20.225	31.4893	2712
Ormara	68.25	26.25	37	18.158	35.8902	7663.97
Quetta	67	30.25	1600	8.8521	25.2631	7649.24
Zhob	69.5	31.25	1405	12.051	26.905	8003.16
Kalat	66.5	29	2015	5.5507	22.2405	6640.7
Khuzdar	66.75	27.75	1231	14.949	29.1668	8152.2
Jiwani	61.75	25.25	56	21.089	30.3304	5511.01
Barkhan	69.75	30	1097	14.783	28.4022	8813.343

Table 2.1: A brief information including latitude, longitude, elevation, maximum and minimum temperature, and mean annual precipitation at each meteorological station of Baluchistan, Pakistan

Chapter 3

Methodology

12 meteorological stations in Baluchistan, Pakistan, were regionalized as part of our study using the Cluster Analysis technique. We were able to group the stations using this method depending on how comparable their meteorological characteristics were based on the meteorological information gathered from the 12 stations and using cluster analysis, we were able to pinpoint certain regions inside Baluchistan. Understanding the geographic patterns and climatic differences in the region was made easier by regionalization. We used L-moments, a statistical technique frequently employed for analysing hydrological and climatological data, to evaluate the homogeneity and discordancy among the indicated locations. L-moments can identify abnormalities or inconsistencies and offer insights into the distributional properties of the data. We were able to evaluate the consistency of the meteorological data within each region by using L-moments. This phase was essential to ensuring that the information utilised to make future estimates and analyses was accurate and devoid of biases or other notable anomalies. In order to incorporate the results of various Global Climate Models (GCMs) for future climate projections, we also used ensemble projection techniques. This method produces more reliable estimates and aids in capturing the uncertainty linked to specific models. To build an ensemble for our projections, we chose eight GCMs with a range of features. We were able to produce a more thorough and trustworthy depiction of potential Baluchistan climatic scenarios thanks to the ensemble approach. We took into account the inherent unpredictability and uncertainty in cli-

mate model simulations by integrating the outputs of many GCMs. This method offers a more accurate spectrum of anticipated future climate conditions while decreasing the biases and limits of individual models. We were also able to produce probabilistic projections of future climate variables for each region of Baluchistan using the ensemble projection technique. The robustness of our projections was increased by these estimations, which took into account the uncertainties related to both the GCMs and the clustering method. Figure (3.1) shows our methodology section.

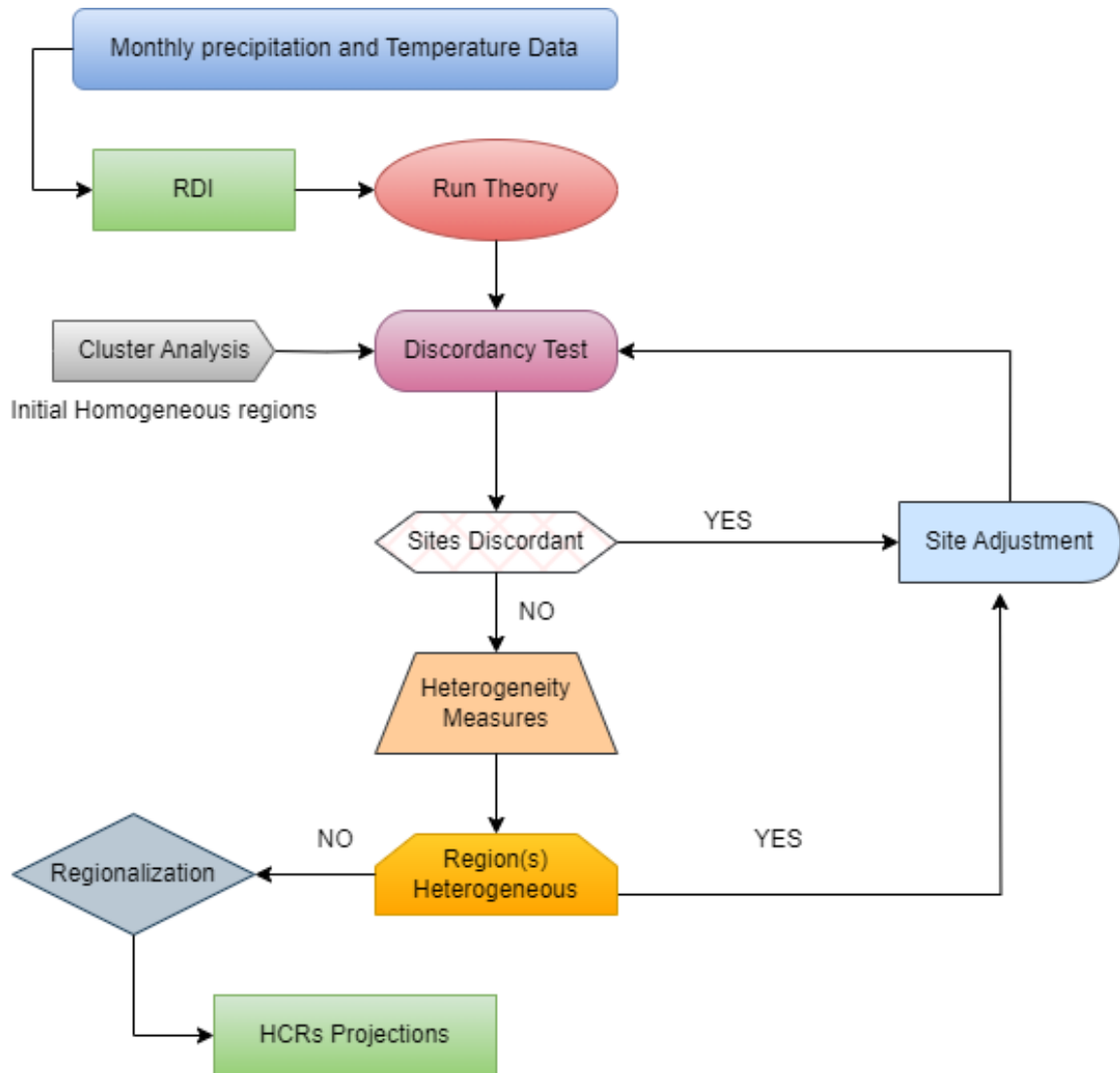


Figure 3.1: Methodology Flowchart

3.1 Droughts and RDI

Tsakiris (2004) [11] explains the concept of Reconnaissance Drought Index (RDI), in his work to analyse and track meteorological drought conditions in the Mediterranean region. A severe water deficit for crops, plants, animals, and human groups results from protracted periods of extremely dry weather or climatic conditions known as droughts. A prolonged absence of precipitation is what defines a drought, and this absence can cause major water deficits in rivers, lakes, reservoirs, and groundwater aquifers. Numerous types of droughts can be distinguished, including metrological, hydrological, agricultural, and socioeconomic droughts.

- Metrological droughts happen when there is a protracted period of precipitation that is below average.
- Agricultural droughts happen when the soil's moisture content is insufficient to sustain the establishment and growth of healthy crops.
- When the water supply in streams, lakes, and reservoirs falls below a crucial threshold, hydrological droughts take place.

When the demand for water outpaces the supply, a socioeconomic drought results, having a detrimental effect on the economy, society, and the environment [12]. Crop failure, water shortages, and wildfires are just a few of the serious effects that droughts can have on agriculture, water resources, human health and safety.

3.1.1 Calculation of RDI

The Reconnaissance Drought Index (RDI), introduced by Hayes et al., (1996) [13], is a commonly employed metric to evaluate drought severity relative to normal weather patterns. It measures the disparity between precipitation and potential evapotranspiration during a specified timeframe, often presented as a percentage of the long-term average precipitation for that period. The RDI presents a mathematical framework for

assessing the intensity of droughts and can be mathematically represented as follows:

$$RDI = \left[\frac{P - PET}{P_{mean}} \right] \times 100 \quad (3.1)$$

In the RDI equation, the Reconnaissance Drought Index (RDI) serves as the indicator for assessing drought severity. P represents the precipitation observed over the designated period. PET corresponds to the potential evapotranspiration occurring during the same timeframe. P_{mean} denotes the long-term average precipitation specifically for that period. To calculate the RDI, the precipitation (P) is subtracted by the potential evapotranspiration (PET), which yields the moisture deficit. This deficit is then divided by the long-term average precipitation P_{mean} and multiplied by 100 to express it as a percentage. The resulting RDI value serves as an indicator of the departure from normal conditions. Positive RDI values indicate wetter conditions compared to the average, while negative values signify drier conditions. Higher positive RDI values suggest relatively wetter conditions, whereas lower negative values indicate more severe drought conditions. The threshold levels for RDI vary depending on the climate and vegetation of the area, although multiple studies have suggested several threshold levels for RDI to signify various degrees of drought severity. According to John A. Kaeyantash (2004) [14], the RDI criterion for California should be between -1.0 and -1.5 for severe droughts, -1.5 to -2.0 for moderate droughts, -2.0 to -2.5 for severe droughts, and below -2.5 for extreme droughts. Similarly, J.D.van Rooyen (2020) [15] suggest the following RDI threshold for South Africa: -0.5 to 1.0 for mild drought, -1.0 to -1.5 for moderate drought, -1.5 to 2.0 for severe drought and below -2.0 for extreme droughts. Overall, threshold value of -0.85 for RDI is one of several possible threshold values used to indicate drought conditions, and Table (3.1) shows it is important to consider multiple factors and indicators when assessing drought severity and impacts.

3.2 L-moment

The L-moments approach involves calculating the L-moments of distribution from ordered data for a drought index, such as the Reconnaissance Drought Indices (RDI).

Categories	RDI
Extremely wet	≥ 2.0
Very wet	1.59 to 1.99
Moderately wet	1.00 to 1.49
Near normal	0 to -0.99
Moderately dry	-1.00 to -1.49
Severly dry	-1.50 to -1.49
Extremely dry	≤ 2.0

Table 3.1: Droughts classification thresholds

The parameters of several probability distributions that are frequently used to describe droughts can be estimated using these L-moments. Droughts have been modelled and examined using the L-moments approach in a number of studies, including those conducted in the Western United States, Australia, and China. This method has the benefit of being able to produce more accurate estimates of the distribution parameters, which is especially useful for data sets that may contain outliers or are not normally distributed. The analysis and estimate of distributions using a linear combination of order statistics were first introduced by [16]. We may determine the PWM using four mathematical equations.

$$\beta_0 = \frac{1}{N} \sum_{i=1}^N z(i) \quad (3.2)$$

$$\beta_1 = \sum_{i=1}^{n-1} \left[\frac{n-i}{n(n-1)} \right] z(i) \quad (3.3)$$

$$\beta_2 = \sum_{i=1}^{n-2} \left[\frac{(n-i)(n-i-1)}{n(n-1)(n-2)} \right] z(i) \quad (3.4)$$

$$\beta_3 = \sum_{i=1}^{n-3} \left[\frac{(n-i)(n-i-1)(n-i-2)}{n(n-1)(n-2)(n-3)} \right] z(i) \quad (3.5)$$

Droughts Event are ranked according to their severity in ascending order for analysis of drought episodes. A rank is given to each event on a scale of 1 to n, where n is the total number of events in a series. The ith event in the ranking series, which has a range from i to n, is represented by the variable z(i). Given below are the 1st four L-moments for population.

$$\lambda_1 = \beta_0 \quad (3.6)$$

$$\lambda_2 = 2\beta_1 - \beta_0 \quad (3.7)$$

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \quad (3.8)$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \quad (3.9)$$

suggested equations for L-moments by Hosking (1990) [16] L-coefficient of variation (L-CV)

$$L - coefficient\ of\ variation(L - CV) : \tau = \frac{\lambda_2}{\lambda_1} \quad (3.10)$$

$$L - skewness(L - Skew) : \tau_3 = \frac{\lambda_3}{\lambda_2} \quad (3.11)$$

$$L - kurtosis(L - kurt) : \tau_4 = \frac{\lambda_4}{\lambda_2} \quad (3.12)$$

3.3 Cluster Analysis

A statistical technique called cluster analysis seeks to arrange objects or observations into relevant groups or clusters according to their traits and features. Cluster analysis is defined as "assignments or a set of observations into subset (called a cluster) so that the observations in the same cluster are similar in some sense", according to Anil K. Jain (1988) [17]. To maximize similarity inside a cluster and reduce similarity across clusters is the goal of cluster analysis. The division of data into discrete groups that are maximally internally homogeneous and maximally outwardly heterogeneous. There are several clustering techniques, each with their own benefits [18]. A dendrogram hierarchy of clusters is produced, for instance, using the hierarchical cluster approach,

which begins with each observation as a distinct cluster and subsequently merges them depending on their similarity [19]. However, the k-means technique by Macqueen (1967) [20] divides the data into a predetermined number of non-overlapping clusters and assigns each observation to the cluster's nearest neighbour. Cluster analysis was used to identify an area of Pakistan with uniform temperatures [12].

3.4 Data screening and homogeneity test

In order to guarantee that climatic data are accurate, dependable, and acceptable for analysis and interpretation, it is crucial to filter data and test for homogeneity. We employ data screening to create climatically homogenous zones, and then we discordant the locations using statistical analysis [12]. The generalised extreme value distribution (GEV), the generalised Pareto distribution (GPD), and the generalised logistic distribution (GLD) are just a few of the probability distributions that is suggested by Hosking (1997) [21] utilising L-moments, L-CV, and L-Kurtosis to estimate their parameters. In addition, they suggested employing a goodness of fit test based on L-moments to assess how well data distributions fit together. In this investigation, N metrological stations are looked at, and a measure of discordancy is computed for each site.

$$D_m = \frac{1}{3}N(v_j - \bar{v})^T S^{-1}(v_j - \bar{v}) \quad (3.13)$$

$$S = \sum_{j=1}^N (v_j - \bar{v})^T (v_j - \bar{v}) \quad (3.14)$$

Hosking and Wallis [22] developed the heterogeneity measure (H) as a test statistic to evaluate the homogeneity of a group of sites that are intended to be integrated into a region. The H statistics employs the L-moments to calculate the standard deviation (V) of the at site sample L-CVs, which in turn determines how homogeneous the sites are. The weighting for V is determined by the historical drought levels at the site, and

the calculation is written as an equation.

$$V = \left\{ \frac{\sum_{i=1}^N N_i (t^i - t^R)^2}{\sum_{i=1}^N N_i} \right\}^{0.5} \quad (3.15)$$

N stands for the overall number of stations, N_i stands for the record duration at the i^{th} station, $t^{(i)}$ stands for the L-CV of stations i-th, and t^R is for the regional average of all station L-CVs within a certain region. Now, to assess the heterogeneity of a specific region, Monte Carlo simulations based on Kappa distributions with four parameters were performed. This created 1000 fictional regions using L-moments ratios if all areas had the same record length as the original data. The variance in L-statistics between the simulated and actual regions was determined when the simulated areas were constructed using the following formulae.

$$H_1 = \frac{(V - u_v)}{\sigma_v} \quad (3.16)$$

Where u_v stands for the median and σ_v for the standard deviation of the simulated and observed V parts, respectively. The evaluation of heterogeneity was given the following criteria by Hosking and Wallis [21].

If $H_1 < 1$, acceptable homogeneous

If $1 \leq H_1 < 2$, possibly heterogenous

If $H_1 \leq 2$, heterogeneous Compared to H_2 and H_3 , H_1 has the greatest ability to discriminate between heterogeneous and homogeneous areas [22]. The test uses L-CV, L-Skew, and L-Kurtosis-based H_1 , H_2 , and H_3 statistics. In this study, we just use H_1 to interpret the regions homogeneity.

3.5 Weighted Average

To describe a model's relative dependability or relevance in the context of model performance, wights might be utilized. These weights are applied to the outputs of multiple

models to produce an ensemble forecast or weighted average. When allocating weights, typically performance or accuracy of each model is taken into account. Models are given larger weights, indicating that their forecasts should have a greater impact on the ensemble forecast as a whole. On the designated assessment criteria, these models have demonstrated more accuracy or superior performance. The values supplied to the weights vary from zero to one.

$$\bar{X} = \frac{\sum_{i=1}^8 w_i N_i}{\sum_{i=1}^8 w_i} \quad (3.17)$$

The weighted average is indicated here by \bar{X} . Some weights may be 0 if models perform poorly.

3.6 Bias Correction

Quantile Delta Mappingsz (QDM) is a statistical technique for projecting changes in the frequency distribution of climatic variables as a result of changes in climate variability and/or averages. The approach, as created by Cannon and associates in 2015 [23], ensures that proportional change ratios are maintained within the quantiles of modeled variables.

Assume that x_0 , $x_{m,h}$, $x_{m,p}$ represent, observed data, historically modeled data, and future modeled data. The cumulative distribution functions (CDFs) of the observed historical data, the modeled historical data, and the modeled future data are each denoted by x_0 , $x_{m,h}$, $x_{m,p}$, respectively. The investigation starts by looking at the changing CDF of the projected series that has been modeled, $x_{m,f}$. Additionally, using the given equation, the variables x and ρ represent the data and their corresponding CDFs.

$$\rho_{m,f}(t) = F_{m,f}(t)(x_{m,f}(t)) \rho_{m,f}(t) \in [0, 1] \quad (3.18)$$

By comparing the ratios of two sets of inverse cumulative distribution functions (CDFs), you can determine the relative change. The model-predicted data from one set is

applied to its own CDFs to calculate it, while the observed data from the second set is applied to the model-predicted data CDFs. You can perform this computation using the given formulae.

$$\Delta_m(t) = \frac{F_{m,f}^{(t)-1}(\rho_{m,f}(t))}{F_{m,h}^{-1}(\rho_{m,f}(t))} = \frac{x_{m,f}(t)}{F_{m,h}^{-1}(\rho_{m,f}(t))} \quad (3.19)$$

$\Delta_m(t)$ denotes the relative change.

To adjust for bias correction, use inverse cumulative distribution functions (CDFs) obtained from observed historical data to modify the quantile of the model's future data $\rho_{m,f}(t)$.

$$\tilde{x}_{0,m}(t) = F_{0,h}^{-1}(\rho_{m,f}(t)) \quad (3.20)$$

You can calculate the updated estimates for future situations using the relative change derived from equation (3.19). The change is applied to the historically bias-corrected data from equation (3.20), as indicated in the following equation:

$$\tilde{x}_{m,f}(t) = \hat{x}_{0,m}(t) + \Delta_m(t) \quad (3.21)$$

The bias-corrected data from the future model, denoted as $\hat{x}_{m,f}(t)$, is useful for further study. To ensure the preservation of relative changes within the data, equations (3.19) and (3.20) are manipulated rather than simple addition, as indicated by Cannon et al. in 2015 [23].

3.7 K-folds cross validation

K-fold cross-validation is a method that is frequently used in statistics and machine learning to evaluate a prediction model's effectiveness and generalizability. By separating the existing data into several subsets, or "folds," and utilizing them for training and testing the model repeatedly, it is possible to anticipate how well a model will perform on unknown data. How k-fold cross-validation operates is as follows:

Data preparation involves first dividing the given dataset into k folds or subsets of equal size.

3.7.1 Model Training and Testing

The remaining fold is utilised for testing once the model has been tested on k-1 folds (along with k-1 subsets). This procedure is done k times, with the testing fold changing each time.

$$y = x\beta + \epsilon \tag{3.22}$$

3.7.2 Performance Evaluation

For each iteration, the performance measure (such as accuracy, precision, or recall) is computed, yielding k performance scores.

3.7.3 Performance Aggregation

The individual performance scores are then added together to provide a single performance estimate, often by figuring out the mean or median of the values.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \tag{3.23}$$

3.7.4 Model Choice

The final model is the one with the best performance estimate. The problem of overfitting, when a model performs well on the training data but fails to generalise to unseen data, is helped by K-fold cross-validation. Multiple fold testing offers a more accurate assessment of a model's performance and aids in the detection of possible issues like overfitting or underfitting. The values of k are frequently between 5 and 10, however they might differ depending on the size of the dataset and the available processing power. In reality, higher values of k yield more accurate performance predictions but also raise the cost of calculation.

It's important to note that the model creation and assessment step uses the k-fold cross-validation approach. After a final model has been chosen, it is usually trained on the whole dataset before being used to make predictions on fresh, unexplored data.

Chapter 4

Results

4.1 Construction of HCRs

The creation of HCRs is dependent on a number of multi-step processes. The first phase is creating HCRs for 12 metrological stations in Baluchistan, Pakistan, between 1987 and 2014 using data on precipitation and average temperature. To do this, latitude, longitude, sum of precipitation, mean average precipitation (MAP), and elevation are used to create subjective climatic regions based on the characteristics of the sites. Euclidean distance and Ward's linkage are then used to determine differences between the sites. The three constructed subjective zones are depicted in Figure (4.1). On their homogeneity metrics, all stations in the connected region remain the same to build adjusted regions.

In the second stage, each station's at-site attributes are used to calculate the RDI-12 series. This series is employed to evaluate L-moments, which aids in the explanation of the probability distribution's form. The Thornthwaite technique and the monthly average temperature are used to get the PET for each station. The RDI-12 series is computed using PET and monthly precipitation data by DRINC software.

In the third step, L-moments from each station are used to calculate L-moments ratios such L-skewness, L-kurtosis, and L-CV. The fourth stage is calculating the discordancy values using L-moment ratios, which aids in spotting the abnormalities in each metrological station. A table containing the analysis's findings demonstrates that

none of the stations are discordant. In addition, the homogeneity of built areas is evaluated using a homogeneous measure (H_1). The H_1 value for each region is shown in Table (4.1), demonstrating the homogeneity of every created region.

Stations	Discordancy Measures D_i		Heterogeneity Measures H_1	
	Observed Data	Baseline Data	Observed Data	Baseline Data
Region 1				
Quetta	0.80	1.24	-0.16	-0.93
Kalat	1.03	1.32		
Khuzdar	0.74	0.89		
Zhob	1.10	0.95		
Barkhan	1.25	0.59		
Region 2				
Ormara	1.00	1.00	-1.10	-0.69
Sibbi	1.00	1.00		
Region 3				
Nokkundi	0.89	1.30	1.39	-0.07
Panjgur	0.58	0.92		
Dalbandin	1.22	1.08		
Jiwani	1.25	0.91		
Passni	0.98	0.75		

Table 4.1: Discordancy values (D_i) and Heterogeneity measures (H_1)

The regions that can be discovered in both the observed data and the model baseline data are depicted in the cluster dendrogram below. It's important to note that no stations show appreciable differences between the observed data and the model baseline data, showing that the two are generally in accord. We were able to correctly identify three separate regions within the dataset by utilizing the variety of their attributes. The spatial distribution of areas within both the observed data and the model baseline data is revealed by the Cluster Dendrogram, a visual representation that aids in understanding the relationships between various data points. We can be sure that our model's baseline is reliable and that it can faithfully represent the real-world observations because there are no discordant stations. We obtain a deeper grasp of the underlying factors driving the data as we go with this thorough study of the regions and their variability. This information offers up new research directions, enabling us to

investigate potential correlations, causes, and implications that could have gone missed in the past. In conclusion, this Cluster Dendrogram has aided in the identification of significant patterns within the data, strengthened the validity of the baseline of our model, and offered insightful information for additional research. We can make more informed decisions and develop our field as we strengthen and broaden our analysis.

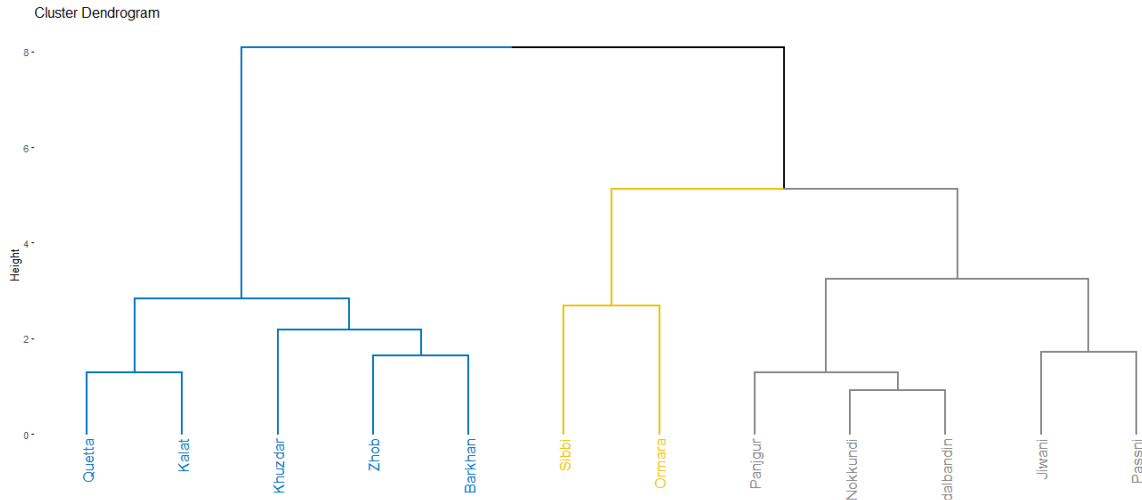


Figure 4.1: A generated dendrogram displaying the HCRs produced using observed data from the Baluchistan region of Pakistan from 1987-2014

4.1.1 Region 1

The region in question comprises five meteorological stations: Quetta, Kalat, Khuzdar, Zhob, and Barkhan, all of which are located in the Baluchistan province. Each station has its own distinct climate characteristics, offering a diverse range of weather patterns and conditions.

Quetta, Kalat, and Khuzdar are situated in the western and central parts of Baluchistan and experience a predominantly dry and arid climate. Summers in these areas can be scorching hot, with temperatures often reaching 40°C-45°C (104°F-113°F). The aridity of the region means that there is limited moisture in the air, resulting in low humidity levels. The lack of cloud cover and sparse vegetation contribute to the intense heat during summer months.

Winters in Quetta, Kalat, and Khuzdar are characterized by cold temperatures that frequently drop below freezing points. The clear skies and absence of cloud cover allow for rapid cooling during the night, leading to chilly conditions. Temperatures can dip to sub-zero levels, and snowfall is not uncommon in these areas during the winter season. The arid nature of the climate means that even during winter, precipitation is limited.

Moving northwards, we find Zhob, which is located in the northern part of Baluchistan. Zhob experiences a more moderate climate compared to Quetta, Kalat, and Khuzdar. Summers in Zhob are still warm, with temperatures ranging from 30°-35°C (86°F-95°F). The region benefits from its proximity to the mountains, which provides some relief from the scorching heat experienced in other parts of Baluchistan. Winters in Zhob are colder, with temperatures dropping below freezing point. However, the severity of the winter season is generally milder compared to the more southern regions.

Finally, Barkhan, located in the southern part of Punjab province, has a subtropical climate. The region experiences hot and humid summers, with temperatures reaching up to 40°C (104°F). The high humidity levels in Barkhan contribute to the discomfort felt during the summer season. Winters in this area are relatively mild, with temperatures ranging from 10-20°C (50-68°F). The proximity to Punjab and the southern region of the country brings a more moderate climate to Barkhan compared to the other Baluchistan stations.

In conclusion, the five meteorological stations in this region showcase a range of climate conditions. Quetta, Kalat, and Khuzdar exhibit a dry and arid climate with scorching hot summers and cold winters. Zhob experiences a more moderate climate with warm summers and colder winters. Barkhan, located in the southern part of Punjab province, has a subtropical climate with hot and humid summers and milder winters. These variations highlight the diverse climatic characteristics found within the Baluchistan province and its surrounding regions.

4.1.2 Region 2

Ormara and Sibbi, two cities in the Baluchistan province, are the subject of the third area. Both cities have a hot, arid desert climate that is characterised by high temperatures. This region's climate presents unique difficulties and has an impact on many facets of daily living. Summertime highs of up to 45°C are not uncommon in Ormara and Sibbi. The sweltering heat becomes a defining characteristic of the season, affecting the neighborhood's environment and the activities of locals. Due to the high temperatures, people must take steps to avoid heat-related illnesses and dehydration. It's important to drink plenty of water, look for cover, and minimise your time in the sun when it's at its hottest. In comparison to Sibbi, Ormara gets higher summertime humidity levels due to its seaside location. The unpleasant nature of the climate can be exacerbated by the combination of high temperatures and humidity. Increased air moisture can make it uncomfortable and necessitate taking extra precautions to be cool and comfortable.

In contrast, the region has cooler temperatures during the winter. Ormara's winter temperatures, which are often around 20°C, offer some respite from the oppressive summer heat. Sibbi enjoys similar weather conditions during the winter, with average temperatures in a similar range. Temperatures in Ormara's winter months can drop as low as 25°C during the day and 15°C at night due to its coastal location. Under comparison to inland places, the neighbouring sea's cooling impact keeps the temperatures under check. Residents might still require a light jacket or a jumper to stay warm in the chilly mornings and evenings. Seasonal temperature changes have an impact on many elements of living in both cities. These temperature swings may have an impact on agriculture and animal management, affecting crop selection and timing of activities. The intense heat and probable cooling needs during the summer months are also taken into account in the design and building of housing.

4.1.3 Region 3

There are five weather stations in the area covered by the paragraph: Dalbandin, Jiwani, Passni, Panjgur, and Nokkundi. These stations provide crucial information on temperature, wind patterns, and other meteorological phenomena, serving as key weather monitoring stations in the region. In the summer, the temperature in this area, especially in Nokkundi, can soar to blistering heights. The summer heat becomes fairly harsh when the temperature soars as high as 47°C. Residents must take care to safeguard themselves from high heat and illnesses associated to heat exposure because the hot weather can have a substantial influence on the neighborhood's environment and human activities.

On the other hand, the area gets significantly colder temperatures in the winter. Particularly in Dalbandin, temperatures may fall as low as 0°C. The cold weather presents its own unique set of difficulties, including the requirement for sufficient insulation, heating systems, and protective apparel to fend off the chill. Additionally, it may have an effect on the region's livestock management and agricultural practises. The predominant wind direction has a significant impact on the local climate and weather patterns in these cities. The majority of the region's cities have prevailing southwesterly winds during the summer. This indicates that the wind typically blows from the southwest towards the northeast and that this affects the temperature, humidity, and general weather conditions. The southwest breeze has the potential to bring dry, warm air, which raises the temperature in Nokkundi and other places.

For most cities, the predominant wind direction changes to the northwest throughout the winter, resulting in a changing weather pattern. Jiwani, where the wind comes from the northeast, stands out as an anomaly. Jiwani's differing wind direction from the other cities can result in various climate traits and perhaps even separate weather events. The five meteorological stations in this area collectively offer vital data on seasonal temperature extremes, wind patterns, and climate dynamics. This information aids in comprehending the local weather patterns, evaluating potential risks, and formulating plans to deal with the difficulties presented by the area's particular climate.

4.2 Evaluation of HCRs using observed data and model baseline data

To organize observed climate data into homogeneous zones, we used a systematic process. This methodology entailed categorizing data points based on a variety of environmental characteristics such as temperature, precipitation, humidity, and others. This approach was designed to capture diverse climatic trends within regions, allowing for a more granular knowledge of local climate changes. We used the same methodology that we used on observed data to apply to model baseline data. This consistency in methodology was necessary in order to allow for a meaningful and fair comparison of observed and modeled climate patterns. We sought to assure the robustness of our evaluation method by using identical criteria and algorithms for region delineation in both datasets. Physical representation of observed data and baseline data is given in Figure (4.2).

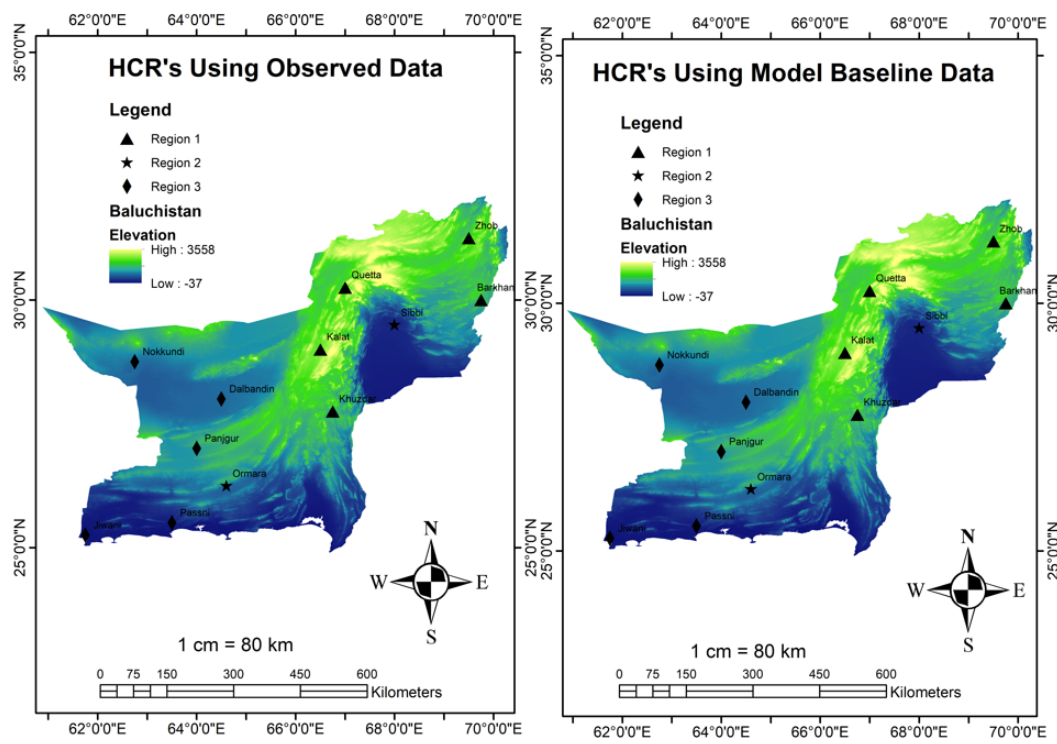


Figure 4.2: Construction of HCRs using observed data and model baseline data

4.3 Assessment of combine model's output and output of GCMs

Taylor diagrams are useful for visually showing and quantifying model or data set performance in comparison to a reference data set. It is frequently used in climate science and gives a comprehensive technique to measure several areas of the model's skill. It uses correlation coefficient, root mean square error (RMSE) and Standard deviation to compare results of ensemble models with other individual models.

4.3.1 Region1

Here is the graphical representation of the maximum temperature, minimum temperature and precipitation of Region 1 from 1987-2014 daily frequency data, where green dot shows the ensemble data and light purple dot represent the observed data. Our first judgment is how this graph is better, so we look our ensemble data which is much more closer to observed data than any other single model. In the maximum temperature Ensemble model follows the trend of observed data. The RMSE of each model (or ensemble) is represented by the distance from the Taylor diagram's origin. Greater agreement with the observed data is indicated by smaller distances. Thus, if your ensemble model has a lower RMSE and is more accurate than the individual models, it is closer to the origin with the correlation rang from 0.95-1 for maximum temperature in Figure (4.3), 0.90-1 for minimum temperature in Figure (4.4) and 0.85-1 for precipitation in Figure (4.5). Using RMSE, correlation coefficient, and Standard deviation (SD) ensemble model give high accuracy and batter performance.

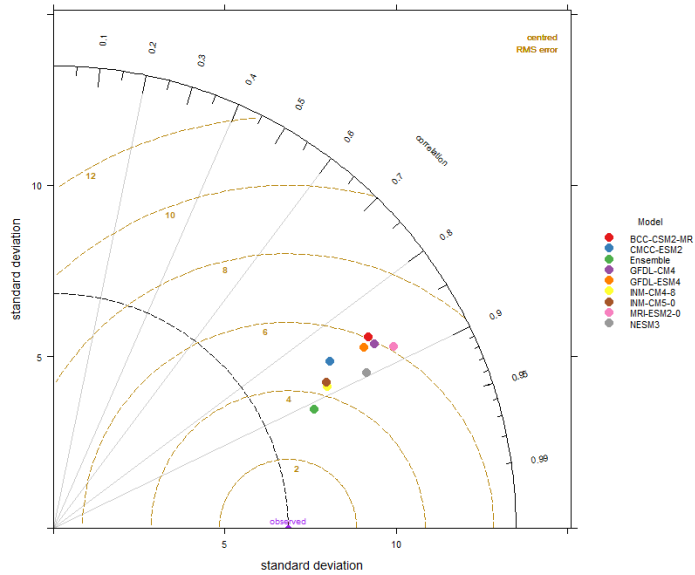


Figure 4.3: Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of maximum temperature of Region 1

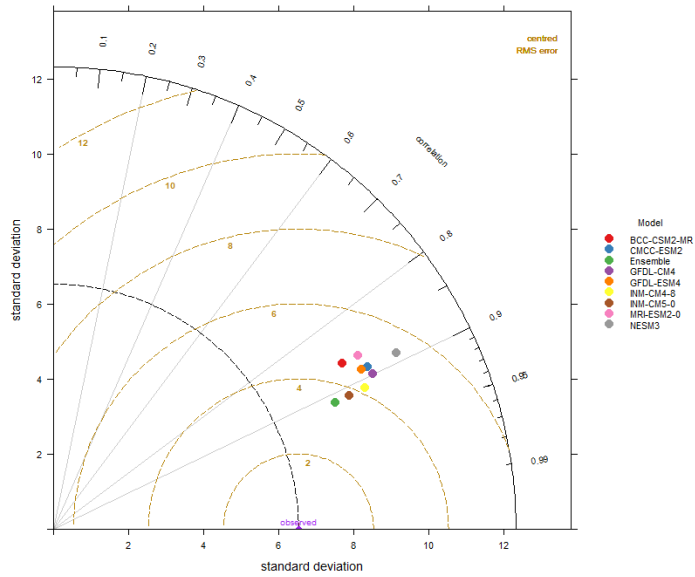


Figure 4.4: Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of minimum temperature of Region 1

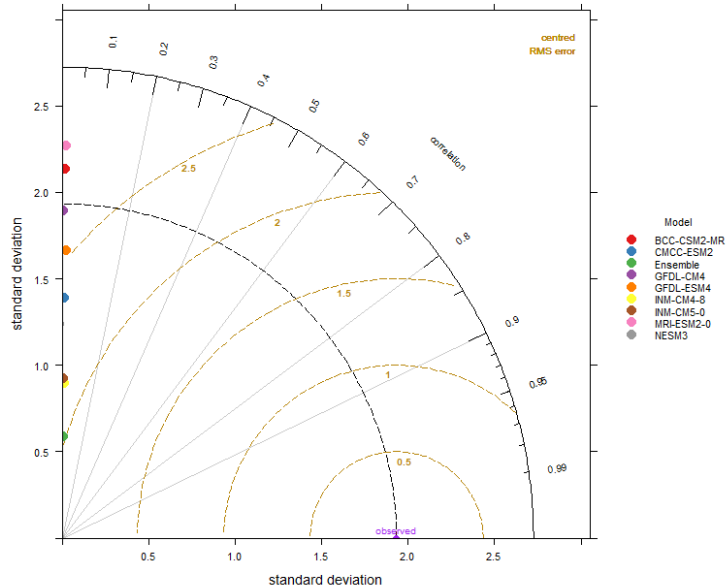


Figure 4.5: Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of precipitation of Region 1

4.3.2 Region 2

Region 2 shows the maximum temperature, minimum temperature and precipitation from 1987-2014 daily frequency data, where green dot shows the ensemble data and light purple dot represent the observed data. Our first judgement is how this graph is better, so we look our ensemble data which is much more closer to observed data than any other single model. In the maximum temperature Ensemble model follows the trend of observed data. The RMSE of each model (or ensemble) is represented by the distance from the Taylor diagram's origin. Greater agreement with the observed data is indicated by smaller distances. Thus, if your ensemble model has a lower RMSE and is more accurate than the individual models, it is closer to the origin with the correlation rang from 0.95-1 for maximum temperature in Figure (4.6), 0.97-1 for minimum temperature in Figure (4.7) and 0.95-1 for precipitation in Figure (4.8) (SD) ensemble model give high accuracy and batter performance.

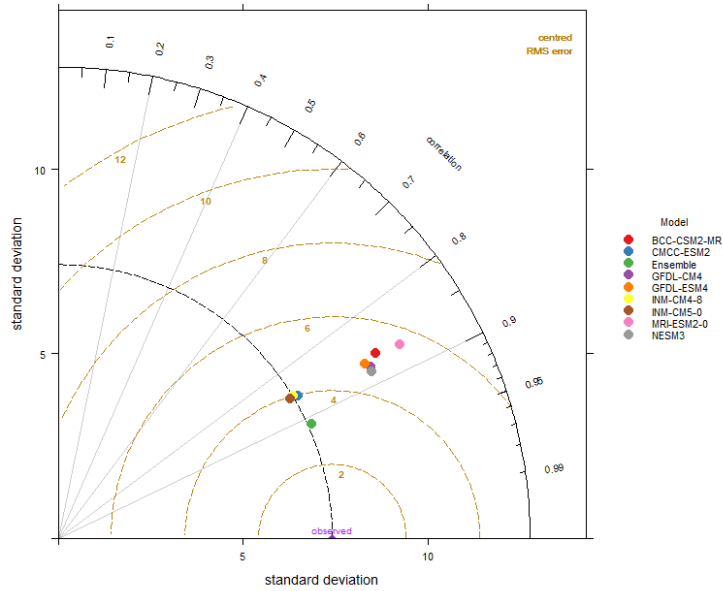


Figure 4.6: Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of maximum temperature of Region 2

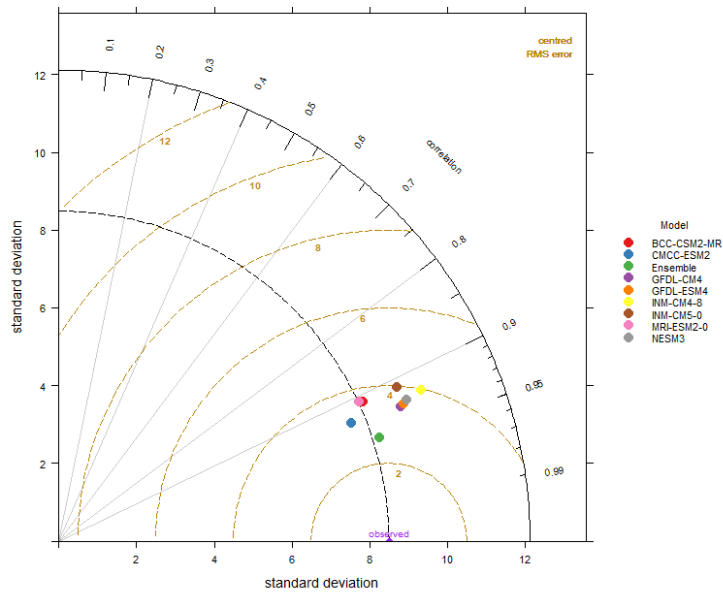


Figure 4.7: Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of minimum temperature of Region 2

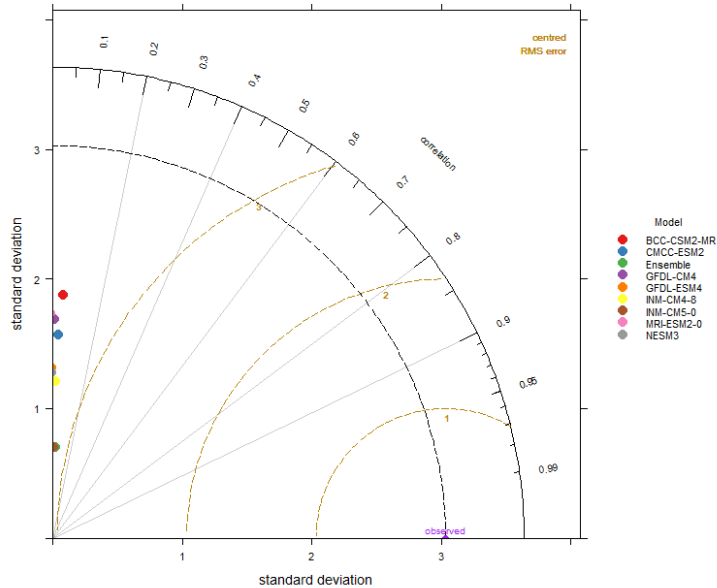


Figure 4.8: Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of precipitation of Region 2

4.3.3 Region 3

Two sets of data points may be seen in these graphs, which shows daily maximum temperature, minimum temperature, and precipitation data for Region 3 from 1987 to 2014. The green dots indicate the ensemble data, while the light purple dots reflect the observed data. According to our preliminary analysis, the ensemble data is far more similar to the observed data than any particular model. In the case of maximum temperature, the ensemble model successfully reproduces the trend of the observed data. We utilize three metrics Root Mean Square Error (RMSE), correlation coefficient, and Standard Deviation (SD) to assess the correctness of the ensemble model and contrast it with the individual models. These measures enable us to evaluate how well the model predicts the observed data. The RMSE of each model or ensemble is represented by the distance from the origin in the Taylor diagram. Greater agreement with the observed data is indicated by a smaller distance, which suggests greater accuracy. The range of the correlation coefficient is 0.95 to 1 in Figure (4.10) , 0.96

to 1 in Figure (4.9), and 0.48 to 1 in Figure (4.11) for maximum temperature, minimum temperature, and precipitation, respectively. By taking into account the RMSE, correlation coefficient, and standard deviation, we discover that the ensemble model outperforms in precipitation pattern is not performing well because it is not giving us a good prediction but overall we shows 9 diagrams and in each diagram ensemble model is on top as compared the individual models in terms of accuracy and performance. This is clear from the ensemble data points' proximity to the Taylor diagram's origin, which shows a good correlation with the actual data.

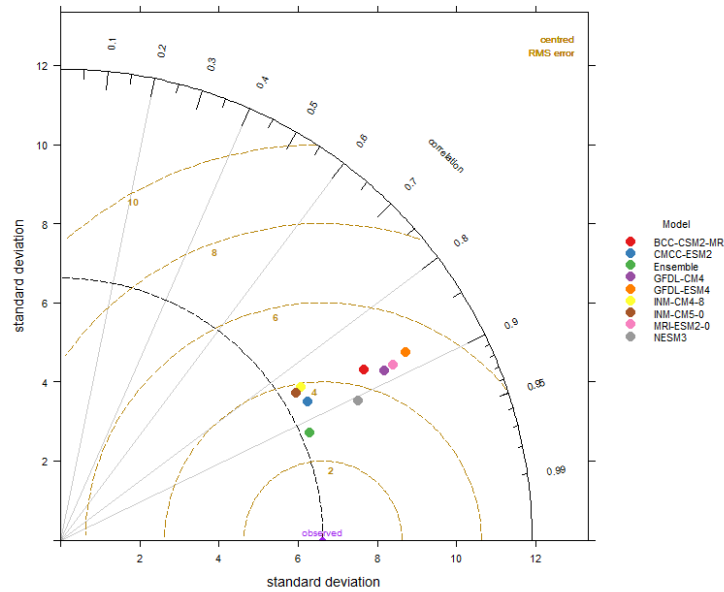


Figure 4.9: Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of maximum temperature of Region 3

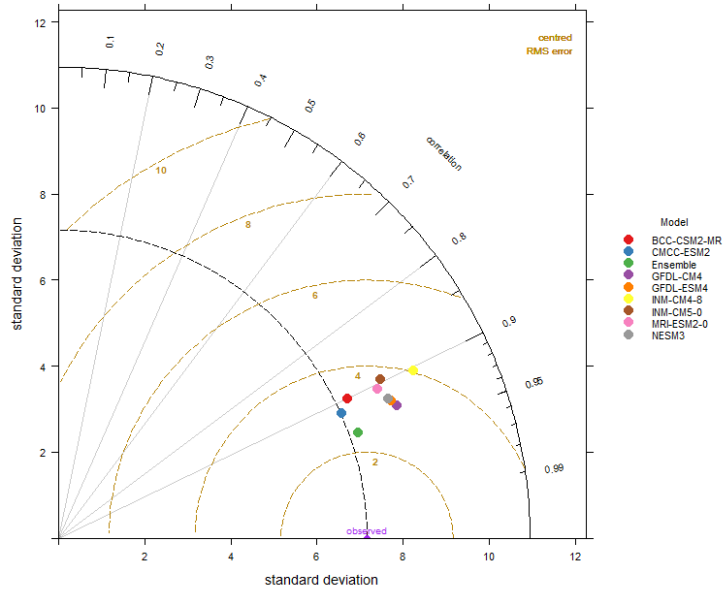


Figure 4.10: Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of minimum temperature of Region 3

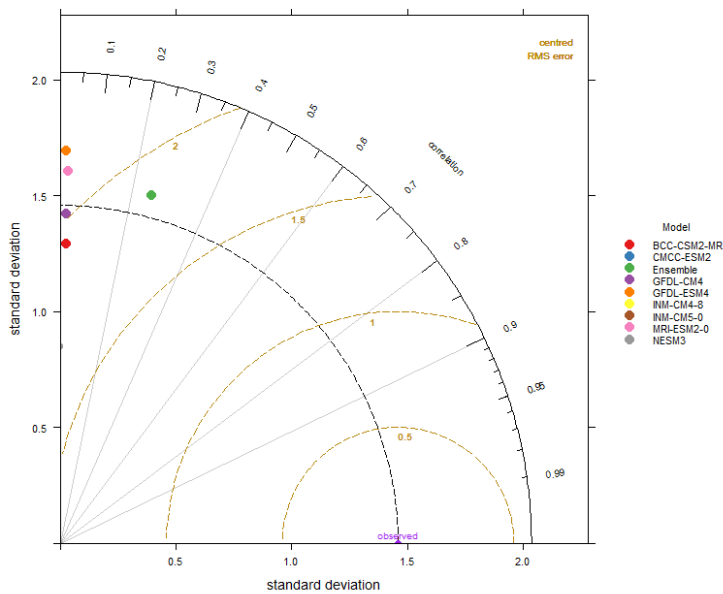


Figure 4.11: Taylor diagram shows the comparison between observed data, eight GCMs data and Ensemble model data from 1987-2014 of precipitation of Region 3

4.4 Evaluation of HCRs using Ensemble Models output

In the aforementioned picture, a comparison between the created HCRs (Homogeneous Climate Regions) and baseline model data is made. HCRs, which are geographical areas with comparable climatic conditions, are frequently utilised in climate studies and impact analysis. This comparison aims to assess if the HCRs created using these two different sources of data are similar or dissimilar. The comparison shows that the areas are built the same way in both scenarios, demonstrating agreement in terms of the broad-scale climate patterns reflected by the observed data and the baseline model data. This shows that the baseline model data can be utilised as a tool for future climate projections and impact analyses because it reasonably captures the observed climatic conditions.

However, the comparison also highlights differences in the heterogeneity and discordancy values between the observed and model data. Heterogeneity refers to the degree of variation or diversity within a region, while discordancy refers to the extent to which a particular weather station deviates from the overall pattern of the region. These values provide insights into the consistency and reliability of the data and their representation of the climatic conditions. The numbers for each station's heterogeneity, which represents the degree of variation within the region corresponding to that station, are presented in the table that goes with the figure. The discordancy values are also given, indicating how much each station deviates from the overall pattern of the area. These numbers can be used to locate stations that may have unusual or distinctive climatic traits when compared to the surrounding area. The heterogeneity and discordancy values for the areas and individual stations differ between the observed and model data. This implies that there can be differences between the observed climatic circumstances and how those conditions are portrayed by the baseline model. These discrepancies could be caused by a number of things, such as resolution constraints in the model, biases in the observed data, or parameterization uncertainties. Despite these variations, the homogeneity of the HCRs created from the baseline model data shows

that the model reasonably represents the large-scale climate patterns. Additionally, the lack of inconsistent stations in the baseline model data suggests that the climatic properties of all stations within the created regions are uniform. Based on these results, the model data is regarded as valuable for upcoming measurements of the effects of climate change on these HCRs. It offers a trustworthy foundation for comparison and can be applied to evaluate how climate change can affect the climatic conditions in these areas. When evaluating the results and developing projections based on the model data, it is crucial to take into consideration the restrictions and uncertainties of the observed data.

4.5 Impact assessment of climate change on HCRs using SSP 4.5 and SSP 8.5 Scenarios

Homogeneous Climate Regions (HCRs) were created using projected data from global climate models (GCMs) in order to analyse the potential effects of climate change. Three unique time periods were assigned to the data: 2017 to 2044, 2045 to 2072, and 2073 to 2100. An evaluation of the 21st century's climate trends is possible using these time frames. The climatic scenarios for the first time period, which ran from 2017 to 2044, were based on Shared Socioeconomic Pathways (SSPs). According to the analysis, every region had a consistent weather pattern, and no station displayed any major variations from the regional pattern. This suggests that during this time, the climate remained largely constant, and the HCRs created using the GCMs' future data were internally consistent.

The similar tendency persisted as the investigation expanded into the second time span, 2045 to 2072. No stations were moved to new locations, demonstrating that the weather patterns remained constant within the defined HCRs. This shows that, at least within the context of the HCRs, the climate projections made during this time did not cause any appreciable changes or shifts in the regional climatic patterns. Through the third time period, which runs from 2073-2100, the pattern of consistent weather patterns and absence of discordant stations maintained. No discordant stations were

found, suggesting that every station continued to display climate traits typical of its region. This implies that the climate projections for the latter half of the 21st century did not indicate any substantial disruptions or departures from the established climate patterns within the HCRs based on the future data from the GCMs.

These discoveries have consequences for understanding and anticipating the effects of climate change. The same weather patterns seen throughout the three time periods suggest some stability and dependability in the GCMs' estimates. Within the designated HCRs, these data can be used to evaluate the prospective impacts of climate change on a variety of industries, including agriculture, water resources, and infrastructure planning. Overall, the results indicate that the established HCRs maintained similar weather patterns and did not undergo notable shifts or discordances during the three time periods examined, according to the GCM's future projections. Understanding the possible effects of climate change and guiding adaptation plans in these places can both benefit from these findings shows in Figure (??)

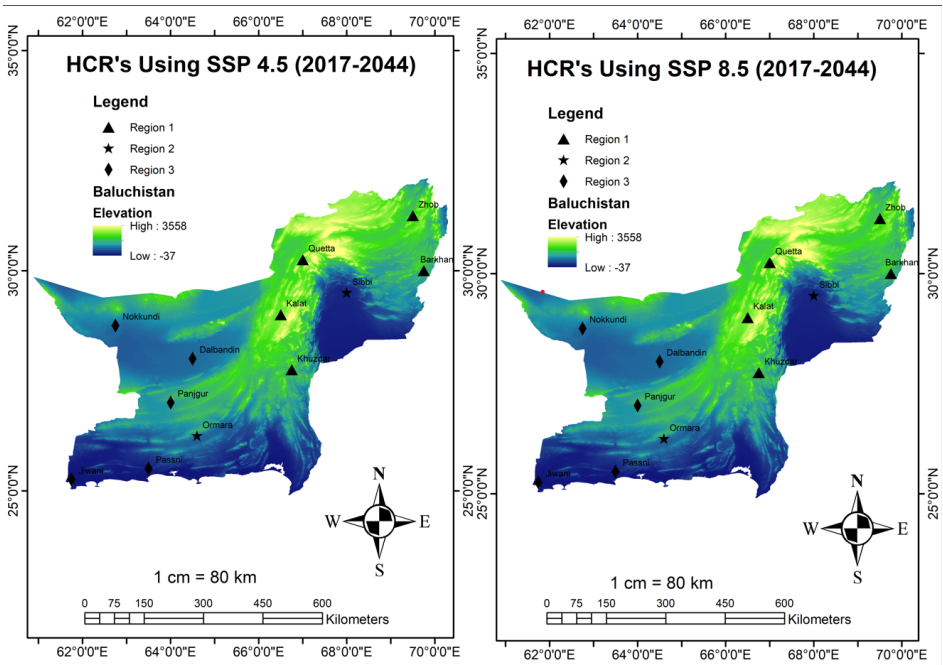


Figure 4.12: Construction of HCRs using SSP 4.5 and SSP 8.5 from 2017-2044

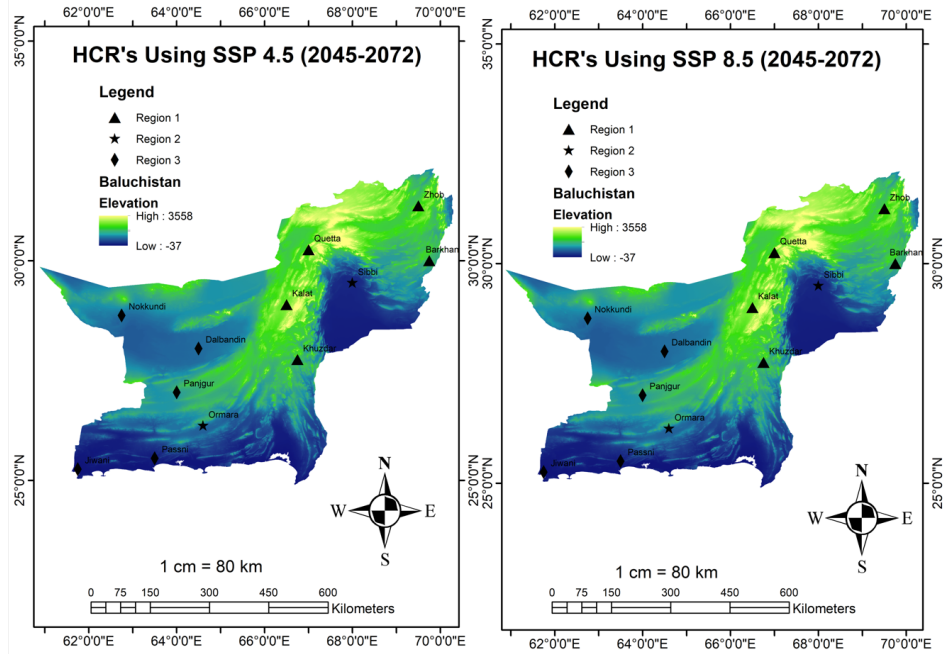


Figure 4.13: Construction of HCRs using SSP 4.5 and SSP 8.5 from 2045-2072

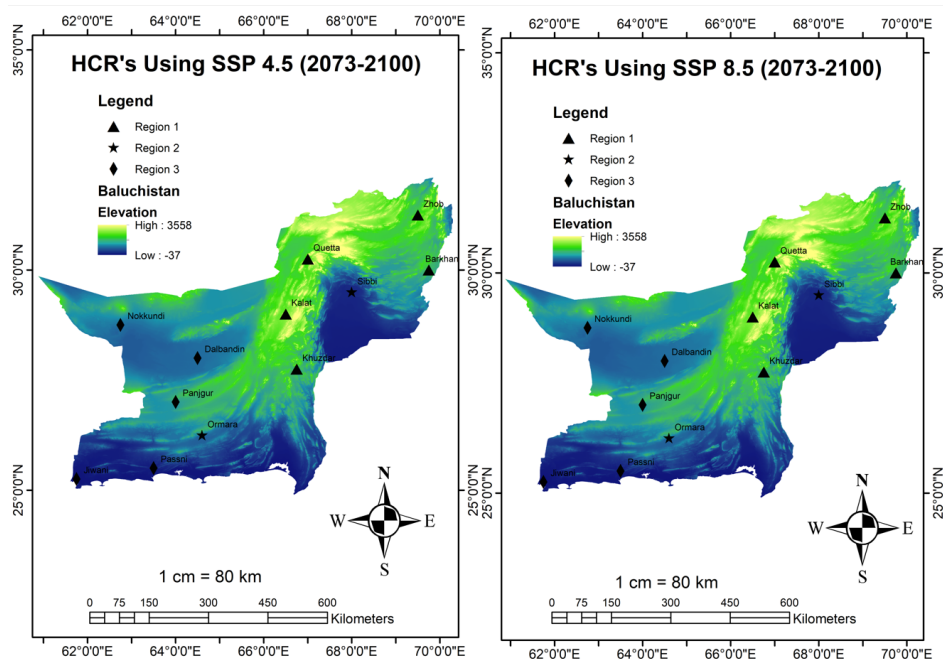


Figure 4.14: Construction of HCRs using SSP 4.5 and SSP 8.5 from 2073-2100

4.6 Evaluation of ensemble model output

Bias correction is a popular tactic used in climatic simulation to reduce the risks related to climate models. Bias correction aims to rectify any potential biases in the climatic data produced by global climate models (GCMs). By closer aligning the GCM data and observed data, this approach seeks to increase the simulation's accuracy and dependability. The bias correction approach is used in this analysis to adjust three variables: rainfall, maximum temperature, and minimum temperature. To guarantee that the generated data appropriately represents the climatic conditions unique to that region, the GCM climate data for each zone is individually treated to bias correction. The main goal of bias correction is to match the observed data with the GCM data. One can see that the median of the model baseline is very different from the actual data by comparing the observed data, model data, and bias-corrected data of maximum temperature using a boxplot. Additionally, compared to the observed data, the model baseline median is much lower. The bias correction method has successfully changed the data to lessen the discrepancy and bring it in line with the observed values, though, as the median of the bias-corrected data is now more closely aligned with the observed data.

The boxplot analysis shows that the median of the bias-corrected data and the observed data for the minimum temperature are relatively similar to one another, demonstrating that the bias correction technique was successful in aligning the data. The model's baseline data for minimum temperature, however, show a median that is somewhat lower than the actual data. The observed data is shown to be superior to the historical model data in terms of precipitation. This suggests that there are biases or differences between the model's baseline precipitation data and the measured values. The bias-corrected data is used to evaluate the effects of climate change in light of these findings. When compared to the model data, the bias-corrected data, which has undergone the appropriate corrections to lessen biases, is thought to deliver more accurate results. The fact that the bias correction method successfully increased the accuracy and reliability of the climate simulation is demonstrated by the resemblance

between the bias-corrected data and the observed data. GCMs used in our study is listed in Table (4.2).

BCC-CSM2-MR	Beijing Climate center, China metrological administration China
CMCC-ESM2	Centro Europe-Mediterraneosui Cambiamentia cliamtic Itlay
GFDL-CM4	Geophysical Fluid Dynamics Laboratory,USA
GFDL-ESM4	Geophysical Fluid Dynamics Laboratory,USA
INM-CM4-8	Institute for numerical mathematics
INM-CM5-0	Institute for numerical mathematics
MRI-ESM2-0	Metrological Research Institute, Japan
NESM3	Neinjing University of Information-Science and Technology, China

Table 4.2: Global Climate Models

Box plots shows the Model data, observed data and bias corrected data of each region and show how biased corrected data is nearly related to Observed data as compared to model data. The Figures (4.15, 4.16 and 4.17) show Region 1 maximum temperature, minimum temperature and precipitation, respectively.

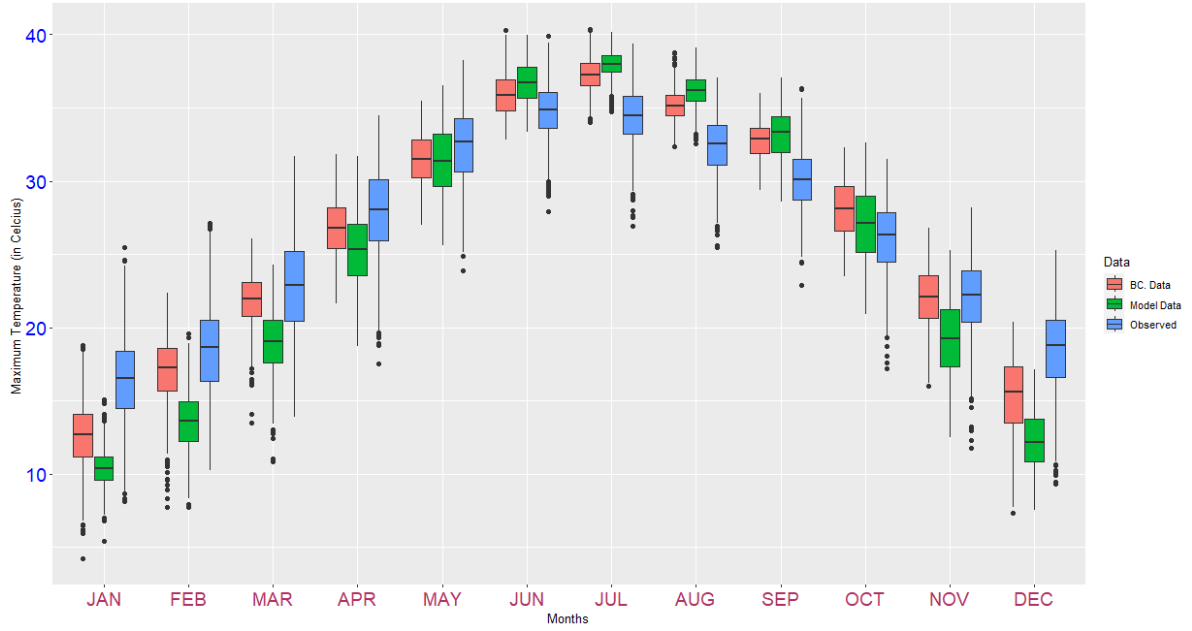


Figure 4.15: This graphs shows maximum temperature of zone 1 using model data, baseline data and observed data.

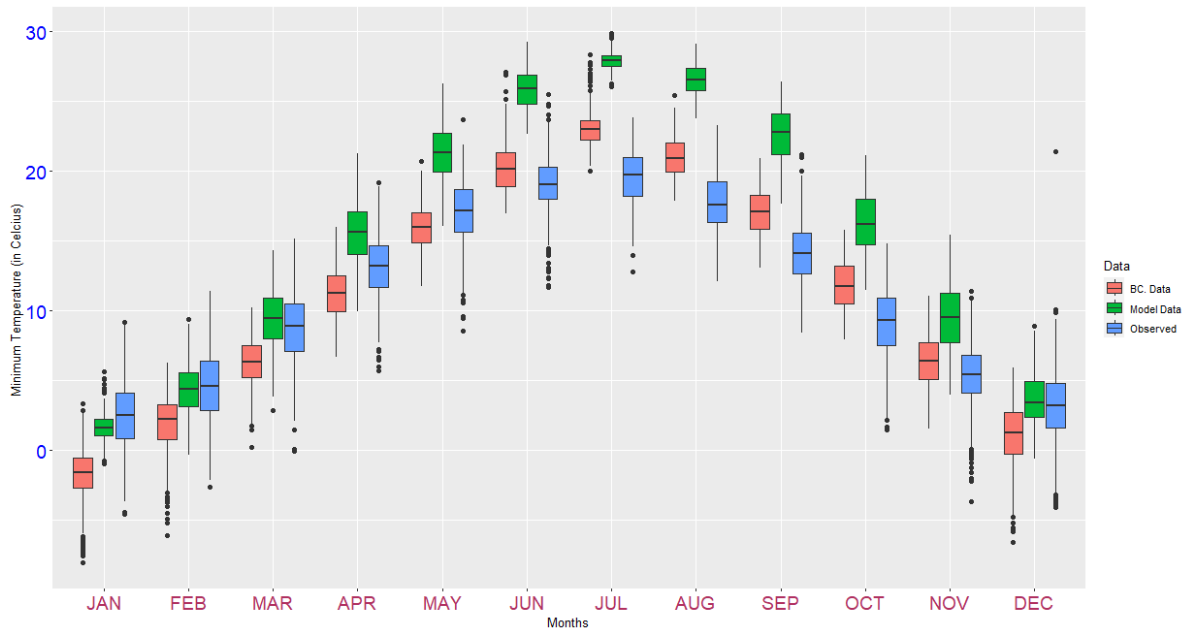


Figure 4.16: This graphs shows minimum temperature of zone 1 using model data, baseline data and observed data.

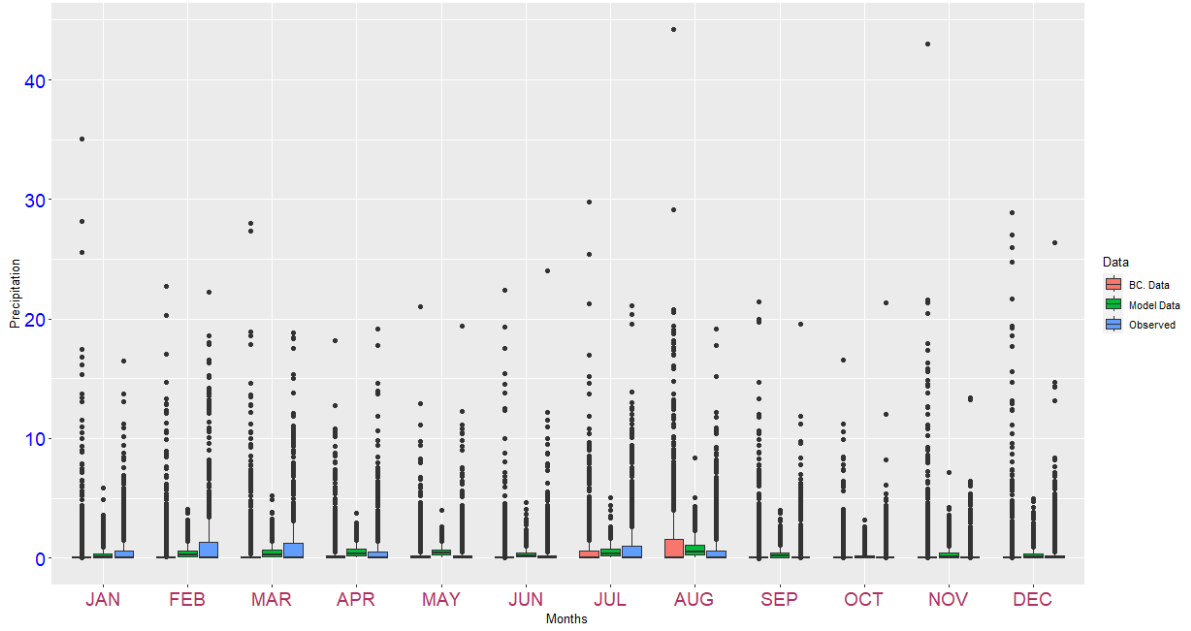


Figure 4.17: This graphs shows rainfall pattern of zone 1 using model data, baseline data and observed data.

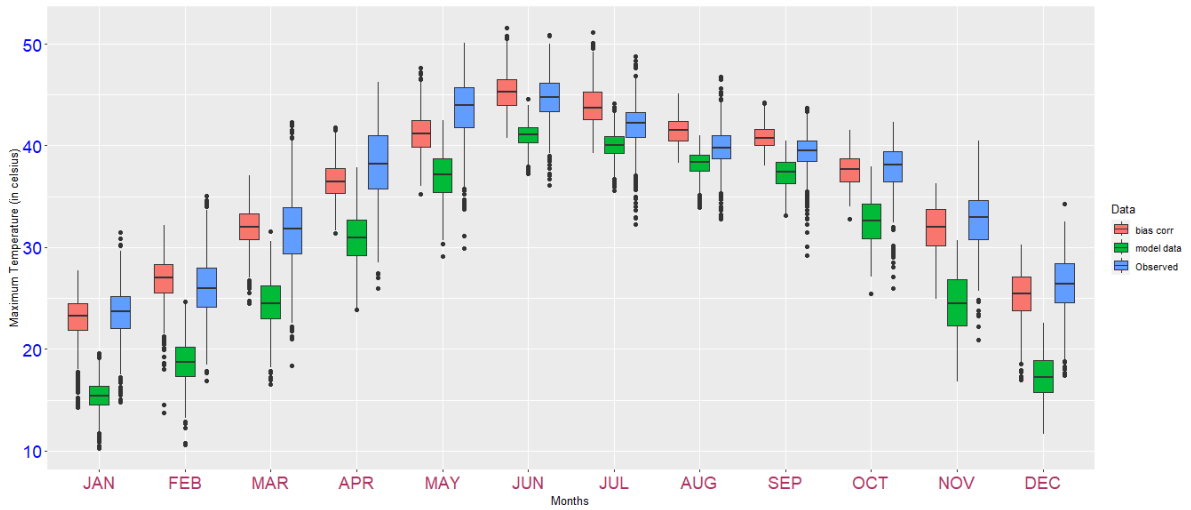


Figure 4.18: This graphs shows maximum temperature of zone 2 using model data, baseline data and observed data.

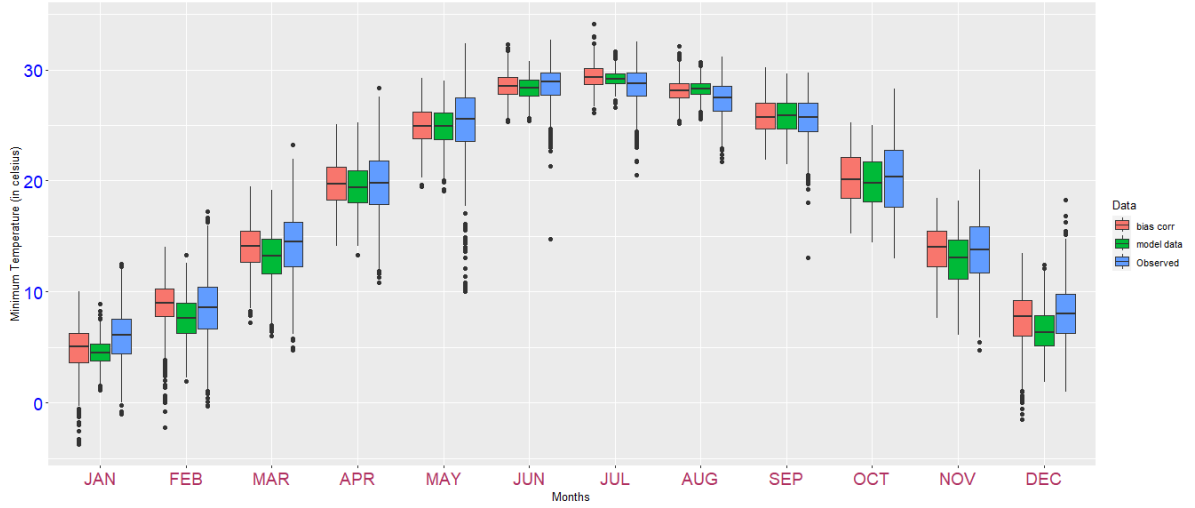


Figure 4.19: This graphs shows minimum temperature of zone 2 using model data, baseline data and observed data.

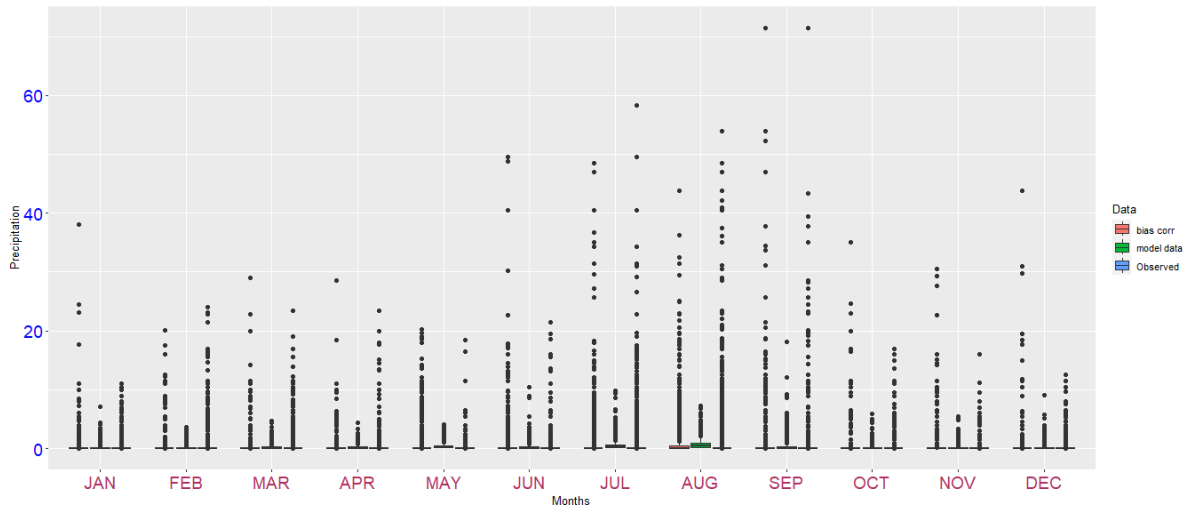


Figure 4.20: This graphs shows rainfall pattern of zone 2 using model data, baseline data and observed data.

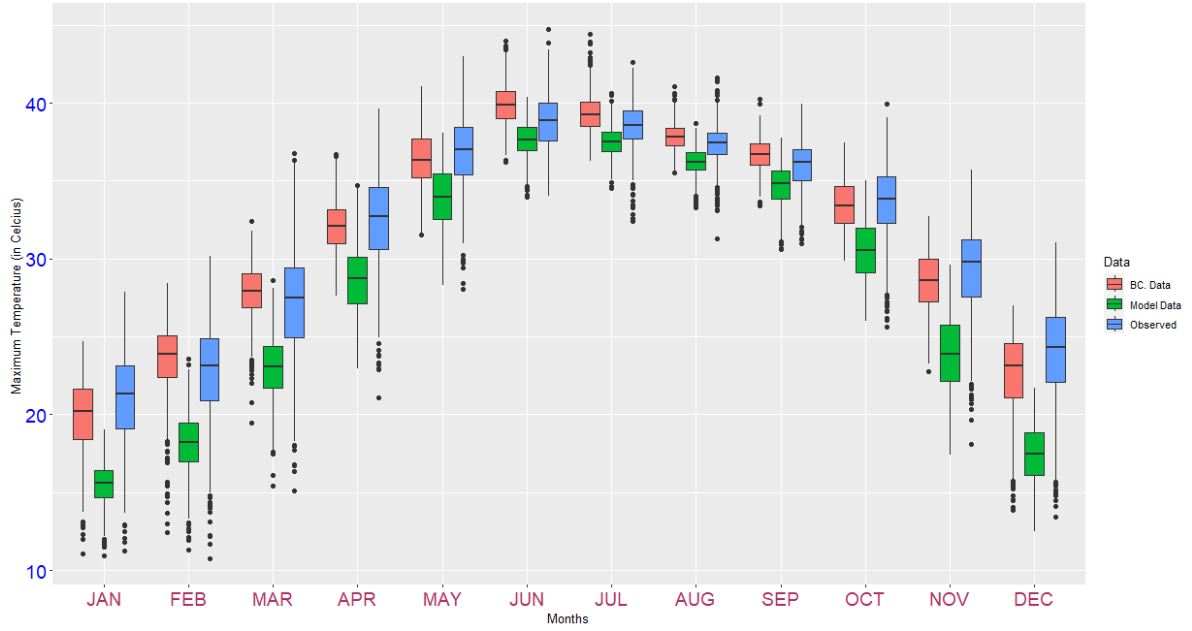


Figure 4.21: This graphs shows maximum temperature of zone 3 using model data, baseline data and observed data.

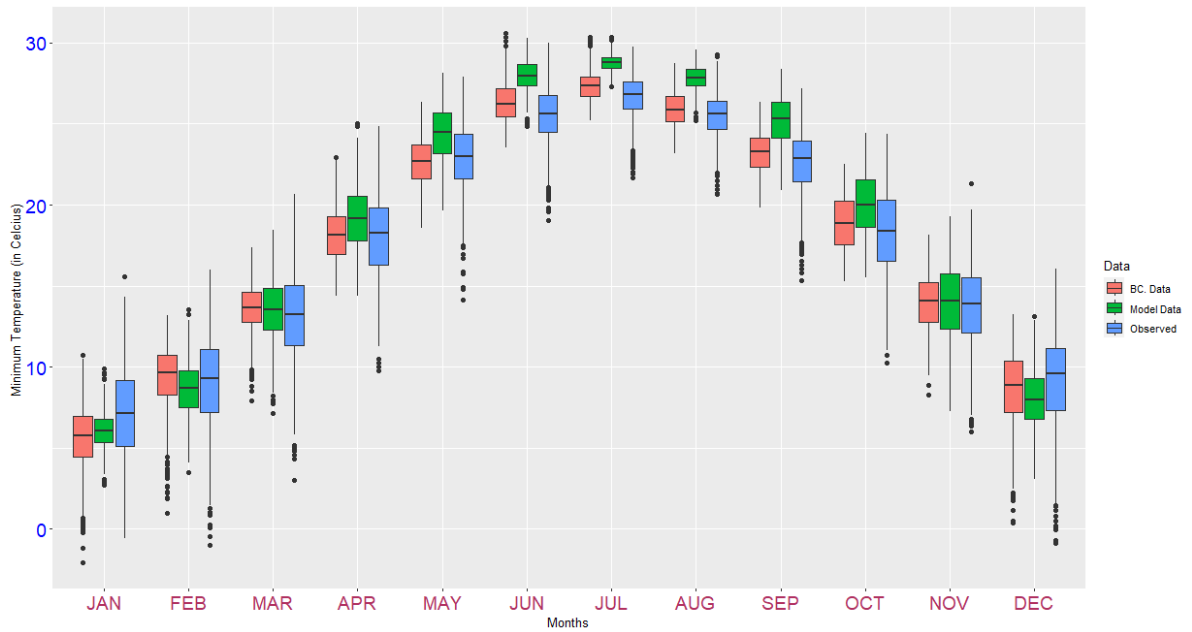


Figure 4.22: This graphs shows minimum temperature of zone 3 using model data, baseline data and observed data.

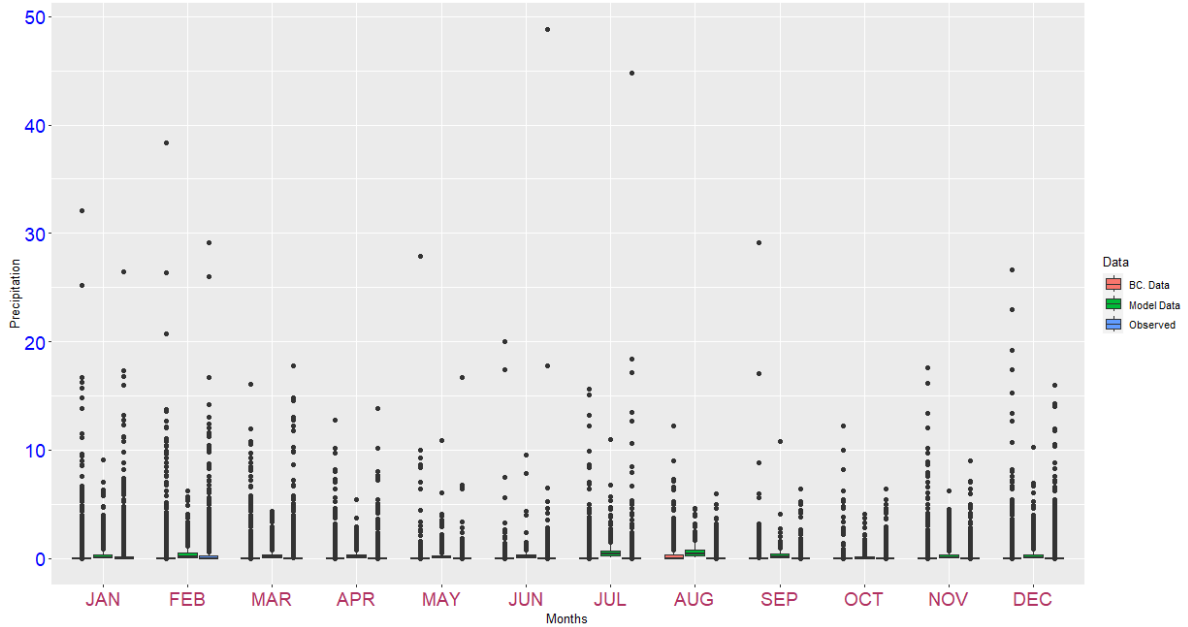


Figure 4.23: This graphs shows rainfall pattern of zone 3 using model data, baseline data and observed data.

4.7 Ensemble Climate Projections

The future climatic data for various regions have been projected by the eight GCMs (Global climate models) spanning three unique time periods: 2017-2044, 2045-2072, and 2073-2100. These forecasts take into account the SSP 4.5 and SSP 8.5 scenarios. The years 1987 through 2014 serve as the comparison's baseline. The monthly precipitation, minimum temperature, and maximum temperature of the climate have all been projected. Understanding and analysing the effects of climate change on various regions requires these variables.

The eight GCMs, which are intricate computer models that replicate the Earth's climate system, were used to produce the projections. These models take into consideration a number of variables, including land surface interactions, ocean circulation patterns, atmospheric conditions, and greenhouse gas emissions. In this case, the period from 1987 to 2014 serves as the baseline for comparison. For assessing how precipitation and temperature patterns have changed throughout time, this baseline period

is used as a point of comparison. Researchers and decision-makers can evaluate the potential effects of climate change on precipitation and temperature in various regions under various scenarios by comparing the forecasts from the GCMs to the baseline period.

4.7.1 Region 1

In region 1, analysis for the SSP 4.5 and SSP 8.5 scenarios compared baseline data with forecasts for rainfall, minimum temperature, and maximum temperature. The distribution of data over 28-year periods was represented using the boxplot approach. Let's look at the specific results for each variable shows in Figures (4.24, 4.25 and 4.26) Maximum Temperature: The box plot analysis showed a considerable change in the median values, or the centre points, of the predicted peak temperatures. The median line for the years 2017 to 2100 under the SSP 4.5 scenario closely resembles the median line seen in the baseline data. This shows that, under the moderate emissions scenario, the central tendency of maximum temperatures is still largely consistent with historical trends. The median value for the majority of the upcoming time periods, however, greatly exceeds the baseline data's median value under the SSP 8.5 scenario. In the years 2045 to 2072, with the exception of December, the median value of SSP 8.5 continuously exceeds the baseline median. This suggests that maximum temperatures in region 1 will significantly rise, especially in the scenario with large emissions.

Minimum Temperature: Both SSP scenarios have a similar pattern for the minimum temperature. When compared to the baseline data, the temperature decreased by a very small amount in September. This suggests that, aside from a slight cooling trend seen in September, the minimum temperature in region 1 will stay largely steady over the coming seasons. Rainfall: Depending on the scenario, the research shows considerable changes in region 1's rainfall patterns. In comparison to the baseline data, precipitation for the period from 2017 to 2044 is noticeably lower under the SSP 8.5 scenario. According to the high emissions scenario, this points to a tendency towards drier conditions earlier in the century. Contrarily, compared to the baseline data, precipitation increases during the period from 2073 to 2100 under the SSP 4.5

scenario. This suggests that, under the moderate emissions scenario, the century's later years will get wetter.

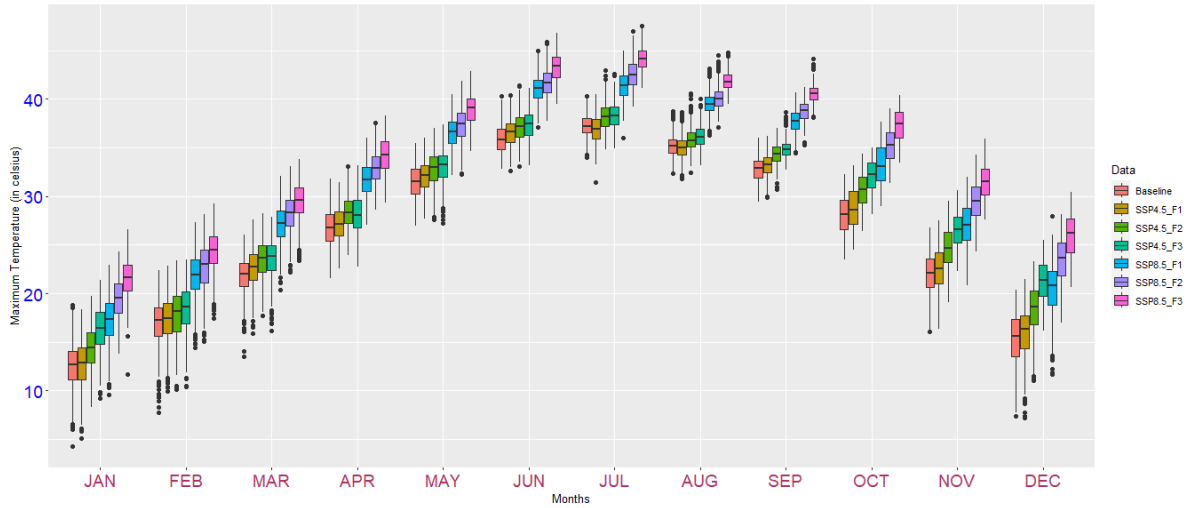


Figure 4.24: Future Maximum Temperature trend since (2017-2100) of region 1 using SSP 4.5 and SSP 8.5

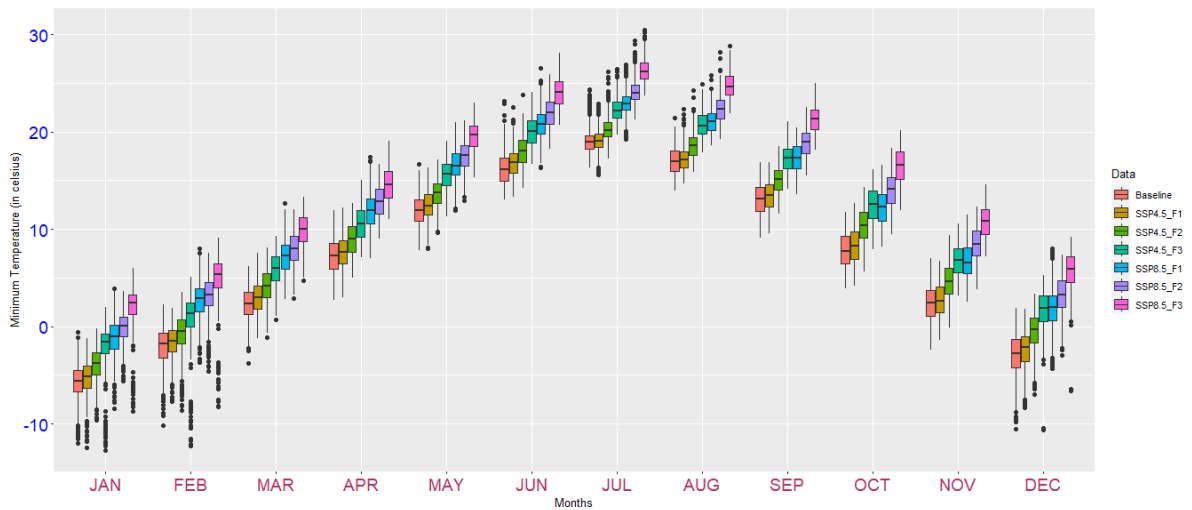


Figure 4.25: Future Minimum Temperature trend since (2017-2100) of region 1 using SSP 4.5 and SSP 8.5

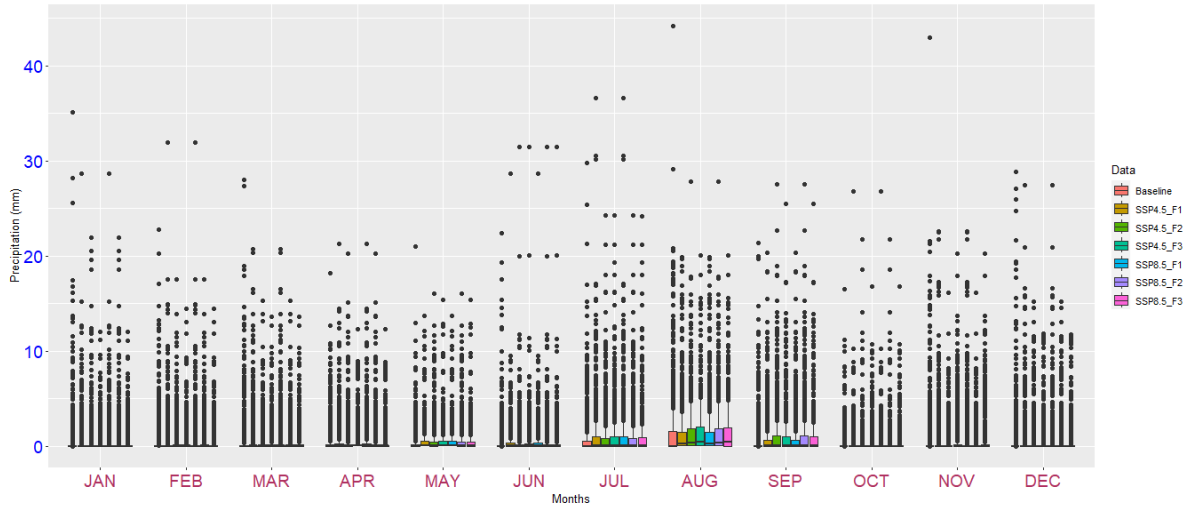


Figure 4.26: Future Rainfall trend since (2017-2100) of region 1 using SSP 4.5 and SSP 8.5

4.7.2 Region 2

When compared to the baseline data, the research shows that maximum temperatures are rising in region 2 for all scenarios, including SSP 4.5 and SSP 8.5. Let's examine the specifics of the findings for rainfall, lowest temperature, and maximum temperature. Maximum Temperature: In comparison to the region 2 baseline data, all scenarios for maximum temperature exhibit an increased trend. The baseline information, SSP 4.5 F1 forecasts, and SSP 4.5 F2 projections are fairly close to the baseline information, indicating a mild rise in temperature. However, there is a noticeable and obvious upward trend in temperature as we approach the later time of SSP 4.5 F3 and SSP 8.5. Figure (4.27) that maximum temperatures in region 2 will significantly rise as time goes on, especially under the high emissions scenario.

Minimum Temperature: Up until the year 2100, there is a markedly rising trend in minimum temperatures for both the SSP 4.5 and SSP 8.5 scenarios, similar to the maximum temperature trend. Figure (4.28) suggests that under both moderate and high emissions scenarios, the region is expected to suffer rising minimum temperatures over the course of future time periods. Rainfall: In terms of rainfall, Figure (4.29)

shows a substantial increase compared to the baseline data. This means that the projected future precipitation rates are higher than what is observed in the baseline data. However, specific values or magnitudes of the increase in rainfall are not provided in the given information.

Overall, in region 2, the analysis suggests an increasing trend in maximum temperatures across all scenarios. The baseline data, SSP 4.5 F1, and SSP 4.5 F2 projections show a moderate increase, while SSP 4.5 F3 and SSP 8.5 indicate a significant and clear rise in maximum temperatures. Additionally, both SSP 4.5 and SSP 8.5 scenarios exhibit a notable increase in minimum temperatures throughout the future periods. Furthermore, the rainfall pattern indicates a substantial increase compared to the baseline data, projecting a higher precipitation rate in region 2 in the future.

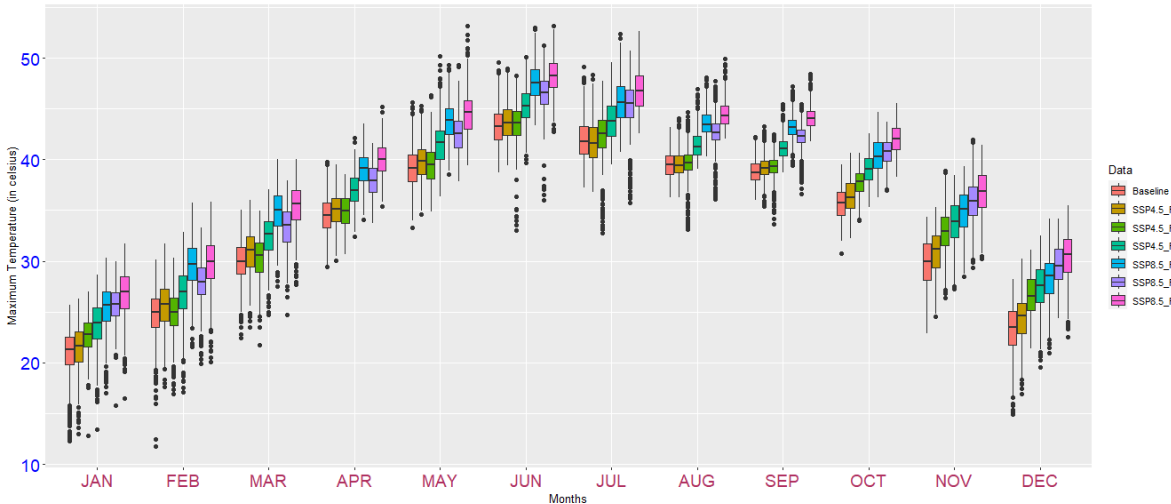


Figure 4.27: Future Maximum Temperature trend since (2017-2100) of region 2 using SSP 4.5 and SSP 8.5

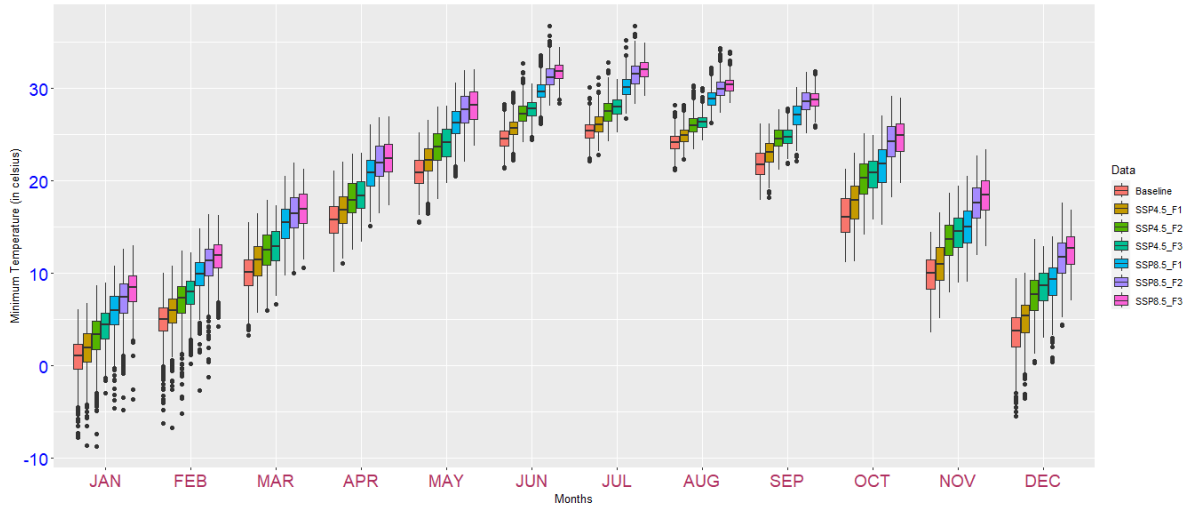


Figure 4.28: Future Minimum Temperature trend since (2017-2100) of region 2 using SSP 4.5 and SSP 8.5

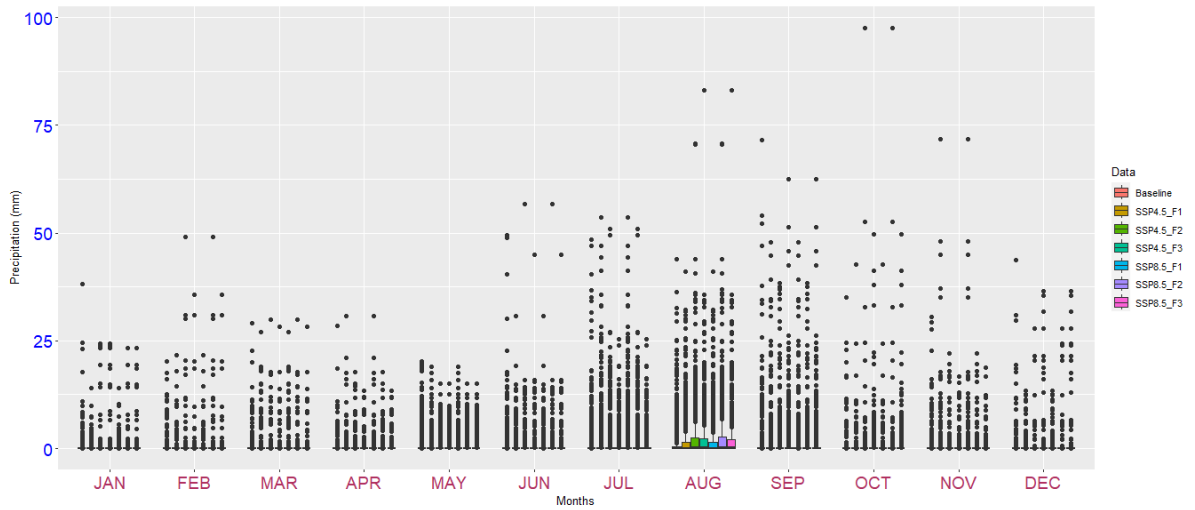


Figure 4.29: Future Rainfall trend since (2017-2100) of region 2 using SSP 4.5 and SSP 8.5

4.7.3 Region 3

According to the analysis, maximum temperatures in region 3 are rising in all time periods and scenarios. Under the SSP 4.5 scenario, there is a noticeable temperature

rise, particularly in the years 2073-2100. Figures (4.30, 4.31 and 4.32) that under the moderate emissions scenario, maximum temperatures in area 3 will rise significantly in the latter part of the century. Across all scenarios and time periods, a considerably rising trend in minimum temperatures has been seen. Comparing the baseline data to all the forecasted scenarios, the baseline data's median is lower. This shows that, independent of the emissions scenario, area 3 is anticipated to have a continuous upward trend in minimum temperatures over the upcoming timeframes. The SSP 8.5 scenario from 2017 to 2100 shows a noticeable variation in precipitation patterns. According to the analysis, there is a noticeable shift in precipitation amounts in the high emissions scenario, with an increase compared to the baseline data. According to SSP 8.5, it appears that area 3 will likely see increased precipitation rates in the coming years. In contrast, the research shows that the precipitation amounts closely match the baseline data when taking into account the SSP 4.5 scenarios. This shows that the precipitation pattern in region 3 would stay consistent with historical observations under the moderate emissions scenarios.

It's crucial to note that the precise magnitudes of changes in precipitation and temperature are not mentioned in the material because they are normally covered in the research or study from which this analysis is produced. Additionally, neither the information nor the comparative data are specific about the difference in precipitation levels under SSP 8.5. Across all scenarios and time periods, the analysis generally shows rising maximum temperatures in region 3. The median values of minimum temperatures show a clear rising trend that is regularly higher than what was seen in the baseline data. The SSP 4.5 scenarios nearly match the baseline data in terms of precipitation amounts, however only the SSP 8.5 scenario reveals a noticeable difference with an increase in rain fall pattern.

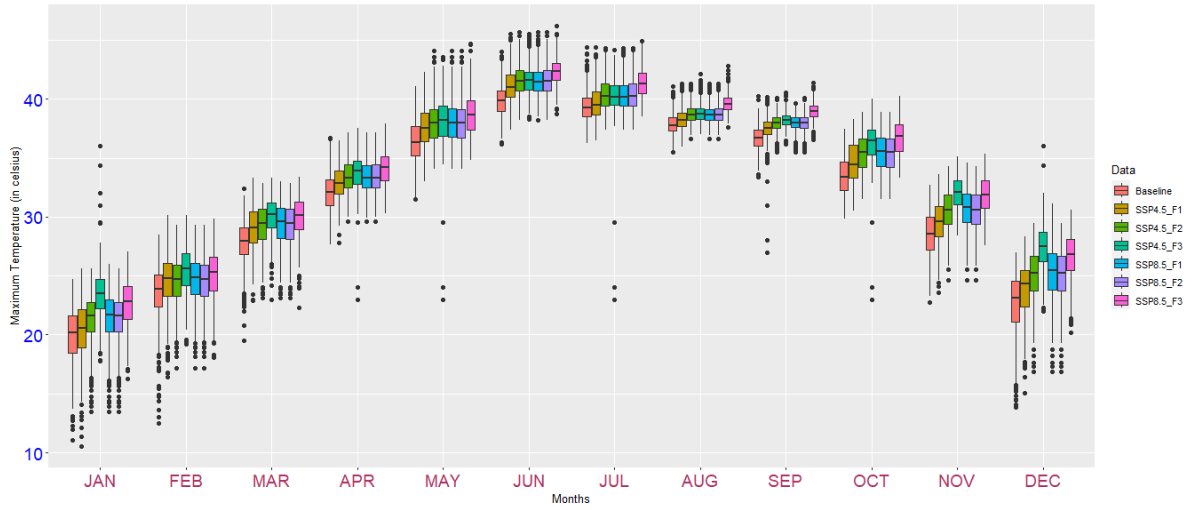


Figure 4.30: Future Maximum Temperature trend since (2017-2100) of region 3 using SSP 4.5 and SSP 8.5

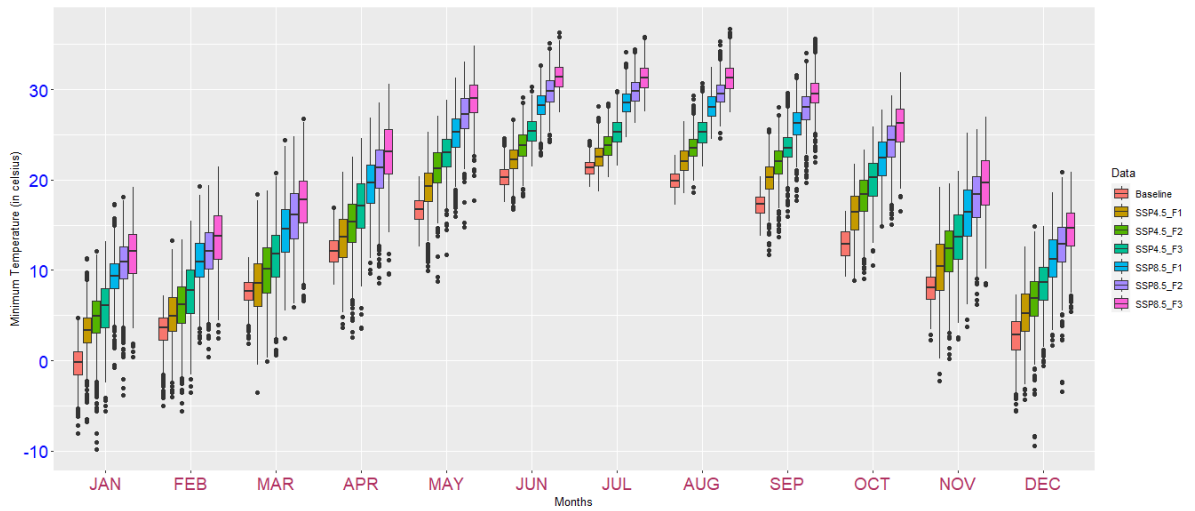


Figure 4.31: Future Minimum Temperature trend since (2017-2100) of region 3 using SSP 4.5 and SSP 8.5

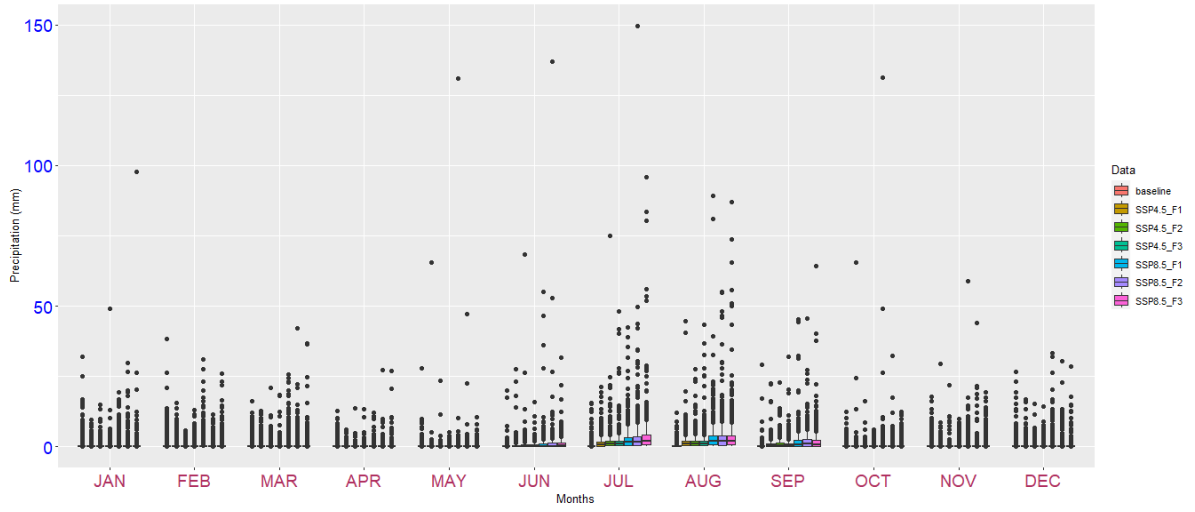


Figure 4.32: Future Rainfall trend since (2017-2100) of region 3 using SSP 4.5 and SSP 8.5

Chapter 5

Conclusions and Recommendations

5.1 Conclusion

A major global problem that has serious effects on both our planet and societies is climate change. A study was carried out in Pakistan to look at how the homogenous climatic regions (HCRs) there were affected by climate change. Based on observed data, the study used a mix of cluster analysis and the L-moments technique to discover and characterise three unique HCRs. Two statistical methods based on L-moments were used to evaluate the consistency within the observed regions and the variation among weather stations. Understanding the patterns and peculiarities of the climatic data within each HCR is aided by these measurements. Data from global climate models (GCMs) was used to forecast the climate of these locations in the future. To ensure consistency between the model output and the observed data, changes were done using bias correction before using the data from the GCMs. The data from the model and the observed data are brought into alignment through this correction process, improving the projections' accuracy.

When the data from the GCMs were compared to the observed and bias-corrected data, it was found that they were tightly aligned, meaning that the data from the models closely matched the observed data for both scenarios and all future time intervals. As a result, there is more reason to believe that the model projections are accurate. According to the study, there have been more fluctuations within some regions as a result of climate change, and differences have been seen at various weather stations. All

future time intervals for both scenarios show the same fluctuations and inconsistencies, indicating that climate change impacts are anticipated to be extensive and long-lasting.

In terms of temperature projections, it is anticipated that maximum and minimum temperatures will rise during the coming time periods in all HCRs. Particularly under the SSP 8.5 and SSP 4.5 scenarios, the rise in temperatures is anticipated to be more pronounced in the period between 2073-2100. This highlights the need for adaptation techniques to deal with the possible effects of increasing temperatures and shows a continuing warming trend. Zone 2 is predicted to have higher precipitation levels than the baseline data in terms of precipitation. The remaining zones, on the other hand, are anticipated to keep their baseline levels of precipitation. This brings to light the regional variation in precipitation patterns and emphasizes the significance of region-specific adaptation strategies. Overall, this study offers insightful information about the various levels of climate change in Baluchistan, Pakistan. Policymakers and stakeholders can create plans and policies to lessen the negative effects of climate change and improve resilience in the impacted regions by analyzing the expected changes in temperature and precipitation.

5.2 Recommendation

To combat the negative impacts of global warming and its influence on the environment and human communities, mitigation measures for climate change are crucial. The suggested strategy for Punjab, Pakistan, combines increasing environmental awareness, improving environmental knowledge, and enforcing laws and regulations. It also places a focus on cooperation and collaboration between different parties.

Raising Awareness and Developing Environmental Knowledge: It's critical to educate and teach the general public about the truth and effects of climate change at the grass-roots level. It is possible to organize awareness programs to spread knowledge about Punjab's rising temperatures' causes and impacts. People will be inspired to take personal action to lower their carbon footprint by seeing how urgent the situation is.

Collaboration: To effectively implement climate change adaptation and miti-

gation strategies, legislators, governmental agencies, authorities, local residents, and environmental nonprofits must work together. An inclusive strategy guarantees that decisions are well-informed, reflect various viewpoints, and receive wider support.

Temperature Monitoring and Research: To comprehend the scope and trends of climate change, it is crucial to continuously record temperature variations at numerous sites across Punjab. The study of the precise effects of rising temperatures on numerous industries, including agriculture, water resources, and public health, should be encouraged.

Strict Regulations on Greenhouse Gas Emissions: The government needs to act right away to rein in the excessive demand and supply of energy sources that contribute to climate change by producing greenhouse gases. The region's carbon footprint may be greatly reduced by enacting tough laws on businesses and cars to cut emissions.

Promotion of Renewable Energy: Promoting the usage of renewable energy sources, such solar, hydroelectric, and wind turbines, can significantly reduce greenhouse gas emissions. In addition to reducing climate change, investing in renewable energy infrastructure also diversifies the energy mix and lessens reliance on fossil fuels.

Economic Benefits: Changing to renewable energy sources may open up new business opportunities and lead to the creation of jobs in the renewable energy industry. Reduced reliance on imported fossil fuels can also increase the nation's energy security and have long-term positive economic effects.

Agriculture Efficiency: The agriculture industry faces substantial difficulties as temperatures rise. To boost agricultural productivity and lessen water stress, however, use climate-resilient agricultural methods and use renewable energy for irrigation.

Stabilizing Earth's Average Temperature: Punjab may support international efforts to stabilize the Earth's average temperature, which is essential for preventing the most serious effects of climate change. This can be done by collectively reducing greenhouse gas emissions.

In conclusion, combating Punjab, Pakistan's rising temperatures, calls for a comprehensive strategy that incorporates numerous stakeholders and makes use of a com-

bination of mitigation and adaptation techniques. The government can guide the area toward a more sustainable and climate-resilient future, which will be advantageous to both the environment and the economy, by prioritizing renewable energy sources and implementing tough controls on greenhouse gas emissions.

Bibliography

- [1] Naomi Oreskes. The scientific consensus on climate change. *Science*, 306(5702):1686–1686, 2004.
- [2] Valérie Masson-Delmotte, Panmao Zhai, Hans-Otto Pörtner, Debra Roberts, Jim Skea, Priyadarshi R Shukla, Anna Pirani, Wilfran Moufouma-Okia, Clotilde Péan, Roz Pidcock, et al. Global warming of 1.5 c. *An IPCC Special Report on the impacts of global warming of*, 1(5):43–50, 2018.
- [3] Priyadarshi R Shukla, Jim Skea, E Calvo Buendia, Valérie Masson-Delmotte, Hans Otto Pörtner, DC Roberts, Panmao Zhai, Raphael Slade, Sarah Connors, Renée Van Diemen, et al. Ipcc, 2019: Climate change and land: an ipcc special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. 2019.
- [4] Ross Garnaut. The garnaut climate change review. *Cambridge, Cambridge*, 2008.
- [5] Mike Bonell and Graham Sumner. Autumn and winter daily precipitation areas in wales, 1982–1983 to 1986–1987. *International Journal of Climatology*, 12(1):77–102, 1992.
- [6] Thomas R Karl and Richard W Knight. Secular trends of precipitation amount, frequency, and intensity in the united states. *Bulletin of the American Meteorological society*, 79(2):231–242, 1998.

- [7] Harry F Lins and Patrick J Michaels. Increasing us streamflow linked to greenhouse forcing. *Eos, Transactions American Geophysical Union*, 75(25):281–285, 1994.
- [8] Carlo Buontempo, Camilla Mathison, Richard Jones, Karina Williams, Changgui Wang, and Carol McSweeney. An ensemble climate projection for africa. *Climate dynamics*, 44:2097–2118, 2015.
- [9] Hayley J Fowler, Stephen Blenkinsop, and Claudia Tebaldi. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 27(12):1547–1578, 2007.
- [10] Rajendra K Pachauri, Myles R Allen, Vicente R Barros, John Broome, Wolfgang Cramer, Renate Christ, John A Church, Leon Clarke, Qin Dahe, Purnamita Dasgupta, et al. *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Ipcc, 2014.
- [11] G Tsakiris, Dialecti Pangalou, and H Vangelis. Regional drought assessment based on the reconnaissance drought index (rdi). *Water resources management*, 21:821–833, 2007.
- [12] Hamd Ullah, Muhammad Akbar, and Firdos Khan. Construction of homogeneous climatic regions by combining cluster analysis and l-moment approach on the basis of reconnaissance drought index for pakistan. *International Journal of Climatology*, 40(1):324–341, 2020.
- [13] Michael J Hayes, Mark D Svoboda, Donald A Wihite, and Olga V Vanyarkho. Monitoring the 1996 drought using the standardized precipitation index. *Bulletin of the American meteorological society*, 80(3):429–438, 1999.
- [14] John A Keyantash and John A Dracup. An aggregate drought index: Assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage. *Water Resources Research*, 40(9), 2004.

- [15] JD van Rooyen, AP Watson, and JA Miller. Combining quantity and quality controls to determine groundwater vulnerability to depletion and deterioration throughout south africa. *Environmental Earth Sciences*, 79(11):255, 2020.
- [16] Jonathan RM Hosking. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the royal statistical society: series B (methodological)*, 52(1):105–124, 1990.
- [17] Anil K Jain and Richard C Dubes. *Algorithms for clustering data*. Prentice-Hall, Inc., 1988.
- [18] Glenn W Milligan. An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *psychometrika*, 45:325–342, 1980.
- [19] T Warren Liao. Clustering of time series data survey. *Pattern recognition*, 38(11):1857–1874, 2005.
- [20] J MacQueen. Classification and analysis of multivariate observations. In *5th Berkeley Symp. Math. Statist. Probability*, pages 281–297. University of California Los Angeles LA USA, 1967.
- [21] Jonathan Richard Morley Hosking and James R Wallis. *Regional frequency analysis*. 1997.
- [22] JRM Hosking and JR Wallis. Some statistics useful in regional frequency analysis. *Water resources research*, 29(2):271–281, 1993.
- [23] Alex J Cannon, Stephen R Sobie, and Trevor Q Murdock. Bias correction of gcm precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes. *Journal of Climate*, 28(17):6938–6959, 2015.

Appendix