

2024-11

Evaluating the Performance of Cmic6 Gcms for Simulation of Rainfall Climatology over Ethiopia

Awoke, Simegnew

<http://ir.bdu.edu.et/handle/123456789/17390>

Downloaded from DSpace Repository, DSpace Institution's institutional repository



BAHIR DAR UNIVERSITY
COLLEGE OF AGRICULTURE AND ENVIRONMENTAL SCIENCES
DEPARTMENT OF NATURAL RESOURCE MANAGEMENT
POSTGRADUATE PROGRAM
EVALUATING THE PERFORMANCE OF CMIP6 GCMS FOR SIMULATION OF
RAINFALL CLIMATOLOGY OVER ETHIOPIA

MSc THESIS

BY

SIMEGNEW AWOKE

NOVEMBER 2024
BAHIR DAR, ETHIOPIA

BAHIRDAR UNIVERSITY
COLLEGE OF AGRICULTURE AND ENVIRONMENTAL SCIENCES
DEPARTMENT OF NATURAL RESOURCE MANAGEMENT

**EVALUATING THE PERFORMANCE OF CMIP6 GCMs FOR SIMULATION
OF RAINFALL CLIMATOLOGY OVER ETHIOPIA**

M. Sc. Thesis

By

SIMEGNEW AWOKE

**A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Science (MSc) in “Environment and Climate Change”**

Advisor: Temesgen Gashaw (Ph.D.)

November 2024

Bahir Dar, Ethiopia

©2024 SIMEGNEW AWOKE

APPROVAL SHEET

As a member of the Board of Examiners of the Master of Science (M.Sc.) thesis open defense examination, we have read and evaluated this thesis prepared by **Mr. Simegnew Awoke** entitled **“Evaluating the Performance of CMIP6 GCMs for Simulations of Rainfall Climatology over Ethiopia”**. We hereby certify that the thesis is accepted for fulfilling the requirements for the award of the degree of Master of Science (M.Sc.) program in **“Environment and climate change”**.

Board of Examiners

Tatek Belay (Ph.D)_____



14 Feb, 2025

Name of External Examiner

Signature

Date

Solomon Addisu (Ph .D)_____



12/07/25

Name of Internal Examiner

Signature

Date

Abebe Aschalew (Ph.D)____



11/07/25

Chair person's name

Signature

Date

DECLARATION

This is to certify that the thesis prepared by Simegnew Awoke Shumetic (ID. No BDU1401868), entitled: **“Evaluating the Performance of CMIP6 GCMs for Simulations of Rainfall Climatology over Ethiopia”** and submitted in partial fulfillment of the requirements for the Degree of Masters of Science in **“Environment and Climate Change”** complies with the regulation of the University and meets the accepted standards with respect to originality and quality. This thesis is my work and has not been presented for a degree in any other university or any other award.



Simegnew Awoke

Dec 12, 2024

Name of the student

Temesgen Gashaw

Name of advisor

Signature



Signature

Date

11/ 9/25

Date

ACKNOWLEDGMENTS

My first thanks extend to the All-Powerful God, who has given me good health throughout this study and my entire life. I sincerely thank Dr. Temesgen Gashaw, who served as my thesis advisor, for his assistance and direction throughout the thesis research process.

I gratefully thank Sellale University for offering the opportunity to attend and paying the necessary expenses. I'm also grateful to thank Ethiopian Meteorology Institute and the creators of the CMIP6 GCMs model datasets for making the information available without charge. Endeg Aniley and Alehegni Belay helped me in all research process; hence I am very delighted to acknowledge their contributions.

I would like to express my sincere gratitude to my parents for their support and moral encouragement throughout my research Process. Finally, but not least, Bahir Dar University, all coworkers, staff members are also gratefully recognized, for their encouragement and assistance throughout this study.

DEDICATION

This thesis is dedicated to my lovely family for their concern, prayer, and unwavering support and encouragement throughout my whole academic career and life.

LIST OF ABBREVIATIONS AND ACRONYMS

AEZs	Argo Ecology Zones
CMIP	Coupled Model Inter-comparison Project
CMIP6	Coupled Model Inter-comparison Project Phase 6
CMIP5	Coupled Model Inter-comparison Project Phase 5
CHCN-M	Global Historical Climatology Network
CHIRPS	Climate Hazard Infrared precipitation with Stations
CMIP6 GCMs	Coupled Model Inter-comparison Project Phase 6 Global Climate Models
CRU	Climate Research Unit
CLIVAR	Climate Variability and Predictability
CPC	Climate Prediction Center
CRI	Critical Rating Index
ENACTs	Enhanced Climate Services for Tropics
ESGF	Earth System Grid Federation
GCMs	Global Climate Models
GPCC	Global Precipitation Climatology Center
IPCC AR5	International Panel on Climate Change Fifth Assessment report
KGE	Kling Gupta Efficiency
MAE	Mean Absolute Error
ME	Mean Error
NSE	Nash-Sutcliffe Efficiency
PBIAS	Percent Bias

PCMDI	Program for Climate Model Diagnosis and Inter-comparison
PRECL	Precipitation Reconstruction over Land
R	Pearson Correlation Coefficient
R2	Coefficient of determination
RMSE	Root Mean Square Error
UDEL	University of Delaware
WCRP	World Climate Research Program
WGCM	Working Group on Coupled Modeling
WMO	World Meteorological Stations

LIST OF TABLES

Table 2. 1 Summary characteristics of global gauge-based Products.....	8
Table 2.2 Summary for previously evaluated rainfall simulation performance of CMIP6 GCMs in the world	15
Table 3. 1 The list of CMIP6 GCMs employed in this study and their characteristics	20
Table 3. 2 Summary of continuous performance metrics used in this study	24
Table 4. 1 Performance of CMIP6 GCMs for simulating daily rainfall over AEZs of Ethiopia..	27
Table 4. 2 Performance of CMIP6 GCMs for simulating monthly rainfall over AEZs of Ethiopia	33
Table 4. 3 Performance of CMIP6 GCMs for simulating annual rainfall over AEZs of Ethiopia	42

LIST OF FIGURES

Figure 3. 1 The location of Ethiopia from Africa and the meteorological stations used in this study.....	17
Figure 3. 2 Research framework of the study	22
Figure 4. 1 Line graph of mean monthly rainfall, (observed with CMIP6 GCMs data from 1995 to 2014)	38

LIST OF APPENDIX TABLE

Appendix Table 1. Description of studied meteorological stations over traditional AEZs of Ethiopia 61

ABSTRACT

Rainfall is one of the most crucial climatic elements to examine and investigate the effects of the changing climate. Simulations of Global Climate Models (GCMs) have become the most useful and commonly used tools for examining changing in climate systems and simulating future climate. This study evaluated the performance of Coupled Model Inter-comparison Project Phase 6 Global climate models (CMIP6 GCMs) for simulations of rainfall climatology over Ethiopia. The study evaluated the rainfall simulation performance of eight CMIP6 GCMs at daily, monthly, and annual temporal scale over five Argo Ecology Zones (AEZs) of Ethiopia for the period 1995-2014 using station data as reference. The performance evaluation was carried out using continuous statistical metrics like Root mean square error (RMSE), Percent bias (PBIAS), and Pearson Correlation Coefficient (r) and line graphs. The performance of each GCM over different AEZs was ranked based on Critical Rating Index (CRI). The result revealed that; models have different simulation performance of daily to annual rainfall total over AEZs of Ethiopia. In addition, they have both overestimation and underestimation problem. For better simulation of daily rainfall total, Ec-Earth3-veg (over tropical, sub-tropical, and temperate AEZs),MRI-CSM2-0 (over desert AEZ),and MPI-ESM-1-2-LR (over alpine AEZ) is high ranked and preferable.BCC-CSM-2-MR performed well over tropical, sub-tropical, temperate, and alpine AEZs,butMRI-ESM2-0 has better performance over desert AEZ. For simulation of annual rainfall total, MRI-ESM2-0 (over desert and tropical AEZs), BCC-CSM-2-MR (over temperate and alpine AEZs), and Ec-Earth3-veg (over sub-tropical AEZ) is preferable and acceptable. Ec-Earth3 is the worst climate model to reproduce mean monthly rainfall of Ethiopia. The performance of models is not consistent in reproducing rainfall distributions at different statistical metrics and timeframes. Since the performance of CMIP6 GCMs different from place to place and time to time, I recommend to evaluate the performance of models to use for specific place and application and apply appropriate bias correction. The author also further recommends, evaluating the performance of CMIP6 GCMs rainfall simulation capability over rainfall regimes of Ethiopia.

Keywords: AEZs, Climate modes, CMIP6 GCMs, Rainfall, Performance evaluation

CHAPTER 1. INTRODUCTION

1.1. Background of the Study

Climate change is the most important and primary topics threatening humanity in the twenty-first century. Extreme climatic phenomena are becoming more frequent and intense, affecting an existential threat to agricultural livelihoods, water resources, human health, and economic growth in different regions of the world (Salman *et al.*, 2020; Shiru *et al.*, 2020; Chang *et al.*, 2024). Rainfall is one of the most crucial climate element to examine and investigating the effects of changing climate (Chang *et al.*, 2024; Ding *et al.*, 2024). The accelerated warming of the earth's surface have a variety of consequences for the amount of rain that recorded in a given area at specific time (Gebrechorkos *et al.*, 2019; Weber *et al.*, 2020; Ding *et al.*, 2024). A variety of natural disasters, such as floods and droughts, are caused directly or indirectly by variations in the distribution and intensity of rainfall (Field *et al.*, 2015; Chang *et al.*, 2024; Ding *et al.*, 2024; Guo *et al.*, 2024). Continued and accelerated greenhouse gas emissions will increase climate change and variability, increasing the frequency and magnitude of droughts and floods and exposing the population to water-related hazards (Almazroui *et al.*, 2020; Mohammed & Yimam, 2022).

In fact, the effects of hydrological and other climate change and variability risks are not uniform around the world. The change in rainfall intensity and frequency has major impact for livelihoods, particularly in poor countries. In line with this, East Africa is one of the most exposed regions to climate change and hydrological disasters due to its high exposure and poor adaptive capacity (Weber *et al.*, 2020; Chang *et al.*, 2024; Ding *et al.*, 2024; Guo *et al.*, 2024). Now a time, climate disasters are also expected to have an impact on Ethiopia (Mohammed & Yimam, 2022; Wubaye *et al.*, 2023). Therefore, understanding past and present rainfall related disasters helped to addressing future rain-related disasters to increase adaptive capacity of consistent and adequate annual and seasonal rainfall trend analysis and rainfall forecasting (Wubaye *et al.*, 2023).

However, simulations of the GCMs have become the most useful and commonly used tools for examining changes in climate systems caused by increases in radiative forcing and simulating historical and future climates(Taylor *et al.*, 2012;Guo *et al.*, 2024). Starting from the 1950s to the present, researchers in many nations of the world from several climate modeling centers have produced GCMs(Edwards, 2011).Due to significant advancements in GCMs, drawbacks arise because of insufficient parameterization, coarse spatial resolution, and the difficulties of addressing national and regional-scale processes(Flato *et al.*, 2014;Diallo *et al.*, 2019).So, these issues in GCMs results aggravated by uncertainty in future climate change estimates, especially in areas with variable topography(Flato *et al.*,2014).

To overcome these challenges, many GCMs are produced and distributed under Coupled Model Inter-comparison Projects (CMIPs). CMIP is a set of GCMs with improvements in resolutions and including many variables, with CMIP6 becoming the most latest phase (Eyring, Bony, *et al.*, 2016). However, the ability of GCMs to simulate current and past climates used as a benchmark for accuracy of future climate projections. Due to this, the performance of GCMs must be evaluated and verified against reliable observed data with adequate spatial coverage(Mohammed, 2024).

Few studies have evaluated the performance of CMIP5 models in Ethiopia (Jury, 2015;Li *et al.*, 2016;Dyer *et al.*,2019)and they indicated that CMIP5 models are still susceptible to significant biases in simulating seasonal and annual rainfall over the country. In addition, few studies have evaluated and utilized the detection capability of CMIP6 models in Ethiopia (Alaminie *et al.*, 2021; Dyer *et al.*, 2022; Fetene *et al.*, 2022).

Ethiopia is one of the most ecologically diverse countries in the world (Kindt *et al.*, 2014). Many factors, including Ethiopia's location in the tropics, its vast elevation range, and its diverse terrain, are responsible for this diversity of ecosystems and agro-ecology. Due to its vast topographical feature, the climate of Ethiopia varies from place to place (Gashaw *et al.*, 2024). Regardless of the high importance of assessing the performances of different CMIP6 models in replicating daily, monthly, and annual rainfall both spatially and temporally, there is limited topography (AEZ) based national study that investigates the performance of different CMIP6 models relevant to impact studies of different sectors.

This study aimed to evaluate the ability of eight Global climate models (GCMs) from the Coupled Model Inter-comparison Project Phase 6 (CMIP6) to simulate the historical rainfall total at daily, monthly, and annual time scales over all five AEZs of Ethiopia, in representation of the spatiotemporal rainfall total before employing them for future climate projections for building strong adaptive capacity and climate resilient strategies at the national scale. The availability of daily data (Gebrechorkos *et al.*, 2023). The criterion applied to use the models were: Their best performance in different parts of Ethiopia, East Africa, and other global areas (Alaminie *et al.*, 2021). Their wider application in the earlier versions CMIP5 in Ethiopia (Worku *et al.*, 2019).

1.2. Statement of the Problem

Rainfall is one of the most significant climatic element to examine and investigating the effects of changing the climate of an area (Chang *et al.*, 2024). Understanding the past and current climate, helped as a benchmark for future climate prediction (Mohammed, 2024). CMIP6 GCMs provides insight in to historical period model performance and offers projections of future scenarios, covering past, present, and future climate (Mohammed, 2024). Even though, CMIP6 GCMs tells us about past, present, and future of an area, the performance of models is inconsistent at space and time and differs from one CMIP6 GCMs to the other (Gashaw *et al.*, 2024). Hence, the simulation performance of CMIP6 GCMs shall be evaluated and bias corrected before use for any application.

Few studies have been evaluated the performance of CMIP6 GCMs at small scale and at national scale. Few studies conducted at national scale were performed by taking the average value and few others by interpolating for the whole Ethiopia using ENACTs data as reference. In fact, ENACTs data has its own performance variation for different areas of Ethiopia, since it is station data incorporated gridded data set based on TAMSAT algorithm (Dinku *et al.*, 2016). But, this study considers by dividing the country in five AEZs (elevation), since the performance of climate models differ by elevation. Due to lack of insufficient in-situ dataset (Omay *et al.*, 2023), used CHIRPS data as a reference, but the performance of CHIRPS differs from low elevation to high elevation areas of the country. Moreover, CHIRPS has high bias at daily, seasonal and annual scales (Aniley *et al.*, 2023; Addisu *et al.*, 2024; Gashaw *et al.*, 2024). Moreover, the daily rainfall total simulation capability of CMIP6 GCMs is not evaluated as a national scale in AEZs of Ethiopia. In fact, daily time step data is highly valuable for future water balance calculation,

potential evapotranspiration estimation, onset date, cessation date, and length of growing period estimation (Dinku *et al.*, 2022).

Therefore, this study aimed to evaluate the performance of eight CMIP6 GCMs those evaluated and recommend to use for application by other researchers in different areas of the country and the globe in general. The study was conducted over five AEZs of Ethiopia to simulate daily, monthly and annual rainfall simulation for the period 1995-2014 using 47-gauge stations as a reference.

1.3. Objectives of the Study

1.3.1. General objective

The general objective of the study was to evaluate the performance of CMIP6 GCMs for simulations of rainfall climatology over AEZs of Ethiopia during 1995-2014 using 47 reference gauge stations as a reference.

1.3.2. Specific objectives

- ✓ To evaluate the performance of CMIP6 GCMs in simulating the daily, monthly and annual rainfall total over each traditional AEZs of Ethiopia.
- ✓ To analyze the seasonal and spatial variability of rainfall climatology over Ethiopia.
- ✓ To assess the accuracy and performance of CMIP6 GCMs in capturing Ethiopia's rainfall patterns.

1.4. Research Questions

- ✓ Which CMIP6 GCMs is best to simulate daily, monthly and annual rainfall total over each traditional AEZs of Ethiopia?
- ✓ Which CMIP6 GCMs can give us a good insight to seasonal and spatial variability of rainfall climatology over Ethiopia.
- ✓ Which CMIP6 GCMs is best to detect annual rainfall total over each traditional AEZs of Ethiopia?

✓

1.5. Significance of the Study

The study is helpful for researchers and decision makers to fill the data gap problem when rainfall data is needed for research and decision-making process. The study also valuable to analyze the past and present climate, and for predicting future climatology, climate change, and climate variability over each AEZ of Ethiopia by considering bias. The findings will also play key roles for applying the most performing CMIP6 GCMs rainfall data for drought monitoring, environmental impact assessment, and hydrological forecasting over the country. Moreover, the best performing CMIP6 GCMs models over each AEZ can be used for designing mitigation and adaptation strategies.

1.6. Scope of the Study

The study was covering all traditional AEZs of Ethiopia using 47 reference gauge station data for the period 1995-2014 at daily, monthly, and annual temporal scale. Performance evaluation of CMIP6 GCMs for rainfall total was done in each AEZ separately and draw conclusion AEZ based and for national level at all temporal scales.

1.7. Limitation of the Study

Due to the presence of small number of stations those have long time series and qualified rainfall data, I only used 47 stations for this study. In line with this, desert and alpine AEZs were only represented by one station each.

CHAPTER 2. LITERATURE REVIEW

2.1. An overview of climate models

Climate models are fundamental tools to acquire past and future climate information through utilization of estimation techniques for the land, ocean, atmosphere, and ice interactions (Randall *et al.*, 2007). Climate models are the primary tools available for investigating the response of the climate system to various forcing's, for making climate predictions on seasonal to decadal time scales and for making projections of future climate over the coming century and beyond. It is crucial therefore to evaluate the performance of these models, both individually and collectively (Kreienkamp *et al.*, 2012). Climate models are important tools for understanding and predicting the complex Earth's climate. Several global climate models (GCMs) have been developed by several research centers around the world.

2.2. Global climate models (GCMs)

Global climate models (GCMs) are mathematical models representing the physics of the climate system, including atmosphere, ocean, land surface, and cryosphere (Randall *et al.*, 2007). GCMs are useful instruments that are often used to research the global climate system for the past, present, and future. Since 1995, scientists from many international institutes have been working to build GCMs in order to produce dependable and effective models for simulating climate systems (Mohammed, 2024). Climate variables of GCMs are provided in grid cell format coverage of the globe (Eyring *et al.*, 2019). GCM outputs from the climate modeling groups around the world have been collected by the World Climate Research Program (WCRP) Working Group on Coupled Modeling (WGCM) to generate climate model experiments that are under the Coupled Model Inter-comparison Project (CMIP).

Global Climate Models (GCMs) are very essential to simulate the earth's climatic conditions and used as an instrument to assess the changes of the earth's climate system as a result of various radiative forcing (Randall *et al.*, 2007; Flato *et al.*, 2014; Dibaba *et al.*, 2019). For this reason, one of the main applications of the GCMs is an investigation with the aim of improving the understanding and knowledge of the climate system. The GCMs are characterized by uncertainties and the climate of mountain and coastal regions of the world (Fowler *et al.*, 2007). The uncertainties in the GCMs simulation could stem from the spatial and temporal coarse

resolution nature of the model, the initial model assumption, and the mathematical formulation considered in the model (Ahmadalipour *et al.*, 2018). Regardless of its uncertainty, GCMs were considered the most important tools for future climate projection (Loukas *et al.*, 2008). However, it is known that GCMs-based future climate variability and trend studies require performance evaluation of GCMs in simulating climate variables (Nasrollahi *et al.*, 2015).

2.3. Gauge-based Rainfall Estimates

The study of climate change and variability commonly depends on surface gauge observation (Strangeways, 2006). Rain gauge is the most commonly tool for directly analyzing point rainfall at the surface measuring the amount of rainfall as it accumulates over time. There are several types of rain gauge, such as accumulation gauges, tipping bucket gauges, weighing gauges, and optical gauges, and all these gauges have their own strengths and weaknesses. Details about gauge instruments are given in (Strangeways, 2006; Tapiador *et al.*, 2012).

Gauge-based observation of precipitation is collected by national weather services. For climate research and other applications, it is necessary to organize all of the data from different countries in to one integrated global dataset. The World Meteorological organization (WMO), which is established in 1873, is an intergovernmental organization with 191 member states and territories (<http://www.wmo.int>). The WMO facilitates the development of observation networks for fields of climatology, hydrology, and physics as well as the exchange, processing, and standardization of related data. The WMO Global Telecommunication System (GTS) is a complex worldwide communication system for the exchange of meteorological data that forms the basis for many applications (Spence & Townshend, 1996). The total number of rain gauges operated worldwide is estimated to be between 150000 and 250000 (Groisman & Legates, 1995; New *et al.*, 2001). Although many gauges exist, not all have operated in a continuous manner or concurrently (Kidd, 2001). Due to the irregular distribution of gauge stations, gridding of data is required for many climate and related applications. The Gridded precipitation datasets based fully on gauge information have been constructed and in wide used. The climate research unit (CRU) dataset comprises a suit of climate variables, including precipitation. The primary sources used for construction of the CRU monthly precipitation dataset were obtained through the help of the

national meteorological agencies (NMAs), the WMO, the CRU, the international Tropical Agriculture Center, the food and Agriculture Organization (FAO), and others (Sun *et al.*, 2018).

The Global Precipitation Climatology Center (GPCC) has established a unique ability to collect, organize, quality control, and analyze rain gauge data from across the globe. The NMAs provide the primary data for the GPCC, with 158 countries and 31 regional suppliers providing the majority of the gauge data in to the GPCC dataset. Details of gauge dataset are given by (Sun *et al.*, 2018). Some characteristics of gauge-based products are listed in Table (2.1).

Table 2. 1Summary characteristics of global gauge-based Products

Dataset	Resolution	Frequency	Coverage	Period	Source	Reference
CRU	0.5°×0.5°	monthly	Global-Land	1901-2015	The CRU of the university of East Angeli a	Harris <i>etal.</i> (2014)
GHCN-M	5°×5°	monthly	Global-Land	1900-present	NCDC	Peterson & Vose (1997)
GPCC	0.5°×0.5° 1°×1° 2.5°×2.5°	monthly	Global-Land	1901-2013	GPCC	Rudolf <i>et al.</i> (2011)
GPCC-daily	1°×1°	Daily	Global-Land	1988-2013	GPCC	Schamm <i>etal.</i> (2014)
PRECL	0.5°×0.5° 1°×1° 2.5°×2.5°	monthly	Global-Land	1948-2012 (1(0.5)) 1948-present	NCEP/NOAA	Chen <i>et al.</i> (2002)

UDEL	0.5°×0.5°	monthly	Global-Land	1900-2014	University of Delaware	Willmott & Matsuura (1995)
CPC-Global	0.5°×0.5°	Daily	Global-Land	1979-2005	CPC	Xie <i>et al.</i> (2010)

2.4. The Coupled Model Inter-comparison Project (CMIP)

The Coupled Model Inter-comparison Project (CMIP) was established to study and compare climate simulations made with coupled ocean, atmosphere, cryo-sphere, land GCMs. The Coupled Model Inter-comparison Project (CMIP) has been coordinating simulations conducted by international modeling groups since the 1990s, with the aim of improving the performance of climate models and enhancing the scientific understanding of the climate system. This project has become a major tool for climate science, and remarkable progress has been achieved for global climate models (GCMs) (Stouffer *et al.*, 2017).

The first gathering of the international global coupled climate modeling community was at a workshop held at Scripps Institution of Oceanography in La Jolla, California, in October 1994 organized under the auspices of the World Climate Research Program (WCRP). The meeting was convened specifically to examine the state of the art in global coupled climate modeling. It was recommended that an “inter-comparison be performed for the set of models in use” (Meehl, 1995). At about the same time, data from many of these models were collected and analyzed by S. Lambert and G.J. Boer for the Intergovernmental Panel on Climate Change Second Assessment Report (Gates *et al.*, 1995).

The Coupled Model Inter-comparison Project (CMIP) was initiated by the Climate Variability and Predictability (CLIVAR) Numerical Experimentation Group 2 (NEG2, subsequently reconstituted as the WCRP Working Group on Coupled Models, WGCM) late in 1995, partially as an outgrowth of these two efforts. The next year under the first phase of CMIP (CMIP1), model data from unforced climate from 21 global coupled atmosphere-ocean models were archived at the U.S. Department of Energy Program for Climate Model Diagnosis and Inter-comparison

(PCMDI) at the Lawrence Livermore National Laboratory. The second phase of CMIP (CMIP2) was designed to compare the climate changes simulated by the models for an idealized change in forcing of 1% per year increase in CO₂. CMIP2 was initiated in early 1997, and data were collected from 17 of the CMIP1 models. The Coupled Model Inter-comparison Project 5 (CMIP5) provided numerous GCMs that were more complex and accurate in depicting climate change than previous-generation models (Hasson *et al.*, 2016). However, bias still existed in providing climate information on a regional scale (Jia *et al.*, 2019).

2.4.1. Coupled Model Inter-comparison Project Phase 6 (CMIP6)

CMIP6 models have typically updated versions of the models that participated in previous phases of CMIP. Most recent, simulations from state-of-the-art versions of GCMs have become available through the Coupled Model Inter-comparison Project Phase 6 (CMIP6), with a substantial increase in the number and scope of experiments (Eyring *et al.*, 2016). These simulations provide a new opportunity to evaluate the Earth system response to change in radiative forcing during the 21st century.

This new set of models has improved spatial resolution, physical parameterizations (i.e., cloud microphysics), and better representations of various earth system processes (such as biogeochemical cycles) and components (such as ice sheets) (Eyring *et al.*, 2016). The evaluation of CMIP6 simulations over several regions suggests that model output exhibits differences from earlier CMIP phases (Almazroui *et al.*, 2020).

2.5. CMIP6 Model Evaluation System

The evaluation of models has advanced significantly over the past few decades. The CMIP community has now reached a key point when a great deal of efficiency is needed to be achieved in many baseline elements of model evaluation. This will allow for a systematic and quick performance assessment of the various models included in CMIP. Such an evaluation system will be implemented for CMIP6 (Almazroui *et al.*, 2020). The initial goal is that two capabilities will be used to produce a broad characterization of CMIP DECK and historical simulations as soon as new CMIP6 model experiments are published to the Earth System Grid Federation (ESGF):

It was decided at the WGCM conference that these tools' outputs might be seen on a public website instead of one that was password-protected. Watermarking will be applied to the results at first till a quality control has taken place. The WGCM approved of this tactic.

(a) the Earth System Model Evaluation Tool (ESMValTool)(Eyring *et al.*, 2016) is a community-developed diagnostic and performance metrics tool for the evaluation of Earth system models with observations. It includes other well-established model evaluation packages such as the NCAR Climate Variability Diagnostics Package (CVDP)(Phillips *et al.*, 2014). The collection of standard namelists for example allows to reproduce the figures from the climate model evaluation chapter of IPCC AR5 (Chapter 9) and parts of the projection chapter (Chapter 12). The ESMValTool is available as open-source software on GitHub. The website here shows results produced with the ESMValTool for CMIP5 simulations. This website will be updated with CMIP6 results as soon as the model output is submitted to the ESGF. All modeling groups are encouraged to check the results for their model.

(b) the PCMDI Metrics Package (PMP) (Gleckler *et al.*, 2016) emphasizes a diverse suite of summary statistics to objectively gauge the level of agreement between model simulations and observations across a broad range of space and time scales. It is built on the Python based Ultrascale Visualization Climate Data Analysis Tools (UV-CDAT), a powerful software tool kit that provides cutting-edge data management, diagnostic and visualization capabilities. The PMP is available as open-source software on GitHub.

Since these tools are freely available on GitHub, modeling groups participating in CMIP can additionally make use of these packages. They could choose, for example, to utilize the tools during the model development process in order to identify relative strengths and weaknesses of new model versions also in the context of the performance of other models or they could run the tools locally before publishing the model output to the ESGF. Mechanisms are in place to enable contributions from the broader community. Both tools are designed to readily work across ESGF nodes with the intent of ultimately expediting routine analysis by alleviating the needs for data transfer. More details on this approach can be found in(Eyring *et al.*, 2016).

2.6. Continuous statistical metrics of model performance evaluation

Statistical metrics are normally used for assessing the differences between model simulations and observed datasets, and each metric can show the overall (relative) performances differently for the model simulations (Moriasi *et al.*, 2007; Gleckler *et al.*, 2008). Additionally, statistical metrics are often found in the comparison of the climate simulation results between GCMs and observed datasets (Kamworapan & Surussavadee, 2019). Continuous statistical metrics have been employed over different study areas to evaluate the performance of rainfall products rainfall total at different spatial and temporal scales (Nashwan *et al.*, 2019). The continuous statistical indices most commonly employed by most studies are: Root mean square error (RMSE), Mean error (ME), Mean absolute error (MAE), Pearson correlation coefficient (r) and Coefficient of determination (R^2), Kling-Gupta Efficiency (KGE), Nash-Sutcliffe Efficiency (NSE), Percent bias (PBIAS), RMSE-observations standard deviation ratio (RSR) and Mean square error (MSE) (Moriasi *et al.*, 2007). The RMSE represents the standard deviation of residuals. The ME and MAE measure the average magnitude of the error between the rainfall products and the ground-based observation. The major differences between the ME and MAE is that ME can give positive and negative results, whereas MAE only gives positive results (M. L. Tan & Santo, 2018). The RMSE, MAE, ME, and MSE value of 0.0 indicates a perfect fit between rainfall products and observed data set (Singh *et al.*, 2004; Moriasi *et al.*, 2007). A study by Singh *et al.* (2004) state that RMSE and MAE values less than half the standard deviation of the measured data may be considered as low and that either is appropriate for model evaluation. The r and R^2 describe the degree of co-linearity between rainfall products and observed (gauge) rainfall data. The correlation coefficient, which ranges from -1 to 1, is an index of the degree of linear relationship between observed and simulated (observed) data. If $r = 0$, no correlation between observed and simulated (rainfall products) data. If $r = 1$ or -1 , a perfect positive or negative relationship exists between the observed and simulated data set. The R^2 is the proportion of the variance in measured data explained by the model. The value of R^2 ranges from 0 to 1, with a higher value indicating less error variance, and typically values greater than 0.5 are considered acceptable (Santhi *et al.*, 2001; Van Liew *et al.*, 2003). Even though r and R^2 have been widely used for model evaluation, these statistics are over sensitive for high extreme values (outliers) and in sensitive to additive and proportional differences between model prediction and measured data (Legates & McCabe, 1999).

The KGE integrates the correlation, bias, and variability of the ground observed and Rainfall products into one index (Gupta *et al.*, 2009). As compared with the Nash-Sutcliffe Efficiency (NSE), the KGE is less sensitive to extreme rainfall and therefore can describe and measure the overall fitness of rainfall having different intensities (Radcliffe & Mukundan, 2017). The NSE is the normalized statistic which determines the relative magnitude of the residual variance compared to the measured data variance. The NSE indicates how well the plot of the observed versus simulated data fits the 1:1 line. The value of NSE ranges from $-\infty$ to 1, with the optimal value of 1. The NSE value of 0.0 to 1.0 are generally acceptable levels of performance, whereas NSE values less than or equal to 0.0 indicate that the mean observed value is a better predictor than the rainfall products (simulated) value (Nash & Sutcliffe, 1970). The RMSE-observations standard deviation ratio (RSR) standardizes RMSE using the observations standard deviation, and it combines both an error index and the additional information recommended by (Legates & McCabe, 1999). Even though the lower the RMSE, the better the model (rainfall products) performance, a study by Singh *et al.* (2004) has published a guide line to qualify what was considered a low RMSE based on the observations standard deviations. The PBIAS measures the average tendency of the rainfall products or models to be larger or smaller than their observed counterparts (Gupta *et al.*, 1999). The optimal value of PBIAS is 0.0; with low magnitude values indicate accurate model or simulated estimation performance. The positive values of PBIAS indicate a model underestimation bias, and negative values indicate model overestimation bias (Gupta *et al.*, 1999).

2.7. Bias Correction of Climate Models

Model has random and systematic errors that related to numerical error, limits of predictability, and chaotic behavior, and is usually assumed to have zero mean and does not necessarily have zero mean and structural deficiencies in the model, imperfect representation, unresolved dynamics, or imperfect parameterizations (Cannon *et al.*, 2020). Bias correction methods are applied for substantial biases reductions and comparable performances among most GCM-RCM, Observed-Model data pairs (Mendez *et al.*, 2020). All bias correction methods have improved the predicted precipitation values to an acceptable level, which increases the efficiency of the precipitation prediction models in the flood warning system (Motlagh *et al.*, 2022). Global and regional climate models are in general biased, their resolution is often lower than desired. Thus, many

users of climate model data apply some form of bias correction. A fundamental assumption of bias correction is that the considered climate model produces skill full input for a bias correction, including a plausible representation of climate change (Maraun, 2016) and to make the GCM output close to the corresponding observed data as much as possible to decrease the errors of GCM outputs (Tan *et al.*, 2020). The CMhyd software selects the closest observed station to the climate model grid cells to compare observed and simulated historical time series, the grid size considered in the bias correction process is the same for observed (Brighenti *et al.*, 2023). For the suitable selection of an ensemble of GCMs to reduce uncertainties in climate projections, (Lei, 2022). To use the climate simulation outputs, one critical issue to consider is the systematic bias affecting the modeled data. The bias correction of modeled data is often used to assess the effect of climate events on human activities (Trentini *et al.*, 2023).

The bias-correction method to be very effective in reducing the model bias: it removed over 80% and 83% of model bias for surface air temperature and precipitation, respectively, during the validation period (*JGR Atmospheres - 2016*).

2.8. Previous Performance Evaluation of CMIP6 GCMs over different areas

The simulation performances of CMIP6 GCMs have been evaluated in different areas of the world for different applications. Literatures showed that different researchers used different number of gauge stations, study time period, and also type and number of continuous statistical metrics used as well as study area coverage and evaluation timescale. For instance, a study by Berhanu *et al.* (2023), evaluate the performance of GFDL-CSM4 and other 36 CMIP6 GCMs over Ethiopia for the period 1981-2014 and study time scales of monthly, seasonal, and annual by compared with gridded rainfall of Ethiopia data (ENACTs) for reference. Another study conducted over Abay Basin (Ethiopia), also evaluate the performance of BCC-CSM2-MR and other seven datasets for the period 1988 to 2014 using correlation coefficient, root mean square error, bias percentage, Kling-Gupta efficiency and Nash-Sutcliffe efficiency and visual compare on using graphs to simulate the performance of annual and seasonal rainfall totals of models (Mohammed, 2024). A study by Omay *et al.* (2023), also evaluate the performance of CMIP6 GCMs over East Africa for the period 1981-2014 to simulate monthly rainfall climatology of East Africa countries. Omay *et al.* (2023) evaluated ACCESS-ESM1-5 and other 22 models to simulate seasonal rainfall pattern of the region. The authors find that, most models capture

monthly, seasonal, and annual rainfall total, but there is overestimation and underestimation problem. In addition, the performance of CMIP6 GCMs differs from place to place and time to time. Moreover, the climate model which is suitable for temperature may be worst to detect rainfall in the same region and time period (Gashaw *et al.*,2024).Summary forrainfall simulation performance of CMIP6 GCMs in the world given in Table (2.2).

Table 2.2Summary for previously evaluated rainfall simulation performance of CMIP6 GCMs in the world

Study area	Evaluated GCMs	Main finding	Reference
Lake Tana (Ethiopia)	Ec-Earth3 and others	Ec-Earth3, MRI-ESM2-0, and ACCESS-ESM1-5 for daily, CNRM-CM6-1, BCC-CSM2-MR, and Ec-Earth3 for monthly, BCC-CSM2-MR, MPI-ESM1-2-LR, Ec-Earth3, and ACCESS-ESM1-5 for annual performed best.	Lebeza <i>et al.</i> (2024)
Abay basin(Ethiopia)	BCC-CSM2-MR and other seven	BCC-CSM2-MR better captured annual rainfall total.MRI-ESM2-0 for <i>bega</i> ,and BCC-CSM2-MR, MRI-ESM2-0,and Ec-Earth3 better for <i>belg</i> , BCC-CSM2-MR and Ec-Earth3 for <i>kiremt</i> season	(Mohammed, 2024)
Bale Eco region,Ethio pia	BCC-CSM2-MR and 9 others	BCC-CSM2-MR, CNRM-CM6-1,and MRI-ESM2-0 best to simulate daily to annual rainfall climatology of the region.	Gashaw <i>et al.</i> (2024)
S/America	Ec-Earth3-veg and other 37	Most models capture main climatic Characteristics, but varying skill in spatiotemporal monthly distribution	Almazroui <i>et al.</i> (2021)

Nigeria	ACCESS-ESM1-5 and 12 Others	ACCESS-ESM1-5 with other three has high rank for monthly rainfall simulation	Shiru& Chung, (2021)
East Africa	ACCESS-ESM1-5 and Other 22	ACCESS-ESM1-5 and MIROC6 overestimated seasonal rainfall. The 21 model Out of 23 overestimated over East Africa. But, most of models capture seasonal rainfall Climatology of East Africa.	Omay <i>et al.</i> (2023)
Guatemala (C/America)	GFDL-ESM4, Ec-Earth3 and Other 36	Most models underestimate observed rainfall. IPSL-CM6A-LR preferred as best for monthly rainfall simulation.	Rivera, (2023)
Ethiopia	GFDL-ESM4 and other 36	The performances of models were not consistent simulated in reproducing rainfall distribution at different statistical metrics and time frames. More than 20 model simulated large dry bias over topographic areas during <i>kiremt</i> season(June -September).The GFDL-ESM4 performed best to simulate monthly, <i>kiremt</i> season (June to Septe mber) and <i>belg</i> season (February to May).	Berhanu <i>et al.</i> (2023)

CHAPTER 3. MATERIALS AND METHODS

3.1. Description of study area

3.1.1. Location

The study is conducted in Ethiopia, which is in the horn of Africa. Ethiopia lies between equator and tropic of cancer between 3° and 15° N and 33° and 48° E and covers 1.12 million squares kilometers(Degefu & Bewket, 2023). The study area is bordered by Eritrea to the north and northeast, Sudan and South Sudan to the west and northwest, and Somalia to the southeast (Wubaye et al., 2023).

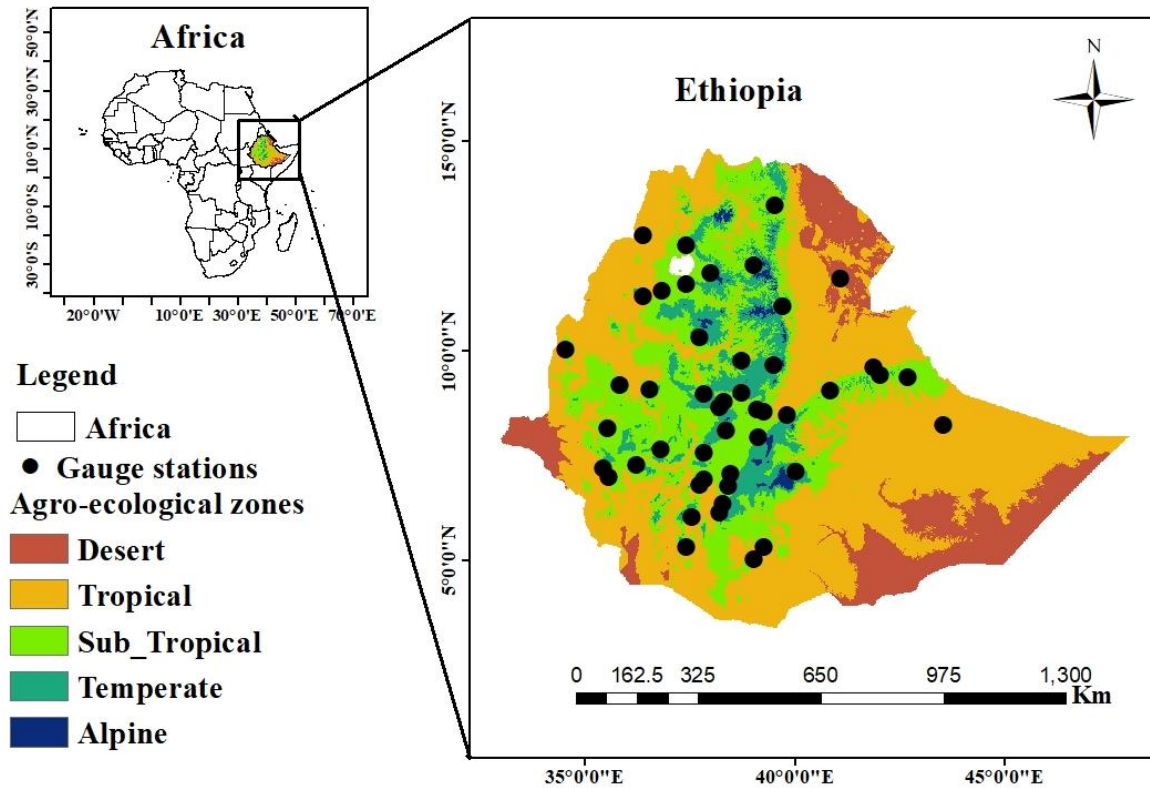


Figure 3. 1The location of Ethiopia from Africa and the meteorological stations used in this study

3.1.2. Topography

Ethiopia is characterized by very complex topography with elevation ranging from 152 meter below sea-level in the north eastern part (Denkel depression) to more than 4600 m.a.s.l at mount Ras Dejen in the northern part of the country (Wubayeet *et al.*, 2023). The country has varied topography including craggy, high mountains, deep gorges with rolling plains, rivers, and plateaus with flat tops (Cheung *et al.*, 2008). About 40% of the country is covered by mountains and plateaus with elevation greater than 1500 m.a.s.l and is classified as highland. The remaining part of the area of less than 1500 m.a.s.l is classified as lowland (Degefu *et al.*, 2022).

3.1.3. Traditional agroecology zone

Ethiopia is characterized by five traditional AEZs(Figure 3.1; Appendix Table 1). These zones are known as *Bereha*/Desert, *Kolla*/Tropical, *WoinaDega*/Sub-tropical, *Dega*/Temperate, and *Wurch*/Alpine. Each AEZ is associated with specific elevations: *Bereha*/Desert encompasses

areas below 500 meters above sea level (m.a.s.l.), Kolla/Tropical ranges from 500 to 1500 m.a.s.l., WoinaDega/Subtropical spans from 1500 to 2300 m.a.s.l., Dega/Temperate covers elevations between 2300 and 3200 m.a.s.l., while Wurch/Alpine represents regions above 3200 m.a.s.l. (MOA, 1998). These AEZs have been distinguished and defined through the examination of altitude, annual rainfall, and temperature patterns in various studies. A brief discussion of the definition and specific characteristics of the traditional AEZs can be found in the work of MOA (1998).

3.1.4. Climate

The complex topographic features have created many local climatic conditions that range from hot deserts in the northeast and southern lowland to very cold mountain ranges in the Semein and Arsi-Bale mountains systems (Degefu *et al.*, 2022). In Ethiopia, gauges which are used to measure rainfall are sparse and uneven. Even those sparse gauges do not have continuous records and lack quality (Dinku *et al.*, 2007). There are different types of gauge stations built in Ethiopia to collect rainfall data. Those are synoptic, principal, ordinary, and rainfall stations. Ethiopia has three distinct seasons, locally known as *bega*, *belg*, and *kiremt*. For the majority of Ethiopia, the short rainy season, known as the *belg* lasts from February to May, while the main rainy season/*kiremt* stays from June to September and the dry season/*bega* lasts from October to January. Geographically speaking, the length of each season on average and the amount of rainfall vary from season to season and place to place (Evangelista *et al.*, 2013).

3.2. Data types and sources

3.2.1. Ground-based (Observed) rainfall data

Daily observed rainfall data from 47 meteorological stations (Figure 3.1; Appendix Table 1) for the period 1995-2014 were obtained from the Ethiopian Meteorology Institute (EMI). The main reason why this study used only 47 stations is due to the absence of other stations which have long data series and less percentage of missing data. In this study, this observed data for 1995–2014 were used to align the climate models from the CMIP6. Of the 47 meteorological stations, 28, 9 and 8 of them are located in WoyynaDega/Sub-tropical, Kolla/tropical and Dega/temperate AEZs, respectively. The remaining two stations are located in Wurch/Alpine and Bereha/Desert AEZs (i.e., one station from each) (Figure 3.1; Appendix Table 1). This study used stations

which have missing values less than 25%, which is reasonable in data scarce regions such as Ethiopia. Due to the absence of many stations which have lower missing values, Gashaw *et al.* (2023) in the Bale Eco-Region, Mohammed *et al.* (2022) in the Upper Blue Nile Basin and Diro *et al.* (2009) over Ethiopia have also used stations that contain up to 20%, 30% and 20% missing data, respectively. Mekasha *et al.* (2014) have also used stations data with up to 25% missing values to study trends and variability of climate in three eco-environments in Ethiopia. The missing values of this study were filled using the Multivariate Imputation by Chained Equations (MICE) package in R software. Then after, the rainfall data have been subjected for quality control in RClimDex package within R software (Zhang *et al.*, 2004) and errors such as negative rainfall were replaced with the nearest station. Outliers were also treated according to the suggestion given by the World Meteorological Organization (WMO, 2009). The details of the data processing and the methods used for filling missing values are given in Wubaye *et al.* (2023). The location and detailed information about the selected meteorological stations can be found in Appendix Table1.

3.2.2 CMIP6 GCMs Model Data

The performance of eight developed by several institutions around the world, are evaluated in this study (Table 3.1). Table 3.1 provides basic information about the eight CMIP6 GCMs, the study was used the outputs of these eight CMIP6 GCMs during 1995-2014 in historical simulations. Data associated with this study have been deposited at <https://esgf-node.llnl.gov/search/cmip6/>, by choosing the features as: Source ID = GCMs' names in Table 3.1, Variable = precipitation, Frequency = day, Experiment ID = amip-historical. The variant labels of each model are used to uniquely distinguish each member of an ensemble based on realizations (r), initialization schemes (i), different physics (p), and forcing (f) indices (Papalexiou *et al.*, 2020).

The GCMs were chosen based on daily data availability and their optimal performance in various regions of Ethiopia (Alaminie *et al.*, 2021), East Africa (Gebrechorkos *et al.*, 2023) and other parts of the world (Iqbal *et al.*, 2021). The models were also chosen since they have been used more widely in various regions of Ethiopia in their previous versions (CMIP5) (Worku *et al.*, 2019). For example, the inclusion of BCC-CSM2-MR the best performance for simulating rainfall in the Upper Blue Nile Basin of Ethiopia (Alaminie *et al.*, 2021) and over Bale-Eco region in Southern

Ethiopia (Gashaw *et al.*, 2024). MRI-ESM2-0,EC-Earth3,and EC-Earth3-Veg were selected since they demonstrated best performance in mainland southeast Asia(Iqbal *et al.*, 2021).

The GCMs rainfall data were extracted for each station to match the GCM data with the observed climate data. In this study, the GCMs data were extracted using the CMhyd tool, which was developed for the extraction and bias correction of GCMs data (Gashaw *et al.*, 2024).

Table 3. 1The list of CMIP6 GCMs employed in this study and their characteristics

Number	CMIP6 model	Institute	Country	Resolution	Variant lable
1	ACCESS-ESM1-5	Australian Community Climate and Earth System Simulator	Australia	1.9°×1.2°	r1i1p1f1
2	BCC-CSM2-MR	Beijing Climate Center (BCC)	China	1.1°×1.1°	r1i1p1f1
3	CNRM-CM6-1	RecherchesMétéorologiques	France	1.4 × 1.4°	r1i1p1f2
4	EC-Earth3	EC-Earth-Consortium	Sweden	0.7° × 0.7°	r1i1p1f1
5	EC-Earth3-veg	EC-Earth-Consortium	Sweden	0.7° × 0.7°	r1i1p1f1
6	MPI-ESM1-2-LR	Max-Planck-Institute für Meteorology	Germany	1.9°×1.9°	r1i1p1f1
7	MRI-ESM2-0	Meteorological Research Institute	Japan	1.1°×1.1°	r1i1p1f1

8	GFDL- ESM4	Geophysical fluid dynamics laboratory	USA	1.25°×1.00°	r1i1p1f1
---	---------------	--	-----	-------------	----------

3.3. Methods of Analysis

3.3.1 Research framework

The overall research framework of this study is shown in Figure 3.2. Performance evaluations of CMIP6 models in this study were undertaken after extracting the CMIP6 GCMs data for the 47 meteorological stations of the study area. Then after, the stations (point) data were changed into areal average. This helps to get a best representative data for the area. The historical data of eight CMIP6 GCMs were evaluated over Ethiopia for the period 1995–2014 over desert, tropical, sub-tropical, temperate, and alpine AEZs separately. The CMIP6 GCMs rainfall data were extracted for each station to match the GCM data with the observed climate data. In this study, the GCMs data were extracted using R-statistical package. The necessary observational climate data cleaning and management task was also implemented using R software. The performance of CMIP6 models under different GCMs were evaluated in simulating historical rainfall over Ethiopia based on efficiency criteria found under “hydroGOF” R package (Zambrano-Bigiarini *et al.*, 2017), then after, the ranks of GCMs in capturing daily, monthly, and annual rainfall total were obtained by aggregating these three metrics using comprehensive rating index (CRI). Details of methods and metrics used in this study are provided as shown in Figure 3.2.

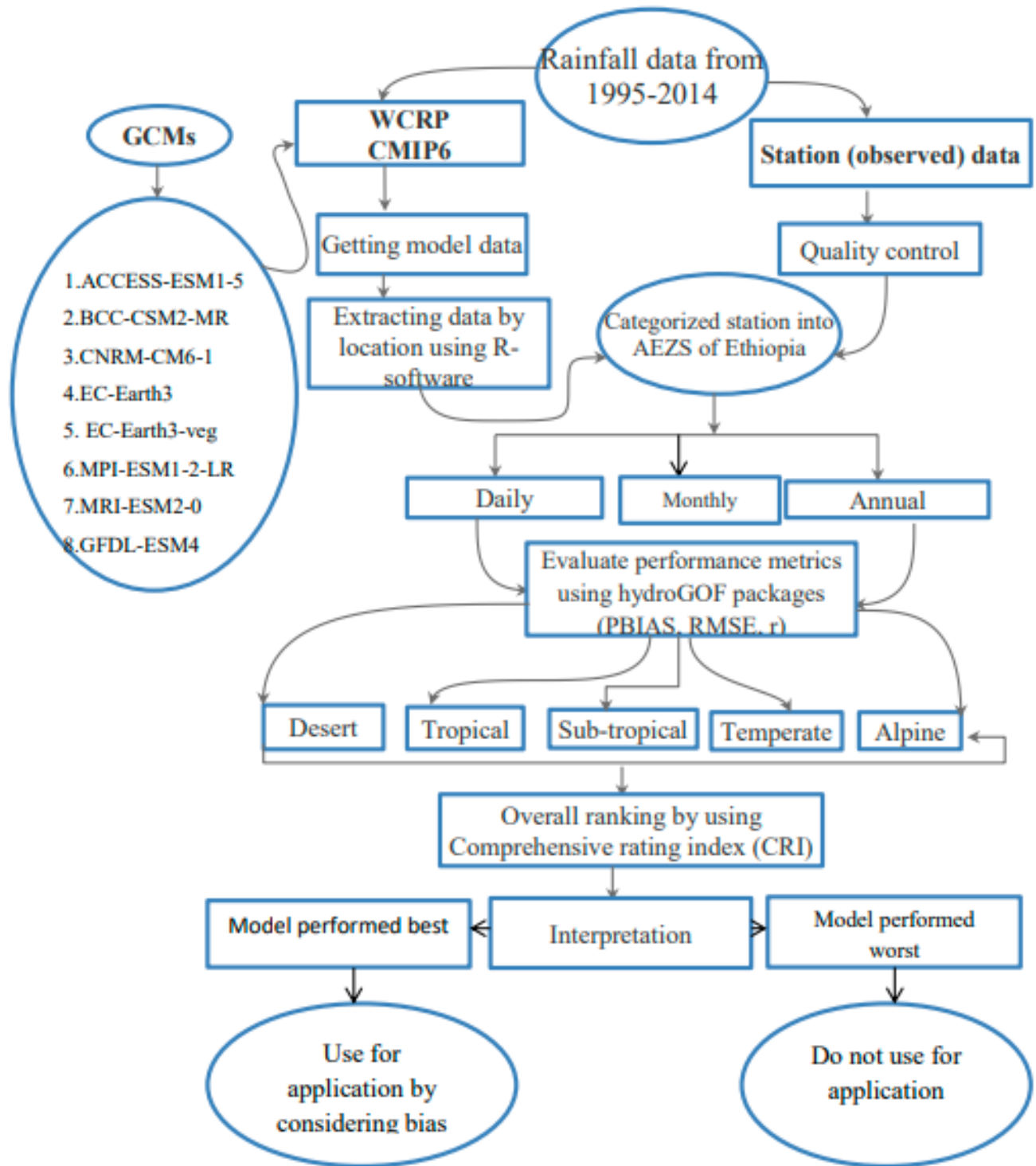


Figure 3. 2Research framework of the study

3.3.2 Performance Metrics for Evaluation of CMIP6 GCMs

Several performance metrics for evaluating the performances of GCMs have been proposed. Statistical metrics are normally used for assessing the differences between model simulations and observed datasets, and each metric can show the overall (relative) performances differently for the model simulations (PGleckler *et al.*, 2008). Additionally, statistical metrics are often found in the comparison of the climate simulation results between GCMs and observed datasets (Kamworapan & Surussavadee, 2019). The eight GCMs were evaluated by three performance metrics for analysis of daily, monthly and annual rainfall data by divided into five AEZs: desert, tropical, sub-tropical, temperate, and alpine. The three-performance metrics (PBIAS, RMSE and r) are used in this study to evaluate the performance of all GCM rainfall data. Comparable to this study, previous studies undertaken in Ethiopia (Worku *et al.*, 2018; Dibaba *et al.*, 2019; Demessie *et al.*, 2023; Gashaw *et al.*, 2024) and elsewhere in the world (Guo *et al.*, 2021) have also used R , RMSE and PBIAS for evaluating performances of climate models. The PBIAS, RMSE and r are calculated to quantify the uncertainties of GCMs to capture the variation of the daily, monthly, and annual time scales. Percent bias (PBIAS) is the models' data tendency to underestimate or overestimate the observed data. A model performance is better when the PBIAS values are closer to zero. A negative PBIAS value is an indication of overestimation, while a positive one indicates underestimation (Gupta *et al.*, 1999). The PBIAS could be calculated using equation (1). Root Mean square error (RMSE) is a statistical metric of the magnitude of the error between the GCM and the observed data. RMSE summarizes errors between the GCM data and reference data (magnitude). The lower RMSE value, the less the error is found from simulation of observed data (Lovino *et al.*, 2018). The RMSE could be calculated using equation (2). Correlation coefficient (r) is employed to assess the relationship degree or to assess the similarity between the spatial patterns in GCM data and in observed data. Between the observed and the modeled climate variables data, -1 and $+1$ correlation coefficient means a perfect negative and positive correlation among the mean of GCM and observational rainfall, respectively. The formula of correlation coefficient (r) is obtained in equation (3).

$$PBIAS = \frac{\sum(G-S)}{\sum(G)} \times 100\% \quad (1)$$

$$RMSE = \sqrt{\frac{\sum(G-S)^2}{N}} \quad (2)$$

$$r = \sqrt{\frac{\sum(G - \bar{G})(S - \bar{S})}{\sum(G - \bar{G})^2 \sum(S - \bar{S})^2}} \quad (3)$$

Where G is observed rainfall, $\bar{G}$ is the average observed rainfall, S is the CMIP6 GCMs derived rainfall, $\bar{S}$ is the average of CMIP6 GCMs of derived rainfall, N is the number of data pairs compared, r is Pearson's correlation coefficient. When multiplied by 100, indicates the degree of overestimation or underestimation as a percentage.

Table 3. 2Summary of continuous performance metrics used in this study

Continuous metrics	Statistical Formula	Value range	Unit	Best value	Reference
RMSE	$RMSE = \sqrt{\frac{\sum(G - S)^2}{N}}$	0 to ∞	mm	0	Ageet <i>et al.</i> (2022)
PBIAS	$PBIAS = \frac{\sum(G - S)}{\sum(G)} \times 100\%$	$-\infty$ to ∞	%	0	(Dangol <i>et al.</i> (2022))
R	$r = \sqrt{\frac{\sum(G - \bar{G})(S - \bar{S})}{\sum(G - \bar{G})^2 \sum(S - \bar{S})^2}}$	-1 to 1	none	1	(Dinku <i>et al.</i> (2007))

3.3.3. Comprehensive Rating Index (CRI)

The CRI is an effective ranking model that has been widely used to rank the performance of GCMs (Zhang *et al.*, 2018). Here, the CRI is used in two ranking processes: Firstly, it is used to rank the ability of GCMs to capture the daily, monthly, and annual temporal scale by combining three indicators (PBIAS, RMSE and r) separately. Finally, the overall performances of each GCM over different AEZs were ranked based on the CRI. The formula of CRI is given in equation 4.

$$CRI = 1 - \frac{1}{nm} \sum_{i=1}^n RANK_i \quad (4)$$

Where n is the number of statistical metrics used to evaluate model performance, m is the total number of GCMs. The closer the value of CRI to 1, the better the model performs. The rank value of the model with the best performance is assigned as 1.

3.3.4. The performance evaluation of CMIP6 GCMs using graphical techniques

Graphical techniques provide a visual comparison of simulated and measured data and is considered a first overview of model performance (ASCE, 1993). According to Legates & McCabe (1999), graphical techniques are essential to appropriate model evaluation. The two commonly used graphical techniques are hydrograph and percent Exceedance probability curves, are especially valuable. Other graphical techniques like bar graphs, box plots, line graphs and others can also be used to examine the variations and data distributions with respect to reference (measured) dataset. For instance, a study by Dangol *et al.* (2022) , used box plots, bar graphs , and Taylor diagrams to evaluate the performance of rainfall products. On the other hand, graphical techniques like line graphs ,and bar graphs were used to evaluate the performance of rainfall products over Bale Eco-region in Southern Ethiopia by Gashaw *et al.* (2023). Moreover, a study by Polong *et al.* (2023) , also used line graph and bar graphs to compare rainfall products with observed measured data with different temporal and spatial scales. In this study line graphs were used to compare mean observed rainfall data with CMIP6 GCMs data over AEZs of Ethiopia at daily, monthly, and annual temporal scales. The closer the models with observed line graph, is considered to be best.

CHAPTER 4. RESULTS AND DISCUSSIONS

4.1. Performance Evaluation CMIP6 GCMs to simulate daily rainfall total

The performance metrics for rainfall at daily time scale were presented for all GCMs and the ranking of GCMs using CRI for all AEZs for the time period from 1995-2014 are presented in Table 4.1. The finding indicated that MRI-ESM2-0 is the best performing climate model at daily timescale over desert AEZ. Followed by MRI-ESM2-0, Ec-Earth-veg and BCC-CSM-2-MR are the second and third ranking CMIP6-GCM models respectively. The poor performed GCMs are CNRM-CM6-1 and GFDL-ESM4 to simulate daily rainfall over desert AEZ. Moreover, most of the PBIAS values of CMIP6 GCMs to simulate daily rainfall over desert AEZ are positive value, which implied that underestimate the daily rainfall total over desert AEZ of Ethiopia. The GFDL-ESM4 models highly overestimate the daily rainfall total by 128.8% over desert AEZ.

Over tropical AEZ, Ec-Earth3-veg, MPI-ESM-1-2-LR, and ACCESS-ESM-1-5 are the first, second, and third ranked performed climate models. On the other hand, BCC-CSM-2-MR and CNRM-CM6-1 are the worst performed CMIP6 GCMs for daily rainfall simulation over tropical AEZ. In this AEZ, four climate models overestimate and four of them underestimate the daily rainfall total. The best performed climate model (Ec-Earth3-veg) underestimate the daily rainfall total for the period 1995-2014 by 11.4%.

Over sub-tropical AEZ, the first, second, and third ranked best performed climate models at daily time scale are Ec-Earth3-veg, MPI-ESM-1-2-LR, and MRI-ESM2-0 respectively. On the other side, BCC-CSM-2-MR and CNRM-CM6-1 are ranked the least performed CMIP6 GCMs model to simulate daily rainfall total. Even though, Ec-Earth3-veg are the best performed model, it underestimates the daily rainfall total by 1.6% (Table 4.1).

Over temperate AEZ, the highest ranked CMIP6 GCMs to simulate daily rainfall total is Ec-Earth3-veg. ACCESS-ESM-1-5 ranked as the least performed climate model to capture daily rainfall total. Most of models considered in the study have overestimation problem over temperate AEZ. Ec-Earth3-veg overestimate daily rainfall total by 9.2% (negative PBIAS) the daily rainfall total for the period 1995-2014 (Table 4.1).

Over alpine AEZ, most of the models underestimate the daily rainfall total. But MRI-ESM2-0, GFDL-ESM4, and Ec-Earth3-veg overestimate the daily rainfall total. Based on CRI value, MPI-ESM-1-2-LR was the highest ranked performed climate model to capture daily rainfall total over alpine AEZ. Whereas, GFDL-ESM4, Ec-Earth3-veg, and CNRM-CM6-1 are the least ranked models. MPI-ESM-1-2-LR (best performed model) underestimates by 24.60% the daily rainfall total.

In every analysis, no single model performed better than all others. Due to this, this finding showed that CRI could evaluate the CMIP6 GCMs uniformly. Overall, over tropical, sub-tropical and temperate AEZs, Ec-Earth3-veg are the best ranked performed climate model to simulate daily rainfall total. But, MRI-ESM2-0 over desert AEZ and MPI-ESM-1-2-LR over alpine AEZ outperformed to detect daily rainfall total. Study by Gashaw et al. (2024), in their conclusion noted that CNRM-CM6-1 and GFDL-ESM4 are ranked as the best and worst performing modes for simulation of daily rainfall total over Bale Eco region, southern Ethiopia. Even though, Ec-Earth3-veg is the most performed model over most AEZs of the study area, Gashaw *et al.* (2024), noted neither best nor worst over their study area. A study by Lebeza *et al.* (2024) also noted that, the superiority of Ec-Earth3, MPI-ESM-1-2-LR, and ACCESS-ESM1-5 for daily, CNRM-CM6-1, BCC-CSM-2-MR, and Ec-Earth3 for monthly, BCC-CSM-2-MR, MPI-ESM-1-2-LR, and Ec-Earth3 for simulation of annual rainfall total. In addition, most of CMIP6 GCMs models have underestimation problem to detect daily rainfall total over most AEZs of Ethiopia. The large bias even in the best performing climate model gives alarm the necessity of bias correction of climate models before using for further application. Most of the CMIP6 models showed lower performance over high topographic (over temperate and alpine AEZs of Ethiopia), while they performed well in low topographic (over desert and tropical AEZs of Ethiopia) areas of Ethiopia (Table 4.1). This indicates that the effect of high topography on precipitation is still a challenge in climate modeling Berhanu *et al.* (2023). So here, models have to be evaluated for Ethiopia's highland areas exclusively.

Table 4. 1 Performance of CMIP6 GCMs for simulating daily rainfall over AEZs of Ethiopia

AEZs	CMIP6 Models	Continuous statistical metrics			CRI	Rank
		RMSE	PBIAS	r		

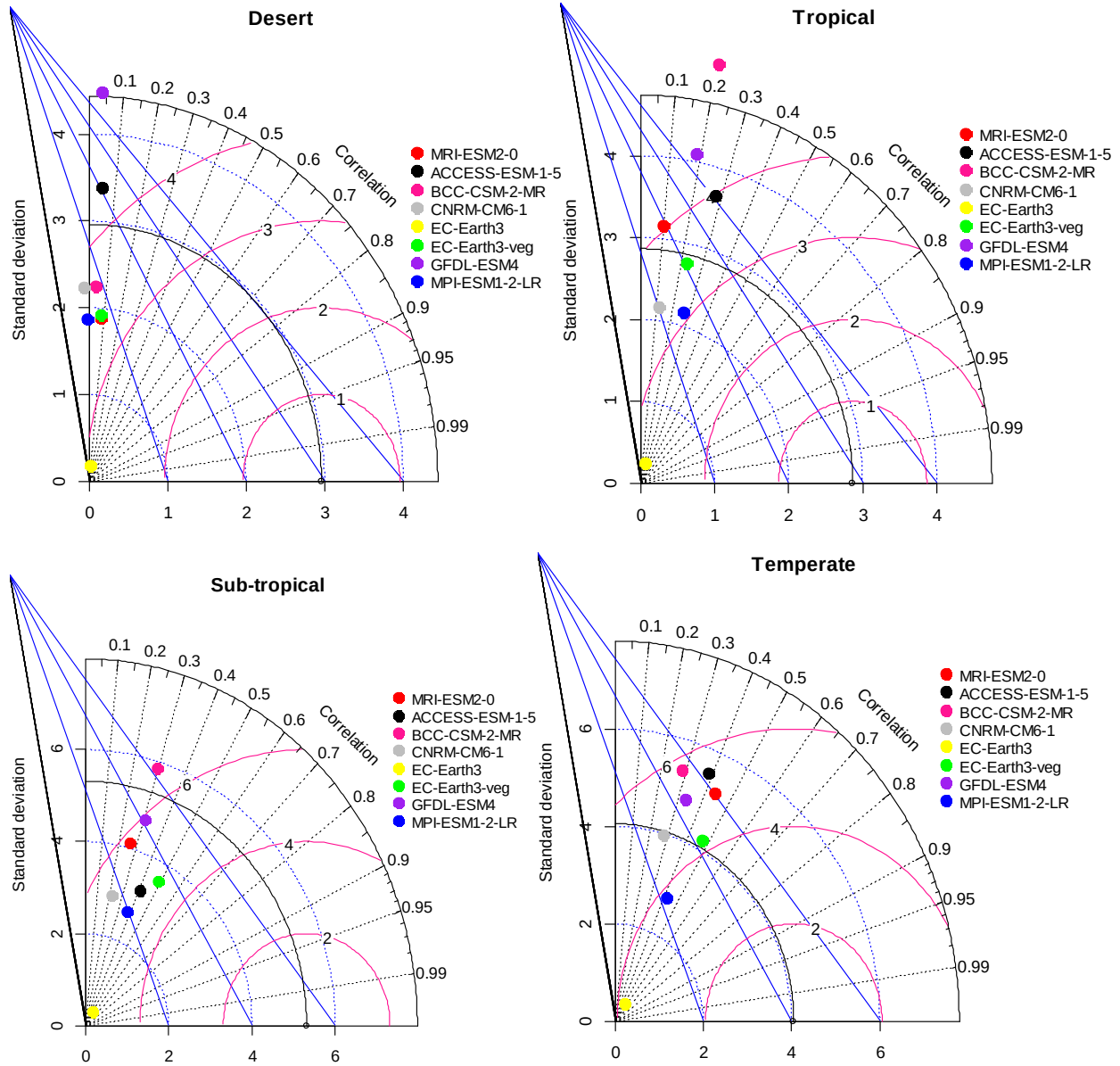
Desert	ACCESS-ESM-1-5	4.49	64.10	0.05	0.42	5
	BCC-CSM-2-MR	4.00	-16.40	0.04	0.58	3
	CNRM-CM6-1	4.07	48.50	0.00	0.21	7
	EC-Earth3	3.43	91.90	0.03	0.46	4
	EC-Earth3-veg	3.82	11.10	0.04	0.71	2
	GFDL-ESM4	5.63	-128.80	0.01	0.08	8
	MPI-ESM1-2-LR	3.86	47.90	0.01	0.42	5
	MRI-ESM2-0	3.81	3.00	0.04	0.79	1
Tropical	ACCESS-ESM-1-5	4.06	-30.80	0.28	0.54	3
	BCC-CSM-2-MR	5.49	-30.30	0.20	0.29	7
	CNRM-CM6-1	3.76	58.40	0.12	0.29	7
	EC-Earth3	3.80	91.60	0.28	0.46	4
	EC-Earth3-veg	3.51	11.40	0.23	0.63	1
	GFDL-ESM4	4.55	-10.50	0.19	0.38	6
	MPI-ESM1-2-LR	3.28	38.40	0.27	0.58	2
	MRI-ESM2-0	4.06	-6.20	0.10	0.42	5
Sub-tropical	ACCESS-ESM-1-5	5.34	37.10	0.41	0.29	6
	BCC-CSM-2-MR	5.74	-19.80	0.32	0.25	7
	CNRM-CM6-1	4.20	46.40	0.19	0.25	7
	EC-Earth3	4.31	90.70	0.55	0.42	4
	EC-Earth3-veg	3.45	1.60	0.49	0.79	1
	GFDL-ESM4	4.69	-3.50	0.35	0.42	4
	MPI-ESM1-2-LR	3.26	32.10	0.47	0.67	2

	MRI-ESM2-0	4.27	-32.10	0.39	0.46	3
Temperate	ACCESS-ESM-1-5	5.66	-50.10	0.38	0.21	8
	BCC-CSM-2-MR	5.73	1.00	0.28	0.33	7
	CNRM-CM6-1	4.85	-17.30	0.28	0.38	5
	EC-Earth3	4.77	89.60	0.53	0.50	3
	EC-Earth3-veg	4.26	-9.20	0.47	0.75	1
	GFDL-ESM4	5.17	-3.90	0.33	0.38	5
	MPI-ESM1-2-LR	3.96	30.10	0.42	0.58	2
	MRI-ESM2-0	5.13	-35.00	0.44	0.42	4
Alpine	ACCESS-ESM-1-5	7.29	23.30	0.28	0.54	2
	BCC-CSM-2-MR	7.76	9.80	0.18	0.42	4
	CNRM-CM6-1	7.51	36.60	0.07	0.29	6
	EC-Earth3	6.56	86.00	0.31	0.50	3
	EC-Earth3-veg	7.67	-48.30	0.29	0.29	6
	GFDL-ESM4	7.81	-15.60	0.17	0.29	6
	MPI-ESM1-2-LR	6.28	24.60	0.33	0.75	1
	MRI-ESM2-0	7.57	-39.00	0.30	0.42	4

Note: that Rank is given based on CRI, and 1st and 8th rank indicates the best and poorest performing models, respectively.

Taylor diagram(Taylor, 2001), has the ability to give a good statistical summary of correlation, standard deviation, and Centered root mean square between the modeled and the observed data in a single diagram. Figure 4.1 showed the performance of daily rainfall data GCMs relative to their observed data using Taylor diagram. The standard deviations, correlations are all different between the modeled and observed over all AEZs of Ethiopia. The highest ranking GCMs were seen to have higher correlations. The individual standard deviations indicate different values in the models, and ACCESS-ESM-1-5, Ec-Earth3-veg, BCC-CSM-2-MR, CNRM-CM6-1, and

BCC-CSM-2-MR over desert,tropical,sub-tropical,temperate,and alpine were close to observed standard deviation respectively. In line with my finding(Rivera, 2023), also indicated the individual standard deviation of the models showed different values, but after bias correction the models close to reference.



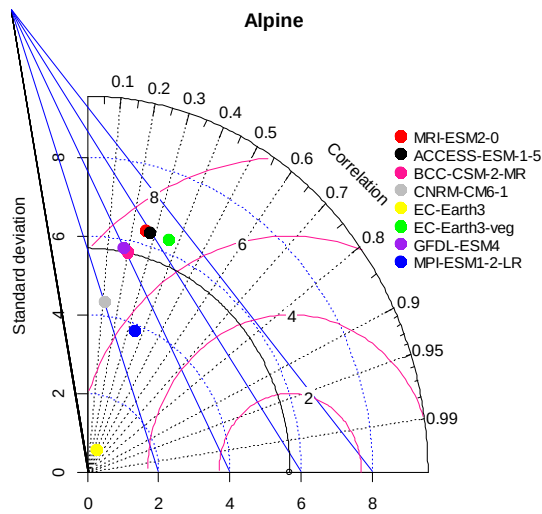


Figure 4.1: Taylor diagram representation of CMIP6 GCMs models daily rainfall against observed daily rainfall over AEZs of Ethiopia for the period 1995-2024. The black line with dot indicates correlation coefficient and standard deviation of observed data.

4.2. CMIP6 GCMs detection capability of Monthly rainfall total

Table 4.2 displayed the performance evaluation metrics and their rank at monthly time scale over five AEZs of Ethiopia. The result revealed that, Ec-Earth3-veg and MRI-ESM2-0 have comparable each other and the best ranked CMIP6 GCMs for monthly rainfall simulation over desert AEZ (Table 4.2). According to performance ranking based on CRI value, ACCESS-ESM1-5 has worst performance to detect monthly rainfall total. The two best performed models that means Ec-Earth3-veg and MRI-ESM2-0 slightly underestimate monthly rainfall total by 11.1% and 3.0% respectively. The least performed climate model (ACCESS-ESM1-5) overestimate monthly rainfall total by 178.4%.

Over tropical AEZ, the first, second, and third ranked CMIP6 GCMs models to simulate monthly rainfall total are BCC-CSM-2-MR, MPI-ESM1-2-LR, and GFDL-ESM4 respectively. On the other hand, CNRM-CM6-1 and Ec-Earth3 performed least to detect monthly rainfall total. The PBIAS value of BCC-CSM-2-MR, MPI-ESM1-2-LR, and GFDL-ESM4 are -30.30% (overestimate), 38.30% (underestimate), and -10.80% (overestimate) of monthly rainfall total (Table 4.2).

Over sub-tropical AEZ, the best performed climate model is BCC-CSM-2-MR. Followed by BCC-CSM-2-MR, Ec-Earth3-veg and MPI-ESM1-2-LR have better performance and

comparable each other. ACCESS-ESM-1-5 has the least ranked performance to simulate monthly rainfall total. Even though, BCC-CSM-2-MR ranked as the best performed model, it overestimates the monthly rainfall total by 19.80% over sub-tropical AEZ (Table 4.2).

Over temperate AEZ, the best performed ranking models are BCC-CSM-2-MR, MPI-ESM-1-2-LR, and GFDL-ESM4 with rank of first, second, and third respectively. The ACCESS-ESM-1-5 has poor performance to capture monthly rainfall total. The BCC-CSM-2-MR underestimate the monthly rainfall total only by 0.9%. But the least performed climate model (ACCESS-ESM-1-5) overestimate by 50.10% of the monthly rainfall total. Some climate models underestimate and some of them overestimate the monthly rainfall total. Generally, the considered CMIP6 GCMs in this study have PBIAS value of -50.10% to 89.60% (Table 4.2).

Over alpine AEZ, ACCESS-ESM-1-5 has best overall performance and comparable with BCC-CSM-2-MR. The two models ranked as the top among the studied CMIP6 GCMs models. Conversely, MRI-ESM2-0 and EC-Earth3-veg have poor performance. ACCESS-ESM-1-5 and BCC-CSM-2-MR (best performed models) overestimate by 18.60% and 18.10% respectively. The PBIAS of studied climate models ranged from -5.0% (GFDL-ESM4) to 87.30% (EC-Earth3).

Generally, over tropical, subtropical, temperate, and alpine AEZs, BCC-CSM-2-MR are the best ranked CMIP6 GCMs model to detect monthly rainfall total. ACCESS-ESM-1-5 has comparable performance with BCC-CSM-2-MR over alpine AEZ. On the other hand, MRI-ESM2-0 has the top ranked climate model over desert AEZ. In lined with my study, Alaminie et al. (2021), also noted the best performance model for monthly rainfall total simulation of BCC-CSM-2-MR. A study by Gashaw et al. (2024) also stated that, BCC-CSM2-MR, CNRM-CM6-1, and MRI-ESM2-0 as the best performing climate models to detect monthly rainfall total over Bale-eco region in Southern Ethiopia. In general, the considered studied climate models have overestimation and underestimation problem to capture monthly rainfall total (Table 4.2). My finding is aligned with study by Berhanu *et al.* (2023), those stated that, most of the models overestimate the reference data. Berhanu *et al.* (2023), also stated that the performance of models is not consistent in reproducing rainfall distributions at different statistical metrics and timeframes.

Table 4. 2Performance of CMIP6 GCMs for simulating monthly rainfall over AEZs of Ethiopia

AEZs	CMIP6 Models	Continuous statistical metrics			CRI	Rank
		RMSE	PBIAS	r		
Desert	ACCESS-ESM-1-5	55.33	-178.40	0.40	0.13	8
	BCC-CSM-2-MR	31.44	-16.40	0.05	0.46	4
	CNRM-CM6-1	38.83	48.50	-0.04	0.21	7
	EC-Earth3	30.08	91.80	0.37	0.63	3
	EC-Earth3-veg	29.98	11.10	0.33	0.75	1
	GFDL-ESM4	50.78	-128.60	0.31	0.25	6
	MPI-ESM1-2-LR	37.11	48.00	0.00	0.33	5
	MRI-ESM2-0	30.18	3.00	0.36	0.75	1
Tropical	ACCESS-ESM-1-5	73.12	-30.80	0.63	0.33	6
	BCC-CSM-2-MR	53.03	-30.30	0.74	0.71	1
	CNRM-CM6-1	70.16	58.40	0.36	0.21	7
	EC-Earth3	89.53	91.60	0.65	0.21	7
	EC-Earth3-veg	57.18	11.40	0.55	0.50	4
	GFDL-ESM4	57.26	-10.80	0.61	0.54	3
	MPI-ESM1-2-LR	48.20	38.30	0.70	0.63	2
	MRI-ESM2-0	70.76	-6.20	0.23	0.38	5
Sub-tropical	ACCESS-ESM-1-5	116.04	-58.90	0.67	0.08	8
	BCC-CSM-2-MR	52.14	-19.80	0.85	0.79	1
	CNRM-CM6-1	85.53	46.40	0.42	0.17	7
	EC-Earth3	115.35	90.70	0.85	0.33	6
	EC-Earth3-veg	60.61	1.60	0.77	0.63	2

	GFDL-ESM4	58.01	-3.50	0.75	0.58	4
	MPI-ESM1-2-LR	52.91	32.10	0.84	0.63	2
	MRI-ESM2-0	68.50	-32.10	0.73	0.38	5
Temperate	ACCESS-ESM-1-5	108.05	-50.10	0.69	0.13	8
	BCC-CSM-2-MR	46.67	0.90	0.85	0.83	1
	CNRM-CM6-1	74.66	17.30	0.61	0.29	6
	EC-Earth3	116.44	89.60	0.89	0.29	6
	EC-Earth3-veg	67.02	-9.30	0.79	0.50	4
	GFDL-ESM4	59.35	-3.90	0.77	0.54	3
	MPI-ESM1-2-LR	58.52	30.00	0.82	0.58	2
	MRI-ESM2-0	75.12	-35.00	0.81	0.33	5
Alpine	ACCESS-ESM-1-5	150.66	-18.60	0.33	0.67	1
	BCC-CSM-2-MR	144.58	18.10	0.14	0.67	1
	CNRM-CM6-1	145.42	42.30	0.13	0.38	5
	EC-Earth3	152.26	87.30	0.26	0.38	5
	EC-Earth3-veg	172.89	-34.90	0.22	0.33	7
	GFDL-ESM4	154.37	-5.00	0.04	0.42	4
	MPI-ESM1-2-LR	147.56	31.40	0.15	0.50	3
	MRI-ESM2-0	182.12	-26.50	0.03	0.17	8

Note that Rank is given based on CRI, and 1st and 8th rank indicates the best and poorest performing models, respectively.

Figure 4.2 showed the mean monthly rainfall of observed and CMIP6 GCMs over AEZs of Ethiopia. Analogously, the CMIP6 GCMs models those are relatively good in representing the long term mean rainfall (1995-2014) of wet months (June, July, August, and September) are Ec-Earth3-veg and GFDL-ESM4 over desert AEZ. Whereas, MPI-ESM1-2-LR and CNRM-CM6-1 are relatively better to detect mean monthly rainfall of the dry months (October, November, December, and January) over desert AEZ. ACCESS-ESM1-5 and GFDL-

ESM4 poorly detect mean wet and dry months respectively. Over tropical AEZ, GFDL-ESM4 better closer with observed to detect mean monthly rainfall. In addition to this, Ec-Earth3-veg is also has a good detection capability of mean monthly rainfall, except April and May months (Figure 4.1). On the other hand, Ec-Earth3 is the worst model to detect mean monthly observed rainfall for the period 1995-2014 (Figure 4.2).

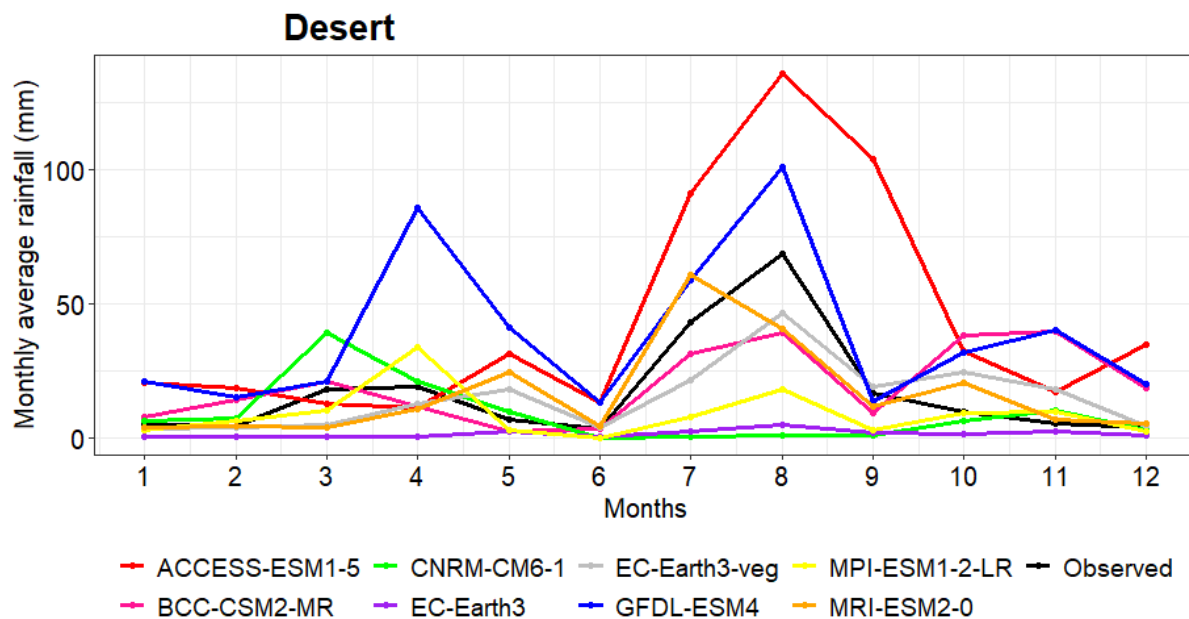
Over sub-tropical AEZ, GFDL-ESM4, Ec-Earth3-veg, and BCC-CSM2-MR climate models are relatively good to capture mean monthly rainfall. But CNRM-CM6-1, Ec-Earth3, and ACCESS-ESM1-5 are not good to simulate mean monthly rainfall, especially wet months (June, July, August, and September) (Figure 4.2).

Over temperate AEZ, the line graph of GFDL-ESM4, Ec-Earth3-veg, and BCC-CSM2-MR climate models are closer with observed mean monthly rainfall. But Ec-Earth3 and ACCESS-ESM1-5 are far away from observed. As the graph depicted, Ec-Earth3 totally underestimates mean monthly rainfall over temperate AEZ (Figure 4.2).

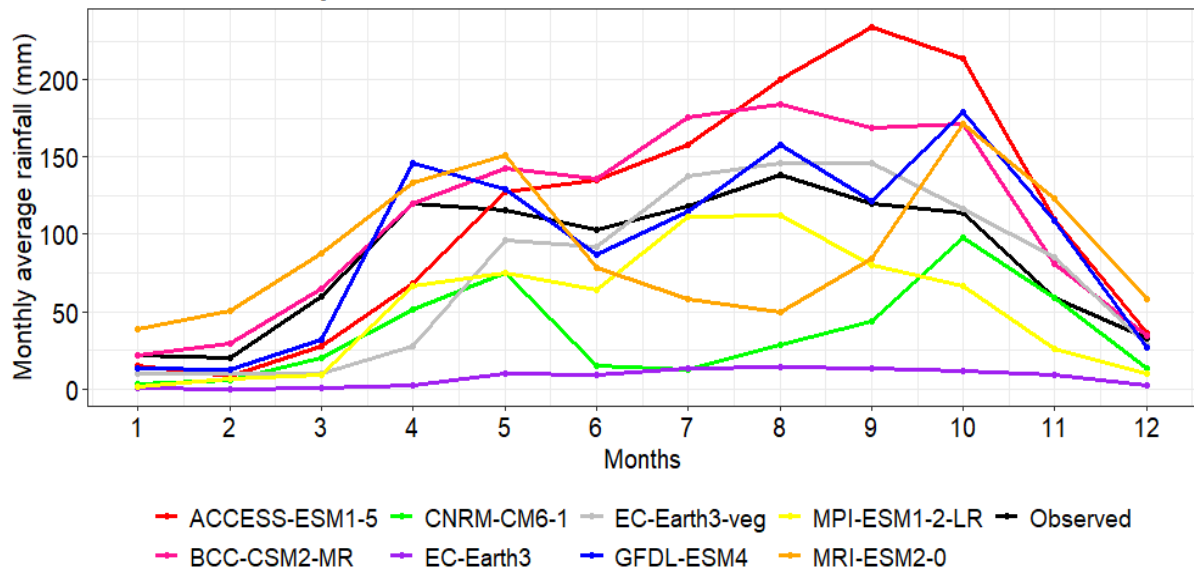
Over alpine AEZ, Ec-Earth3-veg and MRI-CSM2-0 are good climate models to capture mean monthly rainfall of June and July. Whereas, GFDL-ESM4 are relatively best model to simulate mean monthly rainfall of dry months (October, November, December, and January) (Figure 4.2). GFDL-ESM4 overestimates mean monthly rainfall of dry months, but underestimates wet months (June, July, August, and September). Omay *et al.* (2023) stated that, most of models overestimate April to September rainfall over Upper Blue Nile, which is North Wet of Ethiopia. Ec-Earth3 is the worst climate model to detect mean monthly rainfall of the twelve months. It underestimates mean monthly rainfall for all months of the year.

This study revealed diverse outputs of models simulations capability and their performance over five AEZs of Ethiopia. Generally, Ec-Earth3-veg and GFDL-ESM4 are the best performed climate models over all AEZs of Ethiopia to detect mean monthly rainfall. In addition, BCC-CSM2-MR has comparable performance with GFDL-ESM4 and Ec-Earth3-veg over sub-tropical and temperate AEZs. On the other hand, Ec-Earth3 is the worst climate model to detect mean monthly rainfall over five AEZs of Ethiopia. Next to Ec-Earth3, ACCESS-ESM1-5 is the least performed CMIP6 GCMs to detect mean monthly rainfall. Monthly rainfall total for most models shows increasing from January to August months. Omay *et al.* (2023) also noted, the increase in

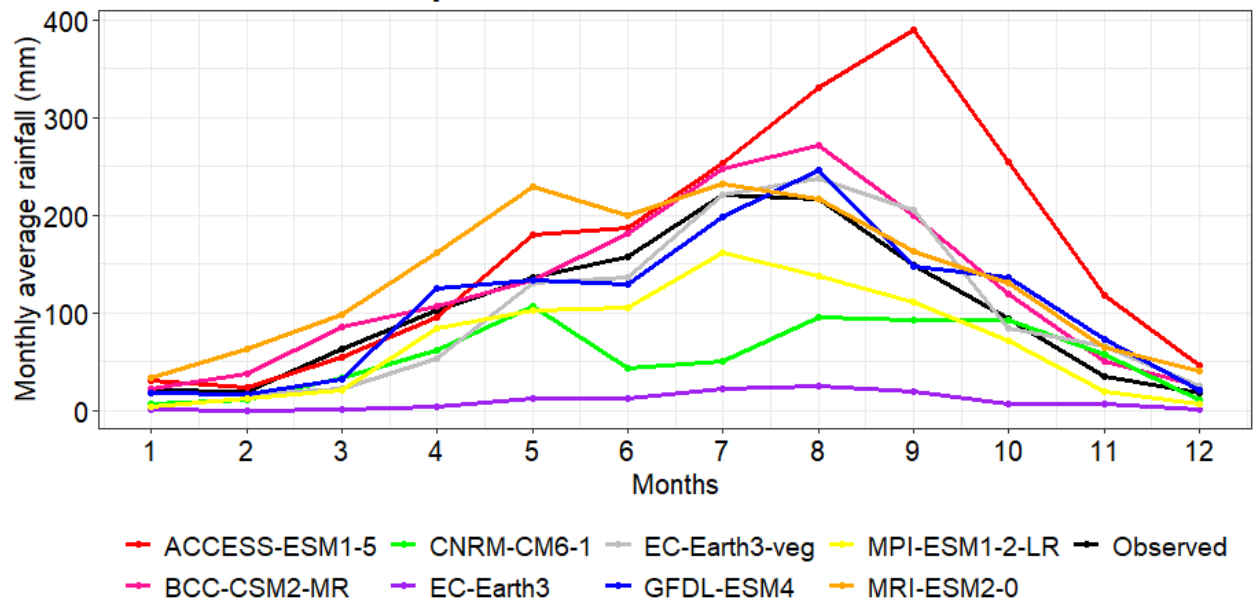
monthly rainfall from April to May and from June to August, due to the north wards migration of Inter Tropical Convergence Zone as the main drivers of simulated and projected rainfall over Ethiopia. That Ec-Earth3 climate model totally underestimates mean monthly rainfall over all AEZs of Ethiopia for twelve months of the year (Figure 4.2). Most models considered in this study cannot simulate peak rainfall of July and August months over all AEZs of Ethiopia. It is in line with Berhanu et al.(2023),those study performance of CMIP6 GCMs in Ethiopia. Best performing models, overestimated in some of the months and underestimated in the remaining months. Gashaw *et al.* (2024) also find similar generalization in their study area over Bale-Eco region, Southern Ethiopia. Models have significant regional variation, which could be caused by errors, uncertainties, and topographic variations(Mohammed, 2024).



Tropical



Sub-tropical



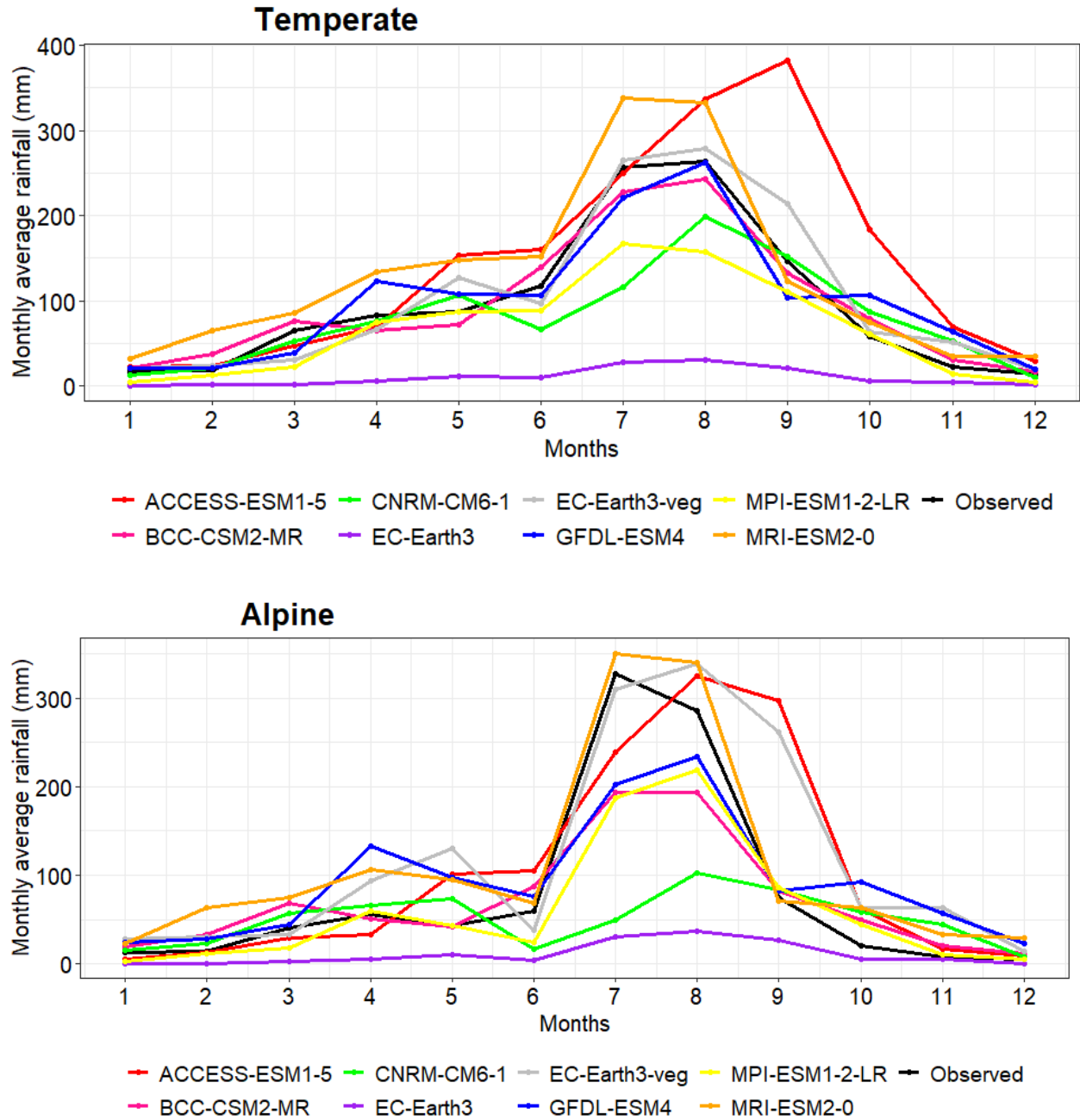
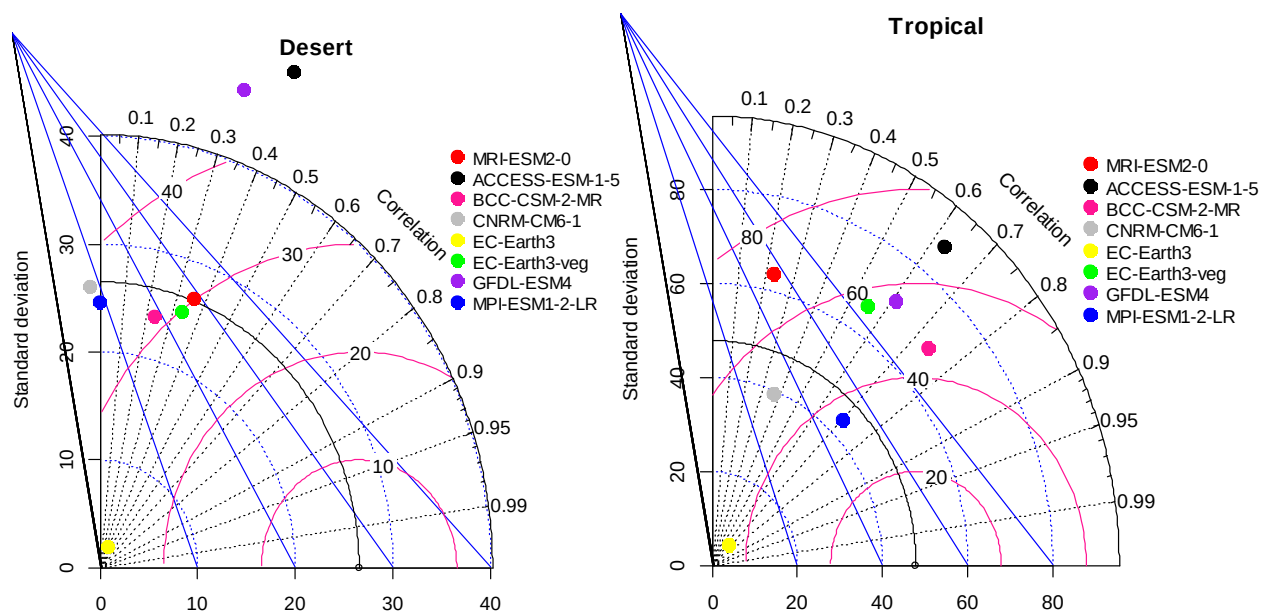


Figure 4. 2 Line graph of mean monthly rainfall, (observed with CMIP6 GCMs data from 1995 to 2014)

Figure 4.3 showed Taylor diagram representation of CMIP6 GCMs monthly rainfall data with observed rainfall data. The result depicted a large and scattered CMIP6 GCMs distribution for monthly rainfall. The CMIP6 GCMs those have maximum correlation with observed were ACCESS-ESM1-5 (0.40 over desert), BCC-CSM2-MR (0.73 over tropical), BCC-CSM2-MR and Ec-Earth3 (0.85 over Sub-tropical), Ec-Earth3 (0.89 over temperate), and ACCESS-ESM1-5 (0.33 over alpine) AEZs. Hence, the monthly rainfall predicted by CMIP6 GCMs had different correlation intervals across the models and AEZs. In addition, standard deviations, for monthly rainfall CMIP6 GCMs deviated more from the observed monthly rainfall. The monthly rainfall totals have higher standard deviation over alpine AEZ than others, indicating higher frequency of extreme rainfall events. However, smaller standard deviation of monthly rainfall was observed over desert AEZ, indicating a decreased probability of extreme monthly rainfall total (Figure 4.3).



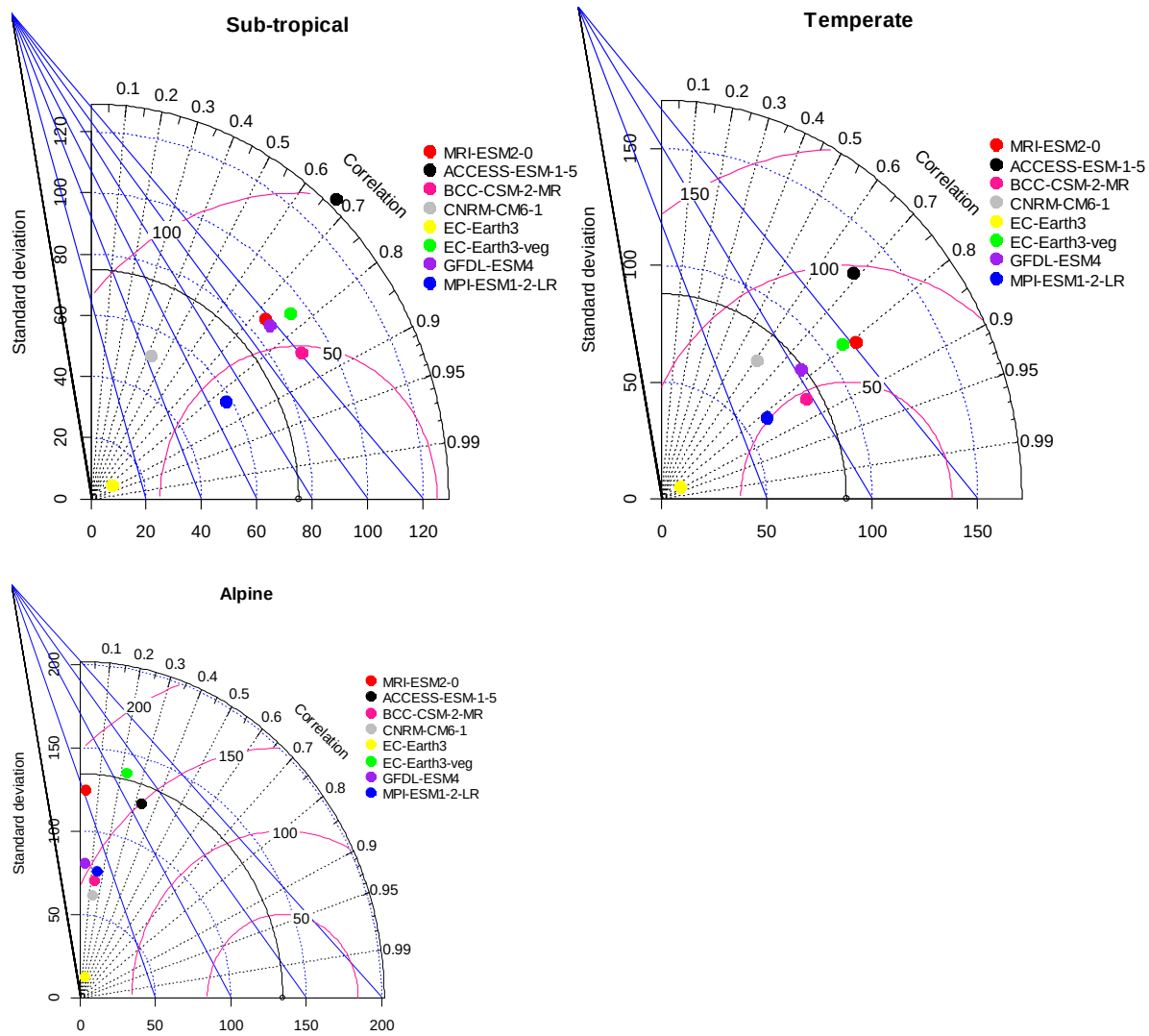


Figure 4.3: Taylor diagram representation of CMIP6 GCMs models monthly rainfall against observed monthly rainfall over AEZs of Ethiopia for the period 1995-2024. The black line with dot indicates correlation coefficient and standard deviation of observed data.

4.3. Performance Evaluation CMIP6 GCMs to simulate annual rainfall total

Table 4.3 described the performance statistical metrics for annual rainfall across AEZs of Ethiopia calculated for all CMIP6 GCMs and ranked them according to their performance using CRI method. Over desert AEZ, MRI-ESM2-0, BCC-CSM-2-MR, and CNRM-CM6-1 are ranked the first, second, and third best performed climate models to capture annual rainfall total. ACCESS-ESM-1-5 is the least performed model. Most of climate models over desert AEZ, seems have underestimation problem to detect annual rainfall total. The PBIAS value of best

performed (MRI-ESM2-0) is 3.10%, which means underestimate annual rainfall total y 3.10% (Table 4.3).

Over tropical AEZ, MRI-ESM2-0 is ranked as the best performed climate model to capture annual rainfall total. It has slight overestimation problem of the observed rainfall for the period 1995-2014. MRI-ESM2-0 overestimates annual rainfall total by 6.20%. On the other side, MPI-ESM1-2-LR has least performance with correlation coefficient of -0.36 and PBIAS of 38.30% (underestimate by 38.30%) (Table 4.3).

Over sub-tropical AEZ, Ec-Earth3-veg, GFDL-ESM4, and BCC-CSM-2-MR are the first, second, and third ranked best performed climate models to simulate annual rainfall total. The best performed CMIP6 GCMs model (Ec-Earth3-veg) has a slight underestimation problem by 1.60%. The lowest ranked climate model based on CRI is Ec-Earth3, which has underestimation problem of annual rainfall total. The Ec-Earth3 underestimates annual rainfall of Sub-tropical AEZ by 90.70%.

Over temperate AEZ, BCC-CSM-2-MR ranked the best performed CMIP6 GCMs to capture annual rainfall total. It underestimates the annual rainfall total by only 0.90%, but the Pearson correlation coefficient is 0.23 (which is weak correlation). On the other hand, Ec-Earth3 climate model ranked the least performance. It has high underestimation problem, which underestimates annual rainfall total by 89.60%. Generally, the considered climate models have both overestimation and underestimation problem. The PBIAS of climate models over temperate AEZ ranges from -50.10% to 89.60%.

Over alpine AEZ, BCC-CSM-2-MR ranked the best performed climate model to detect annual rainfall total. The model slightly underestimates by 9.90%. Surprisingly, next to BCC-CSM-2-MR, ACCESS-ESM-1-5 performed in the ranked as second, which is worst performance over the other four AEZs of Ethiopia. ACCESS-ESM-1-5 overestimates annual rainfall total by 30.40%. Ec-Earth3 ranked the worst performed CMIP6 GCMs climate model to detect annual rainfall total. The model underestimates the observed rainfall of 1995-2014 by 86.0%.

Generally, over desert and tropical AEZs, MRI-ESM2-0 outperformed other climate models to detect annual rainfall. Whereas, BCC-CSM-2-MR ranked the best climate model to simulate annual rainfall total over temperate and alpine AEZs. Over sub-tropical AEZ, Ec-Earth3-veg is

preferable for simulation of annual rainfall total. Overall, the considered CMIP6 GCMs models in this study have both overestimation and underestimation problem over all AEZs of Ethiopia. There is a clear difference between models performance. In line with this finding, Mohammed(2024)also conclude similar statement that, models have clear performance difference. Models have different performance from place to place and time to time. The result from this finding confirms the suggestion given by Gashaw *et al.*(2024),who recommended that in order to use climate models for better decision-making processes, it is necessary to evaluate them across the whole temporal scale range (daily to annual). ACCESS-ESM1-5 which has medium and poor performance in my study, except over alpine AEZ, is preferable over Bale-eco region southern Ethiopia (Gashaw *et al.*, 2024). Models have clear spatial and temporal variation in the performance detection capability of rainfall and other parameters(Gashaw *et al.*, 2024).

Table 4. 3 Performance of CMIP6 GCMs for simulating annual rainfall over AEZs of Ethiopia

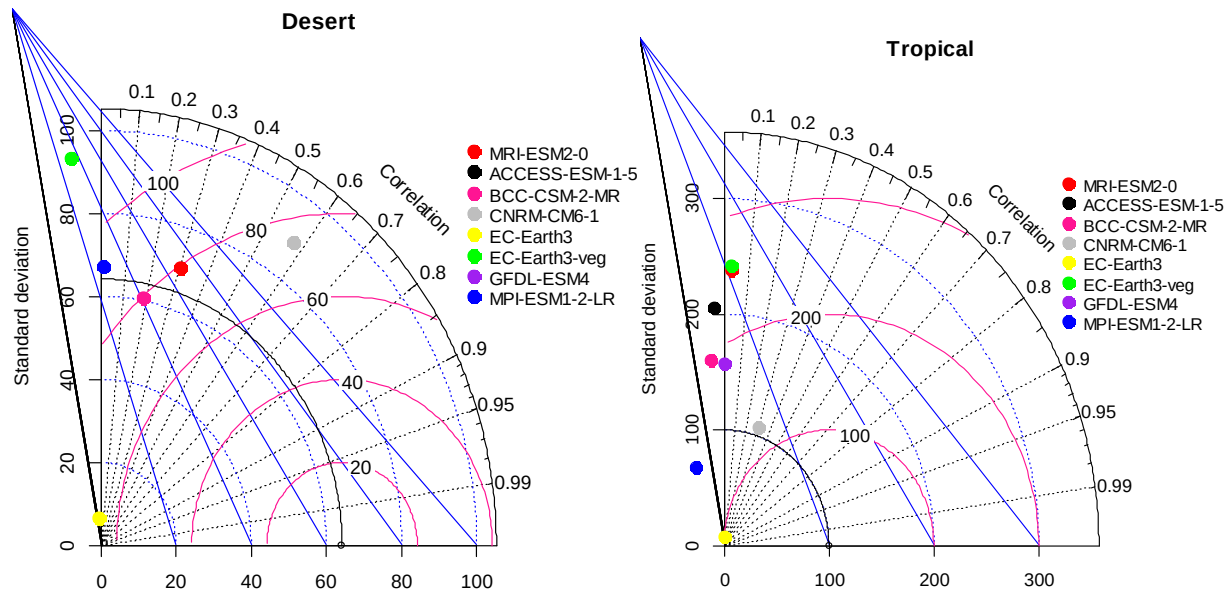
AEZs	CMIP6 Models	Continuous statistical metrics			CRI	Rank
		RMSE	PBIAS	r		
Desert	ACCESS-ESM-1-5	395.63	-178.40	-0.09	0.04	8
	BCC-CSM-2-MR	86.36	-16.40	0.19	0.67	2
	CNRM-CM6-1	122.82	48.60	0.58	0.58	3
	EC-Earth3	196.44	91.90	-0.05	0.29	6
	EC-Earth3-veg	120.08	11.20	-0.08	0.54	4
	GFDL-ESM4	303.17	-128.50	-0.14	0.08	7
	MPI-ESM1-2-LR	133.78	47.90	0.01	0.46	5
	MRI-ESM2-0	79.64	3.10	0.30	0.83	1
Tropical	ACCESS-ESM-1-5	390.48	-30.80	-0.04	0.38	4
	BCC-CSM-2-MR	365.78	-30.30	-0.08	0.33	6
	CNRM-CM6-1	607.75	58.40	0.31	0.38	4

	EC-Earth3	939.32	91.60	0.13	0.25	7
	EC-Earth3-veg	283.79	11.40	0.03	0.63	3
	GFDL-ESM4	214.79	-10.50	0.00	0.67	2
	MPI-ESM1-2-LR	415.80	38.30	-0.36	0.17	8
	MRI-ESM2-0	263.41	-6.20	0.03	0.25	1
Sub-tropical	ACCESS-ESM-1-5	783.92	-58.90	0.01	0.29	7
	BCC-CSM-2-MR	302.63	-19.80	0.00	0.58	3
	CNRM-CM6-1	595.74	46.40	0.25	0.46	4
	EC-Earth3	1124.94	90.70	-0.01	0.13	8
	EC-Earth3-veg	241.70	1.60	0.14	0.79	1
	GFDL-ESM4	182.45	-3.50	-0.09	0.63	2
	MPI-ESM1-2-LR	416.70	32.10	-0.13	0.33	5
	MRI-ESM2-0	465.18	-32.10	-0.11	0.33	5
Temperate	ACCESS-ESM-1-5	640.95	-50.10	0.11	0.25	6
	BCC-CSM-2-MR	146.43	0.90	0.23	0.79	1
	CNRM-CM6-1	281.53	17.30	0.26	0.63	3
	EC-Earth3	1033.63	89.60	0.04	0.13	8
	EC-Earth3-veg	262.16	-9.30	0.24	0.67	2
	GFDL-ESM4	161.48	-3.90	-0.01	0.58	4
	MPI-ESM1-2-LR	362.32	30.00	-0.11	0.29	5
	MRI-ESM2-0	457.47	-35.00	-0.31	0.17	7
Alpine	ACCESS-ESM-1-5	387.40	-30.40	0.04	0.50	2
	BCC-CSM-2-MR	151.93	9.90	0.49	0.83	1
	CNRM-CM6-1	409.78	36.60	0.00	0.38	5

EC-Earth3	821.05	86.00	0.17	0.21	8
EC-Earth3-veg	544.86	-48.30	0.62	0.38	5
GFDL-ESM4	268.02	-15.50	-0.36	0.50	2
MPI-ESM1-2-LR	276.54	24.60	-0.28	0.46	4
MRI-ESM2-0	437.75	-39.10	-0.27	0.25	7

Note that Rank is given based on CRI, and 1st and 8th rank indicates the best and poorest performing models, respectively.

Figure 4.4 showed Taylor diagram representation of the association between CMIP6 GCMs annual rainfall with observed annual rainfall over five AEZs of Ethiopia. The result depicted that there is scattered and weak correlation of CMIP6 GCMs with observed over all AEZs. Most of the models scattered far away from the observed. CNRM-CM6-1 has better correlation with observed than others over desert AEZ. Some models have negative correlation with observed to detect annual rainfall total. Ec-Earth3 has similar performance over all AEZs, with correlation coefficient of near to zero. Ec-Earth3-veg has higher correlation than others over alpine AEZ with correlation coefficient of 0.62.



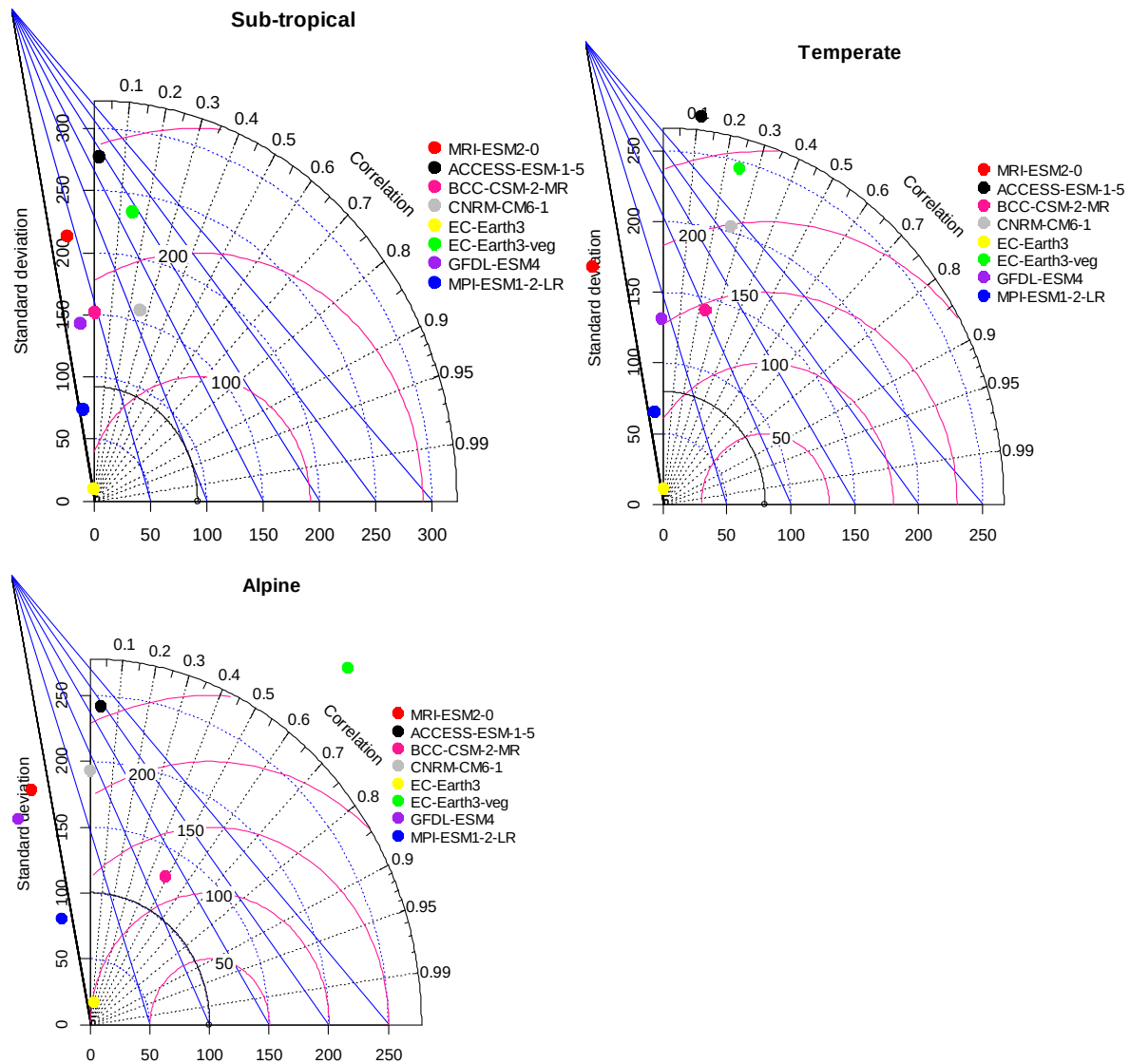


Figure 4.4: Taylor diagram representation of CMIP6 GCMs models annual rainfall against observed annual rainfall over AEZs of Ethiopia for the period 1995-2024. The black line with dot indicates correlation coefficient and standard deviation of observed data.

CHAPTER 5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusion

The study evaluated the performance of eight CMIP6 GCMs in simulating the rainfall climatology of Ethiopia for the period 1995-2014. Ground based products (gauge data) were used to evaluate CMIP6 GCMs. The evaluation used three selected continuous statistical metrics (RMSE, PBIAS, and r) in simulating daily, monthly, and annual rainfall total over five AEZs of Ethiopia. Using CRI, the overall ranking of the CMIP6 GCMs was determined. Furthermore, line graphs and Taylor diagram were used for simulation and visual comparison of observed and CMIP6 GCMs daily, monthly, and annual rainfall total. The result revealed that, models have different performance over AEZs of Ethiopia from daily to annual temporal scales. For better simulation of daily rainfall total, Ec-Earth3-veg (over tropical, sub-tropical, and temperate AEZs), MRI-CSM2-0 (over desert AEZ), and MPI-ESM-1-2-LR (over alpine AEZ) is preferable. Surprisingly, CMIP6 GCMs those simulate daily rainfall total well, is less performed for simulation of monthly rainfall total over all AEZs of Ethiopia, except over desert AEZ. BCC-CSM-2-MR performed well over tropical, sub-tropical, temperate, and alpine AEZs for monthly rainfall total simulation. In addition to BCC-CSM-2-MR, ACCESS-ESM2-0 has also comparable performance for monthly rainfall detection over alpine AEZ. MRI-ESM2-0 has better performance over desert AEZ, not only for daily rainfall total simulation, but also for monthly and annual rainfall total simulation. The mean monthly rainfalls over AEZs of Ethiopia are better captured by Ec-Earth3-veg, GFDL-ESM4, and BCC-CSM2-MR, but Ec-Earth3 is the worst climate model to reproduce mean monthly rainfall of Ethiopia. Annual rainfall total is better simulated by MRI-ESM2-0 (over desert and tropical AEZs), BCC-CSM-2-MR (over temperate and alpine AEZs), and Ec-Earth3-veg (over sub-tropical AEZ). Even though, relatively best models for simulation of daily, monthly, and annual rainfall total have selected for each AEZ, all CMIP6 GCM models have both underestimation and overestimation problem. Moreover, the finding revealed that, the best performing model for daily rainfall total simulation did not perform correspondingly for simulating monthly and annual rainfall total simulation over AEZs of Ethiopia, except over desert AEZ.

5.2. Recommendation

The finding of this study gives important information to model developers and climate model users. The performances of CMIP6 GCMs models are different from place to place and time to time. Hence, evaluating them in our area of interest and consider the bias is a crucial issue. In the studied climate models exhibit both overestimation and underestimation bias. Hence, after evaluate and select the best model, proper bias correction technique shall be applied before proceed for future climate projection and advisory preparation. Since Ethiopia has, different rainfall pattern in different areas, the same AEZ (elevation) in different areas of Ethiopia may have different rainfall distribution. And then, I recommend further CMIP6 GCMs evaluation for rainfall simulation over rainfall regimes of Ethiopia.

REFERENCES

- MOA. (1998). *Agro-ecological Zones of Ethiopia*.
- Addisu, S., Aniley, E., Gashaw, T., Kelemu, S., & Fetene Demessie, S. (2024). Evaluating the performances of gridded satellite products in simulating the rainfall characteristics of Abay Basin, Ethiopia. *Sustainable Environment*, 10(1), 2381349. <https://doi.org/10.1080/27658511.2024.2381349>
- Ageet, S., Fink, A. H., Maranan, M., Diem, J. E., Hartter, J., Ssali, A. L., & Ayabagabo, P. (2022). Validation of satellite rainfall estimates over equatorial east Africa. *Journal of Hydrometeorology*, 23(2), 129–151.
- Ahmadalipour, A., Moradkhani, H., & Rana, A. (2018). Accounting for downscaling and model uncertainty in fine-resolution seasonal climate projections over the Columbia River Basin. *Climate Dynamics*, 50, 717–733.
- Alaminie, A. A., Tilahun, S. A., Legesse, S. A., Zimale, F. A., Tarkegn, G. B., & Jury, M. R. (2021). Evaluation of past and future climate trends under CMIP6 scenarios for the UBNB (Abay), Ethiopia. *Water*, 13(15), 2110.
- Almazroui, M., Ashfaq, M., Islam, M. N., Rashid, I. U., Kamil, S., Abid, M. A., O’Brien, E., Ismail, M., Reboita, M. S., Sörensson, A. A., Arias, P. A., Alves, L. M., Tippet, M. K., Saeed, S., Haarsma, R., Doblas-Reyes, F. J., Saeed, F., Kucharski, F., Nadeem, I., ... Sylla, M. B. (2021). Assessment of CMIP6 Performance and Projected Temperature and Precipitation Changes Over South America. *Earth Systems and Environment*, 5(2), 155–183. <https://doi.org/10.1007/s41748-021-00233-6>
- Almazroui, M., Saeed, F., Saeed, S., Nazrul Islam, M., Ismail, M., Klutse, N. A. B., & Siddiqui, M. H. (2020). Projected change in temperature and precipitation over Africa from CMIP6. *Earth Systems and Environment*, 4, 455–475.
- Aniley, E., Gashaw, T., Abraham, T., Demessie, S. F., Bayabil, H. K., Worqlul, A. W., van Oel, P. R., Dile, Y. T., Chukalla, A. D., Hailelassie, A., & Wubaye, G. B. (2023). Evaluating the performances of gridded satellite/reanalysis products in representing the rainfall climatology of Ethiopia. *Geocarto International*, 38(1), 2278329. <https://doi.org/10.1080/10106049.2023.2278329>

- ASCE Task Committee on Definition of Criteria for Evaluation of Watershed Models of the Watershed Management Committee, I., & Division, D. (1993). Criteria for evaluation of watershed models. *Journal of Irrigation and Drainage Engineering*, 119(3), 429–442.
- Berhanu, D., Alamirew, T., Taye, M. T., Tibebe, D., Gebrehiwot, S., & Zeleke, G. (2023). Evaluation of CMIP6 models in reproducing observed rainfall over Ethiopia. *Journal of Water and Climate Change*, 14(8), 2583–2605. <https://doi.org/10.2166/wcc.2023.502>
- Brighenti, T. M., Gassman, P. W., Gutowski, W. J., & Thompson, J. R. (2023). Assessing the Influence of a Bias Correction Method on Future Climate Scenarios Using SWAT as an Impact Model Indicator. *Water*, 15(4), 750. <https://doi.org/10.3390/w15040750>
- Cannon, A. J., Piani, C., & Sippel, S. (2020). Bias correction of climate model output for impact models. In *Climate Extremes and Their Implications for Impact and Risk Assessment* (pp. 77–104). Elsevier. <https://doi.org/10.1016/B978-0-12-814895-2.00005-7>
- Chang, X., Guo, J., Qin, H., Huang, J., Wang, X., & Ren, P. (2024). Single-Objective and Multi-Objective Flood Interval Forecasting Considering Interval Fitting Coefficients. *Water Resources Management*, 1–20.
- Chen, M., Xie, P., Janowiak, J. E., & Arkin, P. A. (2002). Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations. *Journal of Hydrometeorology*, 3(3), 249–266. [https://doi.org/10.1175/1525-7541\(2002\)003<0249:GLPAYM>2.0.CO;2](https://doi.org/10.1175/1525-7541(2002)003<0249:GLPAYM>2.0.CO;2)
- Cheung, W. H., Senay, G. B., & Singh, A. (2008). Trends and spatial distribution of annual and seasonal rainfall in Ethiopia. *International Journal of Climatology*, 28(13), 1723–1734. <https://doi.org/10.1002/joc.1623>
- D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, & T. L. Veith. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE*, 50(3), 885–900. <https://doi.org/10.13031/2013.23153>
- Dangol, S., Talchabhadel, R., & Pandey, V. P. (2022a). Performance evaluation and bias correction of gridded precipitation products over Arun River Basin in Nepal for hydrological applications. *Theoretical and Applied Climatology*, 148(3–4), 1353–1372.
- Dangol, S., Talchabhadel, R., & Pandey, V. P. (2022b). Performance evaluation and bias correction of gridded precipitation products over Arun River Basin in Nepal for

- hydrological applications. *Theoretical and Applied Climatology*, 148(3–4), 1353–1372.
<https://doi.org/10.1007/s00704-022-04001-y>
- Data, C. (2009). Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation. *World Meteorological Organization*.
- Degefu, M. A., & Bewket, W. (2023). Performance variations of global precipitation products in detecting drought episodes in three wet seasons in Ethiopia: Part II—statistical analysis. *Meteorological Applications*, 30(5), e2154.
- Degefu, M. A., Bewket, W., & Amha, Y. (2022). Evaluating performance of 20 global and quasi-global precipitation products in representing drought events in Ethiopia I: Visual and correlation analysis. *Weather and Climate Extremes*, 35, 100416.
<https://doi.org/10.1016/j.wace.2022.100416>
- Demessie, S. F., Dile, Y. T., Bedadi, B., Gashaw, T., & Tefera, G. W. (2023). Evaluations of regional climate models for simulating precipitation and temperature over the Guder sub-basin of Upper Blue Nile Basin, Ethiopia. *Modeling Earth Systems and Environment*, 9(4), 4455–4476.
- Diallo, I., Xue, Y., Li, Q., De Sales, F., & Li, W. (2019). Dynamical downscaling the impact of spring Western US land surface temperature on the 2015 flood extremes at the Southern Great Plains: Effect of domain choice, dynamic cores and land surface parameterization. *Climate Dynamics*, 53, 1039–1061.
- Dibaba, W. T., Miegel, K., & Demissie, T. A. (2019). Evaluation of the CORDEX regional climate models performance in simulating climate conditions of two catchments in Upper Blue Nile Basin. *Dynamics of Atmospheres and Oceans*, 87, 101104.
- Ding, B., Zhang, J., Zheng, P., Li, Z., Wang, Y., Jia, G., & Yu, X. (2024). Water security assessment for effective water resource management based on multi-temporal blue and green water footprints. *Journal of Hydrology*, 632, 130761.
- Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S. J., & Ropelewski, C. F. (2007). Validation of satellite rainfall products over East Africa’s complex topography. *International Journal of Remote Sensing*, 28(7), 1503–1526.
- Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S., & Ropelewski, C. (2007). Validation of satellite rainfall products over East Africa’s complex topography. *International Journal of Remote Sensing*, 28(7), 1503–1526.

- Dinku, T., Cousin, J., del Corral, R., & Ceccato, P. (2016). The Enacts Approach. *The International Research Institute for Climate and Society (IRI): Palisades, NY, USA*.
- Dinku, T., Faniriantsoa, R., Islam, S., Nsengiyumva, G., & Grossi, A. (2022). The climate data tool: Enhancing climate services across Africa. *Frontiers in Climate*, 3, 787519.
- Diro, G., Grimes, D., Black, E., O'Neill, A., & Pardo-Iguzquiza, E. (2009). Evaluation of reanalysis rainfall estimates over Ethiopia. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 29(1), 67–78.
- Dyer, E., Hirons, L., & Taye, M. T. (2022). July–September rainfall in the Greater Horn of Africa: The combined influence of the Mascarene and South Atlantic highs. *Climate Dynamics*, 59(11), 3621–3641.
- Dyer, E., Washington, R., & Taye, M. T. (2019). Evaluating the CMIP5 ensemble in Ethiopia: Creating a reduced ensemble for rainfall and temperature in Northwest Ethiopia and the Awash basin. *International Journal of Climatology*, 40(6).
- Edwards, P. N. (2011). History of climate modeling. *Wiley Interdisciplinary Reviews: Climate Change*, 2(1), 128–139.
- Evangelista, P., Young, N., & Burnett, J. (2013). How will climate change spatially affect agriculture production in Ethiopia? Case studies of important cereal crops. *Climatic Change*, 119(3), 855–873.
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5), 1937–1958.
- Eyring, V., Flato, G., Lamarque, J.-F., Meehl, J., Senior, C., Stouffer, R., & Taylor, K. (2019). Status of the Coupled Model Intercomparison Project Phase 6 (CMIP6) and goals of the workshop. *Proceedings of the CMIP6 Analysis Workshop, Barcelona, Spain*, 25.
- Eyring, V., Gleckler, P. J., Heinze, C., Stouffer, R. J., Taylor, K. E., Balaji, V., Guilyardi, E., Joussaume, S., Kindermann, S., & Lawrence, B. N. (2016). Towards improved and more routine Earth system model evaluation in CMIP. *Earth System Dynamics*, 7(4), 813–830.
- Eyring, V., Righi, M., Lauer, A., Evaldsson, M., Wenzel, S., Jones, C., Anav, A., Andrews, O., Cionni, I., & Davin, E. L. (2016). ESMValTool (v1. 0)—a community diagnostic and

- performance metrics tool for routine evaluation of Earth system models in CMIP. *Geoscientific Model Development*, 9(5), 1747–1802.
- Fetene, Z. A., Zaitchik, B. F., Zeleke, T. T., Yeshita, B. D., & Vashisht, A. (2022). Coupled Model Intercomparison Project phase 5 and 6 representation of peak and end of rainy season over Upper Blue Nile basin. *International Journal of Climatology*.
- Field, C. B., Barros, V. R., Mastrandrea, M. D., Mach, K. J., Abdrabo, M. A., Adger, W. N., Anokhin, Y. A., Anisimov, O. A., Arent, D. J., Barnett, J., & others. (2015). Summary for policymakers. In *Climate change 2014: Impacts, adaptation, and vulnerability. Part A: global and sectoral aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 1–32). Cambridge University Press.
- Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., & Eyring, V. (2014). Evaluation of climate models. In *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 741–866). Cambridge University Press.
- Fowler, H. J., Blenkinsop, S., & Tebaldi, C. (2007). 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.
- Gashaw, T., W. Worqlul, A., Lakew, H., Teferi Taye, M., Seid, A., & Hailelassie, A. (2023). Evaluations of satellite/reanalysis rainfall and temperature products in the Bale Eco-Region (Southern Ethiopia) to enhance the quality of input data for hydro-climate studies. *Remote Sensing Applications: Society and Environment*, 31, 100994. <https://doi.org/10.1016/j.rsase.2023.100994>
- Gashaw, T., Worqlul, A. W., Taye, M. T., Lakew, H. B., Seid, A., Ayele, G., & Hailelassie, A. (2024). Performance evaluations of CMIP6 model simulations and future projections of rainfall and temperature in the Bale Eco-Region, Southern Ethiopia. *Theoretical and Applied Climatology*. <https://doi.org/10.1007/s00704-024-04904-y>

- Gates, W. L., Henderson-Sellers, A., Boer, G. J., Folland, C. K., Kitoh, A., McAvaney, B. J., Semazzi, F., Smith, N., Weaver, A. J., & Zeng, Q. C. (1995). Climate models—Evaluation. *Climate Change*, 1, 229–284.
- Gebrechorkos, S. H., Hülsmann, S., & Bernhofer, C. (2019). Long-term trends in rainfall and temperature using high-resolution climate datasets in East Africa. *Scientific Reports*, 9(1), 11376.
- Gebrechorkos, S., Leyland, J., Slater, L., Wortmann, M., Ashworth, P. J., Bennett, G. L., Boothroyd, R., Cloke, H., Delorme, P., & Griffith, H. (2023). A high-resolution daily global dataset of statistically downscaled CMIP6 models for climate impact analyses. *Scientific Data*, 10(1), 611.
- Gleckler, P., Doutriaux, C., Durack, P., Taylor, K., Zhang, Y., Williams, D., Mason, E., & Servonnat, J. (2016). A more powerful reality test for climate models. *Eos*, 97.
- Gleckler, P. J., Taylor, K. E., & Doutriaux, C. (2008). Performance metrics for climate models. *Journal of Geophysical Research: Atmospheres*, 113(D6).
- Groisman, P. Y., & Legates, D. R. (1995). Documenting and detecting long-term precipitation trends: Where we are and what should be done. *Climatic Change*, 31(2–4), 601–622.
- Guo, H., Bao, A., Chen, T., Zheng, G., Wang, Y., Jiang, L., & De Maeyer, P. (2021). Assessment of CMIP6 in simulating precipitation over arid Central Asia. *Atmospheric Research*, 252, 105451.
- Guo, S., Zuo, X., Wu, W., Yang, X., Zhang, J., Li, Y., Huang, C., Bu, J., & Zhu, S. (2024). Mitigation of tropospheric delay induced errors in TS-InSAR ground deformation monitoring. *International Journal of Digital Earth*, 17(1), 2316107.
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1–2), 80–91.
- Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1999a). Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *Journal of Hydrologic Engineering*, 4(2), 135–143.
- Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1999b). Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *Journal of Hydrologic Engineering*, 4(2), 135–143.

- Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations—the CRU TS3. 10 Dataset. *International Journal of Climatology*, 34(3), 623–642.
- Iqbal, Z., Shahid, S., Ahmed, K., Ismail, T., Ziarh, G. F., Chung, E.-S., & Wang, X. (2021). Evaluation of CMIP6 GCM rainfall in mainland Southeast Asia. *Atmospheric Research*, 254, 105525.
- JGR Atmospheres—2016—Miao—A nonstationary bias-correction technique to remove bias in GCM simulations (1).pdf. (n.d.).
- Jia, K., Ruan, Y., Yang, Y., & Zhang, C. (2019). Assessing the performance of CMIP5 global climate models for simulating future precipitation change in the Tibetan Plateau. *Water*, 11(9), 1771.
- Jury, M. R. (2015). Statistical evaluation of CMIP5 climate change model simulations for the Ethiopian highlands. *International Journal of Climatology*, 35(1).
- Kamworapan, S., & Surussavadee, C. (2019). Evaluation of CMIP5 global climate models for simulating climatological temperature and precipitation for Southeast Asia. *Advances in Meteorology*, 2019, 1–18.
- Kidd, C. (2001). Satellite rainfall climatology: A review. *International Journal of Climatology*, 21(9), 1041–1066. <https://doi.org/10.1002/joc.635>
- Kindt, R., Van Breugel, P., Lillesø, J. P. B., Minani, V., Ruffo, C. K., Gapusi, J., Jamnadass, R., & Graudal, L. (2014). Potential Natural Vegetation of Eastern Africa (Ethiopia, Kenya, Malawi, Rwanda, Tanzania, Uganda and Zambia) Volume 9. Atlas and Tree Species Composition for Rwanda. *Department of Geosciences and Natural Resource Management, University Of*.
- Kreienkamp, F., Huebener, H., Linke, C., & Spekat, A. (2012). Good practice for the usage of climate model simulation results-a discussion paper. *Environmental Systems Research*, 1(1), 1–13.
- Lebeza, T. M., Gashaw, T., Bayabil, H. K., van Oel, P. R., Worqlul, A. W., Dile, Y. T., & Chukalla, A. D. (2024). Performance of specific CMIP6 GCMs for simulating the historical rainfall and temperature climatology of Lake Tana sub-basin, Ethiopia. *Scientific African*, e02387.

- Legates, D. R., & McCabe Jr, G. J. (1999). Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, 35(1), 233–241.
- Lei, Y. (2022). *Evaluation of global climate models for the simulation of precipitation based on temporal and spatial assessment metrics in China* [Preprint]. In Review. <https://doi.org/10.21203/rs.3.rs-1658240/v1>
- Li, L., Li, W., Ballard, T., Sun, G., & Jeuland, M. (2016). CMIP5 model simulations of Ethiopian Kiremt-season precipitation: Current climate and future changes. *Climate Dynamics*, 46, 2883–2895.
- Loukas, A., Vasiliades, L., & Tzabiras, J. (2008). Climate change effects on drought severity. *Advances in Geosciences*, 17, 23–29.
- Lovino, M. A., Müller, O. V, Berbery, E. H., & Müller, G. V. (2018). Evaluation of CMIP5 retrospective simulations of temperature and precipitation in northeastern Argentina. *International Journal of Climatology*, 38, e1158–e1175.
- Maraun, D. (2016). Bias Correcting Climate Change Simulations—A Critical Review. *Current Climate Change Reports*, 2(4), 211–220. <https://doi.org/10.1007/s40641-016-0050-x>
- Meehl, G. A. (1995). Global coupled general circulation models. *Bulletin of the American Meteorological Society*, 76(6), 951–957.
- Mekasha, A., Tesfaye, K., & Duncan, A. J. (2014). Trends in daily observed temperature and precipitation extremes over three Ethiopian eco-environments. *International Journal of Climatology*, 34(6), 1990–1999.
- Mendez, M., Maathuis, B., Hein-Griggs, D., & Alvarado-Gamboa, L.-F. (2020). Performance Evaluation of Bias Correction Methods for Climate Change Monthly Precipitation Projections over Costa Rica. *Water*, 12(2), 482. <https://doi.org/10.3390/w12020482>
- Mohammed, J. A. (2024). Performance evaluation and ranking of CMIP6 global climate models over upper blue Nile (abbay) basin of Ethiopia. *Natural Hazards Research*, S2666592124000532. <https://doi.org/10.1016/j.nhres.2024.06.004>
- Mohammed, J. A., Gashaw, T., Tefera, G. W., Dile, Y. T., Worqlul, A. W., & Addisu, S. (2022). Changes in observed rainfall and temperature extremes in the Upper Blue Nile Basin of Ethiopia. *Weather and Climate Extremes*, 37, 100468.

- Mohammed, J. A., & Yimam, Z. A. (2022). Spatiotemporal variability and trend analysis of rainfall in Beshilo sub-basin, Upper Blue Nile (Abbay) Basin of Ethiopia. *Arabian Journal of Geosciences*, 15(16), 1387.
- Motlagh, S. B., Honarbakhsh, A., & Azizian, A. (2022). *Bias correction performance of global precipitation forecasting systems in Poldokhtar Watershed in Iran* [Preprint]. In Review. <https://doi.org/10.21203/rs.3.rs-1644511/v1>
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3), 282–290.
- Nashwan, M. S., Shahid, S., & Wang, X. (2019). Assessment of Satellite-Based Precipitation Measurement Products over the Hot Desert Climate of Egypt. *Remote Sensing*, 11(5), 555. <https://doi.org/10.3390/rs11050555>
- Nasrollahi, N., AghaKouchak, A., Cheng, L., Damberg, L., Phillips, T. J., Miao, C., Hsu, K., & Sorooshian, S. (2015). How well do CMIP5 climate simulations replicate historical trends and patterns of meteorological droughts? *Water Resources Research*, 51(4), 2847–2864.
- New, M., Todd, M., Hulme, M., & Jones, P. (2001). Precipitation measurements and trends in the twentieth century. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 21(15), 1889–1922.
- Omay, P. O., Muthama, N. J., Oludhe, C., Kinama, J. M., Artan, G., & Atheru, Z. (2023). *Evaluation of CMIP6 Historical Simulations over IGAD region of Eastern Africa*. <https://doi.org/10.21203/rs.3.rs-2747422/v1>
- Papalexiou, S. M., Rajulapati, C. R., Clark, M. P., & Lehner, F. (2020). Robustness of CMIP6 historical global mean temperature simulations: Trends, long-term persistence, autocorrelation, and distributional shape. *Earth's Future*, 8(10), e2020EF001667.
- Peterson, T. C., & Vose, R. S. (1997). An overview of the Global Historical Climatology Network temperature database. *Bulletin of the American Meteorological Society*, 78(12), 2837–2850.
- Phillips, A. S., Deser, C., & Fasullo, J. (2014). Evaluating modes of variability in climate models. *Eos, Transactions American Geophysical Union*, 95(49), 453–455.
- Polong, F., Pham, Q. B., Anh, D. T., Rahman, K. U., Shahid, M., & Alharbi, R. S. (2023). Evaluation and comparison of four satellite-based precipitation products over the upper

- Tana River Basin. *International Journal of Environmental Science and Technology*, 20(1), 843–858. <https://doi.org/10.1007/s13762-022-03942-1>
- Radcliffe, D., & Mukundan, R. (2017). PRISM vs. CFSR precipitation data effects on calibration and validation of SWAT models. *JAWRA Journal of the American Water Resources Association*, 53(1), 89–100.
- Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., & Srinivasan, J. (2007). Climate models and their evaluation. In *Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC (FAR)* (pp. 589–662). Cambridge University Press.
- Rivera, P. (2023). Evaluation of Historical Simulations of CMIP6 Models for Temperature and Precipitation in Guatemala. *Earth Systems and Environment*, 7(1), 43–65. <https://doi.org/10.1007/s41748-022-00333-x>
- Rudolf, B., Becker, A., Schneider, U., Meyer-Christoffer, A., & Ziese, M. (2011). New GPCC full data reanalysis version 5 provides high-quality gridded monthly precipitation data. *Gewex News*, 21(2), 4–5.
- Salman, S. A., Shahid, S., Afan, H. A., Shiru, M. S., Al-Ansari, N., & Yaseen, Z. M. (2020). Changes in climatic water availability and crop water demand for Iraq region. *Sustainability*, 12(8), 3437.
- Santhi, C., Arnold, J. G., Williams, J. R., Dugas, W. A., Srinivasan, R., & Hauck, L. M. (2001). Validation of the swat model on a large river basin with point and nonpoint sources 1. *JAWRA Journal of the American Water Resources Association*, 37(5), 1169–1188.
- Schamm, K., Ziese, M., Becker, A., Finger, P., Meyer-Christo, A., & Stender, P. (2014). *Global gridded precipitation over land: A description of the new GPCC First Guess Daily product*. 12.
- Shiru, M. S., & Chung, E.-S. (2021). Performance evaluation of CMIP6 global climate models for selecting models for climate projection over Nigeria. *Theoretical and Applied Climatology*, 146(1–2), 599–615. <https://doi.org/10.1007/s00704-021-03746-2>
- Shiru, M. S., Chung, E.-S., Shahid, S., & Alias, N. (2020). GCM selection and temperature projection of Nigeria under different RCPs of the CMIP5 GCMS. *Theoretical and Applied Climatology*, 141, 1611–1627.

- Singh, J., Knapp, H., & Demissie, M. (2004). Hydrologic modeling of the Iroquois River watershed using HSPF and SWAT. ISWS CR 2004–08. *Champaign, Ill.: Illinois State Water Survey*, 41(2), 343–360.
- Spence, T., & Townshend, J. (1996). The Global Climate Observing System (GCOS) An Editorial. *Long-Term Climate Monitoring by the Global Climate Observing System: International Meeting of Experts, Asheville, North Carolina, USA*, 1–4.
- Stouffer, R. J., Eyring, V., Meehl, G. A., Bony, S., Senior, C., Stevens, B., & Taylor, K. E. (2017). CMIP5 scientific gaps and recommendations for CMIP6. *Bulletin of the American Meteorological Society*, 98(1), 95–105.
- Strangeways, I. (2006). *Precipitation: Theory, measurement and distribution*. Cambridge University Press.
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K.-L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, 56(1), 79–107.
- Tan, M. L., & Santo, H. (2018). Comparison of GPM IMERG, TMPA 3B42 and PERSIANN-CDR satellite precipitation products over Malaysia. *Atmospheric Research*, 202, 63–76. <https://doi.org/10.1016/j.atmosres.2017.11.006>
- Tan, Y., Guzman, S. M., Dong, Z., & Tan, L. (2020). Selection of Effective GCM Bias Correction Methods and Evaluation of Hydrological Response under Future Climate Scenarios. *Climate*, 8(10), 108. <https://doi.org/10.3390/cli8100108>
- Tapiador, F. J., Turk, F. J., Petersen, W., Hou, A. Y., García-Ortega, E., Machado, L. A., Angelis, C. F., Salio, P., Kidd, C., Huffman, G. J., & others. (2012). Global precipitation measurement: Methods, datasets and applications. *Atmospheric Research*, 104, 70–97.
- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7), 7183–7192. <https://doi.org/10.1029/2000JD900719>
- Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4), 485–498.
- Trentini, L., Gesso, S. D., Venturini, M., Guerrini, F., Calmanti, S., & Petitta, M. (2023). *A Novel Bias Correction Method for Extreme Events* [Graphic].

- Ul Hasson, S., Pascale, S., Lucarini, V., & Böhner, J. (2016). Seasonal cycle of precipitation over major river basins in South and Southeast Asia: A review of the CMIP5 climate models data for present climate and future climate projections. *Atmospheric Research*, 180, 42–63.
- Van Liew, M., Arnold, J., & Garbrecht, J. (2003). Hydrologic simulation on agricultural watersheds: Choosing between two models. *Transactions of the ASAE*, 46(6), 1539–1551.
- Weber, T., Bowyer, P., Rechid, D., Pfeifer, S., Raffaele, F., Remedio, A., Teichmann, C., & Jacob, D. (2020). Analysis of compound climate extremes and exposed population in Africa under two different emission scenarios. *Earth's Future*, 8(9), e2019EF001473.
- Willmott, C. J., & Matsuura, K. (1995). Smart interpolation of annually averaged air temperature in the United States. *Journal of Applied Meteorology and Climatology*, 34(12), 2577–2586.
- Worku, G., Teferi, E., Bantider, A., & Dile, Y. T. (2019). Observed changes in extremes of daily rainfall and temperature in Jemma Sub-Basin, Upper Blue Nile Basin, Ethiopia. *Theoretical and Applied Climatology*, 135, 839–854.
- Worku, G., Teferi, E., Bantider, A., Dile, Y. T., & Taye, M. T. (2018). Evaluation of regional climate models performance in simulating rainfall climatology of Jemma sub-basin, Upper Blue Nile Basin, Ethiopia. *Dynamics of Atmospheres and Oceans*, 83, 53–63.
- Wubaye, G. B., Gashaw, T., Worqlul, A. W., Dile, Y. T., Taye, M. T., Haileslassie, A., Zaitchik, B., Birhan, D. A., Adgo, E., & Mohammed, J. A. (2023). Trends in rainfall and temperature extremes in Ethiopia: Station and agro-ecological zone levels of analysis. *Atmosphere*, 14(3), 483.
- Wubaye, G. B., Gashaw, T., Worqlul, A. W., Dile, Y. T., Taye, M. T., Haileslassie, A., Zaitchik, B., Birhan, D. A., Adgo, E., Mohammed, J. A., Lebeza, T. M., Bantider, A., Seid, A., & Srinivasan, R. (2023). Trends in Rainfall and Temperature Extremes in Ethiopia: Station and Agro-Ecological Zone Levels of Analysis. *Atmosphere*, 14(3), 483. <https://doi.org/10.3390/atmos14030483>
- Xie, P., Chen, M., & Shi, W. (2010). CPC unified gauge-based analysis of global daily precipitation. *Preprints, 24th Conf. on Hydrology, Atlanta, GA, Amer. Meteor. Soc*, 2.
- Zambrano-Bigiarini, M., Nauditt, A., Birkel, C., Verbist, K., & Ribbe, L. (2017). Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical

- and climatic gradients of Chile. *Hydrology and Earth System Sciences*, 21(2), 1295–1320.
- Zhang, X., Yang, F., & others. (2004). RClimDex (1.0) user manual. *Climate Research Branch Environment Canada*, 22, 13–14.
- Zhang, Y., You, Q., Chen, C., Ge, J., & Adnan, M. (2018). Evaluation of downscaled CMIP5 coupled with VIC model for flash drought simulation in a humid subtropical basin, China. *Journal of Climate*, 31(3), 1075–1090.

APPENDIX

Appendix Table 1. Description of studied meteorological stations over traditional AEZs of Ethiopia

No	Station Name	Geographic coordinate		Elevation	Traditional AEZs	
		Longitude	Latitude		AEZs (Amharic name)	AEZs (English name)
1	Abomsa	39.83	8.47	1630	WoyinaDega	Subtropical
2	Adama	39.28	8.56	1648	WoyinaDega	Subtropical
3	Addis Ababa	38.75	9.02	2386	Dega	Temperate
4	Ambo	37.84	8.99	2068	WoyinaDega	Subtropical
5	Arba Minch	37.56	6.06	1220	Kolla	Tropical
6	AsebeTeferi	40.87	9.07	1796	Kolla	Tropical
7	Asela	39.14	7.96	2420	Dega	Temperate
8	Asgori	38.33	8.79	1979	WoyinaDega	Subtropical
9	Asosa	34.55	10.05	1541	WoyinaDega	Subtropical
10	Bahir Dar	37.42	11.60	1770	WoyinaDega	Subtropical
11	Boditi	37.86	6.96	2043	WoyinaDega	Subtropical
12	Bonga	36.24	7.28	1599	WoyinaDega	Subtropical
13	Butajira	38.38	8.12	2074	WoyinaDega	Subtropical
14	Dangila	36.85	11.43	2116	WoyinaDega	Subtropical
15	Debre Birhan	39.51	9.67	3206	Dega	Temperate
16	Debre Markos	37.74	10.33	2446	Dega	Temperate
17	Debretabor	38.00	11.87	2612	Dega	Temperate
18	Degehabur	43.56	8.23	1070	Kolla	Tropical
19	Dilla	38.31	6.38	1515	WoyinaDega	Subtropical
20	Dire Dawa	41.90	9.61	1045	Kolla	Tropical
21	Dubity	41.09	11.73	381	Bereha	Desert
22	Fiche	38.74	9.77	2798	Dega	Temperate
23	Gimbi	35.84	9.20	1844	WoyinaDega	Subtropical
24	Gonder	37.43	12.52	1973	WoyinaDega	Subtropical
25	Gore	35.55	8.16	2024	WoyinaDega	Subtropical
26	Haramaya	42.04	9.42	2025	WoyinaDega	Subtropical
27	Hawassa	38.48	7.07	1694	WoyinaDega	Subtropical
28	Hosanna	37.86	7.57	2306	Dega	Temperate
29	Jijiga	42.72	9.37	1557	WoyinaDega	Subtropical
30	Jimma	36.82	7.67	1710	WoyinaDega	Subtropical
31	Kombolcha	39.72	11.08	1857	WoyinaDega	Subtropical
32	Konso	37.44	5.34	1431	Kolla	Tropical
33	Lalibela	39.04	12.04	2487	Dega	Temperate
34	Mekele	39.53	13.47	2221	WoyinaDega	Subtropical
35	Metema	36.41	12.77	790	Kolla	Tropical
36	Mizan Teferi	35.58	7.00	1444	Kolla	Tropical

37	Mojo	39.11	8.61	1763	WoyinaDega	Subtropical
38	Moyale	39.05	5.05	1166	WoyinaDega	Subtropical
39	Negele	39.27	5.33	1439	WoyinaDega	Subtropical
40	Nekemte	36.55	9.08	2119	WoyinaDega	Subtropical
41	Pawe	36.41	11.31	1119	Kolla	Tropical
42	Robe	40.05	7.13	2480	Dega	Temperate
43	Teppi	35.44	7.20	1208	Kolla	Tropical
44	Tulu bolo	38.22	8.67	2100	WoyinaDega	Subtropical
45	Wolaita	37.75	6.82	1854	WoyinaDega	Subtropical
46	Yirgachefe	38.21	6.15	1856	WoyinaDega	Subtropical
47	Yirgalem	38.43	6.80	1786	WoyinaDega	Subtropical