Evaluation of CMIP6 models performing on rainfall seasons and moisture tracking simulation in Yangtze River Basin

Wenyu Zhou

TU Delft

Evaluation of CMIP6 models performing on rainfall seasons and moisture tracking simulation in Yangtze River Basin

Thesis report

by

Wenyu Zhou

to obtain the degree of Master of Science
at the Delft University of Technology
to be defended publicly on August 24, 2023 at 10:00

Thesis committee:

Chair:	Dr. Ruud Van der Ent
Supervisors:	Dr. Ruud Van der Ent
External examiner:	Dr. Martine Rutten
Place:	Faculty of Civil Engineering and Geo-science, Delft
Project Duration:	December, 2022 - June, 2023
Student number:	5504449

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

Faculty of Civil Engineering and Geo-science · Delft University of Technology

Abstract

This study evaluates the performance of the CMIP6 models in simulating monsoon rainfall and moisture tracking in the Yangtze River basin. The findings reveal varying degrees of accuracy across different regions of the basin during the monsoon period. Downstream and midstream regions demonstrate higher accuracy, whereas upstream areas exhibit lower precision, along with an overall trend of overestimation. The evaluation encompasses the timing of monsoon months, as well as the peak month, while analyzing the simulation's accuracy for rainfall. It also entails an overarching examination through a Taylor Diagram and Taylor skill scores, which spotlight models with superior and inferior performance. EC-Earth3 exhibits commendable performance, whereas models like IITM-ESM showcase poorer results. Furthermore, moisture tracking assessments, utilizing the WAM2layers model, identify limitations within the CMIP6 model in terms of replicating water vapor sources and pathways, especially in proximity to geographical features such as the Himalayas and the coastline. In addition to the basin itself, the CMIP6 model simulates central Asia as the main source of evaporation, rather than the Indian Ocean, according to the results of ERA5. However, no obvious pattern differences are shown between the different CMIP6 GCMs. Persistent challenges stem from data availability and numerical inconsistencies, necessitating enhancements in both the CMIP6 models and the WAM2layers code.

Contents

List of Figures

List of Tables

1

Introduction

1.1. Yangtze River Basin description

The Yangtze River is the third longest river in the world, with a total length of 6,300 kilometers. Its basin extends about 3,200 kilometers from west to east and more than 1,000 kilometers from north to south, covering an area of 1.8 million square kilometers, or about one-fifth of China's total land area [1]. Often referred to as the "Golden Waterway," the Yangtze River Basin (YRB) plays a crucial role in China's inland water transportation, boasting a comprehensive network of over 57,000 kilometers of primary and secondary water routes, accounting for 52.5% of the nation's waterways[2, 3].

The YRB exhibits a prototypical monsoonal climate and is situated within the East Asian subtropical monsoon region. The interaction of multi-stage topography and different circulation systems (e.g. East Asian monsoon, Indian monsoon, Australian monsoon, mid-latitude westerlies, Plateau monsoon, northern cold invasion, etc.) leads to a complex climate condition in the region with dry winters and wet summers, as well as large spatial differences [4].

The average annual precipitation in the basin is 1100mm, while the average annual precipitation in the southeastern coastal region reaches 1200 to 1400mm, and the annual precipitation in three rainy areas, including the west Sichuan, east Sichuan and Xiang-E-Gan, exceeds 2000mm, and the annual precipitation at Jinshan Station in Xinjing County, Sichuan is as high as 2590mm. The rainy season is from April to October and lasts for 7 months, with precipitation accounting for about 85% of the annual precipitation [5]. Precipitation in the middle and lower streams is mainly influenced by the East Asian monsoon [6, 7], while summer rainfall in the upper streams is mainly related to the intrusion of the Indian summer wind [8, 9].

1.2. Division of Yangtze River Basin

The upstream of Yangtze River is from Sanjiangyuan (the source) to Yichang, with a length of 4,500 km and a basin area of about 1 million square kilometers. Most of the main stream and tributaries flow through plateau, high mountain and canyon areas, especially the Tongtian River, Jinsha River and Three Gorges area, which have obvious characteristics of plateau mountain canyon rivers. The riverbeds here have large specific drop, and the difference between the main stream of Jinsha River alone is 3000 meters. The rivers have abundant water, fast flowing and rich hydraulic resources [11]. For example, The Three Gorges Dam which is the world's largest hydroelectric hub, is located in Yichang and generates about 100 billion kilowatt-hours of electricity annually [12].

In the source area is a plateau area, located in the arid zone, with annual precipitation below 400mm. The semi-humid zone with annual precipitation between 400 and 800mm is mainly located in the central and northern parts of the upstream basin, including the western Sichuan plateau, Qinghai-Tibetan Plateau. There are also some areas belonging to the humid zone with extremely high rainfall, in the southern and eastern Sichuan Basin [11].

The middle segment of the Yangtze River Basin spans around 1000 km from Yichang to Hukou. This stretch of the river courses through the Jianghan Plain, a low-lying expanse carved out by the Hanshui and Yangtze rivers, located upstream of Wuhan. The river's path in this area is winding, resulting in a broad river surface and a steep drop in the riverbed gradient, so the flow is relatively sluggish, maintaining

Figure 1.1: Map showing the relationship between the Yangtze River and its principal tributaries and topography [10].

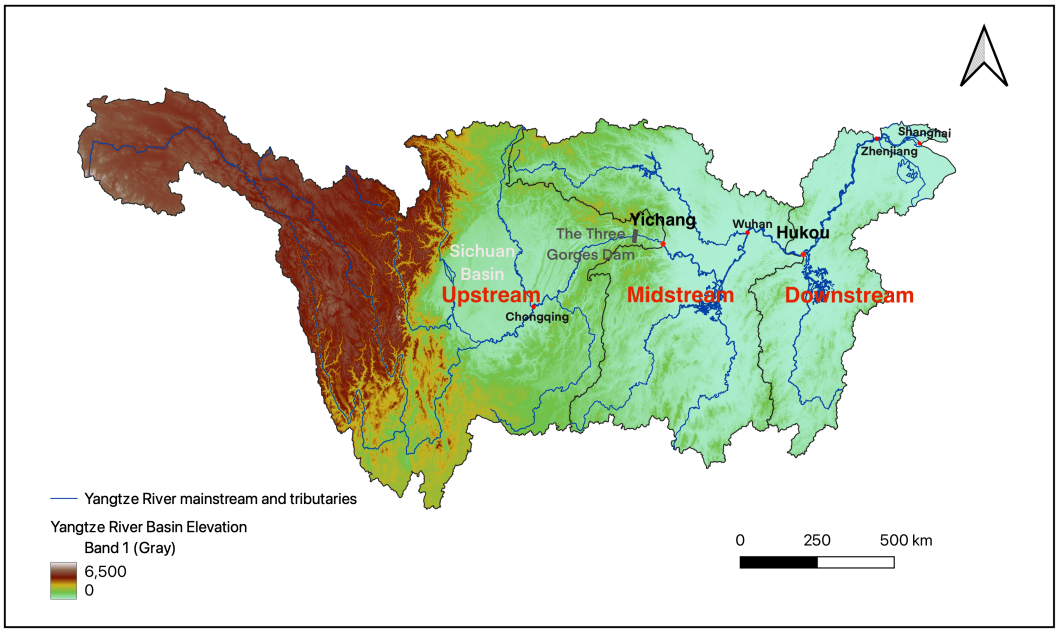


Figure 1.2: The upstream, midstream and downstream basin of Yangtze River Basin, with the dividing point of Yichang and Hukou

an average speed of 1 m/s. This section of the river is dotted with numerous tributaries and lakes. The catchment area of the middle stream contributes to about 40% of the entire basin's catchment, leading to a rapid increase in the Yangtze River's water volume. The intricate water systems in this region play a pivotal role in regulating the water quantity within the main river channel. Hydrological attributes of this river segment, including water levels and flow rates, are notably influenced by rainfall patterns [11].

In this area, most of the basin areas are located in the humid zone, with annual precipitation around 1000mm in the Jianghan Plain area in the northern part of the Yangtze River, and the average annual precipitation can exceed 1200mm in the southern part of the basin.

As we move downstream, the lower reaches of the Yangtze River extend over 800 km, stretching from Hukou to its mouth where it meets the sea. Here, the river widens and deepens, while its tributaries are relatively short and exert minimal influence on the Yangtze River's water volume. Near the mouth of the Yangtze River, the river turns southeast and enters the vast Yangtze River Delta region. The terrain there is flatter, with a dense network of water and many lakes. In the southeastern part of the basin rainfall can exceed 1600mm per year, making it a particularly humid zone [11].

1.3. Introduction of CMIP6 Data

General Circulation Models (GCMs) have the ability to simulate the effects of greenhouse gas (GHG) emissions on climatic systems and realistically predict future conditions [13, 14, 15]. These models are widely used to model past climatic conditions and project future responses to increased GHG emissions and land-use changes [16, 13].

The Coupled Model Intercomparison Project (CMIP) is an evaluation of different emission scenarios leading to climate changes, including historical emissions resulting in climate at the current time, and future climate change based on the currently adopted emission scenarios. It led to a better understanding of past, present and future climate change and variability in a multi-model framework [17].

In CMIP, a variety of GCMs are available. these models contain large uncertainties due to inadequate model descriptions of the physical processes driving the climate system and climate scenarios[18]. Certain models are able to address regional climate events, thus increasing their usefulness in predicting future climate change scenarios for specific regions.

CMIP has developed in phases and the latest phase is the sixth phase, CMIP6, which corrects the long-standing model bias and poor quantification of radiative forcing of CMIP5 [19]. In many studies, CMIP6 has been shown to be better than CMIP5 in predicting rainfall, both in terms of interannual variability of extreme precipitation index (EPI) and current highest 1-day precipitation and highest 5-day precipitation simulations [19, 20]. The uncertainties in simulating seasonal variability of rainfall and temperatures were lower for CMIP6 compared to CMIP5 [21].

1.4. Research Goals

Part 1: Evaluation of CMIP6 models for rainfall simulation in the Yangtze River basin

The GCMs in CMIP6 show different accuracy for rainfall simulations in the Yangtze River basin.

Research Objective

In this part, we will evaluate the models' performance on rainfall using different CMIP6 GCMs datasets, compared with observed data.

Research Question 1

How do the models simulate the rainfall amount, seasonality, and spatial performance of historical rainfall in the YRB?

Research Question 2

Which models show higher accuracy in the YRB? Which models perform poorly?

Part 2: Evaluation of CMIP6 models for moisture source simulation in the Yangtze River basin

Research Objective

In the first part, we obtain the better performing models and the less performing models, so representative models were selected to simulate the water vapor sources separately.

Research Question 1

How accurate are the models simulating water vapor sources of historical data in the YRB?

2

Literature Review

2.1. Background of monsoon rainfall analysis

The study of rainfall patterns and their implications is crucial due to the significant influence of precipitation on various aspects of human society, economic activities, and the natural environment. Precipitation variability, both spatially and temporally, holds immense importance not only for historical analysis but also for understanding potential changes in future rainfall patterns [22].

Monsoonal regions, such as India and China, heavily rely on monsoonal rainfall, which plays a vital role in society, agriculture, and the economy. Monsoon rainfall exhibits notable spatial and temporal variability, influenced by both internal and external factors [23]. In these regions, the amount of rainfall during the monsoon season, particularly from June to September, surpasses other months and constitutes a significant portion of the annual precipitation [24].

The monsoon system in East Asia involves two major monsoon rains occurring between April and September, namely the plum rains in central China and the late July rainfall in northeastern China [25]. The Indian monsoon and East Asian monsoon greatly influence rainfall in the Chinese region, with atmospheric circulation changes at low and mid-latitudes playing a crucial role [24]. The water vapor transport patterns in the Indian monsoon and East Asia exhibit an inverse relationship, affecting precipitation in the middle and lower reaches of the Yangtze River [26].

Various studies have demonstrated changes in monsoon intensity and rainfall patterns in different parts of China, particularly in the southeast and south, since the mid-20th century [27, 28]. These changes in rainfall are attributed to alterations in water vapor supply and the magnitude of water vapor transport from different sources and production areas [29]. The influence of the El Niño-Southern Oscillation (ENSO) phenomenon on rainfall has also been highlighted, explaining significant variations in fall and winter precipitation in southern China [30].

Moreover, the teleconnection between East Asian subtropical monsoon rainfall and rainfall patterns in North America has been observed, suggesting that perturbations in subtropical East Asia may impact rainfall in North America through upper-level Asia-North America (ANA) telecorrelation [31].

In the analysis of monsoon rainfall, diverse methods and indicators are employed. Researchers use various indicators such as the peak month of rainfall, monsoon season onset and withdrawal, and monthly rainfall distribution to examine the spatial and temporal characteristics of rainfall data, providing valuable insights for agriculture and water resources management [32, 33].

Overall, these studies collectively contribute to our understanding of monsoon rainfall patterns, their variability, and the factors influencing them, facilitating improved decision-making in sectors reliant on precipitation and water resource management.

2.2. Background of moisture tracking

The moisture cycle plays a critical role in redistributing moisture from the point of evaporation to precipitation in different regions. Changes in land use and vegetation can significantly impact the hydrological cycle, making it essential to determine the source of rainfall evaporation and the location of water vapor precipitation, particularly for water resources and agricultural planning [34, 35].

In agriculture, alterations in the water cycle due to land cover change can lead to potential crop yield reductions. Studies have shown that land cover change may cause a 1-17% decrease in crop yields in major agricultural regions worldwide [36]. Deforestation in tropical regions has been found to reduce evapotranspiration, increase surface temperatures, and disrupt boundary layer circulation, resulting in changes in rainfall patterns. The impact of deforestation on the water cycle can lead to a regional reduction in rainfall by up to 40%[37]. Additionally, irrigation practices in heavily irrigated areas contribute to both precipitation and evapotranspiration, but the increase in precipitation is often smaller than the increase in evapotranspiration, which can contribute to water scarcity in these areas[38].

Understanding the water vapor cycle is crucial for studying regional weather patterns. Moisture-tracking models have been employed to analyze the moisture sources of precipitation and identify factors contributing to extreme droughts. For example, a study in southwest China found that extreme droughts in 2006 and 2011 were primarily caused by a significant reduction in external moisture sources [39]. In South China, the main source of atmospheric moisture for heavy rainfall events during the pre-flood season is the Indian Ocean, accounting for 30% of the atmospheric water vapor and influencing precipitation in the region. Significant shifts in water vapor transport from the Indian Ocean and the western Pacific Ocean have been observed, indicating changes in regional precipitation patterns [40].

Different atmospheric moisture tracking models exist, each with its own approach. The RCM-tag method, for instance, integrates highly accurate 3-D water tracking within a regional climate model. Other methods, such as WAM and 3D-T, use either the Eulerian or Lagrangian approach to track moisture. The choice of model and the grid box size used can influence the accuracy of results, and appropriate time steps and tracer parcels must be selected to ensure reliable outcomes [41, 42]. However, the computational requirements for Lagrangian models can be significant due to the need to track individual water blocks over time.

2.3. The CMIP models' performance on study region

CMIP models have been widely used to predict future changes in rainfall. Over the past decades, several GCMs have been developed for climate change projections, and simulations of these GCMs have been released in phases within the framework of the Coupled Model Intercomparison Project (CMIP). CMIP3 consists of 25 GCMs, CMIP5 has 40 GCMs, and the more recent CMIP6 has 55 GCMs [43].

Currently, the CMIP6 model has been widely used for the analysis of rainfall climate in various regions. In the South Asian region, many people have used multiple GCM datasets to analyze rainfall, and the results show that the CMIP6 GCMs can accurately capture the monsoon rainfall of Indian Summer Monsoon Rainfall (ISMR) and outperforms CMIP5 [44, 45, 46, 47]. In addition, machine learning has been applied to the spatial distribution of daily rainfall over mainland India during the monsoon, and a major observation is that all CMIP6 models grossly overestimate the spatial correlation of monsoon rainfall, suggesting that rainfall in model simulations is spatially smoother than actual data [48].

The performance evaluation of Iqbal's selected GCMs in the Mainland South-East Asia (MESA) region revealed their ability to simulate mean annual rainfall for the climatology of the central and southern regions of MSEA with a deviation of less than 25%, and it was concluded that the identified CMIP6 GCMs could be used for climate change prediction and impact assessment in MSEA after correcting for the relevant deviations [49].

For Mid-Asia and East Asia such as China, Yan et al. analyzed the performance of CMIP6 in the East Asian Monsoon and found that in eastern and southern China it still exhibits considerable systematic biases in the amplitude and phase of the annual cycle, with the better performing models being BCC-ESM1, CanESM5 and GFDL-CM4 [50]. Xin et al. analyzed eight models of CMIP6 and CMIP5 and concluded that CMIP6 multi-model integration (MME) is more proficient than CMIP5 MME in spatial correlation and standard deviation (SD) for climate precipitation studies in eastern China [51], current CMIP6 models have a nationwide overestimation of precipitation, Jiang et al. found a comparable overestimation of the spatial variability of the interannual variability [52].

For the study area, the Yangtze River basin, the upper reaches are influenced by the Indian summer monsoon and rainfall in the middle and lower reaches is mainly brought by the East Asian monsoon, with an overall overestimate trend [9, 52, 48]. A large number of CMIP model evaluations have been used for the Yangtze River basin. Since CMIP3, most models can reproduce the maximum rainfall in YRB during the

monsoon season, but overestimate the mean annual rainfall [53]. In CMIP5, most models still overestimate precipitation, especially in winter and in the upper Yangtze River region [54, 55]. Even so, CMIP5 models the mean and extreme values of rainfall in the middle and lower Yangtze River better than the intensity and extreme values of precipitation [56]. Therefore, CMIP5 models are still widely considered to reproduce the spatial distribution of seasonal rainfall in the Yangtze River basin, despite the large differences between models [56]. For the CMIP6 model, most models still overestimate the rainfall, especially in the source area, and also simulate large wet deviations in winter. However, the models reproduce the increased precipitation observed in the northwestern part of the basin [4].

On previous studies, there is a lack of accuracy evaluation of CMIP6 for monsoon period simulations, and very few simulations on water vapor sources. In addition, previous studies have not chosen many models, so expand the study in this area to find the best and the worst performing models.

2.4. Evaluation of General Circulation Models (GCMs)

Most researchers use many different metrics to evaluate the model performance of GCMs, including probability distribution functions [57, 58], correlation coefficients [59], Bayesian methods [60], information entropy [61], Recursive Feature Elimination [62], and symmetric uncertainty [63], and also researchers have used WMO extreme rainfall indices to evaluate GCMs simulations of extreme rainfall [64], Goodman-Kruskal's lambda, Mapcurves, Cramer's V, and Theil U based on four spatial, exponential MCDA methods, were applied to calculate the GCM ranking [49]. For temporal variability, trend analysis and selection skill score IVS (Interannual Variability Skill Score) were used to analyze rainfall in Asia [65, 66], and IVS was also considered in the Taylor skill score.

Since Taylor designed the Taylor diagram (TD) in 2001, summarizing multiple aspects of model performance in a single diagram has been widely used [67]. Taylor diagrams are widely used to obtain more intuitive results, and Taylor skill scores are used to obtain overall evaluations [68, 69, 19]. However, Taylor diagrams and Taylor skill scores have limitations, such as the inability to capture nonlinear relationships and sensitivity to outliers. To address these limitations, kernelized Taylor plots have been proposed that can visualize the similarity between groups under minimal assumptions on the data distribution [70]. All these methods have informed the ranking and selection of GCMs.

On the python package, Eyring et al. developed the Earth System Model Evaluation Tool (ESMValTool), which provides evaluation results for CMIP6 model simulations using a large number of diagnostic and performance metrics, and is also widely used today.This tool focuses on selected essential climate variables (ECVs), a range of known systematic biases common to ESMs, such as coupled tropical climate variability, monsoons, Southern Ocean processes, continental drought bias, and soil hydro-climate interactions, as well as atmospheric CO_2 budgets, tropospheric and stratospheric ozone, and tropospheric aerosols, with the addition of effective climate sensitivity (ECS) and transient climate response (TCR) in the newly updated version 2.0, and various climate-related feedback and diagnostic emergency constraints for future projections of ESM [71, 72]. If only TD plots are considered, Rochford provides a python package called skill metrics for plotting TD plots [70].

3

Data Collection

3.1. Data for Rainfall models evaluation

3.1.1. Observation datasets

Gridded precipitation products are widely used to evaluate CMIP models [73, 74, 59, 75], but there are large sources of uncertainty and processing algorithms among different data sets [76]. In order to minimize the errors introduced by uncertainty, I selected four different gridded precipitation datasets and chose their average values as the observation data set, which are APHRODITE, GPCC, ERA5 and CRU. CRU performs best in capturing drought events, GPCC is best at representing heavy rainfall events, ERA5 identifies precipitation distributions and centers but underestimates extremes, and APHRODITE performs well in the relatively flat test basin [76, 77, 78].

APHRODITE

APHRODITE is a long-term, high-resolution, daily rainfall dataset specially developed for the Asian monsoon region (including mountain ranges) covering a period of more than 50 years, from 1951-2007. `https://www.chikyu.ac.jp/precip/` (Retrieved in February 12, 2023). It has been developed using observed data collected by different national organizations and climate data gathered by Global Telecommunication System network (GTS) with the help of various International Agencies [79, 80, 81]. The gridded products are available for four sub-domains–Monsoon Asia, Middle East, Russia, and Japan–as well as a combined domain [49]. APHRODITE has two resolutions, $0.25^\circ \times 0.25^\circ$ and $0.5^\circ \times 0.5^\circ$. In this thesis, the dataset with $0.5^\circ \times 0.5^\circ$ resolution will be used.

Global Precipitation Climatology Centre

Global Precipitation Climatology Centre (GPCC) from Physical Sciences Laboratory (NOAA) is a monthly precipitation dataset from 1891 to present which is calculated from global station data. `https://psl.noaa.gov/data/gridded/data.gpcc.html` (Retrieved in February 20, 2023). Four kinds of resolution were provided and $0.5^\circ \times 0.5^\circ$ resolution is selected.

The data are based on data from 67,200 stations worldwide with a record duration of 10 years or more. the GPCC contains monthly totals on a regular grid with spatial resolutions of $0.25^\circ \times 0.25^\circ$ and $0.5^\circ \times 0.5^\circ$, $1.0^\circ \times 1.0^\circ$ and $2.5^\circ \times 2.5^\circ$ latitude and longitude. In addition, GPCC interpolates precipitation anomalies at the stations and then superimposes them on GPCC Climatology V2020 at the corresponding resolutions [82].

ERA5-Land monthly averaged data

The ERA5-Land provides a consistent view of surface water and energy cycles over several decades. It contains detailed records from 1950 onwards with a temporal resolution of 1 h. The original spatial resolution of the ERA5-Land reanalysis dataset is 9 km on a simplified Gaussian grid (TCo1279). Data have been re-gridded to a regular latitude/longitude grid of $0.1^\circ \times 0.1^\circ$. `https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land-monthly-means?tab=overview` (Retrieved in February 20, 2023) [83, 84]. In the first part of the paper, monthly average rainfall data from 1951-2007 were selected in order to match the time period of APHRODITE Data.

Climatic Research Unit gridded Time Series
Climatic Research Unit gridded Time Series (CRU TS) is a widely used climate dataset on a $0.5°$ latitude by $0.5°$ longitude grid over all land domains of the world except Antarctica. It is derived by the interpolation of monthly climate anomalies from extensive networks of weather station observations [85], `https://crudata.uea.ac.uk/cru/data/hrg/#info` (Retrieved in February 20, 2023).

The CRU TS dataset provides a high-resolution monthly grid of land (excluding Antarctica) observations going back to 1901. CRU TS v. 3.23 from 1901-2007 are selected here.

3.1.2. General Circulation Models (GCMs) from Cmip6 datasets

This CMIP6 coordinates independent model intercomparison activities and their experiments which have adopted a common infrastructure for collecting, organizing, and distributing output from models performing common sets of experiments. The output of the models have been downloaded from the open-access platform `https://esgf-node.llnl.gov`. In this thesis, 42 models are selected for the analysis of GCMs, and the model source IDs, resolutions and responsible institutions are shown in Table 3.2. The monthly average rainfall from 1951 to 2007 was selected.

The CMIP6 dataset is divided into historical data and future data projected according to different emission scenarios. CMIP6 provides scenario called ”Shared Socioeconomic Pathways” (SSP), which takes into account global economic and demographic changes and Green House Gas (GHG) emissions for climate modeling. Of the five SSP scenarios, SSP1 and SSP5 represent a positive shift toward human development through significant investments in health and education, advanced institutions, and rapid economic growth. The main difference between these two scenarios is the rapid shift to sustainable practices in SSP1 and the shift to a fossil fuel-based economy in SSP5. SSP3 and SSP4 represent a pessimistic future with slow economic development and rapid population growth leading to an unequal distribution of resources, with SSP2 showing an intermediate road scenario compared to SSP1 and SSP3 [49].

In the CMIP6 historical data, data on anthropogenic emissions from agriculture, aircraft, energy, industry, international shipping, residential and commercial, solvent production and application, transportation, and waste are provided for each country for the years 1750-2014 [86]. In addition, historical emissions data for open burning as of 2015, as well as emissions of N_2O and fluorinated gases were also considered [87, 88, 89].

Table 3.1: Description of Observation Datasets

Dataset	Resolution	Temporal coverage	url
APHRODITE	$0.5° \times 0.5°$	1951-2015	https://www.chikyu.ac.jp/precip/
GPCC	$0.5° \times 0.5°$	1891-present	https://psl.noaa.gov/data/gridded/data.gpcc.html
ERA5-Land	$0.1° \times 0.1°$	1950-present	https://cds.climate.copernicus.eu/cdsapp#!/dataset
CRU TS v. 3.23	$0.5° \times 0.5°$	1901-2014	https://crudata.uea.ac.uk/cru/data/hrg/#info

3.2. Data for moisture tracking

3.2.1. ERA5 hourly Reanalysis data

Since the water vapor entering the Yangtze River Basin is global, and the global data set is huge. In order to reduce the memory of the data as well as the amount of computation, longitude between 20°E and 160°E and latitudes between 20°S and 50°N should be chosen, this is because previous studies have shown that the main source of water vapor in the YRB lies within this range [90].

In order to calculate moisture source, it is necessary to take into account of different factors such as evaporation, precipitation, wind speed, pressure, specific humidity. Hourly data is considered to provide more precise information, and hence, the ERA5-hourly data set is recommended as the most suitable option. ERA5-Hourly data reanalysis for the global climate and weather for the past 8 decades. Reanalysis combines model data with observations from across the world into a globally complete and consistent data set using the laws of physics [91, 92].

The variables downloaded from ERA5-hourly are shown in the Table 3.3.

For variables specific humidity, the q, u component of wind, the u, v component of wind, the v, data are in different levels, for other variables surface data are downloaded. Due to the wide study area and the high frequency of the data, the size of the data is very large, so in this study, only data from the decade 2005-2014 is selected.

3.2.2. CMIP6 model datasets

To assess the precision of CMIP6 models in simulating water vapor sources, specific CMIP6 models will be chosen for water vapor source simulation in alignment with the outcomes of the initial phase. This selection process is informed by two CMIP6 daily datasets utilized in this study.

Not all models have all the variables needed for all water vapor simulations because of the different experimental conditions and the types of variables obtained for different models. The names and units of the variables required for CMIP6 model are shown in Table 3.4. Since the CMIP6 model does not provide daily evaporation, surface upward latent heat flux is used instead.

All experiments were checked and only EC-earth3 and MPI-ESM-1-2 were found to meet the requirements, providing the following variables.

Table 3.2: Detail description of GCMs selected in this thesis

Source ID	Insitution	Resolution
ACCESS-CM2 ACCESS-ESM1-5	Commonwealth Scientific and Industrial Research Organization/Australia	250km
BCC-ESM1	Beijing Climate Center/China	250km
CAMS-CSM1-0	Chinese Academy of Meteorological Sciences/China	100km
CanESM5 CanESM5-CanOE	Canadian Centre for Climate Modelling and Analysis/Canada	500km
CAS-ESM2-0	Chinese Academy of Sciences/China	100km
CESM2	National Center for Atmospheric Research, Climate and Global Dynamics Laboratory/USA	100km
CESM2-FV2		250km
CESM2-WACCM		100km
CESM2-WACCM-FV2		250km
CMCC-CM2-HR4 CMCC-CM2-SR5 CMCC-ESM2	Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici/Italy	100km
CNRM-CM6-1	Centre National de Recherches Meteorologiques/France	250km
CNRM-CM6-1-HR		100km
CNRM-ESM2-1		250km
E3SM-1-0	E3SM-Project LLNL(Lawrence Livermore National Laboratory), UCI(Department of Earth System Science, University of California Irvine), UCSB(Bren School of Environmental Science and Management, University of California)/USA	100km
E3SM-2-0	E3SM-Project/USA	100km
EC-Earth3 EC-Earth3-AerChem EC-Earth3-CC	EC-Earth-Consortium/Europe	100km
FGOALS-f3-L	Chinese Academy of Sciences/China	100km
FGOALS-g3		250km
FIO-ESM-2-0	First Institute of Oceanography, Ministry of Natural Resources/China	100km
GFDL-ESM4	NOAA Geophysical Fluid Dynamics Laboratory/USA	100km
GISS-E2-1-G GISS-E2-1-H	Goddard Institute for Space Studies/USA	250km
IITM-ESM	Centre for Climate Change Research, Indian Institute of Tropical Meteorology Pune/India	250km
INM-CM4-8 INM-CM5-0	Institute for Numerical Mathematics/Russia	100km
IPSL-CM5A2-INCA	Institute Pierre Simon Laplace/ France	500km
IPSL-CM6A-LR		250km
IPSL-CM6A-LR-INCA		250km
KACE-1-0-G	National Institute of Meteorological Sciences/Korea Meteorological Administration, Climate Research Division/Koera	250km
KIOST-ESM	Korea Institute of Ocean Science and Technology/Korea	250km
MCM-UA-1-0	Department of Geosciences, University of Arizona/USA	250km
MIROC-ES2L	JAMSTEC (Japan Agency for Marine-Earth Science and Technology/ Japan), AORI (Atmosphere and Ocean Research Institute, The University of Tokyo/Japan), NIES (National Institute for Environmental Studies/Japan), and R-CCS (RIKEN Center for Computational Science/Japan)	500km
MPI-ESM-1-2-HAM	Max Planck Institute for Meteorology/Germany	250km
MPI-ESM1-2-HR		100km
MRI-ESM2-0	Meteorological Research Institute/Japan	100km
NorCPM1	NorESM Climate modeling Consortium consisting of CICERO (Center for International Climate and Environmental Research), MET-Norway (Norwegian Meteorological Institute), NERSC (Nansen Environmental and Remote Sensing Center), NILU (Norwegian Institute for Air Research), UiB (University of Bergen), UiO (University of Oslo) and UNI (Uni Research)/Norway	250km

Table 3.3: Description of Varibales from ERA5

Variables	Long Name	Units
q	specific humidity	kg/kg
u	u component of wind	m/s
v	v component of wind	m/s
tp	total precipitation	m/s
e	evaporation	m
sp	surface pressure	Pa
d2m	2m dewpoint temperature	K
u10	10m u component of wind	m/s
v10	10m v component of wind	m/s
tcw	total column water	kg/m^2

Table 3.4: Description of Varibales from CMIP6

Variables	Long Name	Units
hus	specific humidity	kg/kg
ua	u component of wind	m/s
va	v component of wind	m/s
pr	precipitation flux	$kg/m^2/s$
hfls	surface upward latent heat flux	W/m^2
ps	surface air pressure	Pa
tdps	2m dewpoint temperature	K
uas	10m u component of wind	m/s
vas	10m v component of wind	m/s

4

Methodology

4.1. Part 1: Rainfall pattern analysis and Model estimation

4.1.1. Data pre-processing

Since the original data sets I downloaded are almost global data, it needed to be clipped with the Yangtze River Basin geographic shape file. The Yangtze River Basin shapefiles are from the data center of the Institute of Geographical Sciences and Resources, Chinese Academy of Sciences. `https://www.resdc.cn/data.aspx?DATAID=141` (Retrieved in February 12, 2023). Considering that different datasets have different resolutions, all the data sets are processed to $0.5^\circ \times 0.5^\circ$ resolution for the convenience of calculation and comparison. In addition, since different data sets have different units, the monthly rainfall data of all data sets are converted to mm/day before calculation.

To reduce the error, for the data sets of four observations, their average is taken and compared with the CMIP6 GCMs.

4.1.2. Monsoon season deviation

Since the rainfall in the Yangtze River basin is mainly brought by the monsoon, the indicator for classifying the monsoon period is a very important factor. Considering that the hydrological environment and climatic conditions in the upstream, midstream and downstream basins of the Yangtze River basin vary greatly, it is impossible to choose a uniform quantitative standard.

Therefore, the indicator for the months of the monsoon period is chosen as: greater than half of the average rainfall of the two months with the highest rainfall. According to the pattern of the summer monsoon movement, the monsoon months in the Yangtze River basin are about 5-7 months, with the longest monsoon months in the midstream and the shortest in the upstream.

Therefore, the months entering and leaving the monsoon period can be an indicator of the model performance. Another indicator is the time of occurrence of monthly rainfall peaks and the amount of rainfall; if the model simulates closer to the ground, it proves that the simulation is more effective.

4.1.3. Taylor Diagram

To evaluate the performance of CMIP6 GCMs, multiple statistical metrics are applied, plotted as Taylor Diagram (TD).

These metrics measure the difference between the model field ($cmip$) and the observation field (obv), which each represent a vector of data values with the same number of elements [70].

Metrics used for the plots are bias (B), $B = \bar{cmip} - \bar{obv}$, the correlation coefficient (R), the root-mean-square deviation ($RMSD$), and the standard deviations of the model field (σ_{cmip}) and the observation field (σ_{obv}). Their formulas are given by equations below. The correlation coefficient (R) can be written as:

$$R = \frac{\frac{1}{N}\sum_{n=1}^{N}\left(cmip_n - \bar{cmip}\right)\left(obv_n - \bar{obv}\right)}{\sigma_{cmip}\sigma_{obv}} \tag{4.1}$$

And the $RMSD$ is as:

$$RMSD^2 = \frac{1}{N}\sum_{n=1}^{N}\left[(cmip_n - \bar{cmip}) - \left(obv_n - \bar{obv}\right)\right]^2 \tag{4.2}$$

The standard deviations σ_{cmip} and σ_{obv} are:

$$\begin{aligned}\sigma_{cmip}^2 &= \frac{1}{N}\sum_{n=1}^{N}\left(cmip_n - \bar{cmip}\right)^2 \\ \sigma_{obv}^2 &= \frac{1}{N}\sum_{n=1}^{N}\left(obv_n - \bar{obv}\right)^2\end{aligned} \tag{4.3}$$

Taylor diagrams display three statistics that are useful for assessing the degree of correspondence between the modeled and observed behavior: the Pearson correlation coefficient (R), the standard deviations of the model (σ_{cmip}) and observation (σ_{obv}) fields, and the root-mean-square deviation ($RMSD$). Mathematically, the three statistics are related by the following formula:

$$RMSD^2 = \sigma_{cmip}^2 + \sigma_{obv}^2 - 2\sigma_{cmip}\sigma_{obv}R \tag{4.4}$$

Taylor plots have the advantage of illustrating multiple statistics in a single plot and determining how much of the RMSE can be attributed to variance differences and how much is attributed to poor pattern similarity [67].

4.1.4. Taylor Skill Score

Taylor skill score (TSS) was employed to estimate the performance of CMIP6 models in reproducing the spatial rainfall patterns over Yangtze river basin. The TSS is computed using Eq. 4.5:

$$TSS = 4(1+R)^2 / \left[\left(\frac{\sigma_{\text{cmip}}}{\sigma_{\text{cts}}} + \frac{\sigma_{\text{obs}}}{\sigma_{\text{cmip}}}\right)(1+R_0)^2\right] \tag{4.5}$$

Where R is the spatial pattern correlation coefficient between the model outputs and observation. The R_0 is the highest achievable (here, we set the threshold at 1). The score 1 threshold value shows a perfect association between models and observed whereas 0 expresses contrary model performance [69, 67].

4.2. Part 2: Moisture tracking

4.2.1. WAM2layers model

WAM2layers can be used to determine where precipitation originally evaporated (backtracking), or where evaporated moisture eventually ends up (forward tracking) [93]. WAM-2layers uses the Eulerian method to track the flow of water vapor around the Earth [34]. At each time step, WAM-2layers keeps track of how much water enters the atmosphere as evaporation and how much water leaves the atmosphere as precipitation. Between each time step, WAM-2layers calculates how much water moves in each grid cell in the four basic directions.

The basic equation for WAM2layers moisture tracking models is the balance of a water mass in Eq. 4.6:

$$\frac{\partial S_k}{\partial t} = \frac{\partial\left(S_k u\right)}{\partial x} + \frac{\partial\left(S_k \nu\right)}{\partial y} + E_k - P_k + \xi_k \pm F_{\text{v}}\left[\text{L}^3\ \text{T}^{-1}\right] \tag{4.6}$$

Where S_k is the atmospheric moisture storage (i.e., precipitable water) in layer k (either the top or the bottom layer). t is time, u and v are the wind components in x and y direction. E_k is evaporation entering layer k, P_k is precipitation removed from layer k. ξ is a residual and F_v is the vertical moisture transport between the bottom and top layers.

The moisture transport due to horizontal transport can be written as Eq. 4.7:

$$\begin{aligned} \frac{\Delta(Su)}{\Delta x} &= F^{-}_{k,x} - F^{+}_{k,x} \\ \frac{\Delta(Sv)}{\Delta y} &= F^{-}_{k,y} - F^{+}_{k,y} \end{aligned} \tag{4.7}$$

Where F_k is the moisture flux over the boundary of grid cell in the bottom and top layer, positive from west to east and north to south.

The moisture flux can be calculated as Eq. 4.8:

$$F_k = \frac{L}{g\rho_{\mathsf{w}}} \int_{p_{\mathsf{top}}}^{p_{\mathsf{bottom}}} q u_{\mathsf{h}} \mathsf{d}p \tag{4.8}$$

Where L is the length of grid cell perpendicular to the direction of the moisture flux, g is the gravitational acceleration, ρ_w is the density of liquid water, p stands for pressure, q for specific humidity and u_h for horizontal component in either x or y direction.

For the top layer, these following conditions are applied: $p_{top} = 0$ and $p_{bottom} = p_{divide}$. For the bottom layer applies: $p_{top} = p_{divide}$ and $p_{bottom} = p_{surface}$. The p_{divide} is define as:

$$p_{\mathsf{divide}} = 7438.803 + 0.728786 \times p_{\mathsf{surface}} \, [\mathsf{Pa}] \tag{4.9}$$

This division appeared to best capture the division between sheared wind system, where wind in the bottom layer goes in another direction than wind in the top layer [94].

In Eq. 4.6, the E_k only enters in bottom layer, thus $E_k = E$ in the bottom and $E_k = 0$ in the top. While precipitation is assumed to be immediately removed from moisture storage, and assume ”well-mixed” conditions for precipitation by van der Ent [41]:

$$P_k = P\frac{S_k}{S} \tag{4.10}$$

where P is total precipitation and S is total atmospheric storage in the vertical. The residual ξ in 4.6 is the result of data-assimilation in ERA-I and the fact that our offline tracking scheme calculates the water balance on a coarser spatial and temporal resolution.

In the application of the WAM2layers model, the developer found that the the vertical flux F_v in 4.6 was too small to adequately take care of the vertical transport of tagged water, so this term is used as $4F_v$ in the direction net flux and $3F_v$ in the opposite direction in WAM2layer [41].

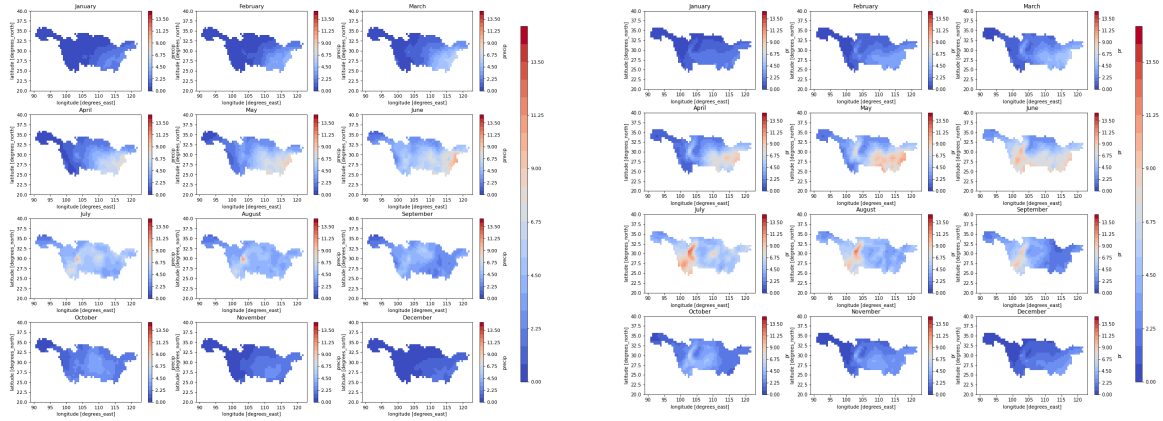
5 Results

5.1. CMIP6 model performance on monsoon rainfall

5.1.1. Model performance for monsoon months

The accuracy of the CMIP6 model in reproducing rainfall patterns in the YRB differs across various research institutions and experimental configurations. To illustrate this, we can examine the APHRODITE dataset (refer to Figure 5.1). The dataset reveals a noticeable spatial pattern in the monthly average rainfall, indicating that the lower Yangtze River region experiences the onset of the rainy season in April, while the upper Yangtze River region encounters it in June. When assessing different CMIP6 datasets, the degree of accuracy in simulating the magnitude, temporal variations, and spatial distribution of monthly average rainfall serves as the basis for evaluating the model's performance.

In the provided link, a collection of spatial distribution maps showcasing the monthly average rainfall for the 42 models is displayed. These maps show the simulation of different models for monthly average rainfall, on each grid of the Yangtze River Basin. For most of the models, they can simulate the difference in rainfall in different months, as well as different rainfall rain patterns within the basin in the same month. Overall, from April to September are the months with concentrated rainfall, in line with the monsoon months. These maps will be further examined and analyzed in subsequent discussions. `https://drive.google.com/drive/folders/1p8gRLnsIawuuC5Q8LCJ932Qg5DEE_01z?usp=drive_link`



(a) (b)

Figure 5.1: Monthly mean rainfall map in YRB, a clear trend of rainfall advancing from downstream to upstream with time can be seen. a. APHRODITE observation dataset, b. EC-Earth3 GCM dataset, legend from 0 to 15.75 mm/day

Different models exhibit varying levels of accuracy when simulating rainfall in the YRB. Some models demonstrate good performance in reproducing both the quantity and timing of rainfall events, which are crucial factors for evaluating model performance.

Figure 5.2, 5.3, and 5.4 display monthly rainfall maps during the monsoon season for the upper, middle, and lower reaches of the YRB. To mitigate the errors inherent in the observed datasets, the top row of the figures presents the average of four different datasets, while the subsequent four rows depict the results from each individual observation dataset.

In the following section, the CMIP6 models are ranked based on their downstream basin TSS scores. This ranking provides a comprehensive overview of the simulation of monsoon monthly rainfall, distinguishing between models that perform exceptionally well and those that exhibit poorer performance.

In the downstream basin, the monsoon season typically begins in March and ends in August, with the peak occurring in June. Figure 5.2 reveals that 16 models accurately simulate the onset and conclusion of the monsoon period, including models such as EC-Earth3 and FLO-ESM-2-0, which simulate the same start month and end month as observation dataset. However, the majority of models show a one-month discrepancy in simulating the monsoon season, either commencing or concluding it a month earlier or later. Poorly performing models like IITM-ESM and MCM-UA-1-0 depict a prolonged monsoon season throughout the year, even featuring a second monsoon.

Among the 42 CMIP6 models, only 8 models accurately simulate the peak occurrence of monsoon rainfall in June, while the majority cluster around May. Overall, the simulated monsoon peaks, especially in the poorly performing models, tend to be higher than the observed values. Figure 5.3-(a) shows the average difference between the twelve months calculated for each modeled data compared to the observed data. It illustrates that models in the top 30 ratings exhibit errors within 20%, whereas some of the poorly performing models with low TSS score such as ACCESS-CM2 and INM-CM4-8 exceed 50% difference.

In the middle YRB Figure 5.3 and 5.3-(b), the monsoon season has a longer duration compared to the downstream region, typically spanning from March to September and the average monthly rainfall is relatively lower. The observed dataset indicates a peak rainfall in June at approximately 6.4 mm/day. Among the models, six successfully replicate the onset and conclusion of the monsoon period, while the majority exhibit a one-month error in either direction. However, the two worst performing models, IITM-ESM and MCM-UA-1-0, simulate monsoon rainfall throughout the year.

Regarding the peak simulation, most models place it in May, one month earlier than the observed June peak. In general, the simulated peak values across all models tend to be higher than the observed values. Figure 5.3-(b) illustrates that poorly performing models display significant differences, with most exceeding 40%, while well-performing models manage to keep the differences below 20%.

Moving to the upper Yangtze River basin, Figure 5.4, the monsoon season spans from May to September, with a shorter duration and less rainfall compared to the downstream and middle stream regions. The observed data indicates a peak monthly rainfall of only 5.6 mm/day in July. While most models accurately simulate the timing of the peak, they consistently overestimate the peak rainfall. Figure 5.3-(c) demonstrates that all models exhibit substantial errors in overestimating upstream rainfall, with only a few below 50% and many exceeding 100%.

In terms of simulating the monsoon months, only 4 models accurately replicate the timely onset and end of the monsoon season. 11 models exhibit a one-month error in either the start or the end month, while other models demonstrate errors of two months or more, indicating highly inaccurate simulations.

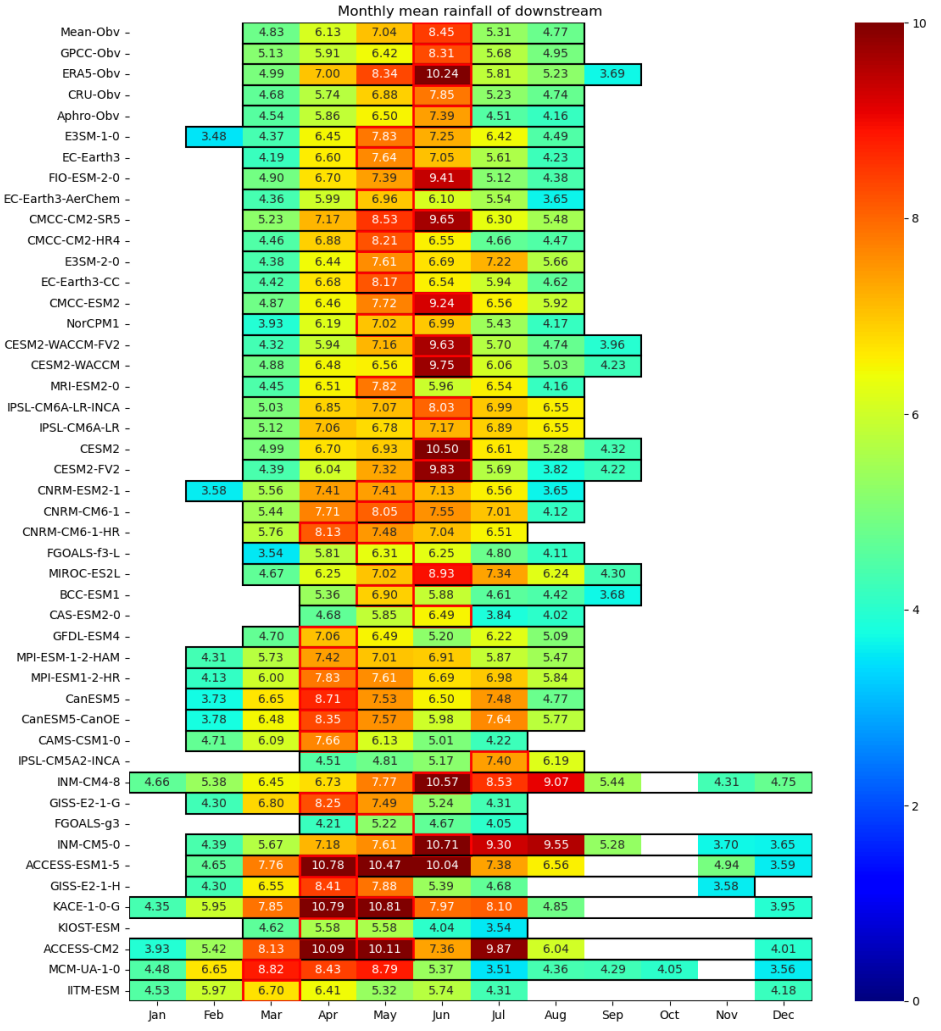


Figure 5.2: Monthly rainfall during the monsoon period in the downstream of YRB

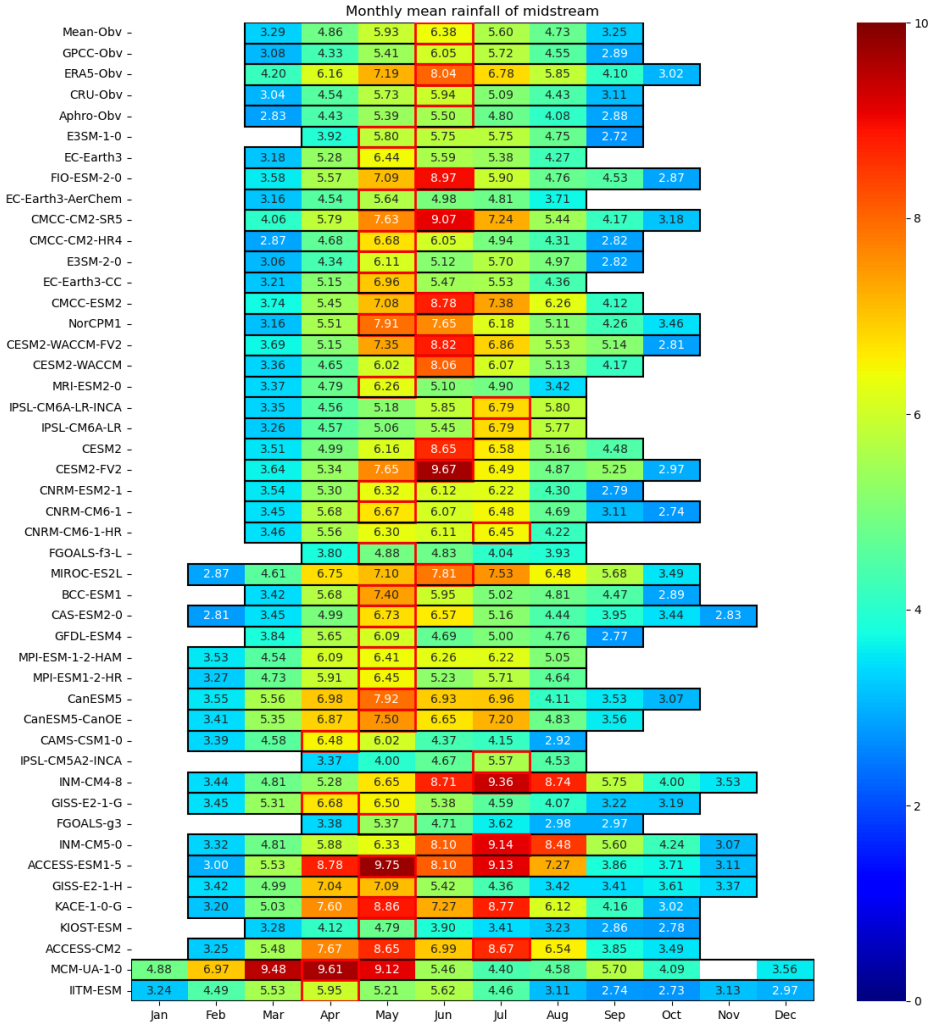


Figure 5.3: Monthly rainfall during the monsoon period in the midstream of YRB

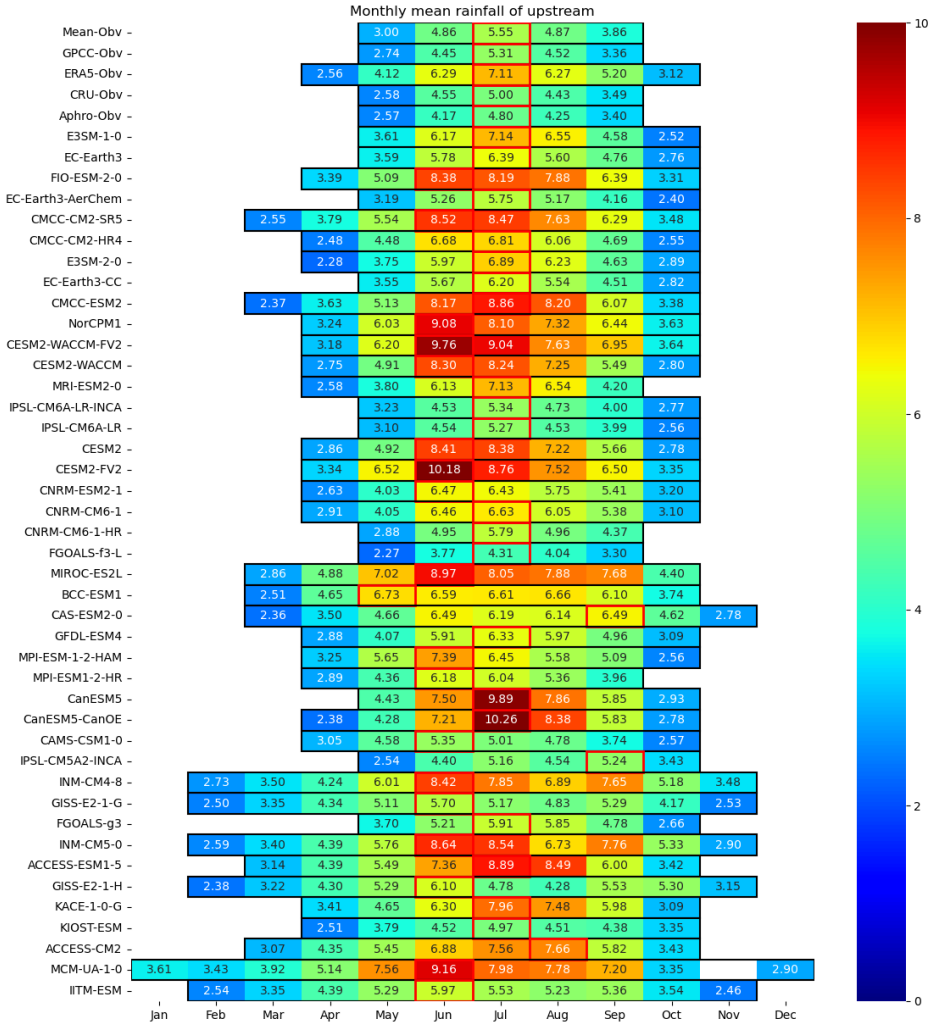
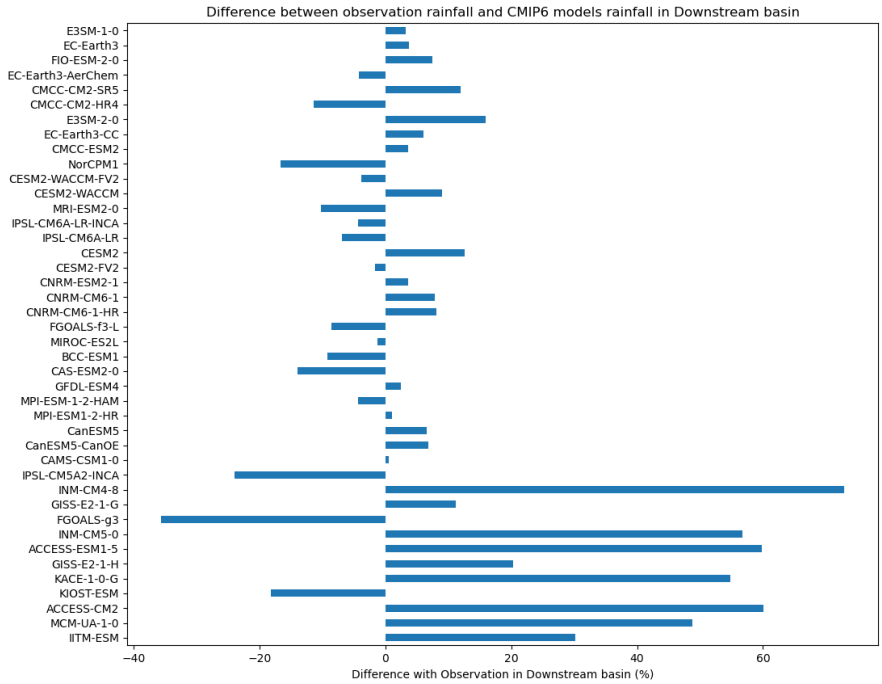
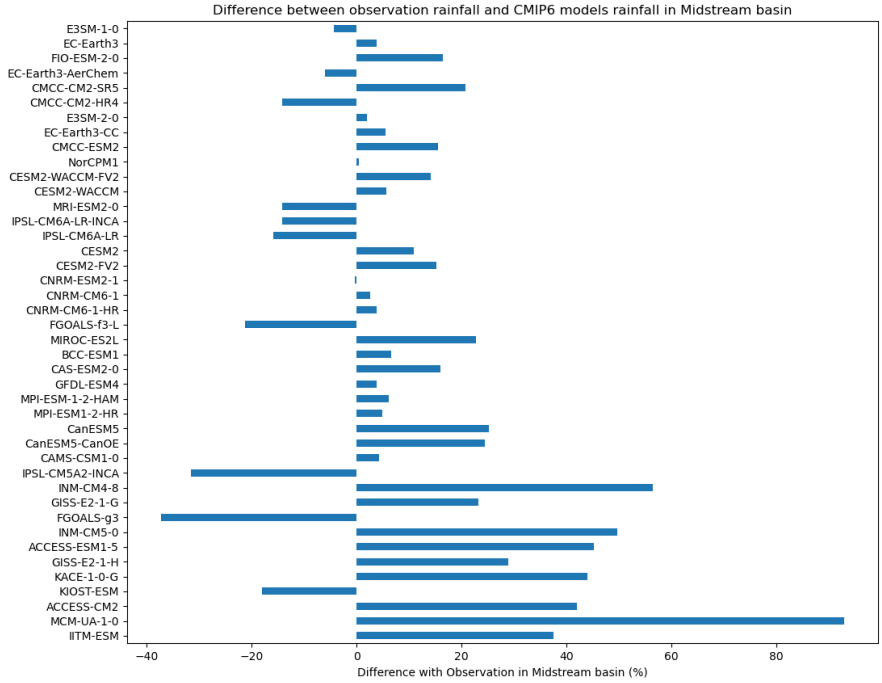


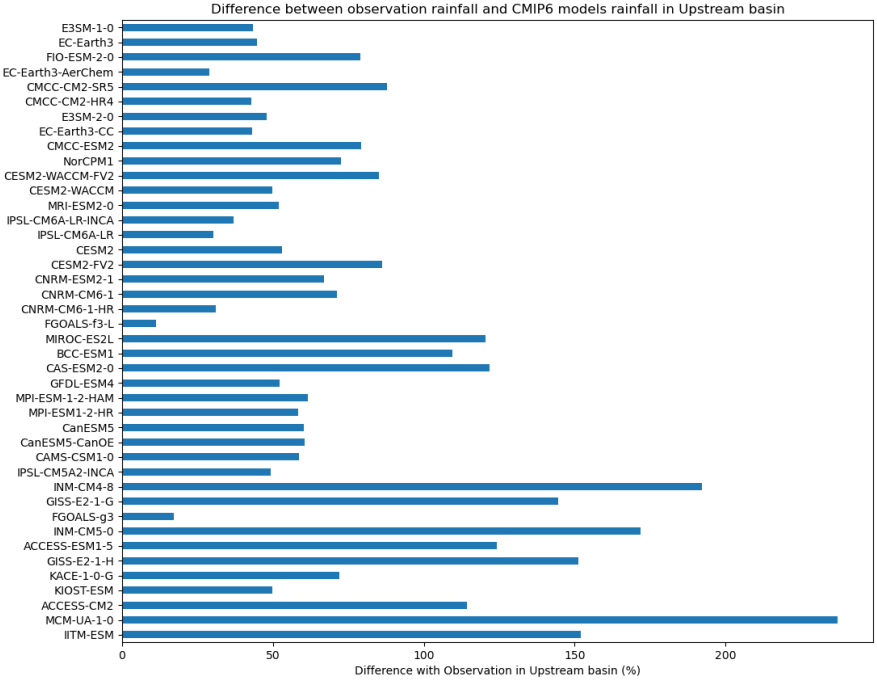
Figure 5.4: Monthly rainfall during the monsoon period in the upstream of YRB


Difference between observation rainfall and CMIP6 models rainfall in Downstream basin
E3SM-1-0
EC-Earth3
FIO-ESM-2-0
EC-Earth3-AerChem
CMCC-CM2-SR5
CMCC-CM2-HR4
E3SM-2-0
EC-Earth3-CC
CMCC-ESM2
NorCPM1
CESM2-WACCM-FV2
CESM2-WACCM
MRI-ESM2-0
IPSL-CM6A-LR-INCA
IPSL-CM6A-LR
CESM2
CESM2-FV2
CNRM-ESM2-1
CNRM-CM6-1
CNRM-CM6-1-HR
FGOALS-f3-L
MIROC-ES2L
BCC-ESM1
CAS-ESM2-0
GFDL-ESM4
MPI-ESM-1-2-HAM
MPI-ESM1-2-HR
CanESM5
CanESM5-CanOE
CAMS-CSM1-0
IPSL-CM5A2-INCA
INM-CM4-8
GISS-E2-1-G
FGOALS-g3
INM-CM5-0
ACCESS-ESM1-5
GISS-E2-1-H
KACE-1-0-G
KIOST-ESM
ACCESS-CM2
MCM-UA-1-0
IITM-ESM
−40
−20
0
20
40
60
Difference with Observation in Downstream basin (%)

(a)


Difference between observation rainfall and CMIP6 models rainfall in Midstream basin
E3SM-1-0
EC-Earth3
FIO-ESM-2-0
EC-Earth3-AerChem
CMCC-CM2-SR5
CMCC-CM2-HR4
E3SM-2-0
EC-Earth3-CC
CMCC-ESM2
NorCPM1
CESM2-WACCM-FV2
CESM2-WACCM
MRI-ESM2-0
IPSL-CM6A-LR-INCA
IPSL-CM6A-LR
CESM2
CESM2-FV2
CNRM-ESM2-1
CNRM-CM6-1
CNRM-CM6-1-HR
FGOALS-f3-L
MIROC-ES2L
BCC-ESM1
CAS-ESM2-0
GFDL-ESM4
MPI-ESM-1-2-HAM
MPI-ESM1-2-HR
CanESM5
CanESM5-CanOE
CAMS-CSM1-0
IPSL-CM5A2-INCA
INM-CM4-8
GISS-E2-1-G
FGOALS-g3
INM-CM5-0
ACCESS-ESM1-5
GISS-E2-1-H
KACE-1-0-G
KIOST-ESM
ACCESS-CM2
MCM-UA-1-0
IITM-ESM
−40
−20
0
20
40
60
80
Difference with Observation in Midstream basin (%)

(b)



(c)

Figure 5.3: The difference between CMIP6 models and mean observation dataset, with a. The difference in downstream river basin, b. The difference in midstream river basin and c. The difference in upstream river basin.

5.1.2. The spatial correlation coefficients between CMIP6 models and mean observation data

The spatial correlation coefficient maps of CMIP6 models with the average observed data can be found in this link `https://drive.google.com/drive/folders/1kEZlN-bmKVJmFhLy19ArH4IUnTCUBDev?usp=sharing`, where the spatial correlation coefficient maps of 42 models are displayed, from which the performance of different models in YRB can be clearly seen. The maps of spatial correlation coefficients of four CMIP6 models with average observed data are shown in Figure 5.4. The EC-Earth3 model of Figure 5.4-(a) has a higher spatial correlation coefficient than other models overall, with an average of 0.96. Figure 5.4-(b) is MPI-ESM-1-2-HAM, which performs slightly worse, showing a lower spatial correlation coefficient in the midstream region and the northern part of the downstream, with an overall coefficient of 0.79. Some models such as MCM-UA-1-0 and GISS-E2-1-H perform relatively poorly, with large negative correlations, as shown in Figure 5.4-(c) and (d). The worst performing model among all models has an average correlation coefficient of 0.5 and is MCM-UA-1-0.

5.1.3. Taylor Diagram and Taylor Skill Scores

The monthly rainfall statistics of the 42 models for the upstream, midstream, and downstream regions are presented in a Taylor diagram, as depicted in Figure 5.5. The Taylor diagram provides a comprehensive overview of model performance by incorporating RMSD, standard deviation, and spatial correlation coefficients in a single plot, enabling a clear assessment of each model's strengths and weaknesses. Models that closely align with the observed data's standard deviation (red line) and exhibit a higher spatial correlation coefficient (closer to the bottom) are considered to perform better.

Overall, the downstream region demonstrates excellent performance, as illustrated in Figure 5.5-(a). The model points are concentrated near the standard deviation of the observed data, with RMSD values close to 1 mm/day and correlation coefficients approaching 0.9, indicating highly accurate simulations.

The midstream region, as shown in Figure 5.5-(b), exhibits slightly lower performance compared to the downstream, but still demonstrates accurate simulations with correlation coefficients close to 0.85.

However, the upstream region displays lower simulation accuracy, as depicted in Figure 5.5-(c). The model points are scattered across the diagram, and although the correlation coefficient can reach 0.8, most models exhibit large errors in terms of standard deviation, consistent with the differences observed in Figure 5.3-(c).

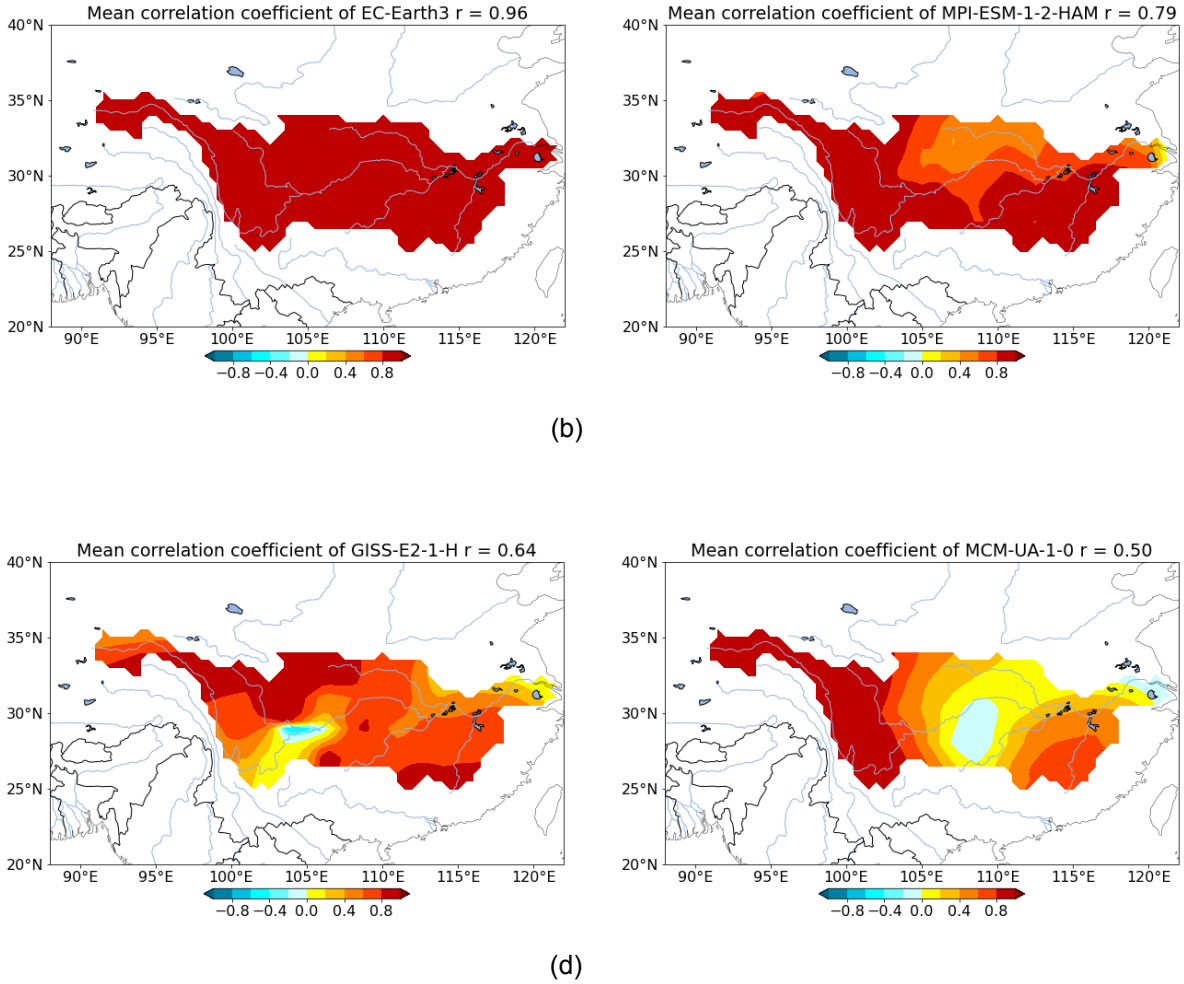


Figure 5.4: Examples of spatial correlation coefficients between CMIP6 data and observation data, a. EC-Earth3 with a mean value of 0.96, b. MPI-ESM-1-2-HAM with a mean value of 0.79, c. GISS-E2-1-H with a mean value of 0.64, d. MCM-UA-1-0 with a mean value of 0.50

For the downstream and midstream regions, the top-performing models are EC-Earth3 and E3SM1-0, while the poorly performing models are IITM-ESM and MCM-UA-1-0. In the upstream region, the better-performing models are EC-Earth3-Aerchem, GISS-E2-1-G, and GISS-E2-1-H exhibit poorer performance.

The Taylor Skill Score quantifies the performance of each model based on various statistics, and a higher score indicates better model performance. Figure 5.6 illustrates the Taylor Skill Scores for the upstream, midstream, and downstream simulations. In general, most models achieve high scores for the midstream and downstream simulations, with the poorer performing models scoring lower. However, for the upstream region, a larger number of models receive lower scores, indicating inadequate simulation performance. The TSS and rankings for the different streams are provided in the Appendix A, Table A.4. The reason for this situation may be that the CMIP6 models do not simulate complex terrain such as high

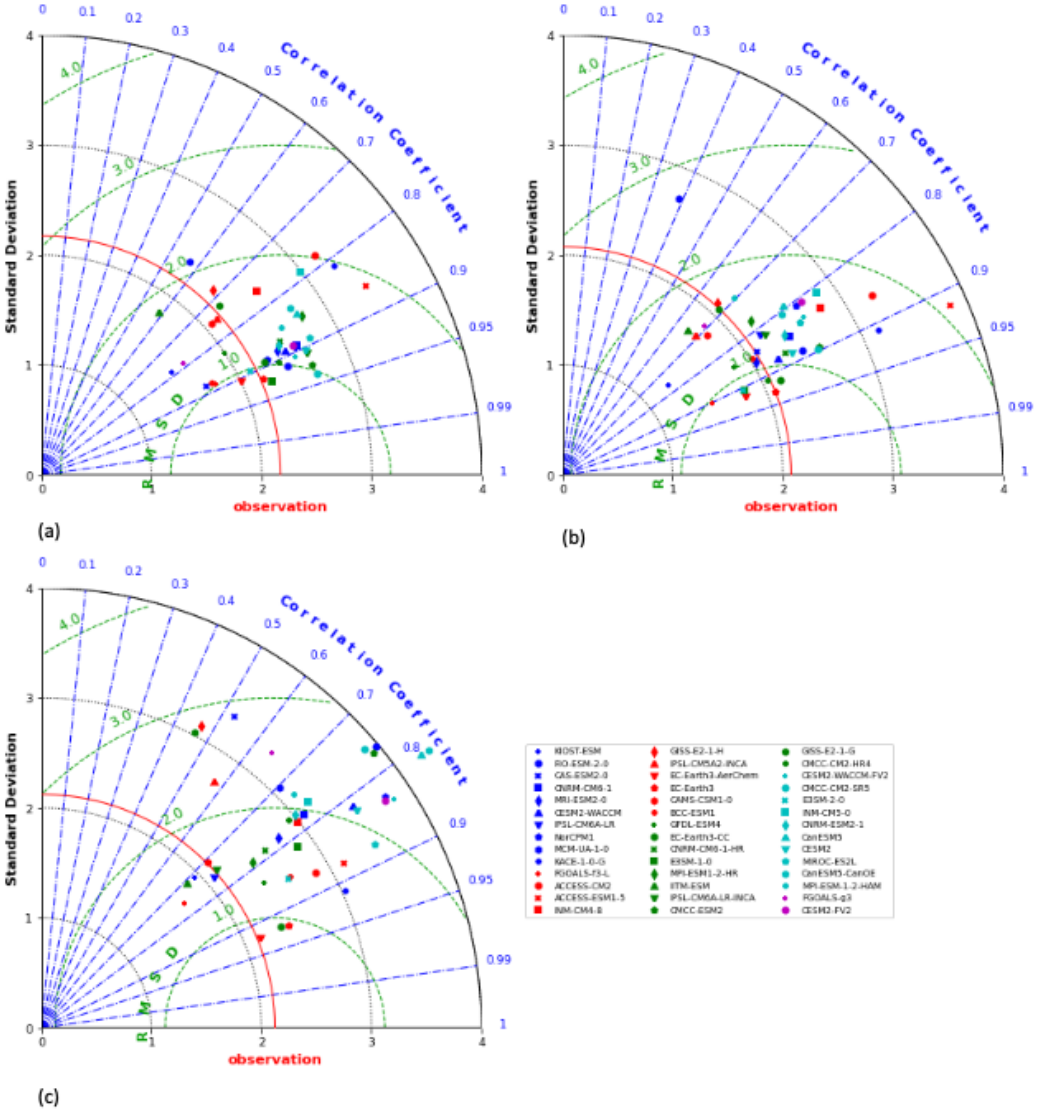


Figure 5.5: Taylor Diagram of 42 CMIP6 models in a. Downstream basin, b. Midstream basin, c. Upstream basin

altitude areas with high accuracy, while the upstream originates from the Tibetan Plateau and the upstream basin is widely covered with plateaus and basins, thus the model performs poorly.

There is a clear convergence of scores between the midstream and downstream simulations, meaning that models performing well in the downstream region tend to also perform well in the midstream region. However, for the upstream region, many models that perform poorly in the midstream and downstream simulations obtain higher scores.

Considering the model scores for all three regions collectively, the top-performing models for YRB simulations are EC-Earth3, EC-Earth3-AerChem, and EC-Earth3-CC. On the other hand, the worst-performing models are MCM-UA-1-0 and GISS-E2-1-H.

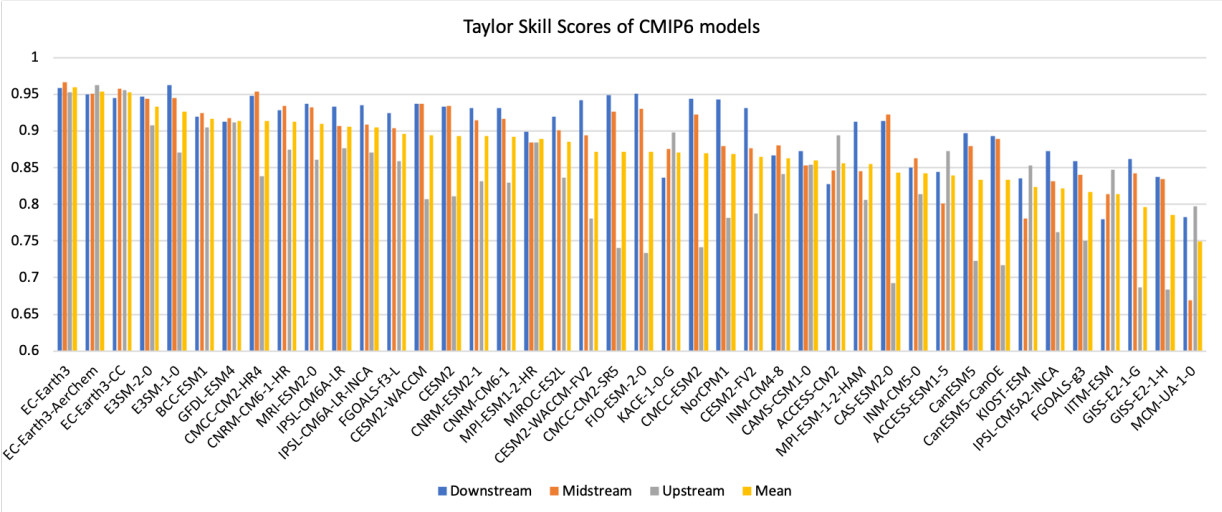


Figure 5.6: Taylor Skill Scores of 42 CMIP6 models in down, mid, upstream basin, sorting by highest value of mean scores

5.2. Moisture tracking in the YRB

5.2.1. Moisture tracking of ERA5 reanalysis data

The Figure 5.7 displays the cumulative average annual evaporation sources from 2005 to 2014. In this figure, we calculated the regional total rainfall for the YRB and evaporation over the study area, i.e., we tracked the source of evaporation for 89% of the water vapor. The remaining 11% may be due to overflow from the boundary, or errors caused by the model. It is evident that the primary origins of water vapor in the YRB are predominantly located in southern China, which are the dark blue and red regions, primarily within the basin itself. The red regions represent the Sichuan basin and the northern plains of Myanmar, both of which contribute significantly to the overall water vapor supply. The outline of the Himalayas is clearly visible to the west of the basin, as the terrain here makes it difficult for water vapor to cross, so the southern Himalayas, the Indo-Ganetic Plain of northern India, provide abundant water vapor.

Additionally, the western Pacific Ocean and the northern Indian Ocean serve as significant sources of evaporation. During the South Asian monsoon season, summer winds transport water vapor from the western coast of the Indian Ocean, including the Red Sea, Persian Gulf, and Arabian Sea, into the basin. Simultaneously, water vapor from the Bay of Bengal in the northern Indian Ocean also contributes to this supply. Furthermore, under the influence of the Southeast Asian summer winds, water vapor from the East China Sea and South China Sea in the western Pacific Ocean, encompassing regions like the Philippines and Thailand, becomes a vital source of water vapor in the YRB.

The arrow depicted in the Figure 5.7 represents the cumulative sum of moisture fluxes in the upper and lower layers, with the length of the arrow indicating the magnitude of the flux. Thus, the arrow indicates the direction of vapor movement. It is apparent that moisture from the western Indian Ocean follows the contour of the coastline, with some of it entering Africa while the rest circulates back to the Arabian Sea and the Bay of Bengal. Subsequently, it enters the YRB through the plains of Myanmar. The flux originating

from the sea off the Philippines in the western Pacific Ocean flows towards the interior of the YRB, partially entering the basin, but also partially moving along the coastline and blowing northward towards Japan. Consequently, despite its proximity to the YRB, the Pacific Ocean does not act as the primary source of moisture.

In general, water vapor flux within the Asian continent is typically minimal, meaning that land evaporation rarely contributes significantly to the water vapor supply in the YRB.

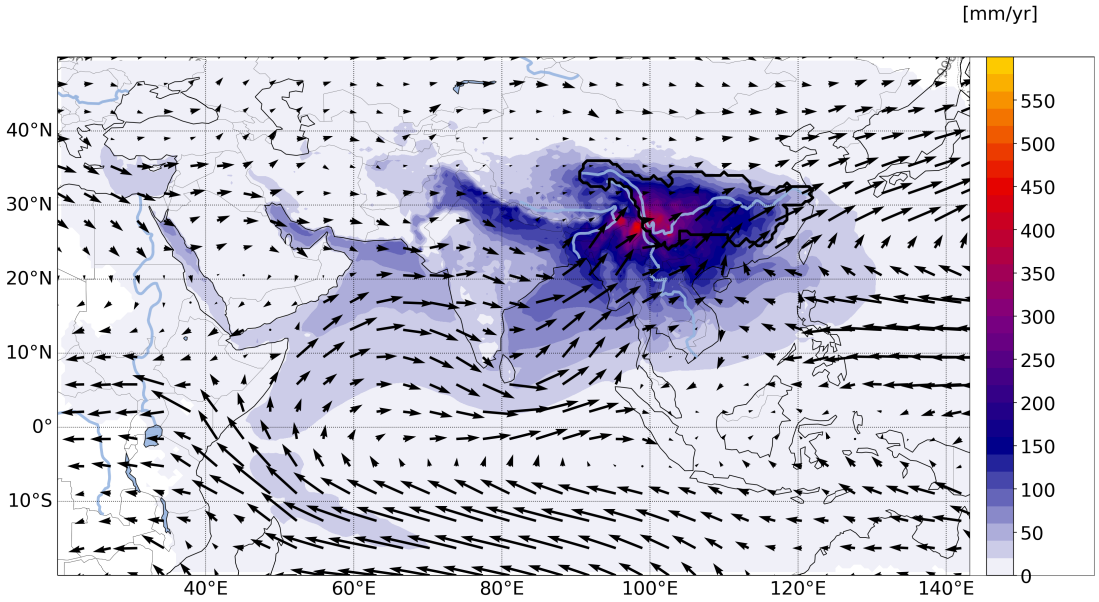


Figure 5.7: Cumulative moisture sources of YRB, data from ERA5 hourly data 2005-2014. The arrows are the total moisture fluxes in the top and bottom layers

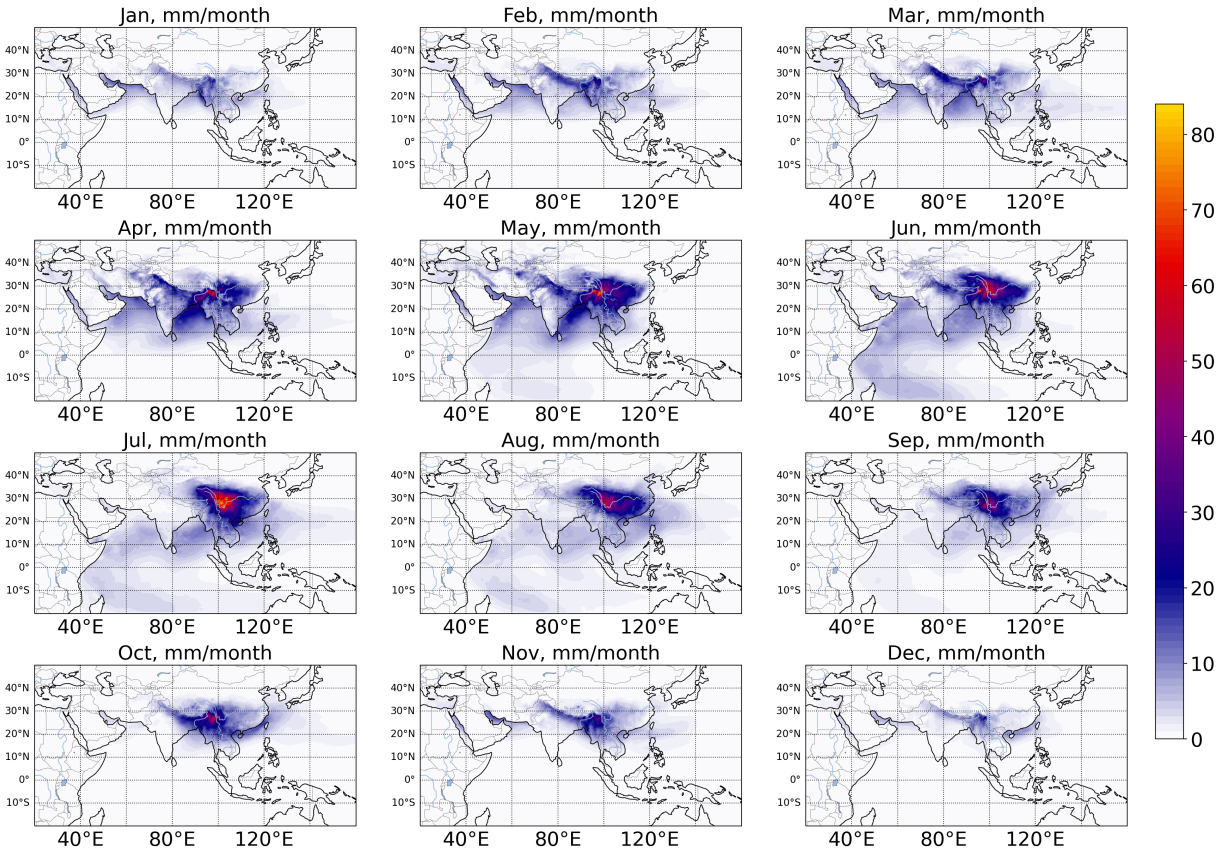


Figure 5.8: Monthly mean moisture sources of YRB in mm/month, data from ERA5 hourly data 2005-2014.

The monthly average evaporation sources for the 12 months of YBR are shown in the Figure 5.8. In different months, the evaporation sources are different. The primary source of rainfall in the YRB is the summer monsoon. The figure indicates that the months from April to September witness a huge influx of evaporated water into the basin, resulting in a vast source of evaporation during this period. In contrast, the supply of water vapor in other months is relatively low.

From January to March, the moisture originates from regions in East and South Asia that are close to the YRB, such as India, Myanmar, and southern China. Starting in April, the Indian Ocean becomes a prominent source, replenishing the basin with a substantial amount of moisture. As the month progresses, the South Asian summer winds from the Indian Ocean intensify, broadening the range of moisture sources. The waters west of the Indian Ocean and east of Africa in the Southern Hemisphere also contribute to the water vapor supply. The source and total amount of water vapor reach their peak during June to July, leading to the highest level of rainfall in the basin during this period.

Additionally, water vapor from the west coast of the Pacific Ocean plays an important role in July, August, and September, while its contribution weakens in other months. After August, as the monsoon weakens, moisture sources converge, and the total amount of moisture decreases, resulting in reduced rainfall within the YRB.

5.2.2. Moisture tracking of CMIP6 data

The Figure 5.9 shows the results of the CMIP6 EC-earth3 model for the years 2005-2014, tracking water vapor with WAM2layers. However, it is important to acknowledge that discrepancies exist in the data regarding the magnitude of the evaporation source, which appears to be nearly double the expected value. These discrepancies may be attributed to potential differences in units or time steps, or other issues that arise when applying the model to the CMIP6 model. In terms of evaporation patterns, the figure shows that the CMIP6 model is difficult to describe the topography. The Himalayas, a very obvious contour line in ERA5, is difficult for water vapor to cross this mountain range directly to YRB but crosses from the

southern plains, which can't be shown on the EC-earth3 model.

For this reason, the traced water vapor comes mainly from the Asian continent, including Central Asia and the interior of the YRB. In this model, the interior of the Yangtze River basin, especially the middle and lower reaches, contributes most of the evaporation, while in normal conditions, the southern Sichuan basin in the upper reaches should be the area contributing the most water vapor. The model likewise fails to describe another major source, the Indian Ocean. Again, because of the poor representation of the CMIP6 model in topography, the ocean, an important source of evaporation, is not simulated.

But the direction of moisture fluxes in the Indian Ocean is almost accurate as seen from the arrows. This is something worth exploring in that the trend around the Indian Ocean can be simulated with the same trend of moisture fluxes. However, looking around the Himalayas, we find that the moisture trend goes directly through the mountains into the basin.

The Figure 5.10 shows the monthly average evaporation sources of the Yangtze River basin simulated by EC-Earth3. It can be seen from the figure that the model also has more difficulty in simulating the temporal variation of water vapor sources. In terms of model performance, the model simulations are rather smaller in the monsoon months, i.e. April to September, and most of the water vapor comes from within the basin itself and apparently the water vapor inflow is even less. On the contrary, in the non-monsoon months, a large amount of water vapor evaporation and a wide range of evaporation sources are simulated. The evaporation pattern as a whole shows a band that extends from the YRB eastward and westward, mainly westward to the West Asia region. However, in reality, water vapor from West Asia does not enter the YRB unimpeded due to the presence of the Himalayas, so this is part of the reason for the error.

The water vapor in the East China Sea and South China Sea region can be simulated from April to October, but the evaporation pattern of the Indian Ocean is not simulated. From the graph, the Indian Ocean is not a major source of water vapor, which is not consistent with the actual situation. The role of South Asian summer winds is not simulated successfully, although the water vapor flux is in the right direction, however, the Indian Ocean does not provide enough moisture.

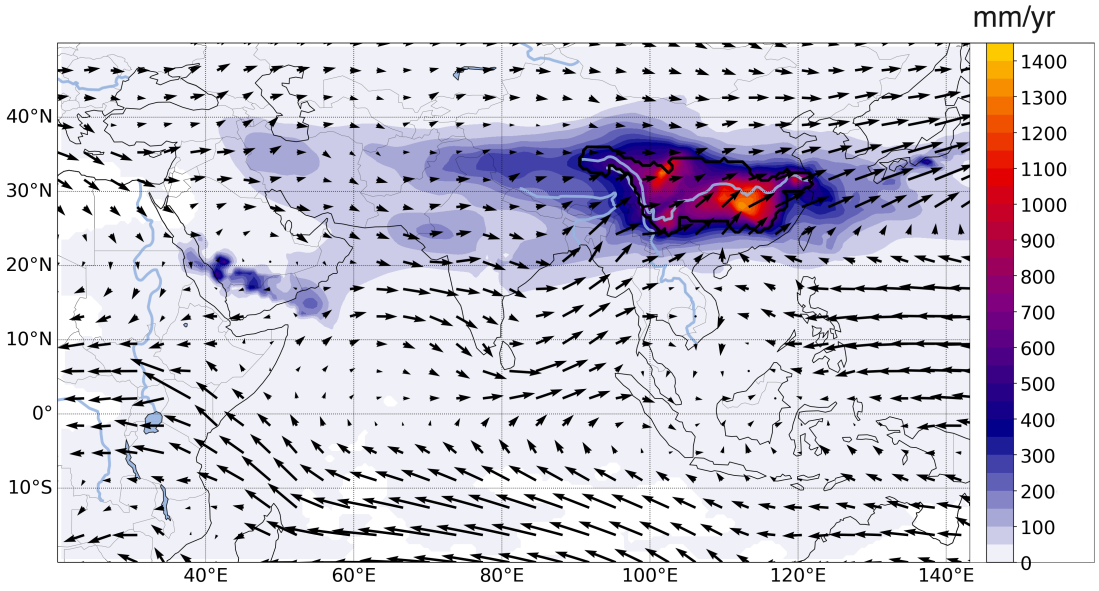


Figure 5.9: Cumulative moisture sources of YRB, data from EC-Earth3 daily data 2005-2014. The arrows are the total moisture fluxes in the top and bottom layers

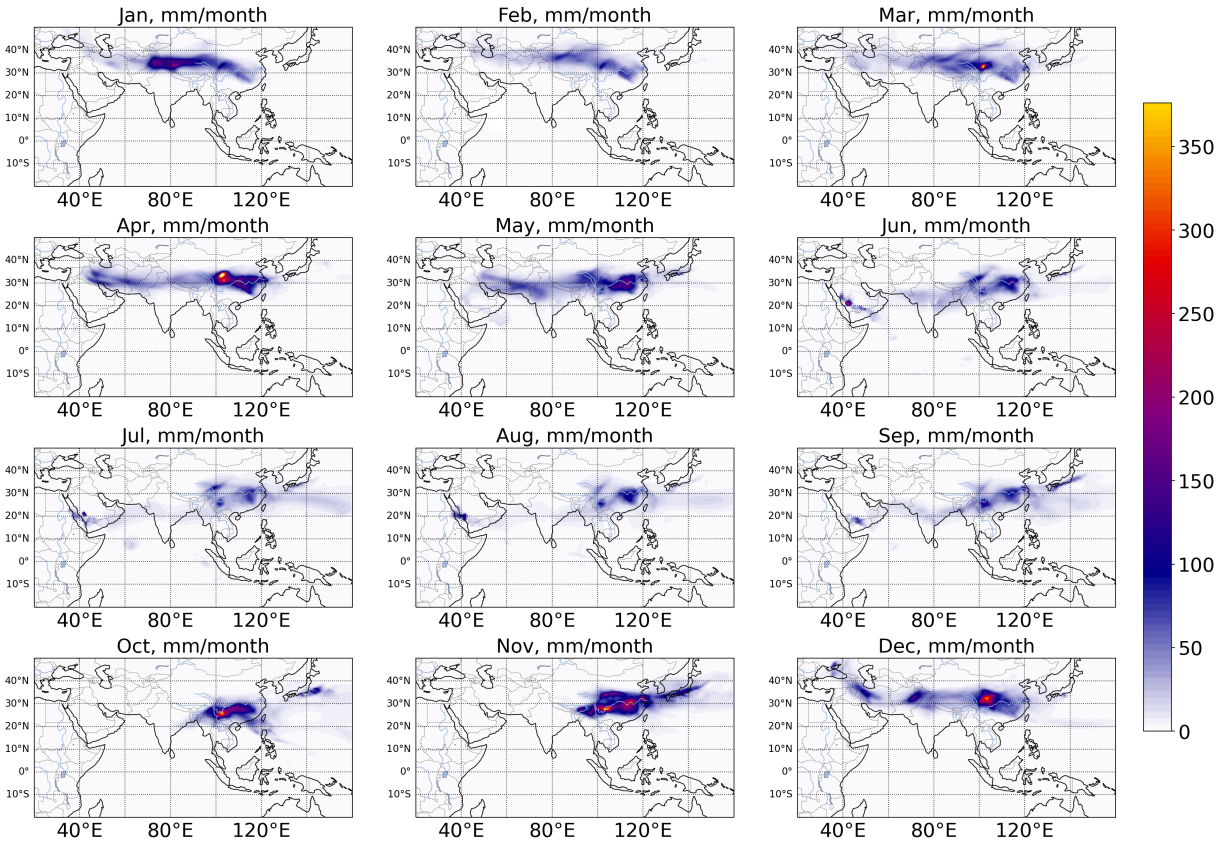


Figure 5.10: Monthly mean moisture sources of YRB in mm/month, data from EC-Earth3 daily data 2005-2014.

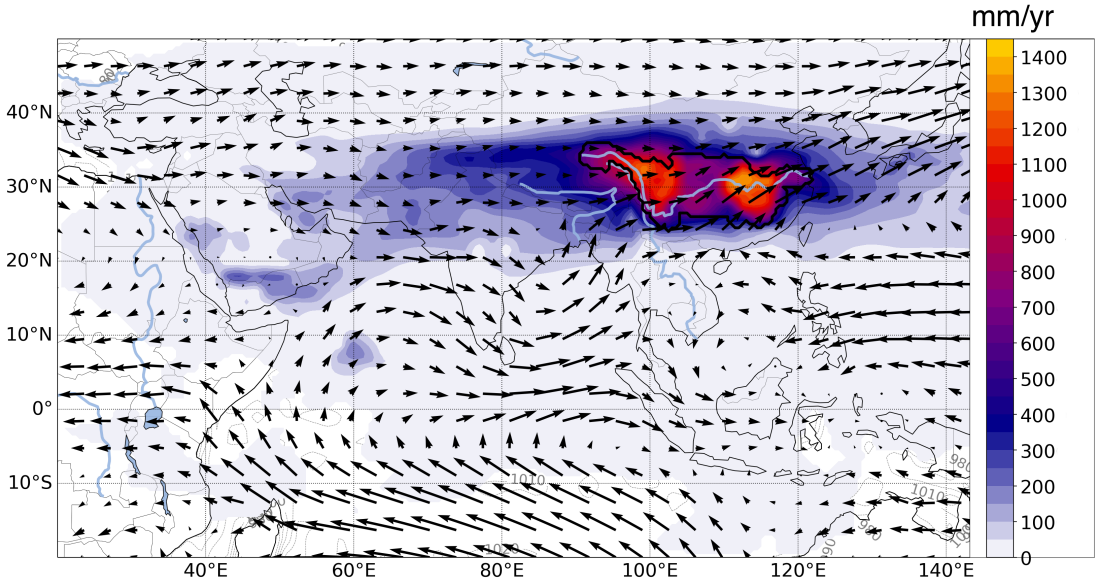


Figure 5.11: Cumulative moisture sources of YRB, data from MPI-ESM-HAM daily data 2005-2014. The arrows are the total moisture fluxes in the top and bottom layers

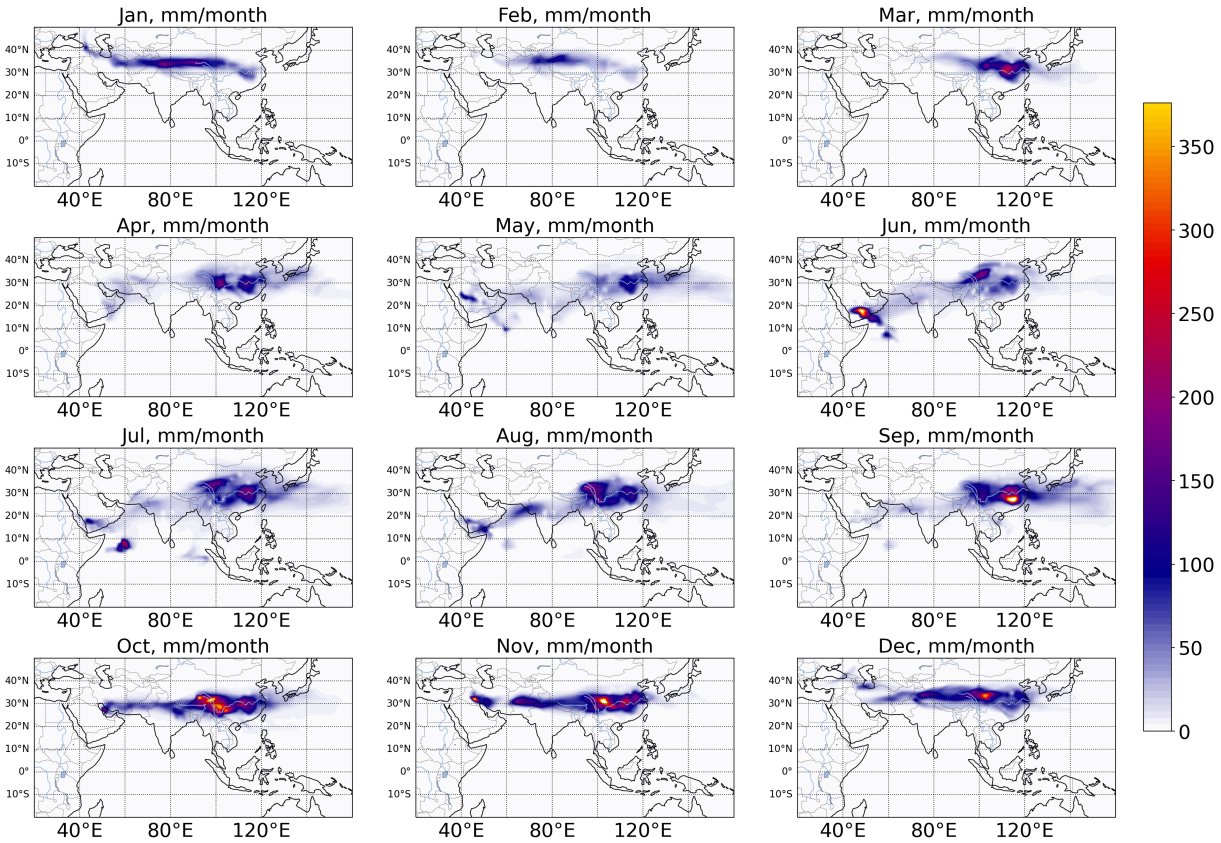


Figure 5.12: Monthly mean moisture sources of YRB in mm/month, data from MPI-ESM-HAM daily data 2005-2014.

Similarly, Figure 5.11 shows the simulation of the MPI-ESM-HAM model for water vapor sources in YRB 2005-2014. the overall spatial correlation coefficient of the MPI-ESM-HAM model with observed rainfall data is 0.79, which is lower than the 0.96 of EC-Earth3. this model overestimates the overall evaporation source more severely than the EC-Earth3 model, which simulates very large evaporation of water vapor in the upper and lower Yangtze River basin. Overall, this model, like EC-Earth3, fails to correctly simulate the role of terrain, ignoring the blockage of water vapor by the Himalayas and the role of the South Asian summer winds on water vapor in the Indian Ocean.

Like the EC-Earth3 model, the MPI-ESM-HAM can simulate the direction of moisture fluxes over the Pacific Ocean, but not the moisture fluxes over the Himalayas and the Tibetan Plateau.

Figure 5.12 shows the simulations of the MPI-ESM-HAM model for the 10-year monthly average source of water vapor. The results of this model simulation also show banding, and the overall model has higher water vapor than EC-Earth3 as a whole due to the overestimation of water vapor evaporation. The model simulates a more divergent source of water vapor than the non-monsoon months from April to September, which performs better.

The water vapor sources in the East China Sea and South China Sea on the west coast of the Pacific Ocean can be simulated from April to October, which means that the influence of the East Asian monsoon region can be successfully simulated. Unlike the dense zonal source of EC-Earth3, MPI-ESM-HAM shows a broader source, covering to the eastern and southern Asia and the northern Indian Ocean during the monsoon period, and remaining zonal in other months, ignoring the influence of the Tibetan Plateau.

Therefore, the reason for the simulation results of the CMIP6 model for the evaporation pattern of water vapor sources may be that the western part of Asia, which is not supposed to be a major source, contributes high amounts of moisture due to the inability to simulate the blocking of water vapor by the plateau and mountains. Both models overestimate the contribution of evaporation in the basin and underestimate the contribution of the Indian Ocean. In addition, the difference in sea paths is also difficult to represent in the model. Looking at Figure5.7 it can be seen that the difference between ocean and land is clearly visible in the results of the ERA5 reanalysis data, i.e., water vapor from the ocean is more abundant than that from land, and the sea-land boundary can be clearly seen, while the two models of CMIP6 Figure5.9 and Figure5.11 both failed to simulate correctly.

6

Discussion

6.1. Evaluation of CMIP6 models performance

Flood control in the Yangtze River Basin has always been an important concern in summer. Due to the influence of summer monsoon, summer rainfall in the Yangtze River Basin, especially in the middle and lower reaches of the concentrated rainfall. Therefore, it is very important to distinguish the peak months and the months with concentrated rainfall. The CMIP6 model performs well in the months with concentrated rainfall, i.e., the monsoon months, but performs poorly in the months with peak rainfall. For monsoon rainfall, as a whole, the upstream model performs worse than the downstream model. This is reflected in the low TSS and large errors in the simulation of monsoon onset and end times as well as monthly mean rainfall.

Topographic factors are an important contributor to the errors, and the complex terrain such as plateaus and basins in the upstream region is a great challenge for the GCMs. Most of the GCMs perform much worse in the upstream than in the middle and lower reaches and show a strong tendency to overestimate.

There are still some models that perform well in the upstream, such as those in the EC-Earth3 series. Although it also does not simulate the water vapor blockage in the Himalayan region, it produces fewer errors in the upstream compared to the other models and has a higher TSS. As a whole, EC-Earth3 performs optimally, both in the upper, middle and lower reaches. If only one model is chosen as the result to simulate, this would be a good choice.

Throughout the previous CMIP phases, many studies have shown that CMIP5 also overestimates rainfall in the upper reaches of the Yangtze River Basin.The CMIP6 model has been developed and proved to have good performance in many studies compared to the previous phases, but it still cannot solve the overestimation performance when there are large geographic gaps.

However, other researches find that the wind induced under-catch of snow is a common phenomenon in the source region of YRB and in the cold season [95]. The ratio of under-catch of precipitation can even beyond 200% in the central to northeastern Tibetan Plateau in the season [96], this may significantly increase the so-called ‘wet biases’ in the source region in relevant studies [4]. In addition, the ground observations are relatively sparse in high elevated source region. The interpolation from stations to surface could not fully capture the orographic lifting on precipitation change over the complex terrain [97]. So dry bias in the observations may also exacerbate the uncertainty in the results.

Another possibility of the trend of overestimation in the upstream of YRB is due to erroneous modeling of the Himalayas. Looking at the pathways of water vapor transport, it is evident that the isolation of the plateau region is not taken into account and the water vapor passes directly through the Himalayas into the upper reaches of the basin. This leads to more water vapor pooling and hence overestimation of rainfall. What’s more, the response of GCMs to complex topography would simulate higher precipitation during terrain uplift, due to condensation of water vapor from terrain elevation. Similar overestimation results appeared in the study of Pan et al.. They found that the model overestimation prediction mainly appeared in the high-elevation areas of the Yellow River Basin where the topography is highly variable [98].

As of now, CMIP is unable to handle the effects of complex terrain. However, it has been noted that the use of The Empirical Quantile Mapping (EQM) approach can be used to correct for model bias and

the averages using multi-model ensemble (MME) all perform better than any individual model [99]. In subsequent studies, the MME mean can be used as a reference value for monsoon rainfall.

For the tracking of water vapor sources by CMIP models, there are not yet enough studies in this direction, and this report only analyzes them in terms of patterns and water vapor trends. From the results, the water vapor evaporation pattern simulated by the CMIP model is more concentrated compared to that of the reanalysis data, and more water vapor from Central Asia is transported into the basin without considering the topographic factors, although this does not correspond to the actual situation. The contribution of the Indian Ocean is underestimated, perhaps because of the extraordinarily high contribution from Central Asia. The source of this error is unknown to us at this time; the CMIP model's water vapor tracking is using the same model as ERA5, and the error may come from CMIP's own data, or somewhere in the model we don't know so far. Since the tracking with WAM2layers uses a very large number of different variables, the source of the error is wide and overall more severe. But in spite of this, the pattern of the CMIP6 model for the simulation of water vapor we think is reasonable because the path of water vapor is almost the same as that of ERA5, which is also in line with the trend of monsoon wind direction. The direction of the moisture flux is almost the same as the ERA5 results except for the error in the Himalayan region.

6.2. Limitation of evaluation of CMIP6 models performance

6.2.1. Limitation of monsoon rainfall evaluation

In this study, I employed a relatively simple yet detailed method to evaluate 42 models by calculating the monthly average rainfall in the three streams of the YRB and comparing it with observed data. This approach provides a straightforward way to assess model performance, but it should be noted that it only evaluates average values and may not capture extreme rainfall events or drought conditions.

In addition, since I averaged the monthly data for all years, I assumed that the temporal distribution of rainfall did not change during the calculation period. However, in reality, the monthly average rainfall is constantly changing due to global climate change. For example, the phenomenon of El Niño-Southern Oscillation, ENSO, has an impact on both the amount and distribution of monsoon rainfall [100], which is not considered in this report.

The classical Taylor diagram and Taylor skill score were used for evaluation, which provide valuable insights. However, it is important to consider the limitations of these methods. Despite using multiple observation datasets to reduce errors, there can still be uncertainties in the observed data. Additionally, the Taylor diagram assumes a Gaussian distribution of the data, which may not hold true for all variables or regions. If the data exhibit a non-Gaussian distribution, the representation in the plot may not accurately reflect the model's skill. Moreover, they do not account for temporal variations and assume that the model's behavior remains consistent over time.

To address these limitations, it would be beneficial to explore additional evaluation methods that can capture the nonlinear relationships, temporal variations, and extreme events in the data. Alternative techniques, such as kernelized approaches [101], may provide a more comprehensive evaluation of model performance. Additionally, considering other statistical metrics that go beyond average values, such as quantile-based measures or extreme event indices, would provide a more comprehensive assessment of model skill.

6.2.2. Limitation of moisture tracking evaluation

Due to the requirement for comprehensive data for moisture tracking, models that provide complete and accurate variables are limited. In this study, out of the 42 models evaluated, only EC-Earth3 and MPI-ESM-1-2-HAM provided the necessary variables accurately. Therefore, the analysis in this study focused on the results from these two models.

Interestingly, despite their different shapes, the water vapor source patterns observed in EC-Earth3 and MPI-ESM-1-2-HAM did not differ significantly. It is not sure that the similarity in the water vapor patterns between these two models is a common characteristic among the CMIP6 models. However, it is important to note that this study did not investigate the performance of other CMIP6 models in terms of water vapor sources, so further research would be needed to determine if this is a widespread issue across the CMIP6 model ensemble.

The observed numerical differences between the ERA5 reanalysis data and the CMIP6 model data

suggest that there may be a need for enhancements in the WAM2layers code when applied to CMIP6 model data. The WAM2layers code is designed to track moisture sources and analyze water vapor fluxes, but it may encounter challenges or discrepancies when applied to different datasets, such as reanalysis data versus model data.

Indeed, the use of hourly data in ERA5 provides higher temporal resolution and potentially more accurate results for simulations. However, it comes with the trade-off of increased data volume and computational requirements, making it challenging to process large datasets and perform simulations efficiently.

To address this issue, reducing the temporal resolution by using larger time intervals could be a potential approach in future runs. By running simulations with coarser time intervals and comparing the results to those obtained from hourly resolution runs, it may be possible to strike a balance between computational efficiency and accuracy. This approach would help to reduce the runtime and computational effort required for processing ERA5 data while still obtaining meaningful results.

When it comes to CMIP6 data, the availability of daily data is more common compared to higher temporal resolutions. However, not all CMIP6 models provide all the necessary variables, such as dew point temperature, surface pressure, and specific humidity. In such cases, additional methods or calculations based on fundamental principles may be required to obtain or estimate these missing variables. However, it is important to acknowledge that introducing these additional steps may introduce additional uncertainties and errors in the analysis.

7

Conclusion

7.1. Research Questions

The research questions posed in Chapter 1 are repeated below for convenience. The primary objective of this study is to assess the performance of the CMIP6 models in simulating rainfall and moisture tracking in the Yangtze River basin, particularly during the monsoon period. The focus is on evaluating the models' ability to accurately reproduce the spatial and temporal patterns of rainfall and moisture transport associated with the monsoon season in the region.

Research Question 1

How do the models simulate the rainfall amount, seasonality, and spatial performance of historical rainfall in the YRB?

Research Question 2

Which models show higher accuracy in the YRB? Which models perform poorly?

Research Question 3

How accurate are the models simulating water vapor sources of historical data in the YRB?

By comparing the model simulations with observed data, I try to determine the strengths and weaknesses of the CMIP6 models in capturing the monsoon characteristics in the YRB. This evaluation includes examining the onset and duration of the monsoon season, the timing and intensity of peak rainfall, the spatial distribution of moisture sources, and the accuracy of moisture fluxes across the region.

The evaluation of model performance provides insights into the models' ability to represent the monsoon rainfall system and their potential for reliable projections of future monsoon behavior in the YRB. This assessment is crucial for understanding the reliability of the CMIP6 models and their utility in climate change studies, water resource management, and related decision-making processes in the region.

7.2. Main conclusions

In summary, this study aimed to evaluate the performance of the CMIP6 model in simulating rainfall and water vapor tracking in the Yangtze River basin, especially during the monsoon period. By analyzing the model simulation results and comparing them with the observed data, several key findings emerged.

First, the CMIP6 model exhibits different degrees of accuracy in simulating the monsoon season in different regions of the Yangtze River basin. The downstream and midstream regions performed relatively well, accurately reflecting the timing and intensity of monsoon rainfall, while the upstream region was simulated with lower accuracy. This discrepancy suggests that the model faces challenges in accurately representing the complex dynamics of the monsoon system, especially in areas with different rainfall patterns and topography.

The performance of the models in simulating monsoon rainfall is evaluated using Taylor plots and Taylor skill scores to provide a comprehensive overview of the strengths and weaknesses of the models. The downstream and midstream regions perform better, with models such as EC-Earth3 and E3SM1-0 excelling. In contrast, models like IITM-ESM and MCM-UA-1-0 exhibit poor simulation accuracy. However, the assessment is based on average rainfall values, which may not fully reflect extreme events or temporal variations.

In addition, the evaluation of the moisture tracking model shows the limitations of the CMIP6 model. While some models successfully replicate the general direction of water vapor fluxes, they often do not accurately represent the sources and paths of water vapor, leading to discrepancies with observed data. Notably, these models have difficulty explaining the effects of geographic features such as the Himalayas, leading to biases in water vapor transport patterns and mismatches with actual observational data.

Despite the progress made by the CMIP6 model, some challenges remain. The limited availability of complete data, especially at higher temporal resolutions, imposes limitations on the accuracy of the simulations. In addition, discrepancies in numerical and evaporation patterns appear when calling WAM2layers during water vapor tracking, which implies that more work needs to be carried out in terms of refinement of the WAM2layers code to the CMIP6 model, as well as improving the accuracy of the CMIP6 model itself.

In summary, this study contributes to our understanding of the performance of the CMIP6 model in simulating monsoon rainfall and moisture tracking in the Yangtze River basin. The findings highlight the need for further improvements in capturing the spatial and temporal variability of the monsoon system and associated water vapor transport in CMIP. Addressing these challenges will improve the reliability of the model and its applicability in climate change assessment, water resources management, and decision-making processes in the Yangtze River basin and similar regions.

8

Recommendations

Based on the evaluation of CMIP6 model performance in simulating monsoon rainfall and moisture tracking in the Yangtze River basin, several recommendations can be derived.

Rec 1
The EC-Earth series, including models like EC-Earth3, EC-Earth3-AerChem, and EC-Earth3-CC, have consistently demonstrated their ability to accurately capture the timing, duration, and intensity of the monsoon season in the upstream, midstream, and downstream regions of the YRB. Their skill scores, Taylor diagram analysis, and Taylor skill scores consistently rank them among the top-performing models.

Thus the data provided by these models can be effectively used to predict different future emission scenarios in the Yangtze River basin. Since muti-model integration generally perform better than individual models [51], by increasing the weight of high-quality output from the EC-Earth series in muti-model integration, researchers and policy makers can gain valuable insights into potential changes in monsoon rainfall patterns and their implications for water resource management, agriculture, flood forecasting, and other socioeconomic sectors.

Rec 2
Although the CMIP6 module is much optimized in terms of terrain representation compared to CMIP5, there is still a need to enhance the representation of monsoon dynamics, taking into account the complex interactions between atmospheric circulation, land-sea interactions, and topographic features, given the challenges observed in accurately simulating the monsoon season in different regions of the Yangtze River basin. This may involve incorporating more realistic representations of key features (e.g., the Himalayas).

In addition, although the CMIP6 model performs relatively well for rainfall simulations, limitations were identified in water vapor tracking, particularly in capturing the sources and paths of water vapor, highlighting the need for model improvements in representing water vapor transport mechanisms. Models should be refined to better account for geographic features, such as mountain ranges, which may significantly affect moisture transport patterns. In addition, efforts should be made to ensure accurate representation of water vapor fluxes from important water vapor sources such as the Indian Ocean.

Rec 3
To improve future monsoon rainfall predictions, it is recommended to further enhance the temporal resolution of climate models. While daily data has been utilized in the CMIP6 models, exploring the potential benefits of using higher temporal resolution, such as hourly data, could provide more accurate and detailed information on monsoon dynamics. Although hourly data requires significant computational resources and storage capacity, advancements in computing technologies and data management systems should be leveraged to enable the incorporation of finer temporal scales in future model simulations.

Acknowledgements

提笔落到论文的致谢，本不打算写太多，但还是想给两年的硕士生涯画下一个句号。

两年求学生涯正赶上了最坏的年份，疫情将我和家人封锁在大陆的两端，小小的屏幕是唯一的联系。两年内发生了许多事，总体来说大都是不幸的事。时而看着临别前的合照喉咙一梗，亲人的过世是一场突如其来的暴雨，而过后便是永远的潮湿。时间匆匆忙忙推着人前行，总觉得什么都没有来得及做就已经长大了。

两个月前看到武大朋友们的毕业典礼，她们穿着学士服抱着花在行政楼前拍照，和两年前本科毕业时一样笑得腼腆又灿烂。恍惚觉得两年时间在我身上停止，内心还以为自己是刚刚毕业的样子，可看到家人眉间的折痕和鬓角的白发，才知道对于我来说白驹过隙一般的两年是如此漫长。

人总是理所应当以为世界就是自己看到的那样，然后忽然看到了不同文化，认识到不同观念，才会豁然开朗。感谢在荷兰认识的中国朋友们，刘译蔓、刘欣遇、魏铭楷、解钰坤、田河、吉雨宁，还有一起玩桌游的朋友们金宇辉、张建彤、辛越捷、楼星皓、贲文浩（以上排名不分先后），和他们一起看到了世界上不同的景色，在异国他乡抱团取暖。还有毕玉涵同学，非常巧合我们能在欧洲相遇。

当然还要感谢远在国内的宁馨儿、郭舟、陈元博，看到她们分享的日常我羡慕又高兴，最终她们都去了想去的地方，成为想成为的人，都有光明的未来，希望我也一样。

最珍重也是最想要感谢的是我的父母，寸草之心难报光辉雨露之恩，感谢他们对我的支持与鼓励，让飘摇的孤舟心中有一个温暖的港湾，无论漂泊多久都会有灯塔在海岸等待归航。

I would like to express my sincere gratitude to my supervisor, Ruud van der Ent, for his dedicated guidance throughout my thesis journey. His patient clarification of my doubts and insightful direction steered me towards the right path. I am also deeply indebted to Peter Kalverla and Vincent de Feiter, whose profound understanding of WAM2layers saved me from many pitfalls and corrected numerous errors.

I extend my heartfelt appreciation to Martine Rutten for her invaluable suggestions in structuring and composing my thesis. Under her guidance, I acquired essential skills in scholarly writing. Special thanks go to Ronald Ligteringen from EWI, who provided substantial assistance when I faced perplexity in my exploration of Delft blue, resolving a significant hurdle.

I would also like to extend my gratitude to my dear friend Nabil. His unwavering support and genuine kindness have been truly heartwarming. I am profoundly appreciative of the times we've shared together. Our friendship, I believe, will remain strong even across the vast expanse of continents and oceans.

思绪万千，不知所言。

不负年华，感谢相遇。

前路尚远，切莫蹉跎。

二〇二三年 八月十八日

于 代尔夫特

References

[1] Charles E. Greer et al. *Yangtze River*. 2023. URL: https://www.britannica.com/place/Yangtze-River (visited on 02/07/2023).

[2] Anonymous. *Basic information on the Yangtze River Basin and its position in China's economic development*. 2007. URL: https://zh.wikipedia.org/zh-hans/%E9%95%BF%E6%B1%9F%E6%B5%81%E5%9F%9F (visited on 02/07/2023).

[3] Anonymous. *Health Yangtze River Q and A*. 2007. URL: https://zh.wikipedia.org/zh-hans/%E9%95%BF%E6%B1%9F%E6%B5%81%E5%9F%9F (visited on 02/07/2023).

[4] Ying Li et al. "Evaluation of precipitation in CMIP6 over the Yangtze River Basin". In: *Atmospheric Research* 253 (2021), p. 105406. DOI: https://doi.org/10.1016/j.atmosres.2020.105406. URL: https://www.sciencedirect.com/science/article/pii/S0169809520313430.

[5] Zhiping Xu. *Hydrological overview of the Yangtze River basin in China*. 2016. URL: https://www.osgeo.cn/post/c5817 (visited on 02/07/2023).

[6] Jiangfeng Wei et al. "Water vapor sources for Yangtze River Valley rainfall: Climatology, variability, and implications for rainfall forecasting". In: *Journal of Geophysical Research: Atmospheres* 117.D5 (2012).

[7] Ning Wang et al. "Quantitative diagnosis of moisture sources and transport pathways for summer precipitation over the mid-lower Yangtze River Basin". In: *Journal of Hydrology* 559 (2018), pp. 252–265.

[8] Xiangde Xu et al. "An important mechanism sustaining the atmospheric" water tower" over the Tibetan Plateau". In: *Atmospheric Chemistry and Physics* 14.20 (2014), pp. 11287–11295.

[9] Bin Chen et al. "Identifying and contrasting the sources of the water vapor reaching the subregions of the Tibetan Plateau during the wet season". In: *Climate Dynamics* 53 (2019), pp. 6891–6907.

[10] A. Lin et al. "Tectonic relationship between the course change of the Yangtze River and Indo-Asia collision". In: *Structural Geology: New Research* (Jan. 2008), pp. 49–62.

[11] Anonymous. *Geographic Overview of the Yangtze River Basin*. 2010. URL: http://www.ziyexing.com/files-7/geography/Geography_inf_11.htm (visited on 02/07/2023).

[12] Anonymous. *The world's largest hydropower plant: The Three Gorges Dam*. 2014. URL: http://www.hydropower.org.cn/showNewsDetail.asp?nsId=14389 (visited on 02/07/2023).

[13] Mohammed Magdy Hamed et al. "Inconsistency in historical simulations and future projections of temperature and rainfall: A comparison of CMIP5 and CMIP6 models over Southeast Asia". In: *Atmospheric Research* 265 (2022), p. 105927. DOI: https://doi.org/10.1016/j.atmosres.2021.105927. URL: https://www.sciencedirect.com/science/article/pii/S016980952100483X.

[14] Dennis L. Hartmann. "Chapter 11 - Global Climate Models". In: *Global Physical Climatology (Second Edition)*. Ed. by Dennis L. Hartmann. Second Edition. Boston: Elsevier, 2016, pp. 325–360. DOI: https://doi.org/10.1016/B978-0-12-328531-7.00011-6. URL: https://www.sciencedirect.com/science/article/pii/B9780123285317000116.

[15] Giovanni Di Virgilio et al. "Selecting CMIP6 GCMs for CORDEX dynamical downscaling: model performance, independence, and climate change signals". In: *Earth's Future* 10.4 (2022).

[16] Huopo Chen et al. "Projection and uncertainty analysis of global precipitation-related extremes using CMIP5 models". In: *International journal of climatology* 34.8 (2014), pp. 2730–2748.

[17] WRCP. 2022. URL: https://www.wcrp-climate.org/wgcm-cmip/wgcm-cmip6 (visited on 10/14/2022).

[18] Jungang Gao et al. “Uncertainty of hydrologic processes caused by bias-corrected CMIP5 climate change projections with alternative historical data sources”. In: *Journal of Hydrology* 568 (2019), pp. 551–561.

[19] Huopo Chen et al. “Comparison of CMIP6 and CMIP5 models in simulating climate extremes”. In: *Science Bulletin* 65.17 (2020), pp. 1415–1418.

[20] Lei Wang et al. “Projection of precipitation extremes in China’s mainland based on the statistical downscaled data from 27 GCMs in CMIP6”. In: *Atmospheric Research* 280 (2022), p. 106462.

[21] Mohammed Magdy Hamed et al. “Inter-comparison of historical simulation and future projections of rainfall and temperature by CMIP5 and CMIP6 GCMs over Egypt”. In: *International Journal of Climatology* 42.8 (2022), pp. 4316–4332.

[22] Jose Carlos González-Hidalgo et al. “A new tool for monthly precipitation analysis in Spain: MOPREDAS database (monthly precipitation trends December 1945–November 2005)”. In: *International Journal of Climatology* 31.5 (2011), pp. 715–731.

[23] P. H. Hrudya et al. “A review on the Indian summer monsoon rainfall, variability and its association with ENSO and IOD”. In: *Meteorology and Atmospheric Physics* 133.1 (2020), pp. 1–14. DOI: 10.1007/s00703-020-00734-5.

[24] Renguang Wu. “Relationship between Indian and East Asian summer rainfall variations”. In: *Advances in Atmospheric Sciences* 34 (2017), pp. 4–15.

[25] K-M. Lau et al. “Seasonal and Intraseasonal Climatology of Summer Monsoon Rainfall over East Asia”. In: *Monthly Weather Review* 116.1 (1988), pp. 18–37. DOI: https://doi.org/10.1175/1520-0493(1988)116<0018:SAICOS>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/mwre/116/1/1520-0493_1988_116_0018_saicos_2_0_co_2.xml.

[26] Zhang Renhe. “Relations of water vapor transport from Indian monsoon with that over East Asia and the summer rainfall in China”. In: *Advances in Atmospheric Sciences* 18.5 (2001), pp. 1005–1017.

[27] Lun Dai et al. “Summer monsoon rainfall patterns and predictability over Southeast China”. In: *Water Resources Research* 56.2 (2020). DOI: 10.1029/2019wr025515.

[28] Bin Wang et al. “Multi-scale climate variability of the South China Sea monsoon: A review”. In: *Dynamics of Atmospheres and Oceans* 47.1 (2009). The South China Sea and its impact on climate, pp. 15–37. DOI: https://doi.org/10.1016/j.dynatmoce.2008.09.004. URL: https://www.sciencedirect.com/science/article/pii/S0377026508000572.

[29] Tian-Jun Zhou. “Atmospheric water vapor transport associated with typical anomalous summer rainfall patterns in China”. In: *Journal of Geophysical Research* 110.D8 (2005). DOI: 10.1029/2004jd005413.

[30] Renguang Wu et al. “Evolution of ENSO-related rainfall anomalies in East Asia”. In: *Journal of Climate* 16.22 (2003), pp. 3742–3758.

[31] Zhiwei Zhu et al. “A New Paradigm for Continental U.S. Summer Rainfall Variability: Asia–North America Teleconnection”. In: *Journal of Climate* 29.20 (2016), pp. 7313–7327. DOI: https://doi.org/10.1175/JCLI-D-16-0137.1. URL: https://journals.ametsoc.org/view/journals/clim/29/20/jcli-d-16-0137.1.xml.

[32] Prabir Kumar Das et al. “Spatial analysis of temporal trend of rainfall and rainy days during the Indian Summer Monsoon season using daily gridded (0.5° × 0.5°) rainfall data for the period of 1971-2005”. In: *Meteorological Applications* 21.3 (2013), pp. 481–493. DOI: 10.1002/met.1361.

[33] Yue Li et al. “Evaluation of monsoon seasonality and the tropospheric biennial oscillation transitions in the CMIP models”. In: *Geophysical research letters* 39.20 (2012).

[34] Patrick W. Keys et al. “Revealing invisible water: Moisture recycling as an ecosystem service”. In: *PLOS ONE* 11.3 (2016). DOI: 10.1371/journal.pone.0151993.

[35] Heinz Lettau et al. “Amazonia’s Hydrologic Cycle and the Role of Atmospheric Recycling in Assessing Deforestation Effects”. In: *Monthly Weather Review* 107.3 (1979), pp. 227–238. DOI: https://doi.org/10.1175/1520-0493(1979)107<0227:AHCATR>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/mwre/107/3/1520-0493_1979_107_0227_ahcatr_2_0_co_2.xml.

[36] Justin E Bagley et al. “Effects of land cover change on moisture availability and potential crop yield in the world’s breadbaskets”. In: *Environmental Research Letters* 7.1 (Jan. 2012), p. 014009. DOI: 10.1088/1748-9326/7/1/014009. URL: https://dx.doi.org/10.1088/1748-9326/7/1/014009.

[37] D.V. Spracklen et al. “The effects of tropical vegetation on rainfall”. In: *Annual Review of Environment and Resources* 43.1 (2018), pp. 193–218. DOI: 10.1146/annurev-environ-102017-030136.

[38] Jiangfeng Wei et al. “Where Does the Irrigation Water Go? An Estimate of the Contribution of Irrigation to Precipitation Using MERRA”. In: *Journal of Hydrometeorology* 14.1 (2013), pp. 275–289. DOI: https://doi.org/10.1175/JHM-D-12-079.1. URL: https://journals.ametsoc.org/view/journals/hydr/14/1/jhm-d-12-079_1.xml.

[39] Chi Zhang. “Moisture sources for precipitation in Southwest China in summer and the changes during the extreme droughts of 2006 and 2011”. In: *Journal of Hydrology* 591 (2020), p. 125333. DOI: https://doi.org/10.1016/j.jhydrol.2020.125333. URL: https://www.sciencedirect.com/science/article/pii/S0022169420307939.

[40] Qu-cheng Chu et al. “Determination of the major moisture sources of cumulative effect of torrential rain events during the preflood season over South China using a Lagrangian particle model”. In: *Journal of Geophysical Research: Atmospheres* 122.16 (2017), pp. 8369–8382. DOI: https://doi.org/10.1002/2016JD026426. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016JD026426. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD026426.

[41] RJ Van der Ent et al. “Should we use a simple or complex model for moisture recycling and atmospheric moisture tracking?” In: *Hydrology and Earth System Sciences* 17.12 (2013), pp. 4869–4884.

[42] Obbe A. Tuinenburg et al. *Tracking the global flows of atmospheric moisture and associated uncertainties*. May 2020. URL: https://hess.copernicus.org/articles/24/2419/2020/#section13.

[43] Veronika Eyring et al. “Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization”. In: *Geoscientific Model Development* 9.5 (2016), pp. 1937–1958.

[44] Anja Katzenberger et al. “Robust increase of Indian monsoon rainfall and its variability under future warming in CMIP6 models”. In: *Earth System Dynamics* 12.2 (2021), pp. 367–386.

[45] A Gusain et al. “Added value of CMIP6 over CMIP5 models in simulating Indian summer monsoon rainfall”. In: *Atmospheric Research* 232 (2020), p. 104680.

[46] B Abida Choudhury et al. “Evolution of the Indian summer monsoon rainfall simulations from CMIP3 to CMIP6 models”. In: *Climate Dynamics* (2021), pp. 1–26.

[47] Kavirajan Rajendran et al. “Simulation of Indian summer monsoon rainfall, interannual variability and teleconnections: evaluation of CMIP6 models”. In: *Climate Dynamics* 58.9-10 (2022), pp. 2693–2723.

[48] Adway Mitra. “A comparative study on the skill of CMIP6 models to preserve daily spatial patterns of monsoon rainfall over India”. In: *Frontiers in Climate* 3 (2021), p. 654763.

[49] Zafar Iqbal et al. “Evaluation of CMIP6 GCM rainfall in mainland Southeast Asia”. In: *Atmospheric Research* 254 (2021), p. 105525.

[50] Yuhan Yan et al. “Annual cycle of East Asian precipitation simulated by CMIP6 models”. In: *Atmosphere* 12.1 (2020), p. 24.

[51] Xiaoge Xin et al. “Comparison of CMIP6 and CMIP5 simulations of precipitation in China and the East Asian summer monsoon”. In: *International Journal of Climatology* 40.15 (2020), pp. 6423–6440.

[52] Dabang Jiang et al. “Differences between CMIP6 and CMIP5 models in simulating climate over China and the East Asian monsoon”. In: *Advances in Atmospheric Sciences* 37 (2020), pp. 1102–1118.

[53] Hui Tao et al. “Assessment of CMIP3 climate models and projected changes of precipitation and temperature in the Yangtze River Basin, China”. In: *Climatic Change* 111 (2012), pp. 737–751.

[54] Fengge Su et al. “Evaluation of the global climate models in the CMIP5 over the Tibetan Plateau”. In: *Journal of climate* 26.10 (2013), pp. 3187–3208.

[55] Qiaohong Sun et al. “Projected changes in temperature and precipitation in ten river basins over China in 21st century”. In: *International Journal of Climatology* 35.6 (2015), pp. 1125–1141.

[56] Zaitao Pan et al. “Current and future precipitation extremes over Mississippi and Yangtze River basins as simulated in CMIP5 models”. In: *Journal of Earth Science* 27 (2016), pp. 22–36.

[57] SE Perkins et al. “Evaluation of the AR4 climate models‘ simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions”. In: *Journal of climate* 20.17 (2007), pp. 4356–4376.

[58] William KM Lau et al. “A Canonical Response in Rainfall Characteristics to Global Warming: Projections by IPCC CMIP5 Models”. In: *National Taiwan University International Science Conference on Climate Change: Multidecadal and Beyond*. GSFC. ABS. 7242.2012. 2012.

[59] Zhihong Jiang et al. “Extreme precipitation indices over China in CMIP5 models. Part I: Model evaluation”. In: *Journal of Climate* 28.21 (2015), pp. 8603–8619.

[60] Seung-Ki Min et al. “A Bayesian approach to climate model evaluation and multi-model averaging with an application to global mean surface temperatures from IPCC AR4 coupled climate models”. In: *Geophysical Research Letters* 33.8 (2006).

[61] Jagadish Shukla et al. “Climate model fidelity and projections of climate change”. In: *Geophysical Research Letters* 33.7 (2006).

[62] Zafar Iqbal et al. “Evaluation of global climate models for precipitation projection in sub-Himalaya region of Pakistan”. In: *Atmospheric Research* 245 (2020), p. 105061.

[63] Saleem A Salman et al. “Selection of climate models for projection of spatiotemporal changes in temperature of Iraq with uncertainties”. In: *Atmospheric research* 213 (2018), pp. 509–522.

[64] Stephan Thober et al. “Robust ensemble selection by multivariate evaluation of extreme precipitation and temperature characteristics”. In: *Journal of Geophysical Research: Atmospheres* 119.2 (2014), pp. 594–613.

[65] Weilin Chen et al. “Probabilistic projections of climate change over China under the SRES A1B scenario using 28 AOGCMs”. In: *Journal of Climate* 24.17 (2011), pp. 4741–4756.

[66] Tianyun Dong et al. “Evaluation of extreme precipitation over Asia in CMIP6 models”. In: *Climate Dynamics* 57.7-8 (2021), pp. 1751–1769.

[67] Karl E Taylor. “Summarizing multiple aspects of model performance in a single diagram”. In: *Journal of geophysical research: atmospheres* 106.D7 (2001), pp. 7183–7192.

[68] Ryo Mizuta et al. “Climate simulations using MRI-AGCM3. 2 with 20-km grid”. In: □□□□. □ *2* □ 90.0 (2012), pp. 233–258.

[69] Hassen Babaousmail et al. “Evaluation of the performance of CMIP6 models in reproducing rainfall patterns over North Africa”. In: *Atmosphere* 12.4 (2021), p. 475.

[70] Peter A. Rochford. *SkillMetrics: A Python package for calculating the skill of model predictions against observations*. 2016. URL: `https://github.com/PeterRochford/SkillMetrics`.

[71] V. Eyring et al. “Earth System Model Evaluation Tool (ESMValTool) v2.0 – an extended set of large-scale diagnostics for quasi-operational and comprehensive evaluation of Earth system models in CMIP”. In: *Geoscientific Model Development* 13.7 (2020), pp. 3383–3438. DOI: 10.5194/gmd-13-3383-2020. URL: https://gmd.copernicus.org/articles/13/3383/2020/.

[72] A. Lauer et al. “Earth System Model Evaluation Tool (ESMValTool) v2.0 – diagnostics for emergent constraints and future projections from Earth system models in CMIP”. In: *Geoscientific Model Development* 13.9 (2020), pp. 4205–4228. DOI: 10.5194/gmd-13-4205-2020. URL: https://gmd.copernicus.org/articles/13/4205/2020/.

[73] Sanjiv Kumar et al. “Evaluation of temperature and precipitation trends and long-term persistence in CMIP5 twentieth-century climate simulations”. In: *Journal of Climate* 26.12 (2013), pp. 4168–4185.

[74] Ali Mehran et al. “Evaluation of CMIP5 continental precipitation simulations relative to satellite-based gauge-adjusted observations”. In: *Journal of Geophysical Research: Atmospheres* 119.4 (2014), pp. 1695–1707.

[75] Zhu Liu et al. “The capability of CMIP6 models on seasonal precipitation extremes over Central Asia”. In: *Atmospheric Research* 278 (2022), p. 106364. DOI: https://doi.org/10.1016/j.atmosres.2022.106364. URL: https://www.sciencedirect.com/science/article/pii/S0169809522003507.

[76] Zengyun Hu et al. “Evaluation of three global gridded precipitation data sets in central Asia based on rain gauge observations”. In: *International Journal of Climatology* 38.9 (2018), pp. 3475–3493.

[77] Qin Jiang et al. “Evaluation of the ERA5 reanalysis precipitation dataset over Chinese Mainland”. In: *Journal of Hydrology* 595 (2021), p. 125660.

[78] Yan Yang et al. “Evaluation of gridded precipitation data for driving SWAT model in area upstream of three gorges reservoir”. In: *PLoS One* 9.11 (2014), e112725.

[79] Sahar Hadi Pour et al. “Physical-empirical models for prediction of seasonal rainfall extremes of Peninsular Malaysia”. In: *Atmospheric Research* 233 (2020), p. 104720.

[80] Akiyo Yatagai et al. “APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges”. In: *Bulletin of the American Meteorological Society* 93.9 (2012), pp. 1401–1415.

[81] Akiyo Yatagai et al. “A 44-year daily gridded precipitation dataset for Asia based on a dense network of rain gauges”. In: *Sola* 5 (2009), pp. 137–140.

[82] Udo Schneider et al. *GPCC Full Data Reanalysis Version 6.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data.* 2011. URL: 10.5676/DWD_GPCC/FD_M_V7_050.

[83] J. Muñoz Sabater. *ERA5-Land monthly averaged data from 1981 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. 2019. URL: 10.24381/cds.68d2bb3.

[84] J. Muñoz Sabater. *ERA5-Land monthly averaged data from 1950 to 1980. Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. 2021. URL: 10.24381/cds.68d2bb3.

[85] Ian Harris et al. “Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset”. In: *Scientific data* 7.1 (2020), p. 109.

[86] R. M. Hoesly et al. “Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS)”. In: *Geoscientific Model Development* 11.1 (2018), pp. 369–408. DOI: 10.5194/gmd-11-369-2018. URL: https://gmd.copernicus.org/articles/11/369/2018/.

[87] M. J. E. van Marle et al. “Historic global biomass burning emissions for CMIP6 (BB4CMIP) based on merging satellite observations with proxies and fire models (1750–2015)”. In: *Geoscientific Model Development* 10.9 (2017), pp. 3329–3357. DOI: 10.5194/gmd-10-3329-2017. URL: https://gmd.copernicus.org/articles/10/3329/2017/.

[88] Guus J.M. Velders et al. “Future atmospheric abundances and climate forcings from scenarios of global and regional hydrofluorocarbon (HFC) emissions”. In: *Atmospheric Environment* 123 (2015), pp. 200–209. DOI: https://doi.org/10.1016/j.atmosenv.2015.10.071. URL: https://www.sciencedirect.com/science/article/pii/S135223101530488X.

[89] J. Gütschow et al. “The PRIMAP-hist national historical emissions time series”. In: *Earth System Science Data* 8.2 (2016), pp. 571–603. DOI: 10.5194/essd-8-571-2016. URL: https://essd.copernicus.org/articles/8/571/2016/.

[90] Chen Bin et al. “Main moisture sources affecting lower Yangtze River basin in Boreal Summers during 2004-2009”. In: *International Journal of Climatology* 33.4 (2012), pp. 1035–1046. DOI: 10.1002/joc.3495.

[91] H. Hersbach et al. *ERA5 hourly data on single levels from 1940 to present*. 2023. URL: 10.24381/cds.adbb2d47.

[92] H. Hersbach et al. *ERA5 hourly data on single levels from 1940 to present*. 2023. URL: 10.24381/cds.adbb2d47.

[93] Ruud J van der Ent et al. *WAM2layers*. Version v3.0.0-beta.4. Aug. 2022. URL: https://github.com/WAM2layers/WAM2layers.

[94] R. J. van der Ent et al. “Should we use a simple or complex model for moisture recycling and atmospheric moisture tracking?” In: *Hydrology and Earth System Sciences* 17.12 (2013), pp. 4869–4884. DOI: 10.5194/hess-17-4869-2013. URL: https://hess.copernicus.org/articles/17/4869/2013/.

[95] Yingzhao Ma et al. “Precipitation bias variability versus various gauges under different climatic conditions over the Third Pole Environment (TPE) region”. In: *International Journal of Climatology* 35.7 (2015), pp. 1201–1211.

[96] Daqing Yang et al. “Compatibility evaluation of national precipitation gage measurements”. In: *Journal of Geophysical Research: Atmospheres* 106.D2 (2001), pp. 1481–1491.

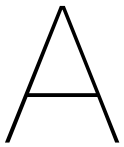
[97] Jennifer C Adam et al. “Correction of global precipitation products for orographic effects”. In: *Journal of Climate* 19.1 (2006), pp. 15–38.

[98] Hong Pan et al. “Comparison of Projections of Precipitation over Yangtze River Basin of China by Different Climate Models”. In: *Water* 14.12 (2022). DOI: 10.3390/w14121888. URL: https://www.mdpi.com/2073-4441/14/12/1888.

[99] Yanlin Yue et al. “Future changes in precipitation and temperature over the Yangtze River Basin in China based on CMIP6 GCMs”. In: *Atmospheric Research* 264 (2021), p. 105828. DOI: https://doi.org/10.1016/j.atmosres.2021.105828. URL: https://www.sciencedirect.com/science/article/pii/S0169809521003847.

[100] Bin Wang et al. “Interactions between the Asian monsoon and the El Niño/Southern oscillation”. In: *The Asian Monsoon* (2006), pp. 479–512.

[101] Kristoffer Wickstrøm et al. “The Kernelized Taylor Diagram”. In: *arXiv preprint arXiv:2205.08864* (2022).



Appendix Table

In this appendix, some tables are provided to complement the main text.

Table A.1: Monthly mean rainfall in the down stream of YRB calculated by observation data and different GCMs, with shades of blue and red representing the amount of rainfall

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mean-Obv	**2.03**	**3.13**	**4.83**	**6.13**	**7.04**	**8.45**	**5.31**	**4.77**	**3.18**	**2.13**	**2.03**	**1.44**
GPCC-Obv	2.34	3.16	5.13	5.91	6.42	8.31	5.68	4.95	2.99	2.15	2.14	1.40
ERA5-Obv	1.96	3.24	4.99	7.00	8.34	10.24	5.81	5.23	3.69	2.23	2.07	1.49
CRU-Obv	1.88	2.94	4.68	5.74	6.88	7.85	5.23	4.74	3.09	2.14	2.02	1.49
Aphro-Obv	**1.93**	**3.16**	**4.54**	**5.86**	**6.50**	**7.39**	**4.51**	**4.16**	**2.98**	**2.01**	**1.88**	**1.37**
E3SM-1-0	2.04	3.48	4.37	6.45	7.83	7.25	6.42	4.49	2.37	1.89	2.18	2.13
EC-Earth3	1.96	2.90	4.19	6.60	7.64	7.05	5.61	4.23	2.29	1.67	2.68	2.76
FIO-ESM-2-0	2.10	3.21	4.90	6.70	7.39	9.41	5.12	4.38	3.35	1.91	2.49	2.17
EC-Earth3-AerChem	2.07	2.92	4.36	5.99	6.96	6.10	5.54	3.65	1.73	1.41	2.96	2.12
CMCC-CM2-SR5	1.88	3.01	5.23	7.17	8.53	9.65	6.30	5.48	3.41	2.06	2.30	2.04
CMCC-CM2-HR4	1.29	2.37	4.46	6.88	8.21	6.55	4.66	4.47	2.39	1.67	1.89	1.39
E3SM-2-0	2.33	3.28	4.38	6.44	7.61	6.69	7.22	5.66	2.97	2.46	2.82	2.65
EC-Earth3-CC	1.86	3.06	4.42	6.68	8.17	6.54	5.94	4.62	2.16	1.67	3.02	2.66
CMCC-ESM2	1.94	3.10	4.87	6.46	7.72	9.24	6.56	5.92	2.75	1.54	1.69	1.90
NorCPM1	0.85	1.83	3.93	6.19	7.02	6.99	5.43	4.17	3.13	2.29	1.48	0.95
CESM2-WACCM-FV2	1.53	2.37	4.32	5.94	7.16	9.63	5.70	4.74	3.96	1.77	1.51	1.61
CESM2-WACCM	2.12	3.08	4.88	6.48	6.56	9.75	6.06	5.03	4.23	2.19	1.53	2.29
MRI-ESM2-0	1.56	3.05	4.45	6.51	7.82	5.96	6.54	4.16	1.13	1.03	1.88	1.93
IPSL-CM6A-LR-INCA	1.41	2.88	5.03	6.85	7.07	8.03	6.99	6.55	3.27	1.14	1.42	1.14
IPSL-CM6A-LR	1.52	2.59	5.12	7.06	6.78	7.17	6.89	6.55	3.21	0.88	1.55	1.02
CESM2	2.17	3.32	4.99	6.70	6.93	10.50	6.61	5.28	4.32	1.76	1.74	2.34
CESM2-FV2	1.51	2.31	4.39	6.04	7.32	9.83	5.69	3.82	4.22	2.11	1.73	1.68
CNRM-ESM2-1	2.06	3.58	5.56	7.41	7.41	7.13	6.56	3.65	2.02	1.65	2.15	2.20
CNRM-CM6-1	1.99	3.37	5.44	7.71	8.05	7.55	7.01	4.12	2.30	2.12	2.18	2.14
CNRM-CM6-1-HR	2.29	3.39	5.76	8.13	7.48	7.04	6.51	3.00	2.17	1.73	2.42	2.60
FGOALS-f3-L	1.42	2.38	3.54	5.81	6.31	6.25	4.80	4.11	3.04	2.58	2.77	1.28
MIROC-ES2L	1.43	2.89	4.67	6.25	7.02	8.93	7.34	6.24	4.30	2.18	1.22	0.72
BCC-ESM1	1.68	1.84	3.36	5.36	6.90	5.88	4.61	4.42	3.68	2.15	1.69	2.06
CAS-ESM2-0	1.91	2.89	3.41	4.68	5.85	6.49	3.84	4.02	2.49	1.75	2.37	1.50
GFDL-ESM4	1.89	2.82	4.70	7.06	6.49	5.20	6.22	5.09	2.16	1.60	2.55	2.69
MPI-ESM-1-2-HAM	2.24	4.31	5.73	7.42	7.01	6.91	5.87	5.47	1.55	0.60	1.21	1.67
MPI-ESM1-2-HR	2.45	4.13	6.00	7.83	7.61	6.69	6.98	5.84	1.53	0.33	1.45	1.87
CanESM5	1.76	3.73	6.65	8.71	7.53	6.50	7.48	4.77	1.75	1.64	2.48	1.63
CanESM5-CanOE	1.93	3.78	6.48	8.35	7.57	5.98	7.64	5.77	1.60	1.52	2.42	1.59
CAMS-CSM1-0	2.84	4.71	6.09	7.66	6.13	5.01	4.22	2.93	1.50	1.67	2.47	1.86
IPSL-CM5A2-INCA	1.32	2.02	3.30	4.51	4.81	5.17	7.40	6.19	1.60	1.26	1.32	0.97
INM-CM4-8	4.66	5.38	6.45	6.73	7.77	10.57	8.53	9.07	5.44	2.75	4.31	4.75
GISS-E2-1-G	2.42	4.30	6.80	8.25	7.49	5.24	4.31	3.20	1.95	2.17	2.75	2.67
FGOALS-g3	0.56	0.89	1.97	4.21	5.22	4.67	4.05	2.62	2.77	2.30	1.71	0.96
INM-CM5-0	3.10	4.39	5.67	7.18	7.61	10.71	9.30	9.55	5.28	3.00	3.70	3.65
ACCESS-ESM1-5	2.98	4.65	7.76	10.78	10.47	10.04	7.38	6.56	3.06	3.22	4.94	3.59
GISS-E2-1-H	2.59	4.30	6.55	8.41	7.88	5.39	4.68	3.03	2.35	2.74	3.58	2.86
KACE-1-0-G	4.35	5.95	7.85	10.79	10.81	7.97	8.10	4.85	2.48	2.26	3.10	3.95
KIOST-ESM	1.66	3.04	4.62	5.58	5.58	4.04	3.54	3.06	2.41	2.40	1.60	1.29
ACCESS-CM2	3.93	5.42	8.13	10.09	10.11	7.36	9.87	6.04	2.91	3.07	3.29	4.01
MCM-UA-1-0	4.48	6.65	8.82	8.43	8.79	5.37	3.51	4.36	4.29	4.05	2.29	3.56
IITM-ESM	4.53	5.97	6.70	6.41	5.32	5.74	4.31	3.20	2.01	1.88	3.46	4.18

Table A.2: Monthly mean rainfall in the middle stream of YRB calculated by observation data and different GCMs, with shades of blue and red representing the amount of rainfall

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mean-Obv	**1.47**	**2.13**	**3.29**	**4.86**	**5.93**	**6.38**	**5.60**	**4.73**	**3.25**	**2.71**	**1.96**	**1.13**
GPCC-Obv	1.55	2.03	3.08	4.33	5.41	6.05	5.72	4.55	2.89	2.65	1.81	1.01
ERA5-Obv	1.78	2.72	4.20	6.16	7.19	8.04	6.78	5.85	4.10	3.02	2.30	1.41
CRU-Obv	1.26	1.80	3.04	4.54	5.73	5.94	5.09	4.43	3.11	2.59	1.90	1.07
Aphro-Obv	**1.30**	**1.98**	**2.83**	**4.43**	**5.39**	**5.50**	**4.80**	**4.08**	**2.88**	**2.59**	**1.83**	**1.03**
E3SM-1-0	1.39	2.34	2.68	3.92	5.80	5.75	5.75	4.75	2.72	2.20	1.83	1.49
EC-Earth3	1.47	2.25	3.18	5.28	6.44	5.59	5.38	4.27	2.35	2.42	2.50	1.85
FIO-ESM-2-0	1.48	2.56	3.58	5.57	7.09	8.97	5.90	4.76	4.53	2.87	2.21	1.46
EC-Earth3-AerChem	1.57	2.23	3.16	4.54	5.64	4.98	4.81	3.71	2.00	1.97	2.49	1.45
CMCC-CM2-SR5	1.33	2.43	4.06	5.79	7.63	9.07	7.24	5.44	4.17	3.18	2.29	1.41
CMCC-CM2-HR4	0.85	1.67	2.87	4.68	6.68	6.05	4.94	4.31	2.82	2.20	1.48	0.91
E3SM-2-0	1.57	2.31	3.06	4.34	6.11	5.12	5.70	4.97	2.82	2.48	2.12	1.70
EC-Earth3-CC	1.50	2.31	3.21	5.15	6.96	5.47	5.53	4.36	2.30	2.47	2.67	1.82
CMCC-ESM2	1.48	2.33	3.74	5.45	7.08	8.78	7.38	6.26	4.12	2.55	1.68	1.39
NorCPM1	0.71	1.55	3.16	5.51	7.91	7.65	6.18	5.11	4.26	3.46	1.56	0.73
CESM2-WACCM-FV2	1.32	2.15	3.69	5.15	7.35	8.82	6.86	5.53	5.14	2.81	1.57	1.32
CESM2-WACCM	1.42	2.36	3.36	4.65	6.02	8.06	6.07	5.13	4.17	2.52	1.35	1.44
MRI-ESM2-0	1.22	2.39	3.37	4.79	6.26	5.10	4.90	3.42	1.55	1.46	1.41	1.30
IPSL-CM6A-LR-INCA	0.98	2.10	3.35	4.56	5.18	5.85	6.79	5.80	2.42	1.66	0.96	0.69
IPSL-CM6A-LR	1.10	1.88	3.26	4.57	5.06	5.45	6.79	5.77	2.57	1.37	1.04	0.63
CESM2	1.48	2.39	3.51	4.99	6.16	8.65	6.58	5.16	4.48	2.36	1.65	1.51
CESM2-FV2	1.20	2.09	3.64	5.34	7.65	9.67	6.49	4.87	5.25	2.97	1.90	1.30
CNRM-ESM2-1	1.26	2.18	3.54	5.30	6.32	6.12	6.22	4.30	2.79	2.24	1.88	1.38
CNRM-CM6-1	1.14	2.05	3.45	5.68	6.67	6.07	6.48	4.69	3.11	2.74	1.84	1.39
CNRM-CM6-1-HR	1.37	2.12	3.46	5.56	6.30	6.11	6.45	4.22	2.71	2.36	2.08	1.70
FGOALS-f3-L	0.95	1.64	2.32	3.80	4.88	4.83	4.04	3.93	2.71	2.32	1.98	0.81
MIROC-ES2L	1.46	2.87	4.61	6.75	7.10	7.81	7.53	6.48	5.68	3.49	1.36	0.83
BCC-ESM1	1.29	1.86	3.42	5.68	7.40	5.95	5.02	4.81	4.47	2.89	1.83	1.54
CAS-ESM2-0	1.84	2.81	3.45	4.99	6.73	6.57	5.16	4.44	3.95	3.44	2.83	1.50
GFDL-ESM4	1.37	2.33	3.84	5.65	6.09	4.69	5.00	4.76	2.77	2.14	2.33	1.83
MPI-ESM-1-2-HAM	1.80	3.53	4.54	6.09	6.41	6.26	6.22	5.05	1.91	1.23	1.48	1.34
MPI-ESM1-2-HR	2.07	3.27	4.73	5.91	6.45	5.23	5.71	4.64	1.40	0.83	1.68	1.69
CanESM5	1.54	3.55	5.56	6.98	7.92	6.93	6.96	4.11	3.53	3.07	2.68	1.22
CanESM5-CanOE	1.62	3.41	5.35	6.87	7.50	6.65	7.20	4.83	3.56	2.65	2.56	1.35
CAMS-CSM1-0	2.03	3.39	4.58	6.48	6.02	4.37	4.15	2.92	1.71	2.22	2.25	1.42
IPSL-CM5A2-INCA	0.80	1.51	2.46	3.37	4.00	4.67	5.57	4.53	1.98	1.59	0.88	0.59
INM-CM4-8	2.23	3.44	4.81	5.28	6.65	8.71	9.36	8.74	5.75	4.00	3.53	2.31
GISS-E2-1-G	2.06	3.45	5.31	6.68	6.50	5.38	4.59	4.07	3.22	3.19	2.53	1.93
FGOALS-g3	0.36	0.68	1.43	3.38	5.37	4.71	3.62	2.98	2.97	2.44	1.22	0.54
INM-CM5-0	1.89	3.32	4.81	5.88	6.33	8.10	9.14	8.48	5.60	4.24	3.07	2.07
ACCESS-ESM1-5	1.41	3.00	5.53	8.78	9.75	8.10	9.13	7.27	3.86	3.71	3.11	1.53
GISS-E2-1-H	2.11	3.42	4.99	7.04	7.09	5.42	4.36	3.42	3.41	3.61	3.37	2.05
KACE-1-0-G	2.18	3.20	5.03	7.60	8.86	7.27	8.77	6.12	4.16	3.02	2.51	2.30
KIOST-ESM	1.20	2.27	3.28	4.12	4.79	3.90	3.41	3.23	2.86	2.78	1.38	0.89
ACCESS-CM2	1.94	3.25	5.48	7.67	8.65	6.99	8.67	6.54	3.85	3.49	2.24	2.09
MCM-UA-1-0	4.88	6.97	9.48	9.61	9.12	5.46	4.40	4.58	5.70	4.09	2.24	3.56
IITM-ESM	3.24	4.49	5.53	5.95	5.21	5.62	4.46	3.11	2.74	2.73	3.13	2.97

Table A.3: Monthly mean rainfall in the up stream of YRB calculated by observation data and different GCMs, with shades of blue and red representing the amount of rainfall

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mean-Obv	**0.46**	**0.58**	**0.93**	**1.65**	**3.00**	**4.86**	**5.55**	**4.87**	**3.86**	**2.18**	**0.90**	**0.44**
GPCC-Obv	0.32	0.38	0.71	1.36	2.74	4.45	5.31	4.52	3.36	1.79	0.64	0.27
ERA5-Obv	0.97	1.19	1.69	2.56	4.12	6.29	7.11	6.27	5.20	3.12	1.57	0.92
CRU-Obv	0.26	0.33	0.66	1.34	2.58	4.55	5.00	4.43	3.49	1.95	0.68	0.28
Aphro-Obv	**0.30**	**0.41**	**0.66**	**1.35**	**2.57**	**4.17**	**4.80**	**4.25**	**3.40**	**1.85**	**0.70**	**0.29**
E3SM-1-0	0.77	1.28	1.37	2.08	3.61	6.17	7.14	6.55	4.58	2.52	1.18	0.81
EC-Earth3	0.82	1.07	1.43	2.15	3.59	5.78	6.39	5.60	4.76	2.76	1.43	0.94
FIO-ESM-2-0	0.80	1.35	2.09	3.39	5.09	8.38	8.19	7.88	6.39	3.31	1.44	0.80
EC-Earth3-AerChem	0.74	0.99	1.28	1.97	3.19	5.26	5.75	5.17	4.16	2.40	1.31	0.74
CMCC-CM2-SR5	0.73	1.48	2.55	3.79	5.54	8.52	8.47	7.63	6.29	3.48	1.53	0.77
CMCC-CM2-HR4	0.65	1.16	1.53	2.48	4.48	6.68	6.81	6.06	4.69	2.55	1.15	0.70
E3SM-2-0	0.90	1.18	1.53	2.28	3.75	5.97	6.89	6.23	4.63	2.89	1.17	0.84
EC-Earth3-CC	0.89	1.09	1.41	2.12	3.55	5.67	6.20	5.54	4.51	2.82	1.41	0.86
CMCC-ESM2	0.77	1.27	2.37	3.63	5.13	8.17	8.86	8.20	6.07	3.38	1.36	0.72
NorCPM1	0.69	1.16	1.87	3.24	6.03	9.08	8.10	7.32	6.44	3.63	1.44	0.66
CESM2-WACCM-FV2	0.84	1.42	1.99	3.18	6.20	9.76	9.04	7.63	6.95	3.64	1.32	0.75
CESM2-WACCM	0.63	1.14	1.67	2.75	4.91	8.30	8.24	7.25	5.49	2.80	0.94	0.49
MRI-ESM2-0	0.93	1.43	1.99	2.58	3.80	6.13	7.13	6.54	4.20	1.94	0.98	0.80
IPSL-CM6A-LR-INCA	0.92	1.25	1.51	2.21	3.23	4.53	5.34	4.73	4.00	2.77	1.16	0.78
IPSL-CM6A-LR	0.84	1.13	1.50	2.16	3.10	4.54	5.27	4.53	3.99	2.56	1.18	0.69
CESM2	0.63	1.13	1.71	2.86	4.92	8.41	8.38	7.22	5.66	2.78	1.05	0.54
CESM2-FV2	0.76	1.33	1.97	3.34	6.52	10.18	8.76	7.52	6.50	3.35	1.63	0.81
CNRM-ESM2-1	0.93	1.32	1.81	2.63	4.03	6.47	6.43	5.75	5.41	3.20	1.62	1.11
CNRM-CM6-1	0.92	1.36	1.98	2.91	4.05	6.46	6.63	6.05	5.38	3.10	1.74	1.08
CNRM-CM6-1-HR	0.71	1.04	1.40	1.98	2.88	4.95	5.79	4.96	4.37	2.18	1.43	0.86
FGOALS-f3-L	0.67	0.91	1.21	1.69	2.27	3.77	4.31	4.04	3.30	2.03	1.45	0.64
MIROC-ES2L	1.03	1.74	2.86	4.88	7.02	8.97	8.05	7.88	7.68	4.40	1.71	0.89
BCC-ESM1	1.20	1.55	2.51	4.65	6.73	6.59	6.61	6.66	6.10	3.74	1.83	1.26
CAS-ESM2-0	1.42	1.72	2.36	3.50	4.66	6.49	6.19	6.14	6.49	4.62	2.78	1.67
GFDL-ESM4	0.76	1.07	1.82	2.88	4.07	5.91	6.33	5.97	4.96	3.09	1.46	0.80
MPI-ESM-1-2-HAM	0.80	1.27	1.87	3.25	5.65	7.39	6.45	5.58	5.09	2.56	1.32	0.80
MPI-ESM1-2-HR	0.92	1.46	2.14	2.89	4.36	6.18	6.04	5.36	3.96	2.03	1.35	0.92
CanESM5	0.78	1.28	1.80	2.14	4.43	7.50	9.89	7.86	5.85	2.93	1.41	0.56
CanESM5-CanOE	0.67	1.28	1.69	2.38	4.28	7.21	10.26	8.38	5.83	2.78	1.36	0.69
CAMS-CSM1-0	0.98	1.47	1.98	3.05	4.58	5.35	5.01	4.78	3.74	2.57	1.52	0.91
IPSL-CM5A2-INCA	1.02	1.54	1.75	1.88	2.54	4.40	5.16	4.54	5.24	3.43	1.56	0.77
INM-CM4-8	2.05	2.73	3.50	4.24	6.01	8.42	7.85	6.89	7.65	5.18	3.48	2.13
GISS-E2-1-G	1.83	2.50	3.35	4.34	5.11	5.70	5.17	4.83	5.29	4.17	2.53	1.75
FGOALS-g3	0.49	0.69	1.04	1.90	3.70	5.21	5.91	5.85	4.78	2.66	1.05	0.61
INM-CM5-0	1.78	2.59	3.40	4.39	5.76	8.64	8.54	6.73	7.76	5.33	2.90	1.63
ACCESS-ESM1-5	1.28	2.17	3.14	4.39	5.49	7.36	8.89	8.49	6.00	3.42	1.83	1.12
GISS-E2-1-H	1.76	2.38	3.22	4.30	5.29	6.10	4.78	4.28	5.53	5.30	3.15	1.78
KACE-1-0-G	0.89	1.37	2.25	3.41	4.65	6.30	7.96	7.48	5.98	3.09	1.19	0.78
KIOST-ESM	0.82	1.47	1.83	2.51	3.79	4.52	4.97	4.51	4.38	3.35	1.52	0.80
ACCESS-CM2	1.23	2.00	3.07	4.35	5.45	6.88	7.56	7.66	5.82	3.43	1.69	1.13
MCM-UA-1-0	3.61	3.43	3.92	5.14	7.56	9.16	7.98	7.78	7.20	3.35	1.77	2.90
IITM-ESM	2.03	2.54	3.35	4.39	5.29	5.97	5.53	5.23	5.36	3.54	2.46	1.93

Table A.4: Taylor Skill Score of Three stream in YRB, with the Resolution and Ranking

	Downstream TSS	Midstream TSS	Upstream TSS	Resolution (km)	Ranking
E3SM-1-0	0.9622	0.9447	0.8707	100	1
EC-Earth3	0.9585	0.9658	0.9528	100	2
FIO-ESM-2-0	0.9508	0.9299	0.7335	100	3
EC-Earth3-AerChem	0.9492	0.9505	0.9624	100	4
CMCC-CM2-SR5	0.9489	0.9262	0.7405	100	5
CMCC-CM2-HR4	0.9475	0.9539	0.8380	100	6
E3SM-2-0	0.9471	0.9442	0.9078	100	7
EC-Earth3-CC	0.9451	0.9571	0.9554	100	8
CMCC-ESM2	0.9441	0.9225	0.7412	100	9
NorCPM1	0.9430	0.8797	0.7820	250	10
CESM2-WACCM-FV2	0.9417	0.8935	0.7805	250	11
CESM2-WACCM	0.9368	0.9373	0.8066	100	12
MRI-ESM2-0	0.9365	0.9320	0.8606	100	13
IPSL-CM6A-LR-INCA	0.9355	0.9088	0.8709	250	14
IPSL-CM6A-LR	0.9331	0.9065	0.8768	250	15
CESM2	0.9329	0.9337	0.8110	100	16
CESM2-FV2	0.9314	0.8761	0.7873	250	17
CNRM-ESM2-1	0.9311	0.9145	0.8318	250	18
CNRM-CM6-1	0.9309	0.9163	0.8297	250	19
CNRM-CM6-1-HR	0.9282	0.9336	0.8748	100	20
FGOALS-f3-L	0.9247	0.9037	0.8584	100	21
MIROC-ES2L	0.9196	0.9005	0.8365	500	22
BCC-ESM1	0.9194	0.9245	0.9046	250	23
CAS-ESM2-0	0.9137	0.9227	0.6922	100	24
GFDL-ESM4	0.9129	0.9175	0.9112	100	25
MPI-ESM-1-2-HAM	0.9129	0.8446	0.8060	250	26
MPI-ESM1-2-HR	0.8992	0.8844	0.8840	100	27
CanESM5	0.8971	0.8791	0.7230	500	28
CanESM5-CanOE	0.8929	0.8894	0.7167	500	29
CAMS-CSM1-0	0.8724	0.8527	0.8540	100	30
IPSL-CM5A2-INCA	0.8721	0.8313	0.7622	500	31
INM-CM4-8	0.8668	0.8801	0.8407	100	32
GISS-E2-1-G	0.8613	0.8417	0.6862	250	33
FGOALS-g3	0.8584	0.8405	0.7500	250	34
INM-CM5-0	0.8496	0.8627	0.8138	100	35
ACCESS-ESM1-5	0.8438	0.8007	0.8728	250	36
GISS-E2-1-H	0.8376	0.8343	0.6836	250	37
KACE-1-0-G	0.8361	0.8758	0.8982	250	38
KIOST-ESM	0.8354	0.7808	0.8531	250	39
ACCESS-CM2	0.8279	0.8458	0.8943	250	40
MCM-UA-1-0	0.7823	0.6695	0.7970	250	41
IITM-ESM	0.7791	0.8135	0.8474	250	42