

Evaluating the Performance of CMIP6 GCMs to Simulate Precipitation and Temperature Over the Vietnamese Mekong Delta

Tran Van Ty¹, Le Hai Tri¹, Nguyen Van Tho², Nguyen Van Toan³, Giap Minh Nhat⁴, Nigel K. Downes⁴, Pankaj Kumar⁵ and Huynh Vuong Thu Minh^{4*}

¹College of Engineering, Can Tho University, Can Tho, Vietnam

²Department of Environment, Mien Tay Construction University, Vinh Long, Vietnam

³Department of Human Resources, Can Tho University, Can Tho, Vietnam

⁴College of Environment and Natural Resources, Can Tho University, Can Tho, Vietnam

⁵Institute for Global Environmental Strategies, Hayama – 240-0115, Japan

✉ hvtminh@ctu.edu.vn

Received March 3, 2023; revised and accepted April 10, 2023

Abstract: This study evaluates the performance of simulated precipitation and maximum and minimum temperatures in the historical runs of the Climate Model Intercomparison Project Phase 6 (CMIP6) for the Vietnamese Mekong Delta (VMD). The precipitation, as well as maximum and minimum temperatures outputs from 16 general circulation models (GCMs), were compared with observations from 12 stations for the period 1980–2014, using a set of statistical metrics, namely, normalised root mean square error (NRMSE), percentage of bias (PBIAS), Nash–Sutcliffe efficiency (NSE), coefficient of determination (R^2), and volumetric efficiency (VE). Finally, ranking (total score - TS) was carried out and the probability distribution function (PDF) and Taylor diagram were used to confirm rankings. The results show that different statistical indicators reveal variation ranking order of the 16 GCMs. Based on RS ranking, it is indicated that each simulation GCM performed differently under the different metrics and no single model performed best for all metrics. The top five highest ranked GCMs based on TS were HadGEM3-GC31-LL, ACCESS-CM2, CanESM5, NESM3 and CanESM5-CanOE for precipitation; and CNRM-CM6-1, CNRM-ESM2-1, GFDL-ESM4, NESM3 and INM-CM5-0 for the maximum; and CNRM-CM6-1, CNRM-ESM2-1, GFDL-ESM4, NESM3 and INM-CM5-0 for minimum temperatures, respectively. We also observed an underestimation of precipitation and an overestimation of temperature over the study area. The TS method demonstrates efficiency to aggregate the multi-model ensemble GCMs based on different statistical indicators which were sometimes contradictory. The findings from this study provide useful guidance in the selection of GCMs for climate change applications in the VMD.

Keywords: General Circulation Models (GCMs); Climate Model Intercomparison Project Phase 6 (CMIP6); Statistical metrics; Ranking score and total score (RS and TS); Vietnamese Mekong Delta.

Introduction

Since their initial publication in 2017, more than 70 results of the Climate Model Intercomparison Project

Phase 6 (CMIP6) GCMs performance were published by the year 2021 (Stouffer et al., 2017; Cui et al., 2021; Earth System Grid Federation, 2021). Understandably, each CMIP6 GCM exhibits different degrees of

*Corresponding Author

performance due to the differences in their model structure, parameterization, and initial conditions setting, as well as differences in climate and topography (Guo et al., 2021; Wang et al., 2021). Despite the same modeling framework, the results of precipitation simulations vary from model to model, and from nation to nation owing to changes in the influencing factors, such as seasons, topographical features, and ocean currents, all of which contribute to the effectiveness of simulation models (Desmet and Ngo, 2022). Furthermore, worldwide testing has shown that CMIP6 GCMs performance may vary due to geographical scope and topography (Rivera and Arnould, 2020; Yazdandoost et al., 2021). However, to date, only a few studies have examined the efficiency of CMIP6 GCM precipitation simulations for the Southeast Asia (SEA) region or the sub-domain partitioning in the SEA region (Iqbal et al., 2021). Moreover, even fewer studies have shown different precipitation trends for individual countries within SEA (Thoeun, 2015; Piman et al., 2016; Phuong et al., 2019; Pimonsree et al., 2022).

Shiru and Chung (2021) highlight the worldwide increase in disaster frequencies, severities, and risks, particularly droughts and floods (Asdak and Supian, 2018; Alamgir et al., 2019; Ayugi et al., 2020; Manawi et al., 2020). These disasters are expected to increase over time under various emission scenarios and GCMs, the main instruments for climate prediction (Shiru et al., 2020; Tan et al., 2020). However, to improve confidence in future climate projections, the evaluation of GCMs performance is strongly needed in order to develop reliable and appropriate adaptation and mitigation measures (Zhao et al., 2020). The evolution of GCMs has been based on the different scenario developments of the United Nations Intergovernmental Panel on Climate Change (IPCC) assessment reports (the coupled model inter-comparison project (CMIP) phase 3, phase 5, and the recently released phase 6). Numerous studies have reported the improvement of the CMIP5 over the CMIP3 and CMIP6 over its predecessor (Taylor et al., 2012; Tanveer et al., 2016; Zhou et al., 2017; Tiwari et al., 2022; Guo et al., 2023; Pimonsree et al., 2023). Other recent studies have demonstrated the supremacy of GCMs in CMIP6 for particular regions, such as South Asia (Zhai et al., 2020), China (Xin et al., 2020; Guo et al., 2023), South Korea (Song et al., 2020), Australia (Grose et al., 2020), and Africa (Ayugi et al., 2021; Shiru and Chung, 2021), as well as SEA (Pimonsree et al., 2023). It is therefore of utmost importance to further assessing their application and performance in

other regions to aid future climate projection (Shiru and Chung, 2021).

Moreover, only a handful of studies have provided an overview of climate change assessments for the SEA region (Tinh et al., 2016; Ge et al., 2021; Ty et al., 2022), reported that the CMIP6 multi-model ensemble medians showed better performances in characterizing precipitation extremes than individual models, while projected changes in precipitation extremes increased significantly over the Indochina Peninsula and the Maritime Continent. Furthermore, Supharatid et al. (2022) used the CMIP6 model to predict changes in temperature and precipitation over mainland SEA, to highlight the future climate risks to Cambodia, Laos, Myanmar, Vietnam, and Thailand. In Vietnam, Khoi et al. (2022) recently assessed the impact of future climate change on river discharges in Ho Chi Minh City (using a calibrated Soil and Water Assessment Tool (SWAT) to simulate the discharge under seven GCMs derived from CMIP6 and suggested that the city's climate will be warmer and wetter by the end of the 21st century.

When reviewing the existing studies a number of important scientific knowledge gaps persist with a lack of research that considers the scale effect of selecting acceptable GCMs for assessing climate change at lower spatial scales, such as the Vietnamese Mekong delta (VMD). As a result, the goal of this study is to select the more applicable GCMs from the CMIP6 profile in order to aggregate a multi-model ensemble across the VMD region. This study uses the monthly precipitation, maximum, and minimum temperature data from 12 stations in the VMD. Five statistical indices, namely, the normalized root mean square error (NRMSE), the percentage of bias (PBIAS), the Nash–Sutcliffe efficiency (NSE), the coefficient of determination (R^2), and the volumetric efficiency (VE), were applied to identify the performances of CMIP6 compared to the observed precipitation, as well maximum and minimum temperatures datasets. Finally, a probability distribution function (PDF) and the Taylor diagram (TD) are used to assess the performances of the GCMs.

Study Area

The VMD covers a large share of Southwest Vietnam with a total area of more than 40,500 km². The VMD is one of the most fertile and intensively cultivated regions worldwide, making it significant for national social-economic development. The VMD shares boundaries with Cambodia to the north, the East

Sea to the southeast, and the Gulf of Thailand to the southwest and Southeast. Administratively, it consists of 12 provinces: Long An, Tien Giang, Ben Tre, Vinh Long, Tra Vinh, Hau Giang, Soc Trang, Dong Thap, An Giang, Kien Giang, Bac Lieu, and Ca Mau, as well as one city, Can Tho City with a combined population of 17,273,630 in the year 2019 (or 18% of the nation's total population), at a population density 1.45 times greater than the national average at 423 people per square kilometer (Figure 1).

The VMD contributes roughly 50% of the nation's rice production, 95% of the nation's export rice production, about 65% of the nation's aquaculture production, 60% of the nation's export fish, and 70% of the nation's types of fruits (Lavane et al., 2023; Ty et al., 2022; Minh et al., 2022). For the years 2016–2018, the regional GDP growth rate for the VMD averaged 6.95%, almost on par with the national average for the same time period (Tran, 2019). The climate of the VMD is influenced by both the northeast and southwest monsoons. The dry season typically lasts from December to April, and the rainy season typically lasts from May to November. The

flood season begins one to two months later than the onset of the rainy season and ends nearly simultaneously with it. During the flood season, the delta facilitates water transport and drainage.

Methodology

GCM Data Collection and Analysis

For the assessment of precipitation and maximum and minimum temperature variability, 16 GCMs belonging to the CMIP6 experiments were utilised (Eyring et al., 2016; Shiru and Chung, 2021). Table 1 provides an overview of the analysed GCMs, with their respective modelling groups, countries, and horizontal resolution. The historical runs of these models were selected based on the observed data availability in VMD - thus the CMIP6 historical simulations that cover the period 1980–2014 were selected. Given that most of the models have different spatial resolutions, the output from each model was then grid-averaged to cover the entire VMD to aid comparison.

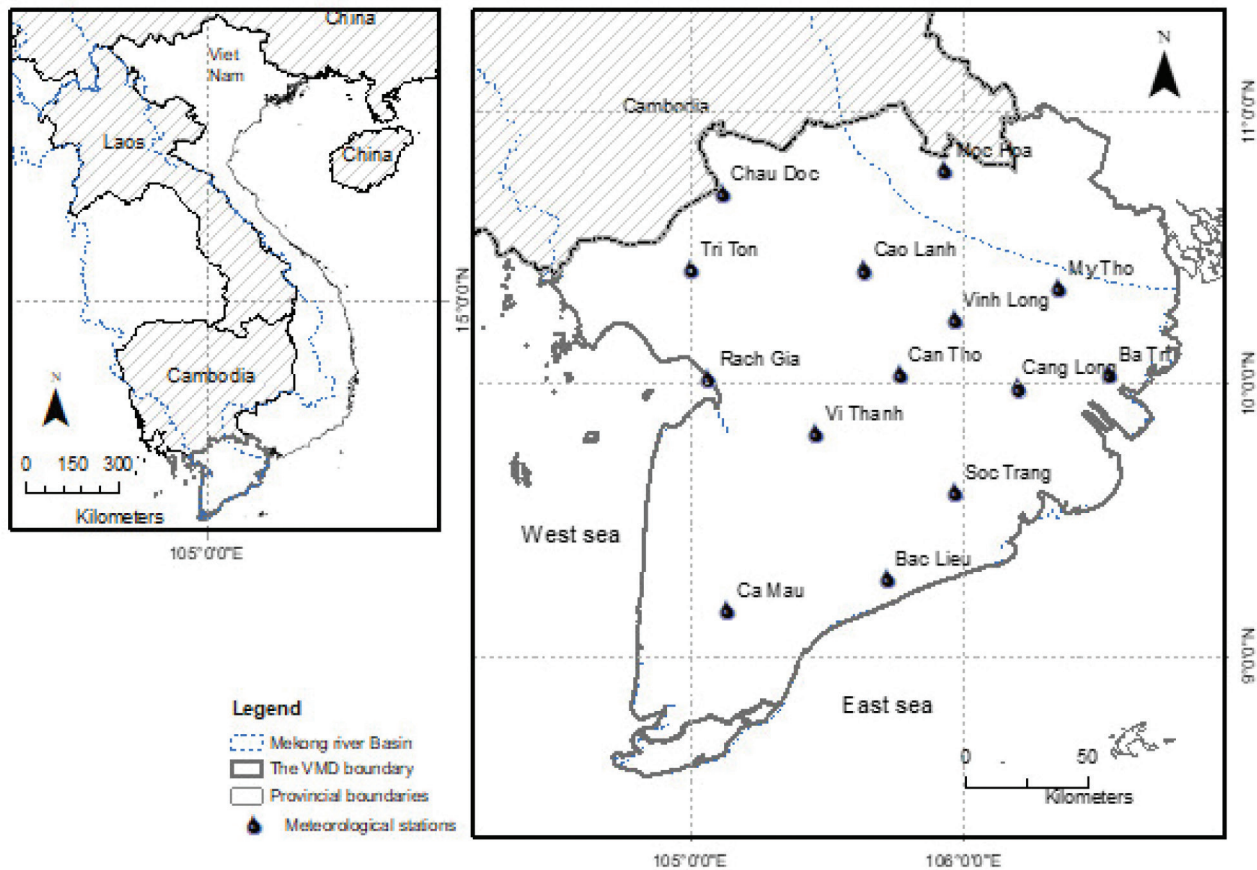


Figure 1: The study area highlighting the location of the meteorological stations.

Table 1: The 16 selected GCMs of the CMIP6 utilized in this study

No	Model name	Modeling agency	Resolution
1	ACCESS-CM2	Australian Community Climate and Earth System Simulator, Australia	$1.88^\circ \times 1.25^\circ$
2	ACCESS-ESM1-5	Australian Community Climate and Earth System Simulator, Australia	$1.88^\circ \times 1.25^\circ$
3	BCC-ESM1	Beijing Climate Center, China Meteorological Administration, China	$2.81^\circ \times 2.78^\circ$
4	CanESM5	Canadian Centre for Climate Modelling and Analysis, Canada	$2.81^\circ \times 2.78^\circ$
5	CanESM5-CanOE	Canadian Centre for Climate Modelling and Analysis, Canada	$2.81^\circ \times 2.81^\circ$
6	CMCC-ESM2	Centre National de Recherches Météorologiques, France	$1.40^\circ \times 1.40^\circ$
7	CNRM-CM6-1	National Centre for Meteorological Research	$1.40^\circ \times 1.40^\circ$
8	CNRM-ESM2-1	National Centre for Meteorological Research	$1.40^\circ \times 1.40^\circ$
9	FIO-ESM-2-0	The First Institute of Oceanography Earth System Model, China	$1.25^\circ \times 0.90^\circ$
10	GFDL-ESM4	NOAA Geophysical Fluid Dynamics Laboratory, United States	$1.88^\circ \times 1.25^\circ$
11	HadGEM3-GC31-LL	Met Office Hadley Centre, UK	$1.88^\circ \times 1.25^\circ$
12	INM-CM5-0	Institute of Numerical Mathematics, Russia	$2.00^\circ \times 1.50^\circ$
13	MIROC6	Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Kanagawa	$1.40^\circ \times 1.40^\circ$
14	MIROC-ES2L	Model for Interdisciplinary Research on Climate - Earth System Simulation version 2, Japan	$2.81^\circ \times 2.81^\circ$
15	MRI-ESM2-0	Meteorological Research Institute, Japan	$1.13^\circ \times 1.13^\circ$
16	NESM3	The Nanjing University of Information Science and Technology, China	$1.88^\circ \times 1.85^\circ$

Statistical Indices

There are a number of uncertainties in applying climate projections which lead to the questionable reliability of impact assessments (Woldemeskel et al., 2014). First of all, the simple ensemble mean of all GCMs may result in very high uncertainty due to the inclusion of poor model performance; Second, considering all models is often difficult due to the constraints of time and computational resources. To overcome these issues, the selection of a small range of GCMs from the large CMIP6 database with good performance is often the most suitable approach (Tiwari et al., 2023). Previous studies have considered various statistical indicators such as RMSE, R^2 , and standard deviation to assess the potential of models (Srivastava et al., 2017; Pandey and Dwivedi, 2021; Kumar et al., 2022). However, the selection of GCMs based on variability/correlation performance in the past may provide reliable future projections in terms of pattern, but miss the possible future mean changes (quantitatively) (Tiwari et al., 2023). Therefore, it is clear that the criteria for selecting the GCMs depend on the characteristics of the study region of interest, the spatial assessment scale, as well as other aspects of impact assessments.

In this study, five statistical indices were used to evaluate the GCM performance including NRMSE, PBIAS, NSE, R^2 , and VE. The statistical metrics used

and their considered acceptable levels of performance in this study are as follows:

- **Normalized root mean square error (NRMSE):** The magnitude of the errors in predictions for various periods was calculated by the NRMSE (Willmott, 1982). It is clearly indicated that the closer the NRMSE value is to zero, the more accurate/the better performance the model is

$$NRMSE = \frac{\left[\frac{1}{n} \sum_{i=1}^n (x_{sim,i} - x_{obs,i})^2 \right]^{1/2}}{\frac{1}{n} \sum_{i=1}^n (x_{sim,i})}$$

where, $x_{sim,i}$ and $x_{obs,i}$ are the i^{th} simulated and observed data; n is the number of observations.

- **Percentage of bias (PBIAS):** The Pbias measured the under or over-estimate of the model to the observed data. Model performance is better when the PBIAS values are closer to zero; while a negative/positive Pbias value indicates overestimation/underestimation (Gupta et al., 1999).

$$Pbias = 100 \times \frac{\left[\sum_{i=1}^n (x_{sim,i} - x_{obs,i}) \right]}{\sum_{i=1}^n (x_{sim,i})}$$

- **Nash and Sutcliffe efficiency (NSE):** The quantitative statistic of Nash and Sutcliffe (1970) was introduced in 1970. It is defined by the following equation. NSE ranges between -∞ and 1.0; and NSE values from 0.0 and 1.0 are considered acceptable levels of performance, whereas values <0.0 are indicative of unacceptable model performance in which the mean observed value is a better predictor than the simulated value (Moriasi et al., 2007).

$$NSE = \frac{\sum_{i=1}^n (x_{sim,i} - x_{obs,i})^2}{\sum_{i=1}^n (x_{obs,i} - \overline{x_{obs}})^2}$$

- **Coefficient of determination (R²):** R² values can range between 0.0 and 1.0, in which a higher value indicates a better agreement (Legates and McCabe, 1999) and R² is defined as follows.

$$R^2 = \frac{\sum_{i=1}^n (x_{obs,i} - \overline{x_{obs}})(x_{sim,i} - \overline{x_{sim}})}{\sqrt{\sum_{i=1}^n (x_{sim,i} - \overline{x_{sim}})^2 \sum_{i=1}^n (x_{obs,i} - \overline{x_{obs}})^2}}$$

- **Volumetric efficiency (VE):** The VE measures the ratio between GCM and observed data over a period, where a VE value of 1 indicates a perfect estimation. It is defined as follows (Shiru and Chung, 2021).

$$VE = 1 - \frac{\sum_{i=1}^n (x_{sim,i} - x_{obs,i})}{\sum_{i=1}^n x_{obs,i}}$$

GCMs Ranking

The normalisation of each of the five metrics mentioned above was used to calculate the score to rank GCMs. The ranking score (*RS*) is defined as (Guo et al., 2022):

If the given metric is smaller the better,

$$RS_i = \frac{\max_j(E_i) - E_{i,j}}{\max_j(E_i) - \min_j(E_i)}$$

If the given metric is larger the better,

$$RS_i = \frac{E_{i,j} - \min_j(E_i)}{\max_j(E_i) - \min_j(E_i)}$$

where $E_{i,j}$ is the value of the i th performance metric of the j th GCM. The $\min_j(E_i)$ and $\max_j(E_i)$ indicate the minimum and maximum values across all models. Finally, the total score (*TS*) for each GCM was obtained

by weighting the sum of all metrics to indicate the credibility of model performance.

$$TS = \sum_{i=1}^P w_i \times RS_i$$

where w_i indicates the weight of each performance metric.

While there may be differences among metrics, however, equal weighting is still a valuable method to assess climate model performance and is used in this study (Guo et al., 2022) and for the case of *TS* - the lower value indicates better simulation performance with the model values closer to observation.

Taylor Diagrams

Taylor diagrams (Taylor, 2001) were used to summarize the degree of correspondence between model simulations and observations considering precipitation, and maximum and minimum temperatures. These diagrams typically combine three metrics: the correlation coefficient, the standard deviation, and the RMSE (Heo et al., 2014; Rivera and Arnould, 2020). The uneven angular coordinate corresponds to the *R*; the radial distance from the origin represents the ratio of the standard deviation of the simulation to that of the observation; and the distance from the observations is a measure of the RMSE. Resultantly, the best model simulation results are when the *R* and the standard deviation are equal to 1 and the RMSE is close to 0.

Results and Discussion

Ranking of GCMs

Precipitation: The performance metrics for precipitation of all CMIP6 GCMs and the ranking score (*RS*) and the total score (*TS*) are shown in Table 2. Based on *TS*, it is seen that the five best-ranked GCMs are HadGEM3-GC31-LL, ACCESS-CM2, CanESM5, NESM3, and CanESM5-CanOE with *TS* of 0.00, 8.76, 13.76, 14.17 and 15.6, respectively. It can also be seen that the better-ranked GCMs in particular metrics (*RS*) may not be the best performance based on *TS*. For example, FIO-ESM-2-0 and ACCESS-ESM1-5 show better performance in terms of *PBIAS* (*RS* of 8.70 and 7.90, respectively) but were classified as the worst performing on other metrics; INM-CM5-0 presents a good performance in terms of *R*², and MRI-ESM2-0 performs good of *VE*. However, these GCMs were not ranked within the top five based on *TS*.

Table 2: Performance metrics and ranking of precipitation CMIP6 GCMs based on *RS* and *TS*

<i>GCMs</i>	<i>Ranking score (RS)</i>					<i>Total score (TS)</i>
	<i>NRMSE</i>	<i>PBIAS</i>	<i>NSE</i>	<i>R²</i>	<i>VE</i>	
HadGEM3-GC31-LL	0.00	0.00	0.00	0.00	0.00	0.00
MIROC-ES2L	35.60	20.40	0.88	0.26	0.20	57.34
MIROC6	26.10	16.30	0.62	0.16	0.15	43.33
MRI-ESM2-0	17.00	11.00	0.39	0.11	0.09	28.59
GFDL-ESM4	19.40	11.90	0.45	0.14	0.11	32.00
BCC-ESM1	25.00	15.10	0.59	0.17	0.13	40.99
NESM3	11.00	2.70	0.24	0.18	0.05	14.17
FIO-ESM-2-0	21.30	8.70	0.50	0.27	0.14	30.91
CanESM5	9.00	4.50	0.20	0.01	0.05	13.76
CanESM5-CanOE	11.00	4.20	0.24	0.11	0.07	15.62
ACCESS-CM2	5.90	2.60	0.13	0.10	0.03	8.76
ACCESS-ESM1-5	13.20	7.90	0.29	0.12	0.07	21.58
CNRM-CM6-1	25.00	14.40	0.59	0.19	0.13	40.31
CNRM-ESM2-1	21.60	11.40	0.50	0.22	0.11	33.83
CMCC-ESM2	27.40	15.10	0.65	0.18	0.16	43.49
INM-CM5-0	33.70	13.40	0.83	0.49	0.20	48.62

Maximum temperature: The performance metrics for maximum temperature for all selected CMIP6 GCMs and their ranking and total scores are shown in Table 3. Based on *TS*, the top five best-ranked GCMs were MIROC6, ACCESS-ESM1-5, MIROC-ES2L, CanESM5, CanESM5-CanOE with *TS* of 0.17; 18.47; 40.26; 70.55; and 72.07, respectively. Whilst, the lowest-ranked GCM for maximum temperature using the *TS* method was NESM3.

Minimum temperature: The performance metrics for the minimum temperature of all selected CMIP6 GCMs and their ranking score and the total score is shown in Table 4. Based on *TS*, it can be seen that for minimum temperature, the top five best-ranked GCMs were CNRM-CM6-1, CNRM-ESM2-1, GFDL-ESM4, NESM3 and INM-CM5-0 with *TS* of 14.78; 28.65; 78.52; 92.93; and 95.76, respectively. The lowest-ranking GCM for minimum temperature using the *TS* method was ACCESS-ESM1-5.

Based on the results presented in Tables 2, 3, and 4, it can be seen that each simulation performed differently over the different metrics. No single model was seen to be best for all metrics. For example, in the study area, INM-CM5-0 showed the best results for best R^2 and *VE* (*RS* of 0.10 and 0.50, respectively), but the worst in terms of *NRMSE* (*RS* of 165.8) and *PBIAS* (*RS* of 8.7) for maximum temperature; however, it was ranked in the

best for minimum temperature (*NRMSE* (*RS* of 85.7), *PBIAS* (*RS* of 7.4)). Moreover, NESM3 was ranked in the best group for precipitation results (*NRMSE* (*RS* of 11) and *PBIAS* (*RS* of 2.7) and *NSE* (*RS* of 0.24)) and minimum temperature (*NRMSE* (*RS* of 82) and *PBIAS* (*RS* of 8.2) and *NSE* (*RS* of 2.36)); however, this model performed worst in simulating maximum temperature (*NRMSE* (*RS* of 255.8) and *PBIAS* (*RS* of 13.4) and *NSE* (*RS* of 15.8)).

It is clearly indicated that the *TS* method was applied in this study to demonstrate its efficiency to aggregate multi-model ensemble GCMs based on the different statistical indicators which are sometimes contradictory as per the above discussion. This suggests the application of other criteria such as PDF curves, TD, and mean monthly precipitation/maximum and minimum temperatures to compare the GCMs with the corresponding observations are shown to be efficient as they all support the findings from the *TS*.

Comparison of the Mean Monthly GCMs Data with the Observed Values

The mean monthly precipitation, maximum temperature, and minimum temperature for the GCMs compared to the observed data from 1980 to 2014 are presented in Figures 2a, b, and c, respectively. For precipitation (Figure 2a), most of the GCMs were found to perform to an acceptable degree during the dry season from

Table 3: Performance metrics and ranking of maximum temperature CMIP6 GCMs based on *RS* and *TS*

<i>GCMs</i>	<i>Ranking score (RS)</i>					<i>Total score (TS)</i>
	<i>NRMSE</i>	<i>PBIAS</i>	<i>NSE</i>	<i>R²</i>	<i>VE</i>	
HadGEM3-GC31-LL	73.00	4.50	3.17	0.21	0.04	80.92
MIROC-ES2L	36.20	2.40	1.44	0.20	0.02	40.26
MIROC6	0.00	0.00	0.00	0.17	0.00	0.17
MRI-ESM2-0	67.20	4.20	2.88	0.17	0.04	74.49
GFDL-ESM4	131.20	7.30	6.47	0.24	0.07	145.28
BCC-ESM1	96.70	4.90	4.43	0.50	0.05	106.58
NESM3	255.80	13.40	15.81	0.24	0.13	285.38
FIO-ESM-2-0	115.60	6.30	5.52	0.42	0.06	127.90
CanESM5	63.00	4.10	2.68	0.15	0.04	69.97
CanESM5-CanOE	64.90	4.20	2.77	0.16	0.04	72.07
ACCESS-CM2	70.80	4.20	3.07	0.41	0.04	78.52
ACCESS-ESM1-5	15.90	1.80	0.60	0.16	0.01	18.47
CNRM-CM6-1	126.10	7.30	6.16	0.00	0.07	139.63
CNRM-ESM2-1	110.90	6.50	5.25	0.09	0.06	122.80
CMCC-ESM2	81.30	4.60	3.60	0.43	0.04	89.97
INM-CM5-0	165.80	8.70	8.75	0.51	0.08	183.84

Table 4: Performance metrics and ranking of minimum temperature CMIP6 GCMs based on *RS* and *TS*

<i>GCMs</i>	<i>Ranking score (RS)</i>					<i>Total score (TS)</i>
	<i>NRMSE</i>	<i>PBIAS</i>	<i>NSE</i>	<i>R²</i>	<i>VE</i>	
HadGEM3-GC31-LL	121.20	4.60	3.96	0.04	0.10	129.90
MIROC-ES2L	95.00	6.70	2.86	0.08	0.08	104.72
MIROC6	87.20	7.40	2.55	0.07	0.07	97.29
MRI-ESM2-0	108.90	5.50	3.42	0.02	0.09	117.93
GFDL-ESM4	67.90	8.70	1.86	0.00	0.06	78.52
BCC-ESM1	139.80	3.30	4.83	0.09	0.11	148.13
NESM3	82.00	8.20	2.36	0.31	0.06	92.93
FIO-ESM-2-0	98.30	6.30	2.99	0.02	0.08	107.69
CanESM5	142.20	2.90	4.95	0.00	0.11	150.16
CanESM5-CanOE	139.70	3.10	4.83	0.00	0.11	147.74
ACCESS-CM2	152.70	2.10	5.47	0.02	0.12	160.41
ACCESS-ESM1-5	181.80	0.00	7.04	0.03	0.14	189.01
CNRM-CM6-1	0.00	14.70	0.00	0.08	0.00	14.78
CNRM-ESM2-1	14.80	13.40	0.33	0.11	0.01	28.65
CMCC-ESM2	133.50	3.70	4.53	0.05	0.11	141.89
INM-CM5-0	85.70	7.40	2.50	0.09	0.07	95.76

December to April, as this is a period of little variability. However, high variability in the GCMs precipitation compared to the observation during the wet season (May to November) in the VMD was seen. It can be seen that all 16 GCMs significantly underestimated the

precipitation during 1980–2014. Compared the results from a previous study in Asia undertaken by Tiwari et al. (2023) evaluating the Northeast monsoon precipitation over India, indicated that monsoon precipitation poses strong month-to-month variability, and changes in the

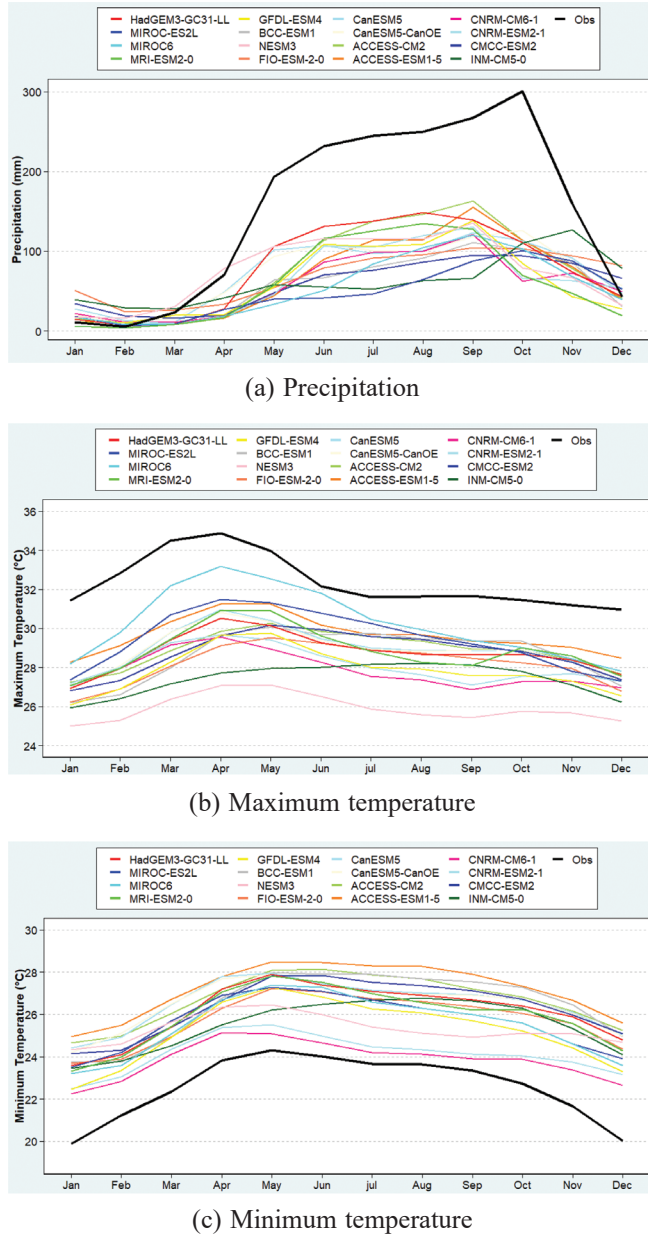


Figure 2: The mean monthly (a) precipitation, (b) maximum temperature, and (c) minimum temperature for the GCMs compared to the observation for 1980–2014.

individual month might also be different. In addition, for maximum and minimum temperatures (Figure 2b and c), it is interesting that all 16 GCMs showed a significant overestimation of both the maximum and minimum temperature for the period 1980–2014.

Performance Assessment

The Taylor diagram provides a good statistical summary of R^2 , standard deviations (SD), and RMSE between the simulation and the observed data. In this study, the performances of GCMs precipitation, maximum

temperature, and minimum temperatures, respectively in comparison to their corresponding observations. Though the SD were all different, the correlation between the simulation and observation of most of the GCMs was found to range from 0.6 to 0.8 for the precipitation, from 0.5 to 0.7 for maximum temperature, and 0.7 to 0.8 for minimum temperature (Figure 3). All these ranges are acceptable for future predictions based on the proposed ranges of Moriasi et al. (2015). It is clearly indicated that the highest-ranked GCMs show higher correlations. It is also shown that most of the GCMs have lower SD compared to their corresponding observations.

Comparison Using Probability Density Function (PDF)

The PDFs of the mean monthly precipitation, maximum temperature, and minimum temperature modeled results compared with actual observation for the period 1980–2014 in the VMD are presented in Figures 4a,b, and c, respectively. It can be seen that most of the GCMs were not able to capture the precipitation and maximum temperature properties accurately, in particular the mean. However, the distribution of the modeled precipitation results relative to the observation data varies more for all 16 GCMs which all showed a significant underestimation of the precipitation values. Additionally, all 16 selected GCMs showed underestimations of maximum temperature. The PDF distributions show that the majority of the maximum temperature distributions are around 29°C. For the minimum temperature, PDFs of the GCMs compared to the observation show an overestimation in all 16 GCMs with the majority of values around 27°C.

Selection of Multi-Models for Future Projection

The different results from each metric have been summed up to provide the overall performance of each model suite using the results of *RS*. GCMs performance was finally ranked by summarising all relative error values derived from all indicators (*TS*) - the lower the value indicates better model performance, with modeled values closer to those observed. The final ranking orders of the 16 GCMs are shown in Figure 5.

It can be seen that for precipitation, the best-ranking GCMs were HadGEM3-GC31-LL, ACCESS-CM2, CanESM5, NESM3 and CanESM5-CanOE; while MIROC6, ACCESS-ESM1-5, MIROC-ES2L, CanESM5 and CanESM5-CanOE were ranked best in terms of modeled maximum temperature. Whereas for minimum temperature, CNRM-CM6-1, CNRM-ESM2-1, GFDL-ESM4, NESM3 and INM-CM5-0 were ranked top.

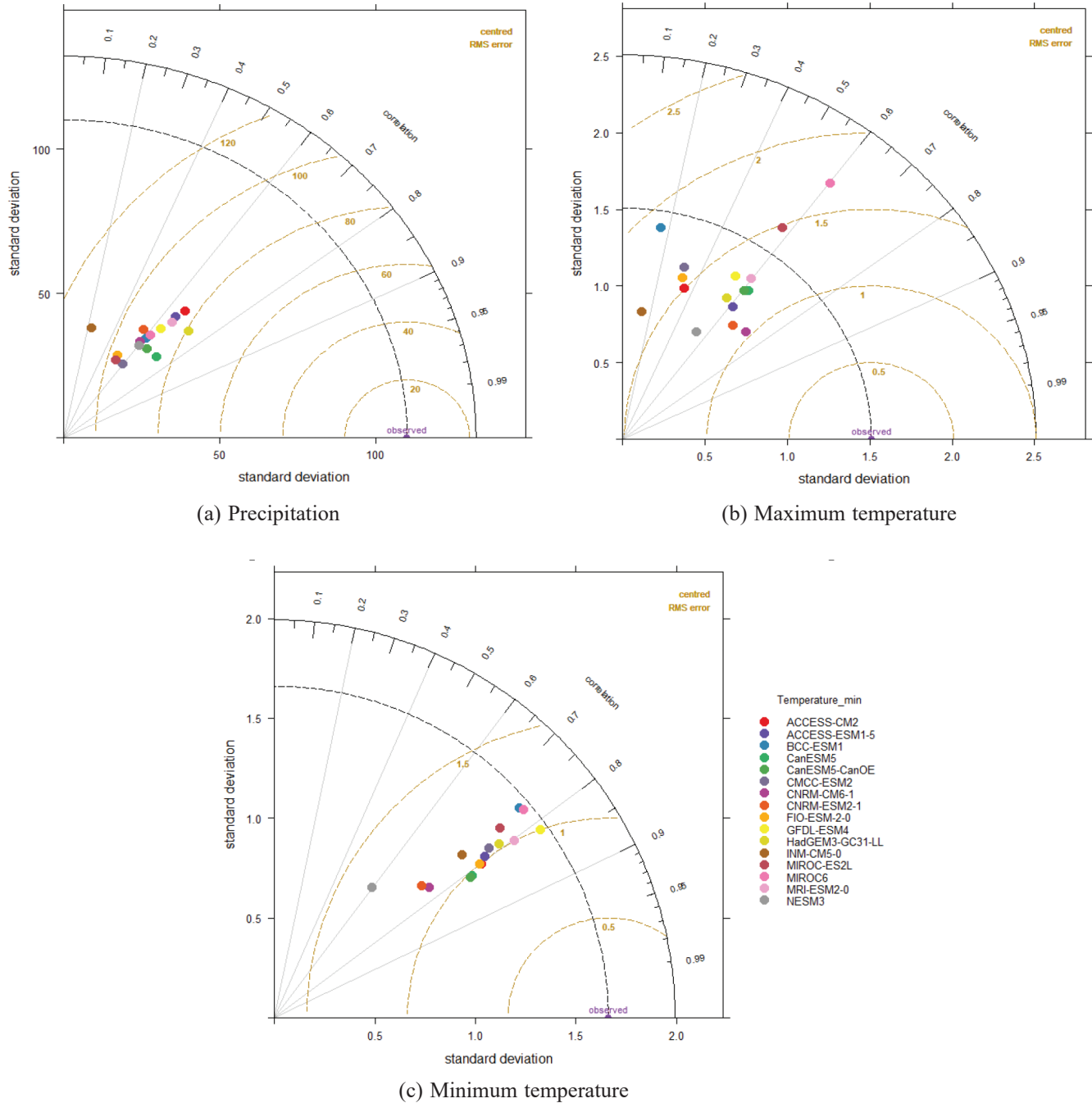


Figure 3: Taylor diagrams represent the correlation between the corresponding observations to (a) GCM precipitation, (b) GCM maximum temperature, and (c) GCM minimum temperature.

Conclusions

This study assessed the performances of 16 selected GCMs of the CMIP6 suite for their modeling performance of precipitation, maximum temperature, and minimum temperature over the VMD using five statistical indicators, NRMSE, PBIAS, NSE, R^2 , and

VE, and supplemented by PDF plots and TD for the period 1980–2014. The results highlight that the different statistical indicators reveal different ranking orders for all 16 GCMs. Based on RS ranking, it can be seen that each simulation GCM performed differently on the differing metrics, and no single model performed best on all metrics.

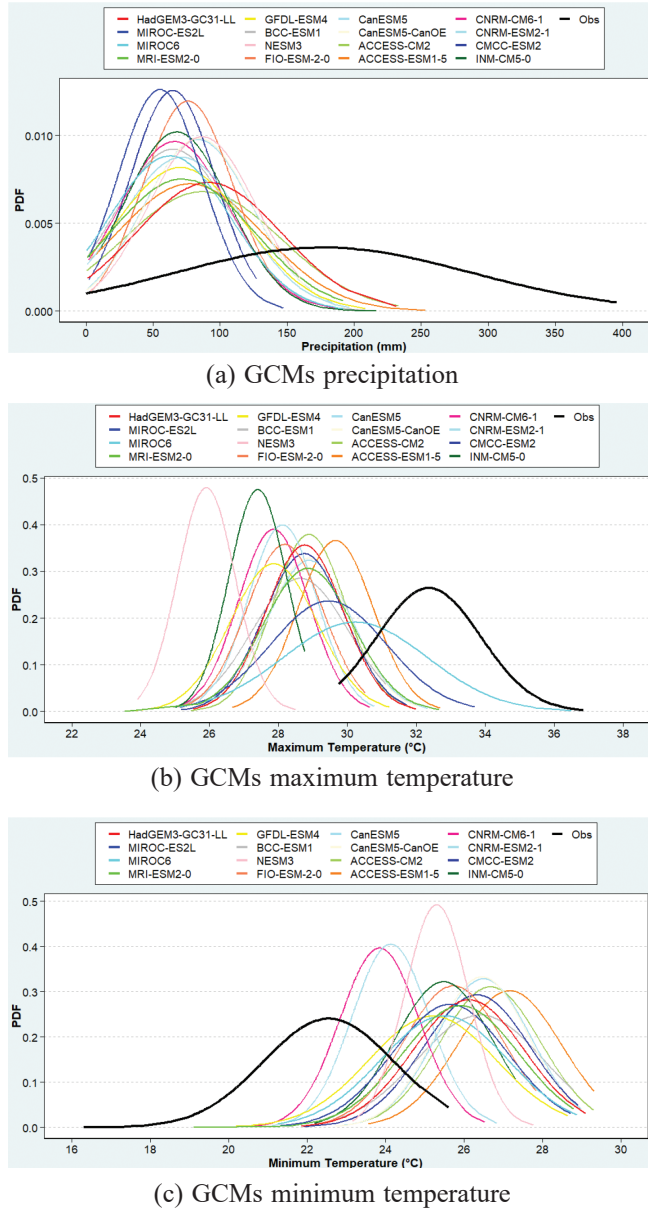


Figure 4: PDF comparison between mean monthly observation and (a) GCMs precipitation, (b) GCMs maximum temperature, and (c) GCMs minimum temperature.

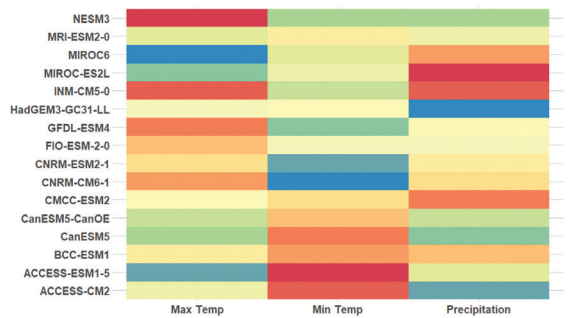


Figure 5: Final ranking of the 16 selected GCMs based on TS results.

The top five highest ranked GCMs-based on TS were HadGEM3-GC31-LL, ACCESS-CM2, CanESM5, NESM3 and CanESM5-CanOE for precipitation; and CNRM-CM6-1, CNRM-ESM2-1, GFDL-ESM4, NESM3 and INM-CM5-0 for the maximum; and CNRM-CM6-1, CNRM-ESM2-1, GFDL-ESM4, NESM3 and INM-CM5-0 for minimum temperatures. It is also observed that generally there was an underestimation of precipitation and an overestimation of temperature. The TS method demonstrated efficiency to aggregate a multi-model ensemble GCMs based on the different statistical indicators which were often contradictory.

The findings from this study provide an understanding and insights into the understanding of future climate variability for the VMD based on CMIP6. Accurate data on future climate change from GCMs are a very important tool in not only assessing risks but also addressing future climate resilience. Going forward, the findings from this study provide useful information for the selection of GCMs for the VMD in particular since the simulation of precipitation values in tropical areas is complex due to both temporal and spatial characteristics.

References

- Alamgir, M., Mohsenipour, M., Homs, R., Wang, X., Shahid, S., Shiru, M.S., Alias, N.E. and Yuzir, A., 2019. Parametric assessment of seasonal drought risk to crop production in Bangladesh. *Sustainability*, **11**(5): 1442.
- Asdak, C. and Supian, S., 2018. Watershed management strategies for flood mitigation: a case study of Jakarta's flooding. *Weather and climate extremes*, **21**: 117-122.
- Ayugi, B., Tan, G., Niu, R., Dong, Z., Ojara, M., Mumo, L., Babausmail, H. and Ongoma, V., 2020. Evaluation of meteorological drought and flood scenarios over Kenya, East Africa. *Atmosphere*, **11**(3): 307.
- Ayugi, B., Zhihong, J., Zhu, H., Ngoma, H., Babausmail, H., Rizwan, K. and Dike, V., 2021. Comparison of CMIP6 and CMIP5 models in simulating mean and extreme precipitation over East Africa. *International Journal of Climatology*, **41**(15): 6474-6496.
- Cui, T., Li, C. and Tian, F., 2021. Evaluation of temperature and precipitation simulations in CMIP6 models over the Tibetan Plateau. *Earth and Space Science*, **8**(7): e2020EA001620.
- Desmet, Q. and Ngo, D.T., 2022. A novel method for ranking CMIP6 global climate models over the southeast Asian region. *International Journal of Climatology*, **42**(1): 97-117.

- ESGF, 2021. Earth System Grid Federation. <https://esgf-node.llnl.gov/search/cmip6/> (accessed 15 December 2022).
- Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J. and Taylor, K.E., 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, **9**(5): 1937-1958.
- Ge, F., Zhu, S., Luo, H., Zhi, X. and Wang, H., 2021. Future changes in precipitation extremes over Southeast Asia: insights from CMIP6 multi-model ensemble. *Environmental Research Letters*, **16**(2): 024013.
- Grose, M.R., Narsey, S., Delage, F.P., Dowdy, A.J., Bador, M., Bosch, G., Chung, C., Kajtar, J.B., Rauniyar, S., Freund, M.B. and Lyu, K., 2020. Insights from CMIP6 for Australia's future climate. *Earth's Future*, **8**(5): e2019EF001469.
- Guo, H., Bao, A., Chen, T., Zheng, G., Wang, Y., Jiang, L. and De Maeyer, P., 2021. Assessment of CMIP6 in simulating precipitation over arid Central Asia. *Atmospheric Research*, **252**: 105451.
- Guo, Y., Xu, Y.P., Yu, X., Xie, J., Chen, H. and Si, Y., 2023. Impacts of GCM credibility on hydropower production robustness under climate change: CMIP5 vs CMIP6. *Journal of Hydrology*, **618**: 129233.
- Gupta, H.V., Sorooshian, S. and Yapo, P.O., 1999. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *Journal of hydrologic engineering*, **4**(2): 135-143.
- Heo, K.Y., Ha, K.J., Yun, K.S., Lee, S.S., Kim, H.J. and Wang, B., 2014. Methods for uncertainty assessment of climate models and model predictions over East Asia. *International Journal of Climatology*, **34**(2): 377-390.
- Iqbal, Z., Shahid, S., Ahmed, K., Ismail, T., Ziarh, G.F., Chung, E.S. and Wang, X., 2021. Evaluation of CMIP6 GCM rainfall in mainland Southeast Asia. *Atmospheric Research*, **254**: 105525.
- Khoi, D.N., Sam, T.T., Chi, N.T.T., Linh, D.Q. and Nhi, P.T.T., 2022. Impact of future climate change on river discharge and groundwater recharge: A case study of Ho Chi Minh City, Vietnam. *Journal of Water and Climate Change*, **13**(3): 1313-1325.
- Kumar, P., Mishra, A.K., Dubey, A.K., Javed, A., Saharwardi, M.S., Kumari, A., Sachan, D., Cabos, W., Jacob, D. and Sein, D.V., 2022. Regional earth system modeling framework for CORDEX-SA: An integrated model assessment for Indian summer monsoon rainfall. *Climate Dynamics*, **59**(7-8): 2409-2428.
- Lavane, K., Kumar, P., Meraj, G., Han, T. G., Ngan, L. H. B., Lien, B. T. B., et al., 2023. Assessing the effects of drought on rice yields in the Mekong Delta. *Climate*, **11**(1): 13.
- Legates, D.R. and McCabe Jr, G.J., 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water resources research*, **35**(1): 233-241.
- Manawi, S.M.A., Nasir, K.A.M., Shiru, M.S., Hotaki, S.F. and Sediqi, M.N., 2020. Urban flooding in the northern part of Kabul City: Causes and mitigation. *Earth Systems and Environment*, **4**: 599-610.
- Minh, H.V.T., Lavane, K., Ty, T.V., Downes, N.K., Hong, T.T.K. and Kumar, P., 2022. Evaluation of the impact of drought and saline water intrusion on rice yields in the Mekong Delta, Vietnam. *Water*, **14**(21): 3499.
- Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D. and Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, **50**(3): 885-900.
- Moriasi, D.N., Gitau, M.W., Pai, N. and Daggupati, P., 2015. Hydrologic and water quality models: Performance measures and evaluation criteria. *Transactions of the ASABE*, **58**(6): 1763-1785.
- Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I-a discussion of principles. *J. Hydrol.*, **10**: 282-290.
- Pandey, L.K. and Dwivedi, S., 2021. Comparing the performance of turbulent kinetic energy and K-profile parameterization vertical parameterization schemes over the tropical Indian Ocean. *Marine Geodesy*, **44**(1): 42-69.
- Phuong, D.N.D., Linh, V.T., Nhat, T.T., Dung, H.M. and Loi, N.K., 2019. Spatiotemporal variability of annual and seasonal rainfall time series in Ho Chi Minh city, Vietnam. *Journal of Water and Climate Change*, **10**(3): 658-670.
- Piman, T., Pawattana, C., Vansarochana, A., Aekakkararungroj, A. and Hormwichian, R., 2016. Analysis of historical changes in rainfall in Huai Luang watershed, Thailand. *International Journal of Technology*, **7**(7): 1155-1162.
- Pimonsree, S., Kamworapan, S., Gheewala, S.H., Thongbhakdi, A. and Prueksakorn, K., 2023. Evaluation of CMIP6 GCMs performance to simulate precipitation over Southeast Asia. *Atmospheric Research*, **282**: 106522.
- Rivera, J.A. and Arnould, G., 2020. Evaluation of the ability of CMIP6 models to simulate precipitation over Southwestern South America: Climatic features and long-term trends (1901–2014). *Atmos. Res.*, **241**: 104953 <https://doi.org/10.1016/j.atmosres.2020.104953>.
- Shiru, M.S. and Chung, E.S., 2021. Performance evaluation of CMIP6 global climate models for selecting models for climate projection over Nigeria. *Theoretical and Applied Climatology*, **146**(1-2): 599-615.
- Shiru, M.S., Shahid, S., Dewan, A., Chung, E.S., Alias, N., Ahmed, K. and Hassan, Q.K., 2020. Projection of meteorological droughts in Nigeria during growing seasons under climate change scenarios. *Scientific Reports*, **10**(1): 1-18.
- Song, Y., Li, X., Bao, Y., Song, Z., Wei, M., Shu, Q. and Yang, X., 2020. FIO-ESM v2. 0 Outputs for the CMIP6 global monsoons model intercomparison project

- experiments. *Advances in Atmospheric Sciences*, **37**: 1045-1056.
- Srinivasa Raju, K., Sonali, P. and Nagesh Kumar, D., 2017. Ranking of CMIP5-based global climate models for India using compromise programming. *Theoretical and Applied Climatology*, **128**: 563-574.
- Stouffer, R.J., Eyring, V., Meehl, G.A., Bony, S., Senior, C., Stevens, B. and Taylor, K.E., 2017. CMIP5 scientific gaps and recommendations for CMIP6. *Bulletin of the American Meteorological Society*, **98**(1): 95-105.
- Supharatid, S., Nafung, J. and Aribarg, T., 2022. Projected changes in temperature and precipitation over mainland Southeast Asia by CMIP6 models. *Journal of Water and Climate Change*, **13**(1): 337-356.
- Tan, G., Ayugi, B., Ngoma, H. and Ongoma, V., 2020. Projections of future meteorological drought events under representative concentration pathways (RCPs) of CMIP5 over Kenya, East Africa. *Atmospheric Research*, **246**: 105112.
- Tanveer, M.E., Lee, M.H. and Bae, D.H., 2016. Uncertainty and reliability analysis of CMIP5 climate projections in South Korea using REA method. *Procedia Engineering*, **154**: 650-655.
- Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, **106**(D7): 7183-7192.
- Taylor, K.E., Stouffer, R.J. and Meehl, G.A., 2012. An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, **93**(4): 485-498.
- Thoeun, H.C., 2015. Observed and projected changes in temperature and rainfall in Cambodia. *Weather and Climate Extremes*, **7**: 61-71.
- Tinh, N.T., Ty, T.V. and Minh, H.V.T., 2016. Evaluation and selection of global climate changes models (GCMs-CMIP5) for the Mekong delta. *Can Tho University Journal of Sciences*, **(42)**: 81-90.
- Tiwari, R., Mishra, A.K., Rai, S. and Pandey, L.K., 2023. Evaluation and projection of northeast monsoon precipitation over India under higher warming scenario: A multimodel assessment of CMIP6. *Theoretical and Applied Climatology*, **151**(1): 859-870.
- Tran, T.A., 2019. Land use change driven out-migration: Evidence from three flood-prone communities in the Vietnamese Mekong Delta. *Land Use Policy*, **88**: 104157.
- Ty, T.V., Lavane, K., Nguyen, P.C., Downes, N.K., Nam, N.D.G., Minh, H.V.T. and Kumar, P., 2022. Assessment of relationship between climate change, drought, and land use and land cover changes in a semi-mountainous area of the Vietnamese Mekong Delta. *Land*, **11**(12): 2175.
- Vietnam News Agency. 2022. Overview of the Mekong Delta region.
- Wang, L., Zhang, J., Shu, Z., Wang, Y., Bao, Z., Liu, C., Zhou, X. and Wang, G., 2021. Evaluation of the ability of CMIP6 global climate models to simulate precipitation in the Yellow River Basin, China. *Frontiers in Earth Science*, **2021**: 1009.
- Willmott, C.J., 1982. Some comments on the evaluation of model performance. *Bulletin of the American Meteorological Society*, **63**(11): 1309-1313.
- Woldemeskel, F.M., Sharma, A., Sivakumar, B. and Mehrotra, R., 2014. A framework to quantify GCM uncertainties for use in impact assessment studies. *Journal of Hydrology*, **519**: 1453-1465.
- Xin, X., Wu, T., Zhang, J., Yao, J. and Fang, Y., 2020. Comparison of CMIP6 and CMIP5 simulations of precipitation in China and the East Asian summer monsoon. *International Journal of Climatology*, **40**(15): 6423-6440.
- Yazdandoost, F., Moradian, S., Izadi, A. and Aghakouchak, A., 2021. Evaluation of CMIP6 precipitation simulations across different climatic zones: Uncertainty and model intercomparison. *Atmospheric Research*, **250**: 105369.
- Zhai, J., Mondal, S.K., Fischer, T., Wang, Y., Su, B., Huang, J., Tao, H., Wang, G., Ullah, W. and Uddin, M.J., 2020. Future drought characteristics through a multi-model ensemble from CMIP6 over South Asia. *Atmospheric Research*, **246**: 105111.
- Zhao, C., Jiang, Z., Sun, X., Li, W. and Li, L., 2020. How well do climate models simulate regional atmospheric circulation over East Asia? *International Journal of Climatology*, **40**(1): 220-234.
- Zhou, T., Chen, X., Wu, B., Guo, Z., Sun, Y., Zou, L., Man, W., Zhang, L. and He, C., 2017. A robustness analysis of CMIP5 models over the East Asia-Western North Pacific domain. *Engineering*, **3**(5): 773-778.