# Bias Correction ofMonthly Precipitation from different General Circulation Models using Cumulative Density Function in Raipur District,Chhattisgarh, India


## ABSTRACT

Accurate representation of precipitation patterns is crucial for understanding and adapting to these impacts. General Circulation Models (GCMs) are essential for projecting future climate scenarios but often exhibit biases in simulating precipitation, undermining the reliability of their outputs. This study focused on bias correction of monthly precipitation data from different GCMs using Cumulative Density Functions (CDFs). Bias correction techniques were employed to align model-simulated precipitation with observed data, revealing significant improvements in the accuracy of future precipitation projections. The study area, Raipur, characterized by diverse topography, served as the location for analysis. Three GCMs were selected based on their availability and participation in the CMIP6 experiment. The bias correction process involved the calculation of CDFs and equiprobability transformations, resulting in a closer match between model predictions and observations. Results showed substantial variability in monthly precipitation values across different climate models and scenarios, with distinct seasonal patterns observed. Inter-model discrepancies underscored the complexities of precipitation simulations, highlighting the need for careful interpretation of model outputs. Continued research efforts were crucial for improving the accuracy and reliability of climate model simulations for informed decision-making and planning in climate-sensitive sectors.



Key words: Bias correction, cumulative density function, General Circulation model


## 1.Introduction


UNDER PEER REVIEW

The ongoing emission of greenhouse gases (GHGs) has led to a global increase in temperatures, resulting in climate change. This shift has profound implications for various sectors crucial to human life, such as water resources, agriculture, health, and energy (IPCC, 2021). The accurate representation of precipitation patterns is crucial for understanding and adapting to the impacts of climate change across various sectors. General Circulation Models (GCMs) are indispensable tools for projecting future climate scenarios, yet they often exhibit biases in simulating precipitation due to complex atmospheric processes, spatial resolution limitations, and parameterizations, undermining the reliability of their outputs for decision-making purposes (Piani *et al.,* 2010).Precipitation plays a fundamental role in shaping ecosystems, water resources, agriculture, and socio-economic activities. Reliable projections of future precipitation patterns are essential for assessing and mitigating the impacts of climate change, such as droughts, floods, and shifts in agricultural productivity. These biases can significantly affect the accuracy of climate projections and hinder effective decision-making in climate-sensitive sectors.It's widely acknowledged that precipitation data from general circulation models (GCMs) requires prior bias correction to be useful for driving hydrological or impact models effectively. Accurate representation of precipitation fields in future climate projections is critical for impact studies and hydrological modelling by utilizing transfer functions derived from observed and simulated cumulative distribution functions (cdfs) (Ines and Hansen, 2006; Piani *et al.,* 2010;). In recent years, bias correction techniques have emerged as essential methods for improving the fidelity of GCM-simulated precipitation by aligning them with historical observations. Bias corrected CMIP6 projection were utilized by various studies. Bias-corrected CMIP6 projections were used to estimate the frequency of rainfall and temperature extremes for Godavari

UNDER PEER REVIEW

river basin during near-, mid- and end-21st century and reported higher frequencies for the far period than the near-term climate (Mishra et al., 2020).Similarly the bias corrected CMIP6 projection were utilized by Saha et al., (2022) for analyzing precipitation extremes.Verma et al., (2023) adjusted raw CMIP5 and CMIP6 outputs before giving inputs in SWAT model.When applying a correction derived from historical data to projected climate simulations, one must assume that the correction remains valid for the future climate. This assumption is more acceptable if the transfer function between raw and corrected GCM output is robust, particularly if it depends on fewer parameters derived from the data.

Cumulative Density Functions offer a robust statistical framework for quantifying and comparing the distribution of precipitation intensities between observed data and GCM simulations. By analyzing discrepancies in the CDFs, researchers can identify systematic biases in GCM-simulated precipitation and develop correction techniques to align the simulated distributions with observed patterns. The use of CDFs facilitates a comprehensive assessment of biases across different precipitation intensity ranges and enables targeted corrections to improve the accuracy of GCM outputs.CDF had been used in recent studies in exploring future trends of precipitation based on CMIP6 projections (Anil et al., 2024 and Wang *et al.,* 2024).

This research focuses on bias correction of monthly precipitation data from different General Circulation Models (GCMs) using Cumulative Density Functions (CDFs). The Cumulative Density Function, which represents the probability distribution of precipitation intensities, offers a valuable framework for comparing and correcting biases in GCM-simulated precipitation against observed data.

UNDER PEER REVIEW

## 2.Materials and methodology

### 2.1 Study area

Raipur, the capital of Chhattisgarh, is situated inthe agro-climatic zone, Chhattisgarh plateau. The region's elevation varies from 244 to 409 meters above mean sea level (AMSL), covering an area of 2892 square kilometers. Being a sub-humid region, Raipur receives an average annual rainfall of 1149 mm, with the southwest monsoon season contributing significantly to total annual precipitation (Khavse*et al.,* 2015). The district is characterized by diverse topography, soil types, and land use patterns.

### 2.2 General circulation models

The study incorporated three General Circulation Models (GCMs): CanESM5, MPI-ESM1.2-HR, and NorESM2-MM. These models were selected due to the availability of predictor datasets required for downscaling in the necessary format. They provided predictor datasets for both historical and future timeframes and encompassed all four considered Shared Socioeconomic Pathways (SSP) scenarios. Additionally, all three GCMs were part of the CMIP 6 (Coupled Model Intercomparison Project) experiment. The projected precipitation data from these models done by Mishra *et al.,* (2020) was downloaded from the website , https://zenodo.org/records/3873998.

Table 1 : General Circulation Models (GCMs)

| GCM | Developed by | Reference |
|---|---|---|
| MPI-ESM1-2-HR | Max Planck Institute for Meteorology, Germany | Müller *et al.,* 2018 |
| CanESM5 | Canadian Centre for climate modelling and | Swart *et al.,* |

UNDER PEER REVIEW

| | | |
|---|---|---|
| | analysis | 2019 |
| Nor ESM2-MM | Norwegian Earth System Model Version 2 | Seland *et al.*, 2020 |

### 2.3 Bias correction using CDF

Cumulative Distribution Function (CDF) and equiprobability transformation are statistical methods used to adjust or correct for any systematic biases or errors in climate models when simulating future rainfall patterns.

In this study bias correction was done by calculating cumulative density function and equiprobability transformation as following steps (Ghosh and Mujumdar 2008; Das and Umamahesh, 2016).

1. Initially, the observed (Yobs), simulated (Ysim), and projected (Ypro) data were sorted in ascending order.
2. Probabilities were assigned to each value in the sorted data, in the second column of the corresponding variable. This probability was calculated as

   $$Y_{Probability} = \frac{i}{(\text{number of data points}) + 1},$$

   where 'i' represented the index of the data point. This step created an empirical cumulative distribution function (CDF) for the data.
3. Repeated values were removed by checking the differences between two consecutive values of each variable.
4. If the difference was zero, it indicated the presence of repeating values, and those were subsequently removed.

UNDER PEER REVIEW

5. Then, plots of the observed, simulated, and projected CDFs were generated. The CDF represented the probability that a rainfall value would be less than or equal to a specific value.
6. Differences in the CDFs between observed and simulated data indicated where the model had a systematic bias.
7. To correct for this bias, the simulated data was adjusted to match the observed CDF. For this, equiprobability distribution was applied to the projected data (Ypro).
8. For each value in Ypro, its cumulative distribution value (CDFpro) was calculated by linearly interpolating between the CDF values of the simulated data (Ysim).
9. Subsequently, the CDFpro value was used to find the corresponding value from the observed data (Yobs) by linear interpolation in the inverse direction.

**3. Results and Discussions**

The cumulative density functions (CDF) before and after bias correction and projected future for MPI-ESM1-2-HR , CanESM5 and Nor ESM2-MM were represented in Fig.1, Fig.2 and Fig.3 respectively. The CDF before bias correction represented the distribution of precipitation values as simulated by the climate models without any adjustments. It reflected the inherent biases or inaccuracies in the model output compared to observational data. The CDF showed differences between model outputs and observed precipitation, indicating potential overestimation or underestimation of precipitation values by the models across different percentiles of the distribution.

The comparison of the cumulative density function (CDF) before and after bias correction revealed significant improvements in the alignment of model-simulated

UNDER PEER REVIEW

precipitation with observational data. Before bias correction, the CDF exhibited notable discrepancies from the observed distribution, indicating systematic biases in the model output. These biases were particularly evident across certain percentiles of the distribution, with the model often overestimating in case of CanESM5 and underestimating precipitation values in other GCMs compared to observations.

Following bias correction, there was a clear convergence between the model simulated CDF and the observed distribution of precipitation. The bias correction techniques effectively mitigated the systematic biases present in the raw model output, resulting in a CDF that closely matched the observed data across a range of percentiles. The corrected CDF demonstrated improved accuracy in capturing the variability and extremes of precipitation events, which were crucial for understanding and predicting regional climate patterns.

Overall, the comparison highlighted the importance of bias correction in enhancing the reliability and utility of climate model simulations for various applications, such as climate change impact assessments and resource management. The alignment of the CDFs before and after bias correction signified a more robust representation of precipitation in the model output, providing researchers and policymakers with more accurate information for decision-making and planning purposes (Piani *et al.,* 2010; Smith *et al.,* 2012).

UNDER PEER REVIEW

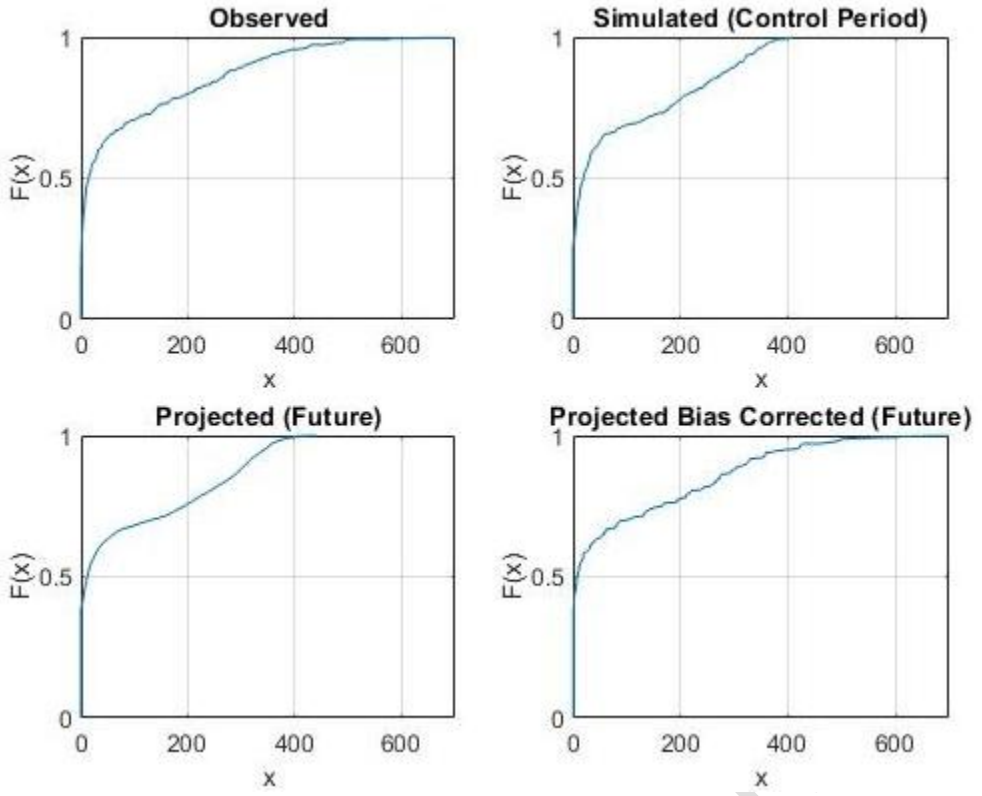


**Figure 1: The probability distribution calculated for MPI-ESM1-2-HR**

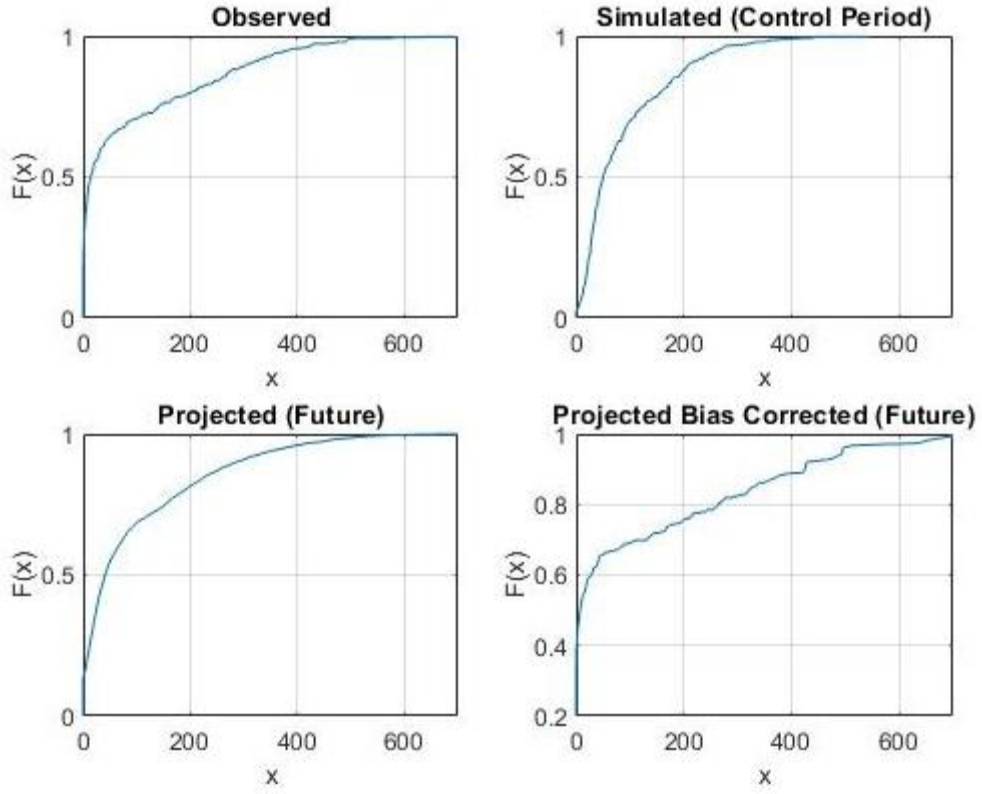


**Figure 2: The probability distribution calculated for CanESM5**

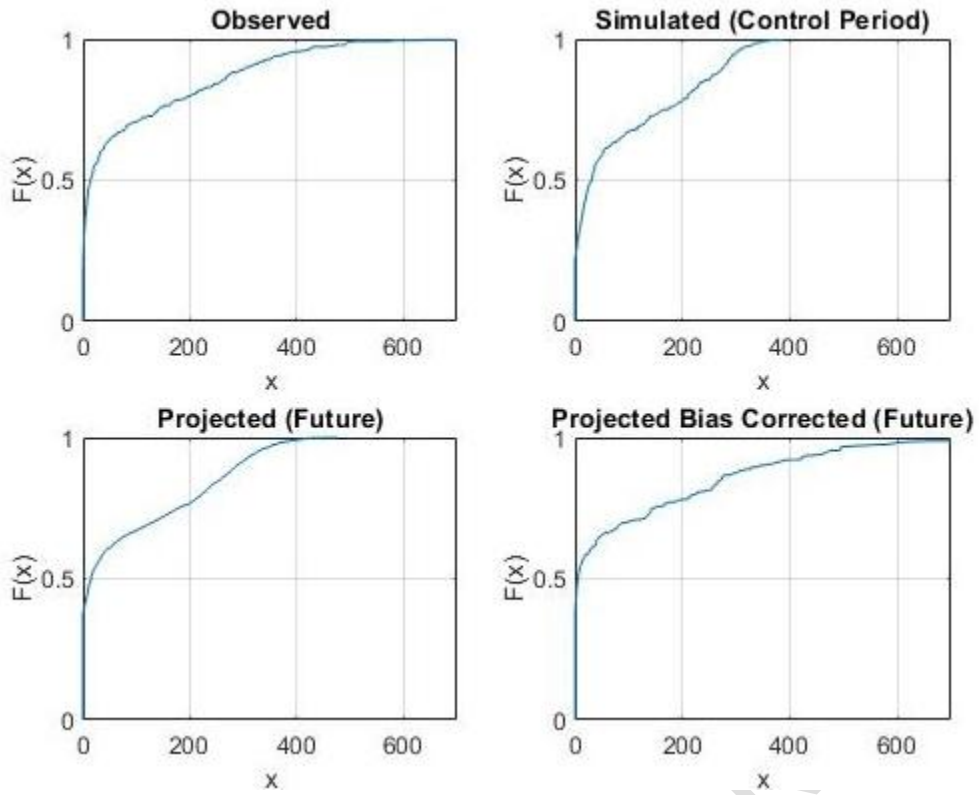


**Figure 3: The probability distribution calculated for Nor ESM2-MM**

Projected rainfall under SSP1 2.6 scenario for end century (2080-2100) after bias correction was given in Figure 4. There was considerable variability in monthly precipitation values across the three climate models. For example, in January, NorESM5 recorded substantially higher precipitation (17.13 mm) compared to MPI-ESM1-2 HR (6.87 mm) and CanESM5 (5.96 mm). Each model exhibited distinct seasonal patterns in precipitation. For instance, CanESM5 showed a notable peak in precipitation during May and June, with values reaching 18.04 mm and 176.64 mm, respectively. In contrast, NorESM5 exhibited a peak in September (234.56 mm), followed by a decrease towards the end of the year. Despite differences in absolute precipitation values, there were instances of consistency across models. For example, all three models showed relatively low precipitation in March and April, followed by an increase in May and June.

UNDER PEER REVIEW

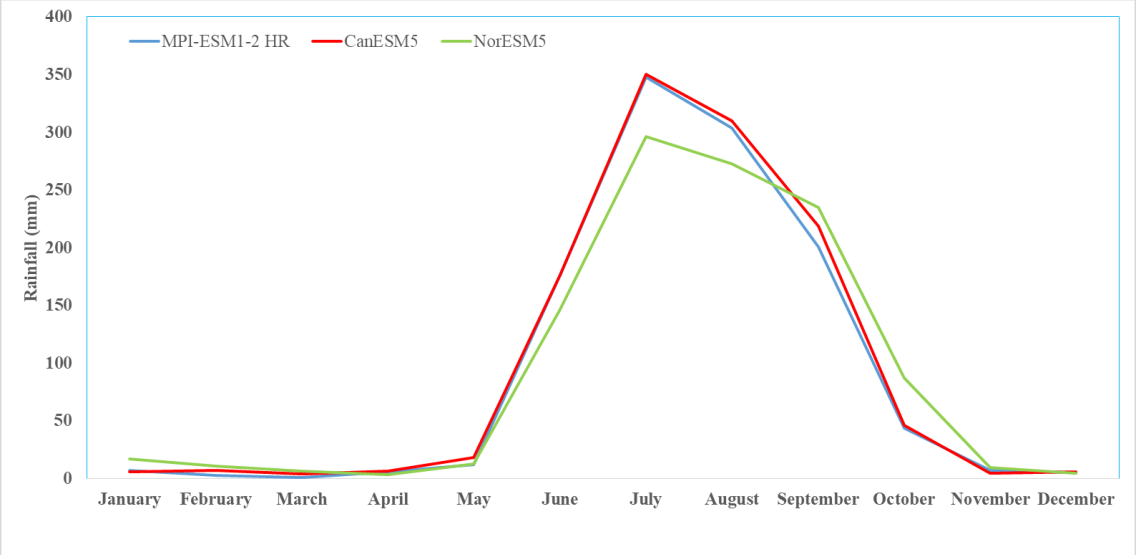


**Fig 4: Projected rainfall under SSP1 2.6 scenario for end century (2080-2100) after BC**

Projected rainfall under SSP24.5 scenario for end century (2080-2100) after bias correction was given in Figure 5. There was substantial variability in monthly precipitation values across the three climate models. For instance, in January, MPI-ESM1-2 HR recorded the highest precipitation (20.62 mm), followed by NorESM5 (4.05 mm) and CanESM5 (1.34 mm). Each model exhibited distinct seasonal patterns in precipitation. CanESM5 showed a significant peak in precipitation during June, with a value of 180.83 mm, while MPI-ESM1-2 HR and NorESM5 showed relatively lower precipitation during the same month (69.33 mm and 106.18 mm, respectively). CanESM5 recorded no precipitation in April, while MPI-ESM1-2 HR and NorESM5 had relatively low precipitation values during this month (6.06 mm and 1.85 mm, respectively). Despite differences in absolute precipitation values, there were instances of consistency across models. For example, all three models showed relatively high precipitation during July and August. However, there were also discrepancies, such as the higher precipitation values recorded by MPI-ESM1-2 HR in several months compared to CanESM5 and NorESM5.

UNDER PEER REVIEW

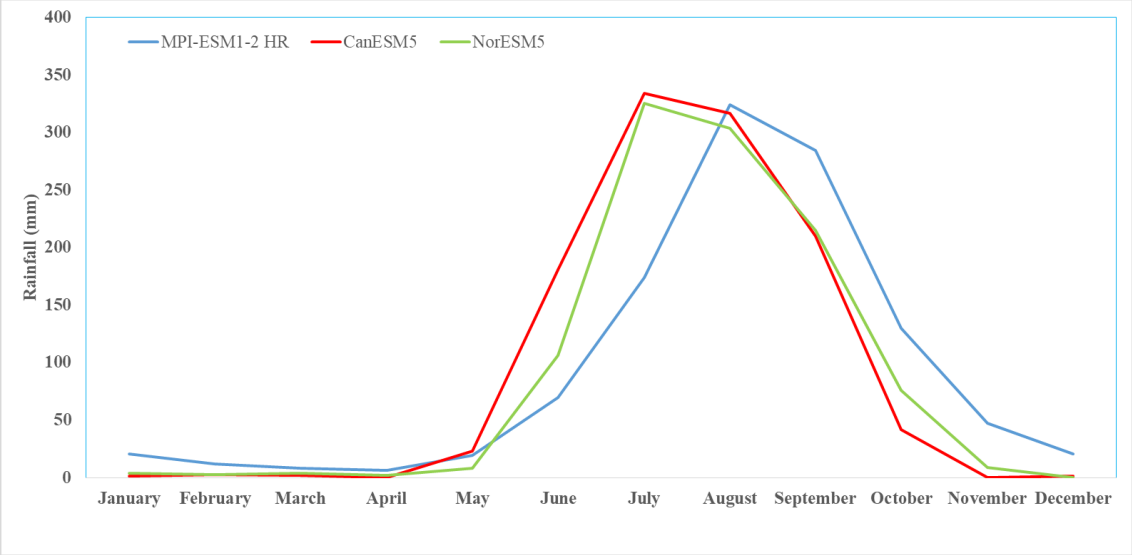


**Fig 5: Projected rainfall under SSP2 4.5 scenario for end century (2080-2100) after BC**

Projected rainfall under the SSP3 7.0 scenario for the end of the century (2080-2100) after bias correction was presented in Figure 6. There was significant variability in rainfall values across both models and months. For example, in June, the MPI-ESM1-2 HR model indicated a very high rainfall amount of 222.476 mm, whereas the CanESM5 model showed 330.386 mm for the same month. This suggested that different climate models might have produced different estimates of rainfall for the same period. Across the models, there were apparent seasonal patterns in rainfall. For instance, in many regions, rainfall tended to be higher during certain months (e.g., June and July) and lower during others (e.g., January and February). These patterns were consistent with typical seasonal variations in rainfall observed in many climates. There were notable differences in rainfall values between the models for certain months. For example, in April, the MPI-ESM1-2 HR model reported a much higher rainfall value (10.900 mm) compared to the CanESM5 model (10.892 mm).

UNDER PEER REVIEW

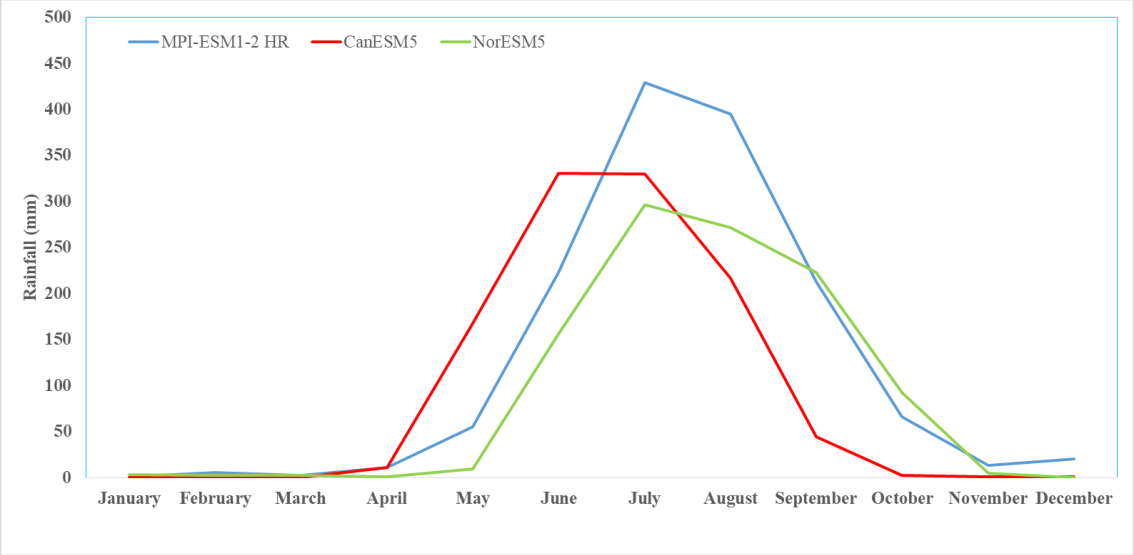


**Fig 6: Projected rainfall under SSP3 7.0 scenario for end century (2080-2100) after BC**

Projected rainfall under the SSP5 8.5 scenario for the end of the century (2080-2100) after bias correction was presented in Figure 7. The MPI-ESM1-2 HR model generally predicted higher rainfall values compared to the CanESM5 and NorESM5 models across most months. For instance, in June and July, the MPI-ESM1-2 HR model estimated much higher rainfall values than the other two models. There was noticeable variability in rainfall values from month to month within each model. For example, in the CanESM5 model, there was a significant increase in rainfall from March to April, followed by a decrease in May. This pattern was also observed, albeit to a lesser extent, in the other two models. Despite some differences, there were also instances where all three models showed similar rainfall values. For example, in December, all models predicted relatively low rainfall amounts.

UNDER PEER REVIEW

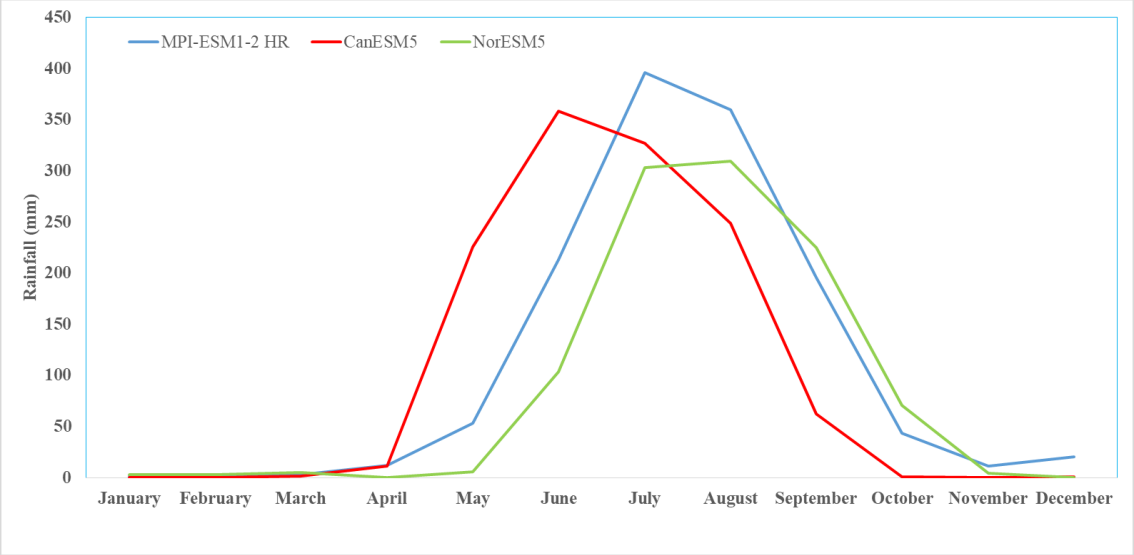


**Fig 7: Projected rainfall under SSP5 8.5 scenario for end century (2080-2100) after BC**

Monthly precipitation values varied widely across models, highlighting the complexity of precipitation simulations in climate models and the need for careful interpretation when using model outputs for various applications. These differences in GCM output might have arisen from differences in model parameters, resolution, or other factors influencing rainfall simulation (Samadi *et al.,* 2010; Huang 2014).The results table provided a general overview of rainfall values, it was important to note that regional variations in rainfall might not have been adequately captured by these global climate models. Local factors such as topography, land use, and proximity to water bodies could significantly influence rainfall patterns, which might not have been fully represented in these global models.

## 4. Conclusion

The comparison of CDFs before and after bias correction indicated significant improvements in aligning model-simulated precipitation with observational data. Systematic biases in the raw model output were corrected through bias correction

UNDER PEER REVIEW

techniques, resulting in a closer match between model predictions and observed precipitation patterns. The bias-corrected future precipitation revealed substantial variability in monthly precipitation values across different climate models and scenarios. Each model exhibited distinct seasonal patterns in precipitation, with notable differences in peak values and timing across months. Inter-model discrepancies highlighted the complexities of precipitation simulations and the need for careful interpretation of model outputs. The findings underscored the importance of continued research and development efforts to improve the accuracy and reliability of climate model simulations. Further investigations into the sources of variability and uncertainty in GCM outputs, as well as advancements in bias correction techniques, could enhance the utility of these models for climate change impact assessments, resource management, and decision-making processes.

## 5. REFERENCES


Anil, S., & Raj, P. A. (2024). An exhaustive investigation of changes in projected extreme precipitation indices and streamflow using CMIP6 climate models: A case study. *Journal of Earth System Science*, *133*(2), 1-22.

Huang, Y. 2014. Comparison of general circulation model outputs and ensemble assessment of climate change using a Bayesian approach. *Global and Planetary Change*, 122: 362-370.

IPCC, 2021. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*[Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield,


UNDER PEER REVIEW

O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Khavse, R., Deshmukh, R., Manikandan, N., Chaudhary, J.L. and Kaushik, D., 2015. Statistical analysis of temperature and rainfall trend in Raipur district of Chhattisgarh. *Current World Environment*, 10(1).305-312.

Mishra, V., Bhatia, U., and Tiwari, A. D. 2020. Bias-corrected climate projections for South Asia from coupled model intercomparison project-6. *Scientific data*, 7(1): 338.

Mishra, V., Kumar, D., Ganguly, A. R., Sanjay, J., Mujumdar, M., Krishnan, R., & Shah, R. D. (2014). Reliability of regional and global climate models to simulate precipitation extremes over India. *Journal of Geophysical Research: Atmospheres*, *119*(15), 9301-9323.

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornblueh, L., Li, H., Modali, K., Notz, D., Pohlmann, H., Roeckner, E., Stemmler, I., Tian, F., & Marotzke. J. 2018. A higher-resolution version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). *Journal of Advances in Modeling Earth Systems*, 10(7): 1383–1413.

Piani, C., Weedon, G. P., Best, M., Gomes, S. M., Viterbo, P., Hagemann, S., & Haerter, J. O. 2010. Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models. *Journal of hydrology*, 395(3-4):199-215.

UNDER PEER REVIEW

Saha, U., & Sateesh, M. (2022). Rainfall extremes on the rise: Observations during 1951–2020 and bias-corrected CMIP6 projections for near-and late 21st century over Indian landmass. *Journal of Hydrology*, *608*, 127682.

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M. 2020. Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geoscientific Model Development*, 13(12): 6165–6200.

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., & Winter, B. (2019). The Canadian Earth System Model version 5 (CanESM5.0.3). Geoscientific Model Development, 12(11), 4823–4873.

Verma, S., Kumar, K., Verma, M. K., Prasad, A. D., Mehta, D., & Rathnayake, U. (2023). Comparative analysis of CMIP5 and CMIP6 in conjunction with the hydrological processes of reservoir catchment, Chhattisgarh, India. *Journal of Hydrology: Regional Studies*, *50*, 101533.

Wang, Q., Sun, Y., Guan, Q., Du, Q., Zhang, Z., Zhang, J., & Zhang, E. (2024). Exploring future trends of precipitation and runoff in arid regions under

UNDER PEER REVIEW

different scenarios based on a bias-corrected CMIP6 model. *Journal of Hydrology*, *630*, 130666.

UNDER PEER REVIEW