Thank you very much for your interest in our paper. As you noted, this study aims to make a significant contribution to determining the reliability of GCMs in climate prediction and impact assessment. Previous bias correction studies have primarily focused on selecting the appropriate method based on performance alone; however, this approach has limitations due to the uncertainties arising from various factors. Our research proposes a comprehensive index that integrates uncertainty quantification with performance evaluation when selecting bias correction methods. This approach provides an opportunity to further enhance the reliability of climate models. Specifically, we expect that the results of this study will be effectively used in the process of selecting GCMs and evaluating the various methods applied in the models, ultimately aiding in the choice of the most suitable approach. Once again, we sincerely appreciate your interest and valuable feedback. Now, I would like to address your comments.

**Comment 1**

L50: What do the authors mean by “which has reduced difference between GCM simulations and observed precipitation? Do the authors mean uncertainty/error in GCMs relative to the observed?

**Answer**

We thank the reviewer for raising this important point. In the original sentence, “reduced the difference between GCM simulations and observed precipitation,” we intended to convey that the bias correction methods decrease the systematic errors namely, the bias present in raw GCM outputs when compared to observed (or reference) precipitation data. Raw GCM simulations often exhibit significant discrepancies from observations due to inherent model limitations and incomplete representations of key physical processes. Therefore, many studies apply various bias correction techniques (e.g., quantile mapping methods) to adjust the raw simulations so that they align more closely with observed precipitation, thereby reducing error metrics such as RMSE or MAE. We have clarified this point in the revised manuscript to emphasize that “reduced difference” refers to the decrease in systematic bias/error rather than a reduction in model uncertainty. In addition, we have revised the sentence to clearly express its meaning as follows:

Many studies have developed appropriate bias correction methods based on various theories, which have reduced the difference between raw GCM simulations and observed precipitation (Abdelmoaty and Papalexiou, 2023; Shanmugam et al., 2024; Rahimi et al., 2021).

**Comment 2**

L57: change “a recent study” to “recent studies” and remove “an”

**Answer**

Thank you for your comment. We acknowledge that there were some errors in writing the sentence. Taking your comments into account, we have revised the sentence as follows:

To address these limitations, recent studies proposed an improved QM approach to reflect future non-stationary precipitation across all quantiles of historical precipitation (Rajulapati and Papalexiou, 2023; Cannon et al., 2015; Cannon, 2018; Song et al., 2022b).

**Comment 3**

L70: There is an abrupt deflection from the bias correction discussion to the MCDA in this paragraph disrupting the flow from the previous paragraph. Authors emphasized the need for selecting appropriate bias correction methods in L67 – 69 and abruptly jumped to GCMs performance evaluation and then again to bias correction. I suggest moving paragraph three to paragraph four and paragraph four to three. Authors should also add a brief importance of TOPSIS to the MCDA method and why it was chosen out of the many other MCDA methods. Nothing about the BMA method was also mentioned in any part of the introduction except at the last paragraph that it was used. While this information is always detailed in the methodology section, briefly introducing while they are used in the introduction section gives comprehensiveness to this section.

**Answer**

We have carefully considered the reviewer’s suggestion regarding the change in paragraph order. However, in paragraph 3 we describe the reasons behind the implementation of existing bias correction methods for GCMs, their limitations, and the various approaches that have been proposed to overcome these shortcomings. In particular, the last sentence emphasizes that the improved QM techniques do not provide uniform results across all grids and points, thereby necessitating a robust evaluation framework for selecting the appropriate QM method. The following paragraph then naturally introduces the MCDA framework. Thus, we believe that the current structure, without any change in paragraph order, already maintains a sufficient logical flow and connectivity.

Nonetheless, to strengthen paragraphs 3 and 4, we have added the following passages regarding TOPSIS and BMA:

Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) is effectively utilized in our study's MCDA framework by integrating multiple evaluation metrics and calculating the distance between each alternative and the ideal solution, thereby enabling clear and intuitive prioritization decisions.

Furthermore, Bayesian Model Averaging (BMA) plays a crucial role in quantifying the predictive uncertainty of multiple climate models and enhancing the reliability of the final predictions, which is why it has been employed as an indispensable tool in our integrated evaluation.

**Comment 4**

L79: This line in part shows the justification for this research. I suggest authors put all justification for the study in the second to the last or last paragraph of this section.

**Answer**

Thank you for your insightful comment. We have incorporated your suggestions by adding the following sentence to the introduction:

In light of the challenges outlined above, including discrepancies among bias correction methods, regional variability in precipitation distributions, and significant uncertainties in GCM outputs, there is a clear need for an integrated framework that evaluates the performance of various QM methods and quantifies their associated uncertainties.

**Comment 5**

L96: change “simulation” to “simulated” or change “simulation precipitation” to “precipitation simulation”

**Answer**

Thank you for your comment. We have revised the text as follows:

However, accurately identifying biases in precipitation simulation remains challenging due to the lack of comprehensive equations reflecting Earth's physical processes.

**Comment 6**

L121: Considering that there are many GCMs, why did the authors consider just 11 in their study?

**Answer**

The reason we used only 11 CMIP6 GCMs in this study is that applying QM methods requires extensive computational time. Generating bias-corrected results with three QM methods across 11 CMIP6 GCMs took more than six months. Nonetheless, we determined that using fewer than 10 GCMs would not adequately justify the study, which is why we employed 11 CMIP6 GCMs. We appreciate your understanding on this matter.

**Comment 7**

L122: mentioning temperature here is not necessary since the study doesn’t consider temperature.

**Answer**

Thank you for your comment. We have removed the sentence from the main text in response to your suggestion.

**Comment 8**

L129: In table 1. I think the climate variables and the variant label columns are not necessary since only single ones were used. Mentioning it in the text is sufficient. Authors should consider including the full name of the institutions or any other information.

**Answer**

Thank you for your comment. In response, we have revised Table 1 and the text as follows:

Table 1. Information of CMIP6 GCMs in this study

|  |  |  |
| --- | --- | --- |
| **Institution** | **Models** | **Resolution** |
| Commonwealth scientific and industrial research organization/ Australia | ACCESS-CM2 | 1.2° × 1.8° |
| ACCESS-ESM1-5 | 1.2° × 1.8° |
| Beijing Climate Center/China | BCC-CSM2-MR | 1.1° × 1.1° |
| Canadian Centre for Climate Modeling and Analysis/ Canada | CanESM5 | 2.8° × 2.8° |
| National Center for Atmospheric Research | CESM2-WACCM | 0.9° × 1.3° |
| Euro-Mediterranean Center on Climate Change coupled climate model/ Italy | CMCC-CM2-SR5 | ~ 0.9° |
| CMCC-ESM2 | 0.9° × 1.25° |
| EC-Earth Climate Model Consortium/ EC-EARTH consortium | EC-Earth3-Veg-LR | 1.0° × 1.0° |
| National Oceanic and Atmospheric Administration/ United States | GFDL-ESM4 | 1.4° × 1.4° |
| Institute for Numerical Mathematics/ Russia | INM-CM4-8 | ~ 0.9° |
| Institute Pierre Simon Laplace/ France | IPSL-CM6A-LR | 1.1° × 1.1° |

**Comment 9**

L136: change “availability” to “purpose”

**Answer**

Thank you for your comment. We attempted to locate the term "availability" in L136, but it was not present. Please review it once again.

**Comment 10**

L146 – 148: were the historical GCMs corrected before comparison to the observed? If so, why was this done since many studies usually compare the raw GCMs first with the reference before bias correcting and then comparing the bias corrected with the reference. If otherwise, authors should rephrase the sentence.

**Answer**

Thank you for your comment. In this study, we compared the raw GCMs and the bias-corrected GCMs with the reference data. In particular, Figure 1 compares the bias-corrected GCMs and raw GCMs using Taylor diagrams.

텍스트, 도표, 라인, 지도이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

Figure 1. Comparison of raw and corrected daily precipitation on six continents using Taylor diagrams (x-axis: standard deviation; y-axis: the correlation coefficient)

**Comment 11**

L148: why was the “frequency-adaptation technique” further needed?.

**Answer**

Frequency-adaptation is used not only to adjust the magnitude of a climate variable but also the frequency of its events. Climate models, particularly for variables such as precipitation, often exhibit biases in both the intensity of precipitation and the number of precipitation days (i.e., the frequency) compared to observations. This technique adjusts the frequency of precipitation in the model output so that it more closely matches the observational data, thereby enhancing the realism of the overall distribution and the representation of extreme events. In addition, frequency adaptation is employed to address the issue that regional climate models often simulate a lower frequency of dry days than observed, which, if left uncorrected, would lead to a systematic wet bias in the bias-corrected precipitation. This technique randomly interpolates a portion of the excess dry-day cases to preserve the observed distribution of dry days and improve the overall accuracy of Quantile Mapping. The effectiveness of this approach has been demonstrated in Themeßl et al. (2011). For example, when correcting precipitation, if a model predicts too many light rain events or too few heavy rain events compared to reality, the frequency-adaptation technique adjusts these frequency differences. As a result, both the magnitude and the frequency of precipitation events are accurately represented, leading to more reliable climate predictions.

**Comment 12**

L187: The authors mentioned 10 metrics and again later mentioned 7 metrics. Table show 10 were used. Authors should correct this in text.

**Answer**

Thank you for your comment. In response to your suggestion, we have revised the sentence as follows:

Ten evaluation metrics used in this study are as follows: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Coefficient of Determination (), Percent bias (Pbias), Nash-Sutcliffe Efficiency (NSE), Kling-Gupta efficiency (KGE), Median Absolute Error (MdAE), Mean Squared Logarithmic Error (MSLE), Explained Variance Score (EVS), and Jenson-Shannon divergence (JSD). **Comment 13**

L187: The authors mentioned 10 metrics and again later mentioned 7 metrics. Table show 10 were used. Authors should correct this in text.

**Answer**

Thank you for your comment. In response to your suggestion, we have revised the sentence as follows:

Ten evaluation metrics used in this study are as follows: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Coefficient of Determination (), Percent bias (Pbias), Nash-Sutcliffe Efficiency (NSE), Kling-Gupta efficiency (KGE), Median Absolute Error (MdAE), Mean Squared Logarithmic Error (MSLE), Explained Variance Score (EVS), and Jenson-Shannon divergence (JSD). **Comment 14**

L206: GEV was not mentioned at all in the introduction section. A brief introduction about it and its usage should be introduced in the introduction section.

**Answer**

Thank you for your comment. In response, we have added the following sentence:

Furthermore, they compared the extreme precipitation of GCMs using the GEV distribution, which allows for more effective estimation of extreme precipitation, and demonstrated that the performance in estimating extreme precipitation varies according to different bias correction methods.

**Comment 15**

L291: The first sentence: “This study applied…………..” is a repetition. It should be removed..

**Answer**

Thank you for your comment. In response to your suggestion, we have revised the sentence as follows:

A Taylor diagram was used to compare the bias-corrected and raw GCM precipitation with the observed data, and Figure 1 presents the results of applying the three QM methods to 11 CMIP6 GCMs.

**Comment 16**

L292: Authors mentioned that figure 1 show the before and after bias correction of the precipitation. There is no indication of the before and after correction in the figure. The figure seems to be showing either of them. Authors should clarify. Figure also shows there seems to be no any changes in the performances of some models based on the three BC methods in some regions. For example, in South America where there are no changes based on the BC methods for models between 45 - 90 o . What could be the reason for these?

**Answer**

Thank you for your comment. In Figure 1, the triangles represent the raw GCM precipitation, while the star symbols represent the bias-corrected GCM precipitation. Furthermore, the black star in Figure 1 denotes the reference data, and most of the bias-corrected GCM precipitation values closely match the reference data across all continents. The models located between 45° and 90° in South America, as you mentioned, represent the raw GCM precipitation. Therefore, all the Taylor diagrams in Figure 1 include both raw and bias-corrected GCM precipitation.

**Comment 17**

L320: This study…….This is not necessary. I t can be removed.

**Answer**

Thank you for your comment. In response to your suggestion, we have revised the sentence as follows:

The spatial patterns of the evaluation metrics computed from the bias-corrected daily precipitation data of GCMs in South America are presented as shown in Figures 2 and S1.

**Comment 18**

L339: did authors meant North America here?

**Answer**

Thank you for your comment. We made some errors while drafting the text, and in response to your suggestion, we have revised the sentence as follows:

Figure 4 shows spatial patterns of evaluation metrics for bias‐corrected daily precipitation in North America.

We utilized supplementary materials to present the results of all evaluation metrics for each continent. Accordingly, the main text explains the outcomes for all 10 evaluation metrics. To further demonstrate that the diversity and computations described in the main text were indeed carried out, we provided various indices. Moreover, the influence of factors such as geographic features (e.g., deserts) on the performance of bias correction methods across regions was primarily discussed in the discussion section with respect to the observations.

**Comment 19**

L339: did authors meant North America here?

**Answer**

Thank you for your comment. In response to your suggestion, we have revised the caption for Figure 7 as follows:

과일, 스크린샷, 레드, 모자이크이(가) 표시된 사진

AI 생성 콘텐츠는 정확하지 않을 수 있습니다.

Figure 8. Performance comparison of DQM, EQM, and QDM for the validation period (1997-2014) using evaluation metrics for daily precipitation in Oceania.

**Comment 20**

L439: removed “biased”

**Answer**

Thank you for your comment. We have revised the sentence as follows:

Figure 10 shows the JSD of GEV fitted daily precipitation for DQM, EQM, and QDM on each continent.

**Comment 21**

L441 – 443: how does this little difference observed for the GEV method compare to the other method in terms of the differences? How does the differences observed in the other methods applied compare to the GEV method?

**Answer**

Thank you for your comment. Figure 9 displays the GEV distributions of the overall precipitation, with distribution differences compared to observed values using the JSD. Subsequently, Figure 10 compares the extreme precipitation values those above the 95th percentile obtained from the three QM methods.

**Comment 22**

L441 – 443: how does this little difference observed for the GEV method compare to the other method in terms of the differences? How does the differences observed in the other methods applied compare to the GEV method?

**Answer**

Thank you for your comment. We have revised the sentence as follows:

In Oceania, high weights were assigned to JSD, KGE, RMSE, and MAE, suggesting that these metrics are critical for evaluating model performance.

**Comment 23**

L42 – 43: This study developed………There is no need for this sentence.

**Answer**

Thank you for your comment. In response, we have revised the sentence as follows:

A daily precipitation ensemble for the historical period was generated using BMA on 11 CMIP6 GCMs, and the standard deviation of daily precipitation by continent is presented as shown in Figure 14.

**Comment 24**

L575: What factors could contribute to the lower index observed in these regions?

**Answer**

These results quantify the uncertainty arising from predicting precipitation in an ensemble form using BMA on GCMs. This uncertainty is primarily attributed to differences in model structure and physical processes, model weight variance, and uncertainties in the input data. Additionally, variations in the spatial and temporal resolutions of the models and the handling of extreme events can further contribute to the overall uncertainty.

**Comment 25**

L615: I think “best” rather than “proper” is more suitable in this context.

**Answer**

Thank you for your comment. In response, we revised the sentence as follows:

QDM was selected as the best bias correction method in western North America, southern and eastern Africa, and northern Europe.

**Comment 26**

L653: All three methods showed strong overall performance…….This statement seems contradictory of the finding of the better performance of the EQM than the QDM and the QDM than the DQM. Authors should clarify what is meant here.

**Answer**

Thank you for your comment. In order to avoid confusion among readers, we revised the sentence as follows:

All three methods, as evidenced by the Taylor diagram, demonstrated overall stronger performance than the raw GCM and consistently produced good results across various regions. Nonetheless, the performance of the bias-corrected GCMs clearly differs. This highlights the need to use multiple performance metrics to fully understand the strengths and weaknesses of the three QM methods, as relying on a single analysis or macroscopic perspective can overlook important details.

**Comment 27**

L766: In conclusion…………I suggest authors merge this part with the conclusion section since these are some if the conclusions from the study.

**Answer**

Thank you for your comment. In response, we have moved the paragraph to Section 5 (Conclusion) and revised it as follows:

In conclusion, EQM has emerged as the preferred method due to its balanced performance, but this study emphasizes the importance of regional assessment and careful consideration of uncertainty when selecting a QM method. Furthermore, EQM is the most balanced method regarding performance and uncertainty and will likely be preferred in future climate modeling studies. However, there may be more suitable QM methods depending on the region, and a comprehensive evaluation with various weights is needed. Therefore, when establishing climate change response strategies or policy decisions, it is essential to take a multifaceted approach that considers uncertainty together rather than relying on a single indicator or performance alone. It will enable more reliable predictions and better decision-making. Future research should integrate greenhouse gas scenarios to improve the accuracy of climate predictions and provide a more comprehensive understanding of future climate risks. Furthermore, more bias correction methods should be used to extend the robustness of CI.

**Comment 28**

Were there any limitations of this study. This should be discussed in a paragraph in this section as it can help guide other researchers interested in similar work in the future.

**Answer**

Thank you for your comment. This study has acknowledged its limitations as follows:

Future research should integrate greenhouse gas scenarios to improve the accuracy of climate predictions and provide a more comprehensive understanding of future climate risks.

**Comment 29**

L794: The previous bullets already talked about the conclusions. Authors can remove “In conclusion” from this line or replace it with “generally”

**Answer**

Thank you for your comment. In response to your comment, we have revised the paragraph as follows:

In conclusion, EQM has emerged as the preferred method due to its balanced performance, but this study emphasizes the importance of regional assessment and careful consideration of uncertainty when selecting a QM method. Furthermore, EQM is the most balanced method regarding performance and uncertainty and will likely be preferred in future climate modeling studies. However, there may be more suitable QM methods depending on the region, and a comprehensive evaluation with various weights is needed. Therefore, when establishing climate change response strategies or policy decisions, it is essential to take a multifaceted approach that considers uncertainty together rather than relying on a single indicator or performance alone. It will enable more reliable predictions and better decision-making. Future research should integrate greenhouse gas scenarios to improve the accuracy of climate predictions and provide a more comprehensive understanding of future climate risks. Furthermore, more bias correction methods should be used to extend the robustness of CI.

.