Paper on methods for assessment of models: response to reviewers

Nguyen et al.

August 19, 2024

Response to Reviewer 1

We are grateful to the reviewers for their insights and believe their comments have substantially improved our manuscript. We address all their comments below, point-by-point, in blue. We trust that our changes to the manuscript will satisfy the reviewers and the Editor.

Replies to general comments

This paper describes a standardized benchmarking framework for selecting CMIP6 GCMs for CORDEX downscaling over Southeast Asia. The topic is important because Southeast Asia faces a high risk of flooding due to climate change, yet fewer models or frameworks are available for characterizing regional changes in precipitation compared to other regions such as Europe and the US. The authors did a great job highlighting the differences between their approach and those in the literature, which mainly rank GCMs according to specific evaluation matrices. The logic of this paper is very clear, and it is very well written. I only have a few minor points for the authors to consider.

Thank you for your positive feedback on the manuscript. We appreciate your recognition of the importance of improving regional climate modelling for Southeast Asia, given the area's vulnerability to flooding due to climate change. The comments have made substantial improvements to this manuscript. Please see the point-by-point responses below regarding your concerns.

Replies to specific comments

Technical corrections

1. L44: GCMs'

Thank you. Revised (L44)

2. L51, L55: should be 'WCRP'?

Revised (L51,55)

3. Section 1: an overview of the paper structure should be added to the end of this section, so the readers know what they expect in each section.

Thanks for suggesting this improvement. We have added the structure-related paragraph below (L116-120)

"The structure of the paper is as follows: Section 2 introduces the data and the benchmarking framework employed in this study. The results are presented in three subsections: Section 3.1 focuses on model assessment using the benchmarking framework; Section 3.2 examines the spread of future climate change among models; and Section 3.3 assesses model dependence through cluster analysis. Finally, we conclude with a discussion of our results in Section 4 and a summary of the main conclusions in Section 5."

4. L120-121: do you mean 'We *do not* consider models which have a horizontal grid spacing greater than…'. Or by 'greater' do you mean finer resolution than 2 degrees?

Thanks for pointing to the need for enhanced explanation here. Yes, we are referring to models with a finer resolution than 2 degrees. We have modified the text accordingly (L125-126).

"We consider only models which have a horizontal grid spacing finer than 2°× 2° which are likely to be more suitable for dynamical downscaling."

5. L123-124: incomplete sentence.

Revised the text accordingly (L129).

"At the time of this analysis, the first member of some models (e.g., CNRM-family models, UKESM1-0-LL and HadGEM3-GC31-MM) was not available so another member was utilized."

6. L174: you may want to remove theta from the first half of the sentence and explain it as wind direction(?)

Good point. We have revised the manuscript accordingly (L199-200).

"Where u_i refers to the simulated wind speed at the grid i, θ_i and $\theta_{i,ref}$ are the wind direction at grid i in the simulated and reference data respectively".

7. L198: did you define DMI somewhere above?

Thank you for raising this point. We have now moved the paragraph that defines DMI (L209-216).

"To track ENSO variability, the Niño3.4 index (5°S-5°N and 160°E-120°W) (Trenberth and Hoar, 1997; Shukla et al., 2011) derived for the 1951-2014 period as area-mean monthly SST anomalies with respect to a 1961-1990 climatology is used. For IOD, we use the Dipole Mode Index [DMI; (Saji et al., 1999; Meyers et al., 2007)] DMI measures differences in monthly SST anomalies between the west equatorial Indian Ocean (50-70° E, 10°S-10°N) and those in the east (90-110°S, 10°S -0°N).

We use a 5-monthly average Niño3.4 and IOD index to remove seasonal cycles. The resulting month time series are detrended using a fourth-order polynomial fit to remove the possible influence of a long-term trend and to better preserve high amplitude (<10 years) variability (Braganza et al., 2003)."

8. L218: what do you mean by 'significant sign'?

We agree that this could be confusing. We are simply referring to the significant correlations that would be obtained from the application of the metrics for assessing agreement in teleconnections. Following the comment, we have amended the text accordingly (L242-243).

"For high-level qualification, we employ spatial correlation and simplified metrics to assess whether there are significant correlations in teleconnections, as recommended by Liu et al. (2024)."

9. L320-322: not sure if I follow the definition or description of the benchmarking threshold. Do you find the six wettest and driest modelled months and require the four wettest and driest months from observations to be within those six modelled months? Then how is the threshold determined?

We appreciate the spirit of this comment. We have defined the benchmarking threshold such that the four driest and wettest observed months must fall within the six driest and wettest months simulated by the models. A model meets the benchmark if the four driest observed months rank between 1 and 6 in the model's simulation. We have addressed your concern and revised the text for a clearer clarification(L357-361).

"According to the benchmarking threshold definitions, all models meet the benchmark for simulating the four wettest observed months. However, six models do not pass the benchmark for simulating the four driest observed months, as highlighted in orange in Fig. 4. Specifically, one of the four driest months according to the APHRODITE dataset (December through March) is ranked as the sixth wettest month (ranked 7th in Fig. 4) by these models."

10. L340-342: did you show the observational trend somewhere or can you cite references for this claim?

The observed trends (e.g., in APHRODITE) are presented in the top panel of Figures 5-6. We have included this information in the text for clearer illustration (L367-368).

"There is a significant decreasing trend in observed total precipitation during the wet season (Figure 5 – top panel) while the dry season has a significant increasing trend (Figure 6 - top panel)."