

Earth Syst. Dynam. Discuss., referee comment RC2  
<https://doi.org/10.5194/esd-2022-31-RC2>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on esd-2022-31

Anonymous Referee #2

---

Referee comment on "Performance based sub-selection of CMIP6 models for impact assessments in Europe" by Tamzin Emily Palmer et al., Earth Syst. Dynam. Discuss., <https://doi.org/10.5194/esd-2022-31-RC2>, 2022

---

Comments on the manuscript entitled "Performance based sub-selection of CMIP6 models for impact assessments in Europe" by Palmer et al. submitted to Earth System Dynamics

Recommendation: Major revision

The authors assessed CMIP6 models in terms of their performance and diversity in simulating several variables, e.g. temperature, precipitation, and circulation, over Europe. Based on the assessment, they created sub-sets of CMIP6 models, which can be used for downscaling or impacts assessments. The approach can also be applied to other regions of the world. The topic is important and falls within the scope of the journal. The manuscript is generally well written. My major concern includes: the assessment of CMIP6 models did not well consider the link between the model's ability to simulate historical climate and future climate change. The assessments are overly dependent on subjective assessment criteria. Detailed comments are laid out below.

Major comments:

- No link was established in terms of the model's ability to simulate the historical climate and the projected changes. Thus, the models that can better reproduce historical climate may not necessarily generate a more reliable projection of future climate. After excluding the least realistic models, the filtered CMIP6 models show higher sensitivity. Is the result reasonable?
- Quantitative measures are preferred for model evaluation. Visual inspection hinders the

inter-comparison of various studies to a certain degree as different people may have different judgments on "satisfactory", "unsatisfactory", and "Inadequate". I'm wondering to what extent the results will be different if the authors use objective assessment criteria only.

- How was the RMSE of the zonal mean track calculated? It seems that the authors calculated the zonal mean track and obtained a time series. The RMSE is calculated using the time series derived from models and observation. Please note it makes no sense by comparing the year-to-year variation of the unforced internal variability derived from AOGCMs against the observed one. In this case, the RMSE is largely determined by the phase discrepancy between simulation and observation. Please also check the use of RMSE elsewhere.

Other comments:

Section 2: It is not clear to me how the CMIP6 models are grouped into classifications. Please clarify how the quantitative and qualitative measures were used and what is the threshold of quantitative measures to group the models. I suggest the authors introduce the "criteria" first and explain the classification definitions based on the criteria.

L64: "processed based" -> "process-based"

L70: How the regional processes are linked to future changes?

L137: "process base" -> "process-based", " does not use and regional or global warming trends"->"does not use regional or global warming trends". Please carefully read throughout the manuscript and correct the typos or grammar mistakes. E.g. L202 ...

L217: What is the temporal resolution of the dataset, monthly mean or daily mean? Which CMIP6 experiment was used for the baseline period? Both the baseline and future periods are only 20 years. The climatological means averaged over 20 years may still contain internal climate variability, e.g. AMO or PDO, which may affect the evaluation and selection of the models to a certain extent.

L225: Please clarify what reanalysis and observational data were used in this study.

L254-255: How the circulation pattern is measured? Is the RMSE calculated using two

wind speed fields or an RMS vector error between two vector fields? If the RMSE is calculated with wind speed, it does not reflect the errors in wind direction. Instead, the RMSE for vector field can reflect both errors in wind speed and wind direction. Therefore, I suggest the authors use the latter one. Similarly, the difference in wind speed illustrated in Fig. 1 can only describe the errors in wind speed. The same wind speed does not mean the same wind direction. The authors may consider using a vector difference between the model and ERA5. The magnitude of vector difference takes both differences in wind speed and wind direction into account.

Xu et al, 2016: A diagram for evaluating multiple aspects of model performance in simulating vector fields. *Geosci. Model Dev.*, 9, 4365–4380

L270: Please explain how the "track density" is defined. Please use the degree symbol "°" to represent latitude and longitude here and elsewhere.

L321: "depending to on" -> "depending on"

L334: "with with" -> "with"

L343-345: How about the range of other quantities, e.g. precipitation and storm track density?

L362: Please clarify what numerical score was given for each group of models.

L644: "35°N-75°" -> "35°N-75°N"

Fig. S4: What does the "???" refer to in the figure caption?