



EGUsphere, referee comment RC1
<https://doi.org/10.5194/egusphere-2022-98-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on egusphere-2022-98

Anonymous Referee #1

Referee comment on "Constraining low-frequency variability in climate projections to predict climate on decadal to multi-decadal timescales – a poor man's initialized prediction system" by Rashed Mahmood et al., EGU sphere,
<https://doi.org/10.5194/egusphere-2022-98-RC1>, 2022

In my opinion this is a very well written and clear paper which presents an interesting new way of combining a large multi-model dataset in order to constrain future projections. As such I have very few comments to make and I believe that this paper should be published with only a few corrections.

The results are impressive and warrant publication. However the one area I would like to see more explanation on is the cause of the skill found. In the summary (line 300) you claim that "these results indicate that there is significant multi-decadal predictability from internal variability". I am not convinced though that as they stand your results can justify this claim (although I'm not suggesting this is necessarily incorrect). By selecting the best 30 (or similar) simulations from the CMIP6 archive based on the 9 most recent years, could it not be possible that what you are doing is weighting your results to certain models (which is in itself a rich area of literature). This could be those models with the most realistic response to forcing, especially when outside the anomaly period at the start or end of the record. Or it could potentially be those with the most realistic modes of variability. Both of which, I would have thought could theoretically give an increase in skill over the next decades. This is especially pertinent, given that you state that the method gives a better constraint than a single model ensemble (although as you say this is expected given the increased number of simulations to sample from). I think that this point needs to be discussed, and preferably investigated. One simple test you could do, without the need for further analysis, would be to check if you are selecting simulations preferentially from one climate model or if the 30 members are selected from the full model range. In addition it would also be interesting to see if the skill varies through time.

I am also somewhat surprised that you have not shown results for future projections, since that would seem to be the logical direction for this type of analysis. Although, perhaps this is being left for future work (which would be perfectly justifiable).

More minor points:

L 84. I think that adding the acronym for the Extended Reconstructed Sea Surface Temperature Version 5 dataset would be helpful.

L88. Why was a reference period of 1981-2010 chosen, and are the results sensitive to this?

L117. Was the forced signal that you removed the multi-model mean?

L151. Why global warming and not external forcing in general? I would have thought that anthropogenic aerosols and volcanic eruptions might also have an impact.

L261. It would be useful to cite papers which have suggested that forcing (particularly natural) could contribute to the slow down, (see e.g. box 3.1 IPCC AR6 WG1)