Shruti Nath
Climate Analytics
Ritterstrasse 3, 10969, Berlin, Germany
Phone: +49 (0) 30 259 22 95 44
E-mail: shruti.nath@climateanalytics.org

Earth System Dynamics editorial board Berlin,
18 March 2022

**MESMER-M: an Earth System Model emulator for spatially resolved monthly temperature**

Dear Dr. Messori,

Please find enclosed the manuscript after implementing the minor revisions from the second round of revisions as well as a version with tracked changes. Point by point answers to the reviews are directly attached with this letter. Most of the changes are small textual edits/clarifications and corrections of grammatical errors and captions (arising from both the reviews and a round of proof reading).

We think that the minor reviews have been extremely helpful in strengthening the quality of the paper and are grateful to all those involved in the review process.

Sincerely,

Shruti Nath

(on behalf of all the co-authors)

**Reviewer 1**

Thank you to the authors for your patience in my somewhat delayed review and for the detailed responses to the questions I raised in my initial review. I believe that the authors have done a reasonable job of responding to the questions I raised and I would be happy to see this article published in more or less its current form. The following are a few comments that might be worth considering, but I believe the authors should have the final say about whether or how best to address items 1-2 below.

We once again thank the reviewer for their input, it has vastly enriched the technical quality of the paper. We hope to have addressed the remaining points listed below sufficiently.

1. While I appreciate that you took my suggestion to also pay attention to the spectrum at high frequencies, I do not understand the choice to switch to looking at only the 50 highest frequencies rather than the 50 lowest frequencies, instead of simply looking at the agreement of the spectra over all frequencies (as was my intended suggestion). I do not think this is such a big deal, but the choice here seems somewhat arbitrary and the low frequencies do matter too.

We recognise that checking for the whole range of frequencies is important, but after further consideration we decided to stay focussed on the higher frequencies since this is also what we are most interested in capturing with the AR(1) process.

2. Choosing the order of the seasonal model to minimize the AIC using a likelihood that assumes temporal independence will, I believe, tend to result in including too many terms in the seasonal model. The reason is essentially that the likelihood assuming independence thinks that the effective sample size is larger than it actually is. Again I don't think this is such a big deal -- and I think it's better to include a few too many terms in the seasonal model than too few -- but just as a point of good statistical practice it might be worth briefly acknowledging this fact.

A line (L110) acknowledging this shortcoming of using the AIC was added

3. Overall I do think that the article could use a round of proofreading. I'm sorry that I don't have the time to point out specific grammatical and spelling issues, but there are a number of them so I'd recommend attending to this.

Proof reading was done and we hope all grammatical errors were addressed.

**Reviewer 2**

I understand and accept the authors' responses to my previous concerns. The draft appears to need minor revisions for clarification. The followings are further comments to be considered including those for editorial suggestion.

We thank the reviewer for their further comments, they definitely improve the quality of the text and its readability. We have implemented most of the suggestions as elaborated below.

L8. It would be informative to notify that the MESMER is a statistical emulator, which is distinct from a physical climate emulator like MAGICC. This is also the case for Introduction.

For ease of readability we do not elaborate on this in the abstract but make it quite clear in the Introduction (L32)

L13-14. Briefly describe how the performance is verified.

We have added this to the line

L93. 'other' should be capitalized.

This has been done

L103. I notice some minor differences between this equation and the one displayed in Figure 1, which should be harmonized if appropriate.

We have corrected this

L106. beta_1 = 1 (not beta_0)

Has been corrected

L117. Throughout the text and figure captions, 'Gaussian' and 'gaussian' are mixed.

We have changed all to Gaussian

L122. This sentence should not be indented.

Indent removed

L130. Here, 'power transformed' may be a bit confusing. It would be good to explicitly mention that Equation (3) is a type of power transformation.

We have changed this to “normalised as according to equation 3.”

L151. The term 'innovations' is a bit confusing. It would be good to rephrase it if possible.

We have changed this to “noise terms”

L192. I notice several sentences including a one-digit number like here. The ESD style guide specifies to use words for cardinal numbers less than 10 for items other than units of time or measure.

We have changed this

L194. I do not think 'pinpoint' is an appropriate wording.

Changed to identify

L197. Is 'physical approach' in the subsection title a suitable wording? The GBR model uses physical variables, but the method itself is a statistical approach.

Changed to physically-informed approach

L200-201. ENSO is a variability mode associated with atmosphere-ocean interaction. It is not suitable for an example of atmospheric processes although resulting global teleconnections occur through atmospheric processes.

We have made this clearer

L245. 'shift' seems vague.

Replaced by “has a relatively low amplitude”

L253. Delete 'lambda'.

Deleted

L241. What is pointed by 'aforementioned' seems vague.

Replaced by “The source of high equatorial $\widetilde{\lambda}_{m,s}$ values “

L266. This sentence should start with 'For July' for clarification.

Sentence has been changed to “July shows similar behaviour for $\gamma_{1,m,s}$ across most ESMs, albeit with a larger spread in values
”

L278. 'by the mean response module and the residual variability module' (maybe 'and' is missing)

And was added

L288. What is described by 'followed by stabilisation' is vague.

Replaced by “non-linear”

L312. 'between the ESM training runs and emulations for a given ESM' (maybe 'and' is missing)

And added

L323-325. It needs to be explained why the value remains high even as the distance increases for the test runs.

“where spatial cross-correlations are not localised” was added

L373-375. Regarding the results from the benchmarking approach shown in Figure 13, it is hard to identify the description about 'with only MIROC6 and MPI-ESM1-LR showing larger distances for the full emulator in the Indo-Gangetic region, South America and Central-West Africa.'

Modified to “MIROC6 and MPI-ESM1-LR showing larger distances for the full emulator in the Indo-Gangetic region and Central-West Africa respectively”

L393-394. ENSO itself is not a suitable example for regional-scale changes. Regional-scale changes occur through the modulation of global-scale atmospheric circulation patterns.

Elaborated upon by adding “ modulation of atmospheric circulation patterns due to changes in ENSO”

L404-405. Given the results described in 5.4, the authors' view as mentioned 'in most cases even better than the physically-based model' is a bit questionable.

We believe that from figure E2 this statement holds (of course with the limitations identified in section 5.4)

L435-443. I suppose that warming-induced changes in major climate modes and their ESM dependencies have been relatively well documented in literature in terms of specific indexes, which may lead to improved emulation of local temperature variability associated with such modal shifts. I will not ask for further revision on this issue, but I leave this comment for clarification.

We changed this to “potentially coupled”

Figure 1. Greatly improved.
As the green panels illustrating the inverse Yeo-Johnson transformation are rather complex, a brief description would be helpful.
In the global maps, I prefer to using a common color scale for the harmonic component and the total, but no problem if the authors' choice is better for some reason.
Period is missing after 'E.g' in the figure text 'E.g Inverse Yeo-Johnson power transformation.'

Periods in the E.G. text are elaborated on and a better caption description given

Figure 2. Consider to use mathematical symbols gamma_{1,m,s} and r_m at 'the local lag-1 autocorrelation coefficients' and at 'the localisation radii' in the caption.

Has been applied

Figure 8-11. For consistency, the number of the training runs should be indicated as for the test runs.

The label 'Test runs' between the top and bottom blocks is not needed.
In the caption, '(colour)', not '(colour(.'

Typo corrected. For consistency with other figures where only test runs are shown we decided to keep only the test run numbers.

Figure B1 and C1. Describe the percentage numbers attached to each panel.

Description provided

Figure B2. Same as Figure B1, not Figure C1.

Figure C2. Same as Figure C1, not Figure D1.

Captions corrected

References. It seems that two discussion papers Beusch et al. and Brunner et al. have already been published or accepted.

Has been corrected

**Reviewer 3**

The authors did a great job to answer reviewer's comments. I am especially pleased with the added discussions to the Conclusion and Outlook section, which now states the limitations of the current approach and proposes direction for further improvements.

We once again express gratitude to the reviewers comments, they have vastly helped the text and improved the outlook.