Earth System Dynamics Discussions

Earth Syst. Dynam. Discuss., referee comment RC2
https://doi.org/10.5194/esd-2021-59-RC2, 2021

Comment on esd-2021-59

Anonymous Referee #2

Referee comment on "MESMER-M: an Earth System Model emulator for spatially resolved monthly temperatures" by Shruti Nath et al., Earth Syst. Dynam. Discuss., https://doi.org/10.5194/esd-2021-59-RC2, 2021

General comments

Climate model emulators are becoming more and more useful in assessing climate change through representing complex Earth system model (ESM) behavior and combining many different ESM simulations with other multiple evidence. The MESMER approach, being developed by the authors, is unique as an emulator for generating spatially-resolved forced and unforced climate realizations based on multi-ESM, initial-condition ensemble simulations. This paper describes the newly developed MESMER-M module for monthly downscaling, which is expected to further expand the emulator's applications.

Although the current manuscript adequately describes the structure and performance of MESMER-M, there is room for improvement as follows, which should be appropriately revised for publication.

(1) The calibration and verification results indicate marked dependency on the number of ensemble members, which may raise concerns about the robustness of the methods.

Such dependency appears in most key parameters and performance aspects: the order of the autoregressive process in the temporal variability module (L218 and L239), the stationarity of the shape parameter λ in the spatial variability module (Appendix D), the scale parameter r_m for the localization in the spatial variability module (L237, L283, Figure 7), and the bench mark test with GBR (L338). Although each of these results is explained in terms of the amount of training runs as input information, it is not much convincing. Implications for the robustness of the methods and a possible guideline of an appropriate size of the training runs should be discussed thoroughly.

(2) The modules and the calibration and verification results lack interpretation from a climatological point of view. Although the seasonality associated with snow cover is frequently mentioned, this is just one aspect. It is necessary to describe and discuss the validity of the modules from the aspect of major variability modes, such as monsoon, ENSO, and AO.

Deviations from the mean seasonal cycle may not necessarily be biophysical feedbacks, as assumed in L184-185. Internal climate variability leading to some deviations, such as jet meandering and blocking associated with the strength of the polar vortex, is hardly

regarded as a biophysical feedback. It is not much convincing that the difference between summer and winter in the mean response verification (5.1) is explained by the snow-albedo feedback only. In the regional-scale verification (5.3), although the increasing deviations in July between the ESMs and the emulator (L294-) is worth being noted, it may also need to be described based on specific natural variability, rather than regarding the tendency as abstract secondary, non-linear responses.

See also the specific comment on the conclusion and outlook below.

(3) From the standpoint of potential users of the series of MESMER modules, who are not necessarily familiar with technical details, it is recommended to devise some descriptions for better understanding.

For example, in Figure 1, adding X-Y plots illustrating a typical seasonal cycle and its variability and skewness would help understand the local variability module. Visual materials would be useful for making sense of technical concepts like the multivariate trans-Gaussian process and the Gaspari-Cohn function.

In terms of consistency between MESMER and MESMER-M, it may also be useful to verify whether the annual average of each element of MESMER-M is consistent with corresponding elements of MESMER.

Specific comments

L59. A brief explanation about limiting the scenarios to high emission SSP5-8.5 and applicability to low emission scenarios would be helpful.

L61-62. The 70-30 train-test split is not consistent with actual split shown in Table A1. It appears that the 70-30 ratio is rather exceptional, and that some models with a large number of members have 50-50.

L65-66. It should be clarified how the anomalies are calculated, i.e., whether they are deviations from the annual climatological mean or from the monthly climatological mean.

L82, 85. The use of the term "forcing" can be a bit confusing. As "other external forcings" imply an underlying primary forcing, "a certain forcing" may be better rephrased in a specific way. Changes in land cover can be anthropogenically forced or induced by climate change and variability. A more specific wording may be necessary to avoid misunderstanding.

L90. The term "monthly cycle's mean response" is a bit, confusing considering the subsequent "seasonal cycle". "Monthly mean response" may communicate its intention without ambiguity.

L96-97. It appears that the need for high-order harmonic terms is not convincing. My understanding is such that up to the second order term representing a bi-modal cycle is enough for the mean monthly response. Are the month-to-month correlations, which is the case for some natural variability modes, out of scope for the mean response?

L111. Check "time-dependent and space-dependent components" is correct wording. They are functions in terms of month, space, and year. Maybe, tempral-variability and spatial-variability components.

L112-114. Is this an appropriate explanation for adopting a autoregressive order-one process model? AR(1) may be suitable when the autocorrelation function of the stochastic process has significant components up to lag three or so.

L130-136. The purpose of localization should be explicitly stated, which would be helpful for the relevant issue described in the paragraph starting L302.

L140. In equation (7), is there a case where the magnitude of γ_1 is greater than 1? Figure 2 shows that some models have means close to ±1.

L145-146. Specify whether area weighting is processed or not.

L168-170. This quantile comparison procedure is unclear.

L194-197. This sentence is complicated and should be clarified more.

L207. The symbols in Equation (8) are not fully described. Instead of this equation, the integral of difference between two CDFs would be more understandable as a definition of the energy distance.

L224-225. Is there anything to be added? Readers would be curious about what kind of characteristics of the two models result in such outlier results.

L233. The description about the equatorial region appears to be limited to January, if so, it should be stated as such.

L248-254. Readers would be curious about the MIROC6 results, in which relatively many ESM points appear outside of the emulator range in WAF. Is there anything to be mentioned in this regard?

L258-260. The authors mention different timing of changes in snow cover among ensemble members for lower correlations in April and October, but this is not the case for several ESMs that have only one training run and zero test run.

L275. Meaning of "such higher power spectra" appears unclear.

L291. It looks mixed positive and negative rather than "generally low values."

L288-289. Is it correct for "deviations of the ESM training (and where available test) runs from the full emulations"? Figure 8 caption writes quantile deviations of the monthly emulated quantile from that of its ESM training and test runs although the embedded figure title is different. Which is the base for the deviation?

L311. Around here, it is hard to trace relations between descriptions in the text and corresponding parts in the figures. Even if the complexity of the figures is unavoidable to some extent, main points and their implications should be clearly indicated.

L341-344. Figure E2 shows an exceptionally bad performance of NorESM2-MM. Is there anything to be mentioned?

L335-338. Regarding greater CDF distance in January, Figure E2 may not clearly indicate such distinctive difference between January and July. Grid lines in the plot space would be helpful for identifying the difference.
Figure E2 also shows that CanESM5 has greater distance in January as well as MIROC6 and MPI-ESM1-2-LR.

Conclusion and outlook. While the emphasis is on future developments that take into account biophysical variables, there remains the question of how similar the variability components by each ESM are to observations, although the latter is out of scope in the current manuscript. What needs to be focused on in this regard would include properly

emulating how major variability modes, such as ENSO, modulate with warming, considering any dependence on ESMs. If a certain variability mode includes some memory effects associated with , for example, land soil moisture they may need to be modeled by higher-order autoregressive processes, and if the variability affects remote climate on a global scale through teleconnections, the localization of the spatial covariance structure may need to be improved. In any case, future developments will need to be described in a broader scope.

Acknowledgement. Refer to https://pcmdi.llnl.gov/CMIP6/TermsOfUse/TermsOfUse6-1.html to confirm whether acknowledging CMIP6 is appropriate, and requirement for citing CMIP6 model output.

Figure 3 and 4. Units are missing. The figure legend implies that plotted data are three members for the ESMs, the three patterns for the sum of the yearly and seasonal cycle, and 50 realizations for each of the three patterns for the emulation, but should be fully described in the caption text. The second sentence of the caption text should be clarified about whether the reference temperature is monthly or yearly.

Figure 6. Although the text in 5.2 (L271) implies that the box plots in Figure 6 show distributions across different realizations by the emulator, the figure caption should be fully described about how different bands (50 elements), different emulator realizations (50 members), and different training and test runs (ESM-dependent ensemble number) are processed to represent the data distribution. Also, if the whiskers in this figure (Figure 7 as well) are drawn as in Figure 2, indicating a min-max range, explicitly state so, otherwise describe it accordingly.

Figure 8. It would be good to have a guide so that readers easily identify representative geographical zones corresponding to the individual regions alighted on the horizontal axis.

Editorial comments

L70. Inline Mathematical symbols should be italic.

L73, L95. Indent is unnecessary.

L105. Section 3.2.1, not section 4.1.1, but this indication within the parenthesis is redundant.

L110, L115, etc. Add comma to the end of the preceding expression and lowercase "Where".

L136-137. "based on", not "based off."

L169, L175, L202, etc. Long dashes (em dashes) are not typesetted correctly.

L183. "properties of the monthly temperature response" (maybe "of" is missing)

L203, 204, 220. "added ontop" and "ontop of" may not be common wording.

L224. HadGEM3-GC31-LL, not HadGeM3-GC31-LL. Put it after ACCESS-CM2 if alphabetically ordered.

L234. "this" in "The source of this" is unclear.

L237-238. In this context, "boreal winter", not just "winter". Pay attention to whether inappropriately referring to specific seasons in terms of the Northern Hemisphere

throughout the manuscript.

L241. "four selected ESMs", instead of "a select 4 ESMs." See the journal's English guidelines for numbers, and, if needed, spell out numerals less than 10 throughout the manuscript.

L244. It is unclear "all ESMs" are the four selected ESMs or all the ESMs used in this study.

L328-329. "latent and sensible heat fluxes", not "latent heat fluxes."

L163. Spell out SREX here. Also in Figure B1 caption.

Figure C1 and C2 captions (also, Figure D1 and D2 captions). For Figure C2, "Same as Figure C1, except for January" would be fine. Referring to the Benjamini/Hochberg correction may be required in the Figure C1 legend.

Figure E1. It might be better to replace X and Y axes for comparison with Figure 12. A more concise caption would be "Same as Figure 12, except that all CMIP6 models are shown for the global land."