Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC1
https://doi.org/10.5194/gmd-2022-144-RC1, 2022

Comment on gmd-2022-144

Anonymous Referee #1

Referee comment on "A new precipitation emulator (PREMU v1.0) for lower-complexity models" by Gang Liu et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-144-RC1, 2022

General Comments:

PREMU presents a novel method for emulating precipitation by considering modes of temperature variability, thus providing a more precise approach towards representing such a dynamic climate variable. I believe this paper would be useful for the GMD reader base and provide new avenues for tackling emulation of other climate variables that do not scale so directly with Global Mean Temperature. In general, the text could undergo some restructuring for understandability and clarity purposes with some suggestions to do so provided in the specific comments below. Some more general comments on the methodological approach that could also be improved upon are as follows:

- The methods overall make sense but could be more explicit in terms of the PCA analysis done: how are the spatial patterns derived, how sensitive are these patterns to sample size, do these patterns make sense/have physical explanations behind them that are identified in previous literature. Additionally, there are parts of the discussion where it is mentioned that the sensitivity of the PCA based coefficient matrix is tested (L271-L281), this could be mentioned before as a subsection under Methods and Results.
- The 3-month mean of temperature does not necessarily mean that lag effects are captured, and the authors should elaborate on this choice as compared to e.g. a multi-linear regression. Again the discussion brings about some tests of lag effects using premu-1mon and premu-6mon but this still considers monthly averages, furthermore this part would also fit better as a subsection under methods and results. In terms of assessing memory effects, the evaluation could be enriched by, for instance, looking at lag-1,2 and 3 correlation coefficients between subsequent month values unless the authors are not interested in preserving month-to-month correlations within PREMU predictions which should then be justified.
- The results only focus on MAP, inter-annual variations and trends and annual spatial patterns, given that PREMU is monthly, readers would also benefit from seeing these analyses done on a seasonal/monthly level which would also be more impact relevant.

Specific Comments:

L8: Elaborate on the term rainfall features, otherwise it is a bit vague

L20: MESMER-M is a monthly extension of MESMER, given that PREMU takes monthly input this may be relevant to mention (Nath et al. https://doi.org/10.5194/esd-13-851-2022)

L23-26: consider moving the policy relevance of LCMs higher up in the abstract, the final sentences should focus on PREMU itself.

Introduction: Generally a good overview into the problem this study is trying to tackle, however it tends to downplay the real value of this work and could benefit from some restructuring to make the points of this study clearer:

- L28-L32: Consider merging this with the next paragraph, otherwise it comes across as a bit redundant.
- L34-L36: Paints a good picture of LCMs, in relevance to this study and generally for LCMs it should be emphasised that they focus on reproducing only a few climate variables that are most impact relevant. This allows them to have such reduced computational expenses as well as differentiates them from ESMs. If this is what is meant by "highly parameterized macro-properties" it should still be elaborated upon, in general this term can imply a lot of things so it should be made more specific.
- L42: The following reference could be insightful for the author/readers when mentioning pattern scaling:
  - Tebaldi, C., & Arblaster, J. M. (2014). Pattern scaling: Its strengths and limitations, and an update on the latest model simulations. *Climatic Change*, *122*(3), 459–471. https://doi.org/10.1007/s10584-013-1032-9
  - Tebaldi, C., & Knutti, R. (2018). Evaluating the accuracy of climate change pattern emulation for low warming targets. *Environmental Research Letters*, *13*(5). https://doi.org/10.1088/1748-9326/aabef2
- L43: It could be good to add that there is already some work that successfully considers modes of variability within emulation, see:
  - Mckinnon and Deser 2018: https://doi.org/10.1175/JCLI-D-17-0901.1
  - Mckinnon and Deser 2021: DOI: 10.1175/JCLI-D-21-0251.1
- L45: I feel like the fact that precipitation is a crucial variable but has statistical features that make its representation within traditional LCM approaches difficult is a strong message to highlight. Consider restructuring the first three paragraphs to emphasise this i.e.
  - First paragraph:
    - There is increasing demand for climate information however lack of sufficient computational power for running ESMs for all potential future emission scenarios etc. .
    - LCMs are a solution to this representing key climate variables, however in terms of gridded data most use pattern scaling based off GMT which is suited for local

temperatures.
  - Second paragraph:
    - Precipitation has high spatio-temporal variability and is affected by atmospheric dynamics, inter-annual modes of variability etc. making representation within LCMs using only GMT as an explanatory variable difficult e.g. OSCAR, IMOGEN
    - Nevertheless precipitation is a crucial component of the water cycle (Eltahir and Bras, 1996; Trenberth et al., 2003; Sun et al., 2018), has key societal implications and is closely associated with the functioning of terrestrial ecosystems
- L65-L67: Here it may also be a good idea to elaborate on the *rainfall features* this study finds especially relevant to preserve and provide relevant literature references (is it just inter-annual variability, spatial and seasonal variance, or are there more to think of under a changing climate? What are the drivers of inter-annual variability in precipitation e.g. El-Nino, Indian Ocean Dipole?) and why basing off gridded temperature provides a more promising avenue as compared to GMT to achieve this.
- L66: The end-use of the resulting emulator should already be introduced, it seems like it can have multiple uses in terms of looking at future emission pathways e.g.
  - Stand in for OSCAR and IMOGEN's precip. pattern scaling (here it should be noted that these LCMs have interactive atmospheric chemistry and endogenous calculations of e.g. biomass burning which may lead to differences from ESMs in the PCA analysis and therefore the resulting emulator is not an ESM one but simply a module within the aforementioned LCMs)
  - An extension of the MAGICC-MESMER emulation chain

L70: Consider making Data and Methods two independent sections, it seems like they have that role in the text anyways.

L71: The text structure could benefit from introducing the approach before going into the subsections i.e. 1) you divide the emulator calibration into that done on the historical period 1901-2016 and that done on the future period 2016-2100 2) for the historical calibration, you use observational data while for future calibration you use ESM data.

L72-73: the use of "first tested" is a bit misleading since you both calibrated and tested PREMU on historical data, maybe replace with "we first demonstrate the applicability of the emulator on observational data provided by the Global Soil Wetness Project Phase 3."

L77: The statement: 'yet provides additional high frequency signals which are lacking in previous products' leaves the reader wanting more, if possible elaborate a bit on what this means.

L78: Elaborate on what Tair is, is it 2m air temperature? For translation to ESMs this is quite important.

L80: Does the author perform regridding to 2.5°x2.5° so that these variables are on the same grid with each other as well as ESMs used? This should be elaborated on and if not done, justified.

L92: Out of curiosity, why first-order? Other archives (e.g. Brunner L., M. Hauser, R. Lorenz, and U. Beyerle (2020).The ETH Zurich CMIP6 next generation archive: technical documentation. DOI: 10.5281/zenodo.3734128) use second order.

L94: When calibrating, do you use all initial condition ensemble member available per ESM for SSP 5-8.5? Same goes for testing on the other SSPs?

L94: Instead of going straight into calibration the author should also first provide the PREMU framework, consider therefore breaking up the text to introduce general approach (gridded temperature, PCR etc.) , emulator framework (here introduce equations and key variables alongside Table 2 and Figure 1), calibration etc.

L113: It is not convincing that taking a 3-month average across the month of interest and the preceding 2 months will sufficiently capture the influence from the previous 2 months. Moreover this could average out the general seasonal transitions in temperature, does the author have a justification for this instead of using e.g. a multiple linear regression

L113-L115: "ESM may under-represent the effects of topography and aerosols on the precipitation from observations" do you have a reference to back this up with? I see this later on in the discussion, consider bringing them up here too.

L120: How the spatial patterns of the PCA components is obtained needs to be elaborated upon

Fig 2: It is strange that there is such consistently high coefficient values occurring in the Arctic, is there a reason to this? If possible, the explanatory power of such spatial patterns towards precipitation should somehow be evaluated in terms of how representative they are of global circulation patterns and inter-annual modes of variability. In such, one suggestion could be to check how the coefficients change when fitted using a leave-one-year out approach.

L141: I find equation 3 redundant as it is the same as equation 2 only with grid-point specific coefficients, for simplicity's sakes perhaps these two equations could be merged

L142-L143: Redundant as the same as before but only for grid-point level

L153, L156, L157: These equations are repetitive, again consider simply referencing to the ones already introduced before.

Validation: Generally this section cuts off to abruptly and should be restructured as well as more detailed. There seems to be too much emphasis on defining the predicted variables and how they are obtained from the test set such that the actual method for validation is lost. The author should try be more explicit on what they are validating for i.e.

- Validation is performed for each predicted variable, T^{PCA}, P^{global}, P^{grid}
- What is validated for? It should be specified that the MAP, the inter annual variance and the trend are checked etc. Evaluation metrics should be provided (e.g. percentage error, correlation) and how they are calculated should also be explained.
- Given the high spatio-temporal variability of precipitation, it would be interesting to see how well PREMU does in preserving spatio/temporal autocorrelations, e.g. lag-1 correlations between months and spatial cross-correlations
- Readers are left wondering if any evaluation on the PCA is done e.g. comparison between ESMs and observations, linkage to existing literature. There should be some physical relevance to the PCA components obtained as well as stability checks e.g. is 70% of variance explained by 10 PCA components within ESMs too, does this apply for all months, what happens to the spatial patterns when some training years are left out.

L162, Eq 7: This is a methodological addition and seems out of place in the validation section, consider moving it to methods (e.g. "from further tests etc. we decided to apply a correction factor etc.")

L170: when taking the areally-average is it a weighted average according to latitude, if not why? Areally-averaged could also be replaced by simply average or average over grid points as areally-averaged reads ambiguously.

L172: How do you check the similarities in inter annual variations i.e. using pearson correlation coefficient between the inter-annual variance or absolute values or applying a low-pass filter to extract the variance before checking the correlations?

Results: Quite nice findings! It should be made explicit the annual precipitation is mainly verified for, given that PREMU is a monthly emulator. It raises the question of the seasonal performance of PREMU, is there a reason the authors did not show this?

L271-L305: These seem to be more points within the results, in general it would be good to first introduce that these things (e.g. memory effect) were tested for an how in the methods and then provide the results in the Results. Otherwise the Discussion is lacking structure and quite long.

L311: Different version of calibrated PREMUs should probably be introduced in their own subsection

Conclusion: Generally the discussion is quite long and covers myriad of possibilities, consider merging some parts into the results (as mentioned above) some parts in conclusion and dividing into further subsections e.g. Further Developments/Versions (PREMU-1MON, PREMU-LAND etc.)

Table 3: It is hard to contrast so many numbers, perhaps providing percentage difference relative to actual trend value or something along those lines would be better?

Figure 1: Having read the text I understand this figure, however by itself it is quite complicated and could benefit from some more graphics and simplification of the text.

Figure 4: Readers would be interested in seeing differences in trends between GWSP and PREMU characterised, it seems like there are areas where the direction in trend is different (e.g. Australia, West Africa) which also matters vs simply looking at over/underestimation of changes. The discussion point that more differences can be seen in this figure as compared to Figure 8 due to topography and aerosols should also be mentioned when describing this figure as it enriches the analysis.

Figure 8: again it may be interesting to characterise the differences: where direction of changes are properly captured and where they are opposite

Editorial Comments:

L28: I would start the sentence off with the subject rather than verb, especially if it is at the beginning of the paragraph, i.e.

Earth system models (ESMs) are the primary tools to study the impact of greenhouse gas emissions on our climate, representing all the important Earth system processes (IPCC, 2013).

L30: run **the**

Figure captions: emulations are referred to as **our** emulations, I am not sure how formal this is.

Section 2.3.2: the Table should be labelled and given a number

Eq 4: T has the superscript Timeseries,val which is not in line with that of eq 1, is this intentional?