

## ***Interactive comment on “PALEO-PGEM v1.0: A statistical emulator of Pliocene-Pleistocene climate” by Philip B. Holden et al.***

### **Crucifix (Referee)**

michel.crucifix@uclouvain.be

Received and published: 18 January 2019

The authors propose a latitude-longitude reconstruction of the climate of the whole Pleistocene, using a Gaussian process emulator calibrated on two experiment designs with the PLASIM-GENIE model. It uses CO<sub>2</sub> and sea-level as inputs, based on an inverse modelling reconstruction provided by Stap et al. and, where available, ice core observations. R code with input files are provided.

The process for designing and calibrating the emulator is largely based on earlier work (experiment design, PCA emulator). There is however a cunning novelty: using two similar experiment designs for isolating the climate anomaly caused by ice sheets.

Although perhaps not in line with the reviewer ‘etiquette’, I wish to make a personal

C1

comment, in the hope that the editor and authors will be forgiving find this intrusion useful for the evaluation of the work under concern here.

After the articles of Araya-Melo et al. (2015), Bounceur et al. (2015), and Lord et al. (2017), we had everything in place at UCLouvain to provide a similar reconstruction, and in fact we tried a few. What stopped us from publishing are:

- that the ocean circulation in LOVECLIM was not behaving adequately, with stronger, deep ocean circulation at glacial maxima, at odds with proxies for ocean ventilation. At some stage we thought of mending the simulation with an additional freshwater perturbation, but this work was never finalised to the point of publishing.
- the strategy used in Araya-Melo et al. 2015 of summarising ice sheet forcing with a single quantity, and which is applied here by Holden et al., can be problematic. There is nothing to guarantee that Weichselian ice sheets were located similarly to those of the Early Pliocene, and also, as the authors rightly acknowledged, the build-up and decay phases of ice sheets are quite asymmetrical during the late Pleistocene. This might not be that much of a problem in certain applications, but it can be heavily misleading to users who would use this product in Europe or in Siberia without much discernment. For example, I would be particularly worried of archaeologists using the provided reconstructions near the limits of ice margins. A 3-D reconstruction of the Pleistocene can be very popular, so it needs to be disseminated wisely.

This little experience brings me to the following, and related comments, about the present contribution by Holden et al.

1. It definitely needs to come up with appropriate health warnings about usage limits of the reconstruction. This is particularly crucial since the introduction presents it

C2

a product oriented to end-users including archaeologists and biodiversity experts. A critical evaluation of the validity of the reconstructions near the North Atlantic (with emphasis on ocean circulation effects) needs to be provided.

2. In the introduction it is clearly said that uncertainty attached to the emulator (here, as a surrogate of GENIE-PLASIM) is distinct from the model uncertainty. This is true, and hence what would think that the evaluation (or “validation”) of the emulator (as a surrogate) should be clearly distinguished from the evaluation of the model as a representation of the real climate. I found that this important distinction is pretty blurred in the section 6 (strangely divided into a section heading an a subsection 6.1). In fact there is very little about the evaluation of the emulator as a surrogate of GENIE-PLASIM. The authors refer to the PMIP ensemble and feel comfortable that the emulator-based reconstruction is in broad agreement with PMIP simulation of the mid-Holocene and the LGM, but in doing this the authors are mainly evaluating the reconstruction, not the emulator as a statistical surrogate. And, as I suggest in point 1. above, this evaluation is not providing with non-climatologist users with enough information about its application domain (the dos and don'ts). It is also quite uncomfortable that the emulator provides so-called bioclimatic variables (MIN and MAX over the seasonal cycle) while the “validation” is made on the basis of seasonal averages.
3. Downscaling. Is this correct that downscaling as presented here assumes constant sub-grid correction anomaly, defined as the difference between the present-day observations and simulated grid-box-mean in a reference experiment (the anomaly being on the log of precipitations in the low-precipitation areas)? This treatment is arguably inadequate in palaeoclimate applications, where topography, surface type (think of Swiss glaciers to take but one example), ice sheet margins, land-sea mask, and ecotone boundaries vary substantially. Again, aren't we misleading the users by providing the illusion of a high-res reconstruction, while it may in fact be quite wrong at places? For reference, Levvasseur and co-authors

C3

have provided some thoughtful contributions about downscaling in palaeoclimate applications (e.g.: Levvasseur et al., 2010, *The Cryosphere*, 10.5194/tcd-4-2233-2010, and subsequent references).

#### 0.1 Minor comments

- L. 280-295: Define the R-squared score and clarify what is mean by “the performance averaged over the eight emulators” (since the P-metric is near one, an arithmetic mean of P may be inadequate).
- Clarify the approach implemented for calibrating the length-scales appearing in covariance functions: do they vary across components, across variables?
- The original reference for Berger and Loutre, as quoted here, is Berger and Loutre (1991), doi 10.1016/0277-3791(91)90033-Q . It could be cited along with the Pangea reference.
- Figures 3 and 4 are a bit overwhelming, with small character size, and only the one with eyes trained in deciphering PMIP-type experiments will understand that the anomalies seen here are reasonably expected from GENIE-PLASIM and understand its limits.

#### 0.2 Conclusion

The article could be a nice addition to current efforts in simulating the Pleistocene climate, but it is ambiguous as to its objective. If the authors ambition is to provide an technical, significant improvement on emulation, then they need to focus more on the evaluation of the emulator as such, and be more thorough in the discussion of the different technical options. If the ambition is to provide a final product to be used by

C4

non-climate users, then I would urge the authors to be much more critical about the pitfalls of the current reconstruction, and in the present state, I would actually discourage dissemination of this product, since the risks of it being misused are too large.

---

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2018-242>, 2018.