Clim. Past Discuss.,
https://doi.org/10.5194/cp-2019-91-AC1, 2019

Climate of the Past
Discussions

CPD

Interactive comment

Interactive comment on "A comprehensive history of climate and habitat stability of the last 800 000 years" *by* Mario Krapp et al.

Mario Krapp et al.

mariokrapp@gmail.com

Received and published: 31 October 2019

We would like to thank Reviewer 1 for his comments on our discussion paper. Below, we have listed the reviewer's comments as quotes, our response in normal text, and action points in *italics*.

> "This paper presents a method for estimating climate and ecological changes over the past 800,000 years. This is incredibly ambitious and I welcome the endeavour and aims of the research. However, I am somewhat perturbed by its execution and cannot presently recommend it for publication. I have some queries that arose from my reading of the work. Some may be due to my own misunderstandings, but as a collection

I feel they bring the validity of the proposed method into question.

The remit of this research is to build an emulator of climate and ecology as simulated by the HadCM3 model. Any errors in HadCM3 will therefore be unavoidably replicated by the emulator: yet this is not an obstacle preventing useful information being gleaned from such a tool.

HadCM3 has previously been used to build emulators - and therefore several training sets exist (e.g. Arayo-Melo et al, 2015, and Lord et al., 2017). Given these have been designed to sample parameter space in a near-optimal form, I was surprised that the current work uses only snapshots of conditions that have existed over the past 120,000 years. The authors provide no explanation for this choice, nor discuss its limitations."

Our response: It is true that previous emulators have been designed to sample parameter space that is near-optimal and that is as large as possible. We are limiting our approach to glacial–interglacial climates. Parameters outside that range are likely to lead to extrapolated climate reconstructions. Fortunately, the last 120ka have been extreme in terms of the glacial-interglacial climate ranges covered. The LGM (26ka BP) was one of the coldest periods and one with the largest glaciations of the Quaternary and likewise the last interglacial (115-130ka BP) was one of the warmest periods. In fact, the parameter sampling (despite being non-random) is much denser for glacial–interglacial climate emulation than any previous emulator designs. We use 72 time slices, more samples than have been used in previous studies, e.g., n=61 in Arayo-Melo et al. (2015) and n=40 in Lord et al. (2017). In summary, the sampling of the forcing is much closer to any forcing of the last 800ka compared to the previous emulator studies using HadCM3.

ACTION: *We will provide more details about our choice of the sampling strategy in the* *introduction and the model description. We will highlight that our emulator approach uses more samples which themselves are much closer to the expected phase space trajectories, i.e., of the last 800ka. This way the emulator is only applicable to boundary conditions which are similar to the Quaternary ice age climates.*

> "A second issue with this training set, as well as others, is the ice-sheet extent. In previous work, ice volume has either been considered an input parameter (Arayo-melo et al., 2015) or emulated through a fixed adjustment (Lord et al., 2017). Ice-dynamics, and the substantial lags that they introduce into the Earth system, are completely neglected in this work. This effectively assumes that ice sheet impacts are wholly and instantaneously determined by CO2 and orbital configuration. I anticipate this would explain elements of the model-data mismatch shown in Figs 6 7. There is also no recognition that last glacail cycle may not represent all glacial cycles (despite the mid- Brunhes transition)."

Our response: We know that our approach has limitations as outlined by the reviewer and indeed, ice-sheets are a factor accounting for the mismatch over ice-covered regions. However, we think this is more of a sampling problem rather than a representation of the temperature–ice sheet feedback. The HadCM3 model has been set up without interactive ice sheet models. As such the ice-sheets are fixed in size as well as in height and surface albedo. Therefore, we understand that the HadCM3 climate response is only being affected by the presence of ice sheets (and any dynamical circulation changes). So he neglect of ice dynamics is due to the HadCM3 setup and not so much due to the emulator itself. As a consequence, we have to accept "that ice sheet impacts are wholly and instantaneously determined by CO2 and orbital configuration", as correctly observed by the reviewer, as a limitation of the underlying HadCM3 simulations (but not the emulator itself).

Regarding the last point, we assume the emulator response to the last glacial cycle (i.e., the last 120ka) is representative of previous glaciations. We don't claim that all glaciations are equal. That is why we use a varying forcing (ice sheet masks, CO2, and orbital parameters) for all the previous glacial cycles for our reconstructions. Previous glacial cycles are therefore not simply repetitions of the last glacial cycle but linear combinations of the different responses to the external forcings. In this way, our emulator can be understood as a decomposition of the climate response into the various forcing components, e.g., $T = T_{orb} + T_{co2} + T_{mask}$.

ACTION: *: We will clarify the limitations of the HadCM3 setup early on in the model descriptions and we will add those raised issues in the discussion, i.e., what does this mean for the reconstructed climate.*

> "A further unanswered question arising from the choice of training data revolves around the ecological reconstruction. Only surface temperature, humidity and precipitation are emulated, and then the biomes estimated off the back of this data. HadCM3 has an dynamic vegetation model (Triffid), although I'm unsure whether it was incorporated in this simulations. Certainly offline simulations of the Sheffield Dynamic Global Vegetation Model using the full HadCM3 climate model output have been performed for a subset of the training simulations (Singarayer et al., 2011, doi:10.1038/nature09739). Using these data could provide a useful comparison to the ecological modelling component in section 5 - i.e. can the emulator replicate the simulator response."

Our response: TRIFFID is part of this HadCM3 model setup, so this question makes perfect sense. Yes, we could use the vegetation model output, e.g., plant functional types, to compare the random forest classifier results with actual model output. Whereas we envisioned the ecological reconstruction as an example for how

Printer-friendly version

Discussion paper

such a dataset could be put to use, we now feel that it is more distracting from the main message of this paper. As the reviewer mentioned a comparison id required to assess the quality of the ecosystem reconstructions which would add another layer of complexity.

ACTION: *We would leave out the ecosystem reconstruction part in a revised paper because we feel that the reader may be overwhelmed with yet another reconstruction approach and its validation/evaluation. Instead, we would focus on the technical aspects of the emulator and provide a more comprehensive validation with HadCM3 model output.*

> "Moving beyond the training data choices, I have four questions about the choice of the method applied. Firstly, the linear (or log-linear) regression used at each grid point is different to previous efforts. Arayo-Melo et al. (2015) spent substantial effort developing an approach that inherenetly builds in the spatail covariances inherent in climate using EOFs dimension reduction. The justification you give to avoiding this approach is that linear regression results in "well-behaved". Isn't this just another way of saying that you avoid non-linear transitions, but are these not a widely-accepted feature of the climate system. Also you should bear in mind that the whilst the functions are well-behaved in time, you have removed any such condition in space. Personally, I prefer the dimension reduction approach, as it pulls out climate features from any grid-point noise."

Our response: We claim that the emulator works perfectly fine for each grid point and that it recovers the climate response as in HadCM3. And yes, this implies that we don't control how grid points interact on different spatial scales. However, we tried to make the argument that the spatial coherence is recovered by the regression coefficients themselves (see Fig. 2 in the discussion paper). To make this point clearer, we have attached plots of the covariance matrices for the (full) data set of both, HadCM3 and the emulator, GCMET (Figs. 1 & 2). The plots show the covariance between each grid point (96x73=7008) for the whole 72 time slices, i.e., a 7008x7008 covariance matrix, and the difference between the covariance matrix for both temperature and precipitation. The covariance matrix is a useful indicator for the representation of spatial structures in the emulator (and HadCM3). Smaller differences mean that GCMET captures the correct spatial covariances. As can be seen in the Figures, the covariance matrix for GCMET output is structurally similar to HadCM3 and the difference is mostly close to zero.

Regarding the non-linear transitions, we can only build on any assumptions that have been put into the HadCM3 simulations, or the climatological output. Although HadCM3 can account for non-linearities (it's a climate model after all), the climatological means (30-yr averages) on which we build GCMET, may not represent any non-linearities found in the climate system. We simply don't account for non-linear transitions because GCMET is supposed to be a surrogate model for the equilibrium climate model response.

A dimension reduction approach such as PCA removes lower-order variability modes in order to make the emulator approach computationally feasible whereas a linear regression model can make use of the full data set. In that sense, the linear regression is less restrictive about the inputs than a dimension reduction approach.

ACTION: *We will expand the discussion around different emulator approaches and point out advantages and disadvantages of using one approach over the other more thoroughly. Specifically, we will add an analysis to show that GCMET preserves spatial covariance sufficiently enough in comparison to the HadCM3 model outputs. A* *preliminary figure is attached to this response.*

> "My second question about the methodology is why the function in equation 5 was selected for precipitation. Specific humidity is strongly related to temperature, unlike relative humidity) so it unclear to me that you can treat them as indepedent variables. The relationship between them can clearly be seen in Fig 2D - where the patterns are approximately opposite to each other."

Our response: This is a valid point. In a strict sense, we shouldn't use specific humidity and temperature because of their collinearity. Admittedly, relative humidity is also a better proxy for precipitation: the closer the air mass is to saturation the more likely it is to rain. However, relative humidity is more difficult to predict.

ACTION: *We will provide the alternative formulation for precipitation based on relative humidity.*

> "My third methodological question revolves around downscaling. I appreciate your effort to downscale the climate results using High-resolution models. However, I wonder if you have applied them is the most optimal method. The (low-res) emulator captures climate changes from a (known) mean state. Your downscaling approach acts to modify those cliamte changes by modelling the resolution dependent aspects of those changes. If you want to convert the emulator output from climate anomalies to absolute climate, you must build back in the known mean state. Your choice of mean state is not explicit, and one wonders whether this might most appropriately be a very-high resolution satellite dataset (see question later about Fig. 6)"

Printer-friendly version

Discussion paper

Our response: That is another very good point. Indeed, the mean state could be based on an observational data set (such as ERA-20C averages, Poli et al., 2016). This is a different way of saying that we need to add a bias correction to the emulator output. And we could. We thought that this is beyond the scope of our paper but we are happy to reconsider. In previous (discussion) paper we show that commonly-used bias correction methods already reduce model biases when comparing to paleo-climate observational data sets (Beyer et al., 2019). We can easily apply such a bias correction for our final data set. We want this paper also to be the reference for a comprehensive reconstruction of the past 800ka. So, it makes sense to have those data available as bias corrected emulator outputs.

ACTION: *We will bias-correct the emulator output and add a paragraph about the final climate reconstructions in the model description and the discussion section.*

> "My final question about the methodology is where are the error bars on your estimate. Whilst I recognise that you cannot capture the error associated with HadCM3's biases, it must be possible to provide error bars of how well your model emulates the simulator. This is surely vital for the verification shown in Fig. 4 - does the true simulator response lie within the error bar estimates? What is the additional errors introudced by the simplicity of the ecosystem model applied to the emulator outputs."

Our response: It is possible to estimate the uncertainty of the emulator in terms of a confidence interval (CI) with level confidence level α, of a parameter p using the standard error (SE):

$$CI = p \pm q \cdot SE \quad with \quad q = CDF - 1(1 - \alpha/2)$$

As we mentioned before, we would drop the ecosystem reconstruction part for this

Printer-friendly version

Discussion paper

paper and, instead, focus on improving the model validation/verification part.

ACTION: *We will provide maps for the confidence interval (e.g., the 95% range) and (optionally) maps of the standard error. For Fig 5A, we propose to add the confidence levels so we can show how well the emulator results compare to HadCM3 and to the global mean temperature time series.*

> "I have an additional question about the validation in Fig. 6 C and D. What is being assessing here? The emulator only models changes in climate, not absolute variables as shown in Fig. 6D. I suspect that the assessment in Fig. 6C is more about the resolution of underlying simulator and little to do with the emulator. In fact, previous efforts (e.g. Lord et al., 2017) have used the present-day climate as the "mean" from which anomalies are calculated - which instead potentially allows the use of ERA-20C as the baseline. Under the test shown in Fig. 6, such a slightly revised emulator would be perfect."

Our response: That is a good point and would fall within the proposed changes of our next to last comment.

ACTION: *see our next to last comment*

> "Given the quantity and importance of these questions about the creation and validation of the emulator, I have chosen not to review the results in any detail."

Our response: We regard the comparison to proxy data as one of the essential pieces of the paper. This part would serve as the ultimate test showing how well our emulator results compare to observations (any previous validation was purely restricted to the "model" world of HadCM3 and GCMET).

References

- Araya-Melo, P. A., Crucifix, M., and Bounceur, N.: Global sensitivity analysis of the Indian monsoon during the Pleistocene, Clim. Past, 11, 45-61, https://doi.org/10.5194/cp-11-45-2015, 2015.
- Lord, N. S., Crucifix, M., Lunt, D. J., Thorne, M. C., Bounceur, N., Dowsett, H., O'Brien, C. L., and Ridgwell, A.: Emula- tion of long-term changes in global climate: application to the late Pliocene and future, Climate of the Past, 13, 1539-1571, https://doi.org/10.5194/cp-13-1539-2017, 2017.
- Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thepaut, J.-N., Tremolet, Y., Holm, E. V., Bonavita, M., Isaksen, L., and Fisher, M.: ERA-20C: An Atmospheric Reanalysis of the Twentieth Century, J. Climate, 29, 4083-4097, https://doi.org/10.1175/JCLI-D-15-0556.1, 2016.
- Singarayer, J. S. and Valdes, P. J.: High-latitude climate sensitivity to ice-sheet forcing over the last 120kyr, Quaternary Science Reviews, 29, 43-55, https://doi.org/10.1016/j.quascirev.2009.10.011, 00092, 2010.

Interactive comment on Clim. Past Discuss., https://doi.org/10.5194/cp-2019-91, 2019.

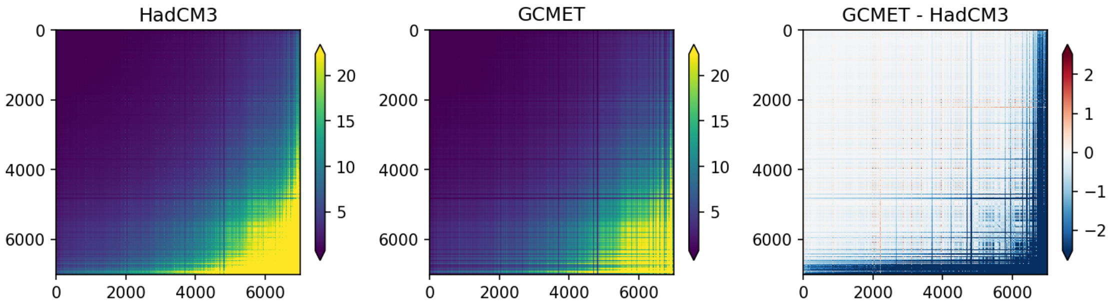

Fig. 1. Spatial covariance matrices of temperature (in units of K2) calculated for all grid points (n = 96x73 = 7008) of HadCM3 (left), GCMET (middle), and the difference of the two covariance matrices (right

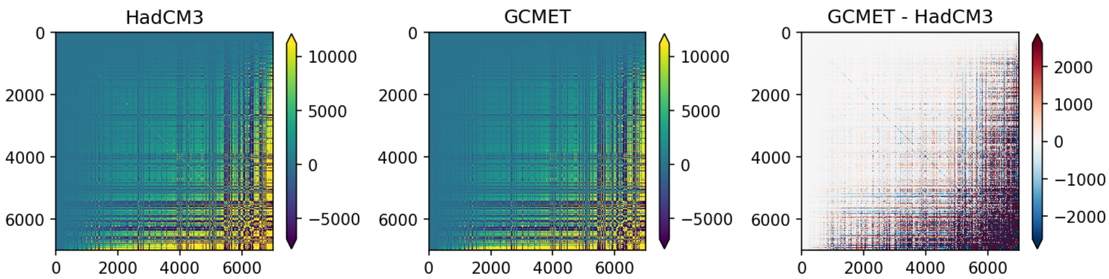

Fig. 2. Spatial covariance matrices of precipitation (in units (mm/a)2), calculated for all grid points (n = 96x73 = 7008) of HadCM3 (left), GCMET (middle), and the difference of the two covariance matrices (

Printer-friendly version

Discussion paper