Clim. Past Discuss.,
https://doi.org/10.5194/cp-2019-91-AC2, 2019

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 2 for the constructive 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 manuscript uses a set of snapshot simulations with HadCM3 over the last 120kyr to build an emulator that then extends into the last 800kyr. In addition a few atmosphere only simulations at higher spatial resolution than the coupled HadCM3 covering the more recent period are used to add more spatial variation into the long time-scale climate emulation. While the approach is novel and seem very promising to me, I have still some diffi-

> culties to understand how well this approach is actually validatable. Maybe this is (partly) due to my limited understanding, but from reading the text and the figures it becomes not clear to me whether the approach is justified and how large potential errors could be. Therefore, in its present form I cannot recommend the paper for publication.
>
> Fig 5 should be the most convincing figure to show that the emulated climate gains quality because it uses - in addition to the HadCM3 snapshots covering the last 120kyr - additional information on spatial patterns from higher-resolution simulations. However, I get very confused by this figure and its discussion: Fig 5B shows GCMET-LO, while Fig. 5C shows LOVECLIM and GCMET. How can I see that GCMET (higher resolution I assume) does better in the spatial patterns than GCMET-LO and LOVECLIM?"

Our response: Apparently, there is a typo in the labels. Sorry about that. Fig. 5C shows how GCMET-LO matches proxy reconstructions in terms of the correlation coefficient compared to LOVECLIM model output. Fig. 5A shows how well GCMET compares with proxy-based global mean temperature estimate throughout the last 800ka. Although it has a higher spatial resolution, this figure was not intended to suggest that this is because of the higher resolution. The whole Fig. 5 is intended to highlight the emulator's capability to reconstruct the temperature change signal realistically enough, i.e., the temporal, climate change signal. Complementary, Fig. 6 is intended to show how much more spatial detail can be recovered by adding the higher-resolution corrections on to of GCMET-LO, and how the different spatial resolutions compare to the observed (i.e., present-day) ranges of temperature and precipitation (from the ERA-20C re-analysis dataset, Poli et al., 2016). In conclusion, the emulator recovers the observed climate change signal through time at different (proxy) locations, and it also recovers the correct shape of the observed spatial climatic heterogeneity across the different continents.

ACTION: *This part of the paper seems to be confusing. Therefore, we would restructure Sections 2 and 3 to clarify that we are showing two different, but unrelated, aspects of the emulator: i) the temporal component, covered by the coarse-resolution HadCM3 emulation and its comparison (currently Section 3), and ii) the spatial component, covered by the dynamical downscaling (currently Section 2.3). The dynamical downscaling section would then follow after the proxy-comparison.*

> "Next, section 3 is meant to compare the emulated climate to proxy data. The overall (global) comparison is made in Fig 5A, while more detailed patterns are evaluated in Figs. 7 and 8. Fig 6 adds a (very useful) comparison to the present day climate by HadCM3. This whole section is again confusing. First of all, during the last 80kyr HadCM3 seems to be at the upper end of the data set (Snyder) in Fig. 5A, while in the reconstructed time period (before 140ka BP) blue and grey lines seem to match rather well. I don't see how we can account for a potential model bias (from the last 80kyr) in the earlier (emulated) period? Moreover, in the other figures it is unclear what is compared to LOVECLIM and what to GCMET-LO or GCMET? Why is the LOVECLIM simulation necessary here, as you also have a low resolution HadCM3 version?"

Our response: The mismatch between HadCM3 output and the global mean temperature (Snyder, 2016) can be explained by two sources of error:

1. GCM model error/bias: this is due to the representation of the climate system and dynamics which is intrinsic to the model. We can't do anything about this type of error.
2. Errors in the boundary conditions: The HadCM3 boundary conditions, such as CO2 or ice sheet configuration, are not the latest, i.e., best estimates of boundary conditions. CO2, for example, has been kept as in the original study by

Singarayer and Valdes, 2010, which is based on the Vostok ice core record. In the meantime, the best CO2 reconstruction for the past comes from the EPICA Dome C ice core (Bereiter et al., 2015). Likewise, the ice sheet extent is based on ICE-5G (Peltier, 2004) which is now superseded by ICE-6G (Peltier et al., 2015). We can do something about this type of error. Because GCMET represents the response of HadCM3 to external forcings, we can use corrected versions thereof. In our 800ka reconstructions, we use CO2 from the EPICA Dome C record, and ice sheet reconstructions based on Ganopolski et al. (2011), ICE-6G covers only the last 122ka.

In general, the flexibility of an emulator approach allows to explore different boundary conditions with their associated uncertainty. Such an exploration is not feasible with state-of-the-art GCMs/EMICs.

The comparison with LOVECLIM is necessary because it is one of very few model results which extend that far back in time in a continuous way. The cross-comparison between GCMET-LO and LOVECLIM allows the reader see that the emulator possesses reasonable skill to model/reconstruct the climate. Again, it seems that most of the confusion is avoidable if we restrict our technical analysis to the aforementioned two aspects of i) temporal climate change signal and ii) the spatial detail which comes along with the dynamical downscaling. The basic idea for bringing those two aspects together is that we can provide the best paleo-climate dataset for the last 800ka, that is both reasonable in time and in space, for the wider paleo-modelling/application community.

ACTION: *We will make our intent clearer: that we want to deliver the best and most comprehensive climate reconstructions for the last 800ka. To reduce the confusion of why we compare different models and different resolutions, we would simply drop* *the comparison with LOVECLIM in the main text and move this cross-comparison part into the supplementary of the paper. For the comparison to global mean temperature (Fig. 5A), we will additionally provide the lower and upper confidence interval of our reconstructions to give the reader an idea about the emulator uncertainty with respect to global mean temperature. The flexibility of the emulator to quickly explore different boundary conditions will be discussed in the discussion section as one of the strengths of such an approach. We will argue that imperfect boundary conditions (such as older CO2 estimates) during the time of the GCM simulations can be corrected afterwards in a modified emulator re-run.*

> "Two more specific (but still general) remarks:
>
> Line 153-54: Some patterns of climate variability most likely don't show up in the lower resolution snapshots. However, these might be very sensitive to orbital forcing and not only depend on just CO2. I don't see how you get a reliable addition to the low-resolution version by just adding the high-low resolution difference based on CO2 concentration."

Our response: That is a good point. Unfortunately, we currently have only 9 high-resolution data points and with too few samples we can't use more than one regressor. We can try to use a second regressor (a rule of thumb is: more than 3*k sample points for k regressors, which means we could use 2 regressors), or alternatively define the first two principal components of all the forcings, CO2 and orbital parameters. This would require further analysis. However, because the spatial downscaling happens after the emulator reconstruction (provided at HadCM3-resolution), we could leave the spatial, dynamical downscaling part for another paper, once we've solved this issue sufficiently. This means that the results shown in Fig. 6C and D would be removed from the paper. Fig. 9 would be updated, depending on which high-resolution maps we generate from the bias-correction.

ACTION: *We will update our analysis based on a higher resolution mean state, i.e., an observational data set (such as ERA-20C, Poli et al, 2016). This dataset will also be used for the bias correction of our climate reconstructions which we want to make publicly available.*

> "Line 258: How important is it for the underlying snapshot simulations to be in 'quasi'- equilibrium? Can you estimate the error due to non-perfect equilibration of the training set?"

Our response: If the snapshot simulations are in (quasi-)equilibrium then we can assume that the climate system response that we capture with GCMET is the equilibrium response. However, if the climate simulations are not in equilibrium than it is possible that the transient response is non-linear. Unfortunately, there is no clear definition of "the climate model being in equilibrium". That means that we can't quantify how far the HadCM3 simulations are away from that equilibrium. Usually, GCMs tend to drift, specifically due to the slow dynamics of the deep ocean. It can take several thousands of simulation years to reach equilibrium, so we can only assume that HadCM3 is sufficiently equilibrated. Here is a quote from Singarayer and Valdes (2010) that summarizes this issue: *"Although not in perfect equilibrium (most GCMs are never truly in equilibrium), we judged that we were close enough for the models to be representative of the time period [...]. The 500 year length of integration was typical of many models used within the PMIP2 project"*. Without a fully transient GCM simulation over a sufficiently long period, from which we can infer the difference between transient and quasi-equilibrium snapshot, it is hard to make a judgment about the importance of this issue.

ACTION: *We will discuss this, as outlined above, in more depth in the discussion* *section of the paper.*

References

- Bereiter, B., Eggleston, S., Schmitt, J., Nehrbass-Ahles, C., Stocker, T. F., Fischer, H., Kipfstuhl, S., and Chappellaz, J.: Re- vision of the EPICA Dome C CO2 record from 800 to 600 kyr before present, Geophys. Res. Lett., 42, 2014GL061957, https://doi.org/10.1002/2014GL061957, 2015.
- Ganopolski, A. and Calov, R.: The role of orbital forcing, carbon dioxide and regolith in 100 kyr glacial cycles, Clim. Past, 7, 1415-1425, https://doi.org/10.5194/cp-7-1415-2011, 2011.
- Peltier, W. R. Global glacial isostasy and the surface of the ice age Earth: the ICE-5G (VM2) model and GRACE. Annu. Rev. Earth Planet. Sci. 32, 111-149 (2004)
- Peltier, W.R., Argus, D.F. and Drummond, R. (2015) Space geodesy constrains ice-age terminal deglaciation: The global ICE-6G_C (VM5a) model. J. Geophys. Res. Solid Earth, 120, 450-487, doi:10.1002/2014JB011176.
- 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.

- Snyder, C. W. (2016). Evolution of global temperature over the past two million years. Nature, 538(7624), 226.

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