Response to Anonymous Referee #2

This paper describes a novel method for ice sheet model forcing using a gaussian process emulator, presents new simulations using this method and performs sensitivity analyses. This is a useful contribution that seeks to overcome the limitations of ice sheet climate coupling when performing multi-million year simulations. It is a well thought out study and I recommend it for publication. I have some minor comments to improve the clarity of the article.

Author's response: We thank the reviewer for the positive evaluation of the manuscript and the suggestions to improve the clarity.

Minor Issues:
L106: I understand how this approach works with an atmosphere-only GCM, how does this differ for a slab-ocean model? If SSTs are prescribed what is the slab ocean doing? Unless I've misunderstood this, in which case clarification would be useful

Author's response: It is clarified in the manuscript why prescribed SST's are needed.

'…in order to calibrate the corrective heat fluxes from the slab ocean model. These corrective heat fluxes represent the seasonal deep water exchange and horizontal heat transport that is present in the real ocean. The oceanic heat fluxes are exchanged between the atmosphere and the slab ocean model in the mixed-layer, which is 50 m thick in our simulations. This way, realistic sea surface temperatures are simulated for the different climate model simulations.'

L106: I'm not suggesting doing this here, but for a true simulation of the EOT would two emulators be needed, one with late Eocene and one with early Oligocene SSTs?

Author's response: Since the simulated SST's are variable, this is not needed. The SST is only fixed in order to calibrate the corrective heat fluxes.

L118: How is ungrounded ice treated? I imagine there isn't much with the Wilson et al., topography that is being used, but this needs stating.

Author's response: Ice shelf formation is included and calculated using the Shallow Shelf Approximation. Ice shelves start to form when the grounding line reaches the coast and the influx of ice from the continent exceeds the ablation (surface ablation and basal melting). A constant basal melt rate of 1 m per year is used in all the simulations. This information is added to the manuscript.

Figure 7: To aid clarity of this figure can you sort the x-axes based on agreement? This would make it easier to compare the different methods.

Author's response: Sorting the x-axis would not be beneficial for the understanding of the figure. Each bar on the x-axis represents one experiment from the model design for the 100 experiments going from experiment 1 (xaemaa) to experiment 100 (xaemdv). When the x-axis would be sorted, it would be impossible to see which experiments perform worse than another.

Instead, because another emulator is introduced to increase the clarity of the manuscript and this would lead to 4 (number of emulators) x 3 (ways to calibrate the ice sheet parameter) subplots, we only show the leave-one-out performance for one emulator with the 3 different ways to calibrate the ice sheet parameter in the manuscript. The figures for the other calibrated emulators have moved to the supplementary information.

L250: Can the emulator be used to predict a spatially and temporally varying lapse rate?

Author's response: This could be done, but in these simulations, the spatially and temporally variable lapse rates are used as they were simulated with the prescribed climate model runs.

L307: Could be worth mentioning that another problem of the single ice sheet parameter (that is common to the matrix method) is that there is no guarantee that ice is growing in the same place in the ice sheet model as is prescribed in the climate model (Figure A2). I.e. a feedback from growing ice on the Antarctic Peninsula could be applied to the Transantarctic Mountains. Is there anyway of overcoming this? E.g. having a regional ice sheet parameter?

Author's response: We thank the reviewer for raising this interesting point. It is indeed true that the single ice sheet parameter does not define in which places the ice would grow. However, since the bedrock topography in the coupled ice sheet-climate simulations is the same as in the prescribed climate model runs, ice naturally starts to grow on the highest elevations and the pattern of ice sheet growth is definitely acceptable.

It is hard to overcome this issue and there are reasons why the implementation of a regional ice sheet parameter would not give the right results. Implementing a regional ice sheet parameter would come with introducing another variable that receives a certain value on whether ice is present in that region or not. That regional parameter is again strongly correlated to the ice sheet parameter and is not a true variable. For that reason, it might be hard to calibrate the emulator (in a similar way as it was hard to calibrate the emulator based on both ice volume and ice area).

The best way to overcome the problem of possible ice sheet growth in different regions than the prescribed input ice sheet geometries, is by implementing enough different input ice sheet geometries for the prescribed climate model runs in order that the climate captures the boundaries of many different ice sheets. We have added this information in the discussion section:

'A common problem for the emulator and the matrix look-up table method, where the ice sheet parameter is defined by a single number, is that there is no control on the regions where ice starts to grow. There are no obvious solutions for this problem. Adding an additional regional ice sheet parameter, that would receive a certain value depending on whether the region is ice-covered or not, would give rise to an ice sheet parameter that is defined in multiple ways. It has been shown that calibrating the emulator on two parameters (ice area and ice volume) that are strongly correlated and not independent, induces problems with the emulation accuracy. The best solution to overcome the lack of spatial control on ice sheet growth is to use enough prescribed ice sheet geometries in the model design, in order to give limited degrees of freedom to the emulator.'

L310: How much slower is the model with these different coupling timesteps?

Author's response: Halving the coupling time step increases the computational time with about 40 %. This information is added to the manuscript.

Figure 14: There are a lot of color-blind unfriendly colors. Here you could just show one of the emulators, rather than two.

Author's response: We assume that this remark was based on Figure 16. We changed the figure and only show the annual cycle of the temperature with respect to the January temperatures. The colour palette has also been changed to increase the figure attractiveness.

Figure 17: Can just show one of these subplots as the impact of the timestep already explored.

Author's response: Since the behaviour of the ice sheet evolution for the emulator including the estimates of variance is also different when using different coupling time steps, we opted for leaving both figures in the manuscript.

L487: This suggestion of using a direct mass balance calculation comes very late, it might be worth expanding on this point more or removing.

Author's response: We decided to remove the suggestion of emulating directly the mass balance for coupled ice sheet-climate simulations.

There are supplementary videos which are not available to review, can these be uploaded to GMD rather than zenodo?

Author's response: Yes. This will be done for the revised manuscript.

Typos, etc:
L9: "considering" to "using"

Author's response: Done.

L12: "from" to "in"

Author's response: Done.

L12: Sentence starting "The sensitivity…" is hard to follow. Suggest rewriting, or adding comma ", and to the coupling time".

Author's response: This phrase has been split into two parts as follows:

'The sensitivity of the evolution of the ice sheet over time is tested with respect to the number of predefined ice sheet geometries the emulator is calibrated on. Additionally, the model performance is evaluated to the formulation of the ice sheet parameter (being either ice sheet volume, either ice sheet area, or both) and to the coupling time.'

L42: "ran" to "run"

Author's response: Done.

L56: "paleoclimate" to "climate during the Pleistocene"

Author's response: Done.

L64: need reference for these CO2 changes.

Author's response: Two references are added that indicate the large CO_2 variations during the late Eocene to early Oligocene: Pagani et al., 2011 and Zhang et al., 2013.

L142: "amount" to "number"

Author's response: Done.

L180: (Eq 2.)

Author's response: Added.

L234: "Poorly"

Author's response: Done.

L235: "ran" to "run"

Author's response: Done.

L235: "warm-biased", "cold-biased"

Author's response: Done.

L236: remove "locally"

Author's response: Done.

L259: "The notion of an"

Author's response: This sentence has been rephrased.

'The ice sheet parameter is a single number representing the shape and area of the prescribed ice sheets. In previous studies (Araya-Melo et a;., 2015; Lord et al., 2017), it is defined as an integer, representing the number of different ice sheet geometries.'

L286: "with up to" to "by up to"

Author's response: Done.

L290: "mostly follows"

Author's response: Done

L425: "GP" define or change to "gaussian process"

Author's response: GP was defined the first time on line 131.