We thank the Editor and the three Reviewers for critical comments and suggestions, which helped to revise this manuscript constructively. We provide our answers to each question below.

Reviewer query: black

Answer: Blue

Added/revised text to the main manuscript: Blue, *italics*

//

Reviewer 3#

A review of "Process-based Estimate of Global-mean Sea-level Changes in the Common Era" by Nidheesh, Goosse, Parkes, Goelzer, Maussion, and Marzeion

The authors use models to quantify thermosteric effects from ocean heat storage and barystatic effects from land ice changes on long-term global-mean sea-level (GMSL) fluctuations during the preindustrial Common Era (PCE). They compare their results to proxy reconstructions of PCE GMSL changes from Kopp, Kemp, and Walker. One of the authors' main conclusions is that glaciers made dominant contributions to GMSL changes during the PCE.

I'm a sea-level scientist with training in physical oceanography. I don't have expertise in modeling land ice. Thus, I restrict my review mainly to sections on thermosteric effects, and recommend the Editor solicits reviews from experts in ice-sheet and glacier modeling.

I really liked this study. Something like it has been needed for a while. The past decade has seen real advances in the community's ability to quantify PCE GMSL changes from proxy reconstructions (from the likes of Kemp, Kopp, Walker, and others). But there's been a total lack of modeling studies to complement those observational studies. This paper starts to fill that gap: it won't be the last word on the subject, but it takes the logical first steps, and therefore deserves to be published after some minor revisions.

We thank the reviewer for recommending the study.

Specific comments

References to Walker et al. (2020) should be to Walker et al. (2021)

Thanks. It has been corrected.

Section 2.1. The authors describe how they estimate global-mean thermosteric sea level from PMIP3/CMIP5 temperature and salinity. Why not just use the zostoga global-mean thermosteric sea-level diagnostic variable made available by several PMIP3/CMIP5 groups? Also, the authors should identify precisely which PMIP3/CMIP5 model simulations they use (they only identify GISS-ES-R; were other models used?).

"ZOSTOGA" available for most of the PMIP3 models exhibits strong climate drift and we do not have a corresponding control run ZOSTOGA available for PMIP models to correct the drift. That is the prime reason we estimated the GMTSL using T, S simulation and restricted the computation to the upper 700 meters (we find that the drift is mostly associated with deep-layer temperature adjustments). In addition to that, understanding regional steric variations is also a perspective of this research and estimating thermosteric and halosteric fields from gridded temperature and salinity is required. Please note that the list of PMIP3/CMIP5 models used in this work is given in the supporting information which is mentioned in the main text (section 2.1; line 89).

The authors consider LOVECLIM, PMIP/CMIP models, and the reconstruction of Zanna et al. (2019). Is that all of the relevant data sources for ocean warming and thermosteric effects during the Common Era? Are there other ocean reconstructions that could also be brought in to corroborate the story they're telling?

Thanks for this question. This point has been clarified in the method section. Hence, we expanded section 2.4 as follows:

We also compare our model GMTSL with the reconstructed GMTSL estimates from Zanna et al. (2019) over 1870 – 2018. Since Zanna et al. (2019) already compared their reconstruction to different observation-based oceanic heat content estimates (e.g. Levitus et al. 2012; Ishii et al. 2017), we do not show all those available products in this paper for the twentieth century comparison. Reconstructions of ocean temperatures over the CE are limited to either sea surface temperature derived from paleoceanography (proxy) data (e.g. PAGES Ocean2k Synthesis Data; McGregor et al. 2015) or spatially averaged oceanic heat content estimates generated through inverse modelling and using available instrumental and paleo-data (Gebbie and Huybers 2019). Though such datasets have been shown useful to understand certain key features of ocean climate variability during the Common Era, they do not provide a direct estimate of the contribution of ocean changes to GMSL. Hence, we do not *attempt to compare our model thermosteric variability with any of those datasets in this paper. Also, as the GMSL reconstruction from Walker et al. (2022) and Kemp et al. (2018) already incorporated the tide-gauge-based twentieth-century GMSL, we do not show those available twentieth-century GMSL reconstructions in this paper.*

Section 2.5.2 Uncertainty on rest of the processes. I find this whole section unclear, ad hoc, and arbitrary. Can the authors please explain more the basic rationale and provide references for their methods when possible? In particular, I'm confused what their uncertainty quantification is supposed to represent. What missing process do they imagine they're accounting for by adding the autocorrelated noise, for example?

As we state in the beginning of section 2.5.2, because of the limited number of independent estimates of GMSL from each contributing process, it is hard to quantify (and show) the uncertainty in a consistent manner. Hence, the basic rationale of generating 1000 synthetic curves from existing / available curves by perturbing them using white noise is to have a consistent set of sea-level time series for each of the contributing processes considered. We have estimated the glacier uncertainty in an independent way (but still by producing thousand members; section 2.5.1). This kind of perturbation is not supposed to bring any additional specific process that misses in our modelling experiments but simply to acknowledge the remaining uncertainty (e.g. uncertainty arises from model initialization, inputs or differences in model physics) using a simple framework and to propagate the overall uncertainty to the final GMSL curve in a consistent way. One specific case, is the addition of the RMSE of thermosteric variability below 700 meters (derived from LOVECLIM) to the perturbation standard deviation of PMIP3 thermosteric curves (upper 700 meters) to specifically account for the uncertainty associated with missing variability below 700 meters (as noted in Line 246). We also agree that the entire procedure of uncertainty estimation might be a bit arbitrary, however, some choices have to be made and a similar approach could be seen in other studies (e.g. Frederikse et al. 2020).

Section 3.1

Line 266ff. Can the authors speculate on the high-frequency (decadal) global-mean thermosteric variability apparent in the LOVECLIM solution that is unrelated to volcanism? Is it related to ENSO or another global mode of natural climate variation (e.g., Hamlington et al., 2020, PNAS)?

We have not addressed specifically this point in our analysis which is focused on longer term changes. Our first guess is that such internal variability such as the one associated with ENSO (whose amplitude is very small in LOVECLIM) should be damped in the *ensemble mean* that we show in the paper. Those decadal signals could also be a response to other forcing, for instance, the small volcanos which are not discussed in the manuscript. We prefer to restrain ourselves commenting more on them without proper diagnosis.

Lines 283ff. Can the authors speculate on the mechanisms of these changes and when or why upper-ocean and deep-ocean effects may be opposing or reinforcing? More generally, some discussion of the physics involved, rather than just a tabulation of numbers, would be informative.

Thanks for this suggestion. We added the following sentences (Lines 283ff) to inform this:

The lag in the lower-layer thermosteric rise compared to recent warming of the upper ocean (~ since 1850 CE; Fig. 1b) could be due to the extending deep-layer cooling from LIA, as shown in Gebbie and Huybers (2019). Similarly, a rise in the upper 700 m thermosteric sea level during 1250 – 1400 CE might be a rebound of the upper ocean from volcanic cooling (strong 1257 eruption), but the deep ocean has still cooled during this period (Fig. 1b). In general, the differing response of thermosteric response in the upper and lower ocean indicates two distinct time scales of ocean response, the deep layer being much slower than the upper ocean.

Line 290. What is the plus/minus values?

As noted in section 2.5.2, we compute the rate, standard deviation and budget of GMSL *for each ensemble member* and subsequently derive the mean and confidence intervals from the large ensemble. Hence, ± indicate the 1-std of the mean standard deviation across members.

Line 302 and elsewhere. Why the italics on Roman Warm Period? Also on the next line it should be Antiquity not Antique.

Thanks. It is corrected.

Line 305ff. Is LOVECLIM model output available to say more about what drove these multi-centennial changes? Were they the long-term effects of volcanism? Changes in insolation? Again, some physical insights would be useful.

The main forcing in LOVECLIM is volcanic and the insolation has a weak impact on oceanic variability in annual mean (see for instance the supplementary document of McGregor et al. 2015).

Line 373ff. The more muted variability in the Walker et al. (2021) results relative to the Kemp et al. (2018) result may be owing to different prior assumptions made in the two studies with regard to dominant timescales of variability (i.e., what time smoothing is implied by the respective versions of the empirical spatiotemporal model).

As we note in the main text, there is an apparent difference between the GMSL curves of Kemp et al. (2018) and Walker et al. (2021) before ~ 600 CE, which may manifest as a unique centennial-scale variability in either of the reconstruction prior to ~ 600 CE. This difference comes because in Kemp et al. (2018), the GMSL during -100 – 100 CE is made equal to GMSL over 1600 – 1800 CE to avoid a spurious regional sea-level trend (arise from GIA) component. However, such a constraint is not employed in Walker et al. (2022) reconstruction. In addition to that, Walker et al. (2021) curve is smoother (less variability) compared to other reconstructions as the reviewer pointed out, and this could primarily due to the fact that Walker et al. (2021) employed the latest updated database of sea-level reconstructions in the model. The regional sea-level signal is expected to possess more lower-frequency variability and including data from more locations could potentially cancel out the variability in the global mean. We would like to restrain ourselves going deeper into features of sea-level reconstruction and focus on our model global-mean signal.

Line 393ff. Have the authors identified why their model results with respect to global-mean thermosteric versus barystatic contributions during the twentieth century differ so greatly from estimates of Slangen et al. (2017; Surveys in Geophysics) and Frederikse et al. (2020; Nature)? The authors point out the differences several times, but it would be good to know why these differences exist, and whether they bear on the confidence we have in their simulations of the PCE.

The underestimation of twentieth-century model GMSL comes from the barystatic components (see for instance the agreement between model and reconstructed twentieth-century thermosteric sea level in fig. 1). We have uncertainties on model initialization, reference climate state used, and forcing fields in the common era. If we consider glacier simulations, for instance, the models state in 1800 results from the simulation since 1CE and thus may integrates biases over this all period, in particular due to model drift and uncertainties in the forcing. We could 'correct' the state in 1800 to have better results over the last 2 centuries but, for example, the glacier distribution around ~1800 to initialize the model is not well-known and of the new model drift it will induce at the start of the simulation is hard to estimate. Such experiments would be interesting but our goal is to provide a consistent set up over the full millennium, with uncertainties clearly highlighted, not to have the most realistic set up for the twentieth-century (which is the aim of other existing studies (e.g. Marzeion et al. 2015; Frederikse et al. 2020). As we have noted in section 4.1: "*Uncertainties on ice sheet simulations are even larger and what we present here is a qualitative description of ice sheet changes in the common era based on Physics but it's quantitative assessments (for example the twentieth-century change) require further improvements (better constraining the climate forcing, developing paleo data etc.)*". We have added a few more sentences to highlight these aspects in the discussion part.