Interactive comment on "Glacial cycles simulation of the Antarctic Ice Sheet with PISM – Part 2: Parameter ensemble analysis" by Torsten Albrecht et al.

Response to Referee Lev Tarasov (lev@mun.ca)

(Received and published: 15 July 2019)

We thank Lev Tarasov for an excellent and detailed review and helpful comments. We learned a lot by working through his ideas and suggestions. Find in **blue** the referee's comments in in **black** the author's response.

This part II submission of Albrecht et al examines a moderately sized ensemble of Antarctic glacial cycle runs with the PIK variant of the PISM ice sheet model. The ensemble runs are scored against a range of paleo and present−day (PD) constraints. The scored ensemble uses a reasonably state of the art model for ensemble glacial cycle contexts and adds to the literature of what various models will do for past AIS glacial cycles. I therefore see the study potentially worthy of publication in TC given the current bar. At some point in the future I hope that model−based studies will have the requisite level of uncertainty quantification to enable much more meaningful inferences about past ice sheet evolution. However even with the current bar, a number of significant deficiencies need to be addressed.

The experimental design has some significant problems that are not even discussed. The study only using 4 ensemble parameters. Briggs et al, (TC, 2013) for instance, have 12 ensemble parameters just for the climate forcing and 31 ensemble parameter in total. At least 1 of the 5 temperature related ensemble parameters in that study (Tmix1) was one of the most sensitive ensemble parameters (with generally more sensitivity to this than to the precipitation related parameter (PdeselevEXP) that best corresponds to the sole climate forcing parameter (PREC) in this submission. The lack of an ensemble parameter relevant to the temperature forcing is especially problematic given the stated context of providing a distribution of present−day ice sheet states for initializing future projection runs. That state will depend significantly in the 3D temperature field of the ice sheet, the uncertainty of which is not probed in this study. Ideally this would be remedied, but that would be a major endeavor. At the very least I expect a clear and complete discussion of model and experimental design weaknesses and associated relevance to given results. A summary of this should also be in the conclusions.

Again, we thank the reviewer and are glad that he considers the study in principle worthy for publication in The Cryosphere (TC). We understand that only 4 selected model parameter cannot map the whole multidimensional phase space of model states and that other independent parameters might be relevant, too. However, given the limited granted computational budget and the minimum in simulation length and resolution (see Sect. 2.2 in companion paper part 1 (Albrecht et al., 2019a)) we were able to run up to around 500 simulations. It is a compromise, but as we decided to use simple averaging instead of advanced statistical emulators that interpolate parameter space (Chang et al., 2016a,b), we were restricted to 4-5 most relevant parameters, in order to privide reasonable results for the ensemble (see Chang et al. (2014) for Greenland application), while more than 30 parameters (also with Latin HyperCube, as in Briggs et al., 2013) would require many thousand simulations to be sufficiently spaced. Briggs et al. (2014) compensated for their „low sample size of model runs relative to the dimension of the parameter space“ with „some emphasis ... on sensitivity to the choice of ensemble sieves.“

Regarding the PREC ensemble parameter in our study, this in fact does not correspond to the desert elevation effect coefficient PdeselevEXP in Briggs et al. (2013), as it only scales with the external temperature forcing, not with changes in the surface geometry. It would be more similar to PphaseEXP, if it would not scale with insolation but with temperature.

We also tested for different temperature forcings in the companion paper part 1 (Albrecht et al., 2019a), and found comparably little influence of the present-day temperature distribution (Sect. 3.1: parameterized or from model output, with or without PDD, different PPD paramters) and for different temperature forcings (Sect. 4.2) on the Antarcrtic Ice Sheet history with less than 1m SLE sensitivity for LGM and about 2m SLE for PD results. In the Briggs et al., 2013 study, Tmix1 showed infact a high variance of 10 mESL, but this was related not only to a present-day temperature parameterization, but also included insolation and sea-level forcing.

We have added a clearer discussion of model and experimental design deficiencies and associated relevance to given results to the revised manuscript.

Another major omission is a comparison of the ensemble results against the paleo data constraints. The chosen normalization of all score components against median scores means that the scores do not convey any information about absolute model fit to paleo data, only relative fit. It is therefore incumbent that the complete set of ensemble fits to paleo constraints be explicitly shown and discussed, eg as done in Briggs et al, 2014.

We added plots of individual paleo score fits as in Fig. 7-10 in Briggs et al, (2014) and respective discussion to Appendix B.

Furthermore, there are a number of claims and statements (detailed below) made that I find are indefensible, misrepresentative, and/or incorrect.

 # Specific comments

Large Ensemble of 256
−> Ensemble of 256
with eg Briggs et al using a 2000 member ensemble, you can hardly call 256 a "Large Ensemble".

With the label „large ensemble“ we here directly referred to the „LE“ definition of ensembles that Pollard et al. (2016) defined (they used 625 ensemble members), „i.e., sets of hundreds of simulations over the last deglacial period with systematic variations of selected model parameters“. We provide a better quantitative classification of ensemble size in the revised manuscript:
„In view of the even larger ensemble by Briggs et al. (2014) with 31 varied parameters and

over 3,000 simulations, our ensemble with only four varied parameters and 256 simulations is of rather intermediate size, although we used a much finer model resolution."

"The model is calibrated against..."
−> The model is scored against...
scoring a moderately sized ensemble is not calibration

Ok, modified.

what is the SSA enhancement factor value? It is never explicitly given in this study nor in the PART I of this submission.

We agree that the SSA enhancement reference value of 0.6 is somewhat hidden in Sect. 2.3, Fig. 3 and Table. 1 of the companion paper (Albrecht et al., 2019a), where we find only little effect on LGM ice volume and almost no difference in deglacial or present-day ice volume, when values of 0.6 and 1.0 are compared. Our reference value agrees with the reference value in Briggs et al. (2014). For clarity we added a sentence to the parameter section 2.1:
„In all ensemble runs we used for the SSA stress balance an enhancement factor of 0.6 (see Sect. 2.3 in companion paper) which is more relevant for ice stream and ice shelf regions."

This analysis further constrains relevant model and boundary parameters by revealing clusters of best fit parameter combinations.
Isn't that already previously stated in different words:
"The model is calibrated[scored] against..."

We emphasize the new finding here by rephrasing:
„This analysis reveals clusters of best fit parameter combinations and hence a likely range of relevant model and boundary parameters, rather than individual best fit parameters."

Our Large Ensemble analysis also provides well−defined parametric uncertainty bounds and a probabilistic range of present−day states that can be used for PISM projections of future sea−level contributions from the Antarctic Ice Sheet.
Kind of meaningless. I can think up a dozen metrics that would provide "well−defined parametric uncertainty bounds", each with different resultant ranges.

We rephrased this sentence more generally as:
„Our ensemble analysis also provides an estimate of parametric uncertainty bounds for the present-day state that can be used for PISM projections of future sea-level contributions from the Antarctic Ice Sheet."

Nonconserving sub−glacial hydrology model
as from review of part I: pretty crude to call this a model −>
Here we use the non−conserving sub−glacial hydrology parametrization

We agree, the term „parameterization" would be more valid, but we actually refer to the non-conserving mode of the sub−glacial hydrology model, as cited in the previous sentence. We modified this in the manuscript as:
„We use the non-conserving mode of sub-glacial hydrology model, which balances basal

melt rate and constant drainage rate, to determine the effective pressure on the saturated till.“

Missing brief (eg 1 sentence) description of bed thermal model.

We have added more explanation to the part 1 companion paper and added also a sentence to the introduction of this study:
„Geothermal heat flux based on airborne magnetic data from Martos et al., 2017 *is applied to the lower boundary of a bedrock thermal layer of 2km thickness which accounts for storage effects of the upper lithosphere and hence estimates the heatflux at the ice-bedrock interface.“*

Sub−shelf melting in PISM is calculated via PICO (Reese et al., 2018) from salinity and temperature in the lower ocean layers on the continental shelf (Schmidtko et al., 2014) in 18 separate basins based on (Zwally et al., 2015) adjacent to the ice shelves around the Antarctic continent
the companion paper states that salinity was not varied:
"While salinity change over time in the deeper layers is neglected in this study"
and this should be made clear here.

This is correct and we are sorry for this misunderstanding. We referred here to the observations of mean salinity and temperature in the lower ocean layers on the continental shelf by Schmidtko et al. (2014), to define the reference ocean state, while PICO in our study responds to changes in external ocean temperature forcing. We rephrased this paragraph as:
„Sub-shelf melting in PISM is calculated via PICO (Reese et al., 2018) from observed salinity and temperature in the lower ocean layers on the continental shelf adjacent to the ice shelves around the Antarctic continent (Schmidtko et al., 2014) and as mean over 18 separate basins based on Zwally et al., 2015. PICO updates melt rates according to changes in ocean temperatures or the geometry of the ice shelves.“

use the Large Ensemble approach
Why is this capitalized? "the" makes no sense as there
are lots of large ensemble approaches. Furthermore, has already
stated above, this is not a large ensemble.

As mentioned above er here referred to the „LE“ definition in Pollard et al., 2016. We make this clearer in the revised manuscript and avoid capital letters.

This method yields as reasonable results for an adequately resolved parameter space as more advanced statistical techniques with means of interpolating results between sparsely separated points in multi−dimensional parameter space.
I would strongly dispute this since full−factorial sampling
restricts one to a relatively small number of ensemble parameters.
The cited Pollard et al (2016) paper used an ensemble of
WAIS only simulations for the last 30 kyr. Ie all ensemble
members had identical initial conditions and identical
time evolving ice boundary conditions at the junction with
the East Antarctic ice sheet. This is a far cry from applying
4 ensemble parameters to the whole AIS for 2 glacial cycles.

Pollard et al. (2016) only simulated WAIS, but Pollard et al. (2017) applied the same ensemble method to the whole Antarctic Ice Sheet, where also mantle viscosity profiles in a coupled GIA model have been varied. We want to emphasize that also our glacial cycle ensemble analysis has some focus on the last 30kyr and its impact on the present-day state, also as most paleo constraints are limited to this period. Hence, the comparison to the latest PSU-ISM model results are not too far off. We added a sentence:

„Yet, the full-factorial simple averaging method strongly limits the number of varied parameters for available computer resources such that only the most relevant parameters for each class of climatic and boundary conditions were pre-selected (in the companion paper) to cover a representative range of model responses.“

It covers uncertainties within the Earth model for values of
1e19, 5e19, 25e19 and 100e19 Pa s.
this study would benefit from better attention to relevant
litterature. While there is local support for viscosity as low as
1e19 Pa s on the Antarctic Peninsula, there is no support for even
5e19 over say the whole WAIS. Furthermore, the upper bound test
viscosity (and please use the more standard X10^21 units as
preferred in the GIA community) is half of the 95% "confidence"
upper bound of 2.0 X10^21 Pa s of Whitehouse et al, 2012 (GJI).

We thank the reviewer for this comment, as recent literature often give the impression that Antarctic upper mantle viscosities have been overestimated previously. We apologize for using the wrong units in this paragraph, the covered range is actually 0.1-10.0 x 10^21 Pa s for the upper mantle viscosity, and hence much closer to the 95%-confidence range of 0.8-2.0 x10^21 Pa s in Whitehouse et al. (2012b) or the spatially-average of 0.2-1.0 x10^21 Pa s beneath whole Antarctica in Whitehouse et al. (2018), as cited in the companion paper. Units have been adjusted throughout the manuscripts.

This compilation also includes records of regional sea−level change (RSL), which has not been considered in this study since most of the sea−level signal is a result of the sea−level forcing with up to 140m rather than the model’s ice dynamical response expressed in terms of sea−level equivalents, as PISM lacks a selfconsistent sea−level model
Since the RSL data for Antarctica is above present−day elevation,
the above statement as written is incorrect. The RSL data is the
signal, and dominance of a far−field sea−level forcing would result
in sealevel below present−day.

Yes, this was a badly formulated argument, we omitted the far-field sea level: *„This compilation also includes records of regional sea-level change above present-day elevation (RSL), which has not been considered in this study as PISM lacks a self-consistent sea-level model to account for regional self-gravitational effect of the order of up to several meters, which can be similar to the magnitude of post-glacial uplift.“*

We have actually tested for the addition of the RSL data type and found only little difference (slightly favoring higher viscosities) in the associated score pattern in parameter space among the ensemble members, when using the product of individual data type scores.

Mean−square−error misfit to observed grounding line location for the modeled Antarctic

grounded mask (ice rises excluded) using a signed distance field
I don't understand what you mean by signed distance as RMSE would
only care about unsigned distance. Or do you mean what we do in my
group: also track mean (ie not RMSE) error and use that to assign a
signed value to the RMSE?

„Signed distance“ is the name of a numerical technique for finding approximate solutions to the boundary value problems of the Eikonal equation using the fast marching method, here in two dimensions. The reviewer is right, that the differentiation between in and out (sign) is not considered in the mean square error in our study. We omit the term „signed“ to avoid confusion: *„Mean-square-error misfit to observed grounding line location for the modeled Antarctic grounded mask (ice rises excluded) using a two-dimensional distance field approximation (https://pythonhosted.org/scikit-fmm).“*

5. UPL: Mean−square−error model misfit to modern GPS−based uplift rates on rock outcrops at 35 individual sites using the compilation by Whitehouse et al. (2012b, Table 2) including individual observational uncertainty
Would be good to update the GPS data−set. Current GPS data versus
that approaching a decade old would make a significant difference in
observations and observational uncertainties.

We thank the reviewer for this suggestion and we certainly consent that open data compilations should be updated and joined to serve the whole community. There are several groups with expertise in GPS processing, but according to Pippa Whitehouse (personal communication) there is no recent publication that documents GPS rates across the whole Antarctic continent, except for Schumacher et al. (2018). However, there are many different choices, which requires expert input and should fill a seperate publication. Also, one would need to bear in mind that simulation results of a coupled solid Earth model would be associated with the viscous dynamics, while the GPS signal also implies the elastic signal due to contemporary surface mass changes.
For this study we preferred to use similar datasets as in previous publications (i.e. Pollard et al., 2016, 2017) in order to have a better comparison between the individual model responses. But even for relatively large uncertainties in the older data, the data type UPL shows strong variations in individual ensemble scores with impact on the aggregated score accordingly.

Then the individual score Si,j is normalized according to the median to
Why the mean versus the median?

We follow closely the definition in Pollard et al. (2016; Sect. 2.4.1), which does not mean that we support all choices they did. The algebraic mean can be inappropriate if the values range over many orders of magnitude. However, in the 9 used datatypes we find similar values for mean and median (except for TROUGH mean, which is 34% larger), such that the effect on the total score is negligible (see Fig. R1). We used median for consistency reasons with Pollard et al. (2016). Also the RSL data type shows large difference between median and mean value, but this data type has not been considered in our score analysis.

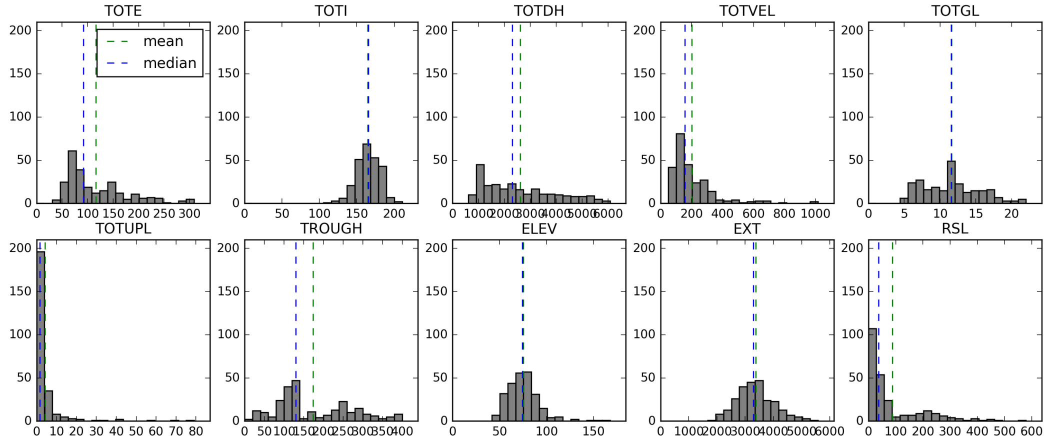

Fig. R1 Histogram of scores per data-type and median (in blue) and mean (green).

As in Pollard et al. (2016) we also assume that each data type is of equal importance to the overall score, avoiding the inter−data−type weighting used by Briggs and Tarasov (2013); Briggs et al. (2014), which would favor data types of higher spatio−temporal density
Would you still do this if say you only had 3 ELEV or EXT
datapoints? If all data were statistically independent, then one
would demand that data types of higher spatio−temporal density would
get more weight since in this case each and every datum should have
equal weight. You need to provide a better justification for this
choice then just blind citation of previous studies.

The reviewer is definitely right, that data with small spatio-temporal influence should weight less. In fact, we have tested for inter-data weighting, similar to Briggs and Tarasov (2013), Briggs et al. (2014) and found only small influence on the overall pattern of the score distribution.

For an interdata-weight of PD(5):TOTUPL:TROUGH:ELEV:EXT of 0.5:0.05:0.15:0.2:0.1 we find that the best 25 unweighted scores (above 0.1 in green in Fig. R2a) also corresponds to the best weighted scores (below 0.75). This means that more than 200 simulations yield a relatively bad score in both definitions. In fact, there are 18 simulations with a weighted score below 0.75, which are below 0.1 in the unweighted case (blue). This distribution does not change much, when RSL is added as dataype with a low interdata-weight of 0.03 (Fig. R2b). In fact, most of those simulations would show similar scores if equal weights were attributed. So this is an effect of product vs. sum of individual scores.

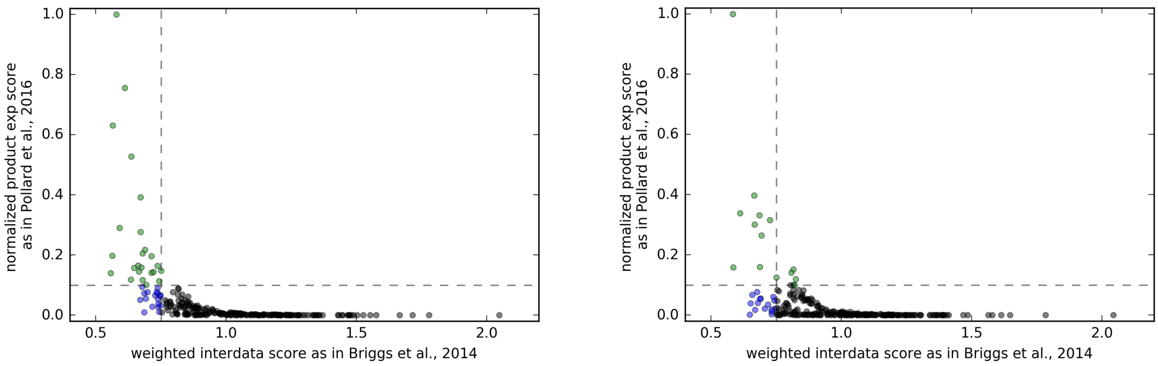

Fig. R2 Scatter plot of total scores in this study (best equals value 1.0) vs. an inter-data weighted score definition in Briggs & Tarasov (2013), with best value around 0.6 and larger scores meaning larger misfit. In the left-hand panel RSL datatype was added.

We will also added supplement plots for individual data types as in Briggs et al. (2014), Fig. 5-10, Fig. R4 for RSL, Fig. R5 for ELEV and Fig. R6 for EXT:

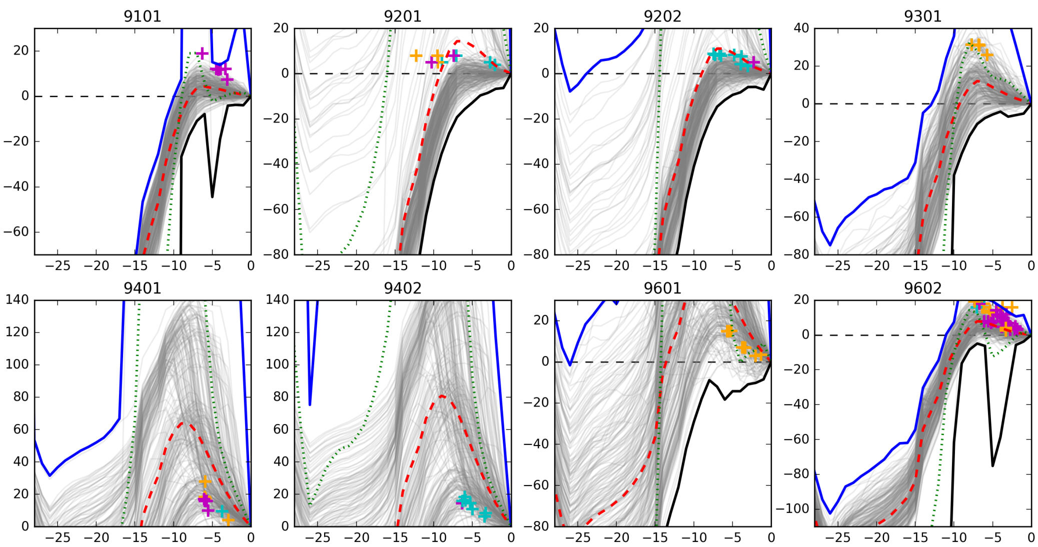

Fig. R3: RSL misfit as in Fig. 5-6 in Briggs et al., 2014, with same axis limits and data points (+ markers). Red dashed is mean of whole ensemble (no sieves applied), black is lower and blue upper relative sea-level history at site location, green dotted is best fit simulation. RSL was not used in the score analysis.

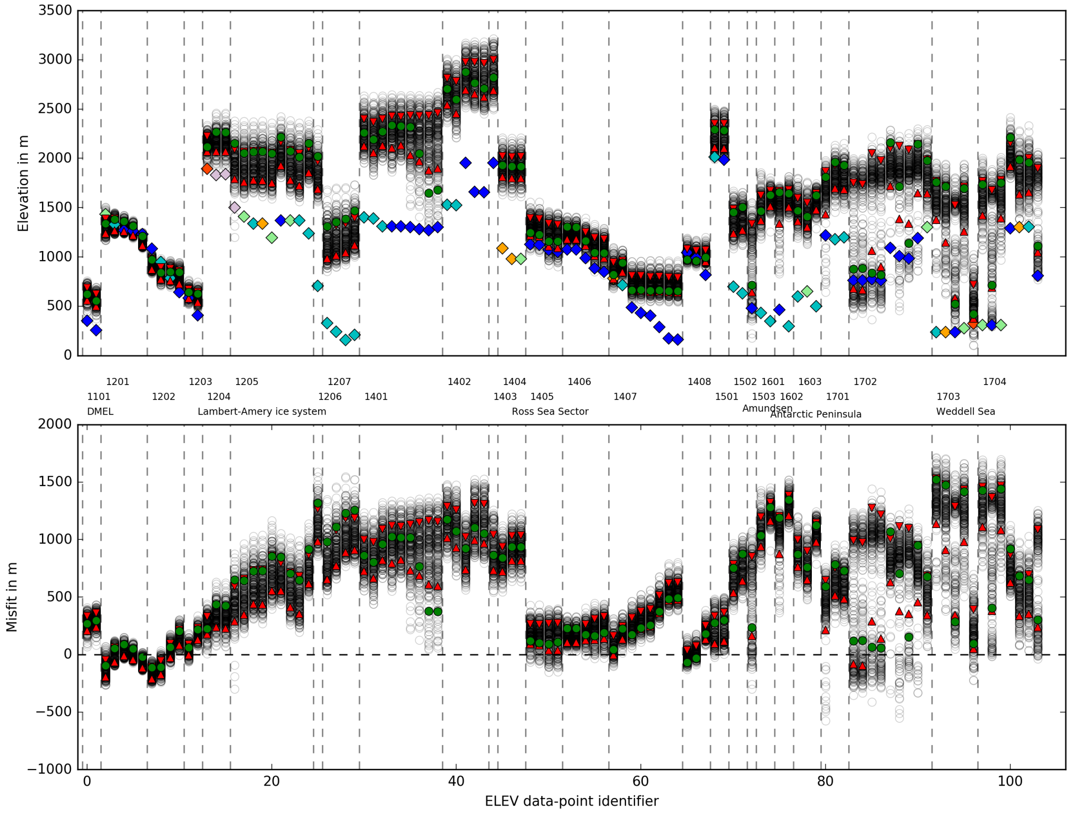

Fig. R4: ELEV misfit as in Fig. 7-9 in Briggs et al. (2014). Green is best fit simulation, blue is reconstruction at site location and date.

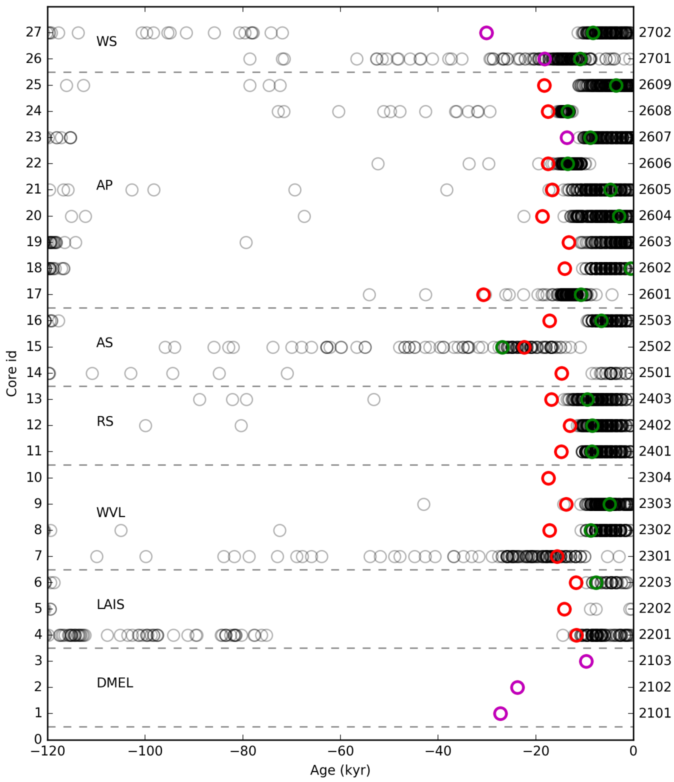

Fig. R5: EXT misfit as in Fig. 10 in Briggs et al. (2014). Green is best-fit simulation, red and magenta is reconstruction of grounding line extent.

The parameter ESIA enhances the shear−dominated ice flow and hence ice thickness
enhanced ice flow will not enhance (ie thicken) ice thickness but thin it

That was a mistake: *„The parameter ESIA enhances the shear-dominated ice flow and hence yields ice thickening particularly in the interior of the ice sheet and therewith the total ice volume.“*

For the upper range of mantle viscosities up to VISC = 1022 Pa s we find a normalized ensemble mean of 27% and 20%,
This contradicts what you previously indicated was an upper bound value of 100 X10^19 Pas on page 4. ???

This issued has been clarified above.

This value is also used in the GIA model ICE−6G (Peltier et al., 2015)
kind of irrelevant since Peltier doesn't do regional tuning of Viscosity profiles.

Has been omitted.

The best score ensemble members are found for intermediate mantle viscosities of VISC=5X10^20 Pa s and VISC=25X10^20 Pa s.
THis again contradicts the values given on page 4. Furthermore,
25X10^20 Pa s is a high viscosity for the upper mantle (upper mantle is what

this half−space model best corresponds to)

The unit question has been clairified above already. And yes, in the two-layer variant of the Lingle and Clark (1985) model, where the lower mantle represents one layer, without the low-viscosity channel beneath the lithosphere. As a half-space model this layer has indefinite thickness and one could think of the influence of the higher viscosity lower mantle, which is not explicitly considered here. We chose the parameter range large enough to find significant shifts in the scores, and therewith parameter values that can be excluded as a result of the analysis.

3.2 Reconstructed sea−level histories
−> ice volume histories or sea−level contribution histories

Yes, we preferred *„sea−level contribution histories"*.

the ensemble mean ice volume is 1.0m SLE below modern with a score−weighted standard deviation of around 2.7m SLE (volume of grounded ice above flotation in terms of global mean sea level equivalent as defined in Albrecht et al. (2019).
Should compare this to published (paleo data−based) inferences of the Eemian high−stand as it makes it hard to fit current proxy−derived estimates for the Eemian high−stand given constraints on what Greenland could have contributed.

This finding shows that the Antarctic Ice Sheet was somewhat smaller at Eemian. The indirect effect of Greenland melt is simply applied as sea-level forcing. Sutter et al., 2016 estimates around 3-4m SLE contribution of Antarctica during LIG, mainly trough WAIS collapse when a certain ocean temperature threshold is crossed. Also, the sea level high stand of the Eemian as a globally integrated signal suggests an Antarctic contribution of at least 1m ESL, and likley significant more (Cuffey and Marshall, 2000; Tarasov and Peltier, 2003; Kopp et al., 2009). This lower bound has been used as sieve criterion in Briggs et al. (2014). This additional information has been added to the manuscript.

The LGM ice volume increases for lower PPQ, lower PREC and lower ESIA values, while it seems to be rather insensitive to the choice of VISC
All these relations would be expected as such.

Added „*As expected, ...*", but we think it is good to remind the reader to this relations.

As MWP1a initiated the Antarctic Cold Reversal (ACR) with about two millenia of colder surface temperatures, a freshening of surface waters leading to a weakening of Southern Ocean overturning, resulting in reduced Antarctic Bottom Water formation, enhanced stratification and sea−ice expansion.
This is not a sentence.

This has been reformulated to:
„The MWP1a initiated the Antarctic Cold Reversal (ACR), a period lasting for about two millenia with colder surface temperatures. This cooling induced a freshening of surface waters and lead to a weakening of Southern Ocean overturning, resulting in reduced Antarctic Bottom Water formation, enhanced stratification and sea-ice expansion."

The modeled range between Last Glacial Maximum and present−day ice volume by Whitehouse et al. (2012a) is about 5.0X10^6 km3 (or 7.5 − 10.5m ESL, eustatic sea−level based on volume above flotation),...

There is no point in listing all the exact ranges here and then showing those ranges in fig 11
Add the conversion factors to the figure key and have the paragraph focus on g eneral comparison.

We have re-formulated the whole paragraph and updated Fig. 11 accordingly.

Briggs et al. (2014) ... from PSU simulations for 40 km resolution
To be accurate PSU + full visco−elastic isostatic adjustment bedrock
response with radially layered earth viscosity profile + different
subshelf melt, basal drag, climate forcing, and calving
treatments. So only PSU ice dynamics and thermodynamics.

We have added more information on varied model parameters, ensemble size, resolution, simulation length and used Earth model to the revised manuscript.

Although the Large Ensemble method is limited to a comparably small number of values for each parameter
no it is not. A full−factorial (grid) ensemble is limited. Not other
large ensemble approaches. And I don't understand why "Large
Ensemble" is capitalized. If you are choosing to equate a grid
ensemble as a "Large Ensemble", that makes no semantic sense. How are
readers supposed to differentiate between this useage and other
modelling studies that will use large ensembles under a different
sampling scheme? What about studies that have O(10) or more larger
ensembles?

This is a good point. We assumed „large ensemble" to be a label for a class of ensembles that cover the whole (chosen) parameter phase space in contrast to sensitivity studies, in which parameter are varied separateley. We omited the term „large" in our studies and reformulated the paragraph as:
„We have run an ensemble of 256 simulations over the last two glacial cycles and have applied a simple averaging method with full factorial sampling similar to Pollard et al. (2016). Although the this kind of ensemble method is limited to a comparably small number of values for each parameter..."

fig 11 plot and caption: there needs to be a note that some of the indicated studies have non symmetry distributions. Eg, Briggs et al, 2014 states : "likely between 5.6 and 14.3 m equivalent sea level (mESL), and with less confidence >10 mESL" = 4.0 X 10^6 km^3 of ice. The whitehouseBently12b datapoint is also problematic as that GJI paper never discusses ice volume or total sea level changes. The whitehouseBently12a does (and is cited in the preceeding discussion) and provides an uncertainty range but the plot only shows a single datapoint with no uncertainty range

OK, we added:
„The provided uncertainty ranges are not necessarily symmetric, e.g. the upper range in Briggs et al. (2014) has less confidence than the lower range." to the figure caption.

Regarding the datapoint in the Whitehouse et al. (2012a) study, we have contacted Pippa Whitehouse and she confirmed that the given range is the total range of simulated ice volumes rather than a standard deviation. The single plotted data point is the best fit simulations and located at the lower end of that range. We have added information on this in the revised manuscript.

fig 12 please increase the font size of the colour key for those of
use with ageing eyes...

We increased the fontsize in Fig. 12-14.

fig 14: please use a higher contrast colour scale, to make the
comparison more discernable. Even better would be the addition
of a difference plot to make clear what the differences are.

We actually tested different color schemes for Fig. 14, and we agree that spectral colormaps may better cover the full range of surface velocity over several orders of magnitude. However, we want to emphasize here that the general arterial pattern of ice streams reaching far into the inland ice sheet is reasonable well reproduced, and preferred this seqential colormap with of model and observations side by side. An anomaly or (root-square-error) plot can be somewhat misleading, as confined ice streams may be slightly shifted in location or ice shelf velocity. In fact, the velocity mismatch is part of the scoring scheme and it can help to identify regions of under- or overestimation, such that we added a difference plot as suggested by the reviewer and increased the contrast and the range of the colormaps.

fig 15−17 are hardly mentioned in the main text, with no consideration
of details. Eg, figure 17 has 7 timeseries, not one of which is individually
refered to in the text. So why is this figure in the paper?

Those figures are made for comparison with a previous study on MWP-1A (Golledge et al., 2014) and have been referenced only once in the text. We added many more information on the distinct deglacial and regrowth phases with figure details to the manuscript.

3.4 Comparison to previous large ensemble study
−> Comparison to Pollard et al. (2016) ensemble study
Your current title implies there was only one large ensemble
and the subsequent text implies it Pollard et al. (2016).
Briggs et al was a much larger ensemble study...

We have drawn a better concerted picture in the revised manuscript.

while other parameters that affect the modern grounded ice areas are sufficiently constrained by earlier studies
This is not a fact, at best state they claim this.

Changed.

In their ensemble analysis Pollard et al. (2016) included an iceberg calving par
ameter. Our ’eigencalving’ model provides a fair representation of calving front dynamics independent of the climate conditions (Levermann et al., 2012).
What does "fair" mean? Be precise.

Changed to: *„Our ‘eigencalving’ parameterization provides a representation of calving front dynamics, which in first order yields present-day calving front positions (Levermann et al.,* *2012). This paramterization is rather independent of the climate conditions, variations of the 'eigencalving' parameter show only little effect on sea-level relevant ice volume (see companion paper Albrecht et al. (2019a)).“*

As we used the PICO model (Reese et al., 2018) that includes physics to adequately represent melting and refreezing also for colder than−present climates, we have chosen other parameters in our ensemble,
What about the large uncertainty is subshelf ocean temperature? Above you invo ke this
to explain the lack of any MWP1a signal compared to eg Golledge et al, 2014

We have actually considered the effect of intermediate ocean warming during ACR on the PICO sub-shelf melt rates in Sect 5.2 in the companion paper. This reference and some more details have been added to the revised manuscript. The Golledge et al. (2014) study used a rather crude estimate of sub-shelf melt rates from a scaling between LGM and modern states, which yields extremely high melt rates of up to 100m/yr for present climate, without considering the overturning circulation in the ice shelf cavity nor boundary effects between ice and ocean.

comparably small mantle viscosity around VISC = 5−25X10^20 Pa s,
small compared to what? I wouldn't call these small upper Mantle viscosities f or Antarctica

We were actually talking about the lower tested range and omitted „comparably small“.

Due to the comparably coarse resolution and the high uncertainty that comes with the strong non−linearity (sensitivity) of the system we here discuss rather general patterns of ice sheet histories than exact numbers.
This non−linearity is another reason to increase the number of ensemble parame ters.

We added: *„..., which would require a larger ensemble with an extended number of varied parameters.“*

Our ensemble−mean lies at the upper range of most previous studies, except for the large ensemble study by Pollard et al. (2017) with only 3−8m SLE since LGM, as their score algorithm favored the more rigid and hence thinner ice sheet configurations.
incorrect in that half of your stated range is below the favoured
range of Briggs et al (2014) who state
"The LGM ice volume excess relative to present−day is likely between 5.6 and 14. 3 m
equivalent sea level (mESL), and with less confidence >10 mESL
Furthermore, your sentence contradicts itself as currently written.

Right, this has been formulated rather crudely. Converted into total ice volume, our score-weighted range of 5.8+-2.0 mio. km3 relative to present overlaps with the less confidence upper range (>4.0 mio. km3) of Briggs et al. (2014) of with 2.2-5.7 mio km3, while there is almost no overlap to the range found by Pollard et al. (2017) of 3.4+-0.7 mio. km3 relative to present (the 3-8m are associated with the approximate range of best fit ensemble members in their Fig. 2, as discussed with Dave Pollard). Their range is quite consistent

with their previous study on WAIS only with 3.2+-1.6 mio. km3 (Pollard et al., 2016). Also Golledge et al., (2012, 2013) are below our range with 2.7 and 3.4 mio km3 respectively, while in contrast the value of 5.8 mio km3 in Golledge et al. (2014) (relative to Bedmap2) is close to our ensemble mean.
We rephrased this paragraph (without the numbers) accordingly: *„Our ensemble-mean ice volume anomaly between LGM and present is close to the best fit value found in Golledge et al. (2014) and Whitehouse et al. (2012a), while the ANICE best fit values (Maris et al., 2014, 2015) lie in the lower uncertainty range of our study. In contrast, the large PSU-ISM ensemble mean by Pollard et al. (2016, 2017) as well as the high confidence (lower) range in Briggs and Tarasov (2013) are found below the uncertainty range of our study. Also the PISM equilibrium values by (Golledge et al., 2012, 2013) are clearly below the uncertainty range of our ensemble.“*

Previous studies with PISM Golledge et al. (2014) suggest
english and punctuation...

„A previous PISM study suggest that the oceanic forcing at intermediate levels can be of opposite sign as compare to the surface forcing, as likely happened during the two millennia of Antarctic Cold Reversal following the MWP1a, causing earlier and larger sea-level contributions from Antarctica (Golledge et al., 2014).“

In this study we used the
Bedmap2 topography remapped to 16 km resolution without local adjustments
Does this actually belong in the "Conclusions"?

This sentence is an explains the different rebound behaviour with respect to the previous study, we switched the order to: *„In contrast to Kingslake et al. (2018), we used the remapped topography without local adjustments, such that in only about 20% of the score-weighted simulations this region re-grounded.“*

provides model and observation calibrated parameter constraints for
projections of Antarctic sea−level contributions
awkward and somewhat indecipherable. Do you just mean
"data−contrained projections of .. using PISM"?

We are sorry for using the terms „calibrated“ and „constrained“ as synonyms, we corrected for this misunderstanding troughout the manuscript.

With the best−fit simulation parameters we have participated in the initMIP−Antarctica model intercomparison (Seroussi et al., 2019, PISMPAL3).
This is not a conclusion

Ok, has been moved to results section and caption of Table 1.

appendix A : is referred two a couple of times, but without any
statement of what the takeaway from the appendix is.

„One key parameter for the onset of retreat could be the minimal till friction angle on the continental shelf with values possibly below 1.0. More discussion of the interference of

basal parameters in terms of an additional ensemble analysis is given in Appendix A.“

„Friction underneath the modern ice shelves is highly relevant, in particular during the deglaciation, as we have discussed in the companion paper Albrecht et al. (2019a). However, instead of choosing the friction coefficient underneath the modern ice shelves as ensemble parameter (Pollard et al., 2016, 2017) we decided on the sliding exponent as uncertain parameter for the entire Antarctic Ice Sheet. In fact we have run an additional ensemble analysis for four basal sliding and hydrology parameters only, including friction underneath modern ice shelves and discussed the results in Appendix A. As the main deglacial retreat (in the basal ensemble mean and in the best fit simulation therein) occurs a few thousand years earlier (closer to MWP-1A) the corresponding scores are even better than for the best fit simulation of the base ensemble (for same sliding exponent but smaller minimal till friction angle)“

In the basal sub−ensemble we find even better scores than for the best fit param
eter combination in the large ensemble (here no. 8102, see Fig. A.
...
However, best fit to the nine constraints are found for
the basal ensemble..
which agrees with the best fit values of large ensemble.
The above two statements contradict each other

Sorry for this ambiguity, we were actually talking about the best fit parameter values and not the scores:
„However, best fit to the nine data constraints are found for the basal ensemble in the middle range of PPQ = 0.5–0.75 and the lower range of till water decay rates of TWDR = 0.5-1 mm/yr (1.55–3.1×10−11 m/s), which agrees with the best fit parameter combination of the base ensemble (PPQ=0.75 and TWDR=1 mm/yr). „

References

Albrecht, T., Winkelmann, R., and Levermann, A.: Glacial cycles simulation of the Antarctic Ice Sheet with PISM – Part 1: Boundary conditions and climatic forcing, The Cryosphere Discuss., https://doi.org/10.5194/tc-2019-71, in review, 2019a.

Briggs, R., Pollard, D., & Tarasov, L. (2013). A glacial systems model configured for large ensemble analysis of Antarctic deglaciation. *The Cryosphere*, *7*(6), 1949-1970.

Briggs, R. D., & Tarasov, L. (2013). How to evaluate model-derived deglaciation chronologies: a case study using Antarctica. *Quaternary Science Reviews*, *63*, 109-127.

Briggs, R. D., Pollard, D., & Tarasov, L. (2014). A data-constrained large ensemble analysis of Antarctic evolution since the Eemian. *Quaternary Science Reviews*, *103*, 91-115.

Chang, W., Applegate, P. J., Haran, M., & Keller, K. (2014). Probabilistic calibration of a Greenland Ice Sheet model using spatially resolved synthetic observations: toward projections of ice mass loss with uncertainties. *Geoscientific Model Development*, *7*(5), 1933-1943.

Chang, W., Haran, M., Applegate, P., & Pollard, D. (2016). Calibrating an ice sheet model using high-dimensional binary spatial data. *Journal of the American Statistical Association*, *111*(513), 57-72.

Chang, W., Haran, M., Applegate, P., & Pollard, D. (2016). Improving ice sheet model calibration using paleoclimate and modern data. *The Annals of Applied Statistics*, *10*(4), 2274-2302.

Cuffey, K. M., & Marshall, S. J. (2000). Substantial contribution to sea-level rise during the last interglacial from the Greenland ice sheet. *Nature*, *404*(6778), 591.

Golledge, N. R., Fogwill, C. J., Mackintosh, A. N., & Buckley, K. M. (2012). Dynamics of the last glacial maximum Antarctic ice-sheet and its response to ocean forcing. *Proceedings of the National Academy of Sciences*, *109*(40), 16052-16056.

Golledge, Nicholas R., Richard H. Levy, Robert M. McKay, Christopher J. Fogwill, Duanne A. White, Alastair GC Graham, James A. Smith et al. "Glaciology and geological signature of the Last Glacial Maximum Antarctic ice sheet." *Quaternary Science Reviews* 78 (2013): 225-247.

Golledge, N. R., Menviel, L., Carter, L., Fogwill, C. J., England, M. H., Cortese, G., & Levy, R. H. (2014). Antarctic contribution to meltwater pulse 1A from reduced Southern Ocean overturning. *Nature Communications*, *5*, 5107.

Kopp, R. E., Simons, F. J., Mitrovica, J. X., Maloof, A. C., & Oppenheimer, M. (2009). Probabilistic assessment of sea level during the last interglacial stage. *Nature*, *462*(7275), 863.

Levermann, A., Albrecht, T., Winkelmann, R., Martin, M. A., Haseloff, M., & Joughin, I. (2012). Kinematic first-order calving law implies potential for abrupt ice-shelf retreat. *The Cryosphere*, *6*, 273-286.

Lingle, C. S., & Clark, J. A. (1985). A numerical model of interactions between a marine ice sheet and the solid earth: Application to a West Antarctic ice stream. *Journal of Geophysical Research: Oceans*, *90*(C1), 1100-1114.

Martos, Y. M., Catalán, M., Jordan, T. A., Golynsky, A., Golynsky, D., Eagles, G., & Vaughan, D. G. (2017). Heat flux distribution of Antarctica unveiled. *Geophysical Research Letters*, *44*(22), 11-417.

Pollard, D., Chang, W., Haran, M., Applegate, P., and DeConto, R.: Large ensemble modeling of the last deglacial retreat of the West Antarctic Ice Sheet: comparison of simple and advanced statistical techniques, Geosci. Model Dev., 9, 1697–1723, doi:10.5194/gmd-9-1697-2016, 2016.

Pollard, D., Gomez, N., & Deconto, R. M. (2017). Variations of the Antarctic Ice Sheet in a Coupled Ice Sheet-Earth-Sea Level Model: Sensitivity to Viscoelastic Earth Properties. *Journal of Geophysical Research: Earth Surface*, *122*(11), 2124-2138.

Schmidtko, S., Heywood, K. J., Thompson, A. F., & Aoki, S. (2014). Multidecadal warming of Antarctic waters. *Science*, *346*(6214), 1227-1231.

Schumacher, M., King, M. A., Rougier, J., Sha, Z., Khan, S. A., & Bamber, J. L. (2018). A new global GPS data set for testing and improving modelled GIA uplift rates. *Geophysical Journal International*, *214*(3), 2164-2176.

Sutter, J., Gierz, P., Grosfeld, K., Thoma, M., & Lohmann, G. (2016). Ocean temperature thresholds for Last Interglacial West Antarctic Ice Sheet collapse. *Geophysical Research Letters*, *43*(6), 2675-2682.

Tarasov, L., & Peltier, W. R. (2003). Greenland glacial history, borehole constraints, and Eemian extent. *Journal of Geophysical Research: Solid Earth*, *108*(B3).

Whitehouse, P. L., Bentley, M. J., & Le Brocq, A. M. (2012a). A deglacial model for Antarctica: geological constraints and glaciological modelling as a basis for a new model of Antarctic glacial isostatic adjustment. *Quaternary Science Reviews*, *32*, 1-24.

Whitehouse, P. L., Bentley, M. J., Milne, G. A., King, M. A., & Thomas, I. D. (2012b). A new glacial isostatic adjustment model for Antarctica: calibrated and tested using observations of relative sea-level change and present-day uplift rates. *Geophysical Journal International*, *190*(3), 1464-1482.

Whitehouse, P. L. (2018). Glacial isostatic adjustment modelling: historical perspectives, recent advances, and future directions. *Earth surface dynamics.*, *6*(2), 401-429.