Response to reviewers on “Glacial cycles simulation of the Antarctic Ice Sheet with PISM – Part 2: Parameter ensemble analysis” by Torsten Albrecht et al.

We would like to thank the anonymous reviewer and Lev Tarasov for the very constructive criticism regarding our manuscript. These reviews have considerably improved the manuscript for which we are grateful. We were able to address all requests.

In order to facilitate the reading of this document, the referee's comments are given in blue and in **black** the author's response.

Response to Anonymous Referee #1
(Received and published: 26 June 2019)

1 Overall assessment

This study presents a large ensemble modelling of the Antarctic ice sheet over the last two glacial cycles with the PISM ice-sheet model. The ensemble reveals clusters of best fit parameters that are evaluated against a series of constraints related to the present-day ice sheet and glacio-geological evidence. Results of the best fit(s) reveal the deglaciation history of the Antarctic ice sheet (in line with previous results reported in Kingslake et al., 2018) and show the major ice loss after MWP-1A.

My first concern is the choice of the sensitivity parameters in the ensemble, which is limited to four factors: ice softness (ESIA), sliding plasticity (PPQ), precipation scaling (PREC) and mantle viscosity (VISC). Similar studies also explore the sensitivity of sub-shelf melting and how it relates to changes in far-field/continental shelf ocean and salinity in terms of oceanic forcing. Especially with respect to the explanation of MWP-1A, ocean forcing and its relation to sub-shelf melt may have played a crucial role (Golledge et al., 2014). Similar sensitivities have been explored in Pollard et al. (2016). Why are such parameters not taken into account, both sensitivity parameters within PICO, but also sensitivity to forcing, i.e. relation between atmospheric/ocean temperature forcing, for instance? As I understand from the paper, ocean temperature/salinity changes in the far field are not considered, neither through an offline ocean model, nor a parameterization that links atmospheric temperature change to oceanic temperature change. This is extremely important, as the conclusions with respect to the deglaciation do not take into account this sensitivity, hence show a large deglaciation pulse significantly later than the occurrence of MWP-1A. Many studies have shown the importance of the ocean in the dynamics of the Antarctic ice sheet, but neither the sensitivity (of PICO) or any ocean forcing has been investigated.

We thank the reviewer for pointing out important aspects of the parameter choice and implications for the last deglaciation. Some of these questions are actually touched in the first part of the study (Albrecht et al., 2019a), which certainly need to be better referenced in this second part. Find our detailed comments below.

A related question is why choosing those four parameters (ESIA, PPQ, PREC and VISC) and not others? Have other studies or previous experience shown that these are the most sensitive/critical? Some explanation should be given.

The choice of the four ensemble parameter is motivated in the companion study (Albrecht et al., 2019a), in which different parameter choices and boundary conditions are compared to a reference model ice volume history to gain some „prior model experience“ and to determine most relevant parameters (for this metric) in each of the different model components (climatic forcing, basal sliding, ice creep and bedrock response). We agree, that this parameter choice is somewhat biased to the modeled total ice volume at LGM and present-day state, while other parameters may be more relevant for the onset and pace of deglaciation. We have added a paragraph on deficiencies of the study. Yet, enhancement factors, sliding coefficient and viscous relaxation time of the bedrock have been also typically varied in previous model studies (e.g., Pollard and DeConto, 2012; Maris et al., 2014; Quiquet et al., 2018). As PISM uses a more generalized basal sliding and bedrock deformation scheme, we have selected different uncertain parameters.

Regarding the reviewer's concern on the sub-shelf melt sensitivity and the MWP-1A, we can state that Pollard et al. (2016) was focussing mainly on ice-oceanic deglacial processes in the WAIS with other relevant parameters fixed, while we consider a broader range of sea-level relevant processes over a longer time scale, such as ice-internal and ice-atmospheric effects, covering both parts of the Antarctic Ice Sheet. Golledge et al. (2014) used an apparently more realistic ocean forcing (from an Earth System Model), but they state „that there is considerable uncertainty in the relationship between ocean temperature and ice-shelf melt“. In fact, much of the oceanic uncertainty of previous models is considerably reduced in our PISM simulations as it uses the PICO module (Reese et al., 2018), in which two uncertain parameters have been constrained by observed melt rates. Of course, we do consider ocean temperature changes, in our case coupled to surface temperature forcing (see Sect. 4.3 in Albrecht et al., 2019a). However, this relationship cannot account for events such as the Antarctic Cold Reversal after MWP-1A, when surface and intermediate water temperatures became rather decoupled. Yet, we have tested our PISM-PICO model for an earlier warming signal in the deeper ocean layers (while the surface was warming at the same time) in the companion paper (see Sect. 5.2 in Albrecht et al. 2019a), which can cause earlier retreat, while we still do not find main deglaciation before MWP-1A. For these reasons we have selected PREC as climate uncertainty instead of an ocean-melt (or calving) related parameter, as it can potentially counteract the other more constrained climatic forcings (see Sect. 4.5 in Albrecht et al. 2019a), and thsi aspect may have been underestimated by previous model studies. We have added some more discussion on the limited parameter choice and consequences for the results in the revised manuscript.

A second concern is about the novelty of the study, that methodologically is heavily relying on Pollard et al. (2016) and is basically performing the same analysis. However, a clear rationale on the choice of the boundaries for the parameter changes is lacking. Moreover, as shown in Figure 4, clear clusterings in misfit show up and best fit results are generally found in a much smaller range of parameter values (basically the range of two values for each parameter. Therefore, it seems to me that a smaller range sub-sampling would lead to an improved fit, hence reduce the uncertainties of the whole ensemble.

Yes, we have been using very similar analysis (and visualization) tools as discussed in Pollard et al. (2016) to allow for better comparison. However, Pollard et al. (2016) used different parameterizations in their model and focussed mainly on ice-oceanic processes in the WAIS over the last 20kyr. We have hence chosen different parameters and parameter boundaries as motivated in the (first part) companion paper (Albrecht et al., 2019a).
A refined analysis could likely provide better constrained (best fit) parameter ranges. But

the high uncertainty in the sea-level history is in fact a result of multiple best-score parameter ensemble members which show quite different sea-level histories.

Also, the best-fit parameters of our paleo study might be shifted slightly for higher spatial resolution, e.g. when performing short-term projections. In this analysis we wanted to consider the broader range of parameter values, covering the wide parameter range used also in other models (implying a rather coarse sampling) to gain a better understanding of (combined) parameter effects in the highly nonlinear model. This also serves as rough constraint for further ensemble simulations and projections with PISM.

2 Specific remarks

Line 153: in -> to

Thanks.

Line 180 and following: All scores are aggregated into one score, thereby giving them an equal weight. However, some constraints are more reliable than others. Would different weighing lead to different results? Is there a certain bias towards one or several parameters; in other words, what is the result if scores would be calculated separately? Does this lead to the same clustering? Which scores are more representative?

This is definitely true, the score aggregation hides lots of information on the individual data types. However, we did not use inter-data-type weighting, e.g. based on spatial and temporal volumes of influence of each data type, as done in previous studies (Briggs and Tarasov, 2013; Briggs et al., 2014). Here, we followed the arguments in Pollard et al. (2016), assuming that „each data type is of equal importance to the overall score, and that if any one individual score is very bad (Si ≈ 0), the overall score S should also be ≈ 0... if any single data type is completely mismatched, the run should be rejected as unrealistic, regardless of the fit to the other data types... The fits to past data, even if more uncertain and sparser than modern, seem equally important to the goal of obtaining the best calibration for future applications with very large departures from modern conditions“. We will refer more clearly to these argumentation in the revised manuscript.
We have also added Supplementary Material with plots of individual paleo data misfits analogous to Briggs et al. (2014). If using the inder-data-type weighting and defining the score as weighted sum as in Briggs et al. (2014), the resultant distribution of best scores (here the smallest values) actually turns out to be very similar as for the product of individual scores, as shown in Fig. R1.

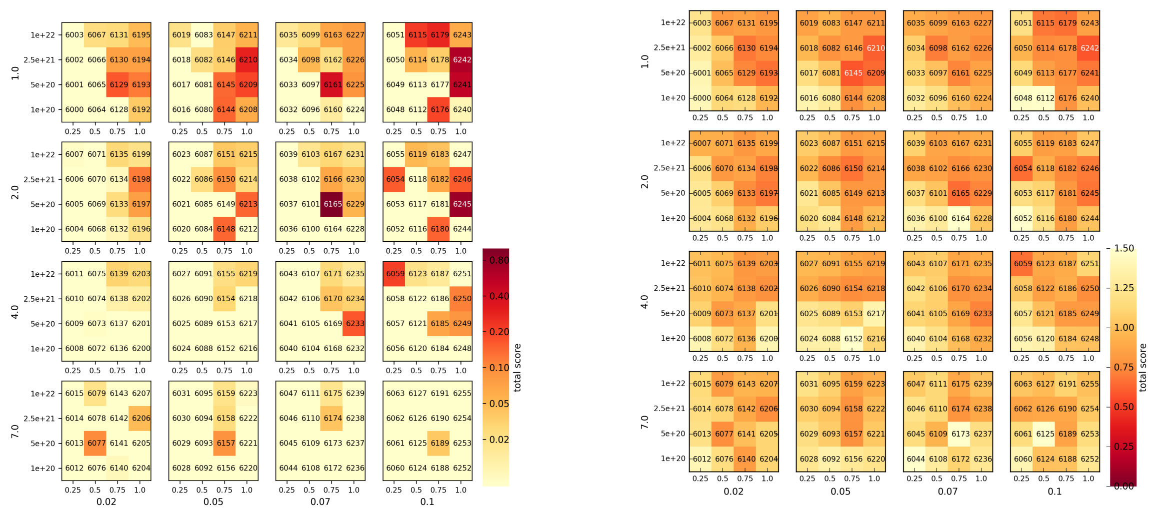

Fig. R1: Aggregated scores as product of individual scores as in Pollard et al. (2016) and used in our study (best fit equals 1, log color scale), compared to the aggregated score as a result of a inter-data-type weighted sum, as in Briggs et al. (2014), with best fits for lowest scores.

In fact, we can learn more about the model's response when discussing statistics on individual scores. Some of these information can be estimated from Fig. 2 or Fig. 5 and are discussed rather qualitatively in corresponding sections. We added ensemble standard deviation (Table R1) for each data type and some more discussion on the stasticial aspects to the revised manuscript. Find also Fig. R2 and Fig. R3 for comparison (analogous to the ones in the manuscript, but separated for individual data-type scores).

	TOT	TOTE	TOTI	TOTDH	TOTVEL	TOTGL	TOTUPL	TROUGH	ELEV	EXT
MAD	0.002	0.123	0.023	0.183	0.144	0.099	**0.292**	0.156	0.049	0.047
SD	0.082	0.156	0.035	0.190	0.179	0.126	**0.300**	0.204	0.075	0.072

Table R1: Medan absolute deviation (MAD) and standard deviation (SD) for each data-type score, SD values are used in the revised manuscript.

In the manuscript (Sect. 3.1) we have discussed for each ensemble parameter how best scores are related to individual data types, as shown in Fig. R2. We want to avoid additional figures in the manuscript, but we added a general comment to the Supplementary Material B:

„The corresponding variability of each of the resultant normalized scores hence contribute different skills to the aggregated score (see Table 2). Generally, grounding-line related (TOTE, TOTGL, THROUGH) and ice volume-related data-types (TOTDH) show similar individual score patterns (not shown here) with ensemble standard deviations of 0.1-0.2. In the aggregated score this patterns becomes even more pronounced, while paleo scores (ELEV and EXT) and ice shelf extent (TOTI) show only little variation (<0.1) among the ensemble, and hence only little effect in the aggregate score pattern.“

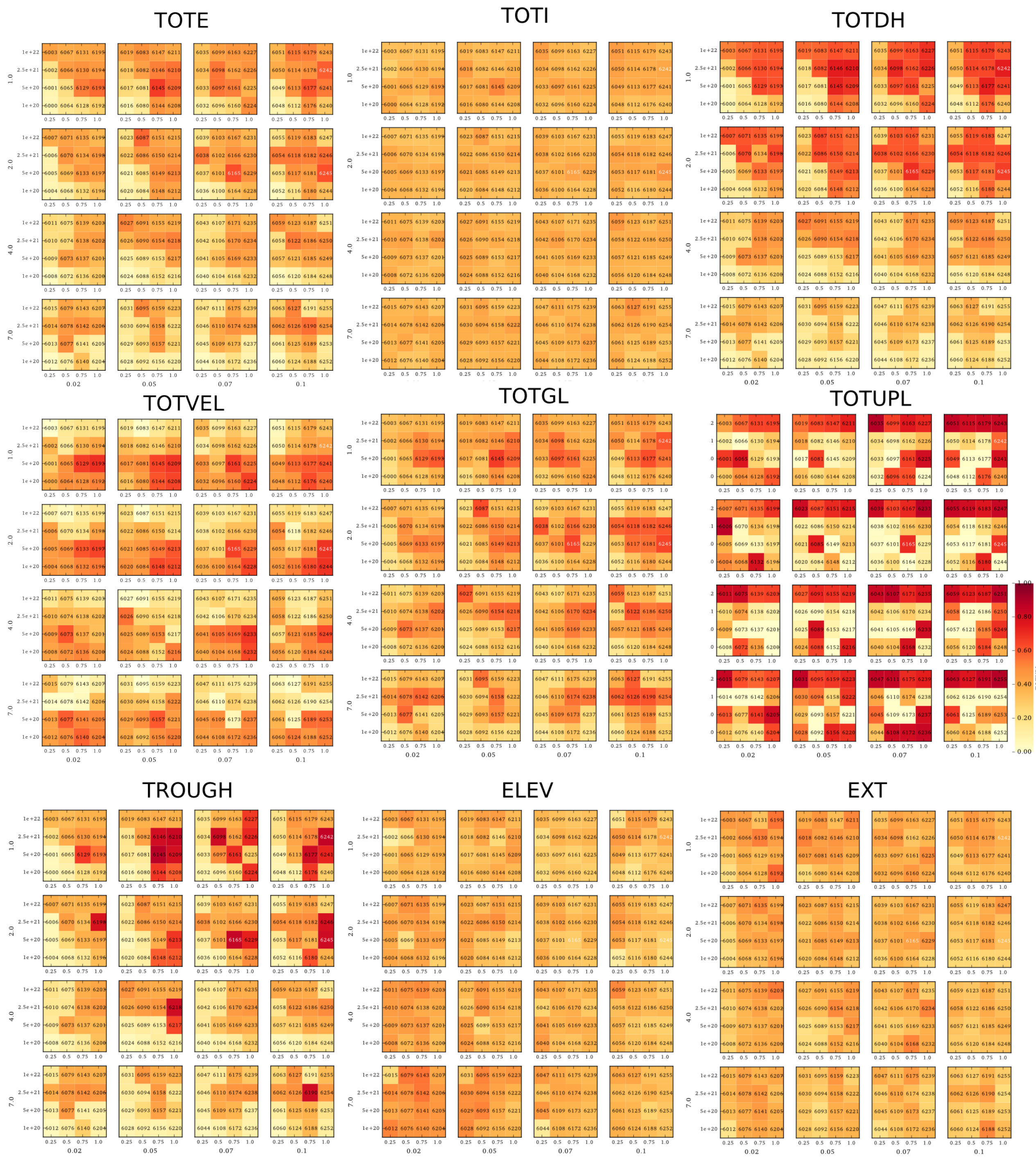

Fig. R2: Individual data-type scores for all 256 ensemble members, as in Fig. 1 in the manuscript, but with linear color scale. Scores in individual data-types are normalized by median.

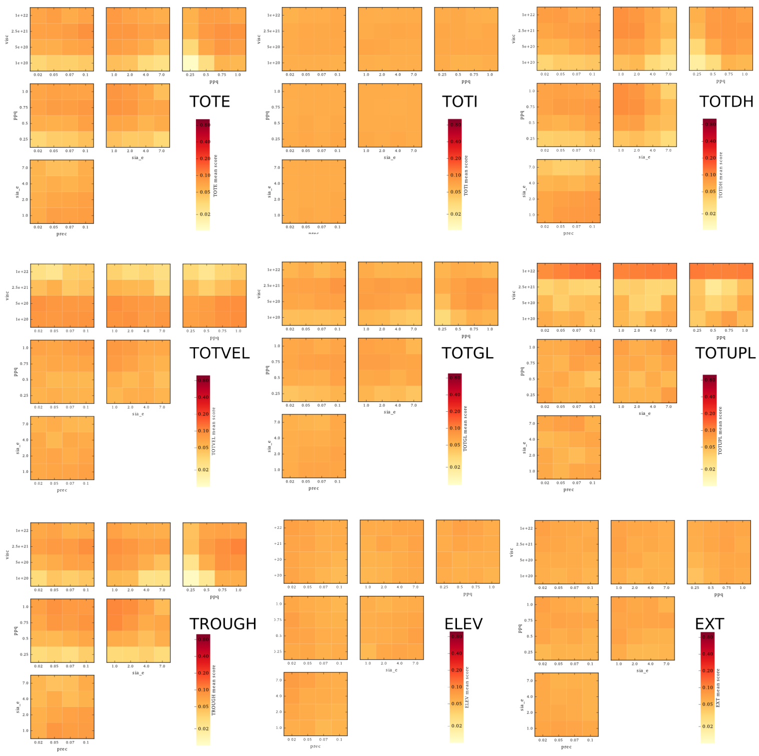

Fig. R3: Individual data-type mean scores for six possible pairs of parameter values, as in Fig. 4 in the manuscript.

The score pattern is also shaped by the uplift-related individual score (TOTUPL), that shows the highest ensemble standard deviation of 0.3 (Table R1) with a clear tendency towards higher VISC values, (see Fig. R3) probably a result of lower sensitivity to fluctuations in grounding line location. In contrast, the velocity-related individiual score (TOTVEL) with ensemble standard deviations of 0.2, favors lower VISC values, probably a result of more advanced grounding line location, which implies lower ice shelf velocities and hence lower chance for misfit (see Fig. R3). In the product formulation of the aggregated score such a reverse pattern can lead to highest total values for intermediate parameter values for VISC. We have discussed this aspect in the revised manuscript:

„As mantle viscosity determines the rate of response of the bed to changes in ice thickness a low viscosity corresponds to a rather quick uplift after grounding line retreat and hence to a retarded retreat, which corresponds to a rather extended present-day state. This implies smaller ice shelves with slower flow and less velocity misfit, such that also TOTVEL favors small VISC values. In contrast, a trend to rather high mantle viscosities in the aggregated score stems mainly from the misfit of present-day uplift rates expressed as data-type score TOTUPL, probably due to reduced sensitivity to fluctuations in grounding line location. High mantle viscosities involve a slow bed uplift and grounding line retreat can occur faster. More specifically, in the partially over-deepened ice shelf basins, which have been additionally depressed at the Last Glacial Maximum by a couple of hundred meters as compared to present, grounding line retreat can amplify itself in terms of a regional Marine Ice Sheet Instability (Mercer, 1978; Schoof, 2007; Bart et al., 2016). In fact, the best score ensemble members are found for intermediate mantle

viscosities of VISC =0.5×10^{21} Pa s, and VISC =2.5×10^{21} Pa s. This could be a result of the product formulation of the aggregated score, in which individual data types scores favor opposing extreme values."

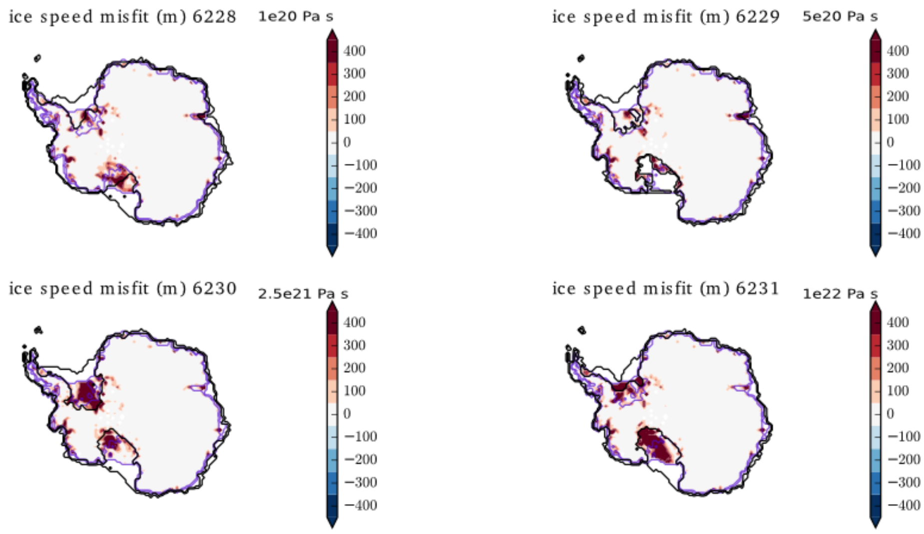

Fig. R4: Map of misfit of modeled modern surface velocity (related to TOTVEL) in four ensemble members with different VISC values indicated in labels (but otherwise identical parameters).

Line 256: intermediate values of mantle viscosity give the best results. However, these are values for the whole Antarctic continent and several studies show that there is a distinct contrast in mantle viscosity between WAIS and EAIS. Would this not explain the best score (mean of both extremes)?

This is a good question. Recent literature suggest comparably small values for the oceanic WAIS plate. As most of the ice volume and grounding line changes occur in WAIS, one would suggest that this regions also leaves the strongest imprint on the individual-scores, that are related to the VISC parameter.

As already mentioned above, in our ensemble it is the TOTVEL data type which favors lower values, likely related to the grounding line location and ice shelf extent (see also Fig. R4), while TOTUPL actually favors large VISC values, which might actually be related to better scores for the EAIS part, where lower bedrock sensitivity and lower measurement uncertainty leads to lowest misfits (see Fig. R5).

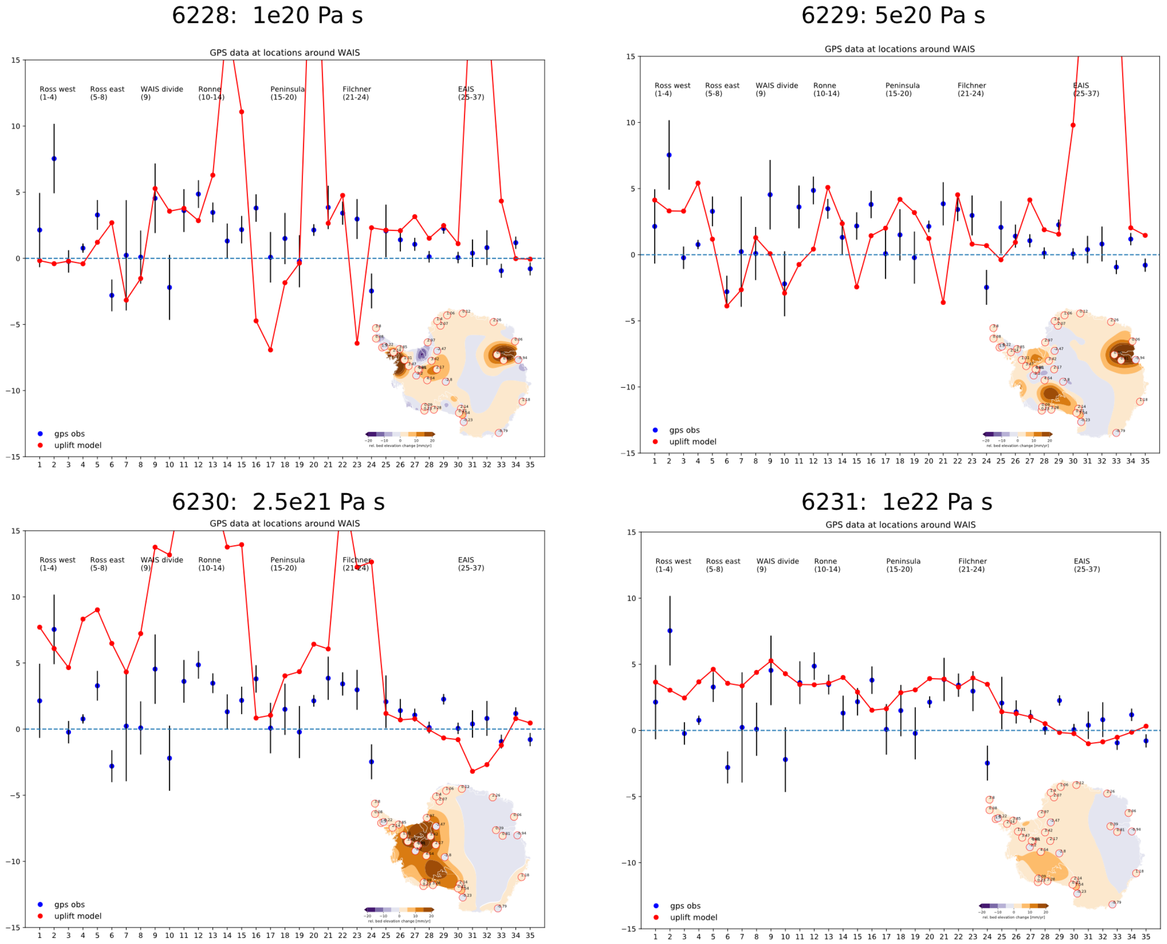

Fig. R5: Misfit of modeled present-day bedrock change rates to GPS measurements (related to TOTUPL) around the Antarctic continent for four different VISC values. Insets show location of PGS sites and map of bedrock change.

Line 278-79: why high basal friction? The power of the friction law only determines how sliding scales with τ_b .

Thanks for pointing out this imprecise formulation. Basal shear stress τ_b balances the driving stress within the SSA stress balance. As in the PISM friction law u_0 is considered as reference velocity (Eq. 2), such that for $q>0$ slower flowing upstream regions experience reduced basal shear stress, while fast flowing regions downstream are subjected to increased basal shear stress. Thus, reducing q from 0.75 to 0.25 produces slower flow in the interior and faster ice stream flow. We omitted this confusing aspect in the paragraph in the revised manuscript:

„In about 10% of the score-weighted simulations grounding line remains at the extended position without significant retreat, linked to ~~high basal friction (PPQ=0.25) and~~ an efficient negative feedback on grounding line motion related to a fast responding bed (low VISC).“

Figure 4: see general remarks: clustering demonstrates that the sampling range is too large and can be refined.

As already stated above we intended to cover a broad range of parameter values typically (and plausibly) used in other ice sheet modeling studies for better comparison (ESIA, PPQ, VISC) and to gain a better understanding of the actual (combined) effects of parameters on the ice sheet dynamics. For follow-up projections (with higher resolution) a similar score scheme may be used, but for different (more recent) data-types in terms of hindcasting, with more refined parameter ranges.

Line 334: sub-surface melt: ambiguous, could point to melt occurring just below the surface. Using a term as sub-shelf melt is more appropriate.

Thanks, has been changed accordingly in the revised manuscript.

Line 378: remove 'with' and add year of communication.

Changed to „(personal communication Dave Pollard, 2017)".

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, 2019.

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.

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.

Maris, M., De Boer, B., Ligtenberg, S., Crucifix, M., Van De Berg, W. J., & Oerlemans, J. (2014). Modelling the evolution of the Antarctic ice sheet since the last interglacial. *The Cryosphere*, *8*(4), 1347-1360.

Pollard, D., and R. M. DeConto. "Description of a hybrid ice sheet-shelf model, and application to Antarctica." *Geoscientific Model Development* 5.5 (2012): 1273-1295.

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.

Quiquet, A., Dumas, C., Ritz, C., Peyaud, V., & Roche, D. M. (2018). The GRISLI ice sheet model (version 2.0): calibration and validation for multi-millennial changes of the Antarctic ice sheet. *Geoscientific Model Development*, *11*(12), 5003.

Reese, R., Albrecht, T., Mengel, M., Asay-Davis, X., & Winkelmann, R. (2018). Antarctic sub-shelf melt rates via PICO. *The Cryosphere*, *12*(6), 1969-1985.

Response to Referee #2: 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.

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 the Supplementary Material 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.

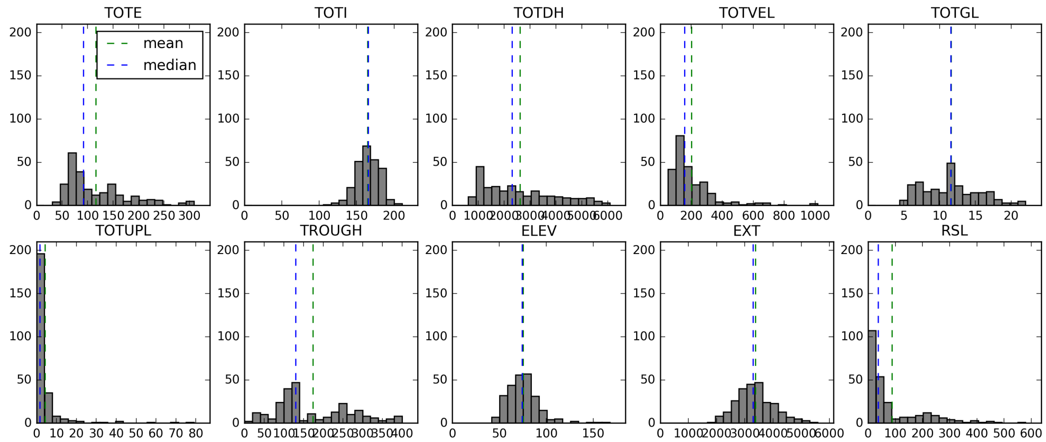

Fig. R1 Histogram of scores per data-type and median (in blue) and mean (green).

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.

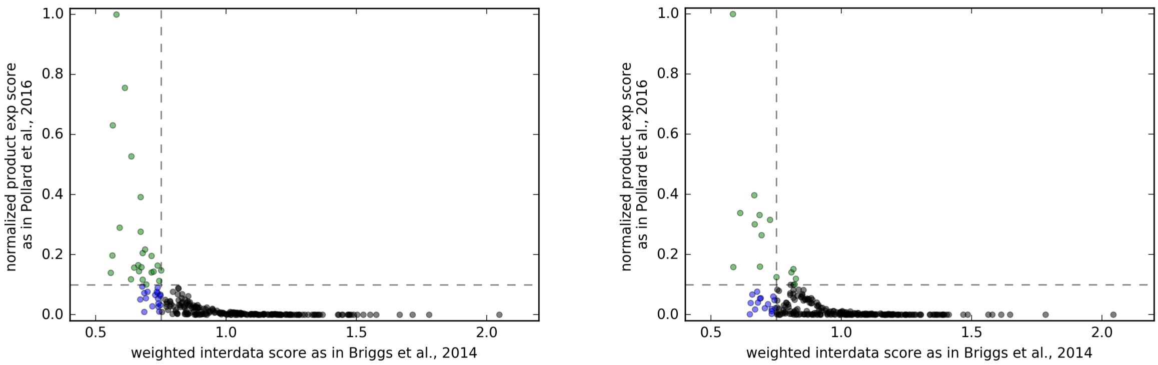

Fig. R2 Scatter plot of total scores in this study (best equals value 1.0) vs. an inter-data weighted score definition in Briggs & Tarasov (2013), with best value around 0.6 and larger scores meaning larger misfit. In the left-hand panel RSL datatype was added.

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:

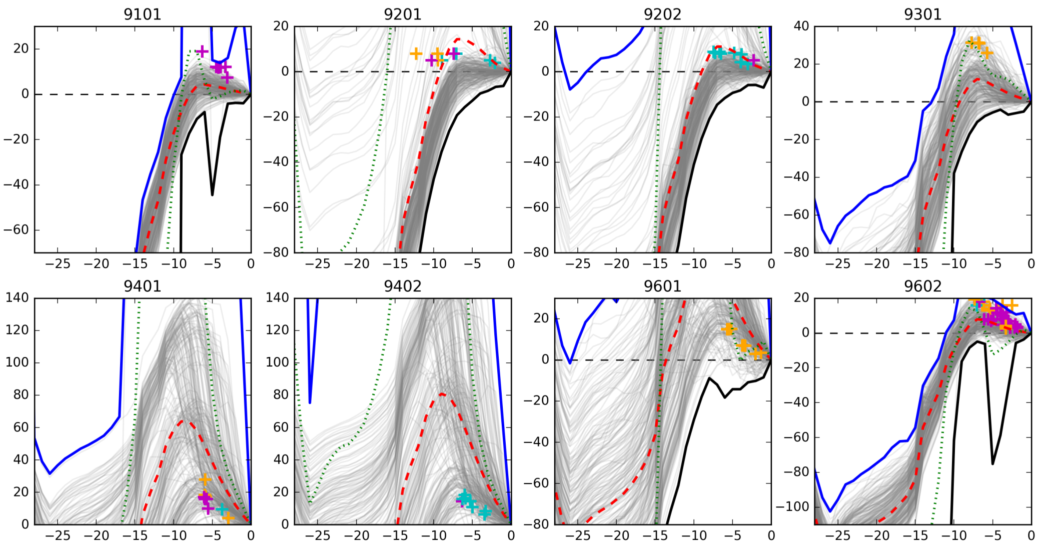

Fig. R3: RSL misfit as in Fig. 5-6 in Briggs et al., 2014, with same axis limits and data points (+ markers). Red dashed is mean of whole ensemble (no sieves applied), black is lower and blue upper relative sea-level history at site location, green dotted is best fit simulation. RSL was not used in the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

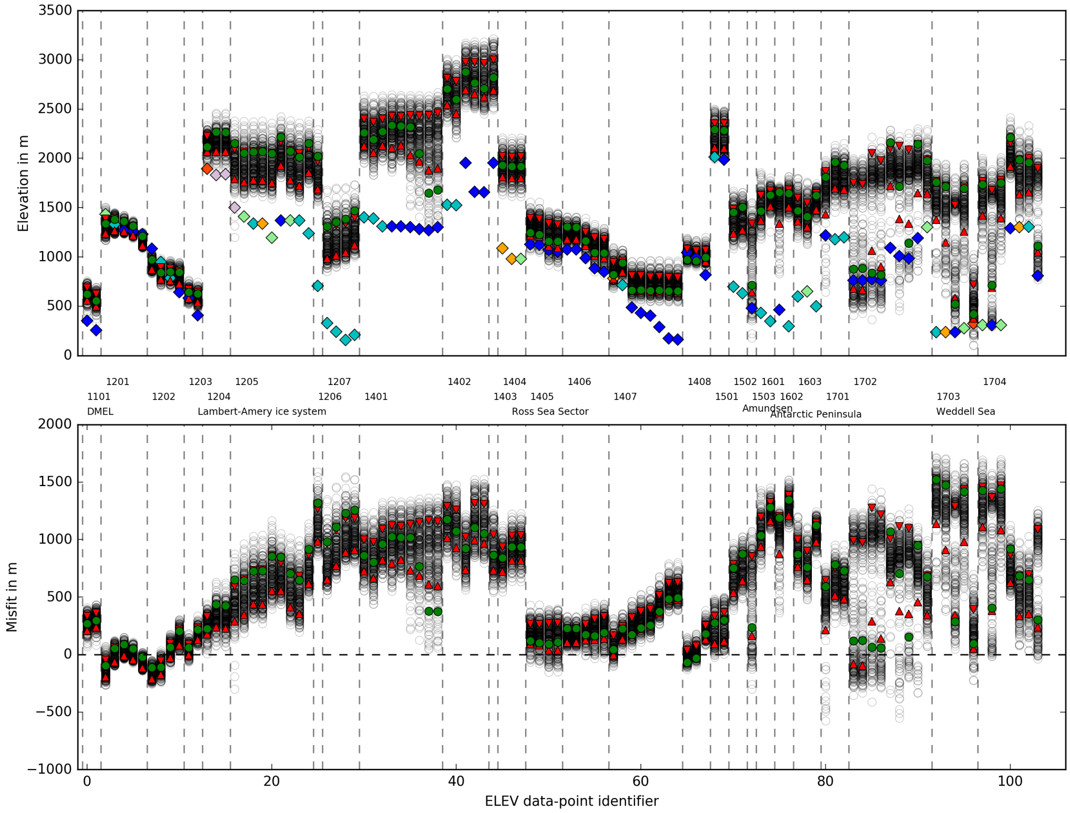

Fig. R4: ELEV misfit as in Fig. 7-9 in Briggs et al. (2014). Green is best fit simulation, blue is reconstruction at site location and date.

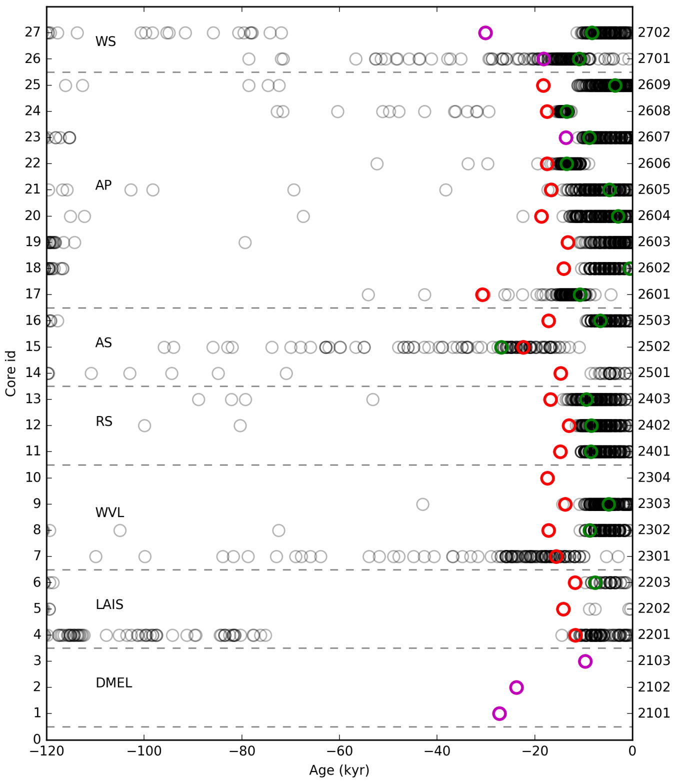

Fig. R5: EXT misfit as in Fig. 10 in Briggs et al. (2014). Green is best-fit simulation, red and magenta is reconstruction of grounding line extent.

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 Supplementary Material 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 the Supplementary Material 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.

Manuscript prepared for The Cryosphere
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 7 November 2019

Glacial cycles simulation of the Antarctic Ice Sheet with PISM - Part 2: Parameter ensemble analysis

Torsten Albrecht [1], Ricarda Winkelmann [1,2], and Anders Levermann [1,2,3]

[1]Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, Potsdam, Germany
[2]Institute of Physics and Astronomy, University of Potsdam, Potsdam, Germany
[3]Lamont-Doherty Earth Observatory, Columbia University, New York, USA

Correspondence to: T. Albrecht (albrecht@pik-potsdam.de)

Abstract. The Parallel Ice Sheet Model (PISM) is applied to the Antarctic Ice Sheet over the last two glacial cycles (≈ 210,000 years) with a resolution of 16 km. ~~A Large Ensemble~~ An ensemble of 256 model runs is analyzed in which four relevant model parameters have been systematically varied using full-factorial parameter sampling. Parameters and plausible parameter ranges have been identified in a ~~companion paper~~ companion paper (Albrecht et al., 2019) and are associated with ice dynamics, climatic forcing, basal sliding and bed deformation and represent distinct classes of model uncertainties. The model is ~~calibrated~~ scored against both modern and geologic data, including reconstructed grounding line locations, elevation-age data, ice thickness and surface velocities as well as uplift rates. An aggregated score is computed for each ensemble member that measures the overall model-data misfit, including measurement uncertainty in terms of a Gaussian error model (Briggs and Tarasov, 2013). The statistical method used to analyze the ensemble simulation results follows closely the simple averaging method described in Pollard et al. (2016).

This analysis ~~further constrains~~ reveals clusters of best fit parameter combinations and hence a likely range of relevant model and boundary parameters~~by revealing clusters of best fit parameter combinations~~, rather than individual best fit parameters. The ensemble of reconstructed histories of Antarctic Ice Sheet volumes provides a score-weighted likely range of sea-level contributions since the Last Glacial Maximum of 9.4 ± 4.1 m (or 6.5 ± 2.0 $\times 10^6$ km^3), which is at the upper range of most previous studies. The last deglaciation occurs in all ensemble simulations after around 12,000 years before present, and hence after the Meltwater Pulse-1A. Our ~~Large Ensemble~~ ensemble analysis also provides ~~well-defined~~ an estimate of parametric uncertainty bounds ~~and a probabilistic range of~~ for the present-day ~~states~~ state that can be used for PISM projections of future sea-level contributions from the Antarctic Ice Sheet.

1 Introduction

Sea-level estimates involve high uncertainty in particular with regard to the potential instability of marine-based parts of the Antarctic Ice Sheet (e.g., Weertman, 1974; Mercer, 1978; Slangen et al., 2017). Processed-based models provide the tools to evaluate the currently observed ice sheet changes (Shepherd et al., 2018a, b), to better distinguish between natural drift/variability and , variability or anthropogenic drivers (Jenkins et al., 2018) and to estimate future changes for possible climatic boundary conditions (Oppenheimer and Alley, 2016; Shepherd and Nowicki, 2017; Pattyn, 2018). Regarding the involved variety of uncertain parameters and boundary conditions, confidence of future projections from such models is strengthened by systematic calibration validation against modern observations and past reconstructions. We can build on experience gained in several preceding Antarctic modeling studies (Briggs et al., 2013, 2014; Whitehouse et al., 2012a; Golledge et al., 2014; Maris et al., 2014, 2015; Poll providing paleo dataset compilations or improved calibration algorithmsadvanced scoring schemes. Modern datasets encompass ice thickness, grounding line and calving front position (Fretwell et al., 2013) (Bedmap2, Fretwell et al., surface velocity (Rignot et al., 2011) as well as uplift rates from recent GPS measurements (Whitehouse et al., 2012b). Reconstructions of grounding line location at the Last Glacial Maximum (LGM) as provided by the RAISED Consortium (Bentley et al., 2014, personal communication Stewart Jamieson) were (Bentley et al., 2014) are used as paleo constraints as well as grounding line locations and cosmogenic elevation–age data from the AntICEdat database (Briggs and Tarasov, 2013) at specific sites during the deglaciation period.

In this study we run simulations of the entire Antarctic Ice Sheet with the Parallel Ice Sheet Model (PISM, Winkelmann et al., 2011; The PISM authors, 2017). The hybrid of two shallow approximations of the stress balance and the comparably coarse resolution of 16 km allow for running an ensemble of simulations of ice sheet dynamics over the last two (dominant) glacial cycles, each lasting for about 100,000 years (or 100 kyr). The three-dimensional evolution of the enthalpy within the ice sheet accounts for the formation of temperate ice (Aschwanden and Blatter, 2009; Aschwanden et al., 2012) and for the production of sub-glacial water (Bueler and van Pelt, 2015). We use a 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. The so-called till friction angle (accounting for small-scale till strength) and the effective pressure enter the Mohr-Coulomb yield stress criterion (Cuffey and Paterson, 2010)and hence . The yield criterion, in turn is part of the the pseudo plastic sliding law, which relates basal sliding velocity to basal shear stress.

PISM comes with a computationally-efficient generalization of the Elastic-plate Lithosphere with Relaxing Asthenosphere (ELRA) Earth model (Lingle and Clark, 1985; Bueler et al., 2007) with spatially varying flow in a viscous upper mantle half-space below the elastic plate, which does not require relaxation time as parameter. Geothermal heat flux based on airborne magnetic data from Martos et al. (2017)is applied, is applied to the lower boundary of a bedrock thermal layer of 2 km thickness, which accounts for storage effects of the upper lithosphere and hence estimates the heat flux at the ice-bedrock interface. Climate boundary conditions are based on mean precipitation from Racmo2.3p2 (Wessem et al., 2018) and a temperature parameterization based on ERA-Interim re-analysis data (Simmons, 2006) in combination with the empirical Positive-Degree-Day method (PDD, e.g., Reeh, 1991). Climatic forcing is based on ice-core reconstructions from EPICA Dome C (EDC, Jouzel et al., 2007) and WAIS Divide ice core (WDC, Cuffey et al., 2016) as well as on sea-level reconstructions from the ICE-6G GIA model (Stuhne and Peltier, 2015, 2017). 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 (Schmidtko et al., 2014) in adjacent to the ice shelves around the Antarctic continent (Schmidtko et al., 2014) . Therein we consider mean values over 18 separate basins based on (Zwally et al., 2015) adjacent to Zwally et al. (2015) . PICO updates melt rates according to changes in ocean temperatures or the geometry of the ice shelves around the Antarctic continent(while changes in salinity are neglected). A description of PISM for paleo applications and sensitivity of the model to various uncertain parameter and boundary conditions are discussed in a companion paper companion paper (Albrecht et al., 2019) .

Here, we explore uncertain model parameter ranges related to ice-internal dynamics and boundary conditions (e.g. climatic forcing, bedrock deformation and basal till properties), and use the Large Ensemble large ensemble approach with full-factorial sampling for the statistical analysis, following Pollard et al. (2016). This 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 use a much finer model resolution. The analysis procedure yields an aggregated score for each of the 256 ensemble simulations, which measures the misfit between PISM simulation and 9 equally weighted types of datasets. Each score can be associated with a probabilistic weight to compute the average envelope of simulated Antarctic Ice Sheet and equivalent sea-level histories and hence providing data-constrained present-day states that can be used for projections with PISM.

2 Ensemble analysis

Ice sheet model simulations generally imply uncertainties in used parameterizations and applied boundary conditions. In order to generate uncertainty estimates for reconstructions of the Antarctic Ice Sheet history and equivalent sea-level envelopes we employ an ensemble analysis approach that uses full-factorial sampling, i.e., one run for every possible combination of parameter values. We follow here closely the simple-averaging approach used in Pollard et al. (2016). This method yields as reasonable results for an adequately resolved parameter space as more advanced statistical techniques ~~with means of interpolating~~ that interpolate results between sparsely separated points in multi-dimensional parameter space. 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 (see companion paper, Albrecht et al., 2019) to cover a representative range of model responses.

2.1 Ensemble parameter

We have identified four relevant independent PISM ensemble parameters with a prior range for each parameter capturing different uncertainties in ice flow dynamics, glacial climate, basal friction and bedrock deformation. The selected parameters passed the two main criteria and (1) showed a relatively high sensitivity of the ice volume to parameter change, while (2) arriving at a present-day state with tolerable anomaly to observations, which is not at all self-evident. The four parameters and the four values used in the ensemble analysis are:

ESIA: Ice-flow enhancement parameter of the stress balance in Shallow Ice Approximation (SIA; Morland and Johnson, 1980; Winkelmann et al., 2011, Eq. 7). Ice deforms more easily in shear for increasing values of ~~1, 2, 4 and 7 (non-dimensional)~~ 1, 2, 4 and 7 (non-dimensional) within the Glen-Paterson-Budd-Lliboutry-Duval law~~connecting~~. It connects strain rates $\dot{\epsilon}$ and deviatoric stresses τ for ice softness A, which depends on both liquid water fraction ω and temperature T (Aschwanden et al., 2012),

$$\dot{\epsilon}_{ij} = \mathrm{ESIA} \cdot A(T,\omega)\, \tau^{n-1}\, \tau_{i,j} \tag{1}$$

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 relevant for ice stream and ice shelf regions.

PPQ: Exponent q used in "pseudo plastic" sliding law which relates bed-parallel basal shear stress $\boldsymbol{\tau}_b$ to sliding velocity $\boldsymbol{u}_b$ in the form

$$\boldsymbol{\tau}_b = -\tau_c \frac{\boldsymbol{u}_b}{u_0^{\,q}\ |\boldsymbol{u}_b|^{1-q}}\,, \tag{2}$$

as calculated from the Shallow Shelf Approximation (SSA) of the stress balance (Bueler and Brown, 2009), for threshold speed u_0 and yields stress τ_c. The sliding exponent hence covers uncertainties in basal friction. Values are ~~0.25, 0.5, 0.75 and 1.0 (non-dimensional)~~0.25, 0.5, 0.75 and 1.0 (non-dimensional).

PREC: Precipitation scaling factor f_p according to temperature forcing ΔT motivated by Clausius-Clapeyron-relationship and data analysis (Frieler et al., 2015), which can be formulated as exponential function (Ritz et al., 1996; Quiquet et al., 2012) as

$$P(t) = P_0 \exp\left(f_p \,\Delta T(t)\right) \approx P_0 \left(1.0 + f_p \,\Delta T(t)\right). \tag{3}$$

For given present-day mean precipitation field P_0, the factor f_p captures uncertainty in climatic mass balance, in particular for glacial periods. Values are 2, 5, 7 and 10/K2, 5, 7 and 10 %/K.

VISC: Mantle viscosity determines the characteristic response time of the linearly viscous half-space of the Earth (overlain by an elastic plate lithosphere) to changing ice and adjacent ocean loads (Bueler et al., 2007, Eq. 1). It covers uncertainties within the Earth model for values of 1×10^{19}, 5×10^{19}, 25×10^{19} and 100×10^{19} Pas0.1×10^{21}, 0.5×10^{21}, 2.5×10^{21} and 10×10^{21} Pa s.

2.2 Misfit evaluation with respect to individual data-types

With four varied parameters with and each parameter taking four values, the ensemble requires 256 runs. For an easier comparison to previous model studies, results are analyzed using the simple averaging method (Pollard et al., 2016). It calculates an objective aggregate score for each ensemble member that measures the misfit of the model result to a suite of selected observational modern and geologic data. The inferred misfit score is based on a generic form of an observational error model assuming a Gaussian error distribution with respect to any observation interpretation uncertainty (Briggs and Tarasov, 2013, Eq. 1).

Present-day ice sheet geometry (thickness and grounding line position) provide the strongest spatial constraint of all data-types and also offer a temporal constraint in the late Holocene. Gridded datasets are remapped to 16 km model resolution. Most of the present-day observational constraints follow closely the definitions in Pollard et al. (2016, Appendix B, Approach (A)), but weighted with each grid-cell's specific area with respect to stereographic projection. We added observed modern surface velocity as additional constraint and expanded the analysis to the entire Antarctic Ice Sheet.

1. TOTE: Mean-square error mismatch of present-day *grounded areas* to observations (Fretwell et al., 2013) assuming uncertainty in grounding line location of 30 km, as in Pollard et al. (2016, Appendix B1). Mismatch is calculated relative to the continental domain , that is defined here as area with bed elevation above -2500 m.

2. TOTI: Mean-square error mismatch of present-day floating *ice shelf areas* to observations (Fretwell et al., 2013) assuming uncertainty in grounding line and calving front location of 30 km, according to Pollard et al. (2016, Appendix B2).

3. TOTDH: Mean-square-error model misfit of present-day state to observed *ice thickness* (Fretwell et al., 2013) with respect to an assumed observational uncertainty of 10 m and evaluated over the contemporary grounded region, close to Pollard et al. (2016, Appendix B3).

4. TOTGL: Mean-square-error misfit to observed *grounding line location* for the modeled Antarctic grounded mask (ice rises excluded) using a signed distance [1] field two-dimensional distance field approximation [1]. This method is different to the GL2D constraint used in Pollard et al.

[1] https://pythonhosted.org/scikit-fmm

[1] https://pythonhosted.org/scikit-fmm

(2016, Appendix B5), and is only applied to the present-day grounding line around the whole Antarctic Ice Sheet according to Fretwell et al. (Bedmap2; 2013)~~and~~ , while considering observational uncertainty of 30 km as in TOTI and TOTE above.

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. Misfit is evaluated for the closest model grid point as in Pollard et al. (2016, Appendix B8), including intra-data type weighting (Briggs and Tarasov, 2013, Sect. 4.3.1).

6. TOTVEL: Mean-square error misfit in (grounded) surface ice speed compared to a remapped version of observational data by Rignot et al. (2011) including their provided grid-cell wise standard deviation, bounded below by 1.5 m/yr.

Paleo-data type constraints are partly based on the AntICEdat compilation by Briggs and Tarasov (2013, Sect. 4.2), following closely their model-data misfit computation. This compilation also includes records of regional sea-level change above present-day elevation (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 self-consistent sea-level model ~~. According in~~ 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. According to Pollard et al. (2016, Appendix B4) we evaluate past and present grounding line location along four relevant ice shelf basins.

7. TROUGH: Mean-square error misfit of modeled grounding line position along four transects through Ross, Weddell and Amery Basin and Pine Island Glacier at the Last Glacial Maximum (20 kyr BP) as compared to reconstructions by ~~Bentley et al. (RAISED Consortium 2014, Scenario A)~~ Bentley et al. (RAISED at present day compared to ~~Fretwell et al. (Bedmap2 2013)~~ Fretwell et al. (Bedmap2; 2013) , both remapped to model grid. An uncertainty of 30 km in the location of the grounding line is assumed as in Pollard et al. (2016, Appendix B4), but as mean over those two most confident dates and for all four mentioned troughs. In contrast to previous model calibrations, reconstructions of the grounding line position at 15, 10 and 5 kyr BP have not been taken into account here, as they would favor simulations ~~which~~ that reveal a rather slow and progressive grounding line retreat through the Holocene in both Ross and Ronne Ice Shelf, which has not necessarily been the case (Kingslake et al., 2018).

8. ELEV: Mean of squared misfit of past (cosmogenic) surface elevation vs. age in the last 120 kyr based on model–data differences at 106 individual sites (distributed over 26 regions, weighted by inverse areal density, see Sect. 4.3.1 in Briggs and Tarasov (2013)). For each data-point the smallest misfit to observations is computed for all past ice surface elevations

(sampled every 1 kyr) of the 16 km model grid interpolated to the core location and datum as part of a thinning trend (Briggs and Tarasov, 2013, Sect. 4.2). A subset of these data has been also used in Maris et al. (2015); Pollard et al. (2016, Appendix B7).

9. EXT: Mean of squared misfit of observed ice extent at 27 locations around the AIS in the last 28 kyr with dates for the onset of open marine conditions (OMC) or grounding line retreat (GLR). The modeled age is computed as the most recent transition from grounded to floating ice conditions considering the sea-level anomaly. The model output every 1 kyr ~~model output~~ is interpolated down to core location and linearly interpolated to 100 yr temporal resolution, while weighting is not necessary here, as described in Briggs and Tarasov (2013, Sect. 4.2). A subset of these data has been also used in Maris et al. (2015)

2.3 Score aggregation

Each of the misfits above are first transformed into a normalized individual score for each data type i and each run j using the median over all misfits $M_{i,j}$ for the 256 simulation. The procedure closely follows Approach (A) in Pollard et al. (2016, Sect. 2.4.1). Then the individual score $S_{i,j}$ is normalized according to the median to

$$S_{i,j} = \exp\left(-M_{i,j} \,/\, \mathrm{median}(M_{i,j=1..256})\right). \tag{4}$$

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. Hence the aggregated score for each run j is the product of the nine data-type specific scores, according to the score definition in Approach (A) by Pollard et al. (2016)

$$S_j = \prod_{i=1..9} S_{i,j}. \tag{5}$$

This implies, that one simulation with perfect fit to eight data types, but one low individual score, yields a low aggregated score for this simulation and hence for instance a low confidence for future applications.

3 Results

3.1 Analysis of parameter ensemble

We have run the full ensemble of PISM simulation over the last glacial cycle. Figure 1 shows the aggregate scores S_j for each of the 256 ensemble members, over the 4-D space spanned by the parameters ESIA, PPQ, PREC and VISC. Each individual sub-panel shows PPQ vs. VISC, and the sub-panels are arranged left-to-right for varying PREC and bottom-to-top for varying ESIA. Scores are normalized by the best score member, which equals value 1 here.

figs/f01.pdf

Figure 1: Aggregated score for all 256 ensemble members (4 model parameters, 4 values each) showing the distribution of the scores over the full range of plausible parameter values. The score values are computed versus geologic and modern data sets, normalized by the best score in the ensemble, and range from <0.01 (bright yellow, no skill) to 1 (dark red, best score)~~(cfs. Pollard et al., 2016, Figs. 2 + C1)~~, on a logarithmic color scale (cfs. Pollard et al., 2016, Figs. 2 + C1) . The four parameters are the SIA enhancement factor ESIA (outer y-axis), the temperature-dependent precipitation scaling PREC (outer x-axis), the mantle viscosity VISC (inner y-axis) and the power-law sliding exponent PPQ (inner x-axis).

The parameter ESIA enhances the shear-dominated ice flow and hence ~~ice thickness~~ yields ice thinning particularly in the interior of the ice sheet and therewith a decrease in the total ice volume. ESIA values of 4 or 7 have been used in other models (e.g., Maris et al., 2015) to compensate for ~~underestimated~~ overestimated ice thickness in the interior of East Antarctic Ice Sheet under present-day climate conditions. In our ensemble, we find a trend towards higher scores for small ESIA with values of 1 or 2 (in the upper half section of Fig. 1). This becomes more prominent when considering ensemble-mean score shares for individual parameter values as in Fig. 3, with a normalized mean score of 46% for ESIA = 1 as compared to a mean score of 6% for ESIA = 7. Most of this trend is a result of the individual data-type score TOTDH (see Fig. 5, column 4, row 3) as it measures the overall misfit of modern ice thickness (and volume distribution). Partly this trend can be also attributed to the TROUGH data-type scores (Fig. 5, column 8, row 3), as for higher ESIA values grounding line motion ~~is slowed downand the time until~~ tends to slow down, such that the time between LGM and present is not sufficient for a complete retreat back to the observed present-day location, at least in some ice shelf basins. The ~~best score~~ best-score ensemble members for small ESIA values are found in combination with both high values of mantle viscosity VISC and high values of friction exponent PPQ (center column panels in Fig. 4).

Regarding the choice of the precipitation scaling PREC the best-score members are found at the upper sampling range with values of 7 %/K or 10 %/K (see right half section in Fig. 1). Considering normalized ensemble-mean score for individual parameter values over the full range of 2–10 %/K, we can find a trend from 13% to 42% (see lowest panel in Fig. 3). Regarding combinations of parameter with PREC (left-hand column in Fig. 4), we detect a weak trend towards lower ESIA and higher PPQ, while individual data-type scores (lower row in Fig. 5) show a rather ~~trendless patterns~~uniform pattern, in particular regarding the misfit to present-day ~~obsrvations~~observations. As the PREC parameter is linked to the temperature anomaly forcing, it affects the ice volume and hence the grounding line location particularly for temperature conditions different from present day. This suggests a stronger signal of PREC parameter variation in the paleo data-types scores.

A more complex pattern is found for PPQ in each of the sub-panels of Fig. 1 with highest scores for values of 0.75 and 1.0. Averaged over the ensemble and normalized over the four parameter choices we find ~~mean scores~~ a mean score of 5% for PPQ = 0.25 (and hence rather plastic sliding) while best scores are found for PPQ = 0.75 and PPQ = 1.0 (linear sliding) with mean scores of 40% and 46%, respectively (see second panel in Fig. 3). ~~The mean score reveal best scores~~ Best scores are found in combination with medium mantle viscosity VISC between 0.5×10^{21} Pa s and 2.5×10^{21} Pa s, as visible in the upper right panel of Fig. 4. As sliding mainly affects the ice stream flux, the ~~score trend~~ trend in aggreagted score over the range of PPQ values mainly results from the velocity misfit data-type TOTVEL and grounding line position related data-types (TOTE, TOTGL and THROUGH), see Fig. 3 (second row).

Regarding mantle viscosity VISC, scores are generally low with 9% for the smallest sampled value of the parameter VISC = ~~10^{20}~~ 0.1×10^{21} Pa s, while best scores are found in the ensemble for the five times larger viscosity of VISC = ~~5×10^{20}~~ 0.5×10^{21} Pa s, with 44%. ~~This value is also used in the GIA modelICE-6G (Peltier et al., 2015) , but~~ In our model, the mantle viscosity parameter has been applied to the whole Antarctic continent, although observations in some localized regions as in the Amundsen Sea ~~,~~ suggest that upper mantle viscosities could be ~~even smaller~~ considerably smaller than the tested range, up to the order of 10^{19} Pa s (Barletta et al., 2018). For the upper range of ~~mantle viscosities~~ tested mantle viscosities, up to VISC = ~~10^{22}~~ 10.0×10^{21} Pa s, we find a normalized ensemble mean of 27% and 20%, respectively. Note that VISC parameter values have been sampled non-linearly over a range of two orders of magnitude. ~~The trend stems mainly from the misfit of present-day uplift rates expressed as data-type score TOTUPL.~~ For the lowest value there is a clear trend towards smaller scores in the grounding-line and ice-thickness related data-types, such as TOTE, TOTGL, TROUGH and TOTDH respectively. As mantle viscosity determines the rate of response of the bed to changes in ice thickness a low viscosity corresponds to a rather quick uplift after grounding line retreat and hence to a retarded retreat~~and hence~~, which corresponds to a rather extended present-day state. This implies smaller ice shelves with slower flow and less velocity misfit, such that also TOTVEL favors small VISC values. In contrast, a trend to rather high mantle viscosities in the aggregated score stems mainly from the misfit of present-day uplift rates expressed as data-type score TOTUPL, probably due to reduced sensitivity to fluctuations in grounding line location. High mantle viscosities involve a slow bed uplift and grounding line retreat can occur faster. More specifically, in the partially over-deepened ice shelf basins, which have been additionally depressed at the Last Glacial Maximum by a couple of hundred meters as compared to present, grounding line retreat can amplify itself in terms of a regional Marine Ice Sheet Instability (Mercer, 1978; Schoof, 2007; Bart et al., 2016). The In fact, the best score ensemble members are found for intermediate mantle viscosities of VISC = 5×10^{20} 0.5×10^{21} Pa s and VISC, and VISC = 25×10^{20} 2.5×10^{21} Pa s. This could be a result of the product formulation of the aggregated score, in which individual data types scores favor opposing extreme values.

The five best-score ensemble members and associated parameter combinations are listed in Table 1. With the best-fit simulation parameters we have participated in the initMIP-Antarctica model intercomparison (Seroussi et al., 2019, PISMPAL3) . The individual scores with respect to the nine data-types are visualized for the best 20 best ensemble members in Fig. 2. The scores associated with the paleo data-types ELEV and EXT show only comparably little variation among the ensemble (in total around 0.4both around 0.07 standard deviation). This also applies for the present-day ice shelf area mismatch TOTI (total spread of 0.20.04), as no calving parameter have has been varied. In contrast, present-day data types associated with velocity (TOTVEL) and uplift rates (TOTUPL) show strong variations among the best 20 best ensemble members, with a spread standard deviation in score across the for entire ensemble of 0.7 and 0.85 0.18 and 0.30 respectively. For data types that are related to grounding line position (TOTGL, TOTE, TROUGH) and ice volume (TOTDH) we find a similar order as for the TOTAL aggregated score (Fig. 2), with individual spread standard deviations in scores of 0.5-0.6 0.12-0.20 across all ensemble members. All data-type specific misfits are visualized as histogram in the Supplementary Material B (Fig. S6).

Simulation No.	ESIA	PPQ	PREC	VISC	Score	Normal. Score
165	**2.0**	**0.75**	**7 %/K**	**0.5×10^{21} Pa s**	6.1×10^{-3}	1.0
245	2.0	1.0	10 %/K	0.5×10^{21} Pa s	4.6×10^{-3}	0.76
242	1.0	1.0	10 %/K	2.5×10^{21} Pa s	3.9×10^{-3}	0.63
241	1.0	1.0	10 %/K	0.5×10^{21} Pa s	3.2×10^{-3}	0.53
261	1.0	0.75	7 %/K	0.5×10^{21} Pa s	2.4×10^{-3}	0.39

Table 1: Five best-score ensemble parameter combinations with parameter values and total scores. The best-fit simulation parameters were used in the initMIP-Antarctica model intercomparison (Seroussi et al., 2019, PISMPAL3) and for the reference simulation in the companion paper (Albrecht et al., 2019) .

Comparing the ensemble-mean present-day ice thickness with observations (Bedmap2; Fretwell et al., 2013) we find regions in the inner East Antarctic Ice Sheet and in parts of the Weddell Sea sector that are about 200 m too thin, while ice thickness is overestimated by more than 500 m in the Siple Coast, in the Amery basin and along the coast line, where smaller ice shelves tend to be grounded in the simulations (Fig. 6a). Ross Sea, Weddell Sea and Amery basins show the largest ensemble-score weighted standard deviation with more than 500 m ice thickness (Fig. 6b). The ensemble spread in those basins can be associated with uncertainties in grounding line position, as grounding line has to retreat in time from . From its extended position at Last Glacial Maximum crossing the grounding line has to retreat across the basins in time, with distances of up to 1000 km long basins, leaving behind the large floating ice shelves (Fig. 7). In about 10% of the score-weighted simulations grounding line remains at the extended position without significant retreat, linked to high basal friction (PPQ=0.25) and an efficient negative feedback on grounding line motion, related to a fast responding bed (low VISC). In contrast, for rather low friction and high mantle viscosities, we find fast grounding line retreat, with a stabilization of grounding line position at our or even inland of the observed location in 50% or 75% of the score-weighted simulations in the Ross and Weddell Sea sector, respectively (Fig. 8, upper panels). Due to the unloading of grounded ice retreat and the consequent unloading across the large ice shelf basinsthe sea floor , the marine bed lifts up by up to a few hundred meterswhich leads , which can lead to grounding line re-advance supported by the formation of ice rises (Kingslake et al., 2018). The ensemble mean re-advance in the ensemble mean is up to 100 km, while some of the best-score simulations reveal temporary ungrounding through the Holocene up to 400 km upstream of the present-day grounding line in the Ross sector. The Amundsen Sea sector and Amery Ice Shelf do not show such rebound effects in our model ensemble (Fig. 8, lower panels) **Simulation No. ESIA PPQ PREC VISC Score Normal. Score**6165 2.0 0.75 7/K 5×10^{20}Pas 6.1×10^{-3} 1.0 6245 2.0 1.0 10/K 5×10^{20}Pas 4.6×10^{-3} 0.76 6242 1.0 1.0 10/K 25×10^{20}Pas 3.9×10^{-3} 0.63 6241 1.0 1.0 10/K 5×10^{20}Pas 3.2×10^{-3} 0.53 6261 1.0 0.75 7/K 5×10^{20}Pas 2.4×10^{-3} 0.39 Five best-score ensemble parameter combinations with parameter values and total scores.

figs/f02.pdf

Figure 2: ~~Nine~~ Aggregated and 9 individual scores for 20 ~~(total)~~ best ensemble members computed versus modern and geologic data sets, divided by dashed line. The score values are normalized by the median misfit, and range from 0 (bright yellow, no skill) to 1 (dark red, best score) on a linear color scale. The ~~ensemble spread of some~~ standard deviation for the individual paleo data ~~types~~ type scores ELEV and EXT, as well as for present-day ice shelf mismatch TOTI, is ~~comparably low with around 0.4 and 0.2 respectively~~below 0.1. In contrast, grounding line location at LGM and present-day along four ice shelf basins (TROUGH) and present-day uplift rates (TOTUPL) have strongest impacts on the aggregated score with ~~more than 0.85 spread in individual scores across the ensemble~~a standard deviation of 0.2 and 0.3, respectively. Intermediate ~~spread~~ variability of individual scores show TOTGL~~with around 0.5~~, TOTE~~and~~ , TOTDH ~~with around 0.6~~ and TOTVEL with ~~around 0.7~~a standard deviation between 0.1 and 0.2.

figs/f03.pdf

Figure 3: Ensemble-mean scores for individual parameter values (normalized such that sum is 1, or 100%). The weighted mean over the four ensemble-mean scores with standard deviation is shown in red (compare Figs. 3 + C2 in Pollard et al. (2016)).

Figure 4: Ensemble-mean scores for six possible pairs of parameter values to visualize parameter dependency (compare Figs. 4 + C3 in Pollard et al. (2016)). Values are normalized such that the sum for each pair is 1. Color scale is logarithmic ranging from 0.01 (bright yellow) to 1 (red).

figs/f05.pdf

Figure 5: Scatter plot of both aggregated score and the nine individual data-type scores (panels from left to right) for each parameter setting (VISC, PPQ, ESIA and PREC as y-axis). Red dots indicate the best-score member, green and blue the second and third best ensemble members (see Table 1). Grey-dashed line indicates mean score tendency over sampled parameter range.

figs/f06.pdf

Figure 6: Score-weighted mean ice thickness anomaly to Bedmap2 (left) and score-weighted standard deviation of ice thickness (right). Ice thickness in coastal regions in West Antarctica but also in the Amery basin are generally overestimated. Amery and Filchner-Ronne Ice Shelves and Siple Coast region reveal the highest standard deviation in reconstructed present-day ice thickness among the ensemble members.

figs/f07.pdf

Figure 7: Ensemble-score weighted grounded mask for 5 kyr snapshots. Mask value 1 (red) indicates grounded area which is covered by all simulations, while blueish colors indicate areas which are covered only by a few simulations with low scores (compare Fig. D4 in Pollard et al. (2016)). For the last two snapshots, grounding line in the Ross Sea and Weddell Sea sector is found in about 50% of score-weighted simulations inland of its present location (Fretwell et al., 2013, grey line) with some grounding line re-advance ~~(Kingslake et al., 2018)~~ (Kingslake et al., 2018; Siegert et al., 2019) . In contrast less than 10% show no grounding line retreat from glacial maximum extent. Black lines indicate reconstructions by the RAISED Consortium (Bentley et al., 2014, Scenario B solid and scenario A dashed).

figs/f08.pdf

Figure 8: Ensemble score-weighted grounded ice cover along transects trough Weddell, Ross, Amundsen and Amery Ice Shelf basins over the last 25 kyr simulation period (left y-axis, compare Fig. D5 in Pollard et al. (2016)). Grounded areas which are covered by all simulations are indicated by value 1 (red), while blueish colors indicate areas which are covered only by some simulations (or those with low scores). Grounding line in the Ross Sea and Weddell Sea sector is found inland of its present location (vertically dotted) within the last 10 kyr simulation time in about 50% and 75% of score-weighted simulations, respectively. The score-weighted mean curve (black) reveals re-advance of the grounding line of up to 100 km in about 20% of the score-weighted simulations, both in the Ross and Weddell Sea sector, as discussed in (Kingslake et al., 2018). Such behavior is not found in the Pine Island trough, where grounding line retreat stops in 90% of the simulations at about 200 km downstream of its present day location. Similar in the Amery Ice Shelf, where in 30% of score-weighted simulations the ice shelf does not retreat at all from its LGM extent. Bed topography (Bedmap2; Fretwell et al., 2013) along the transect is indicated as gray line with respect to the right y-axis. For the four troughs, the data type TROUGH is evaluated for the two time slices corresponding to LGM (20 kyr BP) and present.

3.2 Reconstructed sea-level contribution histories

For the parameter ensemble analysis we have first run The full parameter ensemble is based on four simulations starting in the penultimate interglacial (210 kyr BP). These four simulations use four different values of mantle viscosity covering two orders of magnitude (VISC $= 10^{20} - 10^{22}$ Pa s). They show quite a consistent maximum ice volume at the penultimate glaciation around 130 kyr BP (see violet lines in Fig. 9). Due to the different Earth response times associated with varied mantle viscosities, the curves branch out when the ice sheet retreats. Those four simulations were used as initial states at 125 kyr BP for the other 252 simulations of the large ensemble. At the end of the Last Interglacial (stage (LIG, Eemian) at around 120 kyr BP, when the full ensemble has been run for only 5 kyr, the ensemble mean ice volume is 1.0 m SLE below modern with a score-weighted standard deviation of around 2.7 m SLE (volume of grounded ice above flotation in terms of global mean sea level equivalent as defined in the companion paperAlbrecht et al. (2019)). This corresponds to a grounded ice volume anomaly in relation to present day observations of - $0.3 \pm 1.4 \times 10^6$ km^3. These numbers may not reveal the full possible ensemble spread as simulations still carry some memory of the previous glacial cycle simulations with different parameters. On average, grounding lines and calving fronts retreat much further inland at LIG than for present-day conditions. Yet, complete collapse of WAIS West Antarctic Ice Sheet (WAIS) does not occur in any of the ensemble members, most likely as a result of intermediate till friction angles and hence higher basal shear stress underneath the inner WAIS (see optimization in companion paper). Albrecht et al. (2019)). In the case of triggered WAIS collapse one could expect an Antarctic contribution to the Eemian sea-level high stand of 3–4 m SLE (Sutter et al., 2016) . Also previous paleo model studies estimate the Antarctic contribution to be at least 1 m SLE, based on a globally integrated signal, and likely significantly more, depending on Greenland's contribution (Cuffey and Marshall, 2000; Tarasov and Peltier, 2003; Kopp et al., 2009) . This value has been thus used as lower bound in terms of a "sieve" criterion in previous Antarctic model ensemble analysis (Briggs et al., 2014) .

Assuming, that the memory of the previous spin-up has vanished at the Last Glacial Maximum (in our simulations at around 15 kyr BP), where the model ensemble yields a range of (grounded) Antarctic Ice Sheet volume of 9.4 ± 4.1 m above present-day observations, or $6.5 \pm 2.0 \times 10^6$ km^3. The histogram of score-weighted sea-level anomalies of all simulations at Last Glacial Maximum actually reveals four distinct maxima at around 4.5, 8.1, 9.0 and 13.0 m SLE (Fig. 10 a), which can be attributed to the five best-score simulations in Table 1. The ensemble spread is hence relatively wide, but still quite symmetric, as comparison with the Gaussian normal distribution reveals. The As expected, the LGM ice volume increases for lower PPQ (on average 3 m SLE), lower PREC (on average more than 6 m SLE) and lower ESIA values (more than 12 m SLE on average), while it seems to be rather insensitive to the choice of VISC (less than 0.5 m SLE on average). When comparing simulated volumes at Last Glacial Maximum to modeled present-day volumes (such that model biases cancel out) the model ensemble yields 10.0 ± 4.1 m of global mean sea level equivalents, or 5.8 ± 2.0 × 10^6 km^3.

Most of the deglacial retreat from LGM extent and hence most of Antarctica's sea-level rise contribution occurs in our simulations after 10 kyr BP (cf. Fig. 10 b, c). In particular, for higher mantle viscosities we find episodic self-amplified retreat with change rates of more than 0.5 cm SLE per year ~~change rate~~ in West Antarctic basins (~~as~~ in the best-fit simulation at 7.5 kyr BP, see below in Sect 3.3). This leads in some cases to grounding line ~~migration even upstream of~~ retreat beyond its present location and subsequent re-advance during Holocene~~due to~~ , due to the uplift of the ~~Earth~~ bed (discussed in Kingslake et al., 2018). However, these rapid episodes of retreat occur in our simulations consistently after Meltwater pulse 1A (MWP1a~~)~~ , around 14.5 kyr ~~(~~BP, see dashed line in Fig. 9). This delay supports the idea, that Antarctic Ice Sheet retreat has ~~been not~~ not been a source but rather a consequence of the relatively quick rise in global mean sea-level by about 15 m within 350 yr or ≈ 4 cm/yr at MWP1a (Liu et al., 2016), while core analysis of iceberg-rafted debris suggest earlier and stronger recession of the Antarctic Ice Sheet at the time of MWP1a (Weber et al., 2014). ~~As~~ The MWP1a initiated the Antarctic Cold Reversal (ACR)~~with~~ , a period lasting for about two millenia ~~of~~ with colder surface temperatures~~,~~ . This cooling induced a freshening of surface waters ~~leading~~ and lead to a weakening of Southern Ocean overturning, resulting in reduced Antarctic Bottom Water formation, enhanced stratification and sea-ice expansion. This could have caused an increased delivery of relatively warm Circumpolar Deep Water onto the continental shelf close to the grounding line and hence to stronger ~~sub-surface~~ sub-shelf melt (Golledge et al., 2014; Fogwill et al., 2017). As our sub-shelf melting module is forced with a modified surface temperature anomaly forcing, PICO responds with less melt during ACR period and hence ~~prevents from~~ prohibits significant ice sheet retreat. But even if the intermediate ocean temperature would rise by 1 or 2 K during ACR, the induced additional melt would correspond to less than -1 mm/yr SLE and hence far less than the value of -6 mm/yr SLE found by Golledge et al. (2014) (see also ~~companion paperfor~~ Albrecht et al. (2019) for the corresponding sensitivity analysis). Also MWP1b around 11.3 kyr BP occurred well before deglacial retreat initiated in most simulations of our model ensemble (see Fig.9 c)~~, in contrast to a previous PISM study (Golledge et al., 2014)~~ . The selection criteria for the used ensemble parameters may not sufficiently represent the onset and rate of deglaciation. 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° ~~, see Appendix 4~~ and the availability of till water at the grounding line. More discussion of the interference of basal parameters in terms of an additional (basal) ensemble analysis is given in the Supplementary Material A.

The timing of deglaciation and possible rebound effects can explain a natural drift in certain regions that lasts through the Holocene until present-day. In the score-weighted average the ensemble simulations suggest a sea-level contributions over the last 3,000 model years of about 0.25~~,~~ mm/yr, while for the reference simulation the Antarctic ice above flotation is on average even slightly growing (cf. Fig. 9 c), partly explained by net uplift in grounded areas (Fig. 12).

figs/f09ab.pdf

figs/f09c.pdf

Figure 9: Simulated sea-level relevant ice volume histories over the last two glacial ~~cycle(s)~~ cycles (upper) and for last deglaciation (middle), for all 256 individual runs ~~in~~ of the parameter ensemble, ~~transparency weighted~~ transparency-weighted by aggregated score. Red line indicates the best-score run, the green line and shading indicates the score-weighted ensemble mean and standard deviation

figs/f10.pdf

Figure 10: Equivalent global-mean sea-level contribution (ESL) relative to modern at every 5 kyr over the last deglaciation period. Grey bars show the score-weighted ensemble distribution (0.5 m bins), the red curve indicates the statistically likely range (normal distribution) of the simulated ice volumes with width of 1-sigma standard deviation as for the green envelope in Fig. 9, green gaussian curve from 15k kyr snapshot for comparison (compare to Figs. 6 + C5 in Pollard et al. (2016)).

The simulations are based on the Bedmap2 dataset (Fretwell et al., 2013), remapped to 16 km resolution, which corresponds to a total grounded modern Antarctic Ice Sheet volume of 56.85 m SLE (or 26.29×10^6 km^3). The ensemble mean at the end of the simulations (in the year 2000 or -0.05 kyr BP) underestimates the observed ice volume slightly by 0.6 ± 3.5 m SLE, or in terms of grounded ice volume by $0.7 \pm 1.7 \times 10^6$ km^3 (see Fig. 9). The histogram of score-weighted sea-level anomalies at the end of all simulations can be ~~adequately~~ well approximated by a normal distribution (Fig. 10 d). As for the LGM ice volume the ESIA parameter is responsible for most of the ice volume ensemble range at present day with more than 10 m SLE, while PREC has almost no effect with less than 1 m SLE on average, in contrast to the LGM, as expected. VISC and PPQ reveal on average a range for the present-day ice volume of about 6 m SLE and 5 m SLE, respectively.

3.2.1 Comparison of LGM sea-level estimates to in previous studiesand model observations

For the maximum Antarctic ice volume at the Last Glacial Maximum, the inferred ensemble range of 5.3 - 13.5 m SLE excess relative to observations (or 4.5 - 8.5 $\times 10^6$ km^3) is at the upper range found in the recent literature (Fig. 11). Model reconstructions are basically , except for the "GRISLI" model results (Quiquet et al., 2018) . The other previous model reconstructions are based on four different models: "Glimmer" (Rutt et al., 2009), "PSUPSU-ISM" (or PennState3D) from Penn State University (Pollard and DeConto, 2012a), "ANICE" from Utrecht University (De Boer et al., 2013) and, as in this study, the Parallel Ice Sheet Model (PISM; Winkelmann et al., 2011). This section briefly compares the different model and ensemble approaches with regard to the inferred LGM ice volume estimate.

The modeled range between Last Glacial Maximum and present-day ice volume by Whitehouse et al. (2012a) is about 5.0 $\times 10^6$km^3 (or 7.5 - 10.5m ESL, eustatic sea-level based on volume above flotation), who ran a GIA-model based on a prescribed ice sheet history from different

Whitehouse et al. (2012a) ran 16 Glimmer simulations at 20 km resolution . Golledge et al. (2012) estimated a range of about 2.7 $\times 10^6$km^3 (or 6.7m SLE using a conversion factor of 2.47)from PISM on a with varied sliding and isostasy parameter, and different inputs for the geothermal heat flux, climatic forcing and sea-surface height. They used both geological and glaciological data to constrain the reconstruction and found the best fit simulation at the lower end of their ensemble ice volume range. Golledge et al. (2012) and Golledge et al. (2013) used PISM on a 5 km grid. For the same model and resolutionGolledge et al. (2013) found about 3.35 $\times 10^6$grid for an equilibrium simulation under LGM conditions, while Golledge et al. (2014) retrieved their ensemble mean estimates, relative to observations (Bedmap2), from an ensemble of around 250 PISM deglaciation simulations at 15 km ^{3} (or 8.3m SLE using a conversion factor of 2.47). Briggs et al. (2014) estimated a range between 2.2 and 5.7 $\times 10^6$resolution, with varied basal traction and ice-flow enhancement factors. ANICE simulations have been run on 20 km ^{3} (or 5.6 - 14.3resolution. In a sensitivity study, Maris et al. (2014) varied enhancement factors, till strength and ("ELRA") bedrock deformation parameters, while in Maris et al. (2015) , a small ensemble of 16 simulations with different sea-level and surface temperature forcings have been applied to two different bed topographies over the last 21 mESL, using a conversion factor of 2.52) from PSU simulations kyr. Quiquet et al. (2018) varied four parameters (SIA enhancement, friction coefficient, sub-shelf melt and subglacial hydrology) in 600 equilibrium ensemble simulations with GRISLI for 40 km resolution. Maris et al. (2014) used ANICE on 20km resolution and inferred around 3.8 - 4.8 $\times 10^6$km^3 (or 9.4 - 12.0m s.l. They selected the 12 best thickness fit parameters to run transient simulations over the last four glacial cycles. The relatively high estimate fo LGM ice volume is likely due to the simplified basal drag computation that does not take into account bedrock physical properties (e.using a conversion factor of2.5) . For the same model and resolution Maris et al. (2015) found around 3.5 - 5.2 $\times 10^6$km^3 (or 8.4 - 12.5m s.l.e. using a conversion factor of 2.4). A much higher LGM volume of about 5.8 $\times 10^6$km^3 (or 10.5m SLE ice volume above flotation relative to observations)was retrieved from PISM on 15km resolution by Golledge et al. (2014) . Around 1.6 - 4.8 $\times 10^6$g. sediments). The estimates by Briggs et al. (2014) are based on (the best 178 of) a really large ensemble of more than 3,000 PSU-ISM simulations over the last two glacial cycles with 40 km ^{3} (or 5 - 10m ESL eustatic seal-level change for WAIS only, or 4 - 12m ESL using a conversion factor of 2.48) was found from the PSU model on resolution, coupled with a full visco-elastic isostatic adjustment bedrock response with radially layered earth viscosity profile and different treatments of sub-shelf melt, basal drag, climate forcing, and calving (in total 31 varied parameters). The full ensemble range is certainly much larger, but additional constraints allow for a selection of the most realistic simulations, with most confidence in the lower part of the given range (purple error bar in Fig. 11). Pollard et al. (2016) and Pollard et al. (2017) used the PSU-ISM on 20 km resolution in Pollard et al. (2016) and around 2.8 - 4.1 $\times 10^6$km ^{3} (or 3 - 8m GMSL global mean sea-level change, or 6.9 - 10.2m GMSL using a conversion factor of 2.48) from PSU on a resolution for an ensemble of each 625 simulations over the last 20 km grid in Pollard et al. (2017) . In the Large Ensemble by (Pollard et al., 2017) kyr and varied four parameters related to sub-shelf melt, calving, basal sliding and viscous Earth deformation, while other parameters were supposedly constrained by earlier studies. Pollard et al. (2016) applied an ELRA Earth model applied to the West Antarctic Ice Sheet only, while (Pollard et al., 2017) simulated whole Antarctica coupled to a global Earth-sea level model. In both ensembles, ice volume change since LGM is somewhat biased to comparably low values, as the used scoring algorithm pushed the ensemble to rather slippery basal sliding coefficient on modern ocean beds(personal communication with Dave Pollard).

Ice volume anomaly between Last Glacial Maximum as compared to present in recent modeling studies in units of 10^6km^3. Note that the study by Pollard et al. (2016) only considers the West Antarctic subdomain in their analysis (grey). Golledge and colleagues and this study used PISM, Maris and colleagues used ANICE, Whitehouse and colleagues used GLIMMER and Briggs and Pollard and colleagues used PennState3D as model. Be aware, that ice volume estimates are based on different ice densities in the different models.

3.3 Best-fit ensemble simulation

The best-fit ensemble member simulation (no. 6165, see Table 1) provides an Antarctic Ice Sheet configuration for the present day, which is comparably close to observations. The present-day ice volume of the West Antarctic Ice Sheet is in our ensemble simulations generally overestimated (by around 25), while the much larger East Antarctic Ice Sheet volume is rather underestimated (by around 5). Part of the overestimation can be explained by the relatively coarsely resolved topography of the Antarctic Peninsula and weakly constrained basal friction in the Siple Coast area. This results in a RMSE of ice thickness of 266m (see . As Whitehouse et al. (2012a) , Golledge et al. (2014) and this study provided anomalies based on the volume-above-flotation calculation (VAF), the corresponding SLE values are smaller than the directly converted values (Fig. 12a), a RMSE of grounding line distance of 67km (see Fig. 13) and a RMSE for surface velocities of 6611b). For a conversion factor of $c = 2.5$ our study would yield 12.5 – 16.5 m/yr (see Fig. 14). The best-fit simulation also reproduces the general pattern of observed modern isostatic adjustment rates of more than 10mm SLE instead. For the LGM ice volume excess relative to the modeled present day our study yields 5.9 – 14.1 yr^{-1} (see Fig. 12m SLE (or 3.8 - 7.8 $\times 10^6$ b)with highest uplift rates in the Weddell and Amundsen Sea Region in agreement with GIA model reconstructions (cf. Argus et al., 2014, Fig. 6) . In contrast to these GIA reconstructions, our best-fit simulation shows depression rather than uplift in the Siple Coast regions as grounded ice is still re-advancing and hence adding load.km^3), both indicated in Fig. 11.

figs/f11a.pdf

(a)

figs/f11b.pdf

(b)

Figure 11: Ice volume anomaly between Last Glacial Maximum as compared to present (not observations) in recent modeling studies in units of $10^6\,\mathrm{km}^3$ (a) and in units of meters sea-level equivalents (b). Note that the study by Pollard et al. (2016) only considers the West Antarctic subdomain in their analysis (redish). Golledge and colleagues and this study used PISM (blue and grey), Maris and colleagues used ANICE (orange), Whitehouse and colleagues used GLIMMER simulations (olive), GRISLI simulations by Quiquet and colleagues (green) and Briggs and Pollard and colleagues used PennState3D (or PSU-ISM) as model (blueish) coupled to different Earth models. Be aware, that ice volume estimates are based on different ice densities in the different models and that different conversion factors c have been used. This study, Golledge et al. (2014) , as well as the Glimmer and GRISLI model provided the volume above flotation (VAF), which substracts some portion of the ice volume in panel (b). 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.

3.3 Best-fit ensemble simulation

(a) (b)

Figure 12: (a) Present-day ice thickness anomaly of best fit ensemble simulation with respect to observations (Fretwell et al., 2013), with the continental shelf in grey shades. Blue line indicates observed grounding line, while black lines indicate modeled grounding line and calving front. Large areas of the East Antarctic Ice Sheet are underestimated in ice thickness, while some marginal areas along the Antarctic Peninsula, Siple Coast and Amery Ice Shelf are thicker than observed, with a total RMSE of 266 m. (b) Modeled uplift (violet) and depression (brown) at present-day state as compared to uplift rates from recent GPS measurements (Whitehouse et al., 2012b) in 35 locations (in units mm/yr).

The best-fit ensemble member simulation (no. 165, see Table 1) provides an Antarctic Ice Sheet configuration for the present day, which is comparably close to observations. Yet, the present-day ice volume of the West Antarctic Ice Sheet is overestimated (by around 25%), while the much larger East Antarctic Ice Sheet (EAIS) volume is rather underestimated (by around 5%), which is also valid for the ensemble mean (Fig. 6). Part of the overestimation can be explained by the relatively coarsely resolved topography of the Antarctic Peninsula and weakly constrained basal friction in the Siple Coast and Transantarctic Mountain area. This results in a root-mean-square error (RMSE) of ice thickness of 266 m (see Fig. 12 a), a RMSE of grounding line distance of 67 km (see Fig. 13) and a RMSE for surface velocities of 66 m/yr (see Fig. 14). The best-fit simulation also reproduces the general pattern of observed modern isostatic adjustment rates (see Fig. 12 b) with highest uplift rates of more than 10 mm yr^{-1} in the Weddell and Amundsen Sea Region in agreement with GIA model reconstructions (cf. Argus et al., 2014, Fig. 6) . In contrast to these GIA reconstructions, our best-fit simulation shows depression rather than uplift in the Siple Coast regions as grounded ice is still re-advancing and hence adding load.

Figure 13: Comparison of present-day grounded (left) and floating (right) ice extent in best fit ensemble simulation with respect to observations (Fretwell et al., 2013). Yellow color indicate a match of simulation and observations, orange means grounded/floating in model but not in observations, and blue vise versa. Root-mean-square distance of modeled and observed grounding line is 67 km.

figs/f13.pdf

figs/f14.pdf

Figure 14: Comparison of present-day surface velocity in best fit ensemble simulation (left) with respect to observations (right, Rignot et al., 2011) (middle, Rignot et al., 2011) , all in log-scale. Red Greenish shading indicates slow flowing regions and ice divides, blueish shading indicates regions of fast ice flow with ice shelves and far-inland reaching ice streams, respectively. Model-observations difference is shown for observed glacierized area in right panel, RMSE for surface velocities is 35 m/yr, mean misfit is 66 m/yr.

As At last glacial maximum, before 10 kyr BP, the sea-level curve relevant volume history of the best-score simulation is close to the ensemble mean (Fig.9), a more detailed look into subsequent snapshots since the last glacial termination helps to identify periods of comparably strong changes. Before 10kyr BP our best-fit simulation shows 9). The LGM state is characterized by extended ice sheet flow towards the outer Antarctic continental shelf edges, with more than 2,000 m thicker ice than today in the basins of the largest modern ice shelves (Ross, Weddell, Amery and Amundsen), while the inner East Antarctic Ice was a few hundred meters thinner than today (see Fig. 15). At last glacial maximum around 15 kyr BP our simulations agree well with reconstructions by the RAISED Consortium (Bentley et al., 2014, cf. Fig. 7 a).Last Glacial Termination

Even though this is not the primary focus of this parameter ensemble study, it is worthwhile to have a closer look into the deglacial period. The last glacial termination (also known as Termination I, which is the end of Marine isotope stage 2), and hence the period of major ice sheet retreat~~initiates in~~ , initiates in our best-score simulation in the Ross and Amundsen sector ~~in the best-score simulation~~ at around 9 kyr BP, in the Amery sector at around 8 kyr BP and in the Weddell Sea Sector at around 7 kyr BP. Maximum change rates are found accordingly in the period 10–8 kyr BP with -1.4 mm/yr SLE (or -660 Gt/yr) in the period 8–6 kyr BP with -2.4 mm/yr SLE (or -1,300 Gt/yr, Fig. 16), with a peak of around -5 mm/yr SLE at 7.5 kyr BP (or -3,300 Gt/yr in the 100 yr running mean, compare black and khaki line in Fig. 17). This rate of change is significantly larger than in the ensemble mean, in which some retreat occurs throughout the Holocene (see Fig. 9 c). The total ice volume change during the period 10–5 kyr BP amounts to -9.7 m SLE. Most of this change can be attributed to increased discharge by around -1,000 Gt/yr and increased sub-shelf melting by around -450 Gt/yr (partly due to increased floating ice shelf area), while surface mass balance increased only by around 300 Gt/yr (Fig. 17). Recent proxy-data reconstructions from the eastern Ross continental shelf suggested initial retreat not before 11.5 kyr BP (Bart et al., 2018), likely around ~~9-8~~9–8 kyr BP (Spector et al., 2017), which is consistent with our model simulations. In the reconstructions by the RAISED Consortium most of the retreat in the Ross Sea Sector (almost up to present-day grounding line location) occurred between 10 kyr BP and 5 kyr BP, while major retreat in the Weddell Sea Sector likely happened before 10 kyr BP in scenario A and after 5 kyr BP in scenario B (Bentley et al., 2014, cf. Fig. 7 b,c). ~~In our simulations a~~

A Holocene minimum ice volume is reached in ~~the late Holocene~~ our simulations around 3 kyr BP with slight re-advance and thickening in the Siple coast and Bungenstock ice rise until present-day (see Fig. 16). This regrowth signal cannot be inferred from RAISED reconstructions with ~~only~~ snapshots only every 5 kyr ~~snapshots~~ (Bentley et al., 2014). The corresponding mass change ~~agrees well with sea-level relevant volume change with about 0.07~~is rather small with 60 ~~mm~~Gt/yr ~~SLE ice sheet regrowth (or 60~~(or 0.07 ~~Gt~~mm/yr SLE) in the last ~~3000~~ 3,000 years, see Fig. 17.

During this late Holocene period, surface mass balance of around 3,700 Gt/yr is balanced by approximately -2,600 Gt/yr discharge, while sub-shelf melt plays a minor role with around -1,000 Gt/yr.

figs/f15.pdf

Figure 15: Snapshots of grounded ice thickness anomaly to present-day observations Fretwell et al. (2013) (Fretwell et al., 2013) over last 15 kyr of in best-fit simulation, analogous to Fig. 2 in Golledge et al. (2014) . At LGM state grounded ice extends towards the edge of the continental shelf with much thicker ice than present, mainly in West Antarctica. Retreat of the ice sheet occurs first in the Ross basin between 9 and 8 kyr BP, followed by the Amery basin around 1 kyr later and the Amundsen and Weddel Sea basin between 7 and 5 kyr BP. East Antarctic Ice Sheet thickness is underestimated throughout the deglaciation period (light blue shaded area).Compare Fig. 2 in Golledge et al. (2014) .

figs/f16.pdf

Figure 16: Snapshots of relative ice thickness change rates every 2 kyr over last 16 kyr ~~of~~ in best-fit simulation, analogous to Fig. 4 in Golledge et al. (2014) . Deglaciation starts in the Ross and Amundsen Sector after 10 kyr BP with a mean change rate of -1.4 mm/yr SLE followed by the Amery and Weddell Sea Sector after 8 kyr BP with mean change rates of up to -2.4 mm/yr SLE (with peaks of up to -5 mm/yr SLE at 7.5 kyr BP). In the late Holocene period since 4 kyr BP the best fit simulation shows some thickening in the Siple Coast and in the Bungenstock Ice Rise corresponding to about +0.1 mm/yr SLE.~~Compare Fig. 4 in Golledge et al. (2014) .~~

figs/f17.pdf

Figure 17: Mass fluxes over the last 15 kyr for the best-fit simulation (left axis), with the sum of surface (~~blue~~orange) and basal mass balance (~~orange and green~~blue, sub-glacial melt in light blue is negligible) and discharge (100 yr running mean in ~~red~~violet) ~~as~~ yielding total mass change (~~violet~~khaki). Mass change agrees well with sea-level relevant volume change (100 yr running mean in black, right axis). Main deglaciation occurs between 9-5 kyr BP (black dotted line, right axis) with ~~about 0.07~~on average 2.0 mm/yr ~~SLE ice sheet regrowth (~~or ~~60~~1,000 Gt/yr (blue bar)~~in~~ , significantly after MWP-1A (grey bar). In the last ~~3000 years (indicated~~ 3 kyr of the best-fit simulation, the Antarctic Ice Sheet re-gains mass by ~~dashed vertical line~~about 60 Gt/yr, which equals about 0.07 mm/yr SLE (red bar).

3.4 ~~Comparison to previous large~~ Discussion of individual ensemble ~~study~~parameters

~~Our study follows closely the Large Ensemble analysis method by Pollard et al. (2016) for simulation results of the West Antarctic Ice Sheet with the PSU model for 20km grid resolution~~

In this section we want to discuss the effects of individual ensemble parameters in more detail, also in comparison to previous model studies. We performed our analysis for an ensemble of 256 simulations of the entire Antarctic Ice Sheet ~~for~~ over the last two glacial cycles with 16 km grid resolution using PISM~~and four different~~ . The parameter ensemble is spanned by four model parameters (Sect. 2.1)~~. Pollard et al. (2016) span their ensemble with parameters that involve mainly oceanic properties (ice shelf melting and calving) or properties of modern ocean-bed areas (sliding over ice-shelf basins and bedrock relaxation time~~, two of them are more relevant for glacial dynamics in the West Antarctic Ice Sheet (VISC and PPQ), while ~~other parameters that affect the modern grounded ice areas are sufficiently constrained by earlier studies.~~the other two are more related to glacial ice volume change in the East Antarctic Ice Sheet (ESIA and PREC, see overview in

Sect. 2.1).

For the bedrock response we chose the upper mantle viscosity as one ~~of the parameters in our ensemble with~~ ensemble parameters and found maximum scores around values of VISC = ~~5~~0.5 $\times 10^{21}$ Pa s for whole Antarctica. This corresponds to a rebound time scale of ~~a few~~ 1–3 thousand years, which is in line with the findings in ~~Pollard et al. (2016)~~ Maris et al. (2014); Pollard et al. (2016) for WAIS, using a simplified Earth model (ELRA). Pollard et al. (2017), in contrast, used the same ~~analysis tools but additionally~~ ensemble analysis tools for whole Antarctica, and varied the vertical viscoelastic profiles of the Earth within a gravitationally self-consistent coupled Earth-sea level model. They found only little difference in simulated glacial to modern ice volumes for different viscosity profiles bounded between 1×10^{19} Pa s and 5×10^{21} Pa s. Briggs et al. (2014) have not varied visco-elastic Earth model components, assuming that the impact of for instance climatic forcing is more relevant.

~~Instead of a friction coefficient underneath the modern ice shelves we chose~~ For the basal sliding we decided on the sliding exponent ~~as uncertain parameter for the entire~~ PPQ as uncertain ensemble parameter. A value of 0 corresponds to Coulomb friction as used in the PSU-ISM simulations, while ANICE used a value of 0.3 (Maris et al., 2014) and Quiquet et al. (2018) a linear scaling (1.0). Interestingly, we find best scores for rather high sliding exponents of PPQ with values of 0.75 or 1.0 (rather linear relationship of sliding velocity and till strength).

Briggs et al. (2013) used Coulomb friction and varied instead three parameters that control the basal sliding over soft and hard beds, based on a erosion parameterization. In our study, the till weakness is associated with the till friction angle, which is optimized for the present-day grounded Antarctic Ice Sheet ~~. Regarding sensitivity of the simulated ice volumes for variation of the minimum till friction angle~~ (Pollard and DeConto, 2012b) . In Briggs et al. (2013) , basal sliding additionally accounts for basal roughness and pinnings points (three parameters), which is otherwise underestimated as a result of the coarse model resolution.

Another sliding-related key parameter is the friction coefficient underneath the modern ice shelves ~~see our discussion in the companion paper and the Appendix 4.~~ as varied in (Pollard et al., 2016, 2017) , who found it to be the most dominant ensemble parameter. As discussed in the companion paper (Albrecht et al., 2019) , we also find till properties in the ice shelf regions highly relevant, in particular during deglaciation. As a consequence, we have run an additional ensemble analysis for four parameters associated with basal sliding and hydrology, including friction underneath modern ice shelves and discussed the results in the Supplementray Material A. In the best fit simulations of this “basal ensemble”, we find main deglacial retreat occurring a few thousand years earlier (closer to MWP-1A) than in the base ensemble. Hence, 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 of

$\leq 1°$.

In their ensemble analysis Pollard et al. (2016) included an iceberg calving parameterFor a representation of the ice dynamical uncertainty we chose the ESIA enhancement factor as most relevant ensemble parameter, which mainly affects the grounded ice volume. We find best fits for rather small ESIA of 1–2, while for larger values the modeled EAIS ice thickness underestimates modern observations. Pollard et al. (2016) did not vary enhancement factors in their ensemble and used a rather small enhancement factor of 1 for the SIA, while the value for the SSA enhancement was prescribed to a very low value of 0.3 (Pollard and DeConto, 2012a) . Briggs et al. (2014) varied enhancement factors for both the SIA and SSA in their large ensemble, and determined a rather large reference value of 4.8 for SIA enhancement and a reference value for SSA enhancement close to 0.6 (see Table 1 Briggs et al., 2013) , which we have used in all our ensemble simulations. Maris et al. (2014) determined in their sensitivity study for the SIA enhancement an even larger reference value of 9 and for the SSA enhancement 0.8. In Quiquet et al. (2018) best fits to present-day thickness are found for SIA enhancement between 1.5 and 4, for SSA enhancement between 0.2 and 0.5 respectively.

As climate-related uncertain ensemble parameter we chose a parameter associated with the change of precipitation with temperatures, PREC. The best-fit parameter values of PREC = 7–10 %/K yield for 10 K colder glacial temperatures about 50-65 % less precipitation. This parameter is similar to the insolation scaling parameter in Briggs et al. (2013) , where the best fit value would result in about 60% less precipitation at insolation minimum. In total, Briggs et al. (2014) varied seven precipitation-related parameters based on three different precipitation forcings (one of which is similar to the one we used). Maris et al. (2014) used instead a linear temperature-based scaling between LGM and present-day surface mass balance (with about 58% anomaly) with a fixed parameter.

Beyond the four parameters varied in our ensemble, previous ensemble studies found for instance a high sensitivity in at least one of the five temperature related parameters (Briggs et al., 2013) . In contrast, we found only little effect of temperature on the sea-level relevant ice sheet volume, as discussed in the companion paper (Albrecht et al., 2019) . Concerning iceberg calving, Pollard et al. (2016) included one related parameter in their ensemble analysis, while Briggs et al. (2013) varied three parameters for ice shelf calving and one parameter for tidewater calving. Our ‘eigencalving’ model provides a fair parameterization also uses a strain-rate based calving estimate, combined with a minimal terminal ice thickness and provides a representation of calving front dynamics, which in first order yields calving front locations close to present observations (Levermann et al., 2012) . As this parameterization is rather independent of the climate conditions(Levermann et al., 2012) . Variations , variations of the ‘eigencalving’ parameter show only little effect on sea-level relevant ice volume (see companion

~~paper). Pollard et al. (2016) also included an uncertain parameter for ice shelf melt~~ Albrecht et al. (2019)).

Regarding sub-shelf melting, Pollard et al. (2016) and Quiquet et al. (2018) included one uncertain parameter in their analysis, while Briggs et al. (2013) even varied four melt-related parameters. 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~~ (Reese et al., 2018) . the two key PICO parameters have been constrained for present observations, so that we have preferred other less constrained parameters in our ensemble ~~,~~ that are more relevant for the ice volume history of the ~~eastern part~~ Antarctic Ice Sheet.

The four selected ensemble parameters, representing uncertainties in interacting ice-Earth dynamics, basal sliding as well climate conditions, imply a large range of uncertainty for the total Antarctic ice volume change. They have been chosen, such that the model yields a present-day ice volume close to observations, while the LGM ice volume differs significantly for parameter change. The probed parameter range has been chosen rather wide, which implies a low sampling density of the parameter space. With the knowledge gained in this ensemble analysis, this range could be further constrained in a (larger) sub-ensemble. Also, other parameters may be more relevant for certain regions of the Antarctic Ice Sheet ~~, namely ESIA and PREC (see Sect. 2.1).~~ or for the onset and rate of the last deglaciation, which in our ensemble occurs generally later than suggested by many paleo records. A closer look into the details of the deglacial period and relevant parameters will be discussed in a separate follow-up study.

One deficiency of our model settings is the general underestimation of ice thickness in the inner ice sheet sections (up to -500 m, mainly in the EAIS, which could be a result of the underestimated RACMO precipitation) and an overestimation in the outer terminal regions and at Siple Coast (up to +500 m), where the complex topography is not sufficiently resolved in the model with implications for inferred basal conditions and temperature conditions. Accordingly, we find a considerable misfit to most paleo elevation data (ELEV), which are located mainly in the marginal mountain regions. This could be improved, e.g. by parameterized basal roughness or erosion, as proposed in Briggs et al. (2013) , or by higher model resolution and updated bed topography data sets[2].

The score aggregation scheme according to Pollard et al. (2016) implies that the paleo data types have equal influence as the present-day constraints, although they cover only relatively small regions and periods of the modeled ice sheet history (Briggs and Tarasov, 2013) . However, as the variability in paleo misfits is comparably low among the ensemble, these data types have only relatively small imprint on the aggregated score (see more details in Supplementary Material B). This is also valid

[2]https://sites.uci.edu/morlighem/bedmachine-antarctica/

for a data-type weighted score (Briggs and Tarasov, 2013) , which applied to our ensemble results yields a similar set of best score runs.

Further work will consist in the determination of more realistic climate reconstruction using general circulation model results and in the explicit computation of the local relative sea-level, which could potentially have an strong impact on grounding line migration for glacial cycles (Gomez et al., 2013) .

4 Conclusions

We have run ~~a Large Ensemble of~~ 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 Large Ensemble~~ this kind of ensemble method is limited to a comparably small number of values for each parameter and hence the retrieved scores are somewhat blocky (as compared to advanced techniques that can interpolate in parameter space) we still recognize a general pattern ~~and parameter combination clusters~~ of parameter combinations that provide best model fits to both present-day observations and paleo records. However, the selected ensemble parameters certainly can not cover the full range of possible model response, in particular with regard to the self-amplifying effects during deglaciation.

For the four sampled parameters, best fits are found for ~~comparably small~~ mantle viscosity around VISC = ~~$5–25\times10^{20}$~~ $0.5–2.5\times10^{21}$ Pa s, rather linear relationships between sliding speed and till strength (with exponents PPQ = 0.75–1.0), no or only small enhancement of the SIA derived flow speed (with ESIA = 1–2) and for rather high rates of relative precipitation change with temperature forcing (PREC > 5 %/K). The five best-score ensemble members fall within this range. In comparison to the best-fit member (VISC = ~~5×10^{20}~~ 0.5×10^{21} Pa s, PREC = 7 %/K, PPQ = 0.75, ESIA = 2) slightly more sliding (PPQ = 1) or slower ice flow (ESIA = 1) can compensate for relatively dry climate conditions in colder climates for ~~high~~ higher PREC values, which is associated with smaller ice volumes and hence smaller driving stresses. Strongest effects of varying ESIA and PREC parameters are found for the much larger East Antarctic Ice Sheet volume, while PPQ and VISC have most pronounced effects ~~in~~ for the West Antarctic Ice Sheet dynamics in terms of grounding line migration and induced changes in ice loading.

Grounding line extends at ~~Last Glacial maximum for nearly all simulations~~ last glacial maximum to the edge of the continental shelf for nearly all simulations. The onset and rate of deglaciation, however is very sensitive to the choice of parameters and boundary conditions, in particular those related to basal sliding. 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 reconstructed ice sheet histories than exact numbers, which would require a much larger ensemble with an extended number of varied parameters.

The score-weighted likely range (one standard deviation) of our reconstructed ice volume histories suggest that a contribution of the Antarctic Ice Sheet has contributed to the global mean sea level since the Last Glacial Maximum at around 15 kyr BP of 9.4 $\pm$ 4.1 m SLE (6.5 $\pm$ 2.0 $\times 10^6$ km^3)to the global mean sea level since the Last Glacial Maximum at around 15kyr BP and . The ensemble-mean ice volume anomaly between LGM and present is therewith slightly higher than in most recently published studies. The choice of basal sliding parameterization in the different models seems to have most impact on the corresponding estimates. The ensemble reproduces the observed present-day grounded ice volume with an score-weighted anomaly of 0.6 $\pm$ 3.5 m SLE (0.7 $\pm$ 1.7 $\times 10^6$ km^3) . 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. Our reconstructed and hence serves as a suitable inital state for future projections.

The reconstructed score-weighted ensemble range (1σ) is comparably large with up to 4.3 m SLE (or 2.0 $\times 10^6$ km^3), which can be explained with by a high model sensitivity (see companion paper), (Albrecht et al., 2019) , by a comparably large range of sampled parameters the sampled parameter values and of course due to the choice of the aggregated score scheme(the unweighted ensemble range would be up to 5.4m SLE). A similar large range of is found in Briggs et al. (2014) with . By using "sieve" criteria the ensemble range can be reduced. For the much larger ensemble study by covering 31 parameters Briggs et al. (2014) a narrowed ensemble range of 4.4 mESL (different definition of sea-level equivalent volume change) or 1.8 $\times 10^6$ km^3) but for a different definition of volume change. was found for the best 5% of the ensemble simulations, which is close to the range of our study.

The onset of deglaciation and hence major grounding line retreat occurs in our model simulations after 12 kyr BP and hence considerably well after MWP1a ($\approx$14.3 kyr BP). Previous studies with PISM Golledge et al. (2014) suggest that the A previous PISM study simulated much earlier and larger sea-level contributions from Antarctica for oceanic forcing at intermediate levels can be of opposite sign as compare that is anticorrelated to the surface forcing, as temperature forcing (Golledge et al., 2014) , as likely happened during the two millennia of Antarctic Cold Reversal following the MWP1a.

The PISM model results in Kingslake et al. (2018) are based on this study ensemble study, but have been published before with an a slightly older model version (see data and model code availability therein). Meanwhile, we have improved the bedrock Earth model, which accounts for changes in the ocean water column induced by variations in in bed topography or sea-level changes. Regarding the grounding line migration along a trough through In contrast to Kingslake et al. (2018) , we used the remapped topography without local adjustment in the region of Bungenstock ice rise in the Weddell Sea sectorwe found that Bungenstock Ice Rise is a key region and in only , and found in about 20% of the score-weighted simulations this region regrounded. In this study we used the Bedmap2

~~topography remapped to 16km resolution without local adjustments. Our~~ a extensive retreat of the grounding line and subsequent re-advance in both the Ross and Weddell Sea sector.

The here presented paleo simulation ensemble analysis with PISM provides ~~model and observation calibrated parameter constraints~~ a set of data-constrained parameter combinations, that can be used as a reference for further sensitivity studies investigating specific episodes in the history of Antarctica, such as the the last deglaciation or the Last Interglacial, as well as for projections of Antarctic sea-level contributions. ~~With the best-fit simulation parameters we have participated in the initMIP-Antarctica model intercomparison (Seroussi et al., 2019, PISMPAL3) .~~

Author contributions

TA designed, ran and analyzed the ice sheet model experiments; TA and RW co-developed PISM and implemented processes relevant for application to the Antarctic Ice Sheet. RW and AL contributed to the interpretation of the results.

Code and data availability

The PISM code used in this study can be obtained from https://github.com/talbrecht/pism_pik/tree/pism_pik_1.0 and will be published with DOI ~~link~~reference. Results and plotting scripts are available upon request and will be published in www.PANGAEA.de. For now see jupyter notebook https://nbviewer.jupyter.org/url/www.pik-potsdam.de/~albrecht/notebooks/paleo_paper_final2.ipynb. PISM input data are preprocessed using https://github.com/pism/pism-ais with original data citations.

Competing interests

The authors declare that they have no conflicts of interest.

Acknowledgements. Development of PISM is supported by NASA grant NNX17AG65G and NSF grants PLR-1603799 and PLR-1644277. Open-source software was used at all stages, in particular NumPy (www.numpy.org), CDO (https://code.mpimet.mpg.de/projects/cdo/), NCO (http://nco.sourceforge.net/) and matplotlib (https://matplotlib.org/). The authors gratefully acknowledge the European Regional Development Fund (ERDF), the German Federal Ministry of Education and Research and the Land Brandenburg for supporting this project by providing resources on the high performance computer system at the Potsdam Institute for Climate Impact Research. Computer resources for this project have been also provided by the Gauss Centre for Supercomputing/Leibniz Supercomputing Centre (www.lrz.de) under Project-ID pr94ga and pn69ru. T.A. is supported by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the priority program "Antarctic Research with comparative investigations in Arctic ice areas" by grant LE1448/6-1 and LE1448/7-1. We thank Dave Pollard for sharing ensemble analysis scripts and for valueable discussions and Stewart Jamieson for sharing gridded RAISED datasets. Finally, we appreciate the suggestions and comments by Lev Tarasov and an anonymous reviewer, which led us to considerable improvements of the manuscript.

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 Discussions, 2019, doi:10.5194/tc-2019-71, https://www.the-cryosphere-discuss.net/tc-2019-71/, 2019.

Argus, D. F., Peltier, W., Drummond, R., and Moore, A. W.: The Antarctica component of postglacial rebound model ICE-6G_C (VM5a) based on GPS positioning, exposure age dating of ice thicknesses, and relative sea level histories, Geophysical Journal International, 198, 537–563, 2014.

Aschwanden, A. and Blatter, H.: Mathematical modeling and numerical simulation of polythermal glaciers, Journal of Geophysical Research: Earth Surface, 114, 2009.

Aschwanden, A., Bueler, E., Khroulev, C., and Blatter, H.: An enthalpy formulation for glaciers and ice sheets, Journal of Glaciology, 58, 441–457, doi:10.3189/2012JoG11J088, 2012.

Barletta, V. R., Bevis, M., Smith, B. E., Wilson, T., Brown, A., Bordoni, A., Willis, M., Khan, S. A., Rovira-Navarro, M., Dalziel, I., et al.: Observed rapid bedrock uplift in Amundsen Sea Embayment promotes ice-sheet stability, Science, 360, 1335–1339, 2018.

Bart, P. J., Mullally, D., and Golledge, N. R.: The influence of continental shelf bathymetry on Antarctic Ice Sheet response to climate forcing, Global and Planetary Change, 142, 87–95, 2016.

Bart, P. J., DeCesare, M., Rosenheim, B. E., Majewski, W., and McGlannan, A.: A centuries-long delay between a paleo-ice-shelf collapse and grounding-line retreat in the Whales Deep Basin, eastern Ross Sea, Antarctica, Scientific reports, 8, 12 392, 2018.

Bentley, M. J., Ó Cofaigh, C., Anderson, J. B., Conway, H., Davies, B., Graham, A. G. C., Hillenbrand, C.-D., Hodgson, D. A., Jamieson, S. S. R., Larter, R. D., Mackintosh, A., Smith, J. A., Verleyen, E., Ackert, R. P., Bart, P. J., Berg, S., Brunstein, D., Canals, M., Colhoun, E. A., Crosta, X., Dickens, W. A., Domack, E., Dowdeswell, J. A., Dunbar, R., Ehrmann, W., Evans, J., Favier, V., Fink, D., Fogwill, C. J., Glasser, N. F., Gohl, K., Golledge, N. R., Goodwin, I., Gore, D. B., Greenwood, S. L., Hall, B. L., Hall, K., Hedding, D. W., Hein, A. S., Hocking, E. P., Jakobsson, M., Johnson, J. S., Jomelli, V., Jones, R. S., Klages, J. P., Kristoffersen, Y., Kuhn, G., Leventer, A., Licht, K., Lilly, K., Lindow, J., Livingstone, S. J., Massé, G., McGlone, M. S., McKay, R. M., Melles, M., Miura, H., Mulvaney, R., Nel, W., Nitsche, F. O., O'Brien, P. E., Post, A. L., Roberts, S. J., Saunders, K. M., Selkirk, P. M., Simms, A. R., Spiegel, C., Stolldorf, T. D., Sugden, D. E., van der Putten, N., van Ommen, T., Verfaillie, D., Vyverman, W., Wagner, B., White, D. A., Witus, A. E., and Zwartz, D.: A community-based geological reconstruction of Antarctic Ice Sheet deglaciation since the Last Glacial Maximum, Quaternary Science Reviews, 100, 1–9, doi:10.1016/j.quascirev.2014.06.025, 2014.

Briggs, R., Pollard, D., and Tarasov, L.: A glacial systems model configured for large ensemble analysis of Antarctic deglaciation, The Cryosphere, 7, 1949–1970, 2013.

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

Briggs, R. D., Pollard, D., and Tarasov, L.: A data-constrained large ensemble analysis of Antarctic evolution since the Eemian, Quaternary Science Reviews, 103, 91–115, doi:10.1016/j.quascirev.2014.09.003, 2014.

Bueler, E. and Brown, J.: Shallow shelf approximation as a "sliding law" in a thermomechanically coupled ice sheet model, Journal of Geophysical Research, 114, doi:10.1029/2008JF001179, 2009.

Bueler, E. and van Pelt, W.: Mass-conserving subglacial hydrology in the Parallel Ice Sheet Model version 0.6, Geoscientific Model Development, 8, 1613, 2015.

Bueler, E., Lingle, C. S., and Brown, J.: Fast computation of a viscoelastic deformable Earth model for ice-sheet simulations, Annals of Glaciology, 46, 97–105, 2007.

Cuffey, K. M. and Marshall, S. J.: Substantial contribution to sea-level rise during the last interglacial from the Greenland ice sheet, Nature, 404, 591, 2000.

Cuffey, K. M. and Paterson, W. S. B.: The physics of glaciers, Academic Press, 2010.

Cuffey, K. M., Clow, G. D., Steig, E. J., Buizert, C., Fudge, T., Koutnik, M., Waddington, E. D., Alley, R. B., and Severinghaus, J. P.: Deglacial temperature history of West Antarctica, Proceedings of the National Academy of Sciences, 113, 14 249–14 254, http://www.usap-dc.org/view/dataset/600377, 2016.

De Boer, B., Van de Wal, R., Lourens, L., Bintanja, R., and Reerink, T.: A continuous simulation of global ice volume over the past 1 million years with 3-D ice-sheet models, Climate Dynamics, 41, 1365–1384, 2013.

Fogwill, C., Turney, C., Golledge, N., Etheridge, D., Rubino, M., Thornton, D., Baker, A., Woodward, J., Winter, K., Van Ommen, T., et al.: Antarctic ice sheet discharge driven by atmosphere-ocean feedbacks at the Last Glacial Termination, Scientific reports, 7, 39 979, 2017.

Fretwell, P., Pritchard, H. D., Vaughan, D. G., Bamber, J. L., Barrand, N. E., Bell, R., Bianchi, C., Bingham, R. G., Blankenship, D. D., Casassa, G., Catania, G., Callens, D., Conway, H., Cook, A. J., Corr, H. F. J., Damaske, D., Damm, V., Ferraccioli, F., Forsberg, R., Fujita, S., Gim, Y., Gogineni, P., Griggs, J. A., Hindmarsh, R. C. A., Holmlund, P., Holt, J. W., Jacobel, R. W., Jenkins, A., Jokat, W., Jordan, T., King, E. C., Kohler, J., Krabill, W., Riger-Kusk, M., Langley, K. A., Leitchenkov, G., Leuschen, C., Luyendyk, B. P., Matsuoka, K., Mouginot, J., Nitsche, F. O., Nogi, Y., Nost, O. A., Popov, S. V., Rignot, E., Rippin, D. M., Rivera, A., Roberts, J., Ross, N., Siegert, M. J., Smith, A. M., Steinhage, D., Studinger, M., Sun, B., Tinto, B. K., Welch, B. C., Wilson, D., Young, D. A., Xiangbin, C., and Zirizzotti, A.: Bedmap2: improved ice bed, surface and thickness datasets for Antarctica, The Cryosphere, 7, 375–393, doi:10.5194/tc-7-375-2013, https://secure.antarctica.ac.uk/data/bedmap2/, 2013.

Frieler, K., Clark, P. U., He, F., Buizert, C., Reese, R., Ligtenberg, S. R. M., van den Broeke, M. R., Winkelmann, R., and Levermann, A.: Consistent evidence of increasing Antarctic accumulation with warming, Nature Climate Change, 5, 348–352, doi:10.1038/nclimate2574, 2015.

Golledge, N. R., Fogwill, C. J., Mackintosh, A. N., and Buckley, K. M.: Dynamics of the last glacial maximum Antarctic ice-sheet and its response to ocean forcing, Proceedings of the National Academy of Sciences, 109, 16 052–16 056, 2012.

Golledge, N. R., Levy, R. H., McKay, R. M., Fogwill, C. J., White, D. A., Graham, A. G., Smith, J. A., Hillenbrand, C.-D., Licht, K. J., Denton, G. H., et al.: Glaciology and geological signature of the Last Glacial Maximum Antarctic ice sheet, Quaternary Science Reviews, 78, 225–247, 2013.

Golledge, N. R., Menviel, L., Carter, L., Fogwill, C. J., England, M. H., Cortese, G., and Levy, R. H.: Antarctic contribution to meltwater pulse 1A from reduced Southern Ocean overturning, Nature Communications, 5, doi:10.1038/ncomms6107, 00000, 2014.

Gomez, N., Pollard, D., and Mitrovica, J. X.: A 3-D coupled ice sheet–sea level model applied to Antarctica through the last 40 ky, Earth and Planetary Science Letters, 384, 88–99, 2013.

Jenkins, A., Shoosmith, D., Dutrieux, P., Jacobs, S., Kim, T. W., Lee, S. H., Ha, H. K., and Stammerjohn, S.: West Antarctic Ice Sheet retreat in the Amundsen Sea driven by decadal oceanic variability, Nature Geoscience, 11, 733–738, 2018.

Jouzel, J., Masson-Delmotte, V., Cattani, O., Dreyfus, G., Falourd, S., Hoffmann, G., Minster, B., Nouet, J., Barnola, J. M., Chappellaz, J., Fischer, H., Gallet, J. C., Johnsen, S., Leuenberger, M., Loulergue, L., Luethi, D., Oerter, H., Parrenin, F., Raisbeck, G., Raynaud, D., Schilt, A., Schwander, J., Selmo, E., Souchez, R., Spahni, R., Stauffer, B., Steffensen, J. P., Stenni, B., Stocker, T. F., Tison, J. L., Werner, M., and Wolff, E. W.: Orbital and Millennial Antarctic Climate Variability over the Past 800,000 Years, Science, 317, 793–796, doi:10.1126/science.1141038, 2007.

Kingslake, J., Scherer, R., Albrecht, T., Coenen, J., Powell, R., Reese, R., Stansell, N., Tulaczyk, S., Wearing, M., and Whitehouse, P.: Extensive retreat and re-advance of the West Antarctic ice sheet during the Holocene, Nature, 558, 430, 2018.

Kopp, R. E., Simons, F. J., Mitrovica, J. X., Maloof, A. C., and Oppenheimer, M.: Probabilistic assessment of sea level during the last interglacial stage, Nature, 462, 863–867, 2009.

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

Lingle, C. S. and Clark, J. A.: 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, 1100–1114, 1985.

Liu, J., Milne, G. A., Kopp, R. E., Clark, P. U., and Shennan, I.: Sea-level constraints on the amplitude and source distribution of Meltwater Pulse 1A, Nature Geoscience, 9, 130–134, doi:10.1038/ngeo2616, 2016.

Maris, M., De Boer, B., Ligtenberg, S., Crucifix, M., Van De Berg, W., and Oerlemans, J.: Modelling the evolution of the Antarctic ice sheet since the last interglacial, The Cryosphere, 8, 1347–1360, 2014.

Maris, M., Van Wessem, J., Van De Berg, W., De Boer, B., and Oerlemans, J.: A model study of the effect of climate and sea-level change on the evolution of the Antarctic Ice Sheet from the Last Glacial Maximum to 2100, Climate dynamics, 45, 837–851, 2015.

Martos, Y. M., Catalán, M., Jordan, T. A., Golynsky, A., Golynsky, D., Eagles, G., and Vaughan, D. G.: Heat flux distribution of Antarctica unveiled, Geophysical Research Letters, 44, 2017.

Mercer, J. H.: West Antarctic ice sheet and CO2 greenhouse effect: a threat of disaster, Nature, 271, 321, 1978.

Morland, L. and Johnson, I.: Steady motion of ice sheets, Journal of Glaciology, 25, 229–246, 1980.

Oppenheimer, M. and Alley, R. B.: How high will the seas rise?, Science, 354, 1375–1377, 2016.

Pattyn, F.: The paradigm shift in Antarctic ice sheet modelling, Nature communications, 9, 2728, 2018.

Peltier, W., Argus, D., and Drummond, R.: Space geodesy constrains ice age terminal deglaciation: The global ICE-6G_C (VM5a) model, Journal of Geophysical Research: Solid Earth, 120, 450–487, 2015.

Pollard, D. and DeConto, R.: Description of a hybrid ice sheet-shelf model, and application to Antarctica, Geoscientific Model Development, 5, 1273, 2012a.

Pollard, D. and DeConto, R. M.: A simple inverse method for the distribution of basal sliding coefficients under ice sheets, applied to Antarctica, The Cryosphere, 6, 953–971, doi:10.5194/tc-6-953-2012, http://www.the-cryosphere.net/6/953/2012/, 2012b.

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., and Deconto, R. M.: 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, 2124–2138, 2017.

Quiquet, A., Punge, H. J., Ritz, C., Fettweis, X., Gallée, H., Kageyama, M., Krinner, G., Salas y Mélia, D., and Sjolte, J.: Sensitivity of a Greenland ice sheet model to atmospheric forcing fields, The Cryosphere, 6, 999–1018, doi:10.5194/tc-6-999-2012, https://www.the-cryosphere.net/6/999/2012/, 2012.

Quiquet, A., Dumas, C., Ritz, C., Peyaud, V., and Roche, D. M.: The GRISLI ice sheet model (version 2.0): calibration and validation for multi-millennial changes of the Antarctic ice sheet, Geoscientific Model Development, 11, 5003, 2018.

Reeh, N.: Parameterization of melt rate and surface temperature in the Greenland ice sheet, Polarforschung, 59, 113–128, 1991.

Reese, R., Albrecht, T., Mengel, M., Asay-Davis, X., and Winkelmann, R.: Antarctic sub-shelf melt rates via PICO, The Cryosphere, 12, 1969, 2018.

Rignot, E., Mouginot, J., and Scheuchl, B.: Ice Flow of the Antartctic Ice Sheet, Science, 333, 1423–1427, doi:10.1126/science.1208336, 2011.

Ritz, C., Fabre, A., and Letréguilly, A.: Sensitivity of a Greenland ice sheet model to ice flow and ablation parameters: consequences for the evolution through the last climatic cycle, Climate Dynamics, 13, 11–23, 1996.

Rutt, I. C., Hagdorn, M., Hulton, N., and Payne, A.: The Glimmer community ice sheet model, Journal of Geophysical Research: Earth Surface, 114, 2009.

Schmidtko, S., Heywood, K. J., Thompson, A. F., and Aoki, S.: Multidecadal warming of Antarctic waters, Science, 346, 1227–1231, 2014.

Schoof, C.: Ice sheet grounding line dynamics: Steady states, stability, and hysteresis, Journal of Geophysical Research, 112, doi:10.1029/2006JF000664, http://doi.wiley.com/10.1029/2006JF000664, 2007.

Seroussi, H., Nowicki, S., Simon, E., Abe Ouchi, A., Albrecht, T., Brondex, J., Cornford, S., Dumas, C., Gillet-Chaulet, F., Goelzer, H., Golledge, N. R., Gregory, J. M., Greve, R., Hoffman, M. J., Humbert, A., Huybrechts, P., Kleiner, T., Larour, E., Leguy, G., Lipscomb, W. H., Lowry, D., Mengel, M., Morlighem, M., Pattyn, F., Payne, A. J., Pollard, D., Price, S., Quiquet, A., Reerink, T., Reese, R., Rodehacke, C. B., Schlegel, N.-J., Shepherd, A., Sun, S., Sutter, J., Van Breedam, J., van de Wal, R. S. W., Winkelmann, R., and Zhang, T.: initMIP-Antarctica: An ice sheet model initialization experiment of ISMIP6, The Cryosphere Discussions, 2019, 1–44, doi:10.5194/tc-2018-271, https://www.the-cryosphere-discuss.net/tc-2018-271/, 2019.

Shepherd, A. and Nowicki, S.: Improvements in ice-sheet sea-level projections, Nature Climate Change, 7, 672, 2017.

Shepherd, A., Fricker, H. A., and Farrell, S. L.: Trends and connections across the Antarctic cryosphere, Nature, 558, 223, 2018a.

Shepherd, A., Ivins, E., Rignot, E., Smith, B., van den Broeke, M., Velicogna, I., Whitehouse, P., Briggs, K., Joughin, I., Krinner, G., et al.: Mass balance of the Antarctic Ice Sheet from 1992 to 2017, Nature, 556, pages219–222, 2018b.

Siegert, M. J., Kingslake, J., Ross, N., Whitehouse, P. L., Woodward, J., Jamieson, S. S., Bentley, M. J., Winter, K., Wearing, M., Hein, A. S., et al.: Major ice-sheet change in the Weddell Sector of West Antarctica over the last 5000 years, Reviews of Geophysics, 2019.

Simmons, A.: ERA-Interim: New ECMWF reanalysis products from 1989 onwards, ECMWF newsletter, 110, 25–36, 2006.

Slangen, A., Adloff, F., Jevrejeva, S., Leclercq, P., Marzeion, B., Wada, Y., and Winkelmann, R.: A review of recent updates of sea-level projections at global and regional scales, Surveys in Geophysics, 38, 385–406, 2017.

Spector, P., Stone, J., Cowdery, S. G., Hall, B., Conway, H., and Bromley, G.: Rapid early-Holocene deglaciation in the Ross Sea, Antarctica, Geophysical Research Letters, 44, 7817–7825, 2017.

Stuhne, G. and Peltier, W.: Assimilating the ICE-6G_C Reconstruction of the Latest Quaternary Ice Age Cycle Into Numerical Simulations of the Laurentide and Fennoscandian Ice Sheets, Journal of Geophysical Research: Earth Surface, 122, 2324–2347, 2017.

Stuhne, G. R. and Peltier, W. R.: Reconciling the ICE-6G_C reconstruction of glacial chronology with ice sheet dynamics: The cases of Greenland and Antarctica, Journal of Geophysical Research: Earth Surface, 120, 2015JF003 580, doi:10.1002/2015JF003580, 2015.

Sutter, J., Gierz, P., Grosfeld, K., Thoma, M., and Lohmann, G.: Ocean temperature thresholds for Last Interglacial West Antarctic Ice Sheet collapse, Geophysical Research Letters, 43, 2675–2682, 2016.

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

The PISM authors: PISM, a Parallel Ice Sheet Model: User's Manual, http://pism-docs.org/sphinx/, based on revision stable v.1.0 edn., 2017.

Weber, M. E., Clark, P. U., Kuhn, G., Timmermann, A., Sprenk, D., Gladstone, R., Zhang, X., Lohmann, G., Menviel, L., Chikamoto, M. O., Friedrich, T., and Ohlwein, C.: Millennial-scale variability in Antarctic ice-sheet discharge during the last deglaciation, Nature, 510, 134–138, doi:10.1038/nature13397, 00002, 2014.

Weertman, J.: Stability of the junction of an ice sheet and an ice shelf, Journal of Glaciology, 13, 3–11, 1974.

Wessem, J. M. v., Berg, W. J. v. d., Noël, B. P., Meijgaard, E. v., Amory, C., Birnbaum, G., Jakobs, C. L., Krüger, K., Lenaerts, J., Lhermitte, S., et al.: Modelling the climate and surface mass balance of polar ice sheets using RACMO2–Part 2: Antarctica (1979–2016), The Cryosphere, 12, 1479–1498, 2018.

Whitehouse, P. L., Bentley, M. J., and Le Brocq, A. M.: 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, 2012a.

Whitehouse, P. L., Bentley, M. J., Milne, G. A., King, M. A., and Thomas, I. D.: 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, 1464–1482, 2012b.

Winkelmann, R., Martin, M. A., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C., and Levermann, A.: The Potsdam Parallel Ice Sheet Model (PISM-PIK) – Part 1: Model description, The Cryosphere, 5, 715–726, doi:10.5194/tc-5-715-2011, 2011.

Zwally, H., Li, J., Robbins, J., Saba, J., Yi, D., and Brenner, A.: Mass gains of the Antarctic ice sheet exceed losses, Journal of Glaciology, 61, doi:doi: 10.3189/2015JoG15J071, http://www.ingentaconnect.com/content/igsoc/jog/pre-prints/content-ings_jog_15j071, 2015.

Supplementary Material A: Ensemble of basal parameters

In the sensitivity analysis of parameters in a companion paper various parameters and boundary conditions in a companion paper (Albrecht et al., 2019) , we found that the basal sliding parameterization in conjunction with the sub-glacial hydrology scheme show very diverse simulated ice volume histories for a plausible range of unconfined parameter values. We have chosen the parameter PPQ (Sect. 3.1) as only representative of basal processes uncertainties for the large ensemble analysis.

Here we want to span a sub-ensemble including three other relevant basal parameters. The four parameters and sampled values used in the sub-ensemble basal ensemble analysis are:

- PHIMIN: Minimal till friction angle on the continental shelf, mainly underneath modern ice shelves, where sandy sediments are prevalent (friction coefficient on the continental shelf has been chosen as one of the ensemble parameters in Pollard et al. (2016, 2017)). The tangens of till friction angle enters the Mohr-Coulomb-yield stress criterion. Sampled values are $0.5°$, $1°$, $2°$ and $3°$ $0.5°$, $1°$, $2°$ and $3°$.

- TWDR: The decay rate of the effective water content within the till layer using the non-conserving hydrology model, while basal melt adds water up to a certain threshold. Sampled value are 0.5mm/yr (1.55×10^{-11}m/s), 1mm/yr (3.1×10^{-11}m/s), 5mm/yr (15.5×10^{-11}m/s) and 10mm/yr (31×10^{-11}m/s)0.5 mm/yr (1.55×10^{-11}m/s), 1 mm/yr (3.1×10^{-11}m/s), 5 mm/yr (15.5×10^{-11}m/s) and 10 mm

- FEOP: For this fraction of the effective overburden pressure (for details see Bueler and van Pelt, 2015, Sect. 3.2), excess water will be drained into a transport system in the case of saturated till. Sampled values are 1, 2, 4. 8and 321%, 2%, 4%, 8% and 32%.

- PPQ: as in the large ensemble (see Sect. 3.1)

---

figs/f_a01.pdf

Figure S1: Aggregated score for 318 ensemble members (4 model parameters, 4-5 values each) showing the distribution of the scores over the full range of plausible basal parameter values. The score values are computed versus geologic and modern data sets, normalized by the best score in the ensemble, and range from <0.01 (bright yellow, no skill) to 1 (dark red, best score) (cfs. Pollard et al., 2016, Figs. 2 + C1), on a logarithmic color scale. The four parameters are the effective overburden pressure fraction FEOP (outer y-axis), the minimal till friction angle on the continental shelf PHIMIN (outer x-axis), the tillwater decay rate TWDR (inner y-axis) and the power-law sliding pseudoplasticity exponent PPQ (inner x-axis). Only In the lowest row, only four ensemble scores are shown for 32% of effective overburden pressure fraction, just to ascertain that aggregated scores decline for larger FEOP.

In the basal sub-ensemble we find even better scores than for the best fit parameter combination in the ~~large~~ base ensemble (here no. 8102, see Fig. S1) ~~,~~ that covers also climatic, Earth and ice-internal parameters. Best scores are found in particular for smaller minimal till friction angles PHIMIN = 0.5–1°~~. Best scores are also found~~ , but also for rather high values of the fraction of the effective overburden pressure at which excess water drains, here FEOP = 4–16%. These values are higher than those used in the ~~large~~ base ensemble. 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 $\times 10^{-11}$ m/s), which agrees with the best fit ~~values of large ensemble~~ parameter combination of the base ensemble (PPQ=0.75 and TWDR=1 mm/yr). The LGM volume of the best fit simulation of the basal ensemble is similar to the best fit simulation of the ~~large~~ base ensemble (cf. Figs. S2 and 15), however deglacial retreat occurs a few thousand years earlier for lower PHIMIN.

figs/f_a02.pdf

Figure S2: Snapshots of grounded ice thickness anomaly to present-day observations (Bedmap2; Fretwell et al., 2013) over the last 15 kyr for best-fit simulation in the basal ensemble. At LGM state grounded ice extends towards the edge of the continental shelf, with much thicker ice than present mainly in West Antarctica. Retreat of the ice sheet initiates between 12 and 11 kyr BP and halts already latest 8 kyr in all large ice shelf basins of Ross, Weddell Sea, Amery and Amundsen Sea. East Antarctic Ice Sheet thickness is underestimated throughout the deglaciation period (light blue). Compare Fig. 2 in Golledge et al. (2014).

figs/f_a03.pdf

Figure S3: During deglaciation the score-weighted ensemble mean (green) shows most of the sea-level change rates between 14.5 kyr BP (MWP1a) and 8 kyr BP with mean rates around 1 mm yr^{-1}, while the best-score simulation (red) reveals rates of sea-level rise of up to 4 mm yr^{-1} (100 yr bins) in the same period (cf. Golledge et al., 2014, Fig. 3 d). In contrast to the Large Ensemble including climate and Earth model uncertainty base ensemble (cf. Fig. 9c) the basal ensemble shows a much earlier deglacial retreat and no regrowth during the late Holocene.

Development of PISM is supported by NASA grant NNX17AG65G and NSF grants PLR-1603799 and PLR-1644277.The authors gratefully acknowledge the European Regional Development Fund (ERDF) , the German Federal Ministry of Education and Research and the Land Brandenburg for supporting this project by providing resources on the high performance computer system at the Potsdam Institute for Climate Impact Research. Computer resources for this project have been also provided by the Gauss Centre for Supercomputing/Leibniz Supercomputing Centre (www.lrz. de) under Project-ID pr94ga and pn69ru.T. A. is supported by the Deutsche Forschungsgemeinschaft (DFG)in the framework of the priority program “Antarctic Research with comparative investigations in Arctic ice areas” by grant LE1448/6-1 and LE1448/7-1. We thank Dave Pollard for sharing ensembleanalysis scripts and for valueable discussions.

Supplementary Material B: Misfit to individual paleo data types

This appendix compares model results with corresponding geological data types (AntICEdat from Briggs and Tarasov (2013)) used in the ensemble scoring. This absolute misfit is important information as all scores are normalized against their median (relative fit) in order to calculate the aggregated scores. Thereby, we want to demonstrate how well the ensemble simulations span the data constraints and hence potentially represent reasonably realistic ice-sheet behavior.

Fig. S4 compares elevation vs. age for all 256 runs with cosmogenic data at 26 sites (ELEV; Briggs and Tarasov, 2013) with a median age of constraint of 9.6 kyr. We find a good fit in parts of East Antarctica (e.g. Framnes Mts. (1201-1203)) and in parts of the Ross sector (e.g. Clark Mts. (1405), Allegheny Mts. (1406) or Eastern Fosdick Mts. (1408)), while in the West Antarctic Ice Sheet there is quite a large spread among the ensemble misfit of up to 1,000 m in surface elevation, with ensemble mean misfits of up to 1,000 m. This is due to the fact that in many ensemble simulations the large ice shelves of Ronne-Filchner, Ross and Amery do not become afloat in time, while the best-fit simulation (green markers) shows quite a good fit, although some regions remain thicker than observed until present (Fig. 12).

Fig. S5 shows the misfit of simulated grounding lines retreat for all ensemble simulations at 27 marine core sites (EXT; Briggs and Tarasov, 2013) , which are relatively well distributed around the Antarcric Ice Sheet with a median age of 16.6 kyr, the oldest data point 30.7 kyr. Generally, simulated grounding line retreat occurs later than in most of the observations, less than 5 kyr near Victoria Land, Ross Sea and along the Antarctic Peninsula (2303, 2402-2403, 2602-2608) and less than 10 kyr in the Amundsen Sea, and Weddell Sea (2502, 2609, 2701). At some locations (Dronning Maud-Enderby Land (2101-2103) or at Victoria Land (2304)), however, the ensemble never reproduces the recorded open ocean conditions or grounding line retreat event, respectively.

Although not used as constraint in our scoring scheme, Fig. S7 shows the misfit of modelled relative sea level in all ensemble simulations with respect to 96 RSL proxy records at eight sites (RSL; Briggs and Tarasov, 2013) , with a median age of 5.0 kyr. The data for each site fall well within the overall model envelope (upper and lower bound indicated) with best fits at Syowa Coast (9101), Larsemann Hills (9201), Vestfold Hills (9202), Windmill Islands (9301), and Marguerite Bay (9601) and King George Island (9602), while in Victoria Land the model ensemble generally overestimates regional sea level (Terra Nova Bay (9401) and Southern Scott Coast (9402)).

From each data type misfit we obtain a ensemble distribution of misfits (Fig. S6), which can be rather normal (e.g. for EXT), exponential (e.g. TOTUPL) or long-tail (e.g. TOTDH). In order to calculate aggregated scores we normalize by the median value, which yields for most data types similar results as the mean value, except for TROUGH (34% difference). The corresponding variability of each of the resultant normalized scores hence contribute different skills to the aggregated score. Generally, grounding-line related (TOTE, TOTGL, THROUGH) and ice volume-related data-types (TOTDH) show similar individual score patterns (not shown here) with ensemble standard deviations of 0.1-0.2. In the aggregated score this patterns becomes even more pronounced, while paleo scores (ELEV and EXT) and ice shelf extent (TOTI) show only little variation (<0.1) among the ensemble, and hence only little effect in the aggregate score pattern.

figs/f_b01.png

Figure S4: ELEV observations (colored diamonds, dark and light blue indicate last 10 kyr or 20-10 kyr observational interval) taken from database by Briggs and Tarasov (2013) , ensemble results (black circles), upper and lower bounds from base ensemble (red triangles), and computed misfits (lower panel) for different Antarctic Peninsula sectors, indicated by vertically dashed lines and labels between panels. Green dots correspond to best-fit simulation. Compare to Fig. 7–9 for in Briggs et al. (2014) with same data-point identifiers.

figs/f_b02.pdf

Figure S5: EXT observations and ensemble results as in Fig. 10 in Briggs et al. (2014) . Black circles represent the 256 ensemble simulations with the best-fit simulation in green. Red indicate the grounding line retreat (GLR) two-way constraint types, magenta the open marine conditions (OMC) one-way constraint types. Dashed horizontal lines and associated labels segregate and identify the different sectors.

figs/f_b03.pdf

Figure S6: Histogram of misfits per data-type with median (in blue) and mean (green).

figs/f_b04.pdf

Figure S7: Regional sea level (RSL) data points and ensemble sea level curves for the 8 data sites, analogous to Fig. 5–6 in Briggs et al. (2014) , upper panels in EAIS, lower panels in Antarctic Peninsula and Ross sector. Observed RSL data points are colour coded according to the constraint they provide: two-way (light blue, dated past sea level); one-way lower-bounding (mauve, past sea level above or maximum age of beach) or one-way upper-bounding (orange, past sea level below or minimum age beach). For a detailed description of the RSL datasets and its processing, refer to Briggs and Tarasov (2013) . RSL has not been used as constraint in this study.