Reviewer 1

The work in this paper is motivated by the critical nature of climate change and its effect on the Greenland ice sheet. It is an interesting study presenting some very detailed work on improving a particular (ORCHIDEE) land surface model with the known techniques of parameter optimisation (for that model) by use of external, independent observations. These observations are the retrieved snow albedo product from MODIS. A regional model simulation produced output that was also used to compare with the ORCHIDEE model run with the different parameter sets. I think that the paper should be published, subject to some minor revisions to clarify the work for the readers.

Please see the attached pdf which comments on particular lines to aid the authors in making both small and larger changes. I have made some suggestions on where definitions and citations would be useful among other recommendations and critical comments. I hope that the authors find this useful.
We would like to thank the reviewer for taking the time to read and comment on our manuscript.

I made a comment about the accessibility of colour in one one figure (jet is not a good choice for communicating information). The following sites contain useful information, tips and tools on appropriate use of colour in illustrations:
https://www.nature.com/articles/s41467-020-19160-7
https://www.ascb.org/science-news/how-to-make-scientific-figures-accessible-to-readers-with-color-blindness
https://towardsdatascience.com/two-simple-steps-to-create-colorblind-friendly-data-visualizations-2ed781a167ec
In addition, here is very convincing talk about why the Python standard map is now "viridis" and why it is better than jet.
https://www.youtube.com/watch?v=xAoljeRJ3lU
Thank you for highlighting this issue, we have changed the colour of the plot to viridis as suggested.

In summary, my recommendation is for minor revisions to address issues of clarity in the communication of the work.
We thank the reviewer for this recommendation. We have implemented the suggestions from the PDF, which we agree improves the readability and clarity of the work.

Reviewer 2

OVERVIEW

This paper aims to improve simulating snow albedo by the ORCHIDEE land surface model via assimilation of MODIS snow albedo retrievals and parameters calibration through ORCHIDAS. The domain of study is the Greenland Ice Sheet (GrIS). ORCHIDEE is used in offline mode forced using atmospheric data coming from the MAR model. The optimisation is performed over three random years (2000, 2010 and 2012) taken over the period of study 2000-2017 while the improvement of snow albedo is checked over the whole period of study. An increased weight is given to pixels defining the edges of the GrIS compared to other pixels in order to improve albedo simulations over those areas crucial in term of mass loss. Moreover, a comparison between outputs of ORCHIDEE and MAR on surface mass balance, runoff and sublimation is carried out as well as a Morris sensitivity analysis of various parameters related to snow albedo and snow density relative to several outputs such as surface temperature, sensible and latent heat fluxes in addition to the previously mentioned outputs
We would like to thank the reviewer for taking the time to read the manuscript and comment on it, as well as for their faith in our ability to improve the paper for publication. We believe the different clarifications asked for and addition of an in situ evaluation will significantly strengthen the manuscript.

GENERAL COMMENTS

The paper covers the timely question of snow/ice albedo modelling over the GrIS in the context of increased melt in links with global warming. Here the chosen approach is to assimilate satellite derived albedo retrievals, which only few studies have performed before. As such, the novelty aspects of the paper are self evident. Therefore, such work would clearly deserve to be published. Nevertheless, I have some reservations about the methodology and I especially find the paper difficult to read. As indicated by one previous referee, “The article reads like a description of the research in the way it was conducted.”. Unfortunately, after the first revision, this problem still remains. My main issues with the paper are the following:

- The goal of the paper is straightforward: improve snow albedo by assimilation. But why do we want to do that? In which context? Global land surface modelling, if so, why using MAR instead of ERA5 as atmospheric forcing? Or are the targets are ice sheets and large glaciers? During all the several readings I made, I asked myself what were the purposes of such study and I could not find the answer anywhere in the introduction where normally one would expect to find such information.
  We acknowledge that we need to be more explicit in describing the context and motivations of our study focused on modelling the Greenland ice sheet, as well as our choice of the MAR forcing. Additions to the manuscript are described below in response to the relevant comments.

- The methodology clearly misses validation using independent datasets. Modelled albedos are compared with MODIS data that are partially assimilated (3 years over 18 years). Then ORCHIDEE outputs of surface mass balance, runoff and sublimation are compared with outputs from MAR, while ORCHIDEE is forced using atmospheric data coming from MAR itself. Obviously both model outputs are related, the main difference would come from the modelling approach of snow, ice and albedo in ORCHIDEE and MAR. As such, MAR cannot be considered as a reference as stated by the authors. I do, however, acknowledge that the authors raise some reservations about the limitation of the comparison in the paper.
  We acknowledge that evaluation against independent data is lacking. Therefore, we have added a section evaluation against in situ PROMICE data. This is further discussed in the validation section of the comments below.

- Several peripheral considerations (such as the test of the two minimisation approach, LBFGS-B and genetic algorithm) degrade the readability of the paper while providing very little novelty in terms of science.
  All mentions of L-BFGS-B have been removed from the manuscript for added clarity.

As I indicated beforehand, this paper cannot be published in The Cryosphere in its current version. Nevertheless, I do think the authors, who are renowned experts in their fields of expertise, have the ability to improve the paper to reach the publication stage. It would be unfortunate that such good science would, in the end, not be published. I list my comments and questions section by section below:

SPECIFIC COMMENTS

About the objectives of the paper:

- Please specify loudly the context and the purposes of the study in the introduction. Is the goal of the study to focus on ice sheets and large glaciers? If so, why? If not, do the authors focus on global or large scale climate simulations? If so, again why? Please provide also adequate references to justify your choices.
  Modelling the Greenland Ice Sheet is vital in understanding sea-level rise, amongst other factors impacted by a changing climate. Both dynamical effects and surface processes drive Greenland's evolution. However, recent studies show that surface mass balance changes dominate mass balance changes (The IMBIE team, 2020; Ryan et al., 2019; van den Broeke et al., 2016). To correctly represent the surface mass balance and its components (sublimation and runoff), it is important to simulate the physical processes within the snowpack. These depend on the surface energy balance and, therefore, on the albedo. The snow albedo model in ORCHIDEE has only ever been calibrated for vegetation and bare soil - never for glaciers or ice sheets.

  This work is part of ongoing work at LSCE to integrate an ice sheet model into ORCHIDEE. Currently, when used as part of the larger IPSL Earth System Model, the atmospheric component (LMDZ) treats the ice sheets through a rudimentary one-layer scheme and a fixed albedo. There is a desire for ORCHIDEE to model all types of continental surfaces. The modelling of the snow surface processes as it is done in ORCHIDEE would be a more suitable, coherent and explicit replacement to the original LMDZ scheme. This work will be documented fully in Charbit et al. (prep). Optimising parameters of the albedo of this snow model is a vital step in this development, given the importance of albedo, and this Raoult et al. paper discusses the techniques we can use to perform this parameter optimisation. Therefore this paper acts as a step in the development of the ORCHIDEE snow model with an application over the Greenland ice sheet. As discussed in the conclusions, ORCHIDAS is well adapted to the optimisation purpose, but, given the complexity of modelling the surface mass balance, more structural changes are needed to ORCHIDEE first, and further optimisations should consider more data in addition to the satellite albedo.

The second paragraph of the introduction has been expanded as follows:

> *Both dynamical effects and surface processes drive Greenland's evolution. However, recent studies show that surface mass balance changes dominate mass balance changes (The IMBIE team, 2020; Ryan et al., 2019; van den Broeke et al., 2016). To correctly represent the surface mass balance and its components (sublimation and runoff), it is important to simulate the physical processes within the snowpack. These depend on the surface energy balance and, therefore, on the albedo.* Given the importance of *this* albedo, it is crucial that it is accurately simulated in land surface models.

In addition, the final paragraph of the introduction has been expanded as follows:

> “Using MODIS snow albedo, in this study, we use DA for parameter estimation to improve the albedo parameterisation *over ice sheets* inside the ORCHIDEE LSM (Krinner et al., 2005). *While albedo parameters in ORCHIDEE have been optimised for vegetation and bare soil, this will be the first study optimising them for ice sheets. Our target area from this study is the Greenland Ice Sheet. This study is the first test of applying the ORCHIDEE data assimilation system over ice sheets to improve modelling albedo and, in turn, the surface mass balance of the ice sheet.* Instead…”

Reference:

Van den Broeke, Michiel R., et al. "On the recent contribution of the Greenland ice sheet to sea level change." *The Cryosphere* 10.5 (2016): 1933-1946.

The IMBIE team, "Mass balance of the Greenland Ice Sheet from 1992 to 2018." *Nature* 579, no. 7798 (2020): 233-239.

Ryan, J. C., et al. "Greenland Ice Sheet surface melt amplified by snowline migration and bare ice exposure." *Science Advances* 5.3 (2019): eaav3738.

- Justify more adequately the use of MAR in relations to your goals either in the introduction and section 2.2 describing MAR. MAR is indeed a fine atmospheric model well adapted to Greenland but by using MAR, the authors make their approach less replicable to other parts of the world if their goal is to fine tuning parameters for modelled albedo for ice sheets and large glaciers as optimal parameters found are likely to depend on atmospheric inputs.
  It is true that MAR is less replicable in other parts of the world. However, as now explicitly stated in the introduction as per the previous comment, our target region is Greenland. For this region, MAR has been shown to outperform reanalysis products such as ERA5 (Delhasse et al., 2020). We have added the following sentence to the end of section 2.2. describing MAR:

  > "*MAR was specifically developed for polar regions and offers good performances for the calculation of SMB and its components. Furthermore, it has been shown to outperform reanalysis products such as ERA5 (Delhasse et al., 2020), especially in providing the near-surface temperature in summer which play a critical role in representing snow and ice processes.*"

  Reference:

  Delhasse, A., Kittel, C., Amory, C., Hofer, S., van As, D., S. Fausto, R., and Fettweis, X.: Brief communication: Evaluation of the near-surface climate in ERA5 over the Greenland Ice Sheet, The Cryosphere, 14, 957–965, https://doi.org/10.5194/tc-14-957-2020, 2020.

- The authors starts to talk about runoff, sublimation and surface mass balance (SMB) from L. 92-93. This should be specify in the introduction why those quantities are considered in the discussion section of the paper.
  The ultimate goal of the ORCHIDEE snow model developments is to have the best possible representation of the snow energy budget and SMB (and thus of its components). While accumulation depends directly on forcing, the processes related to compaction and ablation are the ones modelled. Hence our interest is the effect of the optimisation on the modelling of the runoff, sublimation and SMB. We have added the following to the introduction:

  > "*To correctly represent the surface mass balance and its components (sublimation and runoff)...*"
  > "*This study is the first test of applying the ORCHIDEE data assimilation system over ice sheets to improve modelling albedo and, in turn, the surface mass balance of the ice sheet.*"

  And to section 3.2:

  > "*We are especially interested in the impact of the albedo parameters of the surface mass balance and its components (sublimation and runoff)*"

- Also MAR includes a modelled albedo. How does modelling in ORCHIDEE compare with MAR's? Might be worth comparing both if possible

MAR albedo has been added to Fig 2, and the text expanded to discuss how MAR albedo compares to MODIS and ORCHIDEE.

- Mass loss in Greenland is not only due to snow melting but also ice dynamics in outlet glaciers that are located in many edge pixels. This should be mention somewhere as it is a limitation of the authors' work (no ice dynamics considered) in the introduction with adequate references such as:
    - Aschwanden, A., Fahnestock, M. & Truffer, M. Complex Greenland outlet glacier flow captured. Nat Commun 7, 10524 (2016). https://doi.org/10.1038/ncomms10524
    - Khan, S.A., Bjørk, A.A., Bamber, J.L. et al. Centennial response of Greenland's three largest outlet glaciers. Nat Commun 11, 5718 (2020). https://doi.org/10.1038/s41467- 020-19580-5

  The snow model optimised in this study has been developed to simulate mass loss from surface processes only, not to simulate the total mass loss. It is not a dynamic ice sheet model. This has been made clearer in the introduction when stating the context and purpose of the study.

- Darkening of the GrIS is mentioned in the introduction as a very important phenomenon for albedo. Nevertheless, this darkening is not explored anywhere in the paper. The ability of the authors' approach to capture this darkening could be studied in the result section.
  This model does not explicitly take into account the deposition of aerosols, algae and dust. In addition, although snow ageing, linked to the metamorphism of the snow, is taken into account, metamorphism itself is not explicitly modelled. These limitations have been added to the conclusions in paragraph discussing structural changes (starting L359):

  > *"Processes linked to the darkening on the ice sheet (e.g., deposition of aerosols, algae and dust) also need to be considered in future developments of the model."*

  Indeed, work on the deposition of dust and its impact on snow albedo is under development.

- Since this paper is about albedo assimilation, a longer paragraph dedicated to that very subject would be much appreciated. Also several references are missing, among others:
    - Dumont, M., Durand, Y., Arnaud, Y. and Six, D.: Variational assimilation of albedo in a snowpack model and reconstruction of the spatial mass-balance distribution of an alpine glacier, J. Glaciol., 58(207), 151-164, doi: 10.3189/2012JoG11J163, 2014.
    - Cedelnik, J., Carrer, D., Mahfouf, J.-F. and Roujean, J.-L.: Impact Assessment of Daily Satellite-Derived Surface Albedo in a Limited-Area NWP Model, J. Appl. Meteorol. Clim., 51(10), 1835-1854, https://doi.org/10.1175/JAMC-D-11-0163., 2012.
    - Boussetta, S., Balsamo, G., Dutra, E., Beljaars, A. and Albergel, C.: Assimilation of surface albedo and vegetation states from satellite observations and their

impact on numerical weather prediction, Remote Sens. Environ., 163, 111-126, https://doi.org/10.1016/j.rse.2015.03.009, 2015.

- Geppert, G.: Analysis and application of the ensemble Kalman filter for the estimation of bounded quantities, PhD thesis, Universität Hamburg, Hamburg. Doi: 10.17617/2.2161673, 2015.

The albedo DA paragraph in the introduction has been expanded as follows:

"Several studies have used remotely sensed albedo for DA in LSMs. *Due to albedo's influence on the partitioning of the surface energy fluxes and the subsequent effect on the development of planetary boundary conditions and clouds (Pielke & Avisser, 1990), some studies have focused on the impact assimilating surface albedo has on numerical weather prediction (e.g., Cedelnike et al., 2014; Bousseta et al., 2015). Others have mainly used remotely sensed data to derive new vegetation and soil background albedo parameters to use in land surface models (e.g., Liang et al., 2005; Houldcroft et al., 2009). There are also a number of examples of using snow albedo to improve snow models.* Malik et al. (2012) used MODIS-based snow albedo and direct insertion methodology in the Noah LSM over three sites in Colorado to improve simulated snow depth and snow season duration. Satellite-based albedo data was also used by Wang et al. (2015) to calibrate the ORCHIDEE (ORganizing Carbon and Hydrology in Dynamic Ecosystems, Krinner et al., 2005) LSM and investigate the impacts of albedo assimilation on offline and coupled model simulations. *Dumont et al., (2014) assimilated remotely sensed albedo in the Corcus snowpack model (Vionnet et a., 2012) to improve the modelling of the spatial distribution of the glacier mass balance.* Navari et al. (2018) *further* improved the Crocus model using satellite-derived albedo to improve surface mass balance (SMB) along Greenland's Kangerlussuaq transect."

Additional references:

Liang, X.-Z., et al. (2005), Development of land surface albedo parameterization based on Moderate Resolution Imaging Spectroradiometer (MODIS) data, *J. Geophys. Res.*, 110, D11107, doi:10.1029/2004JD005579.

Houldcroft, Caroline J., et al. "New vegetation albedo parameters and global fields of soil background albedo derived from MODIS for use in a climate model." *Journal of Hydrometeorology* 10.1 (2009): 183-198.

- A plan at the end of the introduction describing the sections developed in the paper would improve the readability of the paper to a great extent.

The following has been added:

"The paper is organised as follows. Methods and data, including the details about the ORCHIDEE land surface model and its data assimilation framework, driving and observational datasets, and performed experiments, can be found in Sect. 2. Section 3 lists the results, starting with an assessment of the prior and a sensitivity analysis of the main parameters. These are followed by results of the

main experiments, and an evaluation over PROMICE in situ sites. In Sect. 4, we look at the impact of the optimisation of the modelling of the SMB of the GrIS, as well as the different SMB components. Finally, the discussion and conclusions can be found in the Sect. 5.”

About the methodology:

- The set of parameters is optimized over the whole GrIS but albedo conditions varies greatly between the interior and the edges of the GrIS as stated in the paper. Does it make sense to follow this approach instead of a multisite approach? Also, Figure 4b shows differences between Northern and Southern parts of the GrIS.
  We did one optimisation over the whole of the GrIS because the aim was to have one set of parameters that works across the ice caps and with different climates. An optimisation on selected sites might have given better results, but that could have meant optimising the specificities of those sites and not having a parameter set robust under different climatic conditions. Furthermore, selecting representative sites, the number of sites to use and the number of parameter sets to have in the end comes with its challenges. There would be a lot of additional considerations and choices to be made, which may lead to different outcomes. In ORCHIDEE, one parameter set is currently used and we are able to greatly improve the model albedo by just optimising that one set.

- The Morris sensitivity analysis does not intervene in the choice of optimized parameters. It should be instead include in a discussion section after the analysis of results. Also it involves parameters that are not optimized. The Morris sensitivity analysis and the comparison between MAR outputs and ORCHIDEE outputs could be merged in a same discussion section as they tend to complete each other. This discussion should also be pointed out in the introduction of the paper. Also, the list of parameters of interest should have be provided in the ORCHIDEE land surface subsection. Since the authors focus not only on albedo per se but also on rate of density change and parameters controlling surface mass balance and runoff, authors should provide more information on the snow model employed in ORCHIDEE or at least point towards adequate references. This would help the readers to understand more clearly the authors’ objectives
  After the first round of reviews, we moved the sensitivity analysis to the forefront of the manuscript, as requested by one of the reviewers. We believe that there are arguments to be made in both cases. This type of analysis makes sense at the beginning of the manuscript since it is often the first step in data assimilation. However, since the Morris sensitivity analysis does not intervene in the choice of optimised parameters and the fact that we introduce parameters from additional parameterisations not discussed elsewhere in the manuscript, it makes sense to have it at the end of the manuscript as part of the discussion. After going back and forth on this, we have decided to return the analysis to the discussion since a) it will help alleviate concerns around the choice to keep tau_max, b) allow us to make links between the comparison ORCHIDEE-MAR outputs and parameter sensitivities and posterior values, and c) avoid confusion between the

parameters of the albedo parameterisations and the additional ones considered. Further to this, we believe the additional parameters used in the sensitivity analysis are best kept in the appendix to avoid confusion with the parameters used in the optimisation. During the first round of reviews, we had not yet included the comparison MAR-ORC and so the sensitivity analysis seemed out of place. Now with this comparison, the sensitivity analysis can be used to help explain some of the behaviours observed. The following as been added to the SA section:

> When comparing different ORCHIDEE runs to MAR (Sect. 4.1), we saw that sublimation was the output most impacted by the different parameter sets. This was especially notable at the centre of the ice sheet. This sensitivity analysis highlights that sublimation at the centre of the ice sheet is most sensitive to the Bdec and τdec parameters, which are changed in the optimisation to lower the decay term therefore increase albedo. In contrast, for runoff and SMB, both of which show no spatial variability over the middle of the ice sheet in Fig. 7, the v2 parameter from the viscosity parameterisation is more important. However, this parameter was not optimised in this study. Nor were other parameters from the viscosity parameterizations, to which sublimation, runoff and SMB are sensitive, especially at the edges. Although we do get some variation in runoff and SMB in the different ORCHIDEE runs (Sect. 4.1), since these are concentrated at the edges, it is possible that by optimising these viscosity parameters we would better fit MAR outputs.

The added plan in the introduction highlights the discussion between the MAR and ORCHIDEE outputs and the concluding sensitivity analysis. For a detailed response about concerns linked to the snow model and adequate references, see below in response to the snow and albedo modelling section of the reviews.

- τ_{max} seems to have almost no influence for albedo according to the Morris sensitivity analysis, why keeping it for optimization?
  While the Morris is a useful tool to get an idea of sensitivities, it remains exploratory with some caveats (dependant on ranges used and the number of iterations, its treatment correlated parameters etc). Since we are working with a small number of parameters, we decided to include all from the albedo parameterisation since these are all varied in the manual tuning experiments.

About MODIS data:

- About MODIS albedos, it is unclear to me if there are any reliable observations during winter time. Authors exclude data from November to February in their data assimilation system but later states albedo is improved in winter times in the result sections. This is rather confusing. I suggest you exclude all albedo comparison for winter period as I do not think they make sense (see e.g. sentence L. 280-281 “We also see that the optimisations improve the fit … in the optimisation”). Alternatively, authors can explicit in the manuscript their methodology regarding albedo during winter time.
  Yes, we agree - we forgot to remove this sentence. This has now been removed so that no comparison to winter is considered in the manuscript.

- How do the authors aggregate MODIS data from the original grid to the MAR grid (just an average or something more sophisticated)? That should be made more explicit in the text.
  The MODIS data was aggregated using bilinear interpolation. L116 has been expanded to state this:
  > "...further aggregating these data *using bilinear interpolation* to the resolution of the ORCHIDEE outputs, imposed by the meteorological forcing files (20 km)."

About the snow and albedo modelling and spin up:

- L. 69-71: "For the new icy soil type, … those of the loam soil type because it is the dominant soil type in the non-ice-free reguibs around the GrIS (Fischer et al., 2008)". Does it make sense to assume that basically, the icy soil type is impermeable (porosity and saturation are equal) for the edge of ice sheets?
  Ice is an almost entirely impervious medium, whether in the centre of the cap or at the edges. However, when crevasses are formed at the surface due to surface meltwater or liquid precipitation, surface water may penetrate at depth and flow through channel networks until reemerging at the glacier front. We acknowledge that this process (called a "moulin" in glaciology) is not accounted for in ORCHIDEE as the model does not represent the lateral water transport - snowmelt and icemelt are added to runoff which can either infiltrate (on flat terrains and if a part of the grid is composed of non-glaciers areas) or be transferred to the river network. We have added the following to L71:
  > "This amounts to considering that ice is an impermeable medium. However, it does not allow the representation of processes such as moulins where water seeps through a network of galleries because the model does not simulate the lateral transport of water."

- What is it classically used for ORCHIDEE in CMIP experiments and how do your modifications compare with the usual approach?
  Until now, ORCHIDEE did not consider ice-covered surfaces (including in CMIP simulations). In fact, when coupled with the atmospheric LMDZ model, it is LMDZ with a rudimentary one-layer scheme and a fixed albedo that handles the ice caps.

- I have several questions and comments on the following point "In the absence of fresh snow, snow albedo decreased exponentially from its fresh value" (L. 74-75):
  - According to Table 1, fresh snow has a fixed albedo. This is a rather crude approximation. Does it make sense? Why not instead using an increase such as the linear increase with snowfall intensity implemented by Boone and Etchevers (2001)? Could the authors reflect on that?
    Reference: Boone, A. and Etchevers, P.: An intercomparison of three snow schemes of varying complexity coupled to the same land-surface model: Local scale evaluation at an Alpine site, J. Hydrometeorol., 2, 374–394, 2001

  Two snow albedo schemes were implemented in ORCHIDEE and tested in Wang et al., (2012); the first is the formulation from Boone and Etchevers (2001) and the second the formulation from Chalita and Le Treut (1994) which was used as a standard in the model. This latter formulation was found to work better when evaluated against in situ sites. Therefore, the Chalita and Le Treut (1994) albedo scheme is the one used in ORCHIDEE. This citation has been added to the text.

In our model, the albedo of fresh snow is constant but is different over GIS compared to the land surfaces. The albedo value of old snow differs widely from one location to the other as a function of the environmental conditions and the state of the snowpack. In a very near future, we intend to use satellite observations to evaluate this value, and see if parameterizations such as the one proposed by Boone and Etchevers (2001) could be valuable.

Reference:
Chalita, S., and H. Le Treut. "The albedo of temperate and boreal forest and the Northern Hemisphere climate: a sensitivity experiment using the LMD GCM." *Climate Dynamics* 10 (1994): 231-240.

- Also, fresh snow occurs during winter when no MODIS data are available, how does B_{aged} can be calibrated?
  Since B_{aged} is constant in ORCHIDEE, we only have one value for the whole year.

- A_{aged} is the albedo of pure ice, please indicate it in Table 1.
  A_{aged} is the albedo of old snow, this has been added to Table 1.

- The snow albedo modelling implement an exponential decrease law that is close to the approach of Douville et al. (1995) but involves, contrary to the aforementioned paper, soil temperature. Where does this formulation come from? Has it been previously published and validated before? Has it been used for ice sheet before as well? I think the modelling of albedo deserves more explanation or justification in the paper. Reference: Douville, H., Royer, J., and Mahfouf, J.: A new snow parametrization for the Meteo-France climate model, Clim. Dynam., 12, 21–35, 1995.
  As described in the previous answer, this formulation comes from Chalita and Le Treut (1994) and tested in Wang et al., (2013). This citation has been added to the text. The exponential decrease depends on snow age which is itself dependent on the function fage. The expression of fage was initially implemented in the LMDZ atmospheric general circulation model to take into account the slowing down of metamorphism (and thus of the snow aging) in the polar regions due to the extremely cold temperatures. This formulation has been tested at Dome C (East Antarctica) within the framework of a PhD thesis, but to our knowledge, it has never been published in a peer-reviewed journal. Subsequently, it has been implemented in the ORCHIDEE snow scheme for glaciated areas. In Douville et al. (1995), a weak decrease is also taken into account during cold days, which is somehow equivalent to our formulation, except that “cold days” are explicitly recognized thanks to Tsoil. We would like to remind here, that, so far, the snow

model has never been applied to ice sheets or glaciers. As a consequence, there is no reference related to this formulation.

- L. 180-182 "All the simulations performed in this study include two years of model spin-up to allow the snow to accumulate … ensuring correct initial states": I assume that the two years of model spin-up are 1998 and 1999 since the period of study is 2000-2017. Am I correct? If so, please indicate it in the manuscript. Also what was the snow depth before model spin-up. Was it 0? Did you allow only snow to accumulate or was melting also occurring? Please provide more details. One key point of a scientific paper is that readers can be able to reproduce the described experiment themselves. The authors should focus on that very point. It would have avoided many questions I have listed.
  The two years of spin up were the two years preceding each year simulated; i.e., 1998 and 1999 for 2000, 2008 and 2009 for 2010, and 2010 and 2011 for 2012. The model was run normally, i.e., allowed for accumulating and melting, from a snow depth of 0. The precision has been added to the text:

  > "I*n each case, the two years preceding the years of study were used in the spinup and the model normally over these years (i.e., allowing for accumulation and melting) from an initial snow depth of 0.* "

About the data assimilation approach:

- The approach of selecting three random years (2000, 2010 and 2012) for the optimisation raises the question of the robustness of the approach. While it shows consistent improvement over the whole period of study 2000 – 2017, I wonder if selecting other years would have led to different results especially if, for example, 2000, 2001 and 2002 would have been selected. Could the authors reflect on the robustness of their approach?
  It is true that this is a limitation of the study. To best answer, we would have to repeat the experiment with a different subset of years - however, redoing the simulation would require a lot of computational time. Instead, we have added the following to the discussion:

  > "*For the optimisations, we also selected three random years instead of the full time series. However, it is possible that a different subset of years would give different results. Nevertheless, given the consistent improvement found over the whole period, we do not think the results would be too different.*"

- L. 120-122 "We define the observation error (variance) as the mean-square difference between … but also the model errors" I am partly unsettled by the authors' stance on observation errors. It breaks the underlying assumption of independence between background, model errors and observation errors of the Bayesian statistical formalism. Nevertheless, I can see why the authors adopt such approach but it deserves more justification. Basically, why using this approach instead of a fixed error variance for MODIS albedo data? Also, it induces more uncertainties at the GriS's edge than in its centre that probably comes from the modelling approach. Does this assumption really make sense?

We acknowledge that our method for defining the observation errors is not ideal. It is true that we want the observations errors in **R** to be independent from the parametric error in **B** and by using a prior simulation to describe **R,** this is not assured. However, **R** must contain both the observation error and the model structural error - this estimation of structure error is crucial. If we only used the error variance of MODIS, we would be neglecting this critical structural error. In an ideal world, we would be able to estimate the structural error properly, for example, by comparing a large number of different models. Since we cannot do this, we use the prior model as a proxy. Furthermore, the error in the parameters is small compared to the structural error. Since the edges are where we have the most uncertainty - both in terms of modelled processes and retrievals errors - it makes sense for the errors to be larger there.

We have added the following to the text:

> "Although not ideal, this approach is common since it is one of the only ways we can assess the model structural error, which is a large contributor to the **R** matrix.

- About τ_{max}, it was shown by the Morris sensitivity analysis that it has almost no impact. Figure 6b focus on correlations. What about variances? The cost function's Hessian at minimum can provide local sensitivity analysis (see e.g. . Also its inverse is the analysis covariance matrix. It would be of interest to study analysed variances compared to background variances and evaluate their reduction. Authors show strong anti-correlation between δ_c and τ_{max}, but associated comments should be weighted with the role of variances.

  Thank you for this suggestion. We have now calculated the Hessian at the optimum and used this to calculate the reduction in parameter uncertainty. These values match up with the sensitivity analysis in Fig. 3 where Bdec was found to be the most sensitive parameter and tau_max the least. These results impact the correlation discussion, which has now been caveated by these results. The following text has been added, first to the ORCHIDAS section 2.3:

> 2.3.4 Posterior uncertainty
>
> Assuming Gaussian prior errors and linearity of the model in the vicinity of the solution, the posterior error covariance matrix of the parameters, **A**, can be approximated by
>
> $$\mathbf{A} = [\mathbf{M}^T\mathbf{R}^{-1}\mathbf{M}+\mathbf{B}^{-1}]^{-1}$$
>
> where **M** is the model sensitivity (Jacobian) at the minimum of J($\mathbf{x}$) (Tarantola, 2005).

And second to the Posterior parameter section:

> "However, this relationship is seen to not be critical when we consider the variance at the optimum. We can see that $\boldsymbol{\tau}_{max}$ remains unconstrained by the optimisation. The reduction parameter uncertainty is small - the lowest of all the

parameters. The other parameters show high levels of parameter uncertainty reduction, showing they are highly contained by the optimisation, with B_{dec} reducing the most. This analysis mirrors the results of the sensitivity analysis in Sect. 2.3.2 - the most sensitive parameter shows the largest reduction in uncertainty, and the least sensitive show the smallest reduction."

- Two minimisation algorithms are considered in this study: the gradient-based L-BFGS-B approach and the genetic algorithm. There is no added value of evaluating those two algorithms compared to what was already written in Bastrikov et al. (2018). I think this subject is very much on the side of the authors' work. They should focus instead to the heart of their study: albedo assimilation and remove anything related to the gradient-based LBFGS-B algorithm including appendix A1.
  We have now removed all mentions of L-BFGS-B from the manuscript.

About the validation:

- The authors acknowledge that as MAR being a model this part of the work cannot be considered as validation thus limiting their study to an evolved proof of concept although still better than a twin experiment. Also the comparison is not independent since MAR atmospheric variables are used as inputs of ORCHIDEE. The study therefore shows the difference of modelling approached in MAR and ORCHIDEE (in its various configurations). What is the objective of such comparison in regards with the objectives of the papers? Again, such question occurs because the context of the study is not stated loudly by the authors.
  Hopefully, the addition of PROMICE validation (see below) will help mitigate some of these concerns.

- I do realize it would be hard work to use independent data to validate the authors' approach but the absence of proper validation really weakens the validity of the approach and therefore the paper in its current version. One possibility could be to use data from GRACE about Greenland mass loss as a way to validate your approach by assuming all the mass loss is carried out by melting (in direct link with SMB). While this assumption is probably excessive as it ignores the impact of outlet glaciers, the comparison with independent data coming from GRACE could solve partially my aforementioned reservations about the validation aspects of this paper.
  We have highlighted the potential to use GRACE in the discussion. However, we do believe it is outside the scope of this paper since, as the reviewer acknowledges, this relies on a new set of assumptions and error treatment. Instead, we have chosen to validate with PROMICE, as mentioned in the following comment.

- Another possibility would be to use in situ data from the PROMICE network as mentioned in the conclusion L. 379-380 "One solution would be to run … lead to issues of scale and representativity". In land surface modelling, it is common for example to compare modelled soil moisture at e.g. 0.25° with in situ measurements, see e.g. Kumar et al. (2019) or Albergel et al. (2020). While there are issues of scale, those comparisons are still by far important and very useful. I do not understand why it could not be done in the authors' context.

We agree that we were probably overly cautious in not including the evaluation against the in situ PROMICE data. This has now been added, as well as an introduction to PROMICE in the Methods and Data section, although we still caveat issues linked to scale representation. See below for the full text.

References:

- Albergel, C., Zheng, Y., Bonan, B., Dutra, E., Rodríguez-Fernández, N., Munier, S., Draper, C., de Rosnay, P., Muñoz-Sabater, J., Balsamo, G., Fairbairn, D., Meurey, C., and Calvet, J.-C.: Data assimilation for continuous global assessment of severe conditions over terrestrial surfaces, Hydrol. Earth Syst. Sci., 24, 4291–4316, https://doi.org/10.5194/hess-24-4291-2020, 2020.
- Kumar, S. V., Mocko, D. M., Wang, S., Peters-Lidard, C. D., and Borak, J.: Assimilation of remotely sensed Leaf Area Index into the Noah-MP land surface model: Impacts on water and carbon fluxes and states over the Continental U.S., J. Hydrometeorol., https://doi.org/10.1175/JHM-D-18-0237.1, 2019.

"3.3.2 Evaluation over PROMICE in situ sites

To evaluate the success of the optimisation, it is important to confront the results with different data. Here we look at how the fit against albedo at in situ sites is improved with the optimisation (Fig. 6). Generally, the albedo is found to improve - the fit to the observations results in a lower RMSD compared to when using the prior model. With the exception of UPE, reductions in RMSD are greater for the upper sites (between 11 and 25%) than for the lower sites (between -6 and 8%, where negative means the fit has degraded). For the UPE sites, this is the opposite. Of the 24 sites tested, the fit to the observations is only degraded in three cases. These sites are all lower sites - i.e., where the measurement station is near the ice sheet margin, where processes are harder to model. Two sites are found on the eastern edge of the ice sheet (SCO_L, TAS_L), and the last one is found at the southern tip of the ice sheet (QAS_L). When comparing to Fig. 4b, we can see that the eastern edge of the ice sheet is where the largest errors occur, even after the optimisation. Furthermore, TAS_L and QAS_L are two locations where the smallest amplitude and highest winter temperatures occur (van As et al., 2011, Fig.1) due to being exposed to the relatively warm wintertime atmospheric conditions of the Atlantic Ocean.

Figure 6 also shows us how ORCHIDEE generally performs at these sites - the magnitude of the RMSD remains similar for both parameter sets. Since the sites are mainly found at the edges of the ice sheet, errors are generally high - between 0.15 and 0.32. The two sites with the lowest RMSD for both the prior and posterior models are the ones located near the middle of the ice sheet, in the accumulation area (KAN_U and EGP). There is no obvious link between latitude and the magnitude of the errors.

Instead, elevation due to the position on the edges of the ice sheet is a more important factor.

Overall, this evaluation is encouraging - it shows that the optimisation was successful at improving model albedo when tested against a different data source. Nevertheless, we do need to highlight a couple of shortcomings in this comparison. Firstly, we do not have accurate local forcing data at the sites with which to drive ORCHIDEE. Therefore, the 20km MAR data was used, meaning that we are comparing observations and the model at different resolutions. Secondly, MODIS has been validated, and some of its biases due to the solar zenith angle were corrected for using PROMICE data (see Sect. 2.2.2). As such, the MODIS data used in the optimisation is not completely independent from the PROMICE data used in this evaluation.

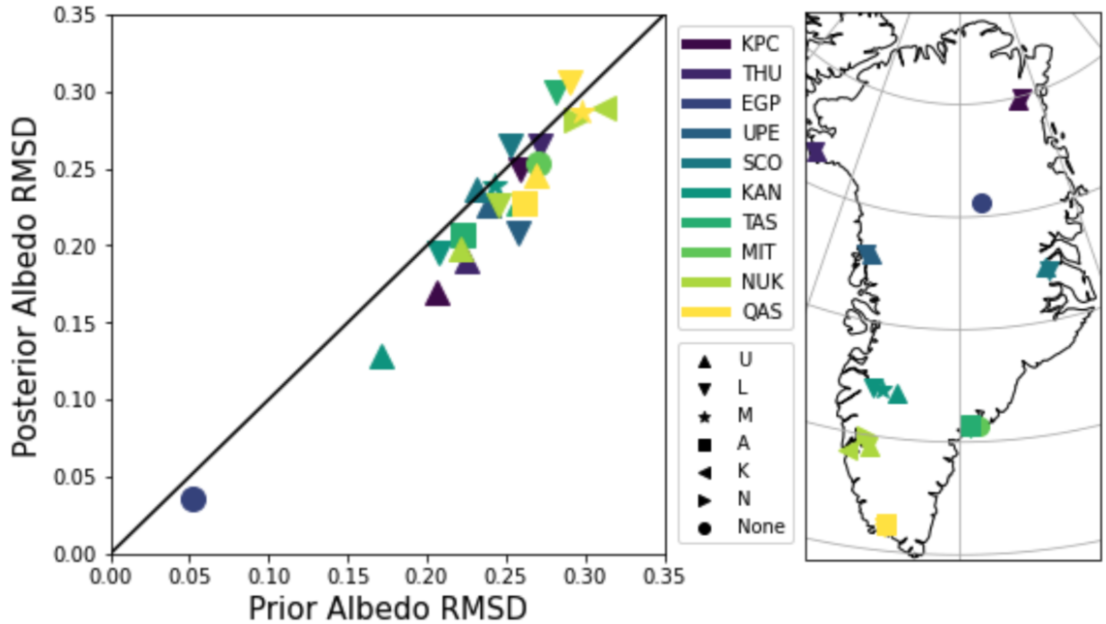

Fig 6.: Evaluation of model-observation fit over PROMICE sites. For each year of available data, the RMSD for the months (Mar-Oct) is calculated. The mean over these RMSD values is shown in the figure. Points below the 1-to-1 line represent sites where the model-data fit is improved by the optimisation.

"

- If no independent data are used, then the part of comparing ORCHIDEE and MAR should be strengthen by highlighting the differences in term of modelling for both approaches (with appropriate bibliographical literature) and this comparison should be strongly linked with the Morris sensitivity analysis and with the background context of the study that is definitely missing. Also for SMB and runoff, the difference between MAR and ORCHIDEE clearly occur at the edges of the GrIS (either with standard, tuned or optimised parameters for ORCHIDEE). The parameter optimization for albedo does not make ORCHIDEE closer to MAR (quite the contrary). The previously performed Morris sensitivity analysis could help to understand the mechanisms behind those increased differences and would make a nice discussion section.

A full comparison of both models, with additional literature and parallels to the SA, would be great. However, this would involve a lot more work in understanding the intricacies of both models. We do have a paper in preparation, Charbit et al., which will discuss a lot more the ORCHIDEE snow model. A more thorough comparison with MAR is undertaken in that study.

Nevertheless, we have added a couple of additional sentences to the MAR comparison:

> "*The fact that MAR has a more complex snow model that works better at capturing the different processes over Greenland leads us to believe structural changes are needed in ORCHIDEE for it to be able to better simulate SMB and its components. Through the optimisation, we have improved the representation of albedo but not of SMB and its components. This is because albedo is not the only important parameter in the modelling of the snowpack evolution. Other processes like melting depend on the snow's temperature profile, compaction, and refreezing, therefore on the thermal and mechanical properties of the snowpack. These processes must be well represented in the model and may require further calibration in future works.*"

MINOR COMMENTS AND TYPOS

L. 8 "This improvement is consistent for all years, even those not used in the calibration step". Could the authors rephrase the sentence as I would expect such result otherwise the methodology would not work?

We believe this statement is necessary since the model's structural errors could still prevent an improvement over evaluation years. This is especially true if different years suffer different conditions, such as increased atmospheric temperature and/or extreme events (e.g., volcanos depositing dust).

L. 14-16 "Increased warming … algae growth (Cook et al., 2020)". Darkening of GrIS has already observed and expect to worsen and increased impact on GrIS melting. Rephrase accordingly. Also, several missing references, among others:

- Dumont, M., Brun, E., Picard, G., Michou, M., Libois, Q., Petit, J.-R., Geyer, M., Morin, S. and Josse, B.: Contribution of light-absorbing impurities in snow to Greenland's darkening since 2009, Nature Geosci., 7, 509-512, 2014.
- Williamson, C. J., Cook, J., Tedstone, A., Yallop, M., McCutcheon, J., Poniecka, E., Campbell, D., Irvine-Fynn, T., McQuaid, J., Tranter, M., Perkins, R. and Anesio, A.: Algal photophysiology drives darkening and melt of the Greenland Ice Sheet, PNAS, 117(11), 5694-5705, 2020.
- Perini, L., Gostinčar, C., Anesio, A. M., Williamson, C., Tranter, M. and Gunde-Cimerman, N.: Darkening of the Greenland Ice Sheet: Fungal Abundance and Diversity Are Associated With Algal Bloom, Front. Microbiol., 10, https://doi.org/10.3389/fmicb.2019.00557, 2019.

Text rephrased and citations added.

L. 18-19 “This, in turn, enhances melting, creating feedback to the atmosphere”. Missing reference to support the statement: •

- Le clec'h, S., Charbit, S., Quiquet, A., Fettweis, X., Dumas, C., Kageyama, M., Wyard, C., and Ritz, C.: Assessment of the Greenland ice sheet–atmosphere feedbacks for the next century with a regional atmospheric model coupled to an ice sheet model, The Cryosphere, 13, 373–395, https://doi.org/10.5194/tc-13-373-2019, 2019.
- Box, J. E., Werhlé, A., van As, D., Fausto, R. S., Kjeldsen, K. K., Dachauer, A., Alhstrøm, A. P. and Picard, G.: Greenland Ice Sheet Rainfall, Heat and Albedo Feedback Impacts From the Mid-August 2021 Atmospheric River, Geophys. Res. Lett., https://doi.org/10.1029/2021GL097356, 2022.

Citations added

L. 22: “… it is crucial that it is accurately simulated in THE land surface models (LSMs) …”
Corrected

L. 39-40: “Examples of DA used for parameter estimation … in snow modelling are less common. Bonan et al. (2014) … ”. The reference is not about snow modelling but on ice sheet initialization and DA or inverse modelling is well known in this field, see the following paragraph in Bonan et al. (2014) “MacAyeal (1992) and MacAyeal (1993) introduced control methods to infer basal drag in ice-stream models, using in particular the self-adjoint property of such models, leading to many application papers (Rommelaere and MacAyeal, 1997; Vieli and Payne, 2003), and later for full Stokes models (Morlighem et al., 2010; Jay-Allemand et al., 2011). Later on, many DA and inverse methods were introduced in glaciology. The Best Linear Unbiased Estimation (BLUE) and Optimal Interpolation (OI) methods were introduced by Arthern (2003) and Berliner et al. (2008). The Robin inverse method due to Chaabane and Jaoua (1999) has been introduced by Arthern and Gudmundsson (2010) for ice sheet models, and finally Heimbach and Bugnion (2009) presented the first adjoint ice sheet model derived automatically.” Most previous references focuses on estimating basal friction or basal velocities as parameters. Regarding parameter estimation for ice sheet mass balance, you can see:

- Bonan, B., Nodet. M., Ozenda, O. and Ritz, C.: Data assimilation in glaciology, in Advanced Data Assimilation for Geosciences, Lecture Notes of the Les Houches School of Physics: Special Issue June 2012 (Edited by Blayo, E., Bocquet, M., Cosme, E. and Cugliandolo, L. M.), 577-584, Oxford University Press, 2014.

We have changed the reference to that of Su et al., 2011 which discusses joint parameter and state estimation in snow modelling.

Reference:

Su, Hua, et al. "Parameter estimation in ensemble based snow data assimilation: A synthetic study." *Advances in water resources* 34.3 (2011): 407-416.

L. 48-49: “However, with large amounts of data, … the multisite approach is common”: Debatable statement. One of the main reason of the multisite approach is commonly used is

that the set of optimal parameters or parametrizations for various sites in LSMs can differ significantly due to soil properties (soil texture, water potential, hydraulic conductivity …) and land cover (vegetation variables). Please soften the previous sentence accordingly.
Removed the part “multisite approach is common” to soften the statement.

L. 64: authors indicate that CMIP 6 version of ORCHIDEE is used in this paper: add reference publication(s) for this version in addition to the historic paper of Krinner et al. (2005). Also authors mention in the code availability section that ORCHIDEE vAR6 is employed. Could the authors harmonize notations between both paragraphs?
The citation Krinner et al., (2015) has been added to this section.The citations Boucher et al., 2020 and Cheruy et al., (2020) can be found on L61. These described the IPSL ESM of which ORCHIDEE is the terrestrial component. Unfortunately, there are no other more recent publications describing ORCHIDEE. We have expanded the code availability section to explain that vAR6 is the version used for the CMIP6 exercice.

L. 73-74 “we computed the mean of albedo in both visible (VIS) and near-infrared (NIR) spectral domains”. Please indicate that this is to be in accordance with MODIS data. Also the description of albedo following this sentence does not distinguish VIS and NIR spectral domains. I do not think spectral domains intervene in the computation of modelled albedo. Rephrase sentence L. 73-74 in order to state that your model does not distinguish VIS and NIR albedo.
We have added the following: *“This is done to be in accordance with MODIS data”.* ORCHIDEE does calculate both VIS and NIR albedo which is why taking the mean was necessary.

Section 2.2: Could the authors cite the paper(s) associated with MAR Version 3.11.4? Gallée and Schayes (1994) is rather outdated for this version.
The following citation has been added:

> Kittel, C.: Present and future sensitivity of the Antarctic surface mass balance to oceanic and atmospheric forcings: insights with the regional climate model MAR, PhD Thesis, Université de Liège, Liège, Belgique, Liège, 2021.

Section 2.3: Please mention that this dataset do not include data from the Aqua satellite as explained in Box et al. (2017). People familiar with MODIS datasets tend to expect data coming from both Terra and Aqua satellites.
Added to L101:

> “Note that this dataset does not include data from the Aqua satellite.”

L. 106-108 “Finally, in this dataset, … in the April values”. This statement is rather confusing, please rephrase.
Rephrased as follows:

> “*Finally, in this dataset, the April values are used for the winter months (January, February, November, and December). This is because there is inadequate solar illumination to compute the albedo during these months.*”

L. 114-120 “Bayesian statistical formalism (Tarantola, 2005)”. The formulation of the cost function can also be seen as an optimal control problem without any assumption on probabilistic distributions. This has given the basis of 3D and 4D-Var approach, see for example Nichols (2010). Could the authors rephrase L.114 to L.120 to play down the emphasis on the Bayesian statistical formalism?

- Reference: Nichols, N. K.: Mathematical concepts of data assimilation, in: Data assimilation: making sense of observations, edited by: Lahoz, W., Khattatov, B. and Menard, R., Springer-Verlag, Berlin, Germany, 13–40, 2010

Thank you for forwarding the Nichols’ paper - it was a very interesting read. This paper describes data assimilation mainly for the correction of model state. The problem of state estimation can effectively be seen as a control problem and it is true that the cost function we define can be used in this different context. Nevertheless, we believe that using data assimilation for parameter estimation is very different. We have comparatively few parameters and these usually represent biogeophysical properties that we can sometimes measure and for which we can set ranges. We, therefore, believe it is essential to add prior information to the assimilation which makes the problem well adapted to Bayes’ theorem. Furthermore, the literature in the field of land-surface modelling parameter estimation uses this terminology; from all the studies from our group (https://orchidas.lsce.ipsl.fr/publications.php) to the early BETHY optimisation (e.g., Knorr and Kattge, 2005). As such, we do not feel comfortable removing the Bayesian term from the ORCHIDAS description. However, we have downplayed it in the rest of the text (see responses below).

Reference

Knorr, Wolfgang, and Jens Kattge. "Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling." *Global change biology* 11.8 (2005): 1333-1351.

L. 163 “they correspond to ablation areas” Most parts of Greenland nowadays experience ablation during summer. Edges are where strong ablation occurs. Please rephrase accordingly.
Changed to: *“to areas of strong ablation*”

L. 168 “… into THE ORCHIDEE, where … ”
Done

L. 174-176 “They were also the pixels with the largest errors when compared … with RMSD greater than 0.1”. When this calculation is performed? Before calibration or after? Can the author explain where does this number comes from? It can be simply done by referring to a subsequent section of the paper.
This is the error in model prior to calibration - this has been clarified in the text. This calculation was done when testing whether the threshold of 0.5 would be satisfactory for defining edges (not shown in the paper).

L. 181 “… to allow THE snow to accumulate …”
Removed

L. 209 “Bayesian framework” see previous comment on the Bayesian term
Replaced with minimisation algorithm

About Figure 2: “currently operational ORCHIDEE version” By currently operational ORCHIDEE version, did the author mean ORCHIDEE with parameters set at default values (as in Table B1)? To my knowledge, the way albedo is modelled and the new “icy” type cannot be called “operational” yet. Could the authors modify the legend of Figure2 to reflect this point?
Changed to “standard ORCHIDEE version (before tuning)”

L. 223 “… affect by THE albedo parameters …”
Removed

L. 297 “Bayesian framework” see previous comment on the Bayesian term
“Bayesian” removed

L. 302 I think the text should refer to Figure 4a instead of Figure 6a here.
Changed

L. 308 Replace “omega” by ω and “beta” by β. L. 320 “… different parameter setS on modelling …”
Done

L. 326-327 “Compared to MAR, the manually … ORCHIDEE performs best at simulating SMB”. Best performance does not really make sense in the context of comparing two models. This sentence and the rest of the section should be rewritten keeping this fact in mind.
Sentence rewritten as follows:

> “*The manually tuned version of ORCHIDEE simulates SMB the closest to MAR's SMB*”

L. 347 “Bayesian optimisation” see previous comment on the Bayesian term
“Bayesian” removed from the line

L. 347-348 “However, we overfitted to albedo with no other data”. This statement raises the question of prescribed observation error variances. Would other prescribed values have made the impact of optimized parameters for ORCHIDEE more in line with MAR? Or is this question related to modelling differences? Could the authors reflect on that question in the paragraph?
Even though we have prescribed relatively large error variances, we are still able to fit the MODIS retrievals reasonably well. We do not think other prescribed values would have made ORCHIDEE more in line with MAR, since there are large modelling differences between MAR and ORCHIDEE. MAR has a much more complex snow model. For example, it works over a number of adjustable layers (generally comprised between 30 and 50) compared to ORCHIDEE’s three. Furthermore, albedo is not the only factor impacting SMB. We fit against albedo in the optimisation but do not include any data for runoff, or sublimation for example. This paragraph has been expanded as follows:

“The fact that MAR has a more complex snow model that works better at capturing the different processes over Greenland leads us to believe structural changes are needed in ORCHIDEE for it to be able to better simulate SMB and its components. Through the optimisation, we have improved the representation of albedo but not of SMB and its components. This is because albedo is not the only important parameter in the modelling of the snowpack evolution. Other processes like melting depend on the snow’s temperature profile, compaction, and refreezing, therefore on the thermal and mechanical properties of the snowpack. These processes must be well represented in the model and may require further calibration in future works.”

L 360-361 “When we cannot further improve … this can point to structural deficiencies in the model”. I tend to disagree with this statement. Parameter estimations can sometimes hide structural model deficiencies, i.e. you may obtain the right results but for the wrong reason. Could the authors weight on that comment?

It is true that parameter estimation can hide structural model deficiencies. It is also true that if we cannot fit the data, it is likely that something is wrong in the model for example a missing process. The two are not mutually exclusive. We have added the following to L360:

> “... it can help identify structural issues in the model - *although we do need to be cautious since parameter estimation can also hide structural model deficiencies*.

L. 361-362 “For example, we cannot capture the different albedos in the north and the south of the ice sheet with the current processes represented” This problem might also come from that the author assume the same set of parameters for the whole Greenland. A multisite approach may have reduce this problem (perhaps for the wrong reasons). Could the authors reflect on that question?

As stated in the reply above, we want to have one set of parameters for the whole of Greenland. If we start having a lot of different parameter sets for different regions, this becomes hard to deal with and to justify when to stop having more cases. Furthermore, since the climate over Greenland is expected to change, so parameters optimised in the south, where it is currently warmer, may be valid later in the north as temperatures change. By having one set of parameters, we are more likely to have averaged out over these different conditions having the set more robust.

L. 375 “There is an urgent need for data producers to provide this uncertainty, ideally at each time step”. I could not agree more. The authors can mention this following reference to strengthen their statement:

- Merchant, C. J., Paul, F., Popp, T., Ablain, M., Bontemps, S., Defourny, P., Hollmann, R., Lavergne, T., Laeng, A., de Leeuw, G., Mittaz, J., Poulsen, C., Povey, A. C., Reuter, M., Sathyendranath, S., Sandven, S., Sofieva, V. F., and Wagner, W.: Uncertainty information in climate data records from Earth observation, Earth Syst. Sci. Data, 9, 511–527, https://doi.org/10.5194/essd-9-511-2017, 2017.

Thank you for the citation; this has been added to the text

L. 379-380 “One solution would be to run … lead to issues of scale and representativity REPRESENTATIVENESS”.
Corrected

L. 395 Rephrase the sentence to replace “better” by “more consistent with MAR outputs”
Done

L. 398-404 I have some reservations with the statements written in the paragraph. Would the idea behind using all these satellite datasets be to replace the modelling of the Greenland ice sheet ice dynamics? Could the authors temper those statements?
These data would be used to optimise the internal parameters of the model, further ensuring that different processes were constrained, not just albedo. To clarify this, we added “*model’s internal parameters*” to L398 and L404.