Responses to Reviewer 1

We would like to thank Michael Lehning for reviewing our manuscript and for his constructive and insightful comments. We addressed all of these comments in the revised manuscript to improve its overall quality in line with the reviewer's recommendations. Our responses are in black, and the comments of the reviewer are presented in blue font.

General:

This paper suggests refined initialization and modelling of snow in a land surface model. The contribution consists of three parts, namely 1) to parameterize an initial snow density profile to replace long spin-up runs, 2) find better models for new snow density and 3) improve the densification routine. The paper describes in detail how these steps are implemented and uses the History Matching (HM) method to calibrate the snow settling routine. An innovative step is the fact from initialization over new snow density to densification a consistent scheme is developed to match available observations in Antarctica and Greenland.

A weakness is that it is not explained how other important snow parameters, notably snow temperature, are initialized along with the density.

First of all, thank you for your feedback on the manuscript. The initialization of snow temperature follows the procedure described in Appendix 1 (Equations A3 and A4). We use a parameterization of snow temperature that depends on latitude and elevation, which is applied uniformly within the snowpack. At the start of the forward simulation, all layers are assigned the parameterized temperature value.

We added a reference to this in the main text to make it clearer (Section Snowpack initialization):

"*The estimated mean annual surface temperature of the snowpack is also used to initialize snow and ice temperatures of all snow and ice layers at the beginning of the simulation.*"

The use of the Ligtenberg firn model for snow density profile initialization is not novel in the current contribution but has already been used in a very similar way for CRYOWRF initialization, which should be acknowledged (https://doi.org/10.1029/2022JD037744).

We included this reference for initializing the density profile in the introduction. In our case, the approach is not exactly the same as the one described in CRYOWRF. In both cases, we use the same equation (the one from Ligtenberg et al. 2011) but the difference lies in the fact that CRYOWRF uses the FDM outputs, whereas we parameterize all the variables necessary to calculate the density profile from the equation of Ligtenberg. Thus, with the various parameterizations introduced (for temperature, accumulation, surface density, etc.), we obtain a density profile that depends solely on elevation and latitude. The philosophy and interest of this method lie precisely in not using another model such as FDM, but rather in using an initialization that could be applied standalone in any other snowpack model.

We added this sentence in the manuscript (Section Snowpack Initialization) :

“This method provides a final parameterization of the density profile that depends only on latitude and elevation. Other studies have used densification equations to initialize snow models, such as CRYOWRF, which relies on the Firn Densification Model to calculate the initial density profile (Gerber et al. 2023). The strength of our method is its independence from additional models, enabling standalone initialisation for any snowpack model.”

This reviewer found it difficult to understand, which observations have been used for model development and calibration and which (only) for evaluation of the overall model performance. It should be made very clear upfront if you used the same data for both or if a clean cross-validation strategy has been adopted.

Thank you for drawing our attention to this point. In the revised manuscript, we have clarified the distinction between the observations used for model development and calibration, and those used for evaluation. All available observations of density, temperature, surface density, and accumulation were used to initialize the snowpack in this study. Regarding the development of the surface density formulation, the full set of surface density observations was used. For the calibration of the viscosity parameters, we relied on a limited subset of observations from Summit and Dome C (five profiles from Summit and two from Dome C). The goal is to capture the physical processes governing compaction in dry snow environments by relying on a limited subset of observations representative of conditions in Antarctica and Greenland. Once the calibration was completed using these sites, the model was evaluated against the full dataset of available density profiles. This approach allows us to assess the ability of ORCHIDEE to reproduce observed density profiles across all sites.

We provided a new table (Table 2 in the revised manuscript, Section Observations):

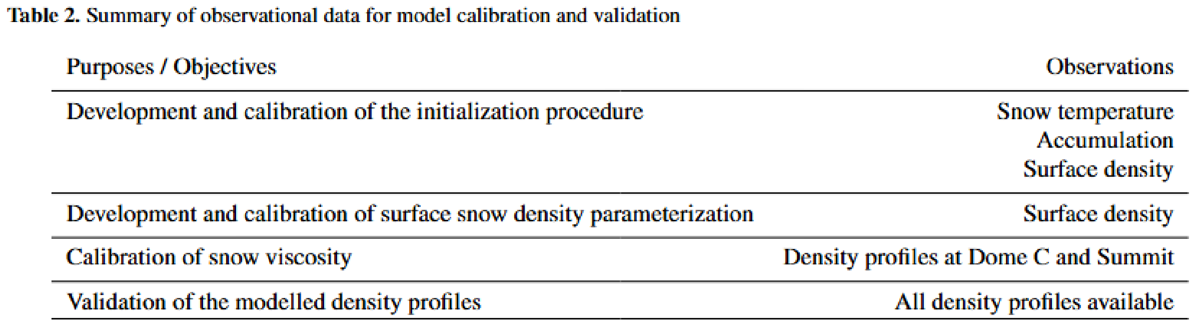


Table 2. Summary of observational data for model calibration and validation

Purposes / Objectives	Observations
Development and calibration of the initialization procedure	Snow temperature Accumulation Surface density
Development and calibration of surface snow density parameterization	Surface density
Calibration of snow viscosity	Density profiles at Dome C and Summit
Validation of the modelled density profiles	All density profiles available

We also adapted the text in the Observations section in order to explain how the observations are used in the study :

“The observations described above are used for distinct purposes. The snow temperature, accumulation, and surface density observations are used to develop and calibrate the initialization procedure. The surface density observations are employed to design and calibrate the parameterization of surface snow densification. Finally for the snow compaction scheme, the vertical density profiles at Summit and Dome C are used for calibration, while the complete set of density profiles serves as an independent validation dataset.”

A further general point is that the description of small improvements of an existing model is better suited for GMD than for TC, especially as there is less focus on new scientific results than on the model implementation.

We agree that this study includes significant model developments. However, these developments have been carried out in such a way that they can be easily transferred to other snow models, in particular simplified snow schemes used in atmospheric models, which generally neglect this aspect. We also place emphasis on the impact of densification on the surface mass balance, particularly on the reduction of refreezing. We therefore believe that our study falls within the scope of *The Cryosphere*.

Specific Points:

l. 8: maybe better "which may be useful other for other modelling chains"

This has been corrected: "*which may be useful for other snowpack models*"

l. 22: "has contributed 10%"

We added this correction to the manuscript.

l. 36 ff: Should discuss explicitly the role of saltation in densification (https://doi.org/10.1017/jog.2017.53), which is confirmed, while the role of vapor transport is still debated (https://doi.org/10.3389/feart.2023.1167760).

Thank you for your comment. In the revised manuscript, we discussed explicitly the role of saltation in densification. We added the following sentence :

"*According to Arctic field studies, the surface density in polar regions* ***can be*** *related to the combined effects of wind compaction and the upward transfer of water vapor induced by the high temperature gradient within polar snowpacks (Barrere et al., 2017; Domine et al., 2016),* ***although this latter effect is still debated (Jafari et al. 2023)****. Moreover, frequent strong winds cause large amounts of snow to be transported, compacted, and sublimated, contributing to the denser surface layers (Liston and Sturm, 2002; Sturm et al., 2001). Recent wind-tunnel experiments further quantified surface snow densification, showing that it strongly depends on wind speed and the duration of wind exposure (Walter et al., 2024). Similarly, observations during a drifting-snow event in Adélie Land revealed rapid post-depositional snow densification, with density increasing from approximately 200 kg m−3 to around 350 kg m−3 in a single day (Amory et al., 2021).* ***Saltation also appears to be a driver of surface densification, as shown by wind-tunnel experiments, emphasizing the role of particle impacts, fragmentation and subsequent sintering in enhancing snow compaction (Sommer et al., 2017).***"

l. 89 ff: This argument of excluding older data for Greenland but not for Antarctica is not convincing. First of all, your new parameterizations should be good enough to be able to handle changes due to climate and then there is growing evidence that large changes are also happening in Antarctica with more precipitation for example.

Thank you for your feedback. We have corrected this limitation and modified the revised manuscript accordingly. We now use all available data for Greenland, just as we do for Antarctica. This ensures a consistent methodology across both ice sheets. The paragraph introducing the observations has been updated as follows: **We delete these sentences:**

“For Greenland, we rely only on temperature and accumulation observations recorded after the year 2000 to ensure that our analysis reflects recent climatic conditions, given the significant impact of contemporary climate warming on the GrIS. Temporal restrictions were not imposed on the AIS dataset, as we assume that, at the ice sheet scale, the surface mass balance (SMB) has experienced only minor variations over recent decades \citep{shepherd2018N}.”

The revised text now reads as:
*From the SUMup collaborative database (Vandecrux et al., 2023), we use **3749** observations of snow temperature collected at different locations in Greenland, between **1930** and **2023**, and **905** in various sites over Antarctica between 1957 to 2021 (Fig. 1a, d). The observations of snow accumulation are sourced, for Antarctica, from the GLACIOCLIM-SAMBA dataset, detailed in Favier et al. (2013) and updated by Wang et al. (2016), with **2981** observations over Antarctica, recorded between **1930** and **2009**, and from the SUMup dataset for Greenland with **358** accumulation observations, collected between **1952** and **2023**.*”
We have updated the figures to incorporate the new observations included in the revised manuscript. An example is provided in Figure 1, showing all the observations considered.

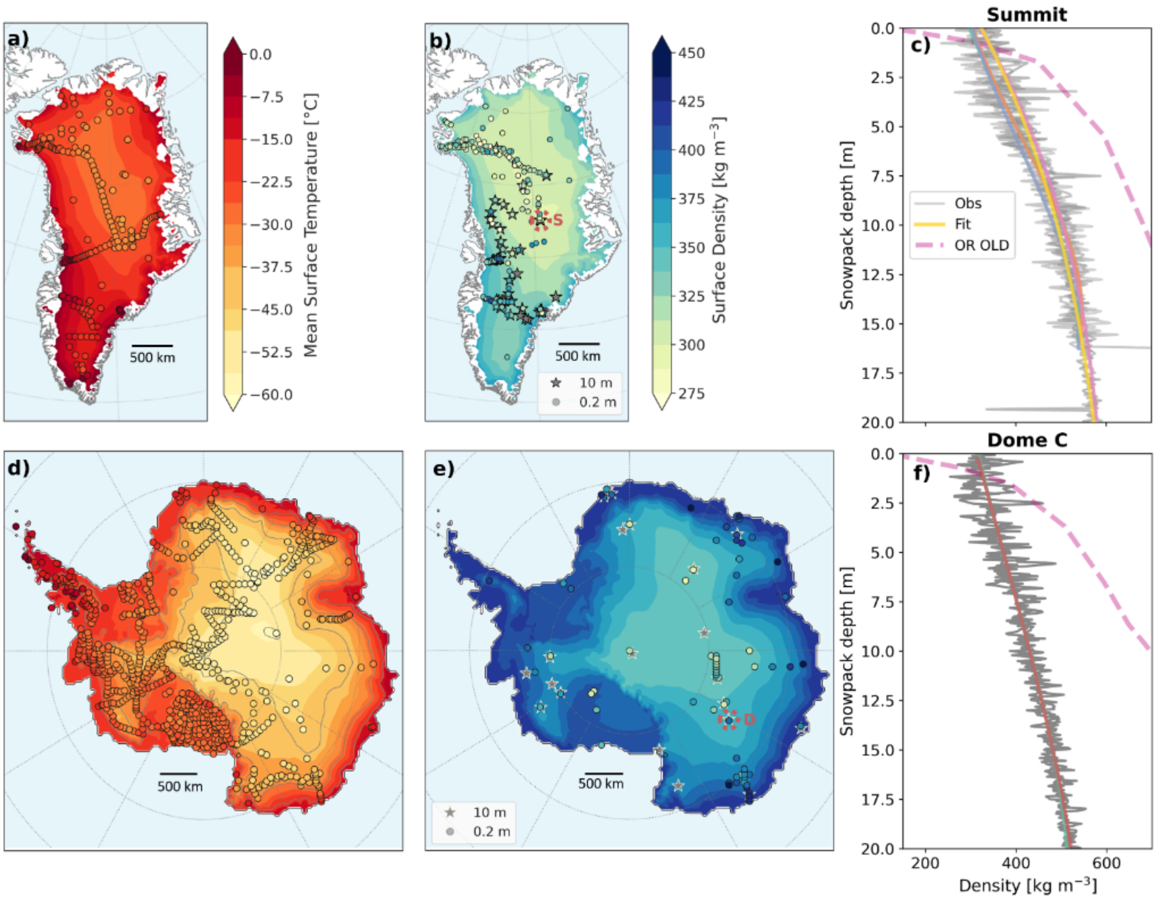


Figure 1. Mean annual surface temperature for Greenland (a) and Antarctica (d), observed at 10 m depth in the snowpack (circles). Color-filled areas are obtained with the initialization procedure (see Section 3.1). Mean surface density for Greenland (b) and Antarctica (e) for the observed top 20 cm (circles) and computed from the initialization procedure (color-filled areas). Star-shaped markers denote density profiles reaching 10 m depth. Red dashed circles indicate the location of the Summit station (’S’) in Greenland (b) and Dome C (’D’) in Antarctica (e). Observed density profiles at the Summit (c) and Dome C (f) station. Grey lines correspond to density observations and colored curves represent the fitted profiles obtained from the observed density profiles. The dashed pink curve shows the density profile computed by ORCHIDEE (OR OLD) with the standard version of the model (Section 2.4).

l. 105 ff: Can you explain what the purpose of grid averaging is over Greenland and why you don’t use it over Antarctica, where the grid is even larger?

We thank the reviewer for pointing this out. We agree that this was a limitation in the original study. To address it, we have now applied the same data processing procedure to both ice sheets. The description of the post-processing has been updated accordingly as follows:

“*To compare point observations and model gridded outputs, the observations are averaged on the ORCHIDEE stereographic grid (15 km resolution in Greenland, 35 km in Antarctica). After averaging temperature observations onto the model grid, we obtain 180 points in Greenland and 770 in Antarctica (Fig. 1a, d). We also compute the grid averaged accumulation, weighted by the corresponding duration of observations. This way, we obtain 76 grid-cell averages over Greenland, and 589 for Antarctica. For surface densities, we obtain 149 grid cell averages for Greenland and 79 for Antarctica (Fig. 1b, e).*”

We have updated the figures to include the processed observations in the revised manuscript. In addition, all metrics for the initialization procedure (bias, RMSE, R²) have been recalculated using the updated datasets. An example is provided in Figure A2, showing the processes observations.

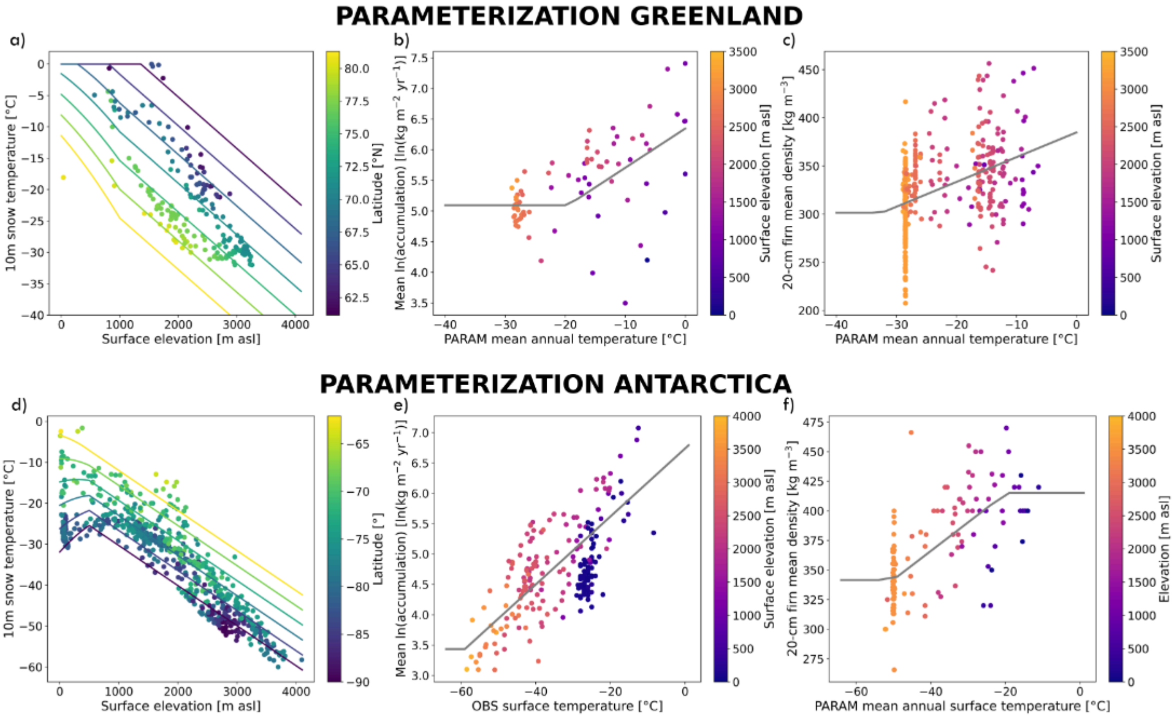


Figure A2. Observed 10 m snow temperature in function of surface elevation and latitude in Greenland (a) and Antarctica (d). Colored lines represent the parameterization of mean annual surface temperature in function of latitude and elevation. Logarithm accumlation observations in function of the parameterized mean annual surface temperature and surface elevation in Greenland (b) and Antarctica (e). Grey line show the final parameterization in function of the parameterized mean annual surface temperature. Surface density observations in function of the parameterized mean surface temperature and surface elevation in Greenland (c) and Antarctica (f). Grey line shows the parameterized surface density in function of the parameterized mean annual surface temperature.

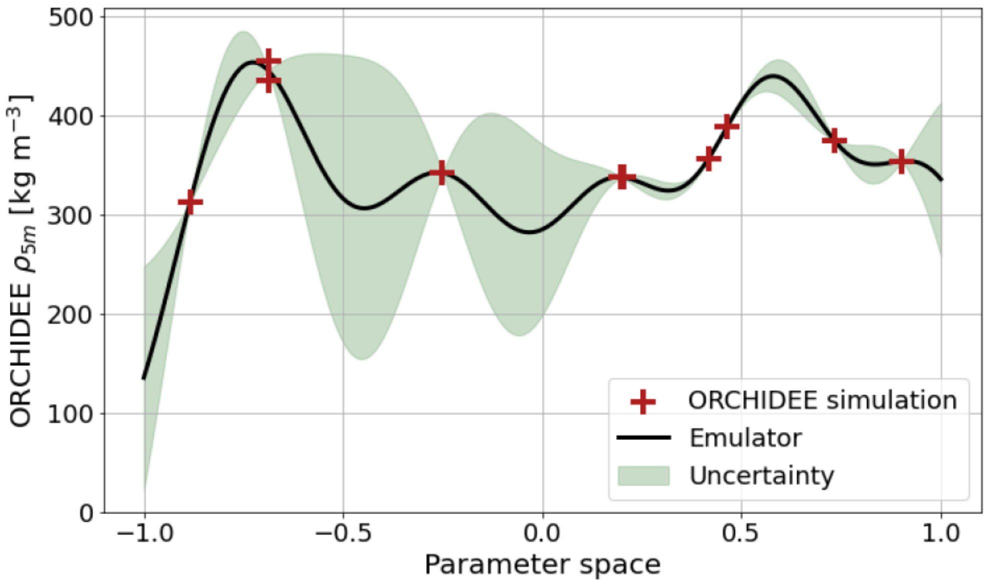
l. 144: Wrongly expressed! Either “negative energy” or cold content or similar

Thank you for your remark. We corrected this by using the “cold content” expression.

l. 182 ff: Can you add some detail to the explanation how the emulator construction works?

We thank the reviewer for this suggestion. The construction of the emulators is indeed a technical process, which we have outlined in detail below. However, given the length of the manuscript, we prefer not to include the full explanation in the paper. Instead, we provide this detailed explanation here in the reply.

To calibrate the snowpack densification, we define two metrics (rho_5m and rho_10m). We then apply a Latin hypercube sampling approach, generating a sample set ten times larger than the number of parameters to be calibrated. Then, we run as many ORCHIDEE simulations as the number of parameter sets. For each sampled parameter set, we compute the model true response for the two chosen metrics.
Once these reference metrics are obtained, the R script from the History Matching procedure implemented in ORCHIDAS constructs one Gaussian process emulator per metric (i.e., two emulators for rho_5m and rho_10m). The emulator seeks the best linear combination of “simple” functions (linear, quadratic, sine, cosine, etc.) to find the best possible interpolation relative to the reference metrics (i.e. those calculated directly by ORCHIDEE). The emulator predicts the metric and an uncertainty derived respectively from the mean and from the covariance matrix from the Gaussian process. Indeed, the kernel of the Gaussian processes defines the covariance between predictions inside the parameter space. The correlation length, a parameter of the kernel, characterizes the distance over which two points remain correlated. As we are moving away from the sampled points, this correlation decreases, leading to an increase in the emulator's prediction uncertainty. The emulator thus assigns a higher uncertainty in sparsely sampled regions of the parameter space, while the uncertainty remains lower in areas well-constrained by the sampling.
The following figure illustrates how the emulators works (not shown in the revised manuscript):



The red crosses represent the true responses of the model. The x axis represents the parameter space for one parameter, normalized between -1 and 1. The y axis represents the metric defined at the beginning of the calibration (here rho_5m). The emulator is represented in black and the uncertainty associated with the emulator prediction is in green.

l. 207: Correct typo!

We corrected this typo in the revised manuscript : "*This method is presented in detail in Appendix A, and summarized hereafter.*"

l. 232 ff: How do you justify this uniform snow depth given massively different temperatures and accumulation rates? And if I understand your introduction correctly, then you have ice below, which is hard to justify as maximum densities (ca. 600 kg m-3) in your 10m column will be very different from ice still.

In this snowpack model, the number of snow layers is fixed (12 layers). In reality, fully representing a snowpack density profile would require considering a snowpack several tens of meters thick. However, within the model framework, this would lead to the formation of unrealistically thick layers. To avoid this limitation, we therefore restrict the model to the upper few meters (~10 m) of the snowpack. In addition, the primary objective of implementing the snowpack model in ORCHIDEE is to represent interactions between the surface and the atmosphere (snow melting, albedo, turbulent heat fluxes, sublimation, etc.), and thus to compute the surface mass balance of ice sheets as accurately as possible. In this context, modeling only the upper part of the snowpack is sufficient to capture the key processes involved in the surface mass balance, such as runoff, melting, and refreezing.

Considering a uniform snow depth at initialization over Antarctica can be justified by the fact that the continent is almost entirely covered by snow. As ablation processes are negligible, we assume at the initialization of the model that Antarctica is characterized by a large and persistent snow reservoir over the ice sheet. In contrast, for the Greenland ice sheet, assuming a uniform snow depth is less appropriate due to significant mass loss at the margins. To account for this, we introduced a parameterization of snow depth as a function of snow temperature (see Appendix A), allowing for spatial variations in snowpack thickness between the interior and the margins of the ice sheet at the initialization.

l. 244 ff: Also in Greenland and Antarctica, snowfalls with very low densities occur regularly and surface snow densities as low as 30 kg m-3 have been measured. Only that they typically don't last very long. This is why event driven accumulation such as in (https://doi.org/10.5194/tc-7-333-2013) are closer to the real process than just a wind dependent parameterization. Have you considered that?

We did not consider a formulation of this type. Therefore, we have implemented in ORCHIDEE the surface densification formulation (Eq. 1) suggested in Groot Zwaaftink et al. (2013). The additional wind-dependent enhancement of the strain rate presented in this study (Eqs. 2, 3) was not implemented, as our primary objective is to compare surface densification formulations of the same type. The results are shown below. We can see that this formulation underestimates surface densification by an average of -38.8 kg/m3. In particular, we observe an underestimation for high densities in Antarctica. This underestimation may reflect the fact that the wind-dependent enhancement of the strain rate was not included in our implementation, as our primary goal was to compare surface densification formulations of the same type.

Here we can also see the difficulty of obtaining a single parameterization that can correctly represent surface densification in both Antarctica and Greenland. We are fully aware that the parameterization we have developed is not perfect and does not directly represent surface densification processes, but as demonstrated in the paper, our parameterization gives satisfactory results in comparison with observations (Fig 6).

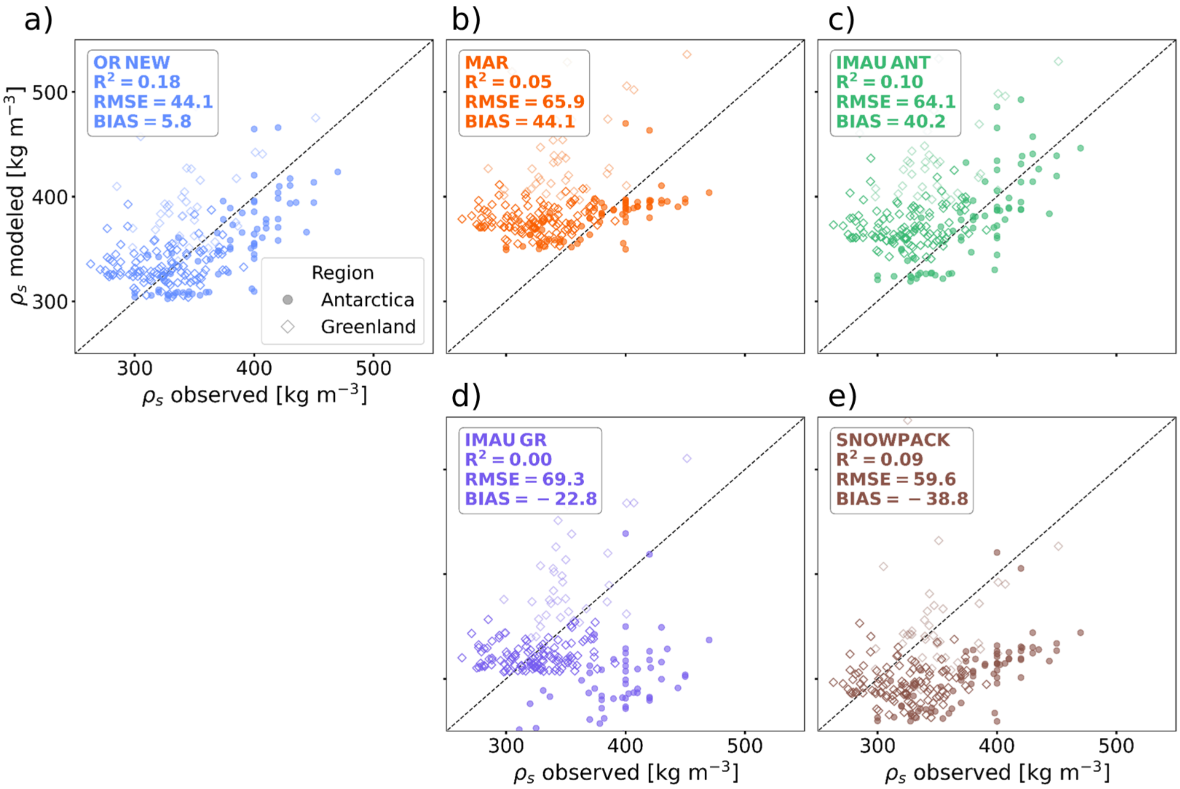


Figure 6. 40y-average surface snow density compared to observations, modeled by (a) the new wind parameterization in ORCHIDEE, (b) the wind parameterization developped for Antarctica in MAR (Kittel et al., 2021), (c) the wind and temperature parameterization developped for Antarctica in IMAU (Lenaerts et al., 2012), (d) temperature parameterization developped for Greenland in IMAU (Ligtenberg et al., 2018) and (e) the wind paramerization developped in SNOWPACK (Groot Zwaaftink et al., 2013). Locations are specified with open diamond and solid circle, for Greenland and Antarctica respectively. Shaded points are locations where the 40y-averaged refreeze in the top 20 cm of the snowpack is larger than 0.1 mm d^{-1}

l. 304 ff: You smooth out also “wanted” variability, for example driven by temperature differences, which has an impact on water transport etc. Why don’t you add a statistical variability to the profiles?

We used the fit to the density profiles to determine the metrics rho_5m and rho_10m. Indeed, this approach smooths small-scale variability. Nevertheless, we account for the variability of the density profile in the uncertainty calculations we developed in Section “Metrics for densification and uncertainty quantification”, in which we specifically account for the variability within the observed profile using sigma_within_profile for rho_5m and rho_10m. These metrics do not explicitly represent density variability, rather reflected in the associated uncertainty.

l. 325: Cancel “to”

We corrected this error in the revised manuscript : “*yielding the observational uncertainty for each metric*”

l. 365 ff: This result could be simply due to equifinality, as the calibrated sensitivity will compensate for different values of the reference viscosity. This should be discussed.

We agree that this behaviour may be attributed to equifinality. This point has now been explicitly discussed in the revised manuscript. In particular, we have added the following sentence:“*This suggests that either the selected metrics are not sensitive enough to constrain η0 or that polar snow densification is primarily driven by the temperature and density sensitivities (aη and bη), with η0 playing only a minor role.* ***Equifinality may also occur, with the calibrated sensitivity to temperature and density compensating for variations in the reference viscosity.***”

l. 545 ff: Certainly a negligible contribution in Greenland and Antarctica (see: https://doi.org/10.5194/egusphere-2025-3035).

We rephrased this sentence to clarify this point and added a reference to the saltation process in relation to missing processes in ORCHIDEE (Section Discussion and conclusion)**:**

“*As an example, saltation can be an important driver of surface densification due to the mechanical impacts on the snowpack when snow crystals fall onto the surface (Sommer et al. 2017, Walter et al. 2024), processes not captured by ORCHIDEE.* “

l. 565 ff: Surface winds and turbulent fluxes are another factor to consider and should at least be discussed next to albedo. A quick comparison against measurements from ablation area could help to better understand

Indeed, we totally agree with this. We completed the sentence by referring to surface fluxes, important to consider and to study:

“*Another perspective to improve the simulation of meltwater production, and therefore refreezing, is to focus on the representation of the surface energy balance, with particular attention to the snow albedo parameterization* ***and surface fluxes such as latent and sensible heat fluxes***.”