This is a review of the work «Implementation of an intermediate complexity snow-physics scheme (ISBA-Explicit Snow) into a sea-ice model (SI³): 1D thermodynamic coupling and validation», which I believe to be a necessary and relevant study. However, there are several concerns regarding the realism and presentation of the model and its outputs. There are large sections of the manuscript that need additional explanation of the choices in parameterization. In addition, the manuscript requires significant formatting improvements as the current presentation of this manuscript gives the impression of a rushed or incomplete submission.

We are very grateful for this thorough and insightful review. We indeed believer this contribution would be highly relevant to the community. We have put a lot of efforts into improving the quality of the manuscript following the advice below.

Note that all line numbers refer to the manuscript version with untracked changes.

The model includes a high level of complexity, incorporating parameters such as albedo, grain size, radiation extinction coefficient, compactive viscosity (accounting for snow overburden), thermal conductivity, snowfall density, and wind-driven snow compaction. Given this complexity, the density outputs appear overly simplified and do not justify the inclusion of so many detailed processes. While the average densities presented appear realistic, the vertical profile of densities on sea ice is not. Typically, a natural sea ice snowpack features a low-density depth hoar layer beneath a higher-density wind slab, which you address later in the manuscript on L403. This stratification is not represented in the model outputs.

Although the inclusion of such a model in SI3 is to improve the representation of the snow cover in the future versions of the CNRM-CM climate model (in which SI3 will be used), ISBA-es already incorporates a higher level of complexity than SI3 slab snow model, it is considered as an “intermediate complexity” (see Barrere et al., 2017, https://gmd.copernicus.org/articles/10/3461/2017/) snow model as it does not explicitly represent more complex, but key snow processes such as snow metamorphism or water vapour diffusion.

As you mentioned, our density profiles shows that the ISBA-es model does not realistically represent the arctic snowpack in detail (wind slab layer + depth hoar), but the average snow densities are realistic. We believe this suggests that the complexity of the model is not sufficient enough to represent the vertical profiles of the snow density. Indeed, ISBA-es is based on Anderson (1976) densification model (as most state-of-the-art snowpack models) that account for compaction due to the weight of the overlying snow, melt and metamorphism. Such a model is able to represent an increase of density with depth, which is typical of alpine snowpacks, but can barely represent accurately the density profiles (i.e: the wind slab and depth hoar layers) typical of arctic snowpacks. This is likely because the model currently lacks some specific arctic snow processes leading to these density profiles such as, for instance, the water vapour diffusion within the snowpack (Domine et al., 2016). We want to include such processes in future versions of ISBA-es, but this will be done in a separate study.

Moreover, it is important to keep in mind that in state-of-the-art climate models, the snowpack on sea ice is typically represented using a simple slab model, which does not allow the conductivity or density to vary in time or space. Therefore, including a snow model that allows these variables to vary is already a significant improvement, as it will facilitate the future inclusion of polar-specific parameterizations. In such, the work presented here could be considered as an intermediate step towards simulating a more realistic snow cover within climate models.

We changed the text in the discussion to make all these statements clearer, see marked version, from line 546 to line 560.

Nonetheless, the model seems to realistically simulate the evolution of snow depth and temperature relative to SHEBA, but since the publication of this SHEBA data, the thermal conductivity measurements have been questioned. I address this again below.

We agree with this comment. The thermal conductivity measured during SHEBA has been questioned (e.g: Riche et al., 2010, Fourteau et al., 2022) since its publication, but the SHEBA average conductivity measured (0.31 W.m-1.k-1) is still used as a constant in climate models. In our paper, we showed that using a parameterization of the thermal conductivity that depends on snow density rather than the constant conductivity value used in most climate models lead to more realistic bottom snow temperatures compared to SHEBA.

Although this does not necessary imply that the ISBA-es in its current version will ultimately outperform the SI3 bulk snow model in all sea-ice areas, having such an intermediate complexity snow model included in SI3 facilitate the inclusion of, for example, more recent sea-ice specific conductivity parameterizations (e.g: Macfarlane et al., 2023, that we plan to include into ISBA-es). Again, this will further facilitate developments that were not achievable with the previous, overly simplistic bulk snow model of SI3.

We also adressed this in the results section from line 433 to 437 and the discussion from line 563 to 566

The model appears tailored specifically to snowpacks on level sea ice, where dynamic processes are minimal. This limited scope should be clearly acknowledged. If this is, in fact, the case, then SHEBA measurements made only on level ice need to be incorporated. I don't believe this is addressed in the manuscript.

The version of the model presented in this paper is a preliminary 1D version that does not account for sea-ice dynamics and therefore misses some details related to ice deformation and non-linearities within a grid cell. Nonetheless, the model is not specifically tailored to snowpacks on level ice since it is included within the SI3 framework that accounts for sea-ice deformation through ridging and rafting and includes several sea-ice categories. Thus, the future versions of SI3+ISBAes (currently under development) will account for snowpacks on level and deformed ice, in particular through the following processes:

- When ridges are formed, part of the snow falls into the ocean.

- Several sea-ice categories will be used.

- The inclusion of a subgrid-scale snow thickness distribution such as in Abraham et al., 2015.

During the SHEBA campaign, most measurements were taken on multi-year ice (Sturm et al., 2002). Thus, we believe that the snow depth averaged over all transect should provide the most accurate estimate possible given the uncertainties of the measurements.

We adressed all this in sec 2.3, line 246- 248 and sec. 2.4 between lines 255 – 258.

In addition, it would be nice to read more about the parameterizations used, e.g., Anderson snow grain size. Can more details be given as to why and how this was used? At the moment, the text is quite vague, and I had to jump between references. Another example is the use of Royner (2021) parametrization for snowfall density, in which the density max is 600. How is this feasible for snowfall? Why was this chosen?

We added more clarifications in the text in sec 2.2.2, lines 160-164. The Royer et al. 2021 maximum density is used for the wind compaction procedure and not on the snowfall parametrization (eq.11).

I would also consider removing the snow-ice conversion section as there is already a lot to unpack in this manuscript and introducing this quickly at the end required new model configurations and new locations is not necessarily contributing to the paper's conclusions. Unless significant snow-ice was measured during SHEBA, but I think this could even be added to another manuscript.

This section was meant solely to illustrate that the model can capture the snow-ice conversion process. However, we agree that this part would be more suited in a future paper using a 3D version of the model (under development). We removed this part.

Please check for colorblind-friendly figures. For example, figures 2,3, 7b, and 8 have green and red plots together.

We changed all colors of all figures to make it more colorblind-friendly according to https://gist.github.com/thriveth/8560036 .

Minor comments

We adressed all minor comment directly in the text.

Please check the formatting of units throughout the text eg. g/m3 should be g/m^3

L23 please add «thermal» conductivity. Here, conductivity can also refer to ionic conductivity

L111 «We assume that snow covers the whole mesh in the presence of snow» seems self-explanatory; consider re-writing this. Do you mean that there is no snow-free ice in winter?

L127 upper layer thickness is bounded so that its thickness does not exceed 0.2m. Please elaborate on this reasoning in addition to referencing Boone and Etchevers (2001)

L134 Why did you choose the Brun et al (1989,1977) scheme? Please explain the reason behind this decision.

Actually we didn't choose this scheme, it was the only one implemented in ISBA-ES (Decharme et al., 2016, https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2018MS001545)

Eqtn 2 W_l seems to be referenced as W_1 ensure the same symbols/fontare used throughout

L147 Tf needs to have subscript f

L154 subscripts needed here

Eq9 c_w incorrectly subscripted

L190 $t\Gamma_{i,drift}$. There is no t and the k is replaced with i in the equation. Please check this.

L194 «which should be more suited to the Arctic region.» why is this? What is your reasoning for this?

We addressed this in the text, line 203-204 *" in order to increase the effect of the wind on compaction, and improve the representation of the upper wind slab layer typical of arctic snowpacks (e.g: Domine et al., 2018; Sturm et al., 2002, and references therein)."*

L198 Since the work of Sturm there have been further work on the snow thermal conductivity. For example Riche 2013 showed there is likely an underestimate in this scheme due to the measurement bias. Is there a reason for this choice of parameterization? Consider looking into Calonne (2019)/ Macfarlane (2023). Would this make a difference in your ice growth estimates?

Actually we used the Sturm et al parameterization to be consistent with the SnowModel-LG model that also uses this parameterization, but the Macfarlane (2023) parametrization will be implemented in the near future. This has been addressed in the discussion part from line 561 "The model could be also further improved by incorporating specific snow-on-sea-ice processes that are not yet represented in SI³+ISBA-ES, such as for example the formation of superimposed ice or by incorporating more recent parametrizations of the thermal conductivity (e.g: (Macfarlane et al., 2023)."

L212 formatting of $\rho_{s,k}$ and $h_{s,k}$

L214 there is no enthalpy in equation 14. Consider moving this to line 218

L218 Soce and Si are the salinities of seawater and ice respectively. If these are fixed values please provide them here

L235 «subscript sn» are you referring to the subscripts s here? There is no sn!

L236 formatting of K_{s-i} from here onwards I will stop pointing out each parameter formatting but please go through the text for the next submission version to correct all these formatting mistakes.

L259 magnaprobes spelling

L261 In Persson et al. 2002, they mention that «Ts was measured at the radiometer stand (or the Barnes radiometer stand) where the snow was significantly shallower, so the Ts values used may be slightly different than the actual surface temperature above the Tice and ds measurements. These factors must be considered when making conductive flux estimates from these measurements». It would be good to mention these uncertanties and how the snow-ice interface temperatures need using with care, it's hard to identify in thermistor profiles. Unless an alternative method was used? Please clarify this

We clarified this line 258: "Note that the two temperatures were not measured exactly at the same exact location, leading to an uncertainty of ~0.6 °C on the temperatures."

L268 Does ERA need caps here?

L276 marginal «ice» zone

Figure 2 top: the different thickness of lines means that it's really hard to see the difference between ERA5 and MERRA2. Please consider changing this or adding a relative difference plot between the two

Figure 2 bottom legend capitalise SHEBA (and fig 3). Are the observed (black dots) a daily average? If so, please include this information in the caption.

L303 Calculation incorrect here, 0.153 x 191 does not equal 50.5. I also got a little lost here with parameters used, why are you referring to mass per unit area? Can you clarify and also bring this into Table 2?

We deleted this part because it was pointless and confusing.

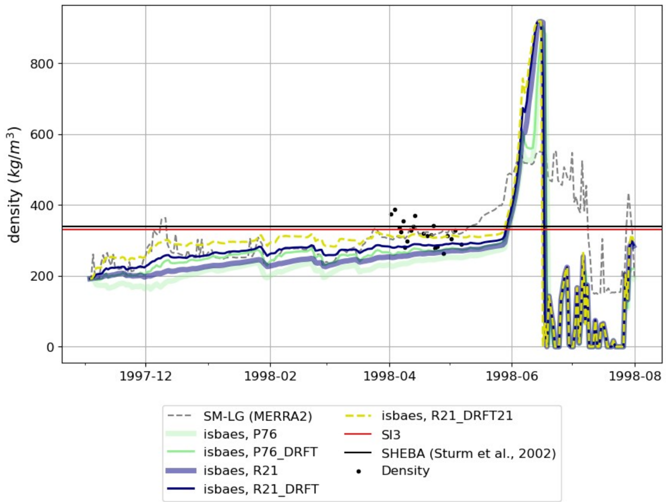
L324 How does it compare well? What are the errors associated with the models? The line seems to lie on the average but the snow melt and the thickness decrease is not well represented. Are all SHEBA datasets on level ice? Or are also the thicknesses measured, including ridged areas? This refers back to my previous comment about this model being tailored to level-ice areas. Maybe only using SHEBA data collected on level ice actually improves the comparison between models as there is likely to be less variability in snow thickness.

The SHEBA dataset measured on level ice was addressed in the text and in the answer to the main comments above. Concerning the snow melt, it has already been addressed in the next section (3.1.2) "*However, the snow tends to disappear in all simulations during the first month of the melting season while it persists until the end of the melting season (roughly, from May to August) in the observations. In the melt season (from May 1998), the excessively fast melting of the snow in all ISBA-ES simulations compared to the SHEBA observations and the SI^3-only simulation is likely due to the ISBA-ES albedo parameterization which seems to underestimate the albedo during the melt season. Indeed, when the snow albedo is forced to 0.8 (which is higher than the albedo simulated with ISBA-ES, see Figure 8), the snow thickness matches the SnowModel-LG simulation during the melt period (see supporting material, Figure S1). This underestimation of the albedo is probably due to its dependency on the density, or to the density itself.*"

L325, if ERA5 underestimates snow thicknesses, then the ISBA-ES looks to also underestimate. Also, the 3.5 cm in the text is listed as 3.4 cm in Table 3. Please check this and also add a reference to Table 3 on L326.

Indeed the snow thickness in ISBA-ES are smaller than SnowModel-LG but are in the range of the observations. The differences with SnowModel-LG have been addressed in the text line 334 "The mean difference of P76_DRIFT simulations with SnowModel-LG outputs are 2.2 cm for ERA5 simulations and 2.1 for MERRA2. "

Figure 4a, can you include the density evolution measurements made on SHEBA rather than the average? This would provide a lot more necessary information to the reader.



The densities at SHEBA were only measured during one month. Here on this figure we added the daily -averaged densities measured at SHEBA (black dots). We have added the daily-averaged densities measured at SHEBA on Figure 4 , in spite of the very short period covered by the measurements.

L369 snow persists until the end of the melting season in the observations. Maybe this includes snow in ridged areas? What about a surface scattering layer that develops? The observations might include this which is trypically approx. 2-10cm. This also relevant for your final sentence in this paragraph, L379. An important process that needs adressing in this discussion!

We adressed this at line 558 in the discussion “The model also seems to overestimate the albedo decrease during the melt season, leading to a rapid snow melt. This underestimation is possibly due to a misrepresentation of the surface scattering layer in our model, that was observed in the melt season at SHEBA (Perovich, 2005).”

L391 Denser snow associated with a thinner snowpack? Really? Can you include a reference for this statement?

Since the snow mass remains unchanged during compaction processes, an increase in density will ultimately reduce the snow thickness in the model.

L410 lacks accurate representation of the wind slabs. In my opinion, the model also lacks accurate representation of the lower depth hoar formation. If this is the case, I’m struggling to convince myself that this 1D model is useful. Please can you comment on this? Why not just look statistically at the data, and are the over-simple representations of snow cover worse than this?

As mentioned in the response to the main comments above, having such a 1D model included in SI3 paves the way to further developments and improvements to the snow that were not feasible before. Thus, although the density profiles of the current version of ISBA-ES are not always realistic, there is room for improvement, which we believe should be treated in a separate study.

=> “are the over-simple representations of snow cover worse than this ?“ In SI3-only previous snow, the density is a unique constant value.

Figure 6 and 7a, it’s hard to see the data in both of these plots, can you consider changing the line thickness or colors or elongating the plot?

L422, include a reference for this statement

L423 Check units throughout W /mK

L425 very low conductivities are used here especially for a bulk conductivity. 0.1 is more likely a conductivity for fresh snow with low density, see Calonne (https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019GL085228) a reason for this is the likely underestimation of k from the snowfork see Riche 2013 (https://tc.copernicus.org/articles/7/217/2013/) which explains why the conductivity estimates are low on SHEBA. I would consider rewording this and including the newer references.

We addressed this in the text from line 433: “Those conductivity estimates are consistent with the measured bulk conductivity at SHEBA (0.14 $W.m^{-1}.K^{-1}$, Sturm et al., 2002) but since the publication of the SHEBA datasets the conductivity measurements have been questioned and appears to be too low (Calonne et al., 2019; Riche and Schneebeli, 2013). Sturm et al. (2002) estimated a thermal conductivity of 0.31 $W.m^{-1}.K^{-1}$, inferred from ice growth and temperature gradients. Thus, the thermal conductivity simulated by ISBA-ES seems slightly underestimated.”

L443 The lower snow surface temperature in Figure 7a (although it is hard to see) combined with the low thermal conductivity means that the snow-ice interface temperatures will be warmer. As a result it is hard to draw direct conclusions about this in the model.

The surface temperature of ISBA-ES is colder than in SI3-only simulation and than the SHEBA observations. Nonetheless, the bottom temperature is much warmer in ISBA-ES than in SI3-only, and is much closer to the SHEBA observations. This suggests than the 0.31 W.K-1.m-1 conductivity used in SI3 is too high.

Can you include a model run with the SHEBA snow surface temperatures as an input?

Unfortunately, the model cannot be forced by surface temperatures at the moment. However, if that was possible we would have expected to have the bottom snow temperatures simulated by ISBA-ES a bit too warm compared to the observations, as the conductivities simulated by ISBA-ES seems a bit low (as you mentionned above).

L458 albedo of 0.83 is consistent with observations at the tower. Where is this seen? Can you include a reference to this statement?

L478 Why is snow-ice conversion important specifically during the Jan 11993-Jun1993 period? Please provide details

This part has been removed.