Dear Editor Masashi Niwano

Thank you for handling our manuscript. We have made further improvements to the manuscript based on comments from you and reviewers

Please see our response (text in blue) below

Please let us know if further improvements are still needed.

Best regards,

Yubing Cheng, Bin Cheng and co-authors

Dear Yubing Cheng, Bin Cheng, and coauthors,

Thank you for submitting the revised manuscript together with the detailed response letter. Anonymous Referee #1 reviewed the revised manuscript, and the referee is generally satisfied with your responses, although the referee has provided some additional minor comments to improve the paper further. Regarding your responses to Anonymous Referee #2, I have checked them along with the relevant revisions and judged that they are convincing. Based on the results, I have judged that this paper can be published after minor revisions. Please consider the constructive comments provided by Anonymous Referee #1 and revise the paper accordingly. Below, I also list some technical comments, which I would like to ask you to consider:

- Figure 3: It seems to me that it is not referred to in the running text. Please check.

The text on lines 97-99 was used to explain Figure 3. The location and the style of expression may be a bit confusing. So, we modified the text:

The relationship between SWE (mm) and snow depth (m) is illustrated in Figure 3. The regression slopes represent mean snow densities for each period: 348 kg/m^3 (entire MOSAiC campaign), 308 kg/m^3 (winter-spring period), and 487 kg/m^3 (MOSAiC legs 4 and 5). These values likely reflect the characteristic snow or surface scattering layer densities at a specified period.

- Tables S1 and S2: It is better to change the order of these tables, because Table S2 is referenced in the running text earlier than Table S1 (L. 105, manuscript v3). If you agree with this point, please revise the following places accordingly: L. 105, L. 166, L. 168, and L. 447.

Agreed, we made changes accordingly.

- Figure 7a: Because the line thicknesses/colors for the linear regression coefficient curves and the error bars are almost the same, it is a bit difficult to distinguish them. Can you improve the presentation of this figure?
- Figure 7b: It is better to add an explanation for the linear regression lines in the caption.

We redraw Figure 7a using the y-axis from both sides of the figure. The regression lines are shown separately with better visualisation.

The figure 7 caption for (b) is updated:

(b) Example for 3 days wind speed accumulation versus surface(red) and bulk(black) snow densities. The solid lines represent the ordinary least squares linear regression fits.

The manuscript addresses an important topic, but several points require clarification before acceptance. First, the use of the term "dry snow" (Line 15) is unclear, as it is not consistently referenced elsewhere; the modelling section should explicitly state the conditions under which the simulations were conducted. Second, while the evaluated snow density schemes reproduce mean values reasonably well, they fail to capture temporal variability. Given that models such as SNOWPACK and CROCUS include more advanced compaction schemes with demonstrated agreement with in-situ data, the authors could clarify why these were not applied. Finally, in the description of MOSAiC snow pit observations (Lines 115–140), the manuscript lacks detail on the distribution of pits across different ice environments. Since snow properties vary substantially between undeformed ice, ridges, and leads, this information is crucial to assess representativeness. Providing the number of pits at each location and their relative areal coverage would strengthen the credibility of the observational dataset.

Line 15 A numerical snow and sea ice model was applied to simulate the sensitivity of sea ice to snow density and snow precipitation during the period when the snow was dry. I may be mistaken, but I could not find any other reference to "dry snow" in the manuscript. Please clarify this in the modelling section, as the condition under which the simulations were conducted is not entirely clear.

We added text in 2.4 Snow and ice model to clarify this matter:

A single-column high-resolution thermodynamic snow and ice model (HIGHTSI) is used to simulate the sensitivity of snow density on the thermal regime and mass balance of snow and ice during the winter-spring period (28 October 2019 - 6 May 2020) when the air temperature was below zero degrees and snow was dry.

Lines 18–19 The examined snow density schemes produced mean snow densities consistent with MOSAiC observations; however, none of the schemes adequately captured the observed temporal variability in snow density.

From the manuscript, it is clear that the densification schemes you selected did not reproduce the temporal evolution of the snowpack. However, snow compaction schemes implemented in models such as SNOWPACK and CROCUS have demonstrated good agreement with in-situ data. Could you clarify why you did not apply those schemes in your study?

This is a good comment, and thank you for pointing out this issue.

The reasons we did not apply those advanced snow schemes are: 1) we are aware of SNOWPACK and CROCUS models. As far as we understand, those are advanced snow process models that have been applied to alpine snow or terrestrial snow where the snow densities differ from those in the polar regions. The parameters applied in the snow compaction schemes may not be valid for the polar conditions; 2) We focus on assessing snow density schemes that are computationally cheap, that can be applied for climate research. So, the snow density schemes we picked up are simple, and they use adequate input snow parameters that are available. On the other hand, we acknowledge that advanced snow models could be valuable for climate research once their computational efficiency is optimised and snow observations are improved.

Lines 116–119, 115–140 During MOSAiC, comprehensive sea ice and snow observations were carried out (Nicolaus et al., 2022). Snow pit measurements were taken at least weekly but often on several days per week, and occasionally more than once a day. Snow pits were dug at various locations on undeformed first-year ice, second-year ice, and places close to open leads and pressure ridges. Most measurements were taken within the central observatories in designated clean, undisturbed snowfields. In this section you describe the frequency and distribution of snow pit measurements.

However, it remains unclear how many pits were dug at each type of location (undeformed first-year ice, second-year ice, ridges, leads, etc.). This information is crucial, as snow properties differ substantially between these environments. For example, repeated measurements near ridges could bias the calculated mean depth and density relative to undeformed first-year ice. I recommend specifying the number of pits at each location and, ideally, providing an estimate of the relative areal contribution of these representative sites to the total study region. This would help readers assess how representative the aggregated values truly are.

We added sample numbers and their percentage to the total snow samples in the revised manuscript:

Snow pit measurements were taken at least weekly but often on several days per week, and occasionally more than once a day. Most measurements were taken within the central observatories in designated clean, undisturbed snow fields. Snow pits were dug at various locations on undeformed first-year ice (107, 142), multi-year ice (20, 76), and near open leads (16, 17), ponds (5, 2), and pressure ridges (61, 35). The numbers in parentheses indicate the sample counts for surface and bulk snow, respectively. More than half of the surface and bulk snow samples (51% and 52%) were collected over the first-year ice (FYI).