

Anonymous referee #1

The following is a review of “Physics-constrained generative machine learning-based high-resolution downscaling of Greenland’s surface mass balance and surface temperature,” by N. Bochow, et al.

This manuscript describes a new machine-learning based method using a consistency model (CM) for downscaling ice sheet surface mass balance. The authors focus on the Greenland Ice Sheet, evaluating their product against output from the regional climate model MAR. They demonstrate the ability of their method to reconstruct high-resolution surface mass balance over the historic period from downgraded input. They also illustrate how their method could be used to downscale low resolution bias-corrected output from an earth system model through the year 2100. They provide examples of different ways, with varying complexities, to inform the CM, and evaluate advancement in the method’s skill. The authors suggest that their approach could be integrated as interface between general circulation models and ice sheet models for improved certainty in sea-level projections.

This work is highly relevant to the ice sheet model community, because high spatial resolution is needed to force accurate ice sheet model estimates of future contribution to sea-level change. However, earth system models, such as those that participate in the CMIP experiments, run at much lower resolutions than needed. This is because they are typically global-scale models faced with significant computational constraints. Current state-of-the-art methods for downscaling to finer spatial scales also have strong computational constraints, because of the high spatial resolutions required to run. The method outlined in this manuscript tackles this issue. It is a novel approach that is capable of successfully downscaling to finer resolutions using a machine-learning based method. It is computationally efficient and has significant skill, on the condition that there are high-quality training sets (e.g., simulations from other methods) available. The emergence of this method is timely in that it could offer support for IPCC efforts, and the outlined approaches are clearly very promising for use by the ice sheet modeling community.

I find that the text is well-written, and the figures are high-quality. However, I do think the manuscript would benefit from restructuring, which I discuss below in general comments. Overall, my impression is that the authors were diligent and attentive to the needs of ice sheet modelers when designing and evaluating these new products. For this reason, I believe it is important that this work be published so that the authors can build upon their collaborations with ice sheet models, perhaps with implications for the next IPCC report.

Thank you for this positive feedback and identifying the bigger implications of our work! In the following we will respond to all comments (blue font).

Having said that, I am not convinced that this work is appropriate for publication in The Cryosphere. I find the manuscript to be highly focused on presenting and evaluating the approach, as it does not provide many results related to a science question. As presented, the main goals of the study are to introduce a method and evaluate its feasibility and robustness. Though the manuscript shows the method has strong potential to impact cryosphere science, there is no specific scientific hypothesis tested. In this way, the manuscript may perhaps be a better fit for GMD? For its publication in TC, on the other hand, I suggest that the authors think about building a more cryosphere-science driven story with their various downscaled products, particularly with development of the discussion section to expand upon for the reader, how the method advances ice sheet science. Focusing more on results pertaining to the projection runs could be a starting point.

We see your point and agree that the structure and presentation of the manuscript was not optimal for TC. Generally, we also agree that we could see this paper published in e.g. GMD, however the topic is most appropriate to this journal. We will modify the manuscript along the suggestions of both reviewers, and expect that the revised version will be suitable for TC. For this, we will (1) reorder the manuscript such that it adheres to the standard order of TC articles, (2) add more experiments showing the advantage of our method for ice-sheet modelling and (3) generally improve the readability of the manuscript.

General comments:

- I suggest that the authors reorganize the sections of the manuscript for readability. The current organization seems to be set up in a letter format, with the Methods at the end. For The Cryosphere, I think it would be helpful to present the methods to the reader before the results, so that it is clear what is being presented in the results, discussion, and conclusion sections.
- We will reorder the manuscript to adhere to the standard section order of TC.
- The downscaling of future projections is buried in the Appendix, but these results are arguably the most interesting and relevant to the study's motivation with respect to cryosphere science (i.e. possible future use for downscaling projections for CMIP/ISMIP). Including a timeseries plot with comparison against GT, for instance, would be helpful to evaluate the model results throughout the projection. Perhaps these projection results could be highlighted more and used to build/expand on a science-related story.
- We agree and will move the figures and add a more cryosphere-driven storyline.
- I find that all figure captions are written in a terse style, with many of the articles removed. Because this makes it difficult for a reader, I would suggest reviewing all the captions and revising as needed so that they flow better and can be more easily understood.

- Thank you for this comment, we will revise the figure captions accordingly.
- For clarity, please make sure all units are consistent. For instance, even if it is stated that the SMB or MAE are over a month period, the units should still reflect this (i.e. mmWE/month) in all references within the text and in the figures.
- We will check the whole manuscript and change the units accordingly.

Specific comments and suggestions:

Page 3, line 2 – “using” openly available ?

Fixed in the revised version.

Results, line 2-3 – Please describe how insolation is defined in your context. For instance, is it downward solar radiation at the ice sheet surface? Is it a climatological monthly average, calculated over a particular period, or is it a timeseries of monthly averages that is used?

We use a monthly- and spatially-varying insolation field over the whole ice sheet. This is just used to give the model some information about the seasonal cycle, the insolation is constant over the years. We show the approach in the manuscript now.

Results, line 7-8 – Should these be “Fig, 2” and “Table 1”?

Thank you, yes, we corrected that.

Section 2.1, line 13 – The phrase “compared to the ground truth” makes it sound like GT is not as accurate, which I don’t think is what the authors mean. Maybe “when compared with ground truth” or something similar that makes this statement clearer.

Agreed, we change the wording.

Section 2.1.1, line 5 – This is an example of where the units are not conveyed per unit time, but if appropriate, should be for clarity.

We will change it.

Section 2.2, line 10 – I may have missed this, but if not included, please add a reference for NorESM and version/output used in this study.

Thank you, we will include a reference.

Discussion, line 9 and Section 5.1, line 19 – Please add a reference for U-Net and perhaps a short definition somewhere in the text, to help a more general audience follow your methods and discussion.

Agreed, we will include a reference and explain it shortly.

Section 5.4, line 2 – Please include a description of which variables are used for precipitation (total precipitation?) and near-surface temperature (2m air temperature?).

We used total precipitation and 2m air temperatures, we will mention that in the revised version.

Figure 1, Caption – Here is an example where the caption is written informally with missing articles, i.e. “at the surface for a random month from the test set” is more appropriate. Also, for example, Figure 2, Caption – “for a warm month”.

Figure 1, Caption – “is visually indistinguishable” is used numerous times throughout the text, but it is not a scientifically precise or quantitative statement. Especially since, when looking at these plots, they have differences that can be detected visually. Figure B.1 for example is clearly visually different. Please rephrase your uses of “is visually indistinguishable” throughout the manuscript. I suggest including difference maps, maybe in the Appendix, to illustrate the method’s skill with respect to the spatial pattern.

Figure 3 (a,f,l), With reference to CM Constrained, it would be easier for a reader to make comparisons between the spatial plots and the line plot if the labeling for the same runs was consistent. Also, PSD is defined in Figure 4 (and in the text), but it should probably also be defined in Figure 3 since it would come before Figure 4.

Figure 5 and Figure B.2 – Instead of time index, could the simulation “date” be given on the x-axis?

Figures C.1, C.2, and C.3 – The letter labels are missing from all panels. In addition, it would be helpful per plot to be more specific about exactly which runs are “CM”, since CM is used in many earlier plots and it gets confusing for the reader to follow.

Figure C.1, Caption – “of precipitation, melt and runoff” -> “of precipitation, evaporation and runoff”?

Figure C.2 and C.3, Caption – Please specify that this is for NorESM in the caption. Also, there is only description included for 6 panels, not 8.

Figure C.4 – Should the units for (a) technically be mmWe/month?

Thank you for all of these suggestions and comments, we will change the captions accordingly, check the units throughout the manuscript, and try to include some difference maps.