Response to the editor (egusphere-2025-3870)

We would like to thank the editor sincerely for taking the time to evaluate our manuscript.

We agree with the general comments from the reviewers and the editor and made substantial changes to the manuscript. Please find attached the tracked changes document that allows to see the amount of change to the manuscript. We streamlined and simplified especially the results and discussion sections. We reduced the word count by ~30 % in these sections by removing Figs. 6 and 9 and their corresponding paragraphs but also by rewriting other paragraphs, removing repetitions and focusing on the most important aspects. We consider this effort to be an extensive rewrite.

In the revised manuscript, we addressed the comments related to potential deviations between the Blatter-Pattyn and full Stokes models. In the Introduction, we have clarified the objective of the paper and elaborated in the Discussion on the limitations of our study. In particular, we discussed deviations between the Blatter-Pattyn and full Stokes models at fine resolutions and emphasized the need for future studies to explicitly assess resolution dependence in full Stokes models (see section 4.4). Please find attached our detailed replies to related comments by the reviewers (referee #2, page 26, and referee #1, page 7).

Comment on egusphere-2025-3870 (referee #1)

We would like to thank reviewer 1 kindly for taking the time to read our manuscript and for the constructive feedback, which helped us to improve the study. In the following, we addressed all comments point by point and point out respective corrections. Please find our answers in blue italic font.

Werner et al. present a study on the impact of spatial resolution when modelling extensive ice cover of mountainous regions. They find that increasingly coarse resolution leads to greater ice volume, slower ice, and slower response times -- at least part of which can be explained by the coarser resolution lowering peaks and filling in valleys.

I think all of the modelling behind the results and discussion is competently and appropriately done (though one or two points about instructed vs. numerical model implementation) and as such don't believe further simulations are necessary. However, I think there is scope (and need) to really simplify, shorten, and tighten the text and the figures and to focus on providing some key takeaways for practitioners (even if that comes with caveats).

For this reason I think this falls into major revisions, even if no new analysis is needed. I also expect that the current presentation means it is possible that I may not have caught all the issues.

We thank the reviewer for the general assessment. In the revised manuscript, we shortened the results and discussion by ~30 %. This reduction was mainly achieved by deleting repetitions and focusing on the most important points. We removed former Figs. 6 (showing velocity difference maps) and 9 (showing resolution effects throughout the cooling phase) and focused on (i) the examination of ice volume changes throughout the run (Figs. 3, 4, 9) as well as (ii) ice thickness, velocity, and basal temperature at the full glaciation (Figs. 5, 6, 7, 8). Please see below our responses to other major and minor comments for more details on this point.

At present the emphasis is on where the highest sensitivity to resolution is (about 400-800 m). This is one way of putting it that could make sense if one has been dealing with interpreting these simulations for a long time. However, if we take 50 m as the 'most realistic' run, then the important value is not so much the rate of change with respect to the previous simulation, but the total drift away from the most realistic simulation. This information then provides an interested reader with information along the lines of 'yes, you can coarsen your resolution to improve runtime, but you will be dealing then with an error of this rough order'. It feels like the authors are maybe avoiding being more definite about this which is perhaps understandable, but I think you have a good case to provide more concrete recommendations (supplied with suitable caveats). For example, Williams et al. (2025, https://doi.org/10.1038/s43247-025-02010-z) suggest 5 km -- maybe this isn't a perfect number or maybe it's spot on, but at the very least it is a working value that can be used in lieu of further investigation.

I would suggest switching to this approach, meaning figures, results, discussions of the actual variable in question (not a difference), or the error relative to that one baseline value. If the authors feel strongly about keeping their current representation I think some strong reasoning for their approach is required.

Thank you for the suggestion, we agree and took the 50 m run as a reference and compared results to the 50 m reference run in the revised manuscript. Instead of emphasizing the "critical mode of resolution sensitivity" at intermediate resolutions, we focused on the contrast of model results at fine and coarse resolutions. We considered the 300 m resolution as "sufficiently fine" to obtain model results in close agreement with the 50 m reference run. We believe that this framing is clearer and more useful to modellers. Especially, the resolution of 300 m for the exemplary ice field in the Western Alps could be a starting point and reference for practical resolution considerations.

Beyond that, I think there is a bit of an oversupply of figures here (and the detail in some, such as 9, is a bit too condensed). I don't want to suggest exactly how this should be done, that requires thought, time, and familiarity with the material that I don't have (nor is it my place) but I would encourage the authors to think about what the key takeaways they want a reader to have are, and emphasise these in a reduced number of figures (and figure panels). Additional information can go into the supplement, even if that is already rather crowded !

As mentioned above, we removed former Figs. 9 and 6. Moreover, we reduced the panels in Fig. 5 and removed the Arrhenius factor analysis (panel f in Fig. 8) from the manuscript. With the newly added Fig. 7, this results in a total number of 9 figures. Fig. 7 shows bedrock and ice surface elevations along a transect and illustrates resolution effects on ice thickness and topography as well as on smaller-scale features that might help modellers to identify resolution issues across a smaller area, which addresses a major comment from reviewer #2.

I was also struggling to get through the results and part of the discussion. I hope that focussing on a reduced number of central points might make it possible to streamline your findings and paint a clearer narrative. This doesn't mean important details have to be lost, just that the ridges and valleys of the text should be a bit more defined to take a mountainous analogy. As with the figures, I won't go into full details on all of specific ways this could be done in the more minor comments.

Thank you for raising this point. In addition to removing repetitions and former Figs. 6 and 9, we removed the term "critical mode" from the manuscript. This enabled us to contrast results at fine and coarse resolutions instead of describing what the "critical mode" is.

My minor comments are not fully comprehensive as I think there is enough work to be done here that that could be a redundant effort, but I have tried to go into some detail !

To end on a positive note, I think the results are really interesting and that this should eventually produce a very valuable paper. So apologies if the below reads as harsh, it's not my intention. I think you have something good here but it needs a lot of filing down and a change in tack on the presentation front.

Abstract:

16 - Opportunity here to say in which direction at no/minor additional word cost.

We removed this sentence from the revised manuscript because it focused on "resolution modes". Instead, we described the direction of changes between coarse and fine resolution results in the rewritten abstract.

18 - It could be me, but here and elsewhere I had to do a bit of thinking about what the critical mode actually is. Following the major comments this could be redundant, but I would suggest using a more easily graspable phrase here.

Thank you for this important comment. Following your suggestions above, we removed the phrase "critical mode" and we now emphasize differences between fine and coarse resolutions and viewing the resolutions at 300 m (and finer) as "sufficiently fine" to accurately capture topography-controlled ice dynamics.

20 - Can say that coarse resolution artificially e.g. lowers slope angles as I'm sure we're all in agreement that coarse resolution is less realistic.

*Thank you for the suggestion, we adopted that in the revised manuscript: "However, at resolutions coarser than ~800 m, topographic resampling **artificially lowers slope angles** and mountain peaks…"*

21 - 'Slower temperature...' I guess this refers to coarse resolution? but should be clearer

*Yes, the reduced hysteresis effects with slower temperature forcing refer to coarser resolutions, we added that to the sentence: "When the rate of temperature forcing is reduced by half, the hysteresis between climate forcing and glacial response **at coarse resolutions** is partially decreased…"*

23 - Critical mode could be either defined, or a different term could be used

We fully agree and removed the term "critical mode" (see also comments above).

24 - 'We expect similar' could go to 'non-linear... are likely in mountain regions worldwide' as I think you can be reasonably sure of this.

*Good point, we have now adopted this: "**Non-linear** and altitudinal-dependent resolution effects are likely in mountain regions worldwide…"*

Here and elsewhere it is not always super clear in what direction the non-linearity is pointing -- as with the major comments I would suggest honing in on a few of the most significant changes and focussing on clear descriptions of these.

*We agree that previous descriptions of resolution effects were not focussed enough. We clarified what the changes at coarse and fine resolutions are: "However, **at resolutions coarser than ~800 m**, topographic resampling artificially lowers slope angles and mountain peaks, providing a larger accumulation area at high altitudes, with **thicker glaciers that are typically warm-based**, while **thinner glaciers at fine resolutions remain cold-based**. Raised valley floors **at coarse resolutions** result in **slower-flowing ice with increased thickness and glacial response times**."*

I might come back to this in the Discussion, but I think the question of whether these changes can be expected in different regions is an interesting one. One thought I had is that an 'idealised' mountain range could be used for widespread applicability, or you could just run this over the Himalayas too. That does arguably 'fall outside the scope of this study', but even without that I think you can be more confident in your assertion that other mountain ranges should follow similar patterns. Certainly the lower peaks and shallower valleys part is self-evident.

Thank you for raising this point. In the discussion, we restructured and shortened the paragraphs and dedicated part of the second paragraph of section 4.3 to resolution effects in other mountain regions. This discussion is based on our experiments in the Western Alps using the original and a smoothed DEM: "We expect similar resolution-dependent patterns at high, mid and low altitudes in other mountainous regions, where specific altitude bands may shift depending on regional topographic features."

Introduction:

I think some broader literature on the importance of resolution is missing from this introduction. In addition to the Williams paper cited above, a quick google scholar search 'importance of resolution for ice sheet modelling' reveals quite a wealth of literature. I think this work should be better situated within this literature, as well as what sets it apart (namely the use of IGGM and more of a focus on mountains). Searching for mountain-glacier specific literature gave less hits, but then you can talk about this being one of the first (if that is the case).

We thank the reviewer for raising this point. In the revised manuscript, we added a short paragraph to the Introduction (second paragraph) to motivate our analysis by referring to resolution effects for ice sheet modelling of Greenland and Antarctica (Williams et al., 2025; Rückamp et al., 2022; Cuzzone et al., 2019; Aschwanden et al., 2016). Moreover, we suggested that resolution effects for modelling mountain glaciers and icefields may be stronger than for ice sheets and pointed out that so far, "a systematic analysis of the impact of grid resolution is lacking".

33 - Overuse of extensive if you feel inclined to change that.

Good point. To avoid double usage of "extensive", we reformulated this to: "In the European Alps, extensive glaciations carved steep and narrow peaks and formed over-deepened and U-shaped valleys during the Quaternary period (Ivy-Ochs, 2015; Liebl et al., 2021; Penck and Brückner, 1909)."

42 - I don't think references are necessary for this statement

We agree and removed the references: "For example, surface elevation controls the mass balance mainly through the temperature lapse-rate, whereas the steepness of the bed controls glacier flow."

42 - about -> for

We changed this accordingly: "topographical details ***for*** *peaks and valleys".*

48 - In The Alps? or in our model domain? Region could be specified, basically.

This sentence refers to the Western Alps and is changed to: "Specifically, at elevations higher than 1000 m a.s.l. in the ***Western Alps****, average slope angles in a 100 m DEM are ~5–10 ° steeper compared to a 500 m DEM and even ~10–15 ° steeper compared to a 1000 m DEM (Fig. 1c)."*

65 - Averaged over the entire domain ? or a local maximum ?

This sentence refers to the Rhône Valley: "In the European Alps, comparisons between modelled ice surface elevations and trimlines suggests ice thickness overestimations during the Last Glacial Maximum (LGM, ~24,000 years ago) in models based on a km-scale resolution, e.g., by up to 1000 m in the ***Rhône Valley*** *(Seguinot et al., 2018; Jouvet et al., 2023)."*

66 - 'A better match...' not clear how this sentence supports your argument and also raises the point, is the 400-1000 m exaggeration due to resolution or sliding, or (probably better), you're suggesting that both have a role to play.

Thank you for raising this point. We added a sentence addressing this comment: " A better match between modelled ice surface elevations and trimlines at the same resolution was achieved by Mey et al. (2016) by adjusting the sliding coefficient. Recent work by Leger et al. (2025) significantly reduced the mean mismatch to ~150 m by using a higher spatial resolution of 300 m in an ensemble of 100 simulations with various parameter combinations. ***These results suggest that the agreement between numerical models and field data can be reduced not only through improved parameter choices but also by increasing the spatial resolution.*** *However, it remains unclear whether even finer resolutions lead to more accurate results, and how spatial resolution influences other aspects of glacier modelling in alpine regions."*

77 - This passage maybe trails off a bit -- I don't think you have to mention everything you do, but at the same time I generally find the use of 'other key variables' a bit vague. If there aren't so many 'other variables', then it's maybe best to simply list all of them, or following the major comments to be more specific about which you take as key indicators (and why).

We agree and clarified which variables are part of our analysis: "Our analysis examines ***ice volume, thickness, velocity, and basal temperature*** *in terms of their spatial and temporal variations."*

Methods:

Good to talk about the implications of using IGGM over a full numerical model. How big could the innacuracies be? Your approach is certainly justified, but as IGGM is quite new, some more background (this part in the intro) would be good!

Thank you for this comment. IGM makes use of a neural network emulator to model ice flow. This neural network is trained to minimize the energy associated with the Blatter-Pattyn high-order stress-balance model. We updated the corresponding paragraph in the methods section to clarify the functionality of IGM. In particular, we distinguished more clearly between pretraining that only influences the initial glacial state and retraining which influences the model during the entire simulation and plays an important role to mitigate inaccuracies (i.e., the degree of assimilation of the Blatter-Pattyn model). Additionally, we gave the retraining frequency in model years (instead of model iterations) which is more meaningful to readers: "The ice flow module in IGM is designed to efficiently simulate 3D velocities using a physics-informed deep learning approach. Instead of solving costly Stokes equations or approximations thereof, ice velocities are simulated by an inexpensive convolutional neural network (CNN) (for a detailed description see Jouvet and Cordonnier, 2023). The ice velocities are simulated as energy-minimizing solutions of the higher-order Blatter-Pattyn model (Blatter, 1995) which has been *shown to reproduce ice velocities that agree well with full Stokes models (Pattyn et al., 2008). […] The initial glacial state is based on a pretrained CNN described in Jouvet and Cordonnier (2023) to advance convergence to the Blatter-Patty model in the first time-steps. For this pretraining, the emulator was trained over a diverse catalogue of glaciers and flow regimes […] For adjustments to new ice field states attained through run time, we retrained the physics-informed neural network multiple times within a single model year (**every ~0.18 years**). The on-the-run retraining in the highest-resolution simulations at 50 and 100 m was performed more often (**every ~0.05 years**). **The retraining frequencies were found to be high enough to ensure that the emulator's deviation from the analytical solver typically remains below 5 m yr^{-1} in ice velocity** (see Figs. 4, 5 in Jouvet and Cordonnier, 2023; Leger et al., 2025)."*

In order to strengthen confidence in IGM, we added that Leger et al. (2025) compared IGM simulations to a traditional model: "A direct comparison of a simulation of the European Alps at 2 km resolution using IGM to the respective simulation using the well-tested and widely used Parallel Ice Sheet Model (PISM; Winkelmann et al., 2011) published by Jouvet et al. (2023) showed minor deviations, confirming that the alpine ice field is a suitable application for IGM (Leger et al., 2025)."

*In the discussion of the revised manuscript, we briefly addressed inaccuracies between the Blatter-Pattyn model and a full Stokes implementation: "**At fine resolutions that are required to accurately model ice field dynamics, the Blatter-Pattyn approximation implemented in IGM may deviate from the full Stokes solution and exhibit mechanical inaccuracies**, particularly in steep terrain (Rückamp et al., 2022; Yan et al., 2023; Hindmarsh, 2004). The presented study deliberately focuses on numerical rather than mechanical errors and demonstrated that substantial numerical errors already occur at spatial resolutions commonly used in large-scale mountain ice-cover simulations (Mey et al., 2016; Seguinot et al., 2018; Jouvet et al., 2023; Golledge et al., 2012; Zhang et al., 2022). **For a Greenland ice stream region modelled by Rückamp et al. (2022), velocity deviations between the Blatter-Pattyn and full Stokes model of less than ~10 m yr^{-1} appear smaller than the resolution-induced differences of up to several tens of meters per year in our analysis.** Therefore, adopting a full Stokes formulation would substantially increase computational costs without addressing the main source of error. Future studies are needed to systematically assess the benefits of full Stokes models at fine resolutions in alpine regions."*

88 - Can you give a more quantitive description of its accuracy ?

*For an extensive discussion of IGM's inaccuracies we refer to Jouvet and Cordonnier (2023). In general, the retraining frequency is the most important parameter to account for inaccuracies. As mentioned above, we added a justification for the choice of retraining frequencies: "The retraining frequencies were found to be high enough to ensure that the emulator's deviation from the analytical solver typically remains **below 5 m yr^{-1} in ice velocity** (see Figs. 4, 5 in Jouvet and Cordonnier, 2023; Leger et al., 2025)."*

89 - End of this sentence a bit unclears

Good point, we replaced the end of the sentence with the following to improve clarity: "The physics-informed deep learning approach is independent of training data from other glacier models, which allows us to run the same experimental set-up at different spatial resolutions and directly compare the model results."

90 - A few more details here about Leger et al. would be good. Particularly the comparisons to actual physics based models, and to observational data. This is the core of your results, so it's good for the reader to know the underlying model can be trusted (even if it is an unseeing unthinking statistical emulation :p).

*A comparison between results from Leger et al. (2025) and observational data is already provided in the introduction: "Recent work by **Leger et al. (2025) significantly reduced the mean mismatch to ~150 m** by using a higher spatial resolution of 300 m in an ensemble of 100 simulations with various parameter combinations." Additionally, Leger et al., (2025) validated an IGM simulation of the European Alps at 2 km resolution by comparing it directly to the simulation (also at 2 km) using PISM, which should strengthen the reader's trust in our setup (see quotation from the revised manuscript above, page 7, 2nd paragraph).*

99 - 'high-order' -> higher-order

*Yes, we changed the sentence accordingly: "The ice velocities are simulated as energy-minimizing solutions of the **higher-order** Blatter-Pattyn model (Blatter, 1995)…"*

100 - hopefully not a killer question, but is the performance a function of the model resolution used in training ? I guess the training takes quite a long time to run so this is a bit difficult to test, but could be quite important. For example, if training at a much lower resolution mitigates many of the problems associated with low actual resolution, that's a pretty important result. Some discussion/thoughts on this would be good !

*Thank you for this interesting question. In general, the idea behind the pretraining in IGM is to make sure that the simulated ice flow converges with the Blatter-Patty solution more quickly. With enough retraining, the ice flow should not be dependent on the pretraining after the beginning of the simulation. We added this to the revised manuscript: "The initial glacial state is based on a pretrained CNN described in Jouvet and Cordonnier (2023) **to advance convergence to the Blatter-Patty model in the first time-steps**."*

Since the pretraining is based on a fine resolution of 100 m and the retraining frequency during run times is quite high, the resolution in the pretraining is not expected to affect model results. Moreover, possible adjustments in the 50 m run to the glacial state from pretraining at 100 m are mitigated during our 1000-year initialization phase. Similarly, deviations from a pretrained emulator at coarse resolution would be accounted for during initialization. We believe an extensive explanation of the implications of spatial resolutions during pretraining is not *necessary, but we mentioned that those possible deviations are accounted for in the initialization in the revised manuscript: "The forcing begins with an initialization phase of 1000 years with no temperature change **to ensure that the initial glacial state is in balance with our setup**."*

104 - It's worth making it clear that Jouvet and Cordonnier did this and not you (if that is the case). See also the comment about introducing Leger et al. in a bit more detail (l. 90)

*We agree and added the references to both sentences: "The initial glacial state is based on a pretrained CNN described in **Jouvet and Cordonnier (2023)** to advance convergence to the Blatter-Patty model in the first time-steps. For this pretraining, the emulator was trained over a diverse catalogue of glaciers and flow regimes at 100 m spatial resolution with 10 vertical layers, fine vertical discretization close to the ice-bedrock interface, 16 CNN-layers, and 32 CNN-output filters (**Jouvet and Cordonnier, 2023**)."*

106 - How was it retrained?

*The retraining takes place during run times and is based on the Blatter-Pattyn model: "The ice velocities are simulated as **energy-minimizing solutions of the higher-order Blatter-Pattyn model** (Blatter, 1995)…" As mentioned above, the retraining frequency plays an important role and we changed the units to retraining frequency in model years, which should be more comprehensive.*

130 - I guess using a spatially inextensive ice sheet ? Shouldn't really matter as you're testing sensitivity only (not gaining 'actual' results) but could be worth mentioning.

*We are not sure that we understand where this question is pointing at. The size of the ice field is described in section 2.2: "Our modelled ice field on this topography covers a **maximum area of ~12,700 km^2**, comparable in size to the present-day Southern Patagonian Icefield (~12,200 km^2; Meier et al., 2018)."*

The gflex module described in lines 126-130 is applied over the whole model domain. In order to lower the computational costs the spatial resolution of the gflex module is lowered to 2 km, which is however fine enough to capture large-scale GIA-induced changes, as shown by Leger et al. (2025).

151 - Not sure I catch the drift of this sentence

Thank you for pointing out that more explanations are needed. This sentence refers to experiments with a smoothed topography that are not shown in the results but are used only as references in the discussion. Therefore, we deleted this sentence from the main text and explained those experiments in the supplements (where corresponding figures S8 and S9 are shown) in more detail: "The smoothed topography is obtained through the following process: We took the DEM of our model domain at 2 km resolution, which was created from the fine resolution DEM using cubic resampling. Therefore, the 2 km DEM is rather smooth with less *topographic detail. We then proceeded to interpolate the 2 km DEM back to finer resolutions, again using cubic convolution.”*

156 - You could specify you start 'warm', or if tied to the present day temperature, how cold exactly. Also good to have a very brief idea of what ice configuration this sets the model up in. Your time step is 0.01-0.04 -- when does this value change? Why do you use this value ?

Thank you for the suggestion. Following the next comment, we specified the present-day climate by giving an annual temperature that remains unchanged during the initialization phase and is the starting point for the cooling phase. As written in line 152, the initial ice configuration is that every simulation starts without any ice.

We specified the ice volume values at the time after the initialization when the simulations are in balance with present-day climate in section 3.2: “After the initialization phase (0–1000 years), the 2000 m simulation has ***~160 km²*** *of ice volume which is more than 10 times that of the 50 m run (****~12 km²****).”*

The time step in IGM is adaptive and not static. IGM computes the time step such that it satisfies the CFL convergence condition to ensure numerical stability for the ice thickness evolution. Therefore, the time step is not a chosen model parameter and instead calculated internally. We removed the details on the time stepping in IGM from the revised manuscript and to focus more on the important retraining frequencies, where we used model years as the unit (see response to comment in line 88).

165 - This lapse rate opens the possibility of reporting an annual temperature at sea level, which is a bit more human redable than degree Kelvin away from the base temperature.

Yes, we added the value for mean annual temperature at sea level to the revised manuscript: “Monthly temperature and precipitation values are derived from climate data averaged over 1981–2010 from a weather station in Modane, France, located at 1228 m a.s.l. within the model domain (Météo-France, 2022). We used a lapse-rate of 6 °C per 1000 m elevation to project the temperatures across the entire model domain. The resulting ***mean annual temperature at sea level is ~14.9 °C****.”*

165 - I don't follow this statement exactly. It was multiplied by 1.6 across all times/temperatures, or just increased from 1 to 1.6 as things got colder? I fully accept the main aim of this paper is to test out resolution, but given you are emulating a realistic event, it would be nice to know a bit more information/rational behind this step.

We agree that more clarification is needed and changed the text passage to: “With an annual sum of 660 mm, the weather station’s precipitation values lie at the lower end of typical values in the European Alps (e.g., Isotta et al., 2014). Therefore, to ensure that the ice field covers the entire mountainous part of the model domain at the time of maximum cooling, ***we multiplied monthly precipitation values by 1.6, resulting in a more realistic annual precipitation*** *of*

1056 mm per year. ***Those monthly precipitation values were kept constant in space and unchanged throughout the warming and cooling phases****."*

167 - This statement is implicit if you don't feel like including it

Thank you for the comment, we believe this statement might be helpful for the reader to make sure they do not confuse our experiments with an attempt for realistic simulations of the LGM.

176 - What percentage of your domain did this influence ? Was this applied for all resolutions if only one resolution qualified ?

Thank you for raising these questions. The mask is based on the 50 m resolution run and it is applied to all simulations after resampling all model outputs to 50 m (see section 2.3). Please note that due to the choice of the model region, the mask is only effective at the left and right boundaries because the ice does not reach top and bottom boundaries at all resolutions. At 50 m resolution, ice-covered pixels (at full glaciation) outside of the mask are ~14 % of the entire model domain. Because the ice field is slightly more extensive towards the top and bottom of the model domain at coarser resolution, the influence of the mask is expected to be lower at these resolutions.

We decided on this approach to make the comparisons between the resolutions more straightforward and not depend on a watershed algorithm that determines the ice area for each resolution individually. Moreover, because the 50 m run serves as our reference, we like to base the model setup (in this case the area considered for our analysis) for all resolutions on the 50 m run. We added a note to the section 2: "We analysed differences ***relative to the highest-resolution simulation at 50 m****, which serves as our reference. Spatially distributed model results are analyzed within a defined mask that is based on glacial drainage basins located well-inside the model domain at maximum ice coverage of the* ***50 m simulation****."*

Results:

Please also see my major comment on this. I think there is significant scope to streamline this, and I don't think it's my place to point out every opportunity so I leave that to you, hence the not huge number of individual comments.

185 - Assumption to -> assessment? I get that this is a bit of a grey area but then again, there is enough of a background that you probably could call this an assessment not an assumption.

Thank you for this comment. To avoid assessment and assess in the same sentence, we used "notion": "Based on the ***notion*** *that resampling to a coarser resolution affects large and deep valleys less compared to small tributary gorges or ravines, we* ***assessed*** *elevation differences for valleys of distinct Strahler stream order (Strahler, 1957)…"*

190 - With reference to the 30 m base map ?

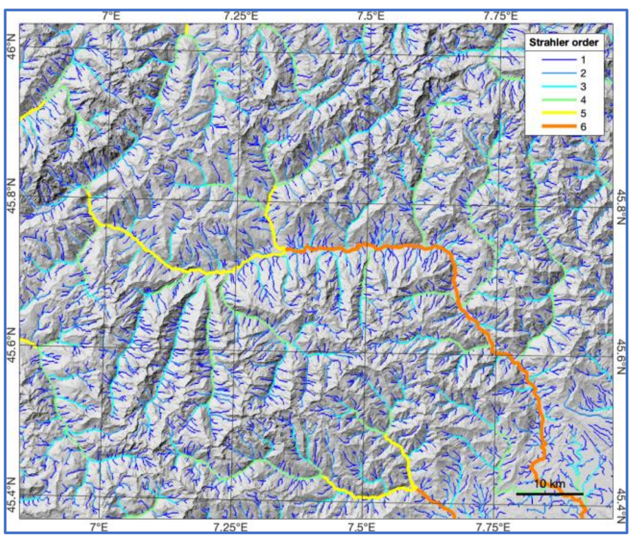
*The reference for all input and output comparisons is the 50 m resolution (see section 2.3). Since this might not be clear from the text, we added that general information to the paragraph: "The comparison of variously resampled DEMs to the **reference DEM at 50 m resolution** shows that elevation differences vary with Strahler order (Fig. 2)." The specific text passage is changed to: "The highest discrepancies are at intermediate Strahler orders (3–4), with median difference values ranging from 150 to over 250 m for DEMs resampled to a resolution of 1000 m and coarser **compared to the 50 m DEM**."*

193 - A little confused by this sentence ! I think it could do with some reorganisation/clarification.

We agree and changed this sentence to: "In contrast, at very low (1) or high (6–7) Strahler orders, corresponding to small and steep as well as large and gentle valleys, respectively, median difference values are less than 165 m, for any resolution."

Fig. 2 - I also think you might need a supplementary figure showing an example of how these stream regions are classified -- from watershed to watershed ?

Thank you for the suggestion, we added such a supplementary figure showing a map of the Aosta Valley with streams colour-coded by Strahler stream order:



194 - I am no master statistician, but if all the changes are positive, but the box and whisker plot goes negative, then perhaps a different range indication is called for?

*We changed the sentence to: "Elevation differences in large valleys of high streams orders (6–7) are small and **strictly positive**, indicating resampling-induced elevated valley floors." Note that this statement refers to high stream orders (6-7), where Figure 2 shows only positive values, especially for the box and whiskers.*

205 - In runs at all resolutions ?

Thank you for the suggestion, this expression appears a few times in the manuscript. We checked the whole manuscript and changed it to “runs at all resolutions” or “simulations at all resolutions” in these cases.

210 - Can this be a range?

Thank you for the comment, this question applies to other passages in the text as well (see also the next comment). This comment refers to: “The finest-resolution simulations (<300 m)…” where we used “<” or “>” for resolutions, which is confusing. We changed all text passages, where we used “<” and “>” for resolutions to a range (e.g., 50-300 m resolution) or write out the comparison (e.g., resolutions finer than 300 m). For this specific text passage, we used “simulations at 300 m resolution and finer”.

215 - This whole paragraph could do with tightening up a bit (sorry I know it's always hard to strike a balance between detail and brevity !). I think the reason it's a little hard to follow is that the values in question are jumping around a bit and the distinctions can feel a bit arbitrary. I.e. sometimes it's 1000 m and the 300 and 400 m, and then it's <500 and >800 . And then at line 221 it goes to 600 and 800. Good to be consistent. May also suggest just three exact values, rather than ranges, which can be a bit harder to describe. I thought about moving this to a major comment but it should be easy to rectify. I do think consistent resolution ranges across the paper are important/useful though !

*Thank you for this comment. We have now structured this paragraph into two smaller paragraphs, where we first focus on differences in ice volume between the resolutions. In the second paragraph, we focus on hysteresis effects. The separation into two paragraphs did not increase the word count. Instead, by removing unnecessary descriptive parts, we reduced the word count from 487 to 341. Moreover, we used consistent resolution ranges for fine (**300 m** and finer) and coarse resolutions (coarser than **800 m**) throughout the paper and in these paragraphs to avoid confusion.*

217 - The sentence beginning on this line is quite unclear

We agree and removed this sentence as it did not add much new insights.

221 - I think I got what's going on here, but it required more interpretation than I would like. the '-' between warming and cooling is a minus sign? At first glance it looks like punctuation. So either an explicit 'minus' or using more obviously mathematical variables for warming and cooling would be good.

*Thank you for the suggestion, we modified the figure accordingly. We structured the text passage into two sentences to enhance clarity: “Larger and persistent **volume differences between the warming (3000–5000 years) and cooling phases (1000–3000 years)** in the coarse resolution simulations occur for temperature forcings stronger than -3 K and thus during large parts of the model run (~1750–4250 years). At resolutions coarser than 800 m, ice volumes* *during the warming phase are more than 60 % higher than during the cooling phase under the same temperature forcing relative to the maximum ice volume (Fig. 3c)." Additionally, we changed the text box in the figure to: "warming* ***minus*** *cooling phase".*

223 - % and deg C brake convection by coming directly after the number

Thank you for the comment, we checked all units and aligned them with the conventions.

225 - So in effect, the higher resolution simulations are much more sensitive to climate change -- that could be a headline finding !

Indeed, we added a concluding sentence to emphasize this finding: "The shorter delays in reaching the maximum ice volume as well as smaller volume differences between the warming and cooling phases at fine resolutions indicate that ***fine resolutions capture changes in temperature forcing more accurately****." Additionally, we also (slightly) expanded on this in the discussion, section 4.2.*

255 - The first two sentences of this paragraph feel a little winding, maybe there is a punchier way to arrive at your descriptions ? I will confess I did not manage to read all of the paragraphs around this point word for word because they are very dense.

We agree and changed the beginning of this paragraph to: "Resolution-related deviations in ice volume that depend on altitude arise from spatially non-uniform ice thickness patterns. At full glaciation, ice thickness maps differences relative to the 50 m reference run reveal distinct spatial patterns between fine and coarse resolutions (Fig. 5; full version in Fig. S4)."

264 - I guess this comes back to my earlier comments, it would be nice to have more consistency in resolutions comparisons, that might make a headline figure /narrative of 'use this resolution if you would like to avoid the worst resolution issues'. I do accept that reality will be a little bit more complicated, but here a little more synthesis is not unwarranted

Thank you for your comment, the first result presented in this paragraph refers to fine resolutions and indicates that 300 m is a sufficiently fine resolution for our exemplary ice field: "Among ***resolutions of 300 m and finer, ice thickness is generally comparable****, and differences only increase slightly with coarser resolution (Fig. 5a, b)." Generally, in the revised manuscript, we now focused on the contrast between model results at fine (300 m and finer) and coarse resolutions and concluded that 300 m is a sufficient resolution for our setup.*

267 - And below? So, it only becomes a clear piedmont glacier as resolution decreases from 900 - 500 m

Thank you for raising this question. At fine resolutions, the length of this specific glacier is too short to form a piedmont-type terminus (Fig. S5a). However, we consider the observations of piedmont-type glaciers a minor detail and removed this part from the manuscript to focus on main aspects.

282 - I would argue this should be reinterpreted. We can say that the lowest resolution is the most accurate, so then the importance is the drift away from that, the difference between the resolutions is then perhaps of more supplementary interest. I would lessen the number of panels (even if having a panel o is cool), but also include one that shows the actual absolute field at 50 m resolution for comparison purposes.

As mentioned above, we removed this paragraph and figure from the main text. All comparisons in the revised manuscript are with reference to the 50 m run and we refrained from comparisons between the resolutions. We reduced the number of panels to 6. A figure with absolute ice thicknesses can be found in the supplements (Fig. S5a).

288 - when compared. This sentence is very long-winded.

We changed Fig. 7 by keeping panel a showing ice volume at bedrock elevations (for all resolutions) and adding ice area and mean ice thickness with respect to bedrock elevation (see updated version at the end of the document). We tightened the beginning of this paragraph to: "A closer look at the distribution of ice across bedrock elevations shows how thickness and area influence ice volume at different spatial resolutions (Fig. 7)." Because we compared to the 50 m reference run and not between resolution, we removed the original sentence.

290-390 - I should be honest and say that I am having a very difficult time reading through all this and I am giving it a good go. I have managed a skim but I think my main recommendation would be a focussed rewrite. For example, you switch to ice temperature at 327 after a lot of text about slope values and DEMs. That certainly should be its own paragraph, or grouped in a shorter paragraph with something clearly closely related. All this to say, I expect this can be really significantly cut down to focus on the main takeaways, and a similar approach applied to the rest of the paper. If you put in a few easy to decipher figures they can do the talking, and the results text can be much reduced.

Thank you for this comment. As mentioned above, we removed Fig. 6, 9, panel 8f and the corresponding paragraphs. Additionally, we removed repetitions and descriptive sentences to present main takeaways more directly. This is reflected in the reduced word count from 3215 to 2449 (~25 % less) in the results of the revised manuscript.

Discussion:

As with the rest of the paper, this could be very much streamlined, with a big focus on what the main takeaways are. I think this could go from 8 to at most 5 paragraphs, hopefully less. Implications, and tying this in to the rest of the literature should really be emphasised. There are few citations in this section.

Thank you for this comment. We reduced the word count in the discussion is reduced from 2295 to 1570 in the revised manuscript (more than 30 % less), while including aspects raised by the *reviewers (e.g., a discussion on deviations between Blatter-Pattyn and full Stokes simulations). Although we did not reduce the number of paragraphs, we believe that the shorter paragraphs strengthen the focus on a few main points. Specifically, we shortened section 4.1 and combined the paragraphs on mid and low altitude.*

405: The last sentence of this paragraph is useful (could actually be expanded so that reference back to Fig. 5 is not necessary). I think the rest can be restructured around distance from 50 m and the points you really want to emphasise .

As mentioned above, all comparisons of resolutions are now with respect to the 50 m run. The spatial pattern of resolution differences between mountain tops and valleys is discussed in section 4.1, which is now shorter and more focused in the revised manuscript.

418 - Here and elsewhere you can drop 'apparent'. You're controlling every variable in this model so you can say this with a modicum of conviction.

Good point, we checked for this term and dropped it where relevant.

421 - Ditto, seem.

We agree and checked the whole paragraph.

423 - Basal melting will only happen if the temperature is at the pressure melting point, so this statement is a bit vague.

*Thank you for this comment. The higher probability for basal melting is just one possible explanation for the connection between higher ice thickness and warmer temperatures at coarse resolutions (lines 423-426). Our results in Fig. 8 show that mean basal temperatures (pressure-adjusted) at coarse resolution are at 0 °C and thus closer to the pressure melting point than at finer resolution. However, since these are mean temperatures, we argue it is justified to be a bit vague about this. For clarification we rewrote this part: "The shift from generally frozen bed conditions at fine resolution simulations to temperate basal ice at coarser resolutions is possibly a consequence of two mechanisms: colder climatic conditions in the fine resolution DEMs due to higher surface elevations, and increased thickness at coarser resolutions resulting in a **higher likelihood for basal melting due to increased basal pressure and thus a lower pressure melting point**."*

424 - Taking an uncomplicated 1D column shear is a \rho g h sin(\alpha), so it's a function of height and slope. Greater thickness doesn't necessarily mean equal or greater slope.

Thank you for the clarification. It is true that we cannot conclude that internal deformation is higher at coarser resolution due to thicker ice, especially because mean slopes are indeed lower (Fig. S6, former S10). We removed the corresponding text passage from the revised manuscript.

425 - Could be clearer.

*We agree and changed the sentence to: "The shift from generally frozen bed conditions at fine resolution simulations to temperate basal ice at coarser resolutions is possibly a consequence of two mechanisms: **colder climatic conditions in the fine resolution DEMs due to higher surface elevations**, and increased thickness at coarser resolutions resulting in a higher likelihood for basal melting due to increased basal pressure and thus a lower pressure melting point."*

426 - Better in results ? Link of temperature to Arrhenius is well known, so you could just focus on temperature.

We agree and as mentioned above, we removed the parts about the Arrhenius factor to enhance the focus of the paper.

428 - This definitely does not need a presumably (softness dependence on temperature)

We agree and we removed the analysis of softness (see comment above).

432 - 437 - Results ? But please think carefully about adding to the results !

This text passage is removed as it is already part of the results.

437 - Repetition ?

Good point, we removed this from the discussion.

445-447 - Results ?

This is already part of the results and thus we removed it here.

447 - 449 - Cool

Thank you for this comment.

452 - 455 Results ? or Not that relevant

We agree and moved the specific number for velocity differences to the results. Additionally, this paragraph is shortened and the discussion on the velocity among fine resolutions is reduced from three sentences to: "Among the finest resolutions, the reduction in ice velocity (Fig. 8) might result from increased lateral shear stress from very steep and narrow valley walls that impose resistance."

459 - Again with seems, hopefully you can be a bit more definite about this. Interesting discussion tough

We agree and changed this sentence: "This feedback is most likely strongest in valleys with steep and narrow sidewalls at fine resolution and partly offsets ice thinning due to acceleration from channelized flow, indicated by consistently thicker ice at low altitudes than at high and mid altitudes - across all resolutions (Figs. 7, 8)."

471 - Half as fast, maybe twice as slow, but not half faster or twice slower, somehow. Probably 'when the rate of temperature forcing is reduced by half'

The corresponding sentence has been removed.

4.2 Opening sentence is results

We agree, the opening sentence has been removed and the paragraph starts now with: "Our simulations show that the glaciers' response to changes in temperature are more accurately captured in the fine resolution simulations, where glaciers react more rapidly to transitions from cooling to warming periods (Figs. 3c, 9)."

I'm going to stop going through this line by line here. Hopefully some of the issues are clear from where I have gone line by line above! There is too much mixing of results and discussion. Each paragraph should have a really clear aim and much more direct. For example the hysterisis effect is cool, but then the main thing to emphasise is that hysterisis is reduced for higher resolutions. That, and its implications should be the main focus of this paragraph, not describing that 'up to 420 years compared to a time lag of less than 200 years at the finest resolutions'

Thank you for the provided line-by-line comments, we checked the rest of the discussion and removed the parts that are describing results to focus more directly on the discussion.

Conclusions:

A conclusion table could be useful. A lot of the bullet points are different ways of saying 'at high resolution this' 'at low resolution this' 'difference is this'. Standard prose styles dictate that we don't repeat language, which I suppose is reasonable, but that really would be useful here because otherwise one has to search through the sentences for the relevant piece of information.

Thank you for the suggestion. We tried to make a compact conclusion table but ended up repeating results for fine and coarse resolutions in respective columns (with one column oppositional to the other).

We instead addressed this comment by rewriting the conclusions and presenting them in a clear continuous text rather than bullet points.

Figures

Please see major comments, but I do feel that the total number of figures can be streamlined into a smaller number of punchier ones. For me the ones which are very useful are 7 and 3 (and maybe something like 6 but for distance from 50 m, and without requiring quite so many panels), but that's just a quick thought !.

Fig. 1 Panel a could be made a bit more recognisable to people through the use of a terrain colourmap and including a couple of city locations -- just a suggestion though.

Thanks for the suggestion, however, we do believe that Figure 1a is already quite crowded. Moreover, the standard terrain colormap (see below) looks less aesthetic for our taste, and is less colour-blind friendly, that is why we use a custom colormap.

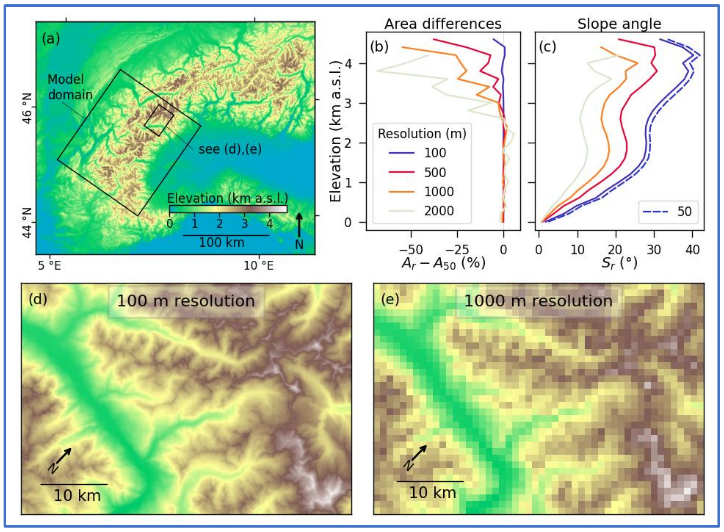


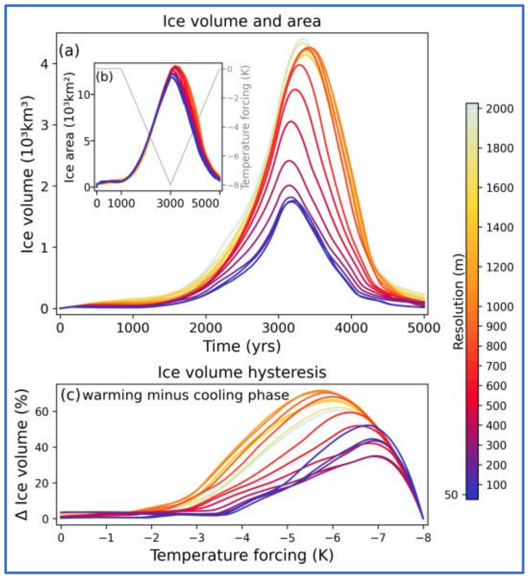
Fig. 2 - It would be nice to see an example (if not the entire region) with labelled stream order so that one can get a feel for which stream order refers to which type of feature.

Thank you for the suggestion. We added such a figure, see above (page 12).

Fig. 3 - I accept this is a bit pedantic on my part, but could the colour bar be flipped to be in ascending order? Also, 2000 is hard to see so minor suggestion to clip at 1,800 colour (don't worry if that's too much hassle)

Panel c, can't this just be described as the difference as the maximum ice volume will cancel if both are relative to that.

Thank you for the suggestion, here is an updated version of Fig. 3 with a flipped colorbar:



We agree that the graph would look the same if we used [ice volume at warming phase "minus" ice volume at cooling phase] compared to that same difference as a percentage of total ice volume. However, we argue that the values of xx % more ice volume in the warming than in the cooling phase relative to total ice volume might be more meaningful than xx km² more ice volume in the warming than in the cooling phase. The reader would have to compare with figure 3b to see if xx km² difference in ice volume is actually a lot or not.

Fig. 5 This is a valuable figure, but I think it probably could/should be transported to the supplementary material -- a reduced version showing the general trend or important jumps with a bit less white space could reside here

We agree and created a reduced version that compares ice thickness differences compared to the 50 m reference run:

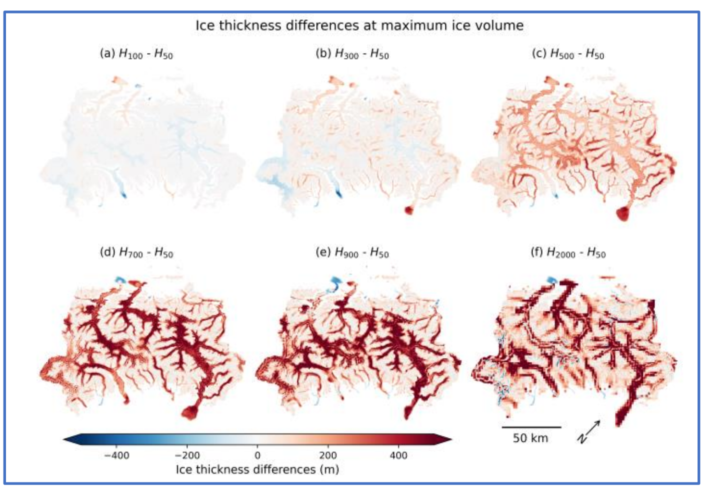


Fig. 6 - The pedant within me would like to see Fig. 5 and 6 following the same layout, and I prefer the more compact Fig. 6 layout.

As mentioned in the comment above, we used a more compact layout for Fig. 5. We removed Fig. 6 from the main text to reduce the number of figures.

Fig. 7 - I don't exactly understand the positive negative distinction here. All the positive and negative cells binned together ? I would consider omitting that.

We changed Fig. 7 and omitted the differences, so there is no distinction in negative and positive values. Here is the updated figure:

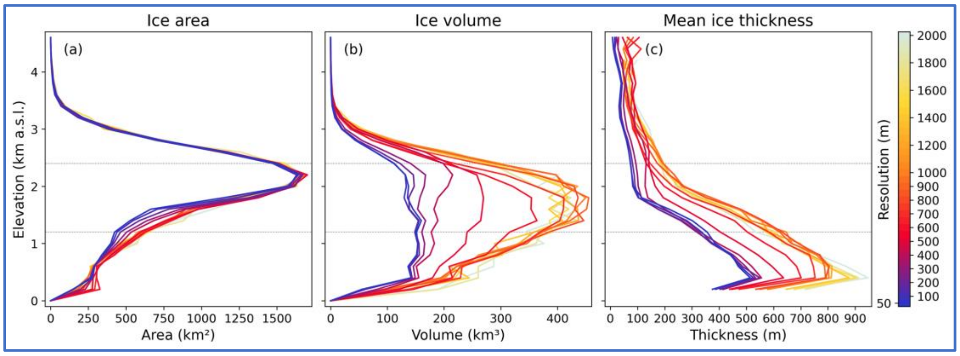
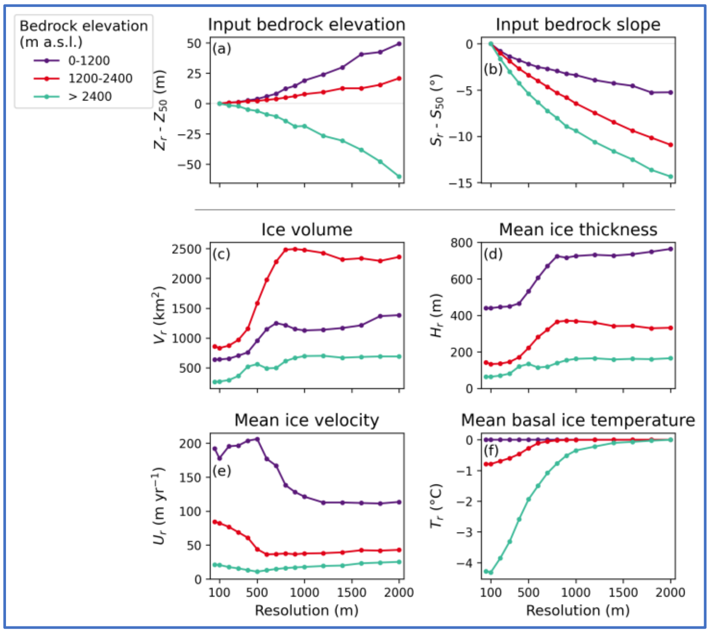


Fig. 8 - Can you not develop some improved notation for r? Such as r∈{50,100:100:1000,1200:200:2000} defined early on, and then saying for each r. Writing out the full range at every mention is a bit much. Why is there a grey horizontal line at the top ? I think to separate model inputs from outputs, but then good to label that in the figure. Otherwise a figure like this really does tell story well (though each title should be 'mean ice thickness' etc.). A panel for total ice volume would be good, too.

Thank you for the suggestion. We agree that listing all resolutions is unnecessary long. Since the individual values for each resolution are visible from the figure (with resolution on the x-axis and values marked as circles) and resolutions used for our simulations are listed in the methods, we do not think it necessary to mention which resolutions are shown.

As mentioned above, we think that the analysis of the mean Arrhenius values does not add much value and instead show maximum ice volume. Here is the updated figure:



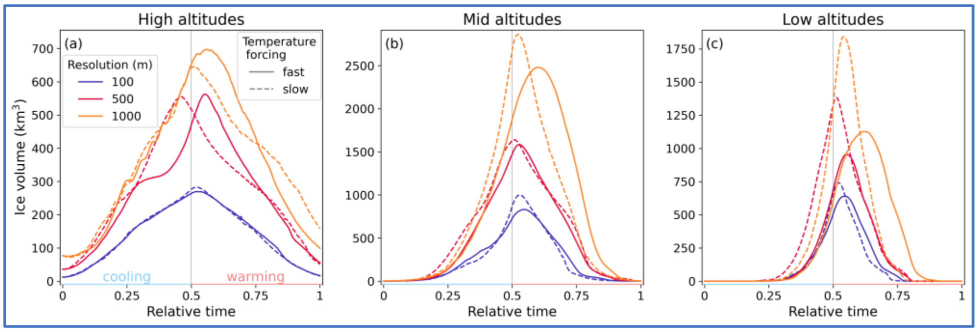
We added a comment in the figure caption that explains that the grey vertical line distinguishes between model inputs (a, b) and outputs (c-f) shown.

Fig. 9 The Hr - H50 approach is good. There should be a description on the colour bar. The images are a bit too small.

Thank you for the suggestions. As mentioned above, we removed this figure from the revised manuscript.

Fig. 10, I'm having difficult with this as it feels like temperature is just a proxy for time here? A data axis should not decrease and then increase again. I would suggest separating this out into two rows and then using two x axis, one for temperature, and another for time. The trend is broadly consistent, so you could also drop the high, mid, low separations.

Thank you for this comment. Since we applied a linear temperature forcing, model years and temperature values are very closely related. Because in this figure we show results for experiments with different temperature rates, the cooling and warming phases are not equally long. For example, with fast temperature forcing, the cooling phase ends at 3000 years and with slow temperature forcing, it ends at 5000 years. Therefore, we put relative time on the x-axis, from 0 (start of cooling phase with a temperature forcing of 0 K) to 1 (end of cooling phase with a temperature forcing of 0 K). At 0.5, half of the simulation is reached, where temperature forcing peaks at 8 K and the warming phases begin.



We argue to keep the distinction between altitudes, which shows interesting details such as different rates of ice volume change with time as well as differences in the timing of maximum ice volume at bedrock altitudes.

Supplement:

Given the quite major comments with the rest of the manuscript I have not extensively explored the supplement. It's nice to see all the parameter values in there. You could put some figures from the main text in here, though even the supplement shouldn't be too long and wavering.

Thank you for this comment. We checked the supplementary and reduced the number of figures from 15 to 9, while adding two supplementary figures to address reviewers' comments.

References

Aschwanden, A., Fahnestock, M. A., and Truffer, M.: Complex Greenland outlet glacier flow captured, Nat Commun, 7, 10524, https://doi.org/10.1038/ncomms10524, 2016.

Cuzzone, J. K., Schlegel, N.-J., Morlighem, M., Larour, E., Briner, J. P., Seroussi, H., and Caron, L.: The impact of model resolution on the simulated Holocene retreat of the southwestern Greenland ice sheet using the Ice Sheet System Model (ISSM), The Cryosphere, 13, 879–893, https://doi.org/10.5194/tc-13-879-2019, 2019.

Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., and Vertačnik, G.: The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, International Journal of Climatology, 34, 1657–1675, https://doi.org/10.1002/joc.3794, 2014.

Ivy-Ochs, S.: Glacier variations in the European Alps at the end of the last glaciation, CIG, 41, 295–315, https://doi.org/10.18172/cig.2750, 2015.

Jouvet, G. and Cordonnier, G.: Ice-flow model emulator based on physics-informed deep learning, J. Glaciol., 69, 1941–1955, https://doi.org/10.1017/jog.2023.73, 2023.

Jouvet, G., Cohen, D., Russo, E., Buzan, J., Raible, C. C., Haeberli, W., Kamleitner, S., Ivy-Ochs, S., Imhof, M. A., Becker, J. K., Landgraf, A., and Fischer, U. H.: Coupled climate-glacier modelling of the last glaciation in the Alps, J. Glaciol., 69, 1956–1970, https://doi.org/10.1017/jog.2023.74, 2023.

Leger, T. P. M., Jouvet, G., Kamleitner, S., Mey, J., Herman, F., Finley, B. D., Ivy-Ochs, S., Vieli, A., Henz, A., and Nussbaumer, S. U.: A data-consistent model of the last glaciation in the Alps achieved with physics-driven AI, Nat Commun, 16, 848, https://doi.org/10.1038/s41467-025-56168-3, 2025.

Liebl, M., Robl, J., Egholm, D. L., Prasicek, G., Stüwe, K., Gradwohl, G., and Hergarten, S.: Topographic signatures of progressive glacial landscape transformation, Earth Surf Processes Landf, 46, 1964–1980, https://doi.org/10.1002/esp.5139, 2021.

Mey, J., Scherler, D., Wickert, A. D., Egholm, D. L., Tesauro, M., Schildgen, T. F., and Strecker, M. R.: Glacial isostatic uplift of the European Alps, Nat Commun, 7, 13382, https://doi.org/10.1038/ncomms13382, 2016.

Penck, A. and Brückner, E.: Die Alpen im Eiszeitalter, Tauchnitz, 540 pp., 1909.

Rückamp, M., Kleiner, T., and Humbert, A.: Comparison of ice dynamics using full-Stokes and Blatter–Pattyn approximation: application to the Northeast Greenland Ice Stream, The Cryosphere, 16, 1675–1696, https://doi.org/10.5194/tc-16-1675-2022, 2022.

Seguinot, J., Ivy-Ochs, S., Jouvet, G., Huss, M., Funk, M., and Preusser, F.: Modelling last glacial cycle ice dynamics in the Alps, The Cryosphere, 12, 3265–3285, https://doi.org/10.5194/tc-12-3265-2018, 2018.

Williams, C. R., Thodoroff, P., Arthern, R. J., Byrne, J., Hosking, J. S., Kaiser, M., Lawrence, N. D., and Kazlauskaite, I.: Calculations of extreme sea level rise scenarios are strongly dependent on ice sheet model resolution, Commun Earth Environ, 6, 60, https://doi.org/10.1038/s43247-025-02010-z, 2025.

Winkelmann, R., Martin, M. A., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C., and Levermann, A.: The Potsdam Parallel Ice Sheet Model (PISM-PIK) – Part 1: Model description, The Cryosphere, 5, 715–726, https://doi.org/10.5194/tc-5-715-2011, 2011.

Comment on egusphere-2025-3870 (referee #2)

We would like to thank reviewer 2 kindly for taking the time to read our manuscript and for the constructive feedback, which helped us to improve the study. In the following, we addressed all comments point by point and point out respective corrections. Please find our answers in blue italic font.

General Assessment

Overall, the manuscript is well-written and the sections are clearly defined. The topic is relevant and timely, given the increasing computational capacity that now allows high-resolution simulations of mountain glaciers. The approach is interesting, and the study addresses an important question: how spatial resolution affects the simulation of ice thickness, flow, and thermal regimes in mountainous regions.

However, while the manuscript is organized into clear sections, some aspects of the results and discussion are difficult to follow, partly due to reliance on references to supplementary material. Certain patterns and the interpretation of figures could be presented more cohesively in the main text.

Besides, there are aspects that require clarification or more explicit justification. While the authors explore a wide range of resolutions (50 m to 2 km), the study relies on a single model (the Instructed Glacier Model, IGM). It is not clear whether generalizations about the impact of resolution on mountain glaciers are fully warranted based on this single framework. Explicit discussion of model limitations and how the resolution interacts with the underlying Blatter-Pattyn approximation, compared to other approaches (full Stokes or others), would strengthen the study.

We thank the reviewer for the general assessment. We tightened and shortened the revised manuscript to enhance clarity. The number of supplementary figures is reduced from 15 to 9, while adding two supplementary figures to address the reviewers' comments. We added slightly more background information on IGM and its functionality to put our results into perspective.

For a detailed description of IGM and comparisons to an analytical solver, we would like to refer to Jouvet and Cordonnier (2023) and point out that the retraining frequency is the most critical parameter that ensures that the ice flow does not deviate too much from the Blatter-Pattyn model. We changed the corresponding text passage to: "For adjustments to new ice field states attained through run time, we retrained the physics-informed neural network multiple times within a single model year (every ~0.18 years). The on-the-run retraining in the highest-resolution simulations at 50 and 100 m was performed more often (every ~0.05 years). ***The retraining frequencies were found to be high enough to ensure that the emulator's deviation from the analytical solver typically remains below 5 m yr^{-1} in ice velocity*** *(see Figs. 4, 5 in*

Jouvet and Cordonnier, 2023; Leger et al., 2025)." To justify using IGM, which depends on the Blatter-Pattyn model, rather than a full Stokes model, we referred to Pattyn et al. (2008) who showed that the Blatter-Pattyn model generally reproduces ice velocities that agree well with full Stokes models.

*We would like to clarify that the IGM framework builds on the Blatter-Pattyn model which cannot be replaced by other ice physics approximations or a full Stokes implementation. Due to substantial higher costs of higher-order models, it is computationally too expensive to run our experiments with a numerical, standard and CPU-based Blatter-Pattyn or with a full Stokes model. For example, a quick calculation based on previous experiments from Jouvet et al. (2022) (Table 4) yields that a single simulation with our 5 kyr-long simulation setup at 50 m resolution would take about one year to run using Elmer Ice, assuming efficient parallelization over 60 CPU cores. Therefore, it would be impossible to directly compare our experiments to corresponding simulations using a full Stokes model. However, we briefly discussed potential deviations between the Blatter-Pattyn and full Stokes model at high resolutions in section 4.3 of the revised manuscript: "**At fine resolutions that are required to accurately model ice field dynamics, the Blatter-Pattyn approximation implemented in IGM may deviate from the full Stokes solution and exhibit mechanical inaccuracies**, particularly in steep terrain (Rückamp et al., 2022; Yan et al., 2023; Hindmarsh, 2004). The presented study deliberately focuses on numerical rather than mechanical errors and demonstrated that substantial numerical errors already occur at spatial resolutions commonly used in large-scale mountain ice-cover simulations (Mey et al., 2016; Seguinot et al., 2018; Jouvet et al., 2023; Golledge et al., 2012; Zhang et al., 2022). **For a Greenland ice stream region modelled by Rückamp et al. (2022), velocity deviations between the Blatter-Pattyn and full Stokes model of less than ~10 m yr^{-1} appear smaller than the resolution-induced differences of up to several tens of meters per year in our analysis.** Therefore, adopting a full Stokes formulation would substantially increase computational costs without addressing the main source of error. Future studies are needed to systematically assess the benefits of full Stokes models at fine resolutions in alpine regions."*

*In order to strengthen confidence in IGM, we added that Leger et al. (2025) compared IGM simulations to a traditional model: "**A direct comparison** of a simulation of the European Alps at 2 km resolution using IGM to the respective simulation using the well-tested and widely used Parallel Ice Sheet Model (PISM; Winkelmann et al., 2011) published by Jouvet et al. (2023) **showed minor deviations**, confirming that the alpine ice field is a suitable application for IGM (Leger et al., 2025)." Additionally, we believe that a more extensive comparison to other studies on resolution effects in Greenland and Antarctica (Williams et al., 2025; Rückamp et al., 2022; Aschwanden et al., 2016; Cuzzone et al., 2019) helped to generalize our findings (see other comments below and the review by referee #1 for further details).*

Introduction

- The introduction's opening paragraph reads as a series of descriptive statements rather than a connected narrative. Consider revising for smoother flow.

We agree and made changes to this paragraph to present the importance of ice models to model present-day climate change as well as paleo-reconstructions with a smoother flow.

- The knowledge gap and justification for the study could be clarified more explicitly. For example, are there other studies on the effect of the spatial resolution using similar or different approaches or applied to a different setting?

Thank you for this raising this important point. We added a paragraph to the Introduction to clarify the knowledge gap more explicitly: "The horizontal resolution of numerical ice models has been shown to influence simulations of continental-scale ice sheet mass loss and estimated sea-level contributions, particularly in regions with high bedrock elevation gradients ***(Cuzzone et al., 2019; Rückamp et al., 2020; Williams et al., 2025; Aschwanden et al., 2016****). In order to address resolution-related model inaccuracies, Williams et al. (2025) suggest resolutions of 5 km or finer for modelling regions in the West-Antarctic ice sheet. For Greenland, Cuzzone et al. (2019) propose a resolution of at least 10 km, Rückamp et al. (2020) report convergence of ISMIP6 (Ice Sheet Model Intercomparison Project; Nowicki et al., 2016) simulations at 1 km or finer, and Aschwanden et al. (2016) finds that a resolution of ~600 m is required to model most Greenland outlet glaciers in agreement with observational data…"*

Methodology

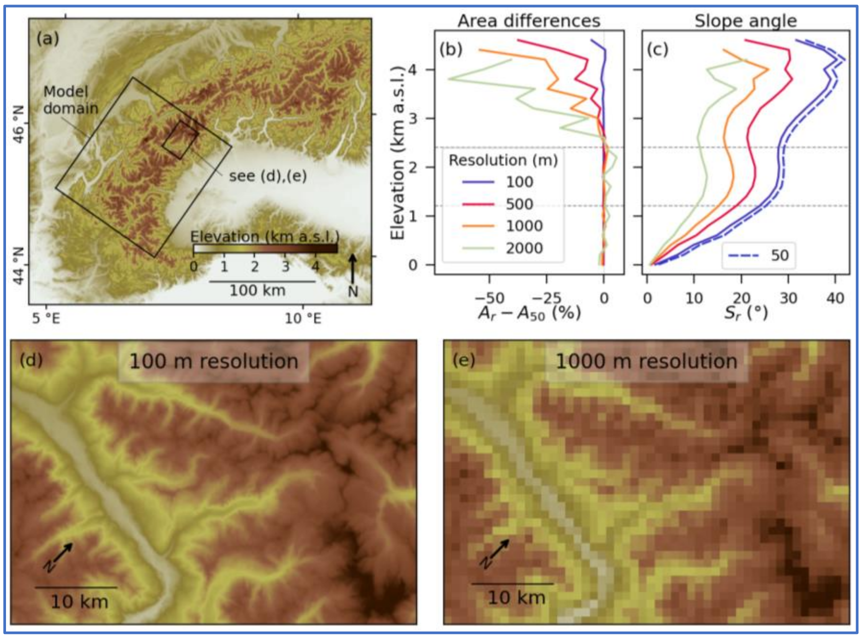
- The choice of model parameters requires more justification. For example:
    - The retraining frequency of the neural network is mentioned, but no explanation about the selection of the values is provided. Explaining why this frequency was chosen would help readers understand its impact.

*Thank you for pointing out the unclarity regarding IGM model choices. All our parameter choices are based on Leger et al. (2025) - except for the higher retraining frequency for the finest-resolution simulations. With IGM's is physics-informed strategy, a higher retraining frequency permits to maintain fidelity level of the Blatter-Pattyn model. In the revised manuscript, we gave the retraining frequency in model years instead of model iterations, which is more understandable. We referred to Jouvet and Cordonnier (2023) and Leger et al. (2025) to justify the parameter choice: "For adjustments to new ice field states attained through run time, we retrained the physics-informed neural network multiple times within a single model year (****every ~0.18 years****). The on-the-run retraining in the highest-resolution simulations at 50 and 100 m was performed more often (****every ~0.05 years****).* ***The retraining frequencies were found to be high enough to ensure that the emulator's deviation from the analytical solver typically remains below 5 m yr^{-1} in ice velocity (see Figs. 4, 5 in Jouvet and Cordonnier, 2023; Leger et al., 2025).****"*

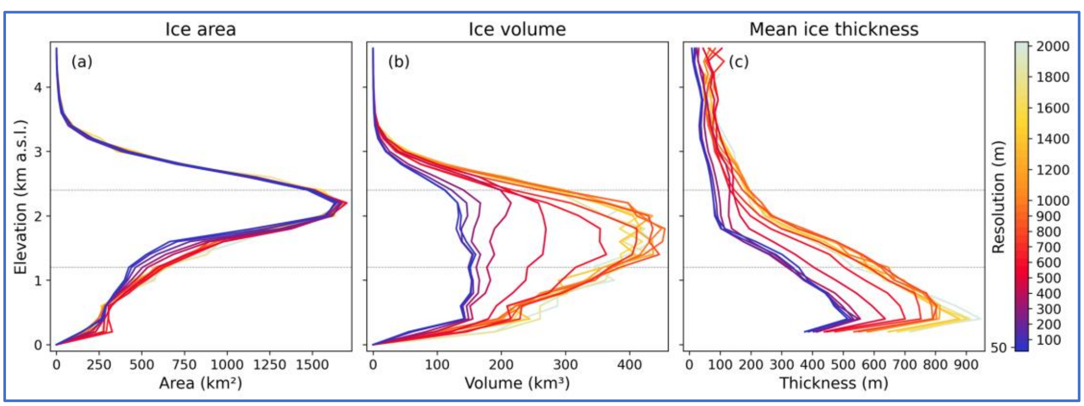
- Elevation bands for low, mid, and high altitudes are defined as 0–1200 m, 1200–2400 m, and >2400 m. A brief justification or reference would improve clarity.

Thank you for raising this point. In the revised manuscript, we clarified the idea behind the elevation bands: "We used elevation bands of 0–1200, 1200–2400, and >2400 m a.s.l. to distinguish between different altitudes that correspond to deeply incised, gently sloping valleys at low altitudes and steep mountain summits at high altitudes, and a transitional domain, where much of the modelled ice volume is located (Figs. 1a, c, S1)." Additionally, we added a supplementary figure S1 showing the distribution of elevation bands of the model domain.

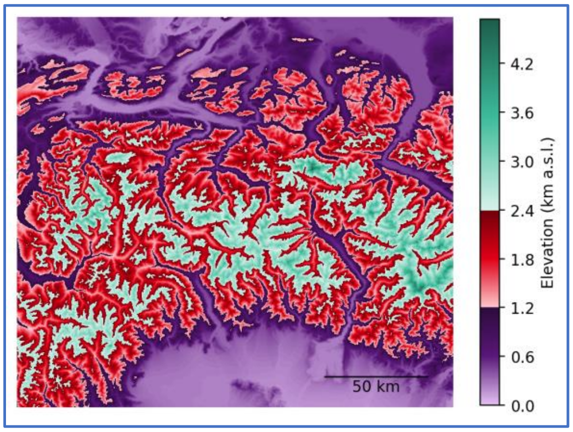
We updated Fig. 1b and c and inserted vertical lines that mark our defines elevation bands. Fig. 1c shows that the defined elevation bands correlate with changes in slope angles, with steep slopes at high altitudes and low slopes at low altitudes.



Additionally, we added the vertical lines to the updated Fig. 7 that shows ice area, volume and mean thickness at bedrock altitudes. This enables us to show results across bedrock elevation instead of using the broad elevation bands for all figures:



We also added a map to the supplements showing the distribution of bedrock across the elevation bands:



- While the study explores a wide range of resolutions, all resampling is performed using cubic interpolation. Would the authors expect differences by using other techniques (bilinear, kriging…)? Could this lead to some misleading results?

We added a justification for the cubic resampling method in the revised manuscript: "For all DEM resampling of input fields, ***we use cubic convolution due to its higher accuracy in resolving surface elevation and slope angles compared to other resampling methods like bilinear or kriging interpolation*** *(Minh et al., 2024)." Consequently, we expect slightly higher differences between fine and coarse resolution DEMs using other resampling methods, which might lead to even stronger resolution effects. However, we consider differences due to resampling methods to be minor compared to the overall resolution effects analyzed.*

Please note that we used cubic resampling for the preparation of the input DEMs, while for post-processing comparisons of our model outputs at different resolutions, we instead used nearest neighbour resampling to preserve the pixel structure at coarse resolutions (see section 2.3).

Results

- Some aspects of the results are difficult to interpret solely from the figures, particularly given the reliance on supplementary material. Consider highlighting the key patterns in the main text to improve readability.

We fully agree and tried to improve the readability of the results section. We reduced the number of figures (removing former Figs. 6 and 9) and generally shorten the paragraphs to highlight the most important aspects. We believe that the key patterns are visible in the figures already provided in the main text. We also reduced the number of supplementary figures from 15 to 9, while adding two supplementary figures to address the reviewers' comments.

- While the manuscript identifies three resolution "modes", the evidence is not entirely clear except for Figure 8 and not for all the outputs and elevation bands.

Thank you for the comment. In line with the comments from reviewer 1, we removed the term "modes" and focused on the comparison to the 50 m reference run in the revised manuscript. We emphasized resolution effects at fine and coarse resolutions (e.g., generally thinner and faster-flowing ice at fine resolutions), instead of mainly focusing on strong changes at intermediate resolutions.

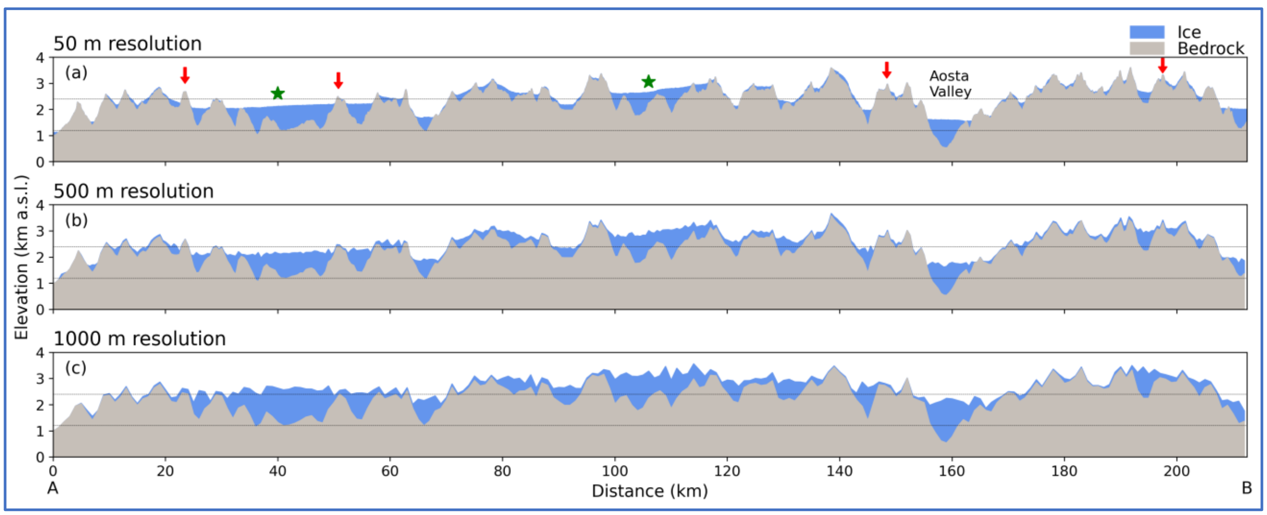
However, we argue that evidence for the three resolution "modes" is shown not only in Figure 8. For example, the maximum ice volume can be divided into rather consistently low ice volume at fine resolutions and consistently high ice volume at coarse resolutions, with high increase in ice volume at intermediate resolutions (Fig. 3).

Discussion

- The discussion is generally relevant but could be more explicitly connected to practical implications. For example, explaining how these findings could guide modelers in selecting appropriate spatial resolutions would increase the impact (other than just going for finer resolution as a general rule).

Thank you for this comment. As mentioned above, our comparisons to the 50 m reference run shows that model results at 300 m and finer are generally comparable. Therefore, we concluded that a resolution of 300 m is sufficiently fine to accurately simulate glacier dynamics in our exemplary ice field in the Western Alps. Removing the confusing "resolution modes" and instead identifying a "sufficiently fine resolution" is already more useful from a practical perspective. Our study might therefore provide a valuable reference for determining a resolution for a mountain region similarly steep to the Western Alps.

Additionally, we added a Figure showing ice surfaces along a transect at different resolutions:



This figure indicates that glacier surfaces might be unrealistically rugged at coarse resolutions, which might be a feature that modellers should pay attention to. We added this idea to the discussion: "Specifically, checking if surfaces of glacier flowing over mountain summits are realistically represented might be an initial indicator, as models at coarse resolutions might show imprints of bedrock topography onto the ice surface (Fig. 7; Fig. 5 in Leger et al., 2025; supplementary data in Jouvet et al., 2023 and Seguinot et al., 2018). Such tests can be performed on small subsets and therefore require limited computational resources."

Moreover, the experiments using a smoothed topography show similar resolution-effects that depend on altitude and are non-linear. From those experiments, we concluded that a resolution of 400-500 m is sufficiently fine to accurately model the icefield in smoother terrain than the Western Alps. We added a discussion of different resolutions for ice-sheets (Greenland, *Antarctica) in context of our experiments for the Western Alps to provide further guidance for modellers.*

- I feel that maybe more detailed exploration of this 400-800 “critical” mode should be addressed before to clearly talk about strong sensibility.

As pointed out above, we removed the “resolution modes”. Strong changes at intermediate resolutions are described more explicitly instead, which adds clarity to the results and discussion. Since we focused more on differences compared to the 50 m reference run, we found it more valuable to contrast ice field characteristics at fine (generally thinner glaciers and faster ice flow) and coarse resolutions – instead of focusing on intermediate resolutions.

Conclusions

- I like the bullet structure, but I think that some points could be rearranged for more clarity.

Thank you for this comment. We rewrote the conclusion section and presented the conclusions in a continuous text, which we think is clearer and easier to follow.

Figures and Supplementary Material

- While the figures themselves are informative, the main text often relies on references to supplementary material to interpret the results. This can make the narrative difficult to follow. Consider ensuring that the key patterns and findings are clearly described within the main text, using the supplementary figures as support rather than as primary evidence.

Thank you for the suggestion. As mentioned above, we limited the number of supplementary figures. We presented our results in a more concise way in the revised manuscript without excessive reference to supplementary figures for details that we won’t show in the main text.

Minor Comments

Here I will try to skip the comments raised from the other reviewer so we don’t need to repeat the same things.

L108: Please, justify this choice of (0.01-0.04).

The time step in IGM is adaptive and not static, i.e., it is not a parameter chosen by the user but instead calculated in IGM based on the CFL condition that ensures numerical stability for the ice thickness evolution. We removed the values for a typical time step and retraining time step frequencies from the manuscript and instead gave the retraining frequencies in units of model years to focus on the retraining frequency instead of the time step. For the choice of retraining frequencies, we added references as justifications: ” For adjustments to new ice field states attained through run time, we retrained the physics-informed neural network multiple times *within a single model year (**every ~0.18 years**). The on-the-run retraining in the highest-resolution simulations at 50 and 100 m was performed more often (**every ~0.05 years**). **The retraining frequencies were found to be high enough to ensure that the emulator's deviation from the analytical solver typically remains below 5 m yr^{-1} in ice velocity (see Figs. 4, 5 in Jouvet and Cordonnier, 2023; Leger et al., 2025).**"*

L178: Please, justify the selection of the elevation bands.

As mentioned above, we clarified the idea behind the elevation bands: "We used elevation bands of 0–1200, 1200–2400, and >2400 m a.s.l. to distinguish between different altitudes that correspond to deeply incised, gently sloping valleys at low altitudes and steep mountain summits at high altitudes, and a transitional domain, where much of the modelled ice volume is located (Figs. 1a, c, S1)." Additionally, we added the elevation ranges to panels b and c of Fig. 1 to visualize that the elevation bands distinguish between gentle-sloping regions at 0-1200 m a.s.l., areas with intermediate slopes at 1200-2400 m a.s.l., and steep regions above 2400 m a.s.l. (see updated Fig. 1 above).

L184: effects -> affects?

Yes, we adopted this in the revised manuscript.

L209: This 10 times more is not clear to me, at least from the cited figures.

*Thank you for highlighting the need for clarification. Due to the y-axis, it is not really visible in the figure that the 2000 m simulation has 10 times more ice volume than the 50 m run at 1000 years of model time. However, it is clear in the figure that at 1000 years, the coarse resolutions have (relatively speaking) much more ice volume than the fine resolutions. We included values for the ice volume in the text: "After the initialization phase (0–1000 years), the 2000 m simulation has **~160 km^2** of ice volume which is more than **10 times** that of the 50 m run (**~12 km^2**)."*

L254: It is not clear to me looking at Fig. S12.

Thank you for pointing this out. We apologize for the typo and refer to Fig. S5a that shows the spatial distribution of ice thickness at full glaciation.

L426-7: I really don't see those greatest relative changes at high altitudes. Probably you have the numbers, but looking at Fig. 8f such a thing is not clear to me.

Thank you for the comment. As mentioned above, the analysis of the Arrhenius factor does not add many insights, and we removed this from our analysis to focus on key variables (thickness, velocity, temperature). We would like to add that the mean Arrhenius factor at high altitudes is ~66 M Pa^{-3} yr^{-1} for the 50 m run and ~112 Pa^{-3} yr^{-1} for the 2 km run, leading to a difference of almost 60 % relative to the 50 m run.

L429-31: Again, this is not obvious to me just looking at Fig. 8.

Similar to the previous comment, we removed the Arrhenius factor parts from the manuscript.

L474-8: Nothing to object here, the opposite, really clear statement and good point resulting from your work.

Thank you very much.

L500: Here you suggest to go to finer than km- and ideally 500-m scale- but your defined critical mode ranges from 400 to 800m. I found this a little bit contradictory.

Thank you for spotting this inconsistency, this should be changed to "400-m-scale".

Overall Recommendation

Even if this is a relevant and well-executed study that addresses an important question in glacier modelling, especially in regional or global studies, I consider the manuscript as major revisions after considering some points:

1. Clarifying model limitations and generalization potential.
2. Justifying some methodological choices, such as elevation bands, retraining frequency, and resampling method.
3. Streamlining results and discussion to highlight the most important findings and improve readability. Exploring a bit more this critical mode sensibility.
4. Enhancing figure clarity and reducing overreliance on supplementary material.

I believe that addressing these points will help the resulting paper become a solid and valuable contribution to the field.

Thanks for the valuable suggestions, which we hope we have addressed in the above-mentioned modifications.

References

Aschwanden, A., Fahnestock, M. A., and Truffer, M.: Complex Greenland outlet glacier flow captured, Nat Commun, 7, 10524, https://doi.org/10.1038/ncomms10524, 2016.

Cuzzone, J. K., Schlegel, N.-J., Morlighem, M., Larour, E., Briner, J. P., Seroussi, H., and Caron, L.: The impact of model resolution on the simulated Holocene retreat of the southwestern Greenland ice sheet using the Ice Sheet System Model (ISSM), The Cryosphere, 13, 879–893, https://doi.org/10.5194/tc-13-879-2019, 2019.

Golledge, N. R., Mackintosh, A. N., Anderson, B. M., Buckley, K. M., Doughty, A. M., Barrell, D. J. A., Denton, G. H., Vandergoes, M. J., Andersen, B. G., and Schaefer, J. M.: Last Glacial Maximum climate in New Zealand inferred from a modelled Southern Alps icefield, Quaternary Science Reviews, 46, 30–45, https://doi.org/10.1016/j.quascirev.2012.05.004, 2012.

Hindmarsh, R. C. A.: A numerical comparison of approximations to the Stokes equations used in ice sheet and glacier modeling, Journal of Geophysical Research: Earth Surface, 109, https://doi.org/10.1029/2003JF000065, 2004.

Jouvet, G. and Cordonnier, G.: Ice-flow model emulator based on physics-informed deep learning, J. Glaciol., 69, 1941–1955, https://doi.org/10.1017/jog.2023.73, 2023.

Jouvet, G., Cordonnier, G., Kim, B., Lüthi, M., Vieli, A., and Aschwanden, A.: Deep learning speeds up ice flow modelling by several orders of magnitude, J. Glaciol., 68, 651–664, https://doi.org/10.1017/jog.2021.120, 2022.

Jouvet, G., Cohen, D., Russo, E., Buzan, J., Raible, C. C., Haeberli, W., Kamleitner, S., Ivy-Ochs, S., Imhof, M. A., Becker, J. K., Landgraf, A., and Fischer, U. H.: Coupled climate-glacier modelling of the last glaciation in the Alps, J. Glaciol., 69, 1956–1970, https://doi.org/10.1017/jog.2023.74, 2023.

Leger, T. P. M., Jouvet, G., Kamleitner, S., Mey, J., Herman, F., Finley, B. D., Ivy-Ochs, S., Vieli, A., Henz, A., and Nussbaumer, S. U.: A data-consistent model of the last glaciation in the Alps achieved with physics-driven AI, Nat Commun, 16, 848, https://doi.org/10.1038/s41467-025-56168-3, 2025.

Mey, J., Scherler, D., Wickert, A. D., Egholm, D. L., Tesauro, M., Schildgen, T. F., and Strecker, M. R.: Glacial isostatic uplift of the European Alps, Nat Commun, 7, 13382, https://doi.org/10.1038/ncomms13382, 2016.

Nowicki, S. M. J., Payne, A., Larour, E., Seroussi, H., Goelzer, H., Lipscomb, W., Gregory, J., Abe-Ouchi, A., and Shepherd, A.: Ice Sheet Model Intercomparison Project (ISMIP6) contribution to CMIP6, Geosci. Model Dev., 9, 4521–4545, https://doi.org/10.5194/gmd-9-4521-2016, 2016.

Pattyn, F., Perichon, L., Aschwanden, A., Breuer, B., de Smedt, B., Gagliardini, O., Gudmundsson, G. H., Hindmarsh, R. C. A., Hubbard, A., Johnson, J. V., Kleiner, T., Konovalov, Y., Martin, C., Payne,

A. J., Pollard, D., and Price, S.: Benchmark experiments for higher-order and full-Stokes ice sheet models (ISMIP–HOM), The Cryosphere, 2008.

Rückamp, M., Goelzer, H., and Humbert, A.: Sensitivity of Greenland ice sheet projections to spatial resolution in higher-order simulations: the Alfred Wegener Institute (AWI) contribution to ISMIP6 Greenland using the Ice-sheet and Sea-level System Model (ISSM), The Cryosphere, 14, 3309–3327, https://doi.org/10.5194/tc-14-3309-2020, 2020.

Rückamp, M., Kleiner, T., and Humbert, A.: Comparison of ice dynamics using full-Stokes and Blatter–Pattyn approximation: application to the Northeast Greenland Ice Stream, The Cryosphere, 16, 1675–1696, https://doi.org/10.5194/tc-16-1675-2022, 2022.

Seguinot, J., Ivy-Ochs, S., Jouvet, G., Huss, M., Funk, M., and Preusser, F.: Modelling last glacial cycle ice dynamics in the Alps, The Cryosphere, 12, 3265–3285, https://doi.org/10.5194/tc-12-3265-2018, 2018.

Williams, C. R., Thodoroff, P., Arthern, R. J., Byrne, J., Hosking, J. S., Kaiser, M., Lawrence, N. D., and Kazlauskaite, I.: Calculations of extreme sea level rise scenarios are strongly dependent on ice sheet model resolution, Commun Earth Environ, 6, 60, https://doi.org/10.1038/s43247-025-02010-z, 2025.

Winkelmann, R., Martin, M. A., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C., and Levermann, A.: The Potsdam Parallel Ice Sheet Model (PISM-PIK) – Part 1: Model description, The Cryosphere, 5, 715–726, https://doi.org/10.5194/tc-5-715-2011, 2011.

Yan, Z., Leng, W., Wang, Y., Xiao, C., and Zhang, T.: A comparison between three-dimensional, transient, thermomechanically coupled first-order and Stokes ice flow models, Journal of Glaciology, 69, 513–524, https://doi.org/10.1017/jog.2022.77, 2023.

Zhang, R., Zhang, Z., Jiang, D., Ramstein, G., Dupont-Nivet, G., and Li, X.: Tibetan Plateau Made Central Asian Drylands Move Northward, Concentrate in Narrow Latitudinal Bands, and Increase in Intensity During the Cenozoic, Geophysical Research Letters, 49, e2021GL093718, https://doi.org/10.1029/2021GL093718, 2022.