Dear Editorial team,
thank you for handling our manuscript with kind attention and care. We appreciated your dedication to it. You can find our replies to your comments below, divided by Editor, Reviewer 1 and Reviewer 2. Your text will be highlighted in *italics.*

We mainly:

- improved results and discussion about data interpretation subjectivity, as well as focus and meaning of the study.
- improved GPR profiles, by indicating where they intersect between each other, and by improving figures' resolution.
- added the original GlabTop2 model coded in IDL, alongside the python implementation of GlabTop2 that we have previously used. The original one performs better near glacier margins (see profile 2022, for example, in the left part).
-performed a general English revision.

Sincerely,
Andrea Vergnano and the Authors

Editor

I have taken over this manuscript from the previous editor because the latter has unexpected urgent work obligations. I have read the manuscript, the two recent reviews and the three initial reviews.
At this point, there are two reviews of which one (#2) recommends minor revision and is generally very positive while the other (#1) is very critical and recommends rejection. I understand the critique raised by #1 and acknowledge that they consider the required revision as too substantial for recommending major revisions. However, I also acknowledge that there are four other reviews who recommend revisions (major or minor). In this challenging situation I take the following decision:
- I ask the authors to revise the manuscript according to both reviews. Review #1 needs to be fully addressed and the authors need to reply to each point raised by the reviewer, explaining how they will address the critique.
- I find the circular reasoning pointed out by #1 particularly concerning. I ask the authors to clearly address this issue. I noticed that Reviewer 3 from the first round of reviews had already commented on that issue ("However, it is important not to turn the logic around [...] reversing the process and leaking the modeling data into the GPR measurements can be delusive. This is the most important comment I have for this work, and I would like to see it discussed further in the manuscript.") It appears to me that the authors have not answered this critique which has now been repeated by reviewer #1 from the second round of reviews.

Dear Editor,
thank you for the time you spent carefully reading our manuscript and previous reviews. We understand the difficult situation, and we are certainly satisfied of this rigorous peer-review process.

It is difficult for us to completely address the critique raised by reviewer 1, because we agree with them: there is leakage of model information into GPR data. Compared to the

first submission, in the last major revision, we had clarified this issue better (maybe not better enough), without denying it, addressing the first-round-reviewer#3 comments.

What we can try to do now, is to explain why we think that our methodology does not indulge in faulty “circular reasoning” thinking, but instead, it is simply a way to interpret the GPR and the models together.

In our opinion, every joint interpretation of two different methodologies, in a context where measurements are of indirect nature and subjective choices must be made, necessarily involves looking at the information of one to understand the information of the other, or, using Reviewer 3 words, leaking data from one to another.

Yes, we use model data to interpret GPR, and GPR to calibrate model data in a sort of basic, manual iterative process. The GPR and model techniques are on the same level in this difficult context: we are not proposing something like "a main technique and a secondary one supporting the interpretation of the main one". They both support each other, in our view, and the proposed methodology is just one way to merge them together.

Finally, we think that, thanks to the “open review” model, this debate can be valuable for the interested readers, to form their personal opinion about the validity of the proposed method. We decided to highlight these main points into the discussion section, in which we would like to express a humble and multilateral perspective about the proposed methodology.

In this regard, we added a little extra part in this revision, to address the lack of inter-analyst comparison raised by Reviewer 1.
We thought a lot about this, and we came up with this idea: in mid September 2025 we went to a scientific conference about GPR (https://www.gpritalia.it/gpritalia2025/), and we made a collective activity in which we asked the listeners to pick the bedrock on our GPR sections, before and after looking at the model estimations. We included a paragraph about this activity, as a qualitative inter-analyst comparison into the paper.

Furthermore, a few editorial comments:

Lines 135 to 138: The text gives the impression that the authors are using the original GlabTop2 code. However, this is not the case as the original code was written in IDL. This code has been presented in the study by Frey et al. (2014) and used extensively in e.g. Farinotti et al. (2019). The authors must make clear that here a code is used which is based on the equations presented in Frey et al. (2014), not the original code. The code used here appears to be an incomplete implementation of the equations and description in Frey et al. (2014). The code appears to simulate large ice thickness directly at the glacier margin, as visible in Fig. 4. The original code contains algorithms to avoid such unrealistic sudden transitions from thick ice to no ice. Being the author of the original code, I would be happy to share the original code with the authors, however, its major limitation is that it is written in IDL and I fully understand if the authors prefer using a more modern Python code.

Thanks for the proposal. We were indeed unsatisfied about the problem at the glacier margins. Unfortunately, we did not manage to run the original code in IDL, due to its closed-source nature, but we thank you to run it for us. We now present the paper as a

four-model comparison, instead of three. We stated in the methodology section that the Glabtop-py model is an open-source, partial implementation of the Original-Glabtop2 model.

Lines 335 – 336: While I do not know how the original GlabTop2 code would have performed on Rutor glacier, this issue would likely have been reduced.
Lines 376 - 377: This appears to be the only mentioning of circular reasoning, although in a single sentence which is kept vague and whose message on human error is unclear to me. To be clear, any circular reasoning needs to be removed from this study for it to become acceptable.

Please, refer to the lines above about what we think about the "circular reasoning". We acknowledge that this issue must be discussed more in depth. Now we included in the main text a more comprehensive discussion about it, mostly taking from the answer to your previous point, above.

Reviewer #1

Dear authors,
Thank you for submitting a manuscript to The Cryosphere. I regret that I have to recommend the rejection of the manuscript after you have already completed a major revision, but it still contains significant weaknesses, particularly in the overall study design. Please find the detailed review below.

Summary
This manuscript presents a case study that integrates Ground Penetrating Radar (GPR) data with three ice-thickness inversion models (GlabTop2, GlaTE, OGGM) to improve the identification of the ice-bedrock interface in the temperate Rutor Glacier in the Italian Alps. The aim is to reconcile discrepancies between past GPR-based ice volume estimates and more recent geodetic observations by visually incorporating model-derived information into the manual interpretation of GPR profiles. The GlaTE model is then re-run using these interpretations as constraints.

The approach is well described, and the manuscript is generally clear in its presentation. However, the study lacks methodological rigor, as it does not sufficiently validate or critically assess the techniques employed. Additionally, the work offers limited originality, with a considerable proportion of the analysis closely resembling existing studies without introducing significant advancements to the field.

Thank you for your clear and constructive feedback. We think that the issues you raised could be discussed and hopefully solved in a reasonable amount of time, therefore, we performed a major revision of our article, which we submit now to the editor and reviewers' attention. Please, see our answers below.

Major Concerns

1. Conceptual Circularity and Confirmation Bias
The study's core methodology—using model outputs to guide the interpretation of GPR data and subsequently constraining one of those same models with those interpretations—is inherently circular. This undermines the objectivity of the interpretation process and risks reinforcing existing biases. Although the authors briefly acknowledge this issue, they do not provide any quantitative means of mitigating it. The absence of an independent validation dataset (e.g., boreholes) impairs this problem, leaving the model outputs effectively validated by themselves.

We cannot object to your observation, all what you said is true. The point that we want to discuss is that we conducted our work with a perspective different from the classical "finding a method (in this case GPR) that can validate models".
Our global point of view is more like "finding an efficient way to merge two indirect, sparse and sometimes subjective methodologies".

In our opinion, every joint interpretation of two different methodologies, in a context where measurements are of indirect nature and subjective choices must be made, necessarily involves looking at the information of one to understand the information of the other, or, using Reviewer 3 words, leaking data from one to another.

Summarizing, yes, we use model data to interpret GPR, and GPR to calibrate model data in a sort of basic iterative process. In our workflow, the GPR and the models are on the same level: we did not want to present "a main technique and a secondary one supporting the interpretation of the main one". They both support each other, in our view, and the proposed methodology is just one way to merge them together. This point of view is not valid for all cases, in our opinion: sometimes you have good GPR data and in these cases you take GPR as the main technique, and others as supporting techniques.

So, in our opinion, we are not in a situation of faulty "circular reasoning", considering the problem from the point of view elaborated above. It's like having 2 indirect clues of a unsolved murder and trying to find the killer: you look at the first clue, and look at the second clue with the first clue in your mind, and vice versa. Then, you get closer to the solution.

After this more "philosophical" answer, we would like to point out some more practical considerations:

About boreholes, it is generally accepted that many glacier studies exist without always showing borehole data, due to their very high economic and logistical costs. We obviously would like to have borehole data supporting our study, but it was economically unfeasible. About the risk reinforcing existing biases, we tried to improve the discussion section in this sense in the "Advantages and limitations of the methodology" paragraph. We think that, however, one result of our work is that, even if the final model could be wrong in some parts, at least it avoided blatant errors that in the past were made, like interpreting reflections due to internal glacier features as bedrock. This kind of mistakes possibly happened also on other temperate glaciers. It seems to us that the GPR interpretation on temperate glaciers can be hard, and similar ideas of interpreting together models and GPR are relevant for scientific community and professionals. We also highlighted in the abstract that we specifically present a point of view of applied geophysicists that have to find ways to handle difficult to interpret GPR data.

2. Lack of Methodological Rigor in GPR Survey Design
The manual classification of radargram features into categories such as “sure,” “model-guided,” and “wrong” pickings is fundamentally subjective. The manuscript does not present any reproducibility metrics, inter-analyst comparisons, or uncertainty assessments based on signal characteristics. Given that these interpreted points are subsequently used to constrain a model and inform conclusions, this lack of rigor critically undermines the study's credibility.
The authors acknowledge widespread scattering and signal attenuation in the GPR data but do not appear to have made any effort to adjust the survey design in response. There is no discussion of antenna frequency selection, signal processing alternatives, or survey geometry optimization. Furthermore, the limited spatial coverage of the GPR lines in high-uncertainty areas is not addressed, nor is any attempt made to densify or extend profiles to improve data quality.

We fear that we are unable to publish quantitative reproducibility metrics on a subjective data interpretation issue. This subjectivity of data interpretation is unfortunately spread across many geophysical techniques, and we have to accept that any GPR data analysis that includes manual selection of interfaces or signal reflections is subjective. We are not willing to publish *the perfect, most statistically correct, subjective data interpretation,* which

actually sounds to us as a strange concept. We are proposing an idea to *reduce subjectivity thanks to model estimations.* Our results show that in this test site this idea is better than just GPR, because only GPR, as performed in the past (Villa et al., 2008), did not pass the geodetic mass balance proof.
About the inter-analyst comparison, we came up with this idea: in mid September 2025 we went to a scientific conference about GPR, and made a collective activity in which we asked the listeners to pick the bedrock on our GPR sections, before and after looking at the model estimations. This experiment was qualitative, but we observed that all the participants (about 10), before seeing the models, picked the internal reflections of the glacier as bedrock, mistakenly. This kind of picking would greatly underestimate the total ice volume of the glacier. After seeing the models, they recognized deeper reflections as the bedrock (unfortunately, only 3 of them. The others did not complete this second part of the exercise). We received positive feedback from the audience, who were particularly concerned that temperate glaciers could be so hard to interpret with GPR, and operators can easily make interpretation mistakes. We included a short report on this in the discussion section, in the "advantages and limitations" paragraph.
About the antenna, we performed the 2022 survey with a lower frequency antenna compared to the 2012 survey (40 vs 70 MHz),and was ground based vs helicopter-based, exactly for this purpose of critically evaluate antenna and survey selection. We noticed no improvement, unfortunately. This was already reported in the methods sections, but we highlighted better that we tested two different antennas.
About signal processing, this was one of the main Ph.D. topics of one of the authors, therefore, we had plenty of time and determination to try every (literally) processing step included in ReflexW software, in various combinations, in order to improve the radargrams. Simple processing as the one shown resulted in being the most effective. We specified better this in the methods section, just after the processing workflow of the GPR.

3. Exaggerated Claims About Accuracy and Reproducibility
The manuscript makes strong claims regarding the accuracy and robustness of its results, particularly the final GlaTE-constrained volume estimate of 450 million m³. However, these claims are undermined by the methodological circularity described above and the limited scope of the GPR data. The substantial discrepancy between this estimate and previous work (e.g., 150 million m³) is not subjected to critical scrutiny, nor is it convincingly reconciled.

The presence of a substantial discrepancy was calculated based on very simple DEM pixel-by-pixel differencing, clearly suggesting that the previous 150 million m^3 estimation is wrong. There is ample space in the introduction where this is discussed (we made a reference to Figure 1, in the introduction, that seems clear to us). In the paper, we investigate and we discover that probably bad GPR data and consequently bad interpretation were the issues. Our 450 million m^3 estimate is certainly closer to the actual value, based indeed on DEM observation, but also by comparing critically this results with other available datasets (e.g. consensus estimate from Farinotti) - which provides an estimate of more than 450 million m^3, but this is justified because they used a 2003 outline, much bigger than that of 2021.

We do not agree that we did a *strong claim regarding the accuracy*. In the discussion, we wrote "It is difficult to say if this very low standard deviation is indicative of real reliability of the GlaTE constrained model; nevertheless, 450 million m3 of ice, distributed as in Figure 6, represents the best estimate of the Rutor glacier thickness we could obtain with this methodology."

This does not seems a *strong claim,* to us at least. We actually paid a lot of attention, given the previous and these reviews, to ensure the reader understands the limitations of the study. We are ready to change this and similar phrases if the reviewer thinks differently.

4. Limited Geographical Scope and Sparse Dataset
The study is based entirely on a single glacier and relies on a limited amount of GPR data, collected along a relatively small number of profiles (relative to the glacier surface area). The narrow spatial focus and sparse dataset therefore constrain the strength of the conclusions. As such, the robustness of the proposed workflow remains difficult to assess.

According to us, the main result of this study should not be linked to the accuracy of the Rutor glacier bedrock model obtained, which, we agree with you, would be quite limited in geographical scope. We reworked the beginning of the discussion to clarify that this work should be viewed more as a methodological tool, than a study about the Rutor glacier. The GPR dataset is sparse, as it was collected from previous available data and we had limited capacity to carry out new surveys. However, we again think that this is not the main point of the paper. Conversely, being able to process difficult-to-interpret GPR data, and notwithstanding their sparse nature (which can be frequent in other surveys on glaciers, due to the often limited budget and logistical constrains), is exactly the scope of the methodology proposed.

5. Originality
The Cryosphere requires research of exceptional originality and novelty. Submissions should therefore "represent substantial progress beyond current scientific understanding." However, there appears to be another manuscript by Strallo et al. (2025) that presents a very similar methodological approach, albeit applied to a different Italian glacier. Despite this change in study site and the integration of other complementary datasets, the overarching design of the study remains largely unchanged. Notably, the Strallo et al. manuscript was submitted on 3 September 2024 — just two weeks after Vergnano et al. (2024) submitted their initial version to TC on 14 August 2024. Given that A. Godio is a co-author on both manuscripts and that A. Vergnano is explicitly acknowledged in the Strallo et al. paper, it is evident that there was awareness of the overlapping work. This raises concerns regarding adherence to established norms of good scientific practice. Furthermore, previous studies from recent years have already combined GPR data with ice thickness models to derive glacier-wide ice thickness distributions over substantially larger and more topographically complex areas (e.g., Grab et al.(2021), Helfricht et al. (2019)). While the current study applies similar principles, these earlier works provide a broader context and may serve as valuable benchmarks in terms of methodological scope.

We are concerned about your first comment about the other regional-scale studies. In fact, it is difficult to compete with such excellent studies. Fortunately, our focus is different: having to deal with a single glacier, (and thanks to the reviewers concerns that made us elaborate better in this direction) we explored much more in depth the subjectivity issue, the intercomparison between models, and we present a point of view of a GPR analyst that have to deal with scattered data, a situation in which many readers of TC may encounter. Let's say that we explore a side-aspect of such more general studies, and given the implications that it had on the Rutor glacier, we think that it's worthy of reading, because similar issues frequently arise on local temperate glaciers.

To better state the literature context in which our paper is situated, we added a short paragraph at the beginning of the discussion, to introduce the aims of our discussion:

"We discuss a methodological approach of interpreting together models and sparse GPR data, with specific reference to the results obtained on the Rutor glacier, which experienced a wrong estimation of glacier thickness due to the scattered GPR data (Villa et al., 2008). GPR has been the favorite tool to investigate glacier thickness in the last decades, however, we observed poor performance on temperate, meltwater-rich glaciers. At the same time, models that estimate glacier thickness based on surface topography and ice-flux equations are now of widespread use, and have been used to perform regional studies, for example on the Austrian glaciers Helfricht et al., 2019) and on Swiss glaciers (Grab et al, 2021) the latter including large-scale GPR surveys.

We focus on a side-aspect of such general studies: the specific, but not uncommon, situation of dealing with very scattered and unclear GPR data on temperate glaciers. We compare different ice-thickness models and literature products, and evaluate how they can help GPR analysts before and after a new GPR survey. Then, we discuss the subjectivity in interpreting GPR data, which can bring to possible human errors, especially in difficult contexts, and how models can be integrated in the interpretation process. This subjectivity is intrinsic in GPR data interpretation, but may raise concerns, therefore, we try to discuss the different aspects of this issue, applied to our situation. "

A few comments about Strallo et al. Methodologically, our and their papers are different, in our opinion, because this paper explicitly focus on the problem of subjective interpretation of "bad" GPR data on temperate glaciers, which is not the main topic of Strallo et al. The present paper focuses on the Rutor Glacier and specifically addresses the challenges of interpreting GPR data in temperate glaciers with high englacial water content. It compares four ice-thickness models (OGGM, GlabTop2-Py, Original-Glabtop2, GlaTE) and emphasizes the synergy between model outputs and GPR data to improve the identification of the ice-bedrock interface. The study also includes a sensitivity analysis and provides an open workflow. It highlights the importance of combining multiple approaches to reduce uncertainty and improve accuracy under challenging conditions. It revises previous Rutor Glacier ice volume estimates upward by integrating model estimates with GPR constraints.

Instead, in the paper by Strallo et al. (2025), the authors adopted an integrated approach for the Indren Glacier, combining historical data, GPR measurements, remote sensing, temperature analysis, and modelling (GlaTE algorithm) to assess glacier thickness changes over two decades. Their results enhance the understanding of glacier dynamics and inform water resource management and hazard mitigation.

The methodologies may be similar, but using established methodologies such as GPR and models does not necessarily mean that the paper is outdated. In fact, we see GPR studies continue to publish, probably because they explore different aspects of the methodologies or new focuses or new applications. For us, the important is that the focus of our and Strallo's papers are completely different. We suppose that the readers may be interested in both discussions, since they are two different focuses and two different points of views.

Minor Concerns
Language and Style: The manuscript requires careful editing for grammar, syntax, and clarity. Specific issues include incorrect sentence constructions, and inconsistent punctuation (e.g., inconsistent use of the Oxford comma). The authors are encouraged to

consult with a native English speaker or use AI-assisted editing tools to improve readability.

Thank you. We carried out a general editing for grammar, syntax and clarity as you suggested.

Recommendation: Reject

This manuscript, in its current form, does not meet the high scientific standards for publication in The Cryosphere. A complete methodological redesign would be required to render the work suitable for reconsideration. Additionally, including a few additional study sites with varying glaciological characteristics would greatly strengthen the credibility and generalizability of the proposed approach. Moreover, glacier thickness models—while valuable—should not be used as a substitute for reliable observational data or to compensate for poor GPR data quality.

After this revision, we hope you could appreciate a little our perspective, which may not be a classical one, but may represent a reasonable point of view to consider when interpreting difficult datasets. Our point of view is that of a geophysicist that have to handle with difficult data to interpret, not of a glaciologist that aims to calculate a regional ice-thickness model.

References:

Grab, M., Mattea, E., Bauder, A., Huss, M., Rabenstein, L., Hodel, E., … Maurer, H. (2021). Ice thickness distribution of all Swiss glaciers based on extended ground-penetrating radar data and glaciological modeling. Journal of Glaciology, 67(266), 1074–1092. doi:10.1017/jog.2021.55

Helfricht et al. (2019): Calibrated Ice Thickness Estimate for All Glaciers in Austria. Front. Earth Sci., 12 April 2019, https://doi.org/10.3389/feart.2019.00068

Strallo, V., Colombero, C., Troilo, F., Mondardini, L., & Godio, A. (2025). Glacier thickness modelling and monitoring with geophysical data constraints: A case study on the Indren Glacier (NW Italy). Earth Surface Processes and Landforms, 50(1), e6068. https://doi.org/10.1002/esp.6068

Reviewer #2

Re-review of "Integrating GPR and ice-thickness models for improved bedrock detection: the case study of Rutor temperate glacier" by Andrea Vergnano et al. (2025)

I am very excited to see this revised version, and commend the authors for the considerable effort they have invested in revising this manuscript. The new manuscript addresses almost all of my earlier concerns, and it is now written much more clear and well structured. Specifically, the research goals, hypotheses and scientific contribution are now stated clearly throughout the manuscript, the manuscript reads well, and the updated figures are much easier to read. I also really appreciate the inclusion of the comprehensive supplementary material, making these codes and the processing flow available is a valuable service to the community.

Lastly, I really appreciate the addition of pick classifications ("sure", "guided by models", "wrong"), as it reveals how manual and model-assisted interpretations interact. I have one suggestion that could further highlight the impact of the model inclusion when interpreting GPR data: I suggest adding a brief quantitative summary that exploits the pick categories, e.g. reporting the percentages of "wrong", "guided" and "sure picks" relative to the total GPR data. These could simply be mentioned in brackets throughout the text that is already there. I believe that the necessary data for this analysis now exists, and this should be straightforward.

Thank you very much for your positive and constructive feedback, yes, we put a lot of effort in revising this manuscript, as also the reviewers did. We included your suggestion about the pick categories into the text.

Apart from this minor suggestion and a few trivial language suggestions below, I believe the manuscript is ready for publication and recommend acceptance after these have been addressed.

Line-by-line comments

Thank you very much for your kind line-by-line comments, for improving our phrases. We appreciated it a lot.

L9-11: I suggest rewording to: "Combined visualization of the GPR and model data in 2D and 3D…, especially where the GPR data contained scattering noise and interpretation was uncertain."

Thank you, done.

L15: Remove "whole", and maybe write in present tense, e.g. is made openly available…

Thank you, done.

L44: I suggest slightly re-wording to: "To study and address the issue in detecting bed returns obscured by scattering noise in temperate glaciers…"

Thank you, done.

L59-60: I suggest clarifying this a bit, e.g. the statement in the Results section is much clearer. "However, … 2008 to 2021. If the previous ice volume estimate was correct, Rutor Glacier would have lost 2/3 of its volume meaning that only 50 million m3 still remain. However, while shrinking is evident, the remaining ice volume is likely larger given the large glacier area size." (please check if this is what you mean)
Thank you, we changed the text as you suggested.

L61: I suggest specifying that the ice volume was underestimated in your hypothesis, e.g. "… ice volume estimate of 150 million m3 is underestimated, …"

Thank you, corrected.

L114: I suggest removing "really" L137: remove "little"

Thank you, done.

L218: replace "whole" with "entire"

Thank you, replaced.

L238: Replace "whole" with "broader"

Thank you, replaced.

L239: I suggest replacing "probably" with "likely"

Thank you, replaced.

L257-258: Could remove that last sentence from the paragraph, as it is a bit of repetition.

Thank you, removed.

L334: I suggest changing to: "… which is close to the GlaTE…"

Thank you, done.

L335: You could replace glacier "outline" with "margin"

Thank you, replaced.

L340: Replace "quite" with "which is"

Thank you, replaced.

L341-345: I suggest reformulating this paragraph a bit to make it easier to read and make the message a bit more clear, for example something along: "Using these existing ice thickness model products to help interpreting the GPR data would be simpler than running the three models as performed in this study. However, these regional scale models may have outdated glacier outlines (which are changing rapidly)." (Please check if this is the message you wanted to convey. In terms of model uncertainty, I would expect that also the regional scale existing models have some sort uncertainty provided, but I am not familiar with them).

Thank you for the suggestion, and we think you are right about the regional-scale uncertainty. We rephrased it to tell that in both cases (literature products or self-run models) it is important to check the model uncertainty.

L353: I suggest being more specific and write: "Therefore, including modeled ice thickness to guide the GPR interpretations could …."

Thank you, done.

L356: I suggest changing to "…recognized to cause less clear GPR sections …"

Thank you, changed.

L368-359: Instead of only explaining that you assigned "wrong" picks, I suggest giving a percentage of them (e.g. compared to the total GPR profile length) to further highlight how the models can help guide the GPR interpretation.
Thank you, we added these percentages and a simple calculation spreadsheet in the supplementary materials, folder number 8.

L373: I suggest rewording to "However, some drawbacks have to be considered."

Thank you, done.

L375-377: Could you clarify these two sentences a bit? What is meant by the "balance between GPR and models", and "leaking the models into the observations"? Do you mean the possibility of incorrect modeled ice thicknesses then leading to incorrect GPR interpretations (e.g. following a reflector that coincides with the model but is not the true bed)? I think this just needs some re-wording and being a bit more specific.

Thank you, it was pointed out also by the Editor and the other reviewer. We reworded and discussed more in depth this phrase. Now it reads:

"
However, the most delicate aspect which may limit this methodology is the balance between the reliability of GPR and that of the models. In fact, our approach is a joint interpretation of two different methodologies, in a context where data are of indirect nature and subjective choices must be made. We need the information of the models to better understand the scattered and sparse GPR data, and vice versa, in a sort of manual iterative interpretation process. We stress that, in this difficult context, we do not see the GPR as the main technique, and the models as a supporting information, but we see that they both support each other. However, the subjective interpretation is potentially prone to human error. In our opinion, the key to minimizing this potential error is to perform simulations on multiple models, display the GPR sections with multiple color scales and in both 2D and 3D environments, and acquire local knowledge about the studied glacier.
"

L407-408: I suggest rewording to: "This study highlights the benefit of combining the two worlds."

Thank you, done.

L408-410: This is just my personal preference, but I suggest instead of saying "GPR practitioners" and "modelers", rewording to e.g. "GPR applications" and "modeling studies".

We agree, changed.

L411-414: I suggest rewording to something along: "Thus, applying the workflow presented here to other glaciers with GPR data available could help calibrate regional models more accurately, …"

Thank you, changed.

L423: Specify which map, I assume the bedrock topography?

Yes, thank you. We specified it.