Dear Editor and Reviewers,

Thank you for the wonderful review you did to our manuscript. As we looked through your in-depth comments, we felt honored of such attention.

We changed the manuscript according to all your suggestions, with very few exceptions we provided explanation for. You will find your comments followed by our replies in italics.

We improved the Supplementary materials and the reproducibility of results, as well as the discussion section and presentation of results. We hope to have improved the manuscript, however, we are conscious that we could still need to further improve it based on your eventual future suggestions. Since now the supplementary materials are well organized to reproduce the workflow, new improvements should be faster and easier to add.

Best regards,
the Authors.

Editor

Dear authors,

thank you for sketching a path by which the comments provided by the three referees will be addressed in your re-submission. In particular, I appreciated the fact that you provided an overarching view for how you intend to address the individual comments, rather than providing one-by-one answers at this stage. Since I herewith invite you to re-submit a suitably revised version of the manuscript, anticipate that you will provide such a point-by-point reply upon resubmission.

In terms of contents of the replies, I would only like to comment on two points:
1) Both Reviewers #1 and #2 take issue with the study's perceived lack of novelty. Here, I would encourage you to address this point somewhat more vigorously than done in the present replies, especially when replying to Reviewer #1.
2) In my understanding of the comment "Ice thickness vs bedrock topography" by Reviewer #2, the reviewer is only asking to subtract the calculated ice thickness from the used surface DEM, as to obtain a bedrock topography. This is different from the request of investigating the potential differences between bedrock topographies obtained from surveys conducted at different times which, I agree with you, would not be feasible within this study. Since the reviewer's request seems reasonable to me and since a dataset on bedrock topography would certainly be valuable for future studies, I would support the reviewer's request.

In terms of further steps, all of the three reviewers recommended the article to be reconsidered after major revisions. In this sense, your revised version of the manuscript – together with your replies to the obtained comments - will be resent to review once available. I anticipate that only two (instead of three) referees will be contacted this time.

Best wishes, and thank you for the constructive way by which you have taken up the referees’ comments.
Daniel Farinotti

Dear Editor,
thank you for handling this review process which has taken us a lot of time; we hope it was well spent.
Thank you for pointing out the main issues raised by the reviewers. Since a lot of rework has been done by our side, let us describe here the main changes:

*- **Supplementary materials**. We based our review (and our answer to reviewers comments) on producing the most reproducible workflow possible. This long work of organizing (and reproducing from scratch) all the workflow, allowed us to understand better the contribution of this work and to present it better, in our opinion. We hope that the reviewers can take a look on the supplementary materials, to understand if everything was done correctly, especially for the OGGM workflow which was more complex for us to setup than the others. We studied all the OGGM tutorials but we are not sure that everything was done correctly.*
We organized the Supplementary materials in 9 folders, each for each step of our workflow, each with a readme.txt file to describe the folder's contents, similar to a little methodological section. The folders contain first the unconstrained models, then the GPR, then the combined visualization in 2D (and 3D, this part is new, as it was suggested by a colleague, Maurizio Ercoli, which we thank), then the model constrained by GPR, and finally a simple sensitivity analysis on this last model. We hope that the supplementary materials can be an added value for the continuation of this review process (for sure, it will be much faster for us to implement eventual new modifications to the workflow based on future reviewers suggestions) and for the paper itself, aiding reproducibility and scaling to other glaciers.

*- **Discussion**. The main points that we want to present, based on the reviewers comments and our increased experience gained during the sensitivity analysis and the organization of the supplementary, are:*
- The models alone present a non-negligible uncertainty given by the choice of the input parameters (glaciological or computational). This was computed by calculating the average and standard deviation on 3 models with default (or suggested by developers) parameters (+-100 millions m3). Also, the sensitivity analysis done on unconstrained GlaTE model with varying glaciological parameters (creep rate and basal sliding) showed a high degree on uncertainty (+-75 millions m3)
- The GPR alone in temperate glaciers can be very tricky due to scattering, nothing to explain here.
- Putting the two worlds together, we obtain a constrained GlaTE simulation that, even changing two main glaciological parameters (creep rate and basal sliding), does not show high standard deviation (+-5 millions m3). In particular, the GPR seems to help a lot in constraining the models near the outline of the

glacier, whereas the models are better than GPR where the ice is thickest.
- The process of putting the two worlds together requires good visualization tools to allow for the most reasonable picking of GPR ice-bedrock interface, especially avoiding gross mistakes. In the previous submission, we presented only a 2D visualization. Now we added also a 3D visualization to allow seeing the GPR profiles as they intersect between each other, as well as the model predicted bedrock, in a Paraview environment. During the building of the supplementary materials and the reproduction of all the workflow, we were sincerely helped due to the combination of 2D and 3D visualization tools and we acknowledged that some of our first pickings in the right part of profile 9 were wrong, inducing an overestimation in the glacier thickness in the final GlaTE model.

Figures: *apart from the reworking of existing figures as requested by the reviewers, we inserted new images that support our discussion, mainly: sensitivity analysis, bedrock topography, comparison with literature product, and we better visualized the manual pickings.*

Abstract and title: *we manually rewrote the whole abstract as suggested by Reviewer 2, and considering the additions made thanks to this revision round. It should be more focused and clear, now. Also we tried to make the title a little more clear on the research objectives.*

A little note: We updated the affiliation of the first author to reflect the current working position. The research work of Andrea Vergnano was carried out thanks to research contracts of both universities, so we feel it's right to include them both.

Thank you again,
sincerely
the Authors.

Reviewer 1

Review of the article entitled:

Ground penetrating radar on Rutor temperate glacier supported by ice-thickness modelling algorithms for bedrock detection

This study addresses the challenges of measuring ice thickness in temperate glaciers, such as the Rutor Glacier in the Southern Alps (Italy), in these glaciers the ice is at or near the melting point throughout its entire mass, including both the surface and deeper layers. This means the glacier contains both ice and meltwater, which makes it sensitive to temperature changes and contributes to faster melting. Meltwater then interferes with the clarity of Ground Penetrating Radar (GPR) signals. To improve the manual selection of bedrock profiles from GPR-based ice thickness measurements in such glaciers, the researchers combined GPR data with three open-source thickness inversion algorithms (GlabTop2, GlaTE, and OGGM), which estimate ice thickness based on surface topography and mass turnover. These models guided the manual selection of unclear or scattered GPR data for the Rutor Glacier. The study analyzed two new GPR datasets and produced a more accurate ice-thickness map using GlaTE (one of the algorithms, after selecting the correct bedrock profile with the aid of outputs from all three models). Authors then conclude that the glacier stored about 515 million cubic meters of ice in 2021, significantly higher than previous estimates. The authors claim that this methodology is replicable and can simplify future GPR surveys of temperate glaciers, particularly in noisy data conditions caused by meltwater.

Overall, the manuscript, methods, and results are well explained, however, I have several corrections to the current text. I find the study very creative and could have potential for the use of this type of data to validate and calibrate ice thickness inversion algorithms. However, my main concern lies with the novelty of the study and significant gaps in the methodology, such as the uncertainty quantification of model results and the use of OGGM in such a small-scale glacier-specific study, as well as not providing details of the set up used for the OGGM inversion. I could consider this study for publication, but only after the authors address my questions and make the necessary changes to the manuscript.

Major comments:

Novelty, Reproducibility, and Scalability

- After reading the manuscript, I find it difficult to see how this analysis effectively contributes to the broader challenge of providing ice thickness observations and distribution products that could be used to constrain, evaluate, or train model simulations, or new deep learning algorithms (e.g., The Instructed Glacier Model).
- Additionally, there is no discussion on how this method could be scaled to a regional level. Expanding the study across multiple temperate environments and numerous glaciers with GPR measurements (e.g., in Alaska, the Alps or South America) raises concerns about the efficiency of manually detecting multiple profiles. Such an expansion would likely require a more robust approach for parameter calibration, validation of ice thickness inversion by each algorithm, and the generation of a final thickness map using more than one algorithm.

  I am concerned about the scalability of this method, as the parameters used for this specific glacier may not be transferable to others. This approach heavily depends on both data quality and the accuracy of model outputs. The examples presented in the manuscript show that the three algorithms perform well on some profiles but less so on others. A sensitivity analysis of the algorithms, with varied parameters to assess their impact on the thickness profiles, would have been valuable. Additionally, a more detailed uncertainty estimation for the final thickness product is needed. This could have been addressed by combining the results from all three algorithms, not just GlaTE, and providing a standard deviation on the final ice thickness map, while also comparing the results to existing ice thickness inversions and volume products (e.g., Millan et al. 2022, Farinotti et al. 2019 and Cook et al 2023 – all available in OGGM).

Thank you for this valuable analysis. We agree with you completely that the scalability and sensitivity analysis of the results were not properly addressed in the previous submission.

We addressed your raised issues in these ways:

- The whole workflow was organized and made available in the Supplementary materials. This should enhance the reproducibility of results making possible to everyone to perform the same steps. We hope that this can be useful for providing more reproducible ice-thickness observations with our methodology, which can be useful for local studies or as calibration of regional studies. We hope that you can take a look to the supplementary materials, hoping to make this review more transparent, especially to discuss better if we handled the issues you raised about the sensitivity and scalability in a proper way.

- The three unconstrained models were rerun (in the meanwhile, a bug in the calculation of tau was found in GlabTop2 by another user, and we updated our OGGM installation to version 1.6, as suggested by the reviewer). The results were averaged and the standard deviation was calculated, obtaining about 100millions m3 of deviation.

- A simple sensitivity analysis was performed on the final GlaTE model, both in its unconstrained and constrained version, by varying its parameters that regulate the ice creep and the basal sliding. We found that the constrained model show excellent reproducibility even with varying the input parameters in a reasonable range, while the unconstrained model showed much higher variability.

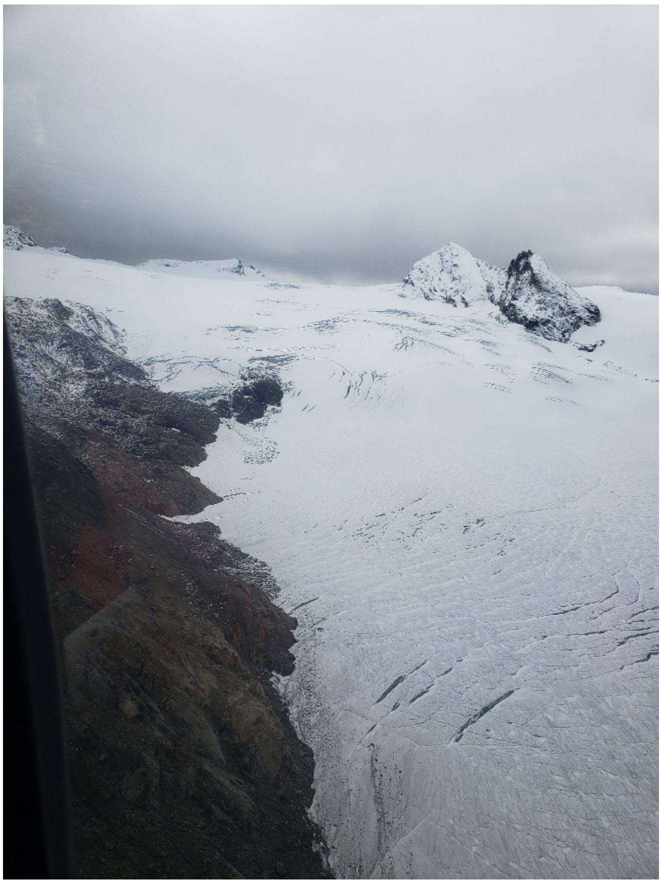
This is what we learned from your review and from our rework of our manuscript:
In our opinion, models alone have a high uncertainty based on which parameters we chose as input, whereas the GPR suffer from scattering in temperate glaciers. Despite the subjectivity and limitations, we think that the manual processing and picking of GPR profiles is a need to accurately depict a glacier geometry, especially near the glacier outline, which often requires local knowledge of the specific glacier and its recent history.
The strength of this methodology is to allow us to analyze a glacier in a more effective way, but requires local knowledge and manual analysis and visualization of results, given the uncertainty in both models and GPR. The outcomes of this laborious work, if performed on a bunch of relevant glaciers with GPR data available, could in turn help calibrate regional models more accurately.

<u>Methodology:</u>
Related to the data input used:

- Regarding the GPR survey conducted by helicopter, I wonder if the authors need to correct for signal reflection from the nearby mountain terrain and elevation changes - i.e., interference caused by the surrounding mountain slopes in the radargram?
- Is the outline from the Randolph Glacier Inventory (RGI)? If authors have used their own glacier outline this might significantly deviate from its RGI counterpart, which could introduce errors and the authors should have computed the calibration steps again in OGGM.
- In the introduction (L69) authors are using their own DEM to predict the ice thickness from all models. In that case, the authors should have re-processed the GIS task of OGGM. A detail on how they use OGGM is missing (see below).
- All data inputs and as well as the model's thickness inversion (glacier initial state) represent different timespans. Why not use the same DEM and glacier outline across all models? You can input your own DEM and glacier outline into OGGM and recompute all steps until the inversion. See the following tutorials.

- Using your own outline in OGGM
- Create local topography maps from different DEM sources with OGGM
- Step-by-Step guide to building preprocessed directories from scratch

Thank you for your comments. Regarding the GPR survey by helicopter, we considered the local morphology and the helicopter altitude above ground, but we did not notice interference caused by surrounding mountain slopes. We think that this is because the glacier sits on the top and only a few slopes are higher than it, as you can see by this photo we took recently.



For the other comments, we used our own outline and DEM, provided by a recent geomatic survey (in 2021).

We carefully read the OGGM tutorials, actually we got a little involved and read every tutorial in the documentation and we found a lot of teaching resources too.

We hope to have understood everything well, therefore we gently ask you to check the OGGM notebook that we included in the supplementary materials, to understand if we did the workflow correctly. We used our own DEM and outline, the same as other models, and a calibration from geodetic mass balance. Thank you!

Algorithms:

- Choice of input parameters in ice thickness inversion models: The authors should clarify that these parameters are not transferable between glaciers (see Zekollari et al. 2022). Additionally, a sensitivity study on the model parameters should have been conducted to assess the impact of parameter variation on ice thickness profiles computed by the models.

  Thank you. We clarified it in the introductory part of the ice thickness models, in the methods sections, and performed a sensitivity analysis on two main glaciological parameters on the final GlaTE model, both constrained and unconstrained.

- How did the authors calibrate surface mass balance and ice thickness inversion in OGGM? or Did they use pre-processed directories? A specific workflow of the steps followed with OGGM is missing. The actual code repository of this study is not shared, thus is not possible to verify.
  Thank you for pointing it out. We performed the informed three step calibration for mass balance parameters, with the tasks.mb_calibration_from_geodetic_mb function.
  For inversion calibration, we used the calibrate_inversion_from_consensus task. You can find the jupyter notebook in the supplementary materials, in the OGGM folder.

Results:

- It would be interesting to see a comparison of ice thickness differences between GlaTE and the other two models, along with a more in-depth discussion of the reasons behind these differences.

  We performed a average and standard deviation comparison on the three unconstrained model, and added it to the results.

- The findings of the paper would be strengthened by comparing the resulting ice thickness map from GlaTE to existing ice thickness inversion and volume products (e.g., Millan et al. 2022, Farinotti et al. 2019, and Cook et al. 2023, all available in OGGM).

Thank you for this point. We compared the results with the ice thickness inversion products downloaded in OGGM. With Farinotti the comparison was possible, even if they used the 2003 RGI outline, whereas Millan product had a much different outline, probably they considered all the nearby little glaciers as part of the Rutor glacier. We clipped the Millan raster using the 2003 RGI outline as mask layer, in QGIS. We could not download Cook dataset from OGGM as it gave us this error that we do not understand completely:

```
from oggm.shop import cook23
cook23.cook23_to_gdir(gdir, vars=['thk', 'divflux'])
```

```
---------------------------------------------------------------------------
AttributeError                            Traceback (most recent call last)
Cell In[18], line 2
      1 from oggm.shop import cook23
----> 2 cook23.cook23_to_gdir(gdir, vars=['thk', 'divflux'])

File ~/miniforge3/envs/oggm_env/lib/python3.11/site-packages/oggm/utils/_workflow.py:482, in entity_tas
k.__call__.<locals>._entity_task(gdir, reset, print_log, return_value, continue_on_error, add_to_log_file
, **kwargs)
    479     task_name += fsuffix
    481 # Do we need to run this task?
--> 482 s = gdir.get_task_status(task_name)
    483 if not reset and s and ('SUCCESS' in s):
    484     return

AttributeError: 'list' object has no attribute 'get_task_status'
```

<u>Discussion:</u>

- I would encourage the authors to provide a stronger justification for how this methodology could be scaled to other glaciers and applied to existing GPR surveys in temperate glaciers. Additionally, it would be helpful to explain how this study could address the under-sampling problem of ice thickness in temperate regions (e.g., the Andes). However, caution is needed, as once models are used to improve observations, they are no longer pure observations and here there is a "human error" element also in place with this method. The authors could emphasize that GPR measurements provide a better representation than models, especially in areas like valley walls where models may struggle due to the simplification of glacier geometry (e.g., elevation flowlines and bed geometry assumptions in the case of OGGM).

Thank you. We agree with the need to improve the discussion about the leak of models into the observations and the human error element. We discussed it in a dedicated paragraph in the discussion section.

- There is little mention about debris cover which is likely not accounted for in the ice thickness inversion algorithms.

We did not consider debris cover, which is not crucial in our glacier (see the *photo above, taken at the end of the ablation season). In the GlaTE model there is the possibility to account for it with a simple parameter, but we are confident that this can be skipped for this glacier.*

Conclusions:

- While the study is well-detailed and clearly explained, it could benefit from a stronger emphasis on its contribution to the broader challenge of ice thickness observations in temperate glaciers. The results, though valuable for this specific glacier, do not provide new insights beyond the updated GPR surveys and improved ice thickness map. To enhance the overall impact, the authors could explore a more quantitative interpretation of the results and better highlight how their findings address larger-scale issues in future research.

Thank you. A more quantitative and reproducible approach was the focus of our revision. For the larger scale issue, refer to our previous discussion. Since the methodology intrinsically requires the human judgment, it is difficult to automatize the whole process and create a regional model out of it, directly. On the other hand, the proposed workflow can be seen as useful to constrain larger-scale models. From what we understood, larger-scale models currently constrain their parameters on mass-balance data or on regional volume estimations, not on detailed thickness maps of some glaciers. Constraining larger models with thickness maps based on GPR observations helped by an ensemble of models is a more complex task, but could render those regional models more accurate, especially near the glacier outlines, where we saw that the models are not always consistent with the GPR data. Or, from the opposite point of view, the regional models could be used not to estimate ice thickness, but to estimate the regional variability of glaciological parameters based on GPR+models observations of ice thickness. It would be a quite big effort, we suppose...

Minor corrections and suggestions:

Thank you for all your very detailed corrections and suggestions. We replied below to every point.

Abstract

L1: Add an example of where temperate glaciers are located (i.e. not at the poles).

Thank you. We added the example of the European Alps.

L10: Authors should explicitly state that they used model output to manually select the best bedrock profile from the GPR data in problematic survey sections, clarifying that the model output is used to fill gaps in the GPR observations along those profiles. This should be stated early on in the abstract and the introduction, to improve the objective of the manuscript.

Thank you. We stated that the model output is used to fill gaps in the problematic GPR profiles, both in the early abstract and in the introduction.

Introduction

L29-30; replace “inner composition.” with “present day ice thickness distribution and geometry”

Thank you, done.

L42; EM to Electro Magnetic.

Done, thank you

L43; Suggestion: Rutishauser et al. (2016) analyzed a large set of GPR data acquire on Swiss glaciers and found that depending on the specific glacier, the bedrock interface could only be successfully detected in 12-69% of the GPR data due to this scattering issue.

Thank you, done.

L48 Remove “Also”. Suggestion: “Air bubbles trapped in ice cause additional scattering of the GPR signal, which helps differentiate between various types of ice...”.

Thank you, done.

L51: replace “are reported in the Study site paragraph” by “are summarised in section X”

Thank you, done.

L56: replace “paper” by “study”.

Thank you, done.

L58: point to a figure to direct the reader to a GPR profile to indicate the issue.

Thank you, we pointed to Figure 3.

L62: replace “help the analysist” with “aid”.

Thank you, done.

L64: I will just call it glacier models or ice dynamical models (they all are ice thickness inversion algorithms of some sort). Authors should pick a single definition throughout.

Thank you. We modified it here and throughout the text.

L67: Please cite the version of OGGM used in the study. See https://docs.oggm.org/en/latest/citing- oggm.html#software-doi

Thank you. We run again the OGGM model according to your suggestions, so we cited the version used now: 1.6.2

L77: statements like “it should be” introduces doubt on the results, try to avoid this type of language and quantify how much the ice thickness product improved via statistics.

Thank you. We simplified and rephrased this paragraph.

Study site

This section is too long and I don't see how past geomorphological events are relevant to this particular study. I would start by describing the site (from L97) and georeferenced so the readers know where the glacier is geographically.

Thank you. We removed the part describing past geomorphological events. We left a reference to a publication of ours, in which we reviewed past studies on the area in detail:

Vergnano, A., Oggeri, C., and Godio, A.: Geophysical–geotechnical methodology for assessing the spatial distribution of glacio-lacustrine sediments: The case history of Lake Seracchi, Earth Surface Processes and Landforms, p. esp.5555, https://doi.org/10.1002/esp.5555, 2023.

L102: Add citation of DEM's used to compute ice thickness losses.

Thank you. We added it and where to download them.

Methods

L111-123: Remove all to".

Thank you, removed.

L113: Remove "and updated".

Thank you, removed

L119: Replace "to perform the manual picking" with "Manually select reflexion events"
Replaced, thank you.

L123: Replace "to draw a final result…" with Produce a map of the glacier ice thickness (Figure 6)

Thank you, replaced.

L124-127: Suggestion:

"Some topographical adjustments were necessary to assist in analyzing GPR observations that span different time periods (2012 and 2022). A 2021 DEM of the glacier surface was used for the GlaTE and GlabTop2 algorithms, while the 2000 DEM was used for the OGGM algorithm. In other words, the GlaTE and GlabTop2 models represent the 2021 situation, OGGM represents 2000, and the GPR data corresponds to 2012."

Thank you. We rephrased as suggested except for the part of OGGM, that now uses the same input DEM as other models.

What about the glacier outline date?

Thank you. The outline date is also 2021, from an ortophoto produced during the same survey of the 2021 DEM. I explained it and put a reference.

L131: Why do the authors not use the same DEM (or the best DEM) for all models? See above.

Since we now use the same DEM, we deleted this phrase.

L124-145: This text seems a bit misplaced, I would divide the text into sections for (i) pre-processing of input data for models and (ii) post processing of model output and the describe (ii) after describing the algorithms.

Thank you, you are right. We reorganized the methods section in order to be smoother to follow and consistent with the supplementary materials. We put this part in a specific section in which we described how we compared models and GPR, in 2D and 3D.

Sect 3.1 Explain if the GPR data collected from a helicopter needed to be corrected for altitude changes in the survey and the scattering effects caused by the nearby terrain. See Church, G. et al. (2018).

Thank you, we explained that we did not noticed any interference from nearby slopes or features, and cited the suggested article.

L154: Add - The GPR data was processed by the following method:

Thank you, added.

L170-175: These lines contain irrelevant text. The increase in usage and citations of a tool or model (e.g., OGGM) does not necessarily indicate it is the best tool for a particular study. A more thorough justification for the choice of tools should be provided here. The OGGM documentation clearly states that it is designed for large-scale or regional glacier modelling. Caution is advised when using OGGM for single glacier studies, and a detailed workflow for producing the thickness inversion should be included in such cases.

Thank you for pointing it out. We think you are right and removed this little part about model usage in the scientific community, and replaced it with some comments on the use targets of each glacier. As per previous comments, we included a detailed workflow for OGGM in the supplementary. Even after reading all the tutorials, we are not completely sure to have done everything correctly on it, but if you have time and could check we would be grateful. Consider, anyway, that OGGM as other unconstrained models are used in our workflow only to provide a general glacier shape, to help the GPR pickings (and to assess the variability among models, which was added to this study following reviewers suggestions), not to provide the best possible solution.

L177: Remove “ice flux mechanics” and replace: ice flow theory and mass conservation.

Thank you, done.

L179: Replace all “picking” with manual selection.

Thank you. We searched throughout the text for all "pick..."words and replaced with manual selection of ice-bedrock interface or similar words, according to the context.

Suggestion replace L181 – L184 with
“The thickness inversion models required specific input parameters to run. These were reviewed for consistency with the physical characteristics of the study area, but unless stated otherwise, default values from similar alpine glacier studies by the algorithm developers were used. See below a summary of all models”.

Thank you, replaced.

Here, authors should state that these parameters are not transferable glacier to glacier and a sensitivity study **on one profile** at least should have been carried out on model parameters to see the impact of parameter uncertainty in ice thickness distribution.

Thank you for your suggestion. After the sensitivity analysis on both unconstrained and constrained GlaTE model, we extracted the ice thickness average and standard deviation of the sesntivitiy analysis in correspondence with the profile 7. We chose profile 7 because it was not one of the best GPR profiles, and the picking was not particularly easy on it. We can see that constraining the model with GPR data helps reducing greatly the variability of the model outcomes, whatever the input glaciological parameters. ("whatever"... considering that we made them vary in a reasonable range). We think that seeing the sensitivity analysis on one profile gives insightful view of the question.

Sect. 3.2
This section would benefit from a table comparing the parameters (and their values) used in all models, providing a quick overview of each model's setup, along with a column citing the publications from which these parameters were sourced.

Thank you for the suggestion. We are a little uncertain about this issue, because the number and kind of parameters is quite different between the three models and we do not think that a table really show a meaningful comparison between the models. We put each set of parameters in the supplementary material, respectively: for OGGM, in a "Show Calibrated Parameters" cell of the Jupiter .ipnyb notebook; for GlabTop2, in a config.cfg file, and for GlaTE, in a .txt file in the matlab folder. We updated the methods section to refer to these files. Can be enough according to you?

Consider also that, apart from the sensitivity analysis, for the three unconstrained models we stuck with the parameters suggested by the developers on alpine glaciers (or they were calibrated during the process), since the primary role of the three unconstrained models is to preliminary give a general guidance on the GPR picking, not to select the best parameters.

For these reasons, we do not think this table deserves to stay in the main text. If you have comments on this or think differently, we are open to change our mind.

L225: Cite OGGM version used.

Done, thank you! We now used the newer version of OGGM compared to the time of first submission.

L227-228: Replace “OGGM bed topography inversion algorithm” with “OGGM ice thickness inversion algorithm” … “which is based in ice flow dynamics and mass conservation (Farinotti, et al. 2009 and Maussion et. al 2019). The ice flux is computed as: … “

Thank you, corrected.

L234: “under the simple assumption of equilibrium”. This is not correct in the case that the latest version 1.6.0 of OGGM was used. Please, note the version used in this study and how the ice thickness inversion was calculated. Do authors calibrate surface mass balance and ice dynamical parameters? Note that it is possible to calibrate OGGM to match geodetic mass balance data which removes the equilibrium assumption. In the latest version is possible to calibrate the glacier mass balance and ice dynamics parameters at the same time using a “dynamic spin-up” see Appendix A and Aguayo et al. (2023) for details and the following tutorials:

https://tutorials.oggm.org/master/notebooks/tutorials/observed_thickness_with_dynamic_spinup. html#dynamic-model-initialization-using-observed-thickness-data

Thank you for pointing it out. At the time of first submission, we used an older version of OGGM. However, now we updated our workflow to include the last version of OGGM (1.6.2 at the time of writing), and the calibration to match geodetic mass balance. We hope to have done everything correctly. We corrected the text accordingly.

Discussion

L276: “ice thickness”? do you mean ice volume (why is this not just stated in Km^3)

thank you, corrected. we prefer the millions m3 because in km3 they became less than 1 and they are poorly readable.

Section 5.1 Here ideally authors should have done a better analysis on the difference between the thickness maps computed by the different models and also show a flowline profile view. Also compare the resultant volume with previous studies and estimates (see references).

Thank you. We computed the bedrock estimations from the three different models, showing an average and standard deviation. We added a flowline profile view. We also showed a comparison with the volume with previous studies and estimates.

A lot of this section could be removed if the authors use the same DEM and there is no need to correct ice thickness changes over time.

We now used the same DEM for every model and we removed the section, and we removed the unnecessary text accordingly. The ice thickness had to be corrected anyway for time change during the workflow, you can see it in the supplementary material QGIS folder in those rasters called "corrected to 2012" or similar ... because GPR profiles were taken in 2012, and DEM is from 2021.

L302-307: Suggestion

"This joint interpretation prevented the mistake of interpreting the first non-reflective layer (white in the GPR sections) as ice and the first reflective zone (scattered black) as bedrock. The deepening reflection on the right side of Figure 3 clearly shows that the ice-bedrock interface is not related to the scattered reflective zone observed at 20-40 m depth. Manually picking the ice-bedrock interface, guided by the estimates from the algorithms, was particularly helpful, especially below 50 m where the GPR signal was too attenuated."

Thank you, we changed the text as suggested.

L309 "This is not far from estimates without GPR data" Quantify such differences.

Done, thank you.

L329. "previous research" add citations.

Thank you, we forgot to put the references. We had in mind the comparison work (ITMIX project) of Farinotti et al., 2017.

Figures
Figure 1.

This figure needs a map of the alps with the location of Rutor glacier. Add RGIID or GLIMS ID.

Thanks, done.

Replace "how many meters it has subsided in the past decade (from 2008 to 2021)" with changes in
ice thickness (m) from 2008 to 2021.
Thank you, done.

Add RGI outline as well as the outlines used in this study with different colours. Add citations.

Thanks, done.

Figure 2.
Dotted survey lines could be thicker.

Thanks, done.

Figure 3.
I would add a point of first guess from authors of where the bedrock might be if they didn't know from the thickness inversion algorithms.

Thank you for this interesting idea! We preferred to present the first guess a little more in depth about this issue, but we worked on the figure with the three GPR plots with models and pickings, instead of working on Figure 3. In the last of the three plots, we divided the pickings in three groups: the points where the picking was quite sure, the points where the picking was made thanks to having seen the models, and the pickings that at first could be considered OK, but after seeing the models they proved wrong (as probably happened in the previous estimation of Rutor glacier volume of Villa 2008). We think that this can help understand better the added value of the proposed methodology.

However, as also suggested by another reviewer, we added an indication of the clutter zone that could be confused by bedrock and the true bedrock.

Figure 4.
Another panel could be added to this figure looking into thickness profiles from models along the main flowline and two more figures showing the ice thickness differences between GlaTe and the other two models.

Thanks. We computed the flowlines with OGGM and, along the main flowline, we extracted thickness profiles from models. We also computed average and standard deviation of models. We added these panels.

Figure 5. (and similar)
Add to the bottom panel the part of the profile that is taken or selected using the ice thickness inversion algorithm (i.e., fill the gap in the profile via another colour)

We did it with different colours, it seems much clearer now. thanks.

Figure 6.
Instead of displaying the GPR data on top of the thickness map, display thickness differences between GlaTE and the GPR. Or plot differences in profiles.

Thank you for the suggestion. We displayed thickness differences between constrained GlaTE and GPR.

Appendix.
Authors should also compare their resultant thickness map with other estimates. See comments above, this could go in the appendix.

We compared the estimations with Farinotti-consensus and Millan thicknesses.

References

Aguayo, R., Maussion, F., Schuster, L., Schaefer, M., Caro, A., Schmitt, P., Mackay, J., Ultee, L., Leon- Muñoz, J., and Aguayo, M.: Assessing the glacier projection uncertainties in the Patagonian Andes (40–56° S) from a catchment perspective, EGUsphere [preprint], https://doi.org/10.5194/egusphere- 2023-2325, 2023.

Church, G. & Bauder, A. & Grab, Melchior & Hellmann, Sebastian & Maurer, H.. (2018). High- resolution helicopter-borne ground penetrating radar survey to determine glacier base topography and the outlook of a proglacial lake. 1-4. 10.1109/ICGPR.2018.8441598.

Cook, S. J., Jouvet, G., Millan, R., Rabatel, A., Zekollari, H., & Dussaillant, I. (2023). Committed ice loss in the European Alps until 2050 using a deep-learning-aided 3D ice-flow model with data assimilation. Geophysical Research Letters, 50, e2023GL105029. https://doi.org/10.1029/2023GL105029

Millan, R., Mouginot, J., Rabatel, A. et al. Ice velocity and thickness of the world’s glaciers. Nat. Geosci. 15, 124–129 (2022). https://doi.org/10.1038/s41561-021-00885-z

Farinotti, D., Huss, M., Fürst, J.J. et al. A consensus estimate for the ice thickness distribution of all glaciers on Earth. Nat. Geosci. 12, 168–173 (2019). https://doi.org/10.1038/s41561-019-0300-3

Zekollari, H., Huss, M., Farinotti, D.,& Lhermitte, S. (2022). Ice-dynamicalglacier evolution modeling—Areview. Reviews of Geophysics,60, e2021RG000754. https://doi.org/10.1029/2021RG000754.

Thank you. We added the suggested references in the text.

Reviewer 2

Review of “Ground penetrating radar on Rutor temperate glacier supported by ice-thickness modeling algorithms for bedrock detection” by Andrea Vergnano et al. (2024)

The manuscript presents airborne and ground-based GPR data collected in 2012 and 2022 over Rutor Glacier, a temperate glacier, which are known for challenges posed by high signal scattering and absorption. The study’s novel approach combines three models (GlabTop2, GlaTE, OGGM) to help with the identification of the ice-bed interface, improving upon prior estimates that likely underestimated ice thickness due to misinterpreted scattering zones near the surface. The study concludes that incorporating the models improves the GPR interpretation in terms of ice thickness. Finally, a new ice thickness map is generated with the new GPR interpretations constraining the GlaTE model.

I think this study presents a creative approach to improve the interpretation of challenging GPR data over temperate glaciers. Overall, the paper fits the scope of the journal and has potential, but in my opinion, several major issues need to be addressed before publication. These include the need for clearer methodological explanations, particularly concerning the use of DEMs in the models. The introduction should more clearly highlight the true novelty of using models to improve GPR interpretation. Furthermore, a deeper analysis of model-assisted picking, including statistical comparisons between model-guided and unguided picks, is necessary to fully support the claim that the models “provided substantial help in manually picking the ice- bed interface”.

Finally, the manuscript requires substantial English language revision to improve clarity, as many sentences are awkwardly phrased or repetitive. I hope the authors find my comments useful and that they can help to improve the manuscript.

Major Issues

- **Language:** The manuscript would benefit from significant English editing. Many phrases are unclear or awkward, and the text could be more concise. Paragraphs often repeat information unnecessarily. I have made specific suggestions in the line-by-line comments.

Thank you for this comment. We are grateful to you for your corrections about language and those of other reviewers, who also highlighted some unclear phrases.

- **Research focus**: The main purpose of this research as stated in the introduction is to "investigate the Rutor glacier thickness with two new GPR datasets" (L56). However, I believe that the manuscript could better highlight the key goal/innovation – using models to assist in identifying the glacier bed in GPR data. This is underemphasized in the introduction, results, and discussion sections.

Thank you. We emphasized more the key goal and innovation, also taking into account the lack of discussion about regional scaling of this methodology, highlighted by Reviewer 1.

- **Abstract**: I find the abstract quite lengthy, and the primary goal and key findings are not clearly conveyed. I recommend revising the abstract after the manuscript has been edited to ensure the message is concise and focused on the main points.

Thank you! We manually rewrote the whole abstract focusing on our research focus as per your previous comment. We feel it is more focused and clear now. We correspondingly updated also the paper title.

- **Methods:**
    - **Ice thickness change:** It is unclear whether the ice thickness change (Figure 1) from the DEM differencing is original to this study or based on previous work. If new, the method should be explained

This is original of this work. It was a standard procedure of differencing two DEMs pixel by pixel, therefore, we did not emphasize it, but we described now it in more detail.

- **DEM use:** The rationale for using different DEMs for different models is unclear, especially why a 2000 DEM was used for OGM. I am not familiar with the models, but is it not possible to run the OGGM model with the 2021 surface topography? Additionally, why was a 2008 and 2021 DEM used for the GlaTE and GlabTop2?

Thank you, this is correct, it is possible to run the OGGM with the 2021 surface topograph: we did so, completely updating this part of our workflow in the revision process. Also Reviewer 1 is of the same opinion. In this sense, we eliminated the part in which we talk about the 2000 DEM, not used anymore.

The 2008 and 2021 DEMs were used because some of the GPR data were collected in 2012 and some in 2022, and we want to provide a picture of the ice thickness in 2021. Therefore, the 2012 GPR data have to be corrected for the loss of ice between 2012 and 2021, which was of dozens of meters in the lowest areas of the glacier. To do so, we calculated an average annual ice loss, pixel by pixel, by subtracting the 2021 and the 2008 DEMs. We tried to clarify this part in the methods section, which now was also reorganized to follow the structure of the workflow and the supplementary materials more closely. It is also described in the supplementary materials, QGIS folder.

- **Ice thickness vs bedrock topography:** I understand that the models output ice thicknesses, but why not compute a bedrock DEM instead? The bedrock topography is not expected to change over the study period, and could directly be compared to the GPR data from any survey time (i.e. 2012 and 2022). Ice thicknesses can still be extracted (subtracting the bed DEM from the surface DEM). This could reduce all the ice thickness corrections that currently need to be applied.

Thank you. Since the models run on different programming languages, it was easier for us to let them compute ice thicknesses and then convert them later to bedrock topography in QGIS. We did so and also added a figure of the bedrock topography vs the surface topography.
For the GPR data, the problem is that if we want to think in bedrock topography instead of thinking in ice thickness, we have to apply the correct surface topography to GPR data, which was not available in the raw GPR data. So, it is not more time-consuming than correcting the ice thickness for the ice lost between 2012 and 2021, in our workflow.
So, we plotted also the bedrock topography as suggested, but we retained the use of the ice thickness for the main workflow, it was just more convenient for us.
However, as we told in the preliminary review report, this comment prompted us for a nice idea for future works: to test these models on a glacier in which we have the DEM in several years to check the robustness in the detection of the bedrock. It would be interesting to build a model that takes into account multiple surface DEMs collected in multiple years, and is forced to calculate always the same bedrock. It could be done on this Rutor glacier using the

2000, 2008 and 2021 DEMs, but this would require an important modification of the code. We believe it is out of the scope of this specific paper, as it may take time to assess it in a rigorous way and change the model code.

We reported these points in the discussion section, prompting future research.

- **Results:**
    - **Ice loss map:** As the ice loss map supports the hypothesis of underestimated thickness, it should be included in the results.

I understand your point of view, We put it before because it seemed to me a good reading flow. We reorganized the paper according to your comment. e

- **Statistical analysis:** A more in-depth quantitative analysis is needed to assess how much the models aid in picking the ice-bed interface. This could include comparisons of ice thickness picks with and without the models, as well as how each individual model was used (e.g. for future recommendation, is there one model that stands out, instead of having to run all three?) The discussion includes some statistics (e.g. "20% of the GPR lines clearly identified the bedrock"), but it is unclear how these were calculated, and they are not included in the results.

Thank you for these interesting ideas on how to present the results more quantitatively. We showed average and standard deviation of the three models, as well as performing a simple sensitivity analysis on two main glaciological parameters (creep rate and basal sliding). We also improved the visualization of the pickings, dividing them in "sure" pickings, pickings helped by models, and pickings that may seem true, but they are probably wrong.
As a future recommendation, since we were able to better see the variability between the models and between the same model with different parameters, we would say that all the three models reconstruct a similar bedrock shape, but the ice thickness may vary of tens of meters. We would say that running a little bunch of models helps in recognizing the model uncertainty, which is helpful in guiding reasonably the pickings avoiding to trust too much one single model.

- **Radargram interpretation:** The manuscript could more strongly emphasize how weak reflectors, identified with low confidence, are validated through model agreement, increasing confidence in identifying the ice-bed interface.

  That being said, I also think there are instances where the selected reflectors appear questionable, which may raise concerns about potential bias in the manual picking process when influenced by model outputs (e.g. picking noise). For example, I have difficulties identifying a reflector that was picked on

  - Profile 2012-7 between ~200-500 m
  - Profile 2012-8 between ~1500-1900 m
  - Profile 2012-9 between 500-1000 m

- Profile 2012-10 between 300-1500 m

This risk should be discussed explicitly (in the discussion section), as it is important to acknowledge the possibility of seeing patterns in noise when guided by models.

You are perfectly right, we discussed the subjectivity of the methodology in a stronger way. To help this, we colored with three colors the pickings, dividing them in "sure" pickings, pickings helped by models, and pickings that may seem true, but they are probably wrong.

- **Figures**

- A study area overview map to see where in the Alps Rutor glacier is would be useful (e.g. integrated in Figure 1 or 2)
  Thank you, added.
- Consider increasing the font size in all figures and remove color scale name information in the figure caption.
  we checked it, thanks.
- *Figure 1*: Add elevation contour lines (or on Figure 2) and a reference to the source of the glacier outline. Also consider labelling the glacier tongues as described in the text.

Thank you, done.

- *Figure 2*: Increase line width, and consider using markers instead of “start” and “end” labels to reduce text and improve readability.

Thank you, done.

- *Figure 3*: I suggest adding arrows to indicate the “clutter zone” and “true bedrock” so the reader can follow what is meant in the text (L256-259). Also, consider removing Figure 3 as it is repeated in Figure 5, or replace it with another example (e.g. Profile 2012-8).

Thank you for the nice addition, we added it. We would stick to the same profile, since it was one of the clearest in showing the difference between clutter zone and true bedrock.

- *Figure 5*: I suggest using different colors instead of line-styles to better distinguish the models.

Thank you for the suggestion, we did it.

- *Figure 6*: I suggest using the same colormap for the GPR and model ice thickness for easier comparison. The GPR data can be surrounded with a white outline for contrast.
  Thank you. Another reviewer suggested to plot the difference between GlaTE constrained by GPR and GPR, and we did so. We think it improved the figure more than just using the same colormap.

- Appendix Figures: I think that some of the description should be moved into the main results/discussion sections.

- *Thank you, we improved the discussion section taking some information from the appendix.*

Minor Issues/Line-by-line comments

L3: I suggest removing the sentence with cold ice, it is rrelevant here.

Thank you, removed.

L8-9: I suggest removing the sentence “Besides, GPR….”

Thank you, removed.

L31, L36, etc.: Consider replacing “meltwater” with “englacial water content” or “water”, to avoid confusion with surface meltwater generation/runoff, englacial water may also result from rain.

Thank you for the clarification. We replaced "meltwater" with "englacial water" throughout the text, except when the context specifically referred to water produced by a melting process.

L32: I believe it is “pressure-melting point”, not “temperature-pressure melting point”.

Thank you. We said "temperature-pressure melting point" to somewhat refer to the fact that the pressure due to the ice above can change the temperature at which the ice below melts, but we probably used an incorrect and misleading term. Anyway, we searched through some books and many just use "melting point". We prefer this simpler term, because Cuffey and Paterson 2010, "The Physics of Glaciers" explain that "the often-used term pressure melting point misleads; because the ice contains impurities it does not have a distinct melting point, determined solely by pressure. Temperate ice is a complex material...." We corrected it throughout the text.

L35-36: This also reads a bit awkward, e.g. we wouldn’t expect a sudden change in geothermal heat flux. I suggest rewording to “Temperate glaciers at the pressure melting point are primed for rapid meltwater production upon small energy or heat inputs…”

Thank you, we reworded the phrase as you suggested.

L37: Specify that while high-quality GPR surveys are possible (e.g. for snow/firn near surface studies),challenges lie in detecting the bed returns. Reword to “…can challenge the interpretation of bedrock returns from Ground Penetrating Radar (GPR) surveys.”

Thank you, we corrected as you suggested.

L40: Clarify “smaller-scale heterogeneities”, e.g. small fractures or sediment grains, smaller than the wavelength (or quarter wavelengths/range resolution)?

Thank you, we corrected it, choosing the quarter wavelength as a thumb rule of resolution.

L42-44: Reword to clarify what was studied, e.g. “Challenges in detecting basal returns over temperate glaciers have been studied …” Additionally, I think it would be good to mention the studies on effects of antenna orientation on detection of the bedrock reflection e.g. (Langhammer et al., 2019).

Thank you. We clarified this and added the reference.

L47-48: Rephrase to clarify that englacial debris may also originate from surface material, not just freeze-on at the bed.

Thank you, we rephrased it.

L49-52: I think this sentence could benefit from directly referencing some of these studies. **Also consider integrating the study site description here**.

Thank you, we integrated the study site description and we rearranged the text.

L52-55: Replace “resolution” with “spatial resolution”. I suggest reformulating to “The spatial coverage of GPR surveys is limited by survey speed, time and access (e.g. crevasses), leading to discrete, limited sampling of the glacier bed. It is therefore possible that the maximum ice thickness remains unknown due to limited survey coverage.”

Thank you very much. We clarified according to your suggestion.

L57-L59: Re-word for clarity, e.g. “These new datasets reveal high scattering of the radar signal over most parts of the glacier, demonstrating the difficulty in detecting the ice-bedrock interface.”

Thank you. We reworded as you suggested.

L60: Include a reference for the “previous doubtful estimates of ice thickness”.

Yes, we forgot to put it also here. We added it. thanks.

L68: I believe the correct reference is (Langhammer et al, 2019a), verify other instances.

Thank you, you are right. We corrected this and another instance.

L68-70: “Thanks to … are extracted.” I suggest rewording to “The ice thickness is predicted using the three models.” (i.e. the DEM part belongs in the methods section).

Thank you. We reworded and removed the part unnecessary here.

L71-72: Rephrase to “… superimposed on the radargram to help identify the most likely ice- bedrock interface…”

Done, thank you.

L74: Replace “inner geometry of the glacier“ with “bedrock topography”

Replaced, thank you.

L93-96: Instead of just mentioning multidisciplinary aspects/different perspectives, provide examples (e.g. glaciology, geomorphology, ecology, hydrology …?).

According to Reviewer 1, this part talking about past geomorphological events and history of the site is not relevant for this particular study. So, we decided to remove it. We left a reference to a publication of ours, in which we reviewed past studies on the area in detail:

Vergnano, A., Oggeri, C., and Godio, A.: Geophysical–geotechnical methodology for assessing the spatial distribution of glacio-lacustrine sediments: The case history of Lake Seracchi, Earth Surface Processes and Landforms, p. esp.5555, https://doi.org/10.1002/esp.5555, 2023.

L96: misspelling of “multidisciplinary”

Same as above, this part was removed. Thank you anyway.

L97: reword to “… the Rutor glacier covers an area of 7.5 km^2 …”

Done, thank you.

L100 and others: Replace “outline” with “margin”

thank you. We replaced outline with margin throughout the text.

L101-108: Moving the ice thickness change discussion to the methods/results sections, or reference to original source if from another study.

Since the ice thickness change - and its lack of agreement with the previous volume estimate - was a starting point and motivation of the study, we prefer to keep parts of the discussion in the introduction. However, we put the figure in the results section.

L107: replace “extension” with “area”.

Done, thank you.

L108-109: Move this sentence to the introduction for better context.

we moved the whole section to the introduction.

L116: Reword to “The results of this step are show in Figure 4.”, or remove this sentence.

Done, thank you.

L119: Replace “reflection events” to “reflectors”

Thank you. We replaced it here and in other parts of the text where related to the ice-bedrock interface. We left "reflection events" when related to other sources of reflection, such as scattering from englacial water pockets or other anomalies.

L120: Replace “limit…” with “reduce the chance of mis-interpreted bedrock reflections”
Thank you, done.

L121: Be more specific: “… surface topography and the GPR-derived bedrock topography.”
Done, thank you.

L123: Step 6 does not contribute to the “overcome the difficulties in interpreting the GPR data…” as stated at the beginning of the methods section. I suggest removing this step.
Removed, thank you.

L124-145: Address comments above and consider moving this section to 3.2. Clarify the glacier outline source (e.g. mentioned in L186)?

we put this section after the GPR section. we clarified the outline source.

L150-151: This is repeated in Step 4, I suggest removing it here.

Removed, thank you.

L154: What was the bandpass filter of for the ground-based survey? I assume it was lower than this.

Yes, for the ground-based survey the bandpass filter was from 0 to 120 MHz. We forgot to mention it. Thank you.

L156: Clarify “correct max phase”, e.g. is it a dewowing process? Also, avoid non-scientific language like “suggested by Reflexw”.

We clarified it. It is just an automatic recognition of the signal direct wave, in order to fix the correct time-zero. In general, it works well after applying the dewow filter.

L178: Replace “drive…” with “help identify the ice-bedrock interface during manual picking…”

Done, thank you!

L181-182: Reword to: “The modeling algorithms required additional input parameters (e.g.*xxx*). These were checked for consistency with the Rutor glacier study area, …”

Done, thank you.

L202: Remove double citation.
Done, thank you.

L207: Clarify that known ice thickness/bedrock points, not GPR data itself, are used as input. Similarly, further down, I assume h_{GPR} is the GPR-derived ice thickness, not the GPR data.
you are right. we corrected it.

L212: I suggest removing “outside”

done, thank you

L213: Clarify “gradient of outside terrain slope”, i.e. is it the slope outside the glacier?

Thank you, yes.

L226: I believe this should be “meltwater runoff”

corrected, thank you.

L228: remove the “is” before"equation”
corrected, thank you.

L233: precipitations (remove s)
removed, thank you.

L232-235: If a mass balance was used to estimate q, include the details on how this was determined for Rutor glacier and the value used.

for geodetic mass balance, the study of hugonnet et al. (2021) was used by the OGGM code, while the climate dataset is the W5E5. We added the citations. thanks!

L245-253: Instead of listing the figures at the start of the results section, I suggest integrating them into the text to improve the flow of the text.

Thank you, we did as suggested.

L257: Replace “black reflection zone” with “strong backscatter zone” or “high amplitude zone”.

replaced, thank you.

L257-258: “However, on the right side of the plot, the clearly submerging ice-bedrock interface shows…” I suggest rewording the interpretation of the submerging ice-bedrcok interface to make it less definitive and more interpretative (e.g. the contrast dipping towards the center on to the left also looks like a bed return, but is not picked as such).

Thank you. We reworded and clarified the phrase.

L260-263: Move the comparison with other studies to the discussion section. Also, the Villa et al. (2008) study used GPR data from 2006, not 2008.

Thank you for noticing the 2006 mistake, and we moved this part into the discussion section.

L270-274: There is a lot of repetition of methods within this section. I suggest focusing on results here only.

Thank you. We rephrased focusing only on the results description.

L277-282: This section is mostly a repetition of the methods part. Move any methods to the methods section and focus the discussion on e.g. how resolution affects the result (e.g. over- deepening being an effect of fine-resolution DEMs?)

Exactly, it produced exaggerated over deepenings if the resolution was too fine. We moved this part to the methods only and focused on the discussion as you suggested.

L288: Explain how the ice thickness near the glacier margin was overestimated, e.g. was it compared to the GPR data?

Yes, compared to the GPR data. We clarified it.

L291: Replace “readability” with “…degree of visibility” or “strength of the ice-bedrock return.”

Thank you, done.

L298: "… more confidence was given…", it is not clear how this was implemented. E.g., do the picks come with a confidence level?

Sorry for the unclear statement. We meant that, when manually interpreting (selecting the ice-bedrock interface on a profile) we looked at the three models. In the parts of the profile where all the three models were consistent, i.e. indicated the same ice-thickness, we were more confident in picking the ice thickness there where all the models indicated. We rephrased it.

L303-305: I suggest including a discussion on the possibility of off-nadir returns (e.g. from valley side walls).

We included it. We do not think that in this glacier this is a critical issue, since the valley side walls are mostly not steep.

L326: What about seismic surveys?

We think that in these very thick glaciers (>100m) one would need very long seismic profiles and very powerful seismic sources (explosives?), making these surveys logistically unconvenient on steep and crevassed glaciers. They are possible and probably effective, though, and we collected some old examples of this in other glaciers nearby, such as the Miage glacier.

L329: I suggest adding this citation here (MacGregor et al., 2021) (relation between frequency and ice thickness)

Thank you. Very helpful publication.

L337: Can we quantify "reasonably comparable models" in the results section, e.g. what is the mean, maximum, standard deviation in the differences in ice thickness predictions?

Thank you. We avoided this term "reasonably comparable" and we provided a more in depth, yet still very simple, analysis of mean and standard deviation of ice thickness predictions.

L339: misspelling of minimizing

Corrected, thank you

L340: It is unclear where these uncertainty estimates come from

They were estimates based on my personal intuition looking at the data. Not exactly scientific... we did not focus enough on this aspect, which is actually very important to understand the relationship between GPR and models. Thank you for pointing it out.

We substituted those estimates with a little discussion on the average/deviation *of the three models (which only coincidentally is exactly our previous estimate of 100 millions m3, imagine how we felt proud to have guessed well) and the sensitivity analysis on the GlaTE model, constrained and unconstrained.*

L343-L360: This section mainly focuses on how the GPR data could be used in the future. However, I think there should be more focus on future applications of this methodology, including whether these models could assist in interpreting GPR data from other glacier surveys.

Thank you for pointing it out. We are very glad to put effort in completely rewriting this part as you suggested. We think that now it conveys the message of the study better.

L376: "… one can choose a lower frequency antenna…", This conclusion is not supported by this study, as the 40 MHz data also did not show improvement regarding ice-bed returns.

You are right concerning no difference was seen between 70 and 40 Mhz antenna. However, what we want to tell was just to avoid high frequencies (for example 400 MHz) on glaciers that are thicker than e.g. 50 m according to the models. The model estimation of thickness can help in this case avoid using the wrong instrumentation. We slightly rephrased it in order to avoid confusion with the study results, let us know if it is clear now. Thanks!

References

Langhammer, L., Rabenstein, L., Schmid, L., Bauder, A., Grab, M., Schaer, P., and Maurer, H.: Glacier bed surveying with helicopter-borne dual-polarization ground-penetrating radar, J. Glaciol., 65, 123–135, https://doi.org/10.1017/jog.2018.99, 2019.

MacGregor, J. A., Studinger, M., Arnold, E., Leuschen, C. J., Rodríguez-Morales, F., and Paden, J. D.: Brief communication: An empirical relation between center frequency and measured thickness for radar sounding of temperate glaciers, The Cryosphere, 15, 2569–2574, https://doi.org/10.5194/tc-15-2569-2021, 2021.

Scanlan, K. M., Rutishauser, A., Young, D. A., and Blankenship, D. D.: Interferometric discrimination of cross-track bed clutter in ice-penetrating radar sounding data, Ann. Glaciol., 61, 68–73, https://doi.org/10.1017/aog.2020.20, 2020.

Reviewer 3

Review of "Ground penetrating radar on Rutor temperate glacier supported by ice-thickness modeling algorithms for bedrock detection"

November 2024

- **General**

The authors demonstrate a model-driven technique for picking points in radargrams corresponding to the glacier bed. They first ran three different models based on surface features, later used to guide the manual picking of Ground-Penetrating Radar (GPR) radargrams of 2012 and 2022. They then estimated the ice thickness in regions without GPR measurements by running the GlaTE model constrained by the GPR measurements.

The manuscript is well-written, and the subject of the work, the difficulty of retrieving the glacier bed, is a hot topic in glaciology, which deserves all the attention of the community. It is one of the most important sources of uncertainty for estimates of the future contribution to sea level rise. Dynamical glacier models are based on reconstructions of bed topography, which are themselves based on *in situ* measurements such as GPR and boreholes. The latter are reliable data, but they are not practical for surveying large areas. In this sense, GPR measurements are the foundation for glaciological studies. For this reason, the manuscript ”Ground penetrating radar on Rutor temperate glacier supported by ice-thickness modeling algorithms for bedrock detection” from Vergnano et al. is very important.

However, it is important not to turn the logic around. Since GPR measurements are an important source of *in situ* information, r**eversing 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**.

- **Major comments**

As mentioned previously, my main concern is related to the leakage of modeling data into measurement data. When inversion modeling is performed, it is crucial to have reliable data to constrain the model and evaluate its quality (see, for example, [Shahateet et al., 2023], where they show the impact of using different thickness maps for ice-discharge calculation and [Shahateet et al., 2024] where they show the importance of reliable thickness measurements). If the measurements are biased toward a specific model, it can highly impact everything that comes after, such as the inversion of the bed and the dynamical models that will use the inversion map.

The methodology is valuable, but the main point is to what extent you can use the picking drove by modeling estimations without data leakage. By analyzing Figure 5 and the appendix, I think you introduced too much bias. Some of the picks are not seen in the radargrams, only through the models.

I think that instead of having the model to then do the picks, the best approach would be to do several different picks and compare them to the models you have. In this case, you have less data leakage and more reliable measurements. In case where you have no reliable pick, leaving it without value is better than filling with model information, since in the future you do the inversion modeling of the ice thickness to cover all the domain. In this way, all the measurements you have are trustful and can be used broadly.

Thank you. We reworked the process of picking taking into account what you suggested. In particular, in all the figures that show pickings, we now show pickings with 3 different colors, to distinguish: 1) pickings that are quite "sure" because the GPR is clear in this point; 2) pickings that were helped by the models and the visualization tools described in the paper (consider that now we added also a 3D visualization to help pick those crossing GPR profiles) - and about those pickings, we were a little stricter than before, to avoid too much leakage of models into GPR as you said; 3) Other possible pickings that proved to be wrong after seeing the models, or let's say after having understood the general shape of the glacier thanks to the models. We considered only pickings of type 1) and 2) to constrain the final GlaTE model. We are perfectly conscious that this subdivision is subjective, and we stressed it more in the discussion. However, in a context where the GPR data are not the best, and the model themselves have uncertainties (see the new standard deviation maps and the sensitivity analysis), allowing a certain degree of subjectivity is necessary if one wants to give a whole picture of a glacier.

In this case, since you use models to support the picking of GPR measurements, you need independent data to validate your method. For this reason, it is desirable to use your method in another glacier with borehole measurements. In this way, you can have an independent validation method. I know it is easy to say and hard to do, but I think it is something important to keep in mind.

We would like to have a borehole on this glacier, but it is hard to do! We discussed it better in the discussion section as a limitation of the study. Unfortunately, the goal of the paper is mainly related to this glacier and the effort of analyzing another glacier during the revision process, with little local knowledge about it, is very hard. We could think of it for further steps of research, where the ground-truth (ice-truth...) calibration is available. Also, sad to say, but with patience, in some years we will have a lot of "boreholes" for free, when the ice melts... in this sense, in our opinion GPR datasets collected in the past will be interesting to analyze in the future, because we will have all the ground-truth data we would ever desire.

In chapter 5.1 (comparison of the three ice-thickness modeling algorithms), you stated that the GlaTE and GlabTop2 had similar results, proving the consistency between the different algorithms. First, you do not provide an overall analysis of their agreement, except by the total volume. To say that, you at least need to show an overall metric to conclude that. Second, it is no surprise that they agree well, since they use the same perfect plasticity method. In my opinion, their agreement is not a proof of the consistency of the method.

Thank you. We had to change a little bit this part, because in the meanwhile a user found a little bug in GlabTop2 code and, running the updated version, the volumes of GlaTE and GlabTop2 are not as similar as before, notwithstanding the "shape" of the glacier is still very similar. But we understand that anyway this kind of comparison just in terms of volumes do not provide an overall analysis of their agreement, as you said. We can now also better compare them to OGGM, since we included the OGGM calculation with the same DEM as GlaTE and GlabTop2A. We included also a new figure with the three models compared along the main glacier flowline, which could also help in this discussion. We mentioned the fact that they both use the perfect plasticity method, thank you for pointing it out.

L287-288 is a warning that something may not be right. Why is the thickness overestimated near the outline of the glaciers? This is the region where you have reliable information from the GPRs, which shows that the measurements do not agree well with the models. Furthermore, L316 stated that 20% of the GPR data was used. Does it mean that 80% of the other points were taken from the models? In this case, it is no surprise that the total mass calculation of your method agrees well with the other models.

Thank you. We generally accepted that near the outline it is more difficult for the models to retrieve the correct thickness, as it is expected for them to provide a general shape of the glacier and not be very accurate in distinguishing if ice is thick 5 or 10 meters at a certain point near the outline. We suppose that this is also due to the fact that the Rutor glacier has not a shape like a very long tongue, which would be modelled better by the algorithms. However, about the percentages, we think that now the picking accuracy is better discussed, after the rework of how we show pickings, divided in 3 types, as per a previous comment. Also, we now were stricter in selecting the pickings as reliable (and we also made use on the new 3D visualization to pick better), and we updated GlabTop2 and OGGM models. With all these additions, we provide a new final estimate of GlaTE model which is, effectively, slightly different than the unconstrained models: about 450 million m3, compared to 510 millions and above of the unconstrained models. We think that this rework results agrees well with your comment: probably, in the first paper submission we leaked too much models into the GPR pickings.

- **Minor comments**

- The description of the homogenization of the different data sources is confusing and hard to follow with so many different years. Consider clarification and reduction of information.

We modified this part after usign the same 2021 DEM for all models. It is surely less complex now.

- Why do you use the GlaTE as your final model? You never gave a complete reason for that. See my comment on L170-L175.

  We used GlaTE as our final model because it was specifically designed to constrain the model estimation with ground-proof data, which is what we are doing in our workflow.

- In the Methods chapter, the figures are not presented in order. Furthermore, Figure 2 is not mentioned in the text except in the first enumeration of the Methods chapter. In the text, you mentioned Figure 1, and the next Figure to be mentioned is Figure 5. In general, I think it is important to improve the way you make references to the figures.

  Thank you, we checked all figures in this sense, after adding some new figures after this revision.

- Where do the other inputs from the OGGM model come from? You did not describe all the inputs.

  Sorry, we were a little superficial in describing the OGGM model. Now, we used the same input data as the other models also in OGGM, so all the presentation of methods should be simpler.

- You don't need the enumeration from L245-253. This information is contained in the legend of the figures. Also, it is better to start talking about the figures before showing them out of the blue.

  Ok, we reworked the beginning of the Results section following your advice.

- You cite a personal communication twice. If you do not have a regular citation for that, rephrase it. For example, changing the word "considered" in L129 to "...showed to be..." avoids the need for a citation of a personal communication.

  Thank you for the suggestion, we rephrased L129 and removed this second citation of the personal communication.

- For the OGGM model, you assumed that the glacier was in equilibrium to infer the ice volume flux (q), which according to the section of the study site is wrong. You can easily use a geodetic mass balance (the one you mentioned) to account for this mass change.

  Thank you. With the updated version of OGGM (1.6.2), and after reworking the OGGM workflow to include our own DEM and outline, we calibrated with geodetic mass balance eliminating the equilibrium *assumption.*

- Several times, you should change to ”Rutor Glacier”, with capital ”G”.

Thank you, we modified it throughout the text.

- **Specific comments**

- L42: The acronym EM is not defined, and it is the first and only time that you use it. So, it is not needed.

  Thank you. We changed it to "electromagnetic".

- L52: Change ”paragraph” to ”section”.

  Done, thank you.

- L96: Change ”multiisciplinary” to ”multidisciplinary”.

  done, thank you.

- L118: I think mentioning the v.sample tool in this overview of the methodology is not necessary and can distract from the main point.

  Sorry, we searched for other places in the text to mention it but this seems anyway the best fit...

- L123: The information in this line is not needed.

  Removed, thank you.

- L151: The sentence ”according to the following steps,” made me get lost. It looks like you are going to explain the steps, but you start to talk about the software. Only in the next page you are actually explaining the steps. Consider passing the sentence to the end of the paragraph: ”The raw data were processed using the commercial. . . open source software (Huber and Hans, 2018), according to the following steps:”.

  We did as you suggested, thank you.

- L170-L175: I think this paragraph is not necessary. It seems to me that you try to give a reason to use them because of their popularity. I would try to address this question with a more objective reasoning.

  Thank you. We removed the paragraph and explained better why we chose GlaTE as the final model.

- L177: ”Ice dynamics” instead of ”ice flux mechanics”.

Corrected, thank you.

- L183: Change ”writers” to ”authors”.

  Thank you, changed to "developers" since we are talking about software.

- L188-L190: How do you avoid the glacier flow line computation? Furthermore, in L190 you say that h_f is the mean ice thickness along the central glacier flow line. So do you actually not avoid it?

  Sorry, there was a typo: we wanted to say "to avoid the laborious process of manually drawing branch lines", as reported in Frey et al., 2014.

- L190: You say what is f , but no further explanation is given. What value did you use? It highly impacts the final result, since it accounts for lateral drag. I presume that in an alpine glacier, this value is important to discuss.

  We understand this issue about f. The original paper presenting the GlabTop2 model says that it was fixed at the value of 0.8 according to previous papers of Haeberli et al.. We reported in the text this value, even if maybe it is not sufficient to answer your question. In this revision, we made a simple sensitivity analysis only on the GlaTE model. However, if you think it is important to do so, we can perform a sensitivity analysis also on the main parameters of GlabTop2, including f, which is indeed among the input parameters in the config.cfg file.

- L198-199: You can exclude this line and pass only ”Further details are provided in the appendix of Frey et al. (2014)” and a good reference of the code (see my next comment).

  Thank you. We formatted the reference of the software as in the submission guidelines of the Cryosphere.

-

- L199: several times you wrote the URL link as reference. I think it is not the right way of referencing a webpage.

  We now properly referenced webpages, datasets and software throughout the text. Thank you for pointing it out.

- L202: ”Clarke et al., 2013” should be ”Clarke et al. (2013)” and ”(Clarke et al., 2013)” is duplicated.

  Thank you, corrected.

- L223: Same as L199.

  corrected, thank you

- L228: ”is” should be removed.

  corrected, thank you

- L238: Same as L199.

  corrected, thank you

- L239: The data is not well cited. It is (NASA JPL, 2020). Also, the reference is wrong in L465.

  Thank you for noticing. Since we changed the workflow using the same 2021 DEM for every model, this DEM is not more used and we deleted the whole phrase.

- L240: Better ”(based on DEM differencing)”.

thank you, corrected.

- L280: Same as L239.

  Thank you. We deleted this phrase since we do not use anymore the DEM from NASA.

- 311: How the bias can not be considered significant? You said that the interpretation of GPR measurements below 50 m was difficult and only 20% of the GPR data was clearly identified (presumably in shallow regions, considering the previous statement). It means that in 80% of the time you used ice thickness from the models, or at least driven by it (it is not clear to me when it is driven and when you simply used the same thickness), especially in the regions
  where it accounts more to the total volume (deep regions). For me, this bias is the major concern regarding the methodology used, and need to be addressed in more details

  In this phrase, we were too soft, you are right. What we wanted to say is that in Profile 8, which has the clearest GPR, the GPR and the models fairly agree. However, this is not sufficient to say that the methodology has no significant bias, it was an overstatement.

  During this revision, thanks to the further improvements in the methodology (3D views, updated codes, for example), we recognized that indeed the constrained model is a little different from the unconstrained one, especially because in this quite wide glacier the models did not estimate the thickness very well near the outline, providing an overestimation. This is corrected thanks to the GPR data in the constrained model.

  We rephrased this phrase accordingly (but we heavily updated the whole discussion section).

- L337: ”from” is duplicated in ”. . . from starting from. . . ”

corrected, thank you

- L339: Change ”miimizing” to ”minimizing”.

  corrected, thank you

- L357-360: It is a conclusion.

  thank you for pointing it out. We moved the phrase to the end of the conclusion section.

- L365: 17.5 m ”on average”.

  added, thank you.

- **Figures**
  - Figure 1: In the legend, change ”areas” to ”categories”. Furthermore, remove the parentheses from ”(Crameri, 2021)”.

  Thank you, done

  - Figure 2: The legend is confusing. Why not numbered from 1 to 5? Also, it is better to number at the end also (e.g.: end-1).

    We did it to be consistent with the original dataset numbering. We specified it better in the caption.

  - Figure 4: Same comment as in Figure 1 regarding ”(Crameri, 2021)”.

    Thank you, corrected.

  - Figure 5: The legends of GlaTE and OGGM are indistinguishable. It would be clearer if you used different colors for the different models.

    We used different colours as suggested, it seems improved now.

  - Figure 6: Same as in Figure 4.

    done, thank you

  - All the Appendix Figures: Same as in Figure 5.

    done, thank you

  - Is Figure A2 the same as Figure 5. If so, no need to show it again.

    Yes it was the same, but it was to have all the complete dataset in the appendix... isn't it ok?

References

[Shahateet et al., 2024] Shahateet, K., Fürst, J. J., Navarro, F., Seehaus, T., Farinotti, D., and Braun, M. (2024). A reconstruction of the ice thickness of the antarctic peninsula ice sheet north of 70°s. EGUsphere, 2024:1–29.

[Shahateet et al., 2023] Shahateet, K., Navarro, F., Seehaus, T., Fürst, J. J., and Braun, M. (2023). Estimating ice discharge of the antarctic peninsula using different ice-thickness datasets. Annals of Glaciology.