Review of Pang et al. (2026) 'ITMSL: an improved ice thickness inversion model integrating basal sliding dynamics for High Mountain Asia (v1.0.0)'

Summary

This paper presents a new ice-thickness-inversion model called ITMSL, and explores its performance across a dataset of 16 glaciers in High Mountain Asia compared to two existing thickness datasets inverted using different methods. The novelty of ITMSL is that it includes explicit consideration of glacier sliding, unlike most previous regional- or global-scale inversion methods that rely on the shallow ice approximation (SIA) and therefore neglect sliding entirely. The authors show ITMSL performs better than the reference datasets on the dataset included in this paper, and that it is generally less prone to producing extreme thickness values, as well as performing much better in faster-flowing areas.

I reviewed this paper when it was initially submitted and had substantial problems with the manuscript at that point, to the extent that I felt rejection was the best option as I thought that the amount of work required to revise the manuscript was too extensive to be completed within a reasonable timeframe. The authors have undertaken substantial further work on the paper and it is much-improved, but I still have some reservations, chiefly about how the results are presented and the effective-pressure parameterisation. I am now considerably happier that ITSML is a useful contribution to the field, so I congratulate the authors on having done the work to persuade me of that, but this could be presented much more effectively, still, I feel. Therefore, I recommend major revisions on this occasion.

Page and line numbers refer to those in the clean version of the submitted manuscript.

Response to Reviewer: We sincerely thank the reviewer for taking the time to re-evaluate our manuscript and for acknowledging the substantial improvements made during the revision. We are particularly encouraged that the reviewer now recognizes ITMSL as a useful contribution to the field and appreciates the additional work we have undertaken.

We fully agree with the reviewer that the presentation of the results and the parameterization of effective pressure still require further clarification. In the following, we provide point-by-point responses to the remaining concerns and have implemented corresponding revisions in the manuscript (with line numbers indicated below).

Major Comments

1. Presentation of results: See below, but I think a lot of the results, particularly those based on glacier- mean quantities, don,t present ITSML in the best light. I suggest that the authors consider presenting their results based on a split between slow- and fast-flowing areas, which is done at one point and is much more convincing in showing that ITSML is an improvement over other SIA-based models and doing something novel.

   Response: We thank the reviewer for the suggestion that glacier-mean statistics may not fully capture the advantages of ITMSL, explicit consideration of basal sliding, is expected to make the difference. We agree that velocity-based provides a much more convincing demonstration of the model's novelty and improvement.

   In response, we have added a dedicated Section 5.1 ("Influence of Surface Velocity on Ice-Thickness Model Performance") in the revised manuscript. In this section, we systematically classified surface velocity into intervals of 5 m/yr and calculated the ice-thickness bias, RMSE for the three models (Millan, Gantayat, and ITMSL), as well as the RMSE differences of ITMSL relative to the other two models (Figure 8). Key findings from this new analysis include:

   (1) The ice-thickness bias of all three models increases approximately linearly with surface velocity, with R^2 values of 0.76, 0.78, and 0.87, respectively, confirming that velocity is a

dominant factor governing inversion accuracy;

(2) In the low-velocity regime (≤50 m/yr), ITMSL consistently yields lower RMSE than the Millan and Gantayat models, and the biases relative to GPR measurements remain relatively small for all models;

(3) When surface velocity exceeds 70 m/yr, both bias and RMSE rise to over 100 m for all models. However, ITMSL exhibits a progressively larger advantage over Millan with increasing velocity; relative to Gantayat, the difference shows a two-stage pattern—below 50 m/yr, the difference increases with velocity; around 50 m/yr, ITMSL performs comparably to or slightly worse than Gantayat; and beyond this threshold, the difference resumes its upward trend;

(4) Overall, ITMSL outperforms or matches the existing models across most velocity intervals, with only a narrow transitional zone (~50 m/yr) where it is on par with Gantayat.

The revised analysis is detailed in Section 5.1 (L410-429) and Figure 8 of the main text.

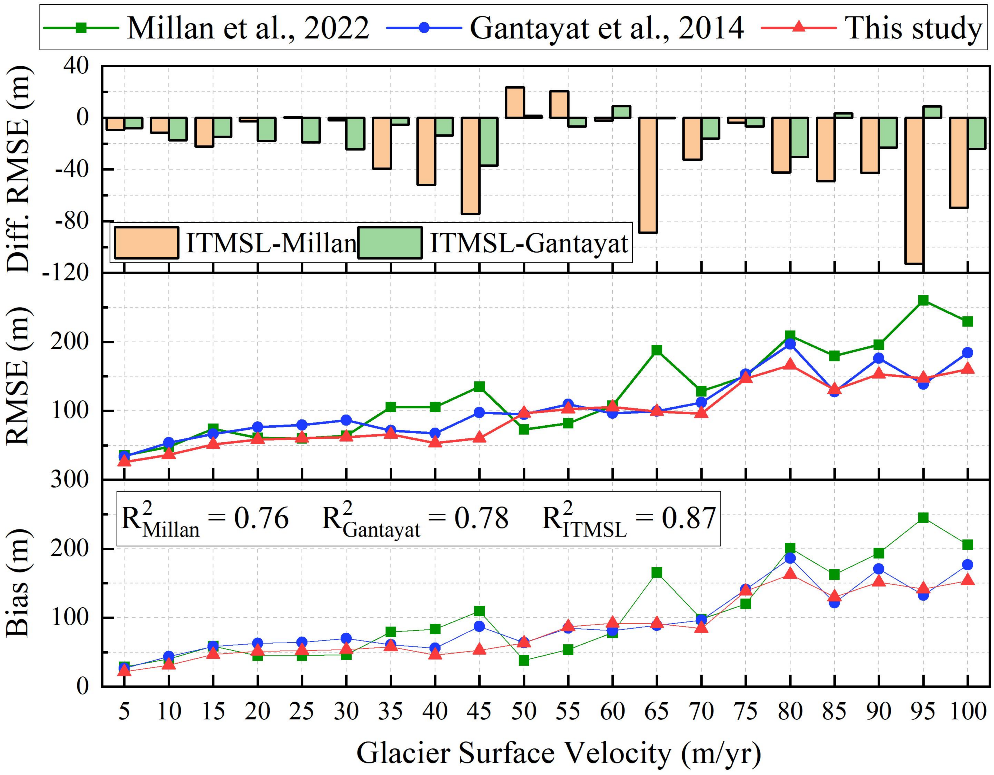

Figure 8: Comparison of ice-thickness inversion bias, root-mean-square error (RMSE), and their differences among ITMSL, Gantayat, and Millan models across surface velocity intervals.

2. Effective pressure: It,s unclear in the manuscript which choice the authors have made for their values of N, but it is either set to be 0 or equal to the ice overburden, neither of which are very good assumptions to make for mountain glaciers. I suspect this will not, in the end, greatly change the results, but I think the authors need, at least, to clarify and justify which choice they've made and subsequently discuss its potential impacts on the results, or make a more sensible choice and re-run the model.

   Response: We thank the reviewer for raising this critical issue. We fully agree that setting the effective pressure to the ice overburden pressure is physically unreasonable for alpine glaciers. In accordance with the reviewer's constructive suggestion, we have re-evaluated and revised our parameterisation of effective pressure. Given the difficulty in precisely quantifying subglacial hydrological conditions at the regional scale, we now approximate the effective pressure as 0.9 times the ice overburden pressure ($N = 0.9\rho g H cos\alpha$). This revision has been explicitly stated and justified in the theoretical section (**Line 211-213**).

We have re-run the ITMSL model with the revised effective-pressure value ($N = 0.9\rho gH cos\alpha$) and updated all inversion results accordingly, including all relevant figures and tables (bias, RMSE, MAE, and all model-comparison plots). ITMSL reduces the RMSE to 39.12 m, which is 32.65% lower than Millan's 58.08 m and 20.62% lower than Gantayat's 49.28 m.

Importantly, the main conclusions remain entirely consistent with those before the revision: ITMSL outperforms or matches the existing models across most velocity intervals. This confirms that while the specific value of effective pressure has a certain impact on the absolute accuracy, it does not fundamentally affect the relative performance ranking among models or our core findings.

Minor Comments

1. p.1, l.21-22: Repeat of the preceding sentence?

   Response: We thank the reviewer for catching this oversight. We have carefully checked the manuscript and removed the duplicated sentence (Line 21).

   "~~However, this method simplifies the parameterization of basal sliding, leading to uncertainties and significant biases in thickness estimates.~~"

2. p.2, l.53-54: Another repeat of the previous sentence?

   Response: We thank the reviewer for pointing out this redundancy as well. We have removed the second sentence, as it is semantically identical to the preceding sentence. The retained first sentence now reads more concisely and sufficiently conveys the logistical challenges.

   "However, due to the harsh high-altitude environments and logistical difficulties, in-situ methods typically require substantial manpower and material resources~~. Many glaciers are remote, making substantial manpower and material resources~~ (Wu et al., 2020; Li et al., 2022; Pang et al., 2023)."

3. p.2, l.54: Bit of a jump here from GPR straight into modelling approaches. Maybe just add a half-sentence introducing the switch from observations to modelling.

   Response: We thank the reviewer for pointing out this transitional issue. We agree that the jump from in-situ GPR measurements directly into the introduction of modelling approaches was somewhat abrupt. To address this, we have added the following bridging sentence in the revised manuscript (Line 53-55s):

   "Therefore, current estimates of glacier ice volume at regional and global scales rely primarily on empirical methods or physical models."

4. p.3, l.63: The way this is written implies the types of model listed here were all considered as part of ITMIX, when the CNN-based approach of Jouvet et al. (2021) postdates ITMIX and is not part of the suite of models assessed there. Just rewrite this a little to make it clear whether this is meant to be discussing the types of model specifically included in ITMIX or ice-thickness models more generally.

   Response: We have revised the relevant paragraph to clearly distinguish between the model types included in ITMIX (i.e., minimization, mass conservation, shear stress-based, and ice velocity-based approaches) and subsequent deep-learning developments such as IGM. The changes have been made in the revised manuscript (Line 65-67).

   "In addition, convolutional neural network (CNN) approaches have been applied to ice-thickness modelling. Currently, an extensively applied CNN-based method is the Instructed Glacier Model (IGM) (Jouvet et al., 2021)."

5. p.3, l.69: 'robust' in what sense?

Response: We agree with the reviewer that our original use of "robust" was ambiguous. To clarify, we have revised the wording to "numerically stable," specifying that we refer to the stability of the numerical solution, rather than making a general statement about the method's performance (**Line 72**).

"...shallow ice approximation (SIA) approaches are numerically stable and computationally fast..."

6. p.4, l.102: Strictly speaking, IGM uses the Blatter-Pattyn approximation (first-order approximation), not full-Stokes

Response: We thank the reviewer for catching this factual error. We agree that IGM is based on the Blatter-Pattyn approximation (first-order approximation), rather than the full-Stokes equations. We have corrected the wording accordingly in the main text, which now reads (**Line 102-105**):

"This approach employs the Blatter-Pattyn approximation (a first-order approximation) to construct input-output mappings across a wide range of glacier geometries and dynamic regimes, thereby establishing itself as an efficient and accurate ice-thickness inversion model."

7. p.5, l.117: 'were validated,

Response: We have corrected the subject-verb agreement by changing "was" to "were" (Line 118-120).

"The inversion results were validated, and the accuracy of the estimated parameters was assessed against in situ observations."

8. p.9, l.205: I,m unclear about whether this means Pw has been set to ice overburden everywhere (and therefore N = 0), or whether N has been set to ice overburden everywhere (and therefore Pw = 0). Either way, this is a very strong assumption that needs to be justified. If the former, physically, this means the authors are assuming all the weight of the ice is supported by basal water, implying that there is zero ice-bed friction. This is true in some places, some of the time, but I think assuming it as a default is perhaps a bit much. Something like N = 0.9(ice overburden) might be a little more defensible and better representative of the mean state of most glaciers? If the latter, the authors are ignoring hydrology entirely and therefore making it very difficult for any sliding to occur, which would seem a bit of a self-destructive assumption given the point of the model is to include sliding. Please clarify what,s happening here.

Response: We fully agree that the original wording—"approximating the effective pressure as ice overburden"—was ambiguous and that either interpretation (setting 'N = 0' or 'N = ice overburden') represents an extreme assumption lacking physical justification. Our original intention was indeed to set the effective pressure 'N' equal to the ice overburden ('Pw = 0'), completely ignoring subglacial hydrology. However, as the reviewer correctly pointed out, this assumption contradicts the very purpose of ITMSL, explicitly accounting for basal sliding, since sliding can hardly occur without hydrological modulation of effective pressure.

Following the reviewer's specific suggestion, we have re-set the effective pressure to 0.9 times the ice overburden pressure (N = 0.9 × Nice) and re-run the ITMSL model, updating all inversion results and associated figures and tables. Compared to the end-member values (0 or 1.0), the 0.9 assumption is physically more plausible and is a widely adopted approximation in regional-scale studies lacking in situ subglacial measurements.

The recalculated results show a slight improvement in ITMSL's accuracy: ITMSL achieves higher correlation coefficients with GPR-measured thickness than Millan in 9 of 16 glaciers and surpasses Gantayat in 13 of 16; the RMSE is reduced to 39.12 m, which is 32.65% lower than

Millan's 58.08 m and 20.62% lower than Gantayat's 49.28 m. Importantly, the main conclusions remain entirely consistent with those before the revision, ITMSL outperforms or matches the existing models across most velocity intervals. This confirms that the specific value of effective pressure does not fundamentally affect the relative performance ranking among models or our core findings. **Such as:**

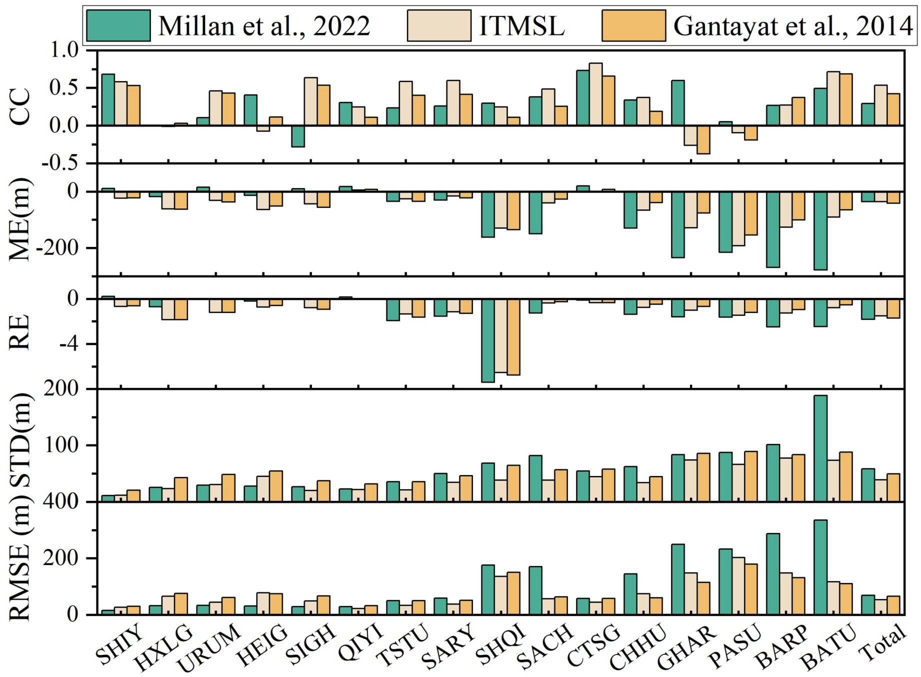

Figure 7: Comparison of ice-thickness inversion accuracy among three SIA-based models (Millan, Gantayat, ITMSL), referenced against GPR measurements. Metrics: correlation coefficient (CC), mean error (ME), relative error (RE), standard deviation (STD), and root mean square error (RMSE). All axis abbreviations are defined herein.

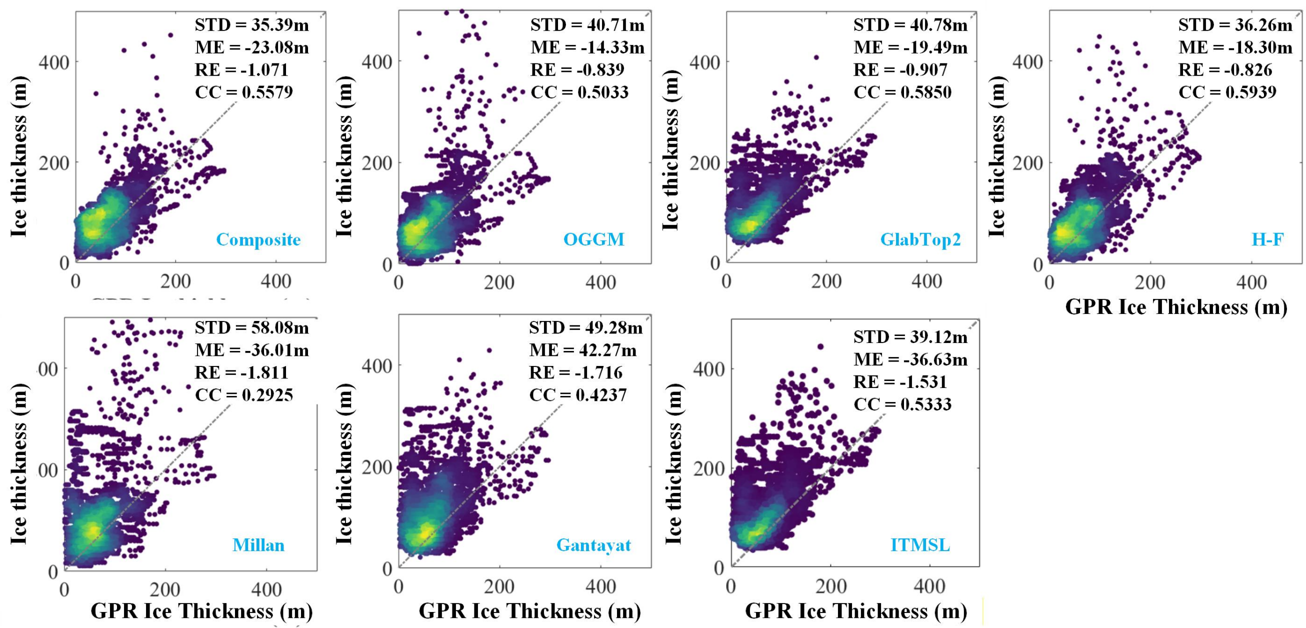

Figure 13: Evaluation of model-inverted ice thickness accuracy using GPR-measured ice thickness.

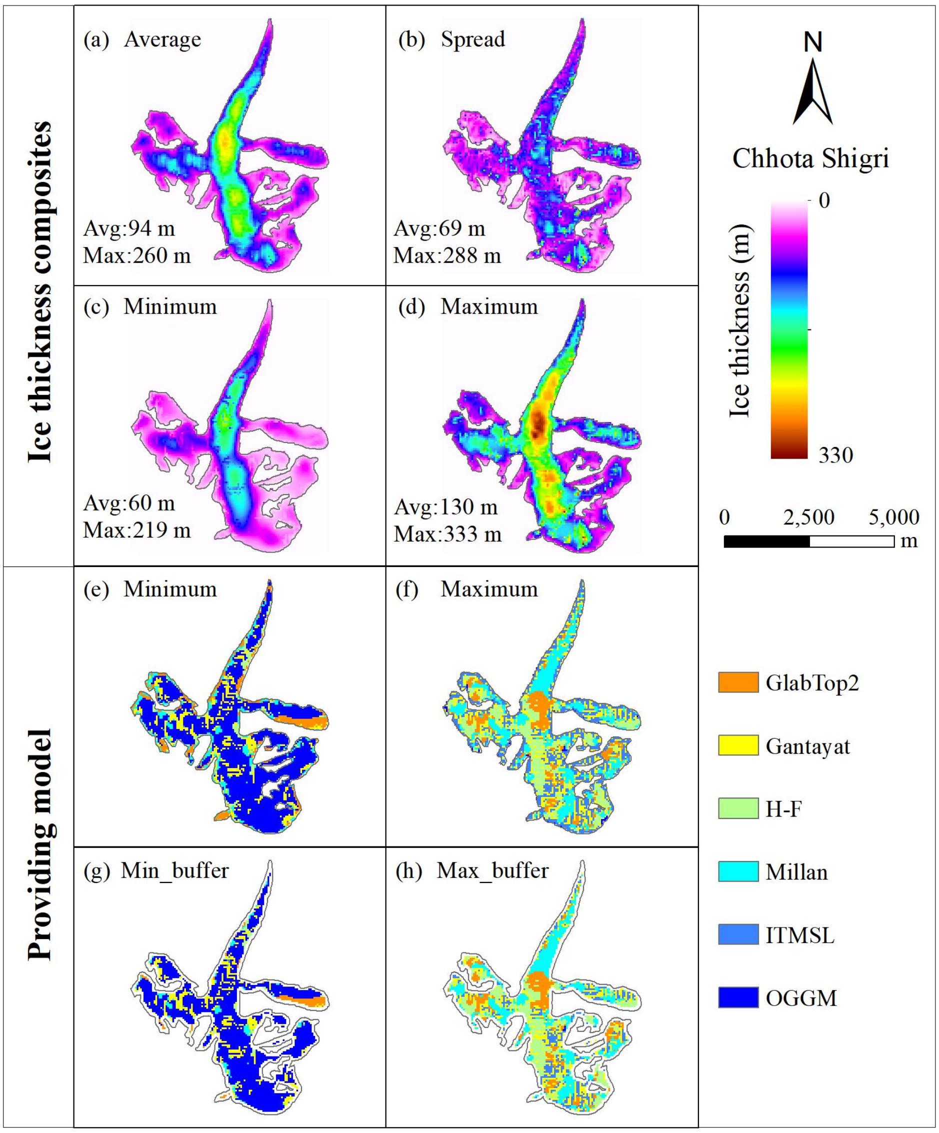

Figure 12: Comparison of simulated ice thickness among models. The first four panels present composite results from the models: (a) average, (b) spread, (c) minimal, and (d) maximal ice thickness distribution. The models corresponding to the minimum and maximum ice thickness at each pixel are shown in (e) and (f), respectively; Models corresponding to the minimum (g) and maximum (h) ice thickness at each pixel after excluding the 3×30 m buffer zone.

9. Figure 3: I,m very glad the authors have extended the data range a bit here, and I,m aware of the points they make in the response to reviewers on this. Still, as presented in the paper itself, 13/16 glaciers have the final, penultimate or antepenultimate value tested selected for B, which very much gives the impression that RMSE just drops as B increases and that a larger range of higher values for it should be included. I think the authors need to perhaps add a couple of sentences discussing this: presumably the apparent relationship is due to B effectively being a tuning parameter that compensates for other inaccuracies in the model, and that what is being sought is the minimum RMSE within the range of physically plausible B values, not the global minimum RMSE (which is what l.293 currently implies). Or, if the point is to find the global minimum RMSE, B really is just a tuning parameter, and we don,t need to worry about the physically plausible range of values, and therefore the authors should show the graphs with the extended axes that are in the response to reviewers.

Response: We thank the reviewer for revisiting this point and for acknowledging the extended B-value range we have implemented. Among the 16 glaciers, 13 have their optimal B values falling within the last three increments of the tested range. To address this, we have added the following discussion in the revised manuscript:

"The values of the ice fluidity parameter B () were initially adopted from the literature (Louis and Lliboutry, 1987). Given that the glaciers in this study are distributed across various regions of HMA, we expanded the parameter's range to in order to obtain locally optimized values within a reasonable range. In addition, the parameter B has been calibrated using ice thickness measurements derived from GPR."

In addition, we have included in the main text the full sensitivity curves with extended axes, to more clearly demonstrate that STD does not decrease significantly beyond the upper end of the selected range, thus justifying the chosen upper bound (Line 313-316).

Table 2. The optimal fluidity parameters of each glacier calculated by GPR.

Glacier	B ($MPa^{-3}yr^{-1}$)	Glacier	B ($MPa^{-3}yr^{-1}$)
BARP	745	SARY	470
BATU	740	SHQI	645
CHHU	735	CTSG	700
GHAR	735	SHIY	345
HXLG	685	SIGH	745
HEIG	715	TSTU	590
PASU	745	URUM	670
QIYI	725	Total	730
SACH	715		

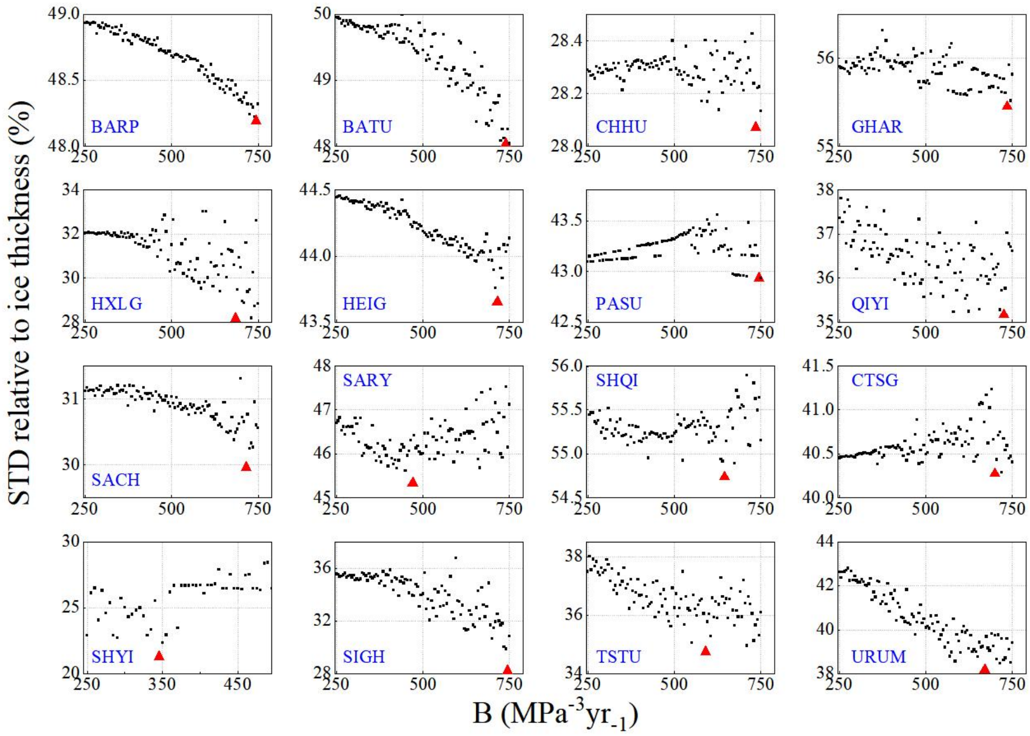

Figure 3: Ratio of the STD of ice thickness inversion to ice thickness for different values of parameter B in the ITMSL model.

10. p.16, l.351: 'regional,

Response: We have revised this sentence (Line 365-367):

“Consequently, incorporating the physically-based sliding law enhances laminar flow equation by explicitly representing subglacial processes, thereby improving regional ice thickness estimation accuracy.”

11. Table 3: I think the authors might want to think about their accuracy metrics here. This table shows that for the vast majority of the 16 (11/16) glaciers, ITMSL is out-performed by either Millan or Gantayat, even on nearly all the faster-flowing glaciers where one would expect inclusion of sliding in the model to improve the results. Of the 5 glaciers where this table shows it,s got the closest match to the validation thicknesses, 2 are only very slightly in its favour (to be fair, one of the 11 it,s not the best on is also only just not in its favour). As presented, it,s therefore not convincing me that ITMSL is really making an advance on existing thickness models, which is undermining exactly what Section 4.3, immediately following, is purporting to demonstrate. This is a tension that needs to be resolved to make the paper credible. Presenting Table 3 as mismatches and then showing the average mismatch (maybe ITSML is better there?) might also be more intuitive. I think it,s also worth pointing out that the modelled ub/us ratios are relatively similar, varying much less than the observed ice velocities, even for glaciers of seemingly similar thickness. As the ice thickness controls the amount of ice creep, the glaciers with higher velocities for the same ice thickness should have much more sliding and thus much higher ub/us ratios, and this is not seemingly something that the model is recognising. Or, the GPR ice thicknesses are very unrepresentative of the mean ice thicknesses at each glacier. Or the mean ratio isn,t very meaningful, which I think is what,s actually going on. This is also something that maybe needs to be considered when redesigning the table.

Response: We thank the reviewer for the thorough analysis of Table 3 and its implications for the overall argument of the paper. We agree that the original presentation could indeed lead readers to conclude that ITMSL does not outperform the existing models on most glaciers, contradicting the argument we intended to make in Section 4.3. To resolve this tension, we have made the following revisions to the table and the accompanying discussion:

1. Redesigned presentation of Table 3

We have restructured Table 3 to show the mean ice thickness and MAE for all three models (Millan, Gantayat, and ITMSL) for each glacier, rather than focusing on which model is the “best” for each glacier. This presentation allows readers to directly compare the absolute errors across models and to more comprehensively assess the error-distribution characteristics of ITMSL.

2. Added discussion on the “higher mean but lower MAE” pattern (**Line 357-365**):

“It is worth noting that, although the mean ice thickness estimated by our model is slightly higher than that of the other two models, it achieves a superior MAE. This indicates that our model captures the mechanical coupling between surface velocity and ice thickness more effectively and exhibits greater resilience to data noise, thereby substantially reducing extreme outliers. Nevertheless, the systematic overestimation in the mean estimates suggests that the current physical parameterization does not fully account for the diversity of subglacial hydrological conditions across the study area, leading to a general upward bias. Unlike the competing models, whose favorable mean performance stems from mutual error cancellation, our model produces a more structured error distribution, making it better suited for regional-scale ice-thickness inversion.”

Table 3: Comparison of mean ice thickness from GPR measurements and retrievals by three models (Gantayat, Millan, ITMSL), along with their respective mean absolute errors (MAE), for the 16 glaciers.

Glacier	GPR (m)	Gantayat et al., 2014		Millan et al., 2022		ITMSL	
		Ice thickness (m)	MAE (m)	Ice thickness (m)	MAE (m)	Ice thickness (m)	MAE (m)
BARP	120.95	143.86	104.81	191.74	269.21	161.88	127.07
BATU	127.06	142.37	89.75	211.63	285.34	156.54	93.47
CHHU	103.93	113.63	45.55	142.55	132.14	124.33	66.47

GHAR	152.94	114.62	91.82	153.15	234.64	135.43	128.52
HXLG	43.96	83.91	62.57	42.60	24.46	82.98	61.61
HEIG	114.87	94.88	57.55	68.75	24.52	103.58	64.52
PASU	145.55	156.36	154.82	185.24	215.10	163.64	191.79
QIYI	77.81	63.58	24.58	53.26	23.67	62.29	18.10
SACH	121.79	123.40	48.54	185.03	153.84	128.42	47.26
SARY	75.52	77.80	39.99	81.28	48.44	75.69	31.94
SHQI	38.68	71.31	135.45	80.88	162.02	70.21	130.26
CTSG	172.88	104.90	47.54	98.39	48.62	110.70	36.52
SHIY	39.47	49.18	24.99	30.09	12.07	53.09	23.55
SIGH	68.93	71.99	55.89	42.28	26.30	70.94	44.49
TSTU	51.63	64.63	38.22	63.55	38.53	59.88	27.33
URUM	40.05	78.45	45.97	45.61	28.50	80.53	34.77

3. Discussion on the ‘ub/us’ ratio and model physics

Regarding the reviewer‘s observation that the modelled ‘ub/us’ ratios exhibit relatively little variation compared to the observed velocity range, and that glaciers with higher velocities for similar thicknesses should have higher ‘ub/us’ ratios, we agree that this is a valid concern and suggests that ITMSL’s capture of this relationship is not yet perfect. GPR transects are typically located along glacier flowlines, where ice thicknesses may not be representative of the glacier-wide mean (often being thicker than the average), whereas the model outputs represent the mean over all glacier pixels, introducing a mismatch in the comparison baseline; The glacier-wide mean of the ‘ub/us’ ratio is statistically not robust enough for such comparisons—it is highly sensitive to measurement point locations, spatial resolution of the velocity data, and local variations in sliding distribution. We therefore suggest that readers focus instead on the systematic error structure, which provides a more robust basis for model evaluation.

Overall, we believe that ITMSL’s advantage lies in its improved error structure rather than in simply “winning” on individual glaciers. This distinction is more meaningful for regional-scale inversion, where systematic errors are more amenable to calibration.

12. p.17, l.357: Presumably, these are the correlation coefficients between the GPR thickness and the modelled thickness? It,s not entirely clear as written.

Response: We have revised the sentence in the main text to explicitly state that these are the correlation coefficients between the modelled ice thickness and the GPR-measured ice thickness. The revised sentence reads (Line 372-375):

“ITMSL achieves higher correlation coefficients between the modelled and GPR-measured ice thickness than Millan in 8 of 16 glaciers, and surpasses Gantayat in 12 of 16, indicating stronger linear relationships with the observed data.”

13. Figure 7: The caption needs to be expanded a bit. At the very least, the abbreviations on the figure axes should be stated in full here, so the reader doesn,t have to find them in the text. Based on this figure, I make it that ITSML has the highest CC on 9 glaciers, lowest ME on 4, lowest RE on 3 or 4, and lowest RMSE on 7. I get that ITSML is individually better than either Gantayat or Millan, but I,m not being convinced it,s a major advance, especially when it seems to be clearly worse than Gantayat on the most heavily sliding glaciers (assuming the fastest glaciers have the most sliding) where inclusion of the sliding law should really matter. As the discussion then goes on to point out, ITSML doesn,t work so well on small, non-sliding glaciers either, so it seems that what it,s actually good at is medium-sized, slightly slidey glaciers, which isn,t nothing, but also not quite the slam dunk the paper wants it to be. Figure 9 and the associated text do a much better job of presenting the advantages of ITSML, so I,m wondering if it would be better to present results generally in terms of a split between slower- and faster-flowing areas, because the mean figures offer only very lukewarm support for the authors, contention that ITSML is doing well. This would also be more intuitive: including the sliding law shouldn,t make any difference in slow-flowing areas, because there,s no sliding, but should matter in fast-flowing ones, and rethinking the results section to highlight that might make it more convincing.

Response: Thank you for your valuable feedback. We have made the following revisions:

（1）The revised caption for Figure 7 is as follows:

“ Figure 7: Comparison of ice-thickness inversion accuracy among three SIA-based models (Millan, Gantayat, ITMSL), referenced against GPR measurements. Metrics: correlation coefficient (CC), mean error (ME), relative error (RE), standard deviation (STD), and root mean square error (RMSE). All axis abbreviations are defined herein.”

（2）Re-evaluating the statistical message:

We acknowledge that ITMSL does not “win” on the majority of individual glaciers for every metric in Figure 7. However, we argue that ITMSL‘s strength lies not in maximizing the count of “best-performing” glaciers, but in improving the error structure—i.e., suppressing extreme over- and under-estimates. As noted in our discussion, while competing models achieve favorable means through error cancellation, ITMSL produces a more concentrated and systematic error distribution, which is more valuable for regional-scale calibration.

（3） Addressing the “poor performance on fastest glaciers”:

Following the reviewer’s explicit suggestion that presenting results via a split between slow-and fast-flowing areas is “more intuitive,” we have added a dedicated section 5.1 where we classify pixels into 5 m/yr velocity intervals. **Slow regime (≤50 m/yr):** All three models perform similarly. **Fast regime (>70 m/yr):** ITMSL shows a significantly lower RMSE and a much gentler R^2 (0.87 for ITMSL vs. 0.76 for Millan and 0.78 for Gantayat). (Line 410-429)

“5.1 Influence of Surface Velocity on Ice-Thickness Model Performance

To evaluate the applicability of various ice-thickness inversion models to glaciers with different surface velocity, we classified surface velocity into intervals of 5 m/yr and calculated the ice thickness bias, RMSE for the models of Millan et al. (2022), Gantayat et al. (2014), and our proposed ITMSL, as well as the RMSE differences of ITMSL relative to the other two models (Figure 8). The results show that the ice-thickness bias of all three models increases approximately linearly with surface velocity, with R^2 of 0.76, 0.78, and 0.87, respectively. In the low-velocity regime (≤50 m/yr), ITMSL consistently yields lower RMSE than the Millan and Gantayat models, and the biases relative to GPR measurements remain relatively small for all models. However, when surface velocity exceeds 70 m/yr, both bias and RMSE rise to over 100 m. Analysis of the RMSE differences reveals that, with increasing velocity, ITMSL exhibits a progressively larger advantage over Millan. In contrast, the difference relative to Gantayat shows a two-stage pattern: below 50 m/yr, the difference increases with velocity; around 50 m/yr, ITMSL performs comparably to or slightly worse than Gantayat; and beyond that threshold, the difference resumes its upward trend. Overall, ITMSL outperforms or matches the existing models across most velocity intervals, with only a narrow transitional zone (~50 m/yr) where it is on par with Gantayat. These findings confirm the robustness and demonstrated superiority of ITMSL under varying flow conditions.

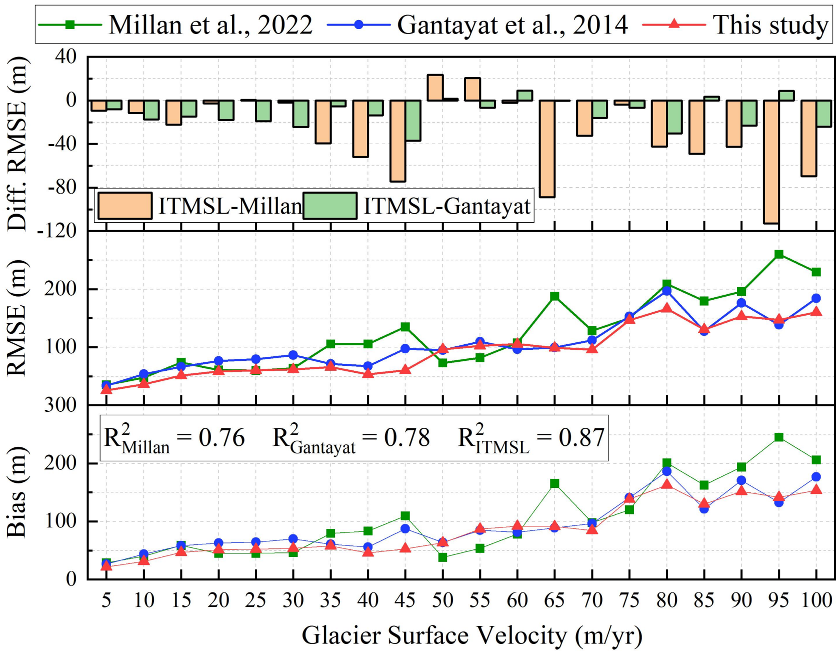

Figure 8: Comparison of ice-thickness inversion bias, root-mean-square error (RMSE), and their differences among ITMSL, Gantayat, and Millan models across surface velocity intervals.

We believe this regime-based presentation is far more convincing than glacier-wide averages, as it directly validates the physical rationale for including a sliding law. We are grateful to the reviewer.

14. p.22, l.455-462: I feel this section could be more impactful if the authors presented it as ITMSL is producing more moderate, sensible ice thicknesses, rather than very thin or very thick ones, and is therefore a bit more reliable than at least some of the other models. As it,s currently written, that is implied, but not explicitly stated, which means it pulls its punches a bit.

Response: We agree that the original conclusion in this section was somewhat understated and did not explicitly highlight ITMSL's advantage in reducing extreme thickness values. We have revised the concluding sentence in the main text to state more directly that ITMSL yields more moderate and physically plausible ice thicknesses in glacier interiors, with substantially fewer extreme thin or thick values than most competing models—demonstrating its greater reliability for regional-scale ice-thickness estimation. We believe this revision makes the take-home message of Figure 12 more impactful and better aligns with the overall narrative of ITMSL's systematic error-structure advantages. The revised sentence now reads:

"As described above, in the ITMSL data processing workflow, ice bed roughness is quantified using the RMSH method with a 3 × 3 pixel moving window. This configuration is likely the primary cause of the extreme errors observed at glacier margins in the ITMSL outputs. To mitigate boundary effects, we applied a buffer distance of 3×30 m to exclude marginal pixels (Figures 12g and 12h). After this exclusion, the proportions of minimum extreme values among the models' thickness estimates over interior pixels rank as follows: ITMSL (0.25%) < H-F (2.38%) < Millan (3.75%) < GlabTop2 (3.91%) < Gantayat (15.45%) < OGGM (74.28%); while the maximum extreme values rank as: OGGM (0.11%) < GlabTop2 (14.71%) < ITMSL (15.14%) < H-F (21.21%) < Gantayat (21.76%) < Millan (27.06%). These results indicate that, although ITMSL yields relatively large errors at glacier margins, its thickness estimates for interior pixels are more concentrated and reasonable."

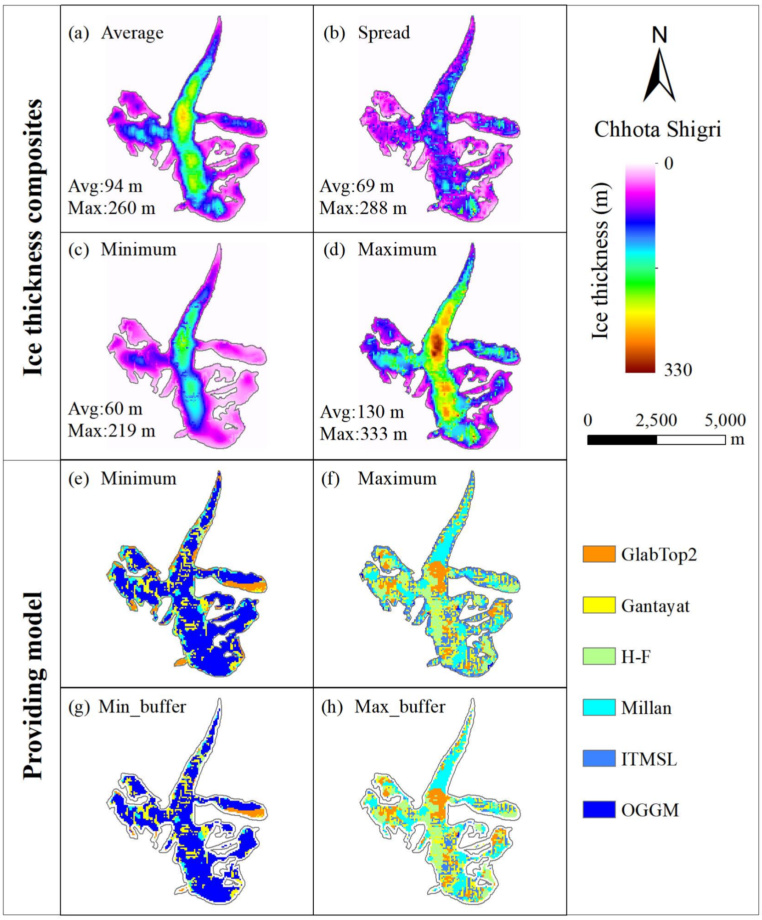

Figure 12: Comparison of simulated ice thickness among models. The first four panels present composite results from the models: (a) average, (b) spread, (c) minimal, and (d) maximal ice thickness distribution. The models corresponding to the minimum and maximum ice thickness at each pixel are shown in (e) and (f), respectively; Models corresponding to the minimum (g) and maximum (h) ice thickness at each pixel after excluding the 3×30 m buffer zone.

15. Figure 12: Again, this isn,t super-convincing that ITSML is doing notably better than any of the other models, largely because all it,s showing is the ice thickness, not the mismatch. ITSML generally seems to predict higher ice thicknesses, but the figure on its own doesn,t provide any information about whether that,s at all realistic. I,d suggest doing the same plot, but graphing the mismatches, as the text states (l. 464) that Figure 12 is related to an accuracy assessment on these data. At the moment, all it shows is the data, not anything about the accuracy, which is the more useful part for the reader to know about. Though, to be honest, with Figure 13 there, that might not be that useful either, as that largely shows that.

Response: We agree that displaying ice thickness distributions without reference to the observed data does not convey meaningful information about model accuracy. Accordingly, we have made the following revisions:

(1) Streamlining of the main text:

We have removed the earlier comparison of model-mean ice thicknesses, as it overlapped with Figure 13, and the use of mean values tended to obscure important error-structure information. This deletion makes the manuscript more concise.

(2) Revision of the main content of Figure 12:

We showing the model assignments for minimum and maximum ice thickness at each pixel after excluding the 3 × 30 m marginal buffer. These panels demonstrate that ITMSL has the lowest proportion of extreme values among interior pixels—the key finding of this figure.

(3) Bias maps moved to the Supplementary Material:

To avoid overcrowding the main figure, we have placed the spatial distribution of biases (modelled minus GPR-measured thickness) for all models in the Supplementary Material (Figures S2–S4).

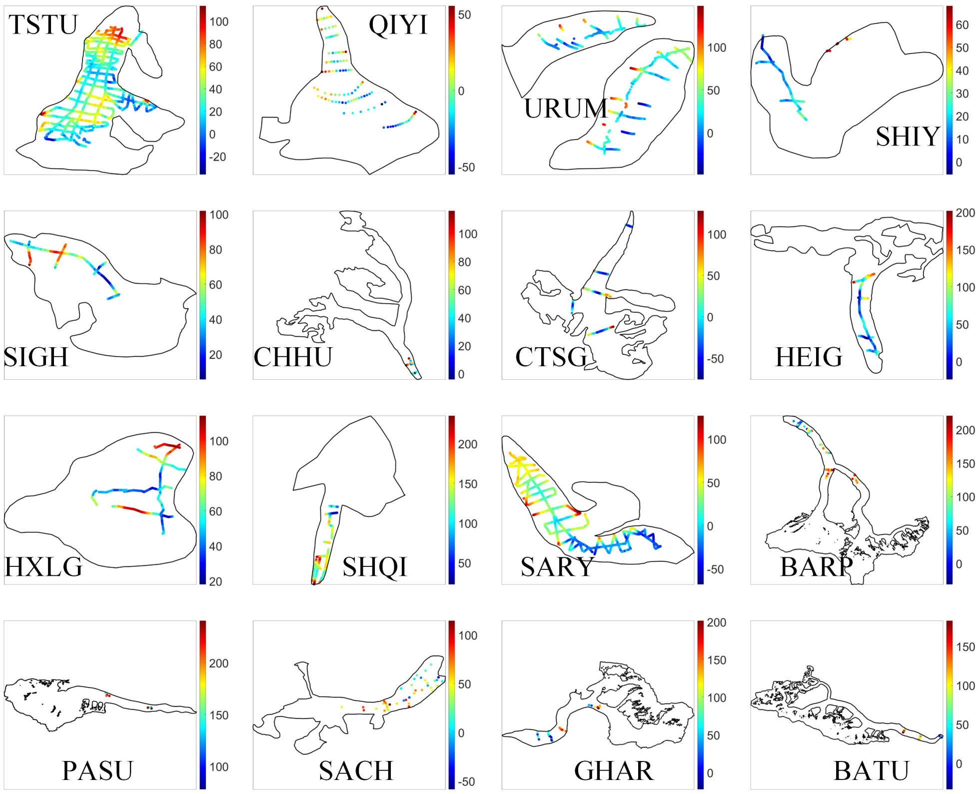

Figure S2: Glacier thickness inverted by Gantayat (unit: m).

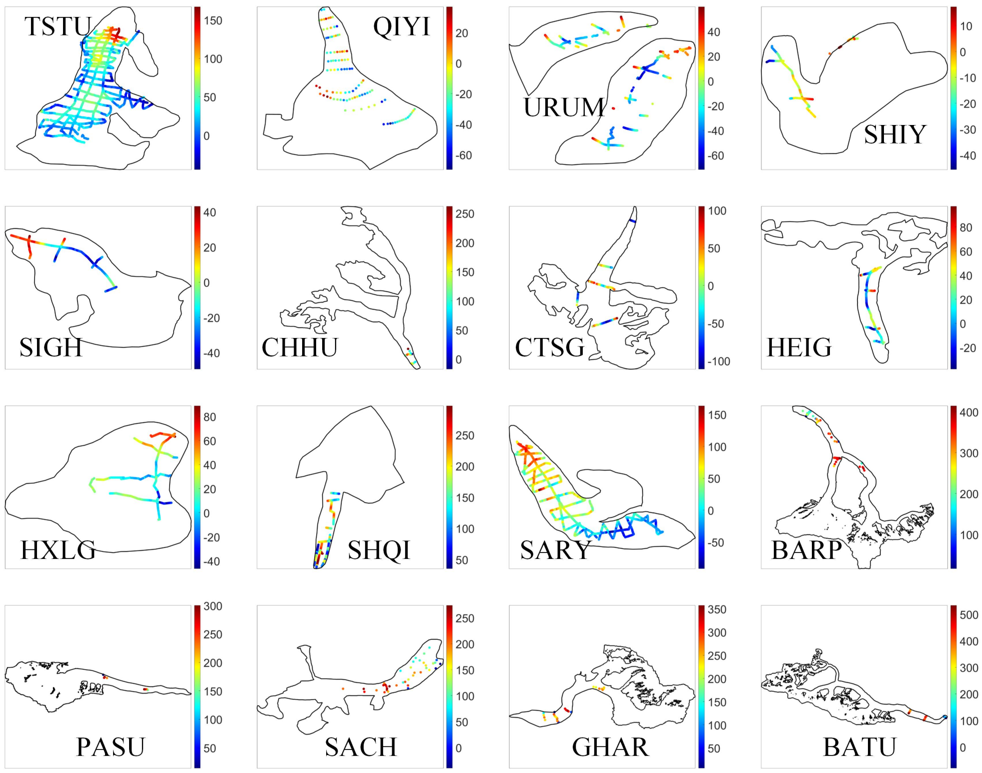

Figure S3: Glacier thickness inverted by Millan (unit: m).

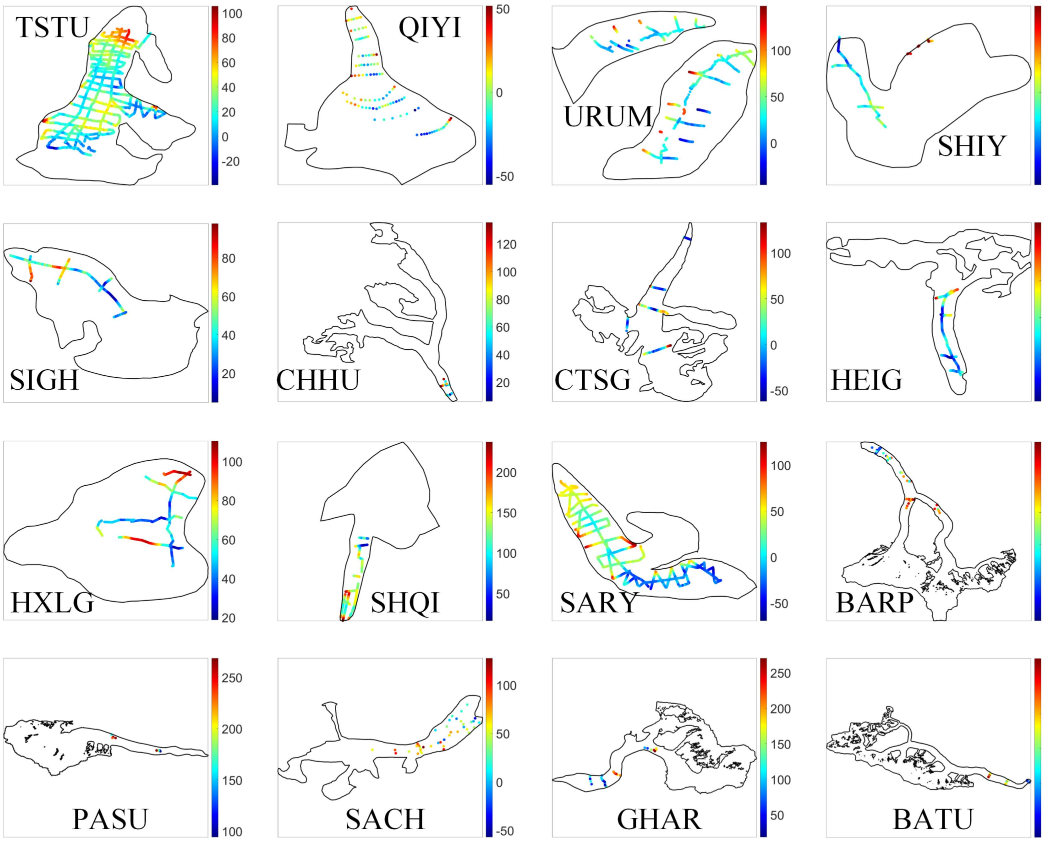

Figure S4: Glacier thickness inverted by ITMSL (unit: m).

Xiaoguang Pang and colleagues develop an ice thickness inversion model based on laminar-flow theory that explicitly incorporates a sliding law, whereas basal sliding parametrization is often neglected or oversimplified in comparable approaches. They apply their model (ITMSL) to 16 glaciers across High Mountain Asia (HMA). First, they compare ITMSL with two other laminar- flow-based inversion methods (Gantayat and Millan). Second, they compare ITMSL with four additional ice thickness inversion models that are well established in the literature.

I received now a new version of the manuscript after a first tour of review. I thank the authors for their work to account for my comments on their first manuscript. However, after reading the new version of the study, I think some of my initial comments were not well addressed and the manuscript still requires improvement.

I want also to inform the authors that the document provided to track changes between versions of the manuscript is not usable in its current form. New text is highlighted in red with a line to the margin of the pages where the previous sentence should be written. Most of the time the previous content is not written and only "Times New Roman" is displayed.

The line numbers refer to the newest version of the manuscript.

Response: We sincerely thank the reviewer for their continued engagement with our manuscript during this second round of review.

1.Technical issue with the tracked-changes PDF

We sincerely apologize for this issue. It arose because numerous formatting traces (e.g., font styles, layout adjustments) were retained during our revision process, which occupied the sidebar space and compressed the substantive changes, making it difficult for the reviewer to track the modifications properly. To thoroughly resolve this problem, we have taken the following remedial measures:

(1) We are submitting, alongside this revised manuscript, a detailed summary of changes that lists each reviewer comment, our corresponding response, and the exact location of each revision, allowing the reviewer to verify all modifications efficiently;

(2) In addition, for non-substantive formatting or stylistic adjustments, we have ensured that they do not appear as tracked changes, thereby avoiding redundant clutter in the sidebar and ensuring that all substantive modifications are fully retained in the margin, so that the reviewer can directly identify them within the PDF of the main text.

2. Regarding the concern that "some initial comments were not well addressed"

We take this concern very seriously. In fact, we had addressed each of the reviewer's original comments point-by-point during the first revision and incorporated the corresponding changes into the manuscript. However, due to the technical problem described above, some of these modifications may not have been visible or clear to the reviewer. To fully address this, we have included in the supplementary materials a detailed response matrix that maps each comment to the specific changes we made, including new analyses, updated figures and tables, and the exact line numbers in the revised manuscript. This will allow the reviewer to verify each point efficiently. We regret the inconvenience caused by the technical issue with the tracked-changes file. We trust that the newly submitted documents and the detailed summary table will resolve this problem and enable the reviewer to assess our revisions smoothly and thoroughly.

Major Comment

While reading the paper, it is difficult to be convinced of the added values of ITMSL.

1. First, it slightly outperforms Millan and Gantayat over these 16 glaciers. However ITMSL was calibrated with ice thickness measurements glacier by glacier while Millan calibrated his model at regional scale. This difference in calibration (and thus in application at large scale or not) could also explain the better performance of ITMSL versus Millan.

   Response: We thank the reviewer for raising this critical issue. We agree that the difference in calibration scale is an important factor in interpreting ITMSL's relative performance against Millan, and we have given this careful consideration. Our view on this is as follows:

   (1) Acknowledging the calibration difference

   We acknowledge that ITMSL was calibrated on a glacier-by-glacier basis, whereas Millan et al. (2022) calibrated their model regionally across the entire domain. This difference may indeed contribute to ITMSL's better performance on the glaciers used for calibration.

   (2) Justification for glacier-by-glacier calibration

   However, we would also emphasize that glacier-by-glacier calibration is a common and legitimate approach in ice-thickness modelling studies, given the substantial climatic, geological, and thermal heterogeneity among glaciers in HMA. The optimal parameter combination for one glacier may not be appropriate for another. This approach is widely used precisely to test whether a model's physical structure is sound.

   (3) ITMSL's improvement also stems from its physics, not calibration alone

   We believe that the improvement of ITMSL is not solely attributable to glacier-specific calibration. The reviewer's observation of the pronounced advantage in fast-flowing regimes (>70 m/yr) is illustrative: if calibration were the only reason for ITMSL's better performance, the gain would be expected to affect slow and fast glaciers similarly. Yet, ITMSL's advantage is most evident precisely in fast-flowing regions where sliding is significant (see new Section 5.1 and Figure 8)—which directly reflects the contribution of the sliding law, not merely calibration.

   (4) Clarifying the practical value of this study

   We agree that glacier-by-glacier calibration is not feasible for glaciers entirely lacking measured thickness data. However, the contribution of our study lies in demonstrating that when limited thickness observations are available for a glacier, ITMSL can make more effective use of that information to provide more accurate estimates than existing regional models. This remains a practically useful scenario in many high-mountain regions where some in-situ measurements are available, even if not globally.

2. Second, I still see no explanation in the manuscript to motivate the use of ITMSL instead of any other inversion methods. Some of these approaches do not require more data that what is needed for ITMSL and seems to better performs as you show in section 5.3. Maybe the reduced complexity reduces strongly the computational time and this is clearly an advantage but this should be referenced in the manuscript. You indeed mention : "Although complex ice thickness inversion models offer advantages in physical realism and inversion accuracy, their high computational cost, large number of parameters limit their application in regional and global scale glacier volume studies" while we see several of these methods applied at large scale (Farinotti et al.,2019, Cook et al. 2023, Frank et al., 2026). Or maybe the simplicity of SIA allows for physical understanding of glaciological phenomena. Again the added value of ITMSL should be more emphasized.

   Response: We agree that the manuscript should more clearly articulate why we developed ITMSL rather than adopting existing methods. This question is closely related to a similar concern raised by another reviewer in the first round. We provided a detailed response at that time, and we now restate its core content here, with additional clarifications specifically addressing this reviewer's comments:

(1) Focused positioning within the ITMIX framework

The ITMIX experiment includes 17 ice-thickness models across five categories. Our goal was not to develop a "best-overall" model across all categories, but rather to address a specific deficiency within the ice-velocity-based approaches, namely, that existing models handle the basal sliding parameter in an oversimplified manner. In the ITMIX framework, Gantayat et al. (2014) assume a fixed ratio between basal and surface velocities (ub = α·us), which is physically crude. Millan et al. (2022) improved this by deriving a spatially variable ub/us ratio from surface velocity and subglacial topography, but the parameterization still lacks a direct physical constraint. We selected Gantayat and Millan as our comparison targets for two reasons:

(a) all three models share the same SIA theoretical basis, ensuring comparability;

(b) both existing models have deficiencies in determining the sliding parameter, which ITMSL aims to address by introducing the Coulomb friction law to explicitly link sliding velocity to effective pressure and basal shear stress.

(2) Why not adopt the "better-performing" complex models?

The reviewer correctly notes that some complex models outperform ITMSL on certain glaciers. We fully acknowledge this, but it does not diminish ITMSL's value, because ITMSL is not intended to replace these complex models, but to improve accuracy within the SIA-based ice-velocity category. The added value of ITMSL lies in:

(a) Accuracy improvement: ITMSL outperforms Gantayat and Millan across most of the 16 glaciers within the same model category, with a particularly pronounced advantage in fast-flowing regimes (>70 m/yr) (Section 5.1, Figure 8), demonstrating that the explicit sliding law improves the physical representation.

(b) Interpretability: The sliding parameters in ITMSL (e.g., effective-pressure coefficient, friction coefficient) have clear physical meanings, making them easier to calibrate with field data and to interpret, compared to black-box or highly parameterised models.

(c) Modest data requirements: ITMSL requires only surface velocity, DEM, and glacier outlines, the same data as Gantayat and Millan—imposing no additional data burden.

3. Thanks you for introducing ITMIX in the introduction. However I'm not convinced by the analysis performed around fig.11. I do not see the point of this analysis in term of bench marking the performance of ITMSL with respect to other models. Analysis of the performance of ITMSL without (or with a reduced set of) ice thickness measurements compared to other models should be a way to answer this issue. This will also enhance the added value of ITMSL if working without ice thickness measurement.

Response: We thank the reviewer for the careful scrutiny of Figure 11. We agree that the comparison presented there suffers from an inherent asymmetry, ITMSL was calibrated against GPR-measured ice thicknesses, whereas the other models (GlabTop2, H-F, OGGM) provide independent predictions. A direct accuracy comparison under such conditions would indeed be unfair.

(1) Clarifying the original purpose of Figure 11

The purpose of introducing the ITMIX model comparison in this section was not to compete with the complex models in a "benchmarking" sense (that task was already accomplished through the comparisons with Gantayat and Millan in Figures 7 and 8). Rather, it was to illustrate the spatial error structure of ITMSL, specifically, that ITMSL produces far fewer extreme values (both minimum and maximum) in glacier interiors than the other models, indicating that its predictions are spatially more uniform and physically more plausible.

(2) Regarding the reviewer's suggestion on "data-holdout experiments"

The reviewer raises a highly valuable point: the real value of ITMSL lies in whether it can maintain reasonable accuracy in the absence of measured ice thickness data. We have not conducted a systematic data-holdout experiment in the current study, but we have added a paragraph in the Conclusions and perspectives section that explicitly acknowledges this limitation and identifies it as a key direction for future work.

(3) Reinforcing the added value of ITMSL

Even without a dedicated holdout experiment, the potential advantages of ITMSL in data-scarce settings can be inferred from our existing results: (1) Identical data requirements: ITMSL requires the same input data (surface velocity, DEM, glacier outlines) as Gantayat and Millan, imposing no additional data burden. Given the same data, ITMSL outperforms both. (2) Physically transferable parameters: The key parameters in ITMSL (e.g., effective-pressure coefficient) have clear physical meanings. This means that parameters calibrated on a limited set of glaciers with measured thickness may be transferable to unmeasured glaciers with similar characteristics.

(4) Future work

As already noted in the Conclusions and Future Perspectives, we anticipate that the growing abundance of remote sensing data will enable regional-scale ice-volume inversion, and the results will be thoroughly compared with existing ice-thickness products in subsequent work.

"Previous studies have demonstrated that multi-model ensemble approaches offer significant advantages in ice thickness inversion. Therefore, ITMSL should be used as a member of such an ensemble, rather than as a replacement for other methods. With the growing availability of high-resolution remote sensing data, ITMSL shows great potential for glacier thickness estimation. However, the ITMSL does not account for the effects of glacier temperature and surface debris on ice thickness estimation, and lacked verification of modeled glacier sliding measurements. Despite these limitations, ITMSL model establishes a promising foundation for estimating ice thickness in complex mountain environments, particularly where traditional SIA assumptions prove inadequate."

Minor Comments

0) Abstract

1. L.18-20: "*Currently, ice thickness inversion models based on ice velocity have achieved significant progress in regional and global studies of glacier thickness and volume. However these methods simplify the parameterization of basal sliding […]*" . I would prefer to limit this statement to SIA based inversion model not on all inversion model using ice velocity.

   Response: We have revised this sentence (L18-20).

   "Currently, ice thickness inversion models based on shallow ice approximation (SIA) have achieved significant progress in regional and global studies of glacier thickness and volume."

2. L.21-22: this sentence is doubled.

   Response: We thank the reviewer for catching this oversight. We have carefully checked the manuscript and removed the duplicated sentence (L 21).

   "~~However, this method simplifies the parameterization of basal sliding, leading to uncertainties and significant biases in thickness estimates.~~"

1) Introduction

1. L.59 : You must introduce and expand the acronym ITMIX.

Response: We have added the full spelling of ITMIX (L 59-62).

“ To evaluate the accuracy and limitations of ice thickness inversion models, the ice thickness models intercomparison experiment (ITMIX) systematically assessed 17 models that infer ice thickness from glacier surface characteristics (Farinotti et al., 2017).”

2. L.64 : “typically requiring two or more types of data from digital elevation models (DEMs)”. Only 1 DEM is provided in ITMIX where most of the state of the art inversion models participated. Where does this statement come from ?

Response: We thank the reviewer for pointing out this imprecise statement. We acknowledge that in the ITMIX framework and most standard inversion experiments, only a single DEM is provided as input. Our original wording (“from DEMs”) incorrectly implied the use of multiple DEM sources, which was not our intention. We have revised the sentence to clarify that models typically require a DEM in combination with at least one additional type of ancillary data (e.g., glacier outlines, surface velocity fields, or mass balance observations), rather than multiple DEM sources. The correction has been made in the revised manuscript (L 66-69).

“Different models have varying input data requirements, typically requiring a digital elevation model (DEM) combined with at least one additional dataset, such as glacier outlines, surface velocity fields, or mass balance data (Farinotti et al., 2017; Farinotti et al., 2021).”

3. L.73-74 :”with uncertainties reaching tens to even hundreds of meter” should be express in % of mean ice thickness to generalize across various glacier size.

Response: We agree that expressing uncertainties in absolute values (tens to hundreds of meters) is not generalizable across glaciers of varying scales. We have revised the statement to express uncertainties as a percentage of the mean ice thickness (L 75-77):

“Substantial disparities exist among different models in their thickness estimates for the same glacier, with relative uncertainties typically exceeding 30% of the mean ice thickness.”

4. L.76 : simulations → inversions

Response: We have revised this sentence (L77-79).

“Weighted ensemble averaging of multiple models can effectively offset systematic biases inherent in individual models, thereby ameliorating the accuracy of ice thickness inversions (Farinotti et al., 2017).”

5. L.78-80 : I’m wondering if this conclusion from ITMIX II is useful for the purpose of your study.

Response: We removed the sentence (L 81).

“~~At the same time, the experiment highlighted the critical role of sparse in-situ data and recommended prioritizing measurements in the thickest location of glaciers.~~”

6. L. 199 : When introducing the model of Millan and Gantayat you will use as benchmark, you should give the references.

Response: We thank the reviewer for this comment. Upon checking, we found that no direct introduction or first citation of Millan et al. (2022) and Gantayat et al. (2014) occurs around Line 199 in the current version of our manuscript. Descriptions of these two models are instead distributed across several paragraphs in the Introduction and Methods sections, and appropriate references were provided at their first occurrence in each of those sections. To avoid any confusion, we have reviewed the entire manuscript to ensure that all references to these two models are properly cited wherever they appear. We appreciate the reviewer's careful reading and helpful suggestion.

2) Methods

1. L.150 : SIA acronym was already introduced before.

   Response: We thank the reviewer for pointing this out. We have removed the redundant full spelling at this location and retained only the abbreviation "SIA" to avoid repetition (L 152-153).

   "Ice thickness models based on the SIA estimate glacier thickness using ice motion characteristics (Gantayat et al., 2014; Millan et al., 2022)."

2. Table 1 : what are the values of u_s displayed in the table ? Mean, Median, max ? From which observations ?

   Response: We have added this information (L158-159).

   "Based on the above data, we calculated the mean surface velocity for each glacier (Table 1)."

3. L.189-201 : This paragraph reads strange. You introduce Weertman and Budd's sliding law but you did not introduce the Coulomb's sliding law. Then you say you use the law of Coulomb.

   Response: We have added this information (L202-207).

   "According to the Coulomb friction law, basal sliding velocity is positively correlated with basal shear stress and negatively correlated with effective pressure. Sliding accelerates once the shear stress exceeds the frictional threshold, with lower effective pressure further promoting motion. Nevertheless, the shear-stress-to-effective-pressure ratio, after peaking, declines with increasing sliding velocity and ultimately stabilizes (Zoet and Iverson, 2020; Helanow et al., 2021)."

3)Model instructions

/

4) Model experiments

1. L.298-300 : " This indicates that the B value range is not universally suitable for all glaciers, [...]". I'm wondering if you could not argue the opposite. For most of the 16 glaciers, the RMSE difference between ice thickness inverted with different values of B and measured ice thickness is smaller than 2%. It never exceeds 4%. This small sensitivity to the parameter B makes the model relatively easy to extent to other glaciers without ice thickness measurements. This is encouraging even if it needs to be further investigated.

   Response: We thank the reviewer for this constructive suggestion. We fully agree that our original interpretation of this result was overly negative and failed to highlight its encouraging implication. Although the optimal B value varies among glaciers, the standard deviation of the inverted ice thicknesses across different B values remains consistently small (<2%, never exceeding 4%). This low sensitivity to B suggests that ITMSL is robust and relatively easy to apply to other glaciers, even in the absence of ice thickness measurements.

2. Table 2 : It is a very personal choice. I would like to see visual separation between the 2 first columns and the 2 last columns since they content the same variables.

   Response: We thank the reviewer for this suggestion. We explored the reviewer's suggested horizontal/vertical rearrangement, but found that a horizontal layout exceeded the page width due to the 16 glaciers, while a vertical layout left excessive blank margins on both sides and made cross-glacier comparisons difficult.

3. L.315-316 : mean ice thickness of a glacier from GPR measurement is the mean of the measurements ?

   Response: We thank the reviewer for pointing out this imprecision. Our original statement, that 172.88 m is the mean ice thickness derived from the 315 GPR measurement points, rather than the spatially averaged thickness of the entire glacier, was indeed unclear. We have revised the text (L 323-324):

   "The mean ice thickness derived from the GPR validation points at the CTSG glacier is 172.88 m."

4. L.321 : lowest → thinnest ?

   Response: We have revised this sentence (L327-330).

   "In terms of mean ice thickness, ITMSL (110.70 m), Millan (104.90 m), and Gantayat (98.39 m) differ from one another, with Gantayat being the thinnest. Crucially, ITMSL exhibited closest agreement with GPR data, indicating superior accuracy relative to the comparison models."

5. L.321-322 : "ITMSL exhibited closest agreement with GPR data, indicating superior accuracy relative to the comparison models.". Are they other models calibrated against individual ice thickness measurements ?

   Response: We fully acknowledge that the difference in calibration scale (glacier-specific vs. regional) may contribute to ITMSL's better performance on the glaciers used for calibration. This is precisely why we have repeatedly emphasised in the revised manuscript that ITMSL's core value lies in its ability to make more effective use of limited measured thickness data when such data are available. At the same time, we note that ITMSL's improvement is not solely attributable to calibration. As shown in the newly added Section 5.1 (Figure 8), ITMSL's advantage is most pronounced in fast-flowing regimes (>70 m/yr), a pattern that cannot be explained by calibration alone. If calibration were the only factor, its effect would be expected to apply similarly across all velocity regimes. The pronounced improvement in fast-flowing areas instead reflects the improved physical representation afforded by the explicit inclusion of the sliding law.

6. Fig.6 : I'm not sure the colormap of panel d correspond to its associated colorbar. I'm not able to localize this green in this colorbar.

   Response: We thank the reviewer for pointing out this issue. Indeed, in the previous version of the manuscript, we inadvertently placed an incorrect colorbar for panel (d), which caused the colors, particularly the green areas, to be unidentifiable in the corresponding colorbar. In the revised manuscript, we have regenerated both panel (d) and its associated colorbar to ensure that the colour mapping is fully consistent. We have also carefully rechecked all subplots and their respective colorbars throughout the figure to prevent similar issues. We appreciate the reviewer's careful reading and constructive feedback.

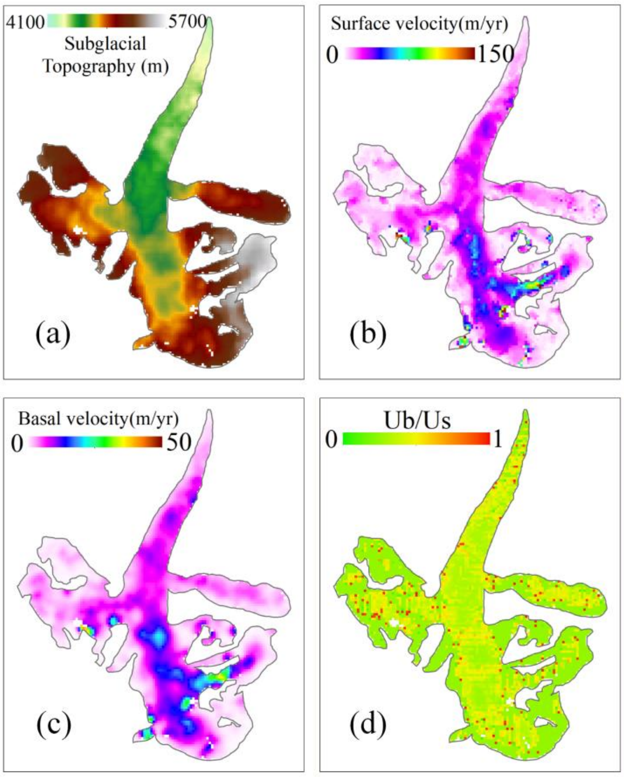

Figure 6: (a) CTSG glacier subglacial topography, (c) basal sliding, and (d) the ratio of sliding velocity to surface velocity inverted by ITMSL; (b) surface velocity used as input to ITMSL.

7. L.349 : "Gantayat's assumed 0.25, which induces substantial thickness overestimation." is not in contradiction with this previous sentence : "In terms of mean ice thickness, ITMSL (114.63 m), Millan (104.90 m), and Gantayat (98.39 m) differ from one another, with Gantayat being the lowest"

Response: We have revised this sentence (L356-365).

"The ratio ub/us ranges from 0.258 to 0.355, exceeding the value of 0.25 assumed by Gantayat et al. (2014). It is worth noting that, although the mean ice thickness estimated by our model is slightly higher than that of the other two models, it achieves a superior MAE. This indicates that our model captures the mechanical coupling between surface velocity and ice thickness more effectively and exhibits greater resilience to data noise, thereby substantially reducing extreme outliers. Nevertheless, the systematic overestimation in the mean estimates suggests that the current physical parameterization does not fully account for the diversity of subglacial hydrological conditions across the study area, leading to a general upward bias. Unlike the competing models, whose favorable mean performance stems from mutual error cancellation, our model produces a more structured error distribution, making it better suited for regional-scale ice-thickness inversion."

Table 3: Comparison of mean ice thickness from GPR measurements and retrievals by three models (Gantayat, Millan, ITMSL), along with their respective mean absolute errors (MAE), for the 16 glaciers.

Glacier	GPR (m)	Gantayat et al., 2014		Millan et al., 2022		ITMSL	
		Ice thickness (m)	MAE (m)	Ice thickness (m)	MAE (m)	Ice thickness (m)	MAE (m)
BARP	120.95	143.86	104.81	191.74	269.21	161.88	127.07
BATU	127.06	142.37	89.75	211.63	285.34	156.54	93.47
CHHU	103.93	113.63	45.55	142.55	132.14	124.33	66.47

GHAR	152.94	114.62	91.82	153.15	234.64	135.43	128.52
HXLG	43.96	83.91	62.57	42.60	24.46	82.98	61.61
HEIG	114.87	94.88	57.55	68.75	24.52	103.58	64.52
PASU	145.55	156.36	154.82	185.24	215.10	163.64	191.79
QIYI	77.81	63.58	24.58	53.26	23.67	62.29	18.10
SACH	121.79	123.40	48.54	185.03	153.84	128.42	47.26
SARY	75.52	77.80	39.99	81.28	48.44	75.69	31.94
SHQI	38.68	71.31	135.45	80.88	162.02	70.21	130.26
CTSG	172.88	104.90	47.54	98.39	48.62	110.70	36.52
SHIY	39.47	49.18	24.99	30.09	12.07	53.09	23.55
SIGH	68.93	71.99	55.89	42.28	26.30	70.94	44.49
TSTU	51.63	64.63	38.22	63.55	38.53	59.88	27.33
URUM	40.05	78.45	45.97	45.61	28.50	80.53	34.77

8. Table 3 : the caption should be reformulated. I do not understand it. Mean ice thickness and motion characteristics extracted at GPR point locations estimated from GPR, Gantayat, ... for the 16 glaciers.

Response: We agree that the original caption was ambiguous and failed to clearly convey the content of the table. We have reformulated the caption of Table 3 to concisely indicate what is presented in each column—namely, the GPR-measured mean ice thickness, the modelled mean ice thicknesses from the three models (Gantayat, Millan, ITMSL), and their corresponding mean absolute errors (MAE) for each of the 16 glaciers. The revised caption now accurately reflects the information shown in the table and avoids vague terms such as "motion characteristics". We have also adjusted the column headers accordingly to ensure consistency with the revised caption, so that readers can easily identify what each column represents. The revised caption now reads (Line 368-369):

"Table 3: Comparison of mean ice thickness from GPR measurements and retrievals by three models (Gantayat, Millan, ITMSL), along with their respective mean absolute errors (MAE), for the 16 glaciers."

9. Table 3 : what are the values of u_s displayed in the table ? Mean, Median, max ? From ITMSL ?

Response: We thank the reviewer for pointing out this ambiguity. The us/ub values displayed in Table 3 represent the glacier-wide mean ratio of basal sliding velocity (ub) to surface velocity (us) simulated by the ITMSL model for each glacier. The surface velocity data used for this ratio (from remote sensing observations) and the associated velocity statistics for each glacier are summarised in Table 1 for reference. In light of the reviewer's suggestion, and in conjunction with related comments from other reviewers on Table 3, we have thoroughly revised this table, including restructuring the layout, reducing the number of columns, and reformulating the caption to improve clarity and readability. Details of these revisions are provided in our response to Comment 7 and are reflected in the updated manuscript.

10. L356 : What are the GPR-validated thickness data ? measurement non used to calibrate ITMSL ?

Response: We thank the reviewer for raising this critical methodological question. We need to be transparent: in this study, all GPR-measured ice thickness data were used for calibrating (parameter optimisation of B) the ITMSL model. Consequently, the term "GPR-validated thickness data" in Line 356 was imprecise and may have misleadingly implied the existence of a fully independent validation dataset. This decision was made for the following reasons: (1) some glaciers have relatively limited GPR measurement points, and further splitting would have compromised the reliability of the parameter optimisation; (2) the primary focus of this study is the physical improvement of incorporating a sliding law within the SIA framework, rather than developing a purely data-driven prediction model—therefore, using all data for calibration serves the purpose of obtaining optimally constrained parameters under physically reasonable assumptions, rather than avoiding overfitting. To eliminate ambiguity, we have revised the wording in the manuscript, replacing "GPR-validated" with "GPR-measured" (L371-372).

"Comprehensive evaluation of the ITMSL model against Gantayat, Millan using 5 statistical metrics applied to GPR-measured thickness data (Figure 7) demonstrates significant improvements."

11. L362-363 : "exhibit higher deviation in ice thickness inversion using ITMSL compared to Millan et al., 2022, and the changes in mean and relative errors are limited, the improvement in accuracy still confirms the […]" . I do not understand "and the changes... limited". I would cut the sentence before it and add between parenthesis "see RMSE, ME and RE in fig7" . Improvement with respect to what ?

Response: We thank the reviewer for the careful reading and constructive suggestion. We agree that the phrase "and the changes in mean and relative errors are limited" was ambiguous, as the changes were not clearly defined. We have removed this phrase entirely, following your suggestion, and added a parenthetical reference ("see RMSE, ME, and RE in Fig. 7") to direct readers to the relevant accuracy metrics. We have also explicitly clarified that the "improvement" refers to a comparison with both Millan et al. (2022) and Gantayat et al. (2014), rather than leaving the reference unspecified. The revised sentence now reads (L378-382):

"Although 4 glaciers (SHIY, HXLG, URUM, HEIG) exhibit higher deviation in ice thickness inversion using ITMSL compared to Millan et al. (2022) (see RMSE, ME, and RE in Figure 7), the overall accuracy of ITMSL remains superior to that of both Millan et al. (2022) and Gantayat et al. (2014) across the majority of the 16 glaciers, confirming its reliability."

12. L370 : Both ME and RE → Positive values of ME and RE ?

Response: As defined in the Methods section (Section 2.3), both Mean Error (ME) and Relative Error (RE) are calculated as GPR-measured minus model-simulated values, i.e., ME = (GPR_measured - Modelled), and RE = (GPR_measured - Modelled) / GPR_measured × 100%. Therefore, positive values indicate that the modelled ice thickness is lower than the GPR-measured thickness (i.e., underestimation); negative values indicate that the modelled ice thickness is higher than the GPR-measured thickness (i.e., overestimation). To avoid any confusion, we have revised the relevant text in the manuscript to explicitly state the sign convention of ME and RE, rather than presenting them ambiguously as "Both ME and RE" (Line 390-392).

"In this study, both Mean Error (ME) and Relative Error (RE) are defined as "GPR-measured minus model-simulated values". A positive value thus indicates that the model underestimates the observed ice thickness, whereas a negative value indicates overestimation."

5) Discussion

1. L421-427 : I would reformulate here since you binned the ratio distribution in 10 bins of 0.1. From this bin analysis, 2 regimes can be defined. The bias is reduced with Millan in the 0.1-0.4 regime (median closer to 0) but even in this regime, the interquartile range and the outliers of ITMSL are better than those of other models.

Response: We agree that, after binning the ratio distribution into 10 intervals of 0.1, two distinct regimes can indeed be identified. Following the reviewer's suggestion, we have reformulated the analysis, with the revised key conclusion as follows (L467-471):

"The comparative analysis at these GPR points (Figure 10) reveals distinct performance regimes:

for the 0.1-0.4 bin, ITMSL outperformed other models in terms of quartiles and outliers; for the 0.4-1.0 bin, ITMSL reduces the absolute deviation relative to both Gantayat and Millan. Crucially, the overall ice thickness deviation of ITMSL was 13.35% lower than that of Gantayat and 0.2% lower than that of Millan."

2. Fig.10 : colorbar does not use the correct colormap. ylabel miss a "space" between ice and thickness.

Response: We made the following modifications to the y-axis and the legend representing point density (L483-485):

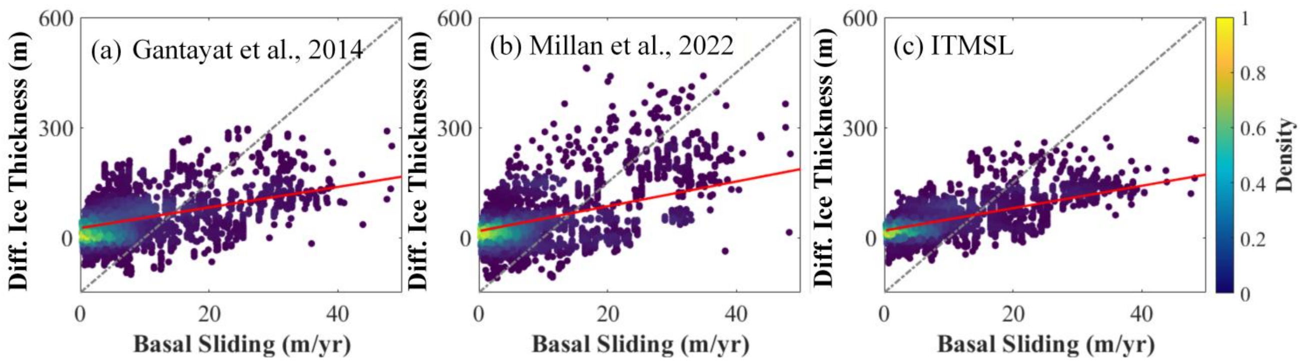

Figure 11: Deviations of ice thickness estimates from three laminar flow models at different basal sliding. (a) Gantayat; (b) Millan; (c) ITMSL.

3. L.445 : We compared → We compare now ?

Response: We thank the reviewer for the careful attention to tense usage. The original sentence used "we compared" because it describes the completed comparison analysis in the methodology section. However, we agree that the sentence also introduces the overall framework of this study. Following the reviewer's suggestion, we have revised the tense to **"are compared"** to better reflect the study's broader purpose. The revision has been made in the revised manuscript (Line 488-491).

"To place the ITMSL, Gantayat, and Millan models within a broader glaciological model framework, their thickness estimates are compared with four established models, namely Composite, GlabTop2, H-F, and OGGM (Huss and Farinotti, 2012; Frey et al., 2014; Farinotti et al., 2019; Maussion et al., 2019)."

4. L463-464 : What is the value of ice thickness displayed in fig 12 ? why only one value per glacier. I'm not sur comparing only the mean ice thickness is a good indicator. Again you should refer to what ITMIX has done so far.

Response: The values displayed in Fig. 12 represent the mean ice thickness for each glacier (the glacier-wide average from each model's inversion results). Our original intention was to use this as a straightforward summary metric; however, we agree with the reviewer that comparing only mean ice thickness may obscure important structural information about error distributions and is not the most informative indicator of model performance.

Taking into account related comments from other reviewers, we have carefully considered this issue and decided to remove Fig. 12 from the revised manuscript for the following reasons: (1) the glacier-mean comparisons in this figure largely overlap with the accuracy metrics (RMSE, MAE, etc.) already presented in Fig. 13, which offer a more informative characterisation of model performance; (2) retaining two partially overlapping figures would make the paper unnecessarily redundant and may distract readers from the core findings.

Regarding the reviewer's suggestion to "refer to what ITMIX has done so far," we have revisited the ITMIX study (Farinotti et al., 2017), which similarly prioritises error distributions (e.g., boxplots, cumulative distribution functions) over simple mean values in model intercomparisons. We follow this principle: accordingly, while removing Fig. 12, we retain Fig. 13 to more comprehensively capture the differences in error structure across models.

5. Fig.12 : Caption : mean, median, max ice thickness per glacier ?

   Response: This figure presents the mean ice thickness for each glacier. In response to the suggestions from other reviewers, we have removed it from the revised manuscript, as its content largely overlaps with that of Figure 13.

6. L.484 : All ME but one are negatives in fig.12. How can you say that :" The ME quantification confirms a systematic positive bias" ?

   Response: As defined in the Methods section (Section 2.3), both Mean Error (ME) and Relative Error (RE) are calculated as GPR-measured minus model-simulated values, i.e., ME = (GPR_measured - Modelled), and RE = (GPR_measured - Modelled) / GPR_measured × 100%. Therefore, positive values indicate that the modelled ice thickness is lower than the GPR-measured thickness (i.e., underestimation); negative values indicate that the modelled ice thickness is higher than the GPR-measured thickness (i.e., overestimation). To avoid any confusion, we have revised the relevant text in the manuscript to explicitly state the sign convention of ME and RE, rather than presenting them ambiguously as "Both ME and RE" (Line 390-392).

   "In this study, both Mean Error (ME) and Relative Error (RE) are defined as "GPR-measured minus model-simulated values". A positive value thus indicates that the model underestimates the observed ice thickness, whereas a negative value indicates overestimation."

7. Fig.13 : Space between Ice and thickness in xlabel.

   Response: We have revised this figure accordingly (L533-534).

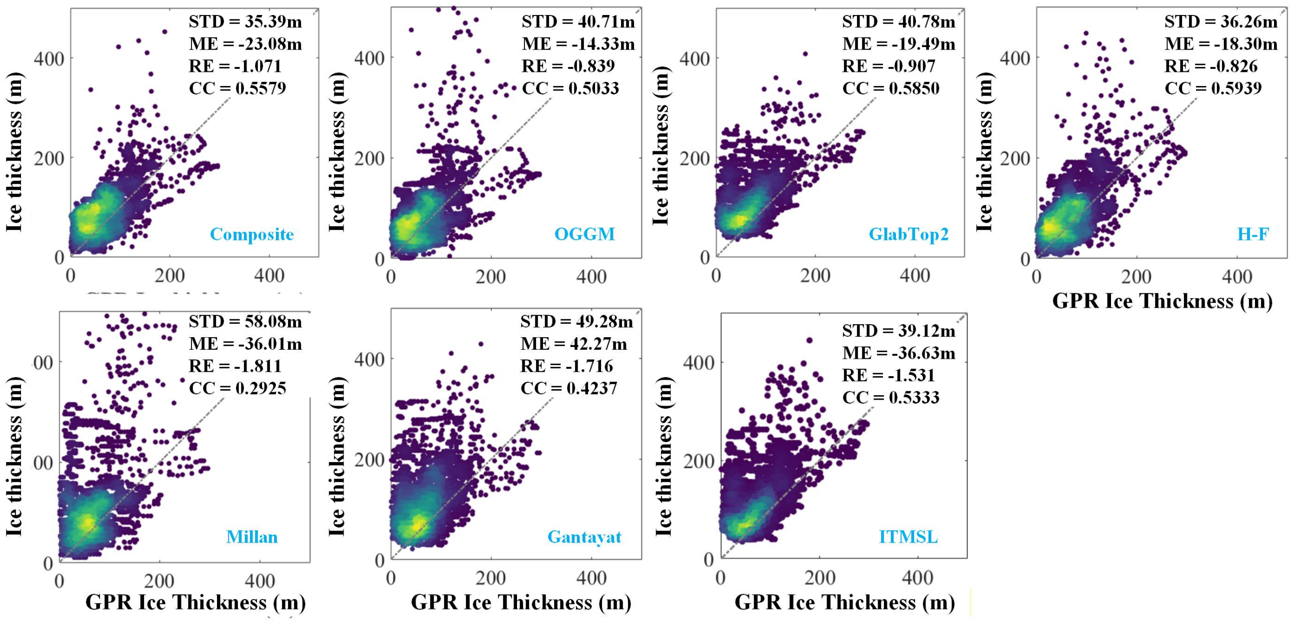

Figure 13: Evaluation of model-inverted ice thickness accuracy using GPR-measured ice thickness.

8. What is the difference between the GPR ice thickness measurement subtracted to the DEM and the GPR derived bed measurement ? This should be explained tounderstand the difference between fig 12 and 13. I repeat this question from my previous review since I'm still not able to understand why fig.12 and fig.13 are different. Inverted bed = DEM – inverted thickness. GPR bed = DEM – GPR thickness. In that case the scatter plot between inverted and GPR quantity should be the same for bed and for thickness.

Response: We thank the reviewer for revisiting this question. We understand the reviewer's logic: if both the inverted bed and the GPR bed were derived from the same DEM and the same thickness data, then the two scatter plots would indeed be mirror images. However, in this study, the two datasets differ fundamentally in their data sources and derivation:

The inverted bed is derived by subtracting the ITMSL-inverted ice thickness from a remote-sensing DEM (GLO-30). It is a gridded dataset covering the entire glacier. The GPR bed is obtained directly from the surface elevation acquired simultaneously with the GPR thickness survey minus the GPR-measured ice thickness. This surface elevation is a field measurement and is not derived from the remote-sensing DEM. The GPR bed is therefore an independent in-situ measurement.

Because of this fundamental difference, the inverted-bed vs. GPR-bed comparison (Figure 14) reflects the agreement between the model output and an independent field dataset, whereas the inverted thickness vs. GPR thickness comparison (Figure 13) directly quantifies the model's thickness prediction accuracy. The two figures provide complementary information: Figure 14 validates the spatial pattern of the ice bed, while Figure 13 quantifies the magnitude of prediction errors at measurement points. They are not redundant. We have added a clarification in the revised manuscript to help readers understand the distinct data sources and purposes of the two figures (Line 536-540).

"There is a discrepancy between the two: the former is the bed elevation indirectly derived from the inverted ice thickness, while the latter is obtained by subtracting the measured ice thickness from the surface elevation that was concurrently recorded during the thickness survey, yielding the bed elevation at each measurement point."

6) Conclusion

/

Reference :

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, https://doi.org/10.1029/2023GL105029

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

Frank, T., van Pelt, W.J.J., Rounce, D.R. *et al.* Global glacier-free topography reveals a large potential for future lakes in presently ice-covered terrain. *Nat Commun* **17**, 3985 (2026). https://doi.org/10.1038/s41467-026-72548-9