Response to Reviewer 3

We sincerely thank the reviewer for the careful assessment of our manuscript and for the constructive and insightful comments. We have revised the manuscript and the associated dataset documentation to improve the transparency, reproducibility, and robustness of the 2mKGI inventory and its comparison with previous glacier inventories.
Major revisions include clarifying the spatial coverage and characteristics of the ZY-3 and Sentinel-2 mapping subsets, refining the criteria for debris-covered glacier delineation, reconsidering the influence of surge-type glaciers and acquisition-date differences, improving the DEM comparison, and clarifying the training–validation–test workflow. We have also expanded the methodological documentation and made the training/inference code and training labels publicly available. Detailed responses to each comment are provided below. The reviewer's original comments are shown in blue and our responses in red..

Major Comment 1

At L305-307 the authors state that approximately 21.5% of the glacier-outline coverage was derived from Sentinel-2 imagery in areas without ZY-3 coverage, with the remainder mapped from ZY-3. Please provide a map showing the spatial coverage of each source. This would allow the reader to assess how outline quality varies between the two, and I would encourage the authors to report glacier number, size distribution, and debris fraction separately for the ZY-3- and Sentinel-2-mapped areas.

Response

We thank the reviewer for this constructive suggestion. We have revised the study-area map (Fig. 1) to explicitly distinguish the spatial coverage of the ZY-3 and Sentinel-2 imagery and have added source-stratified inventory statistics. Of the 13,272 glaciers in the revised inventory, 7,121 (53.7%) were mapped from ZY-3 imagery and 6,151 (46.3%) from Sentinel-2 imagery. The corresponding glacierized areas are 15,853.83 km² (81.0% of the total inventory area) and 3,728.46 km² (19.0%), respectively.
The glaciers mapped from Sentinel-2 imagery are generally smaller, with a median area of 0.082 km², compared with 0.169 km² for those mapped from ZY-3 imagery. Debris-covered ice accounts for 1,905.14 km² (12.02%) of glacier area in the ZY-3 subset and 150.83 km² (4.05%) in the Sentinel-2 subset. In terms of glacier number, 1,068 glaciers in the ZY-3 subset and 323 glaciers in the Sentinel-2 subset contain mapped supraglacial debris. These source-stratified statistics and the corresponding interpretation have now been added to the revised manuscript.

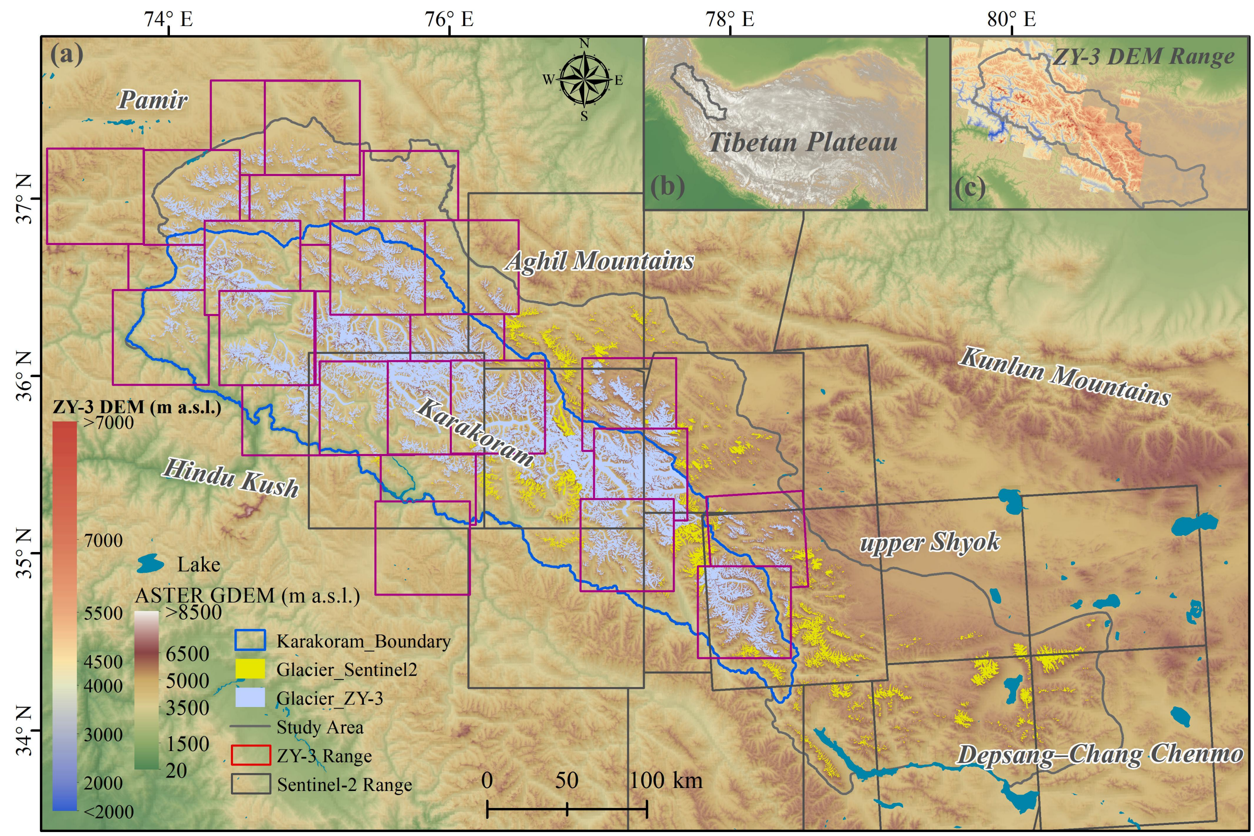

Figure 1. Study area.

Major Comment 2

At L445-448 the authors argue that RGI v7.0 and KGI-2020s overgeneralise debris-covered termini. However, 2mKGI excludes stagnant debris and non-glacial debris-covered surfaces, whereas RGI and KGI-2020s follow the convention of including debris-covered ice that remains connected to the glacier. Under these two criteria, 2mKGI will be systematically smaller than the other inventories, so the difference at debris-covered termini may reflect the mapping criterion rather than an error in the earlier products. Please state explicitly how debris-covered ice was identified and removed in 2mKGI, and what evidence was used to decide that a given debris surface is no longer ice-covered.

Response

This is an important distinction. We agree that differences at debris-covered termini can arise not only from mapping error, but also from differences in delineation criteria among glacier inventories. In the revised 2mKGI, debris-covered ice was identified using multiple geomorphological and topographic indicators rather than surface debris appearance alone. These included the presence and continuity of ice cliffs, glacier meltwater outlets, lateral and terminal moraines, glacier-like surface morphology, and local slope or topographic breaks derived from the ZY-3 DEM. The example figure added to the revised manuscript illustrates how these features were used jointly to support debris-covered glacier-boundary interpretation (Fig.2 in this Document).

Importantly, stagnant or debris-covered surfaces were not excluded simply because they appeared inactive. A debris-covered area was removed from the glacier extent only when the available evidence indicated that it was more consistent with non-glacial terrain or moraine/forefield material than with debris-covered glacier ice. Where the presence of underlying ice could not be determined confidently from surface morphology alone, the boundary was treated cautiously rather than being removed solely on the basis of apparent stagnation. Multi-temporal

Sentinel-2 imagery was also used where temporal information helped distinguish persistent glacier-related morphology from surrounding non-glacial debris surfaces.

We have therefore revised the manuscript to state explicitly that part of the difference between 2mKGI and earlier inventories may reflect differences in mapping criteria at debris-covered termini, in addition to genuine boundary-interpretation differences.

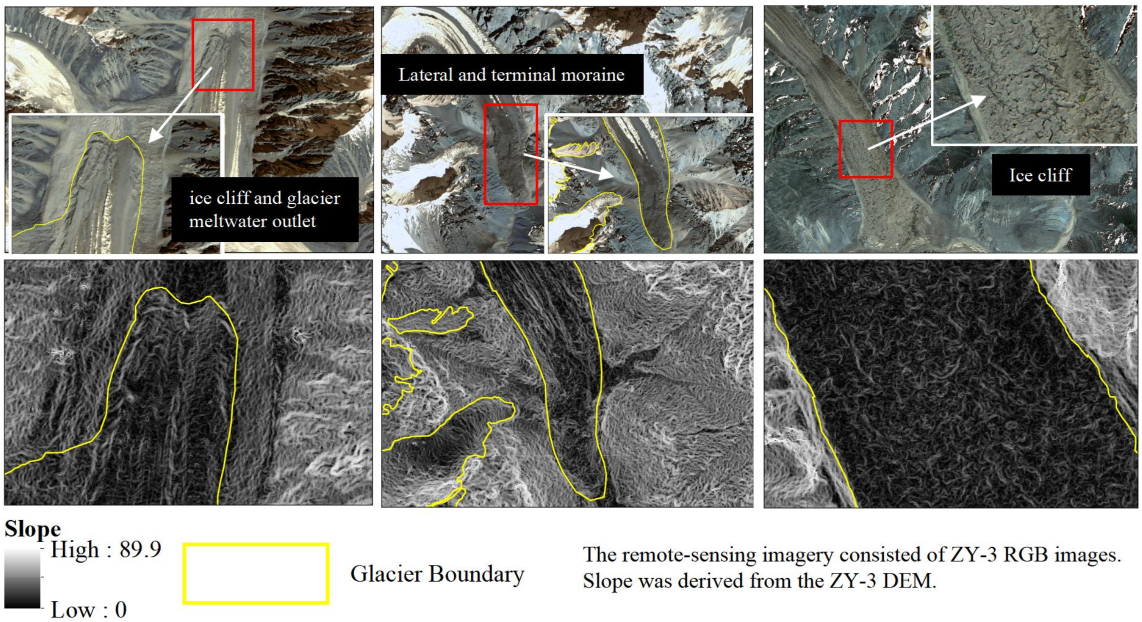

Figure 2. Revision criteria for the boundaries of supraglacial debris.

Major Comment 3

The comparison between inventories spans a wide range of data acquisition periods. The authors identify 192 surge-type glaciers in 2mKGI, and rapid terminus advance during surge events would produce area differences of similar magnitude to the "credible differences" described at L474. Comparisons with other inventories should therefore exclude surge-type glaciers, or at least report the results with and without them.

Response

We accept this concern. Surge-related advance or retreat can produce substantial local area differences when inventories from different acquisition periods are compared. However, the comparisons in this study are not primarily based on glacier-by-glacier area differences. The comparison with KGI-2020s involves inventories representing broadly similar conditions around 2020, which substantially reduces the temporal mismatch. The comparison with RGI v7.0 is conducted mainly at the grid-cell scale, so the influence of individual surge events is spatially aggregated rather than interpreted as a direct glacier-level mapping discrepancy. We have also added a comparison with the recent Pakistan glacier inventory derived mainly from Sentinel-2 imagery acquired in 2022, providing another temporally close reference. We have therefore revised the manuscript to clarify that local differences may still include genuine surge-related geometric changes, particularly where acquisition dates differ, and these should not be interpreted solely as mapping discrepancies.

Major Comment 4

In Section 5.2, it is not surprising that the 4 m ZY-3 DEM resolves higher roughness and more short-wavelength variability than the 30 m ASTER DEM. This is largely caused by the grid spacing and is not by itself evidence of a data-quality difference. I recommend aggregating the ZY-3 DEM to 30 m and repeating the comparison at matched resolution. Any residual difference can then be attributed to ASTER error.

Response

We thank the reviewer for this constructive suggestion. We agree that the native-resolution comparison confounded differences in grid spacing with differences between the DEM products. We therefore aggregated the original 4 m ZY-3 DEM to 30 m by area averaging and bilinearly aligned ASTER GDEM V3 to exactly the same projected grid. Identical terrain operators were then applied to 1,796 glacier units ≥1 km² with at least 95% valid coverage in both DEMs. After resolution matching, the median paired difference (ZY-3 minus ASTER) was −17.1 m for glacier mean elevation, +1.06° for mean slope and +1.39 m (11.9%) for the 90 m local-elevation standard-deviation metric. The 90 m local relief was also 4.46 m (12.2%) higher in ZY-3. Thus, matching the grid spacing removed the direct resolution confound, but measurable residual differences in elevation, slope and short-wavelength terrain variability remained.

Section 5.2 has been revised accordingly. In the absence of an independent reference DEM, we interpret the remaining discrepancies as differences between the two DEM products, which may arise from their effective spatial resolution, acquisition dates, co-registration, vertical reference systems, and inherent uncertainties.

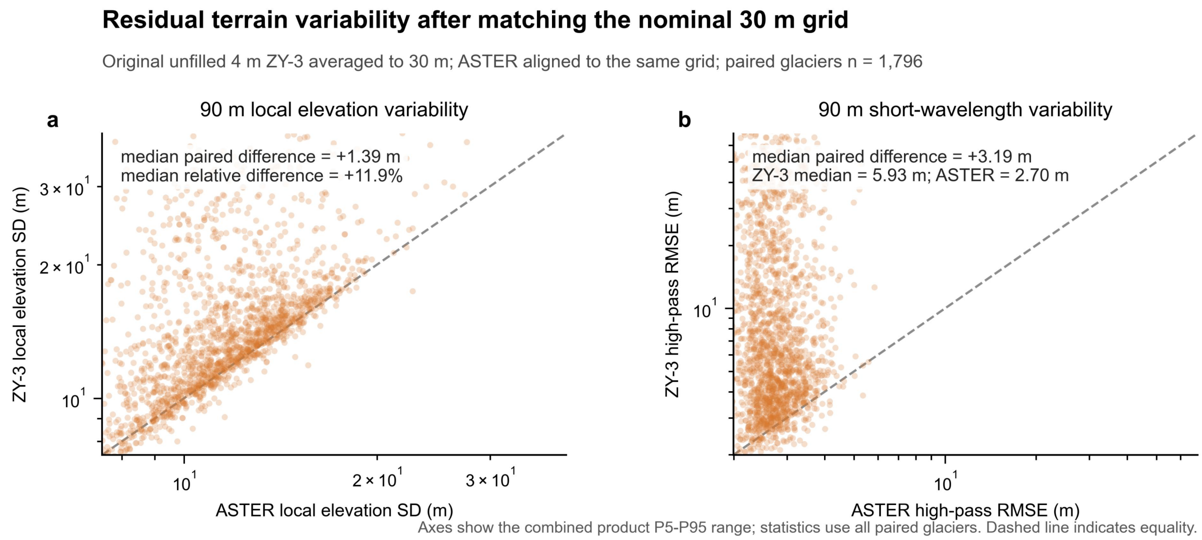

Figure3. Matched-resolution DEM comparison

Major Comment 5

L215-217. The text implies that the training, validation, and test subsets were split after data augmentation, which would place augmented copies of the same source tile on both sides of the split and lead to data leakage and an optimistic evaluation. Please clarify the order of these steps and describe the split protocol in detail. I would also encourage a spatially blocked split rather than a random one, since adjacent tiles from the same scene are strongly correlated. In addition, please confirm that the 4,595 km² test set is independent of the four scenes used for label generation, and clarify its relationship to the 10% test subset described here.

Response

Thank you for identifying this ambiguity. The original description of the workflow was imprecise and could indeed be interpreted as if data augmentation had been performed before the training–validation–test split. This was not the case. The original image tiles were first assigned to the training, validation, and test subsets, and augmentation was subsequently performed within each subset. Augmented versions of a given source tile therefore remained in the same subset as the original tile, and no augmented copy of the same source tile was shared across the training, validation, and test subsets. We have revised the Methods section to make this order explicit.

We nevertheless acknowledge the reviewer's concern that a random tile-level split does not provide a fully spatially independent evaluation, because neighbouring tiles from the same scene can remain strongly correlated even in the absence of direct data leakage. We have therefore revised the manuscript to avoid interpreting the performance obtained from the 10% test subset as evidence of spatial transferability. These metrics are now presented as an internal evaluation of model performance against manually labelled reference data. The deep-learning model is used primarily as an auxiliary tool for generating initial glacier predictions for inventory production, rather than as the principal methodological contribution of this study; the final glacier outlines were subsequently subjected to systematic manual inspection and correction.

We also clarify that the 4,595 km² assessment area is distinct from the 10% test subset described in the model-development workflow. The 10% test subset was drawn from the manually labelled dataset and contains pre-existing reference labels; it was used after model training to calculate the standard confusion-matrix-based performance metrics. In contrast, the 4,595 km² area is geographically separate from the four scenes used for training-label generation and was not included in the training, validation, or 10% test subsets. It was initially mapped without pre-existing reference labels as part of the regional application of the trained model.

For the 4,595 km² area, the model output was subsequently inspected and manually revised to produce the final glacier boundaries. The reported comparison for this area therefore represents agreement between the initial model extraction and the final manually revised mapping, rather than an independent pre-labelled test in the same sense as the 10% held-out subset. We have revised the terminology in the manuscript accordingly to clearly distinguish the internal test subset from the regional assessment area.

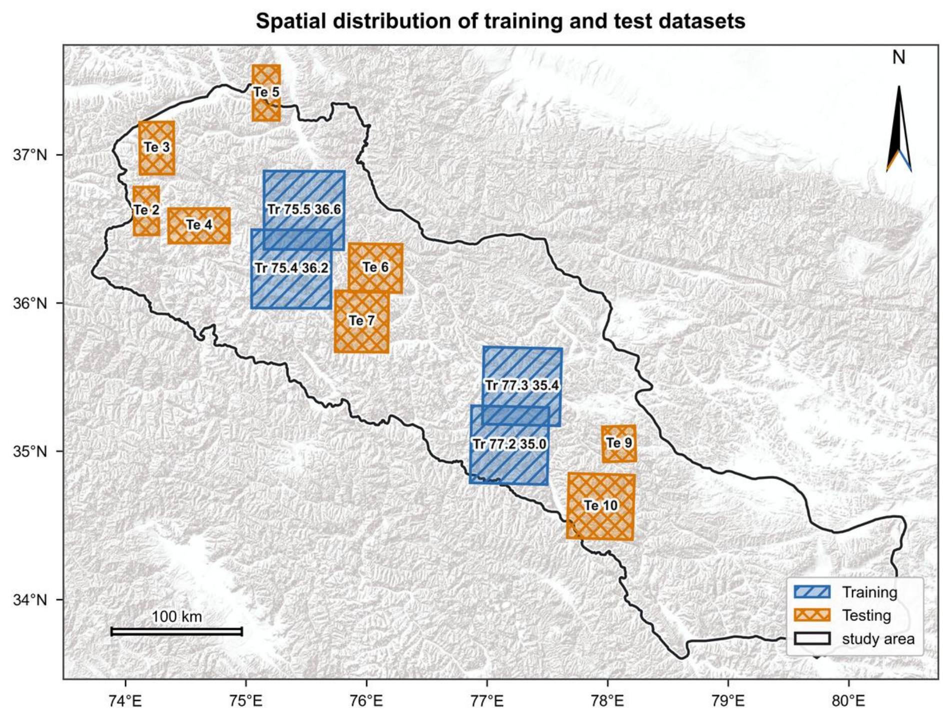

Figure 4. Training and Testing Areas.

Specific Comment 6

L120-125. Please give the exact number of Sentinel-2 images used rather than "approximately 10", and provide a table listing all ZY-3, Sentinel-2, and Landsat scene IDs with their acquisition dates.

Response

Thank you for this suggestion. We have corrected the previous description of “approximately 10” Sentinel-2 images. A total of 12 Sentinel-2 Level-2A granules were used, acquired between 7 August and 24 September 2020. We have also compiled the complete acquisition information for all source imagery used in the study, including 36 ZY-3 scenes, 12 Sentinel-2 granules, and 12 Landsat 8 OLI/TIRS Collection 2 Level-2 scenes, together with their corresponding scene IDs and acquisition dates. The complete list is provided in this response document and will also be included as a supplementary table in the revised manuscript to ensure full traceability of the source imagery.

	acquisition dates	IDs
	2020-06-23	ZY302_TMS_E75.7_N35.0_20200623_L1A0000995617
	2020-06-23	ZY302_TMS_E75.8_N35.4_20200623_L1A0000995614
	2020-08-16	ZY302_TMS_E75.3_N35.8_20200816_L1A0000995613
	2020-08-16	ZY302_TMS_E75.4_N36.2_20200816_L1A0000995612
	2020-08-16	ZY302_TMS_E75.5_N36.6_20200816_L1A0000995577
	2020-08-16	ZY302_TMS_E75.6_N37.0_20200816_L1A0000995611
	2020-08-16	ZY302_TMS_E75.7_N37.4_20200816_L1A0000995562
	2020-09-10	ZY302_TMS_E77.2_N35.0_20200910_L1A0000995568
	2020-09-10	ZY302_TMS_E77.3_N35.4_20200910_L1A0000995573
	2020-09-10	ZY302_TMS_E77.4_N35.8_20200910_L1A0000995576
	2020-09-19	ZY302_TMS_E73.6_N36.2_20200919_L1A0000995574
	2020-09-20	ZY302_TMS_E77.9_N34.3_20200920_L1A0000731295
	2020-09-20	ZY302_TMS_E78.0_N34.6_20200920_L1A0000731267
	2020-09-20	ZY302_TMS_E78.1_N35.0_20200920_L1A0000731332
ZY-3	2020-09-20	ZY302_TMS_E78.2_N35.4_20200920_L1A0000731317
	2020-09-20	ZY302_TMS_E78.4_N36.2_20200920_L1A0000995570
	2020-09-24	ZY302_TMS_E74.1_N36.6_20200924_L1A0000995567
	2021-07-16	ZY302_TMS_E74.6_N36.6_20210716_L1A0000866396
	2021-07-26	ZY302_TMS_E74.6_N36.2_20210726_L1A0000871241
	2021-07-26	ZY302_TMS_E74.8_N37.0_20210726_L1A0000871285
	2021-08-10	ZY302_TMS_E75.9_N35.8_20210810_L1A0000877185
	2021-08-20	ZY302_TMS_E77.2_N35.8_20210820_L1A0000882427
	2021-08-29	ZY302_TMS_E73.5_N37.0_20210829_L1A0000885652
	2021-09-03	ZY302_TMS_E73.8_N37.0_20210903_L1A0000887742
	2021-09-08	ZY302_TMS_E74.0_N36.2_20210908_L1A0000889787
	2021-09-08	ZY302_TMS_E74.2_N37.0_20210908_L1A0000889762
	2021-09-13	ZY302_TMS_E74.6_N37.4_20210913_L1A0000891503
	2021-09-18	ZY302_TMS_E74.7_N36.2_20210918_L1A0000894516
	2021-09-18	ZY302_TMS_E74.8_N36.6_20210918_L1A0000894503

	2021-09-18	ZY302_TMS_E74.9_N37.0_20210918_L1A0000894460
	2021-09-18	ZY302_TMS_E75.0_N37.4_20210918_L1A0000894465
	2021-09-23	ZY302_TMS_E75.3_N37.0_20210923_L1A0000896518
	2021-09-28	ZY302_TMS_E75.4_N35.8_20210928_L1A0000898630
	2021-09-28	ZY302_TMS_E75.5_N36.2_20210928_L1A0000898629
	2021-09-28	ZY302_TMS_E75.6_N36.6_20210928_L1A0000898604
	2021-09-28	ZY302_TMS_E75.7_N37.0_20210928_L1A0000898616
Sentinel2	2020-08-07	S2B_MSIL2A_20200807T052649_N0214_R105_T43SGU_20200807T110831
	2020-08-07	S2B_MSIL2A_20200807T052649_N0214_R105_T43SGV_20200807T110831
	2020-08-07	S2B_MSIL2A_20200807T052649_N0214_R105_T44SKE_20200807T110831
	2020-08-25	S2A_MSIL2A_20200825T053651_N0214_R005_T43SFA_20200825T074030
	2020-08-25	S2A_MSIL2A_20200825T053651_N0214_R005_T43SFV_20200825T080607
	2020-08-25	S2A_MSIL2A_20200825T053651_N0214_R005_T43SGA_20200825T074030
	2020-09-01	S2A_MSIL2A_20200901T052651_N0214_R105_T43SGT_20200901T081543
	2020-09-01	S2A_MSIL2A_20200901T052651_N0214_R105_T44SKC_20200901T081543
	2020-09-01	S2A_MSIL2A_20200901T052651_N0214_R105_T44SKD_20200901T081543
	2020-09-01	S2A_MSIL2A_20200901T052651_N0214_R105_T44SLD_20200901T081543
	2020-09-03	S2B_MSIL2A_20200903T051659_N0214_R062_T44SMD_20200903T074638
	2020-09-24	S2A_MSIL2A_20200924T053651_N0214_R005_T43SFU_20200924T081357
Landsat 8 OLI/TIRS Collection 2 Level-2	2020-09-10	LC08_L2SP_149034_20200910_20200919_02_T1
	2020-09-28	LC08_L2SP_147035_20200928_20201006_02_T1
	2020-10-07	LC08_L2SP_146036_20201007_20220526_02_T1
	2020-10-07	LC08_L2SP_146037_20201007_20220526_02_T1
	2020-10-12	LC08_L2SP_149035_20201012_20201104_02_T1
	2020-10-14	LC08_L2SP_147036_20201014_20201105_02_T1
	2020-10-16	LC08_L2SP_145036_20201016_20201105_02_T1
	2020-10-16	LC08_L2SP_145037_20201016_20201105_02_T1
	2020-10-19	LC08_L2SP_150034_20201019_20201105_02_T1
	2020-10-19	LC08_L2SP_150035_20201019_20201105_02_T1
	2020-10-21	LC08_L2SP_148035_20201021_20201106_02_T1
	2020-10-21	LC08_L2SP_148036_20201021_20201106_02_T1

Specific Comment 7

L196-199. Please state the ATL06 date range used for validation, confirm whether the DEM was co-registered before the assessment, and define the slope thresholds separating the flat, hilly, mountainous, and alpine terrain classes.

Response

Thank you for this helpful comment. We have revised the manuscript to clarify that the ICESat-2 ATL06 observations used for DEM validation span from 1 July to 28 September 2021. The ZY-3 DEM was georeferenced during the stereo-processing workflow and aligned to the common projected coordinate system. No additional horizontal co-registration or vertical bias correction was performed using ATL06 prior to the accuracy assessment, ensuring that the ATL06 observations remained an independent validation dataset.
Terrain classes were assigned according to slope derived from the 4 m ZY-3 DEM. We adopted the slope thresholds used in GB/T 12343.1-2008, with flat, hilly, mountainous, and alpine terrain defined as <2°, 2–6°, 6–25°, and >25°, respectively. These details have now been added to the revised manuscript.

Specific Comment 8

L397. Please give the source for the known surge-type glaciers.

Response

Thank you for pointing this out. The surge-type glaciers used in this study were identified with reference to previously published regional inventories, primarily Bhambri et al. (2017), Guillet et al. (2022), and Guo et al. (2023). We have now added these sources to the revised manuscript and clarified the provenance of the surge-type classification.

References:

Bhambri, R., Hewitt, K., Kawishwar, P., and Pratap, B. (2017). Surge-type and surge-modified glaciers in the Karakoram. Scientific Reports, 7, 15391.

Guillet, G., King, O., Lv, M., Ghuffar, S., Benn, D., Quincey, D., and Bolch, T. (2022). A regionally resolved inventory of High Mountain Asia surge-type glaciers, derived from a multi-factor remote sensing approach. The Cryosphere, 16, 603–623.

Guo, L., Li, J., Dehecq, A., Li, Z., Li, X., and Zhu, J. (2023). A new inventory of High Mountain Asia surging glaciers derived from multiple elevation datasets since the 1970s. Earth System Science Data, 15, 2841–2861.

Specific Comment 9

Table 1. Please give the resolution of KGI-2020s and RGI v7.0. The date of "2000" for RGI v7.0 is also inaccurate, as the source imagery for this region spans several years.

Response

Thank you for pointing this out. We have revised Table 1 to provide more accurate information on the source-image resolution and acquisition periods of the two reference inventories. KGI-2020s was derived primarily from 30 m Landsat imagery (Xie et al., 2023). For the Karakoram, which lies within RGI Region 14 (South Asia West), the RGI v7.0 outlines are based on the GAMDAM Glacier Inventory v2 (RGI Consortium, 2023; Sakai, 2019), which was compiled from 30 m Landsat TM/ETM+ imagery acquired between 1990 and 2010. Most of the source imagery used for this region dates from around 2002, with earlier or later scenes used where suitable imagery was unavailable. We have therefore revised Table 1 to list the source-image resolution of RGI v7.0 as 30 m and replaced the previous single date of "2000" with "1990–2010 (mostly 2002)".
References:

Consortium, R. G. I.: Randolph Glacier Inventory - A Dataset of Global Glacier Outlines, Version 7. NASA National Snow and Ice Data Center Distributed Active Archive Center, 2023.

Sakai, A.: Brief communication: Updated GAMDAM glacier inventory over high-mountain Asia, The Cryosphere, 13, 2043-2049, 2019.

Xie, F., Liu, S., Gao, Y., Zhu, Y., Bolch, T., Kääb, A., Duan, S., Miao, W., Kang, J., Zhang, Y., Pan, X., Qin, C., Wu, K., Qi, M., Zhang, X., Yi, Y., Han, F., Yao, X., Liu, Q., Wang, X., Jiang, Z., Shangguan, D., Zhang, Y., Grünwald, R., Adnan, M., Karki, J., and Saifullah, M.: Interdecadal glacier inventories in the Karakoram since the 1990s, Earth Syst. Sci. Data, 15, 847-867, 2023.

Specific Comment 10

Figure 1. Please give the source of the Karakoram boundary polygon.

Response

Thank you for pointing this out. We have revised Fig. 1 to clarify the sources of the geographic boundaries used in this study. The Karakoram boundary polygon has been updated following the regional extent defined by Bhambri et al. (2022), which we now use to represent the Karakoram proper. The broader study-area extent, which includes the Karakoram and adjacent high-mountain regions, follows the regional framework defined by Xie et al. We have added the corresponding references to both the figure caption and the revised study-area description to make this distinction explicit.

References:

Bhambri, R., Chand, P., Nüsser, M., Kawishwar, P., Kumar, A., Gupta, A. K., Verma, A., and Tiwari, S. K.: Reassessing the Karakoram Through Historical Archives, in: Environmental Change in South Asia: Essays in Honor of Mohammed Taher, edited by Saikia, A. and Thapa, P., Springer International Publishing, Cham, 139–169, https://doi.org/10.1007/978-3-030-47660-1_8, 2022.

Xie, F., Liu, S., Gao, Y., Zhu, Y., Bolch, T., Kääb, A., Duan, S., Miao, W., Kang, J., Zhang, Y., Pan, X., Qin, C., Wu, K., Qi, M., Zhang, X., Yi, Y., Han, F., Yao, X., Liu, Q., Wang, X., Jiang, Z., Shangguan, D., Zhang, Y., Grünwald, R., Adnan, M., Karki, J., and Saifullah, M.: Interdecadal glacier inventories in the Karakoram since the 1990s, Earth System Science Data, 15, 847–867, https://doi.org/10.5194/essd-15-847-2023, 2023

Specific Comment 11

Figure 5. The red ridgelines derived from the ZY-3 DEM are difficult to distinguish from the background DEM. I recommend a different colour.

Response

Thank you for this suggestion. We have revised Fig. 5 by changing the colour of the ZY-3 DEM-derived ridgelines to improve their contrast against the DEM background and make the drainage divides easier to distinguish.

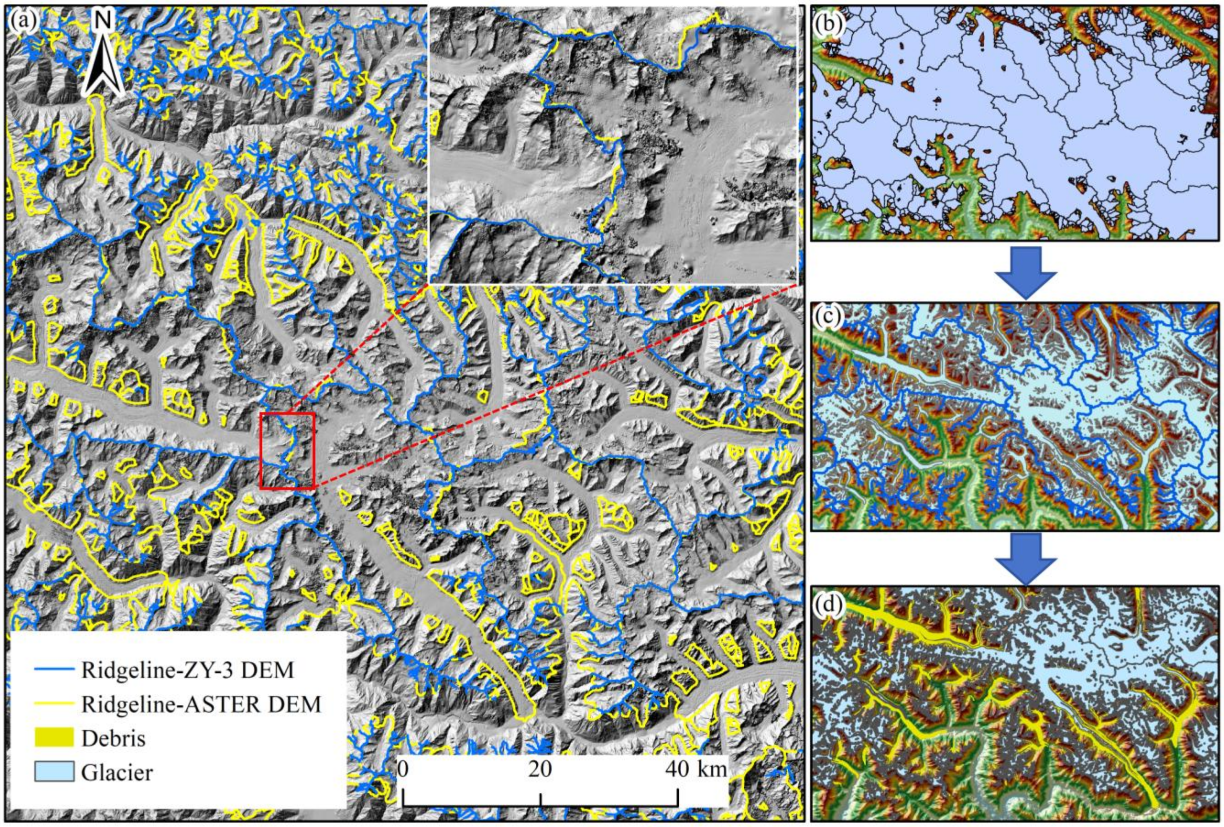

Figure5. Ridgelines derived from the ZY-3 DEM and ASTER GDEM and used to partition glacier units.

Specific Comment 12

Please release the training and inference code in a public repository and include the code link in the data availability section, together with the training labels, so that the study can be reproduced.

Response

We agree that access to both the implementation and reference labels is important for reproducibility. We have therefore created a public GitHub repository containing the training and inference scripts for the U-Net+CBAM workflow, together with instructions for input-data preparation, example input data for running the model, and reproduction of the main processing steps (https://github.com/YNUYangX/Debris-covered-glacier-extraction.git). The corresponding training labels have also been included in the revised data release. The repository link and label-access information have been added to the revised Data and Code Availability statements.