

Response to Reviewer #3,

We thank reviewer for the detailed reviews, and we made all suggested corrections. In this response, the reviewer's comments are in black standard font. Our response is in standard blue font and the modifications to the manuscript are in blue bold font.

Discussing (or at least mentioning in the introduction) current state-of-the-art fusion algorithms

Introduction : Here you need to mention current state-of-the-art fusion algorithms

Charrier et al., 2025 : <https://doi.org/10.5194/tc-19-4555-2025>, 2025.

Greene et al., 2021 : <https://doi.org/10.5194/tc-14-4365-2020>

Derkacheva et al 2020 : <https://doi.org/10.3390/rs12121935>

Thank you for this helpful suggestion. We agree that the Introduction should better reflect recent state-of-the-art developments in multi-sensor and multi-temporal glacier-velocity integration and regularization. In the revised manuscript, we have expanded the Introduction to include the suggested studies, namely Derkacheva et al. (2020), Greene et al. (2020), and Charrier et al. (2025). These references help position our work more clearly with respect to recent advances in multi-sensor velocity processing, seasonal velocity detection, and temporally regularized glacier-velocity time-series reconstruction.

“Recent studies have further advanced the integration and regularization of multi-sensor glacier-velocity observations. For example, Derkacheva et al. (2020) explored statistical and regression-based approaches for reducing and combining ice-velocity information derived from Landsat-8, Sentinel-1, and Sentinel-2. Greene et al. (2020) developed an approach for detecting seasonal ice dynamics from satellite image pairs, highlighting the importance of resolving sub-annual velocity variability. More recently, Charrier et al. (2025) proposed the TICOI framework to generate temporally regularized glacier-velocity time series from heterogeneous and overlapping multi-sensor observations. These studies demonstrate the growing importance of multi-source integration and time-series regularization for improving the continuity and interpretability of glacier surface velocity products.”

Explaining why velocities were computed between images, spaced by one month only. This is surprising, since it drastically reduces the number of image-pair velocities.

Thank you for this important comment. Our intention was not to maximize the number of available image-pair velocities, but to construct a nominal monthly glacier-velocity product with a consistent temporal meaning. For this reason, we preferentially selected image pairs with an interval close to one

month. Using substantially longer baselines would indeed increase the number of valid pairs, but it would also smooth short-term variability and mix velocity signals from different parts of the melt season, thereby reducing the interpretability of the monthly time series. In other words, the one-month strategy was adopted as a compromise between image-pair availability and the temporal resolution required for monthly monitoring.

At the same time, we did not impose a rigid 30-day rule. In practice, we preferentially used low-cloud scenes near the beginning and end of each month, and in months with poor optical conditions the pairwise interval was extended when necessary to maintain data usability.

Give a more detailed explanation of your methodology, including the name of the null hypothesis used, the reason for resampling the images and the presence of gaps after enhancement-based infilling.

Thank you for this helpful suggestion. We agree that the methodological description in the previous version was too concise in several places. In the revised manuscript, we have therefore provided a more detailed explanation of the workflow. Specifically, we now clarify:

- (1) the null hypothesis used in the statistical analysis, namely the default two-sided test in `scipy.stats.pearsonr`, for which the null hypothesis is that the two variables are uncorrelated;
- (2) the reason for resampling the images/products, which is to harmonize the different native grids of the multi-sensor datasets onto a common reference grid prior to pixel-wise comparison, coefficient estimation, and fusion; and
- (3) the reason why some gaps remain after enhancement-based infilling, namely that the enhancement coefficient is computed within a 10×10 sliding window, and when all pixels within a local window are NoData, no valid coefficient can be estimated for that location.

These points have now been clarified explicitly in the revised manuscript to improve reproducibility and reduce ambiguity.

Provide a comparison of your fused velocities with the UAV velocities, even if it's only covering a part of the glacier, and have been used for validation, this is still a way to validate your results.

Thank you for this helpful suggestion. We agree that, even though the UAV data cover only part of the glacier area, they still provide an important local reference for validating the fused velocity product. Following your suggestion, we have added an additional comparison within the UAV survey area in the revised manuscript (Figure 7).

Figure 7 presents the uncertainty analysis results for the UAV-, Landsat-, Sentinel-2-, and fused velocity products within the UAV-covered area from June to November 2023. The results show that the fused product has lower uncertainty than the individual optical products, indicating that the fusion improves

local velocity retrieval quality. At the same time, the fused result still remains notably more uncertain than the UAV-derived reference, which further demonstrates that the UAV comparison provides a meaningful local validation of the fused product. We have incorporated this new analysis and its discussion into the revised manuscript.

In sect 4.2 “Figure 7 presents the uncertainty analysis results for the Landsat-derived velocities, Sentinel-2-derived velocities, UAV-derived velocities, and fused velocities within the UAV survey area. The results show that the fusion method effectively reduces uncertainty relative to the individual optical products. However, a substantial gap still remains between the fused results and the UAV-derived reference. This indicates that, although data fusion can reduce errors to some extent, it still does not achieve the accuracy level of high-precision UAV observations. Moreover, uncertainty estimates based on stable-ground pixels and subsequently propagated to glacier surfaces have inherent limitations. Therefore, the uncertainty obtained from the present data-fusion framework is likely to be underestimated.

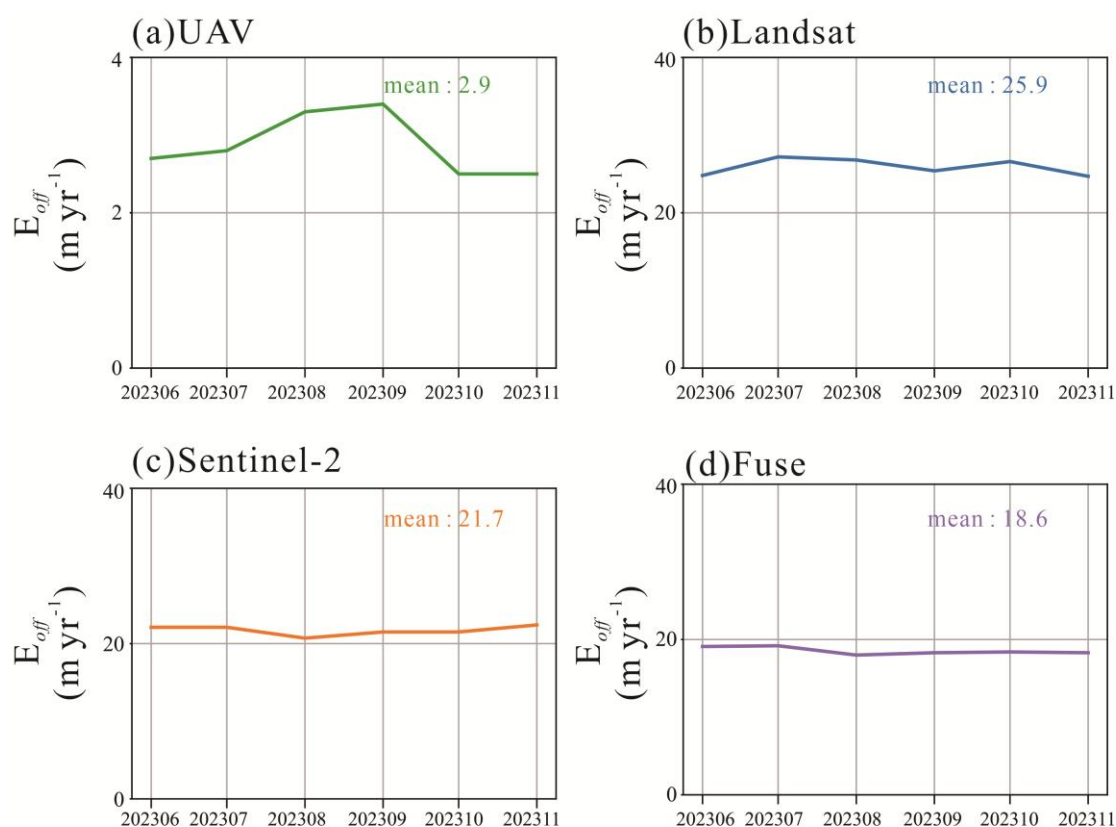


Figure 7: Uncertainty analysis of the four velocity products within the UAV survey area from June to November 2023: (a) UAV-derived velocity uncertainty; (b) Landsat-derived velocity uncertainty; (c) Sentinel-2-derived velocity uncertainty; and (d) fused velocity uncertainty.”

l. 128 : do you convert range/azimuth to E-W / N-S component ? if yes, it needs to be specified

Thank you for this helpful comment. In our Sentinel-1 workflow, the offsets were first estimated in SAR image geometry (range/azimuth) using SNAP offset tracking, and the results were then geocoded through terrain correction to the map-projected ground geometry used in this study. We have clarified this processing sequence in the revised manuscript.

“(3) Radar velocity retrieval: Use intensity-based pixel-offset tracking on Sentinel-1 to estimate range/azimuth offsets in SAR image geometry, then apply terrain correction to geocode the results to map-projected ground geometry and convert them to velocity, providing stable coverage under cloudy conditions. “

Fig. 2 : why are you resampling the images ?

Thank you for this comment. The images/products were resampled because the different sensors have different native grids, and a common reference grid is required for the subsequent pixel-wise WLS fitting and fusion. In the revised manuscript, we have clarified this point and added the detailed resampling procedure in Sect. 3.3.

“ To support pixel-wise WLS fitting and the subsequent fusion, all velocity products were harmonized to a common 30 m reference grid prior to WLS. We used the Sentinel-2 COSI-Corr velocity image as the reference geometry, and co-registered the Landsat-, Sentinel-1-, and UAV-derived velocity image to this grid using nearest-neighbour resampling, so that the original velocity values are preserved without interpolation smoothing. “

l.174 why 1.5 or 0.5 ? why not (annual mean + 2 std mean) which is the zscore ?

Thank you for this important question. We did not use a z-score criterion based on the annual mean \pm 2 standard deviations because our purpose here was not to detect statistical outliers relative to an assumed approximately normal annual distribution, but to remove obviously anomalous mismatches in monthly glacier-velocity time series derived from optical imagery. In this study, the annual sample size for each pixel is only 12 monthly values, and the distribution may be affected by seasonality as well as by occasional cloud-/snow-related mismatches. Under these conditions, the mean and standard deviation can themselves be sensitive to the outliers that we aim to remove.

We therefore adopted a ratio-based threshold relative to a robust annual reference value derived from the intra-annual α -trimmed mean. The $1.5\times$ and $0.5\times$ thresholds were selected as practical empirical bounds to identify values that deviate strongly from the central, year-representative velocity level while retaining normal seasonal variability. In other words, this procedure is intended as a robust quality-control filter rather than a formal distribution-based anomaly test.

eq 1 is not a weighed least square because you did not weight the residuals. There is also a problem in the summa on : it should be over i and j.

Thank you for this careful comment. We agree that Eq. (1), as originally written, should not have been referred to as weighted least squares. In the revised manuscript, we have therefore renamed the method as a constrained least-squares-based weighted fusion framework and revised the text accordingly. We have also clarified the summation notation explicitly over both i and j .

“In this section, we describe a constrained least-squares-based weighted fusion framework in which fusion weights are estimated using co-temporal UAV-derived velocities. Monthly velocity maps from Landsat, Sentinel-1, and Sentinel-2 are generated independently, and UAV orthomosaics are processed to provide a high-accuracy reference for the same periods. A constrained least-squares formulation is then used to estimate the optimal mixing coefficients for the three satellite products, which are subsequently applied in the weighted fusion of the monthly velocity maps.

Let $V_{\text{uav}}(i, j)$ denote the UAV-derived velocity at pixel (i, j) , and $L(i, j), S_1(i, j), S_2(i, j)$ denote the Landsat, Sentinel-1, and Sentinel-2 velocities, with corresponding fusion weights w_L, w_{S_1}, w_{S_2} . The constrained least-squares objective is:

$$f(w) = \sum_{i=1, j=1}^n [V_{\text{uav}}(i, j) - (w_L \cdot L(i, j) + w_{S_1} \cdot S_1(i, j) + w_{S_2} \cdot S_2(i, j))]^2 \quad (1)$$

After defining the objective in Eq. (1), we estimate the fusion weights by minimizing $f(w)$ subject to the constraints $w_L \geq 0$, $w_{S_1} \geq 0$, $w_{S_2} \geq 0$ and $w_L + w_{S_1} + w_{S_2} = 1$. This simplex constraint ensures that the weights are directly interpretable as non-negative mixing fractions. The resulting optimal weight triplet is then applied to fuse the remaining monthly velocity maps.”

Eq 1 do you have one weigth for the entire me-series of each sensor ?

Thank you for this question. Yes, in the present implementation we estimate one set of sensor-level mixing coefficients from the co-temporal UAV-overlap periods, and this same weight triplet is then applied to the monthly velocity maps throughout the time series. In other words, the weights are treated as sensor-specific but temporally invariant in the current framework, rather than being re-estimated separately for each month.

l. 205 Why are you filling the gaps based on Sen nel-1 velocities mainly (even with the enhancement factor) ?

Thank you for this important question. Sentinel-1 is used as the main auxiliary source for gap filling primarily because it provides the most spatially continuous and weather-independent observations in this

cloud- and snow-prone mountain region. In contrast, the Landsat- and Sentinel-2-derived velocity products, although generally more reliable under clear conditions, often contain large gaps caused by persistent cloud cover, seasonal snow, and illumination effects.

In our framework, Sentinel-1 is not used to directly replace missing values in the fused product. Instead, it is used to derive a local enhancement coefficient from the relationship between Sentinel-1 and the preliminary fused velocities in overlapping valid areas. This coefficient field is then smoothed and used to guide the reconstruction of missing pixels. In this way, Sentinel-1 contributes spatial continuity and local flow-pattern information, while limiting the direct propagation of its own retrieval errors into the final product.

In addition, although Sentinel-1 offset tracking is generally less reliable in steep and deeply incised mountain terrain, glacier surfaces are usually smoother and less topographically extreme than the surrounding stable slopes. This makes Sentinel-1 suitable as an auxiliary source for gap filling within glacierized areas. We have clarified this rationale in the revised manuscript.

In sect 3.4 **“Although weighted fusion can effectively integrate multi-source information, some areas may still contain NoData. To further fill these gaps and enhance data continuity, this study introduces a sliding-window enhancement-coefficient infilling method. Because Landsat- and Sentinel-2-derived velocity maps often exhibit large gaps in this cloud- and snow-prone region, and simple interpolation can be unreliable in areas with spatially variable glacier motion, we use Sentinel-1 SAR velocities as a more spatially complete auxiliary field to guide gap repair. This choice is also supported by the topographic characteristics of glacierized terrain: although Sentinel-1 offset tracking is generally less reliable in steep and deeply incised mountain areas, glacier surfaces are usually smoother and less topographically extreme than the surrounding stable slopes. Therefore, within glacierized areas, Sentinel-1 can still provide useful motion-pattern information for gap filling, while avoiding an over-reliance on simple interpolation. In this framework, Sentinel-1 is used as an auxiliary constraint for reconstructing missing values rather than as the primary source for direct interpretation of the final fused velocities. Specifically, Sentinel-1 is used to...”**

1. 220 clouds are also present in Landsat images

Thank you for this helpful comment. We agree that cloud contamination also affects Landsat imagery, and the original wording was not sufficiently precise. In the revised manuscript, we have corrected this sentence to clarify that both Landsat and Sentinel-2 are affected by cloud interference in the study region. **“To visualize quality changes before and after fusion, we compared the fused results separately with velocities from two mainstream optical sources, Landsat and Sentinel-2. Both are widely used**

for glacier-velocity retrieval, but both are affected by cloud contamination in this region. Such contrasts help verify whether fusion effectively integrates complementary strengths and mitigates the weaknesses of the individual optical products. Note that although Sentinel-1 SAR is included, it is not contrasted separately.”

l. 224 the enhancement factor is used only for Nodata no ? why do you say that sentinel-1 contribution on is further modulated by the enhancement-coefficient ?

Thank you for this comment. We agree that our previous wording was not sufficiently precise. In the revised manuscript, we now clarify that the enhancement-coefficient field is used only in the subsequent gap-filling stage for residual NoData pixels, rather than as a modulation of the standard weighted fusion for all valid pixels. The text has been revised accordingly.

“Note that although Sentinel-1 SAR is included in the framework, it is not contrasted separately: (i) it mainly supplements optical observations during cloudy periods and has limited standalone reliability in steep, deeply incised terrain; and (ii) **its contribution to the main weighted-fusion result is limited by its relatively low sensor-level fusion weight (≈ 0.08). The enhancement-coefficient field is applied only afterward, in the gap-filling stage for residual NoData pixels.** Accordingly, this section focuses on before–after comparisons for Landsat and Sentinel-2 to highlight systematic improvements to the primary optical observations.”

Fig4 . why do you still have gaps after the enhancement-based infilling ?

Thank you for this question. The enhancement-based infilling in our framework is not able to fill all gaps unconditionally. Specifically, the enhancement coefficient is estimated within a 10×10 sliding window. When all pixels within a given window are NoData, no valid local relationship can be established and the enhancement coefficient cannot be estimated for that location. As a result, some residual gaps may remain even after the enhancement-based infilling step.

Fig 5. What’s the unit of the axes ? You need to explain why ITS_LIVE have variance than your fuse products : because you used the annual product probably.

Thank you for this helpful comment. We agree that the axis units should be stated explicitly, and we have revised Figure 5 accordingly.

We also agree that the lower variance of ITS_LIVE relative to our fused product requires clarification. In this comparison, ITS_LIVE represents an annual-scale product, whereas our fused velocities are monthly products. Because the ITS_LIVE annual product effectively integrates velocity information over a much longer temporal window, short-term fluctuations are smoothed, resulting in lower apparent

variance. By contrast, our fused product preserves month-to-month variability and seasonal signals more explicitly, which naturally leads to higher variance. We have added this explanation in the revised manuscript.

“Overall, the fused product exhibits a more stable, continuous, and low-fluctuation variance curve in both areas, indicating effective suppression of local noise and improved image consistency. At Yanong, the fused mean variance is $13.2 \times 10^3 \text{ m}^2 \cdot \text{yr}^{-2}$ -lower than GoLIVE ($30.6 \times 10^3 \text{ m}^2 \cdot \text{yr}^{-2}$), comparable to Sentinel-2 ($13.3 \times 10^3 \text{ m}^2 \cdot \text{yr}^{-2}$), and lower than Landsat ($14.6 \times 10^3 \text{ m}^2 \cdot \text{yr}^{-2}$). At Xirinongpu, the fused mean variance is $16.0 \times 10^3 \text{ m}^2 \cdot \text{yr}^{-2}$, below Landsat ($18.7 \times 10^3 \text{ m}^2 \cdot \text{yr}^{-2}$) and Sentinel-2 ($19.9 \times 10^3 \text{ m}^2 \cdot \text{yr}^{-2}$), and also better than GoLIVE ($37.3 \times 10^3 \text{ m}^2 \cdot \text{yr}^{-2}$). Moreover, relative to the two pre-fusion optical products, the fused series shows markedly reduced variance fluctuations, indicating greater smoothness and stability. **The lower variance of ITS_LIVE compared with the fused product because ITS_LIVE is used here as an annual-scale product, whereas the fused results are monthly velocities. The longer temporal aggregation of ITS_LIVE smooths short-term fluctuations and therefore tends to yield lower apparent variance.**

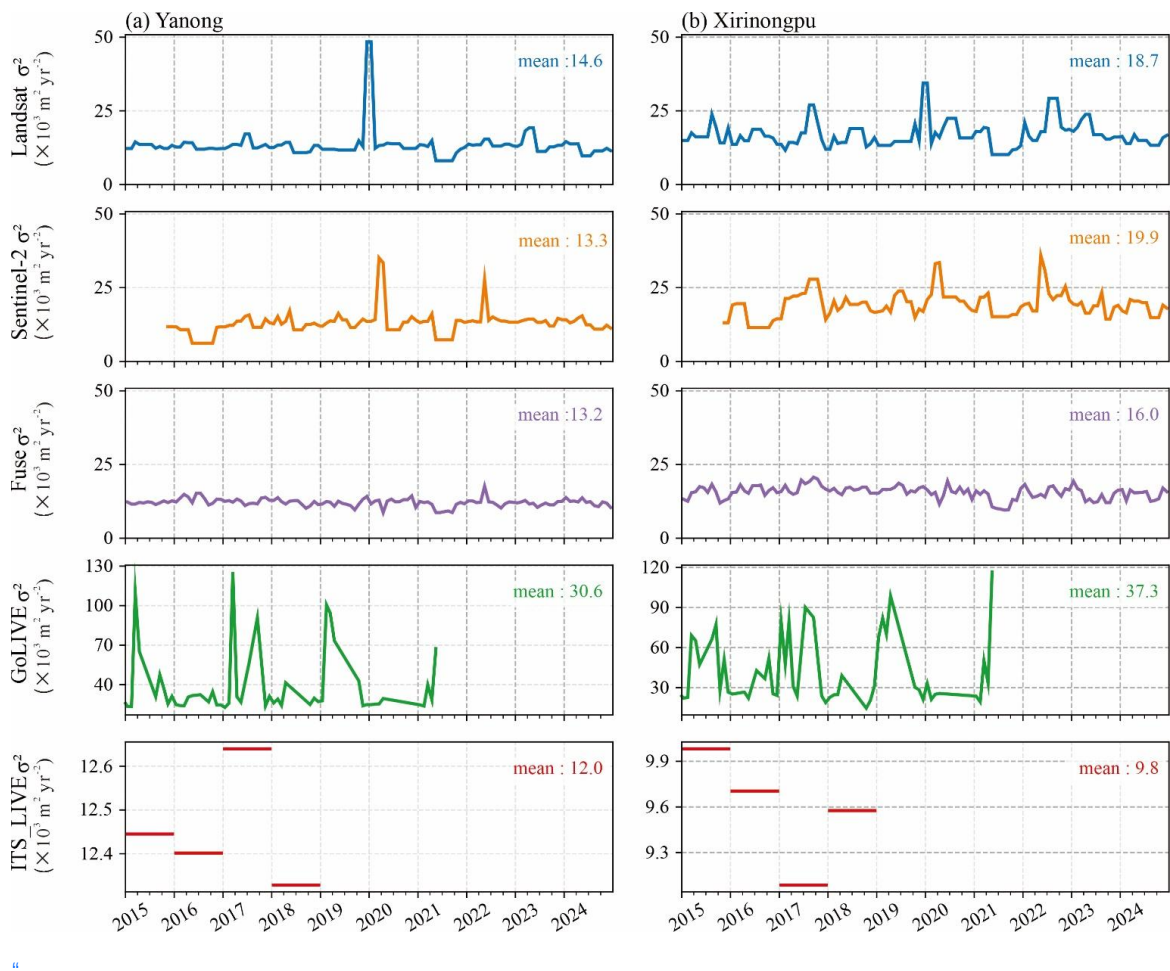
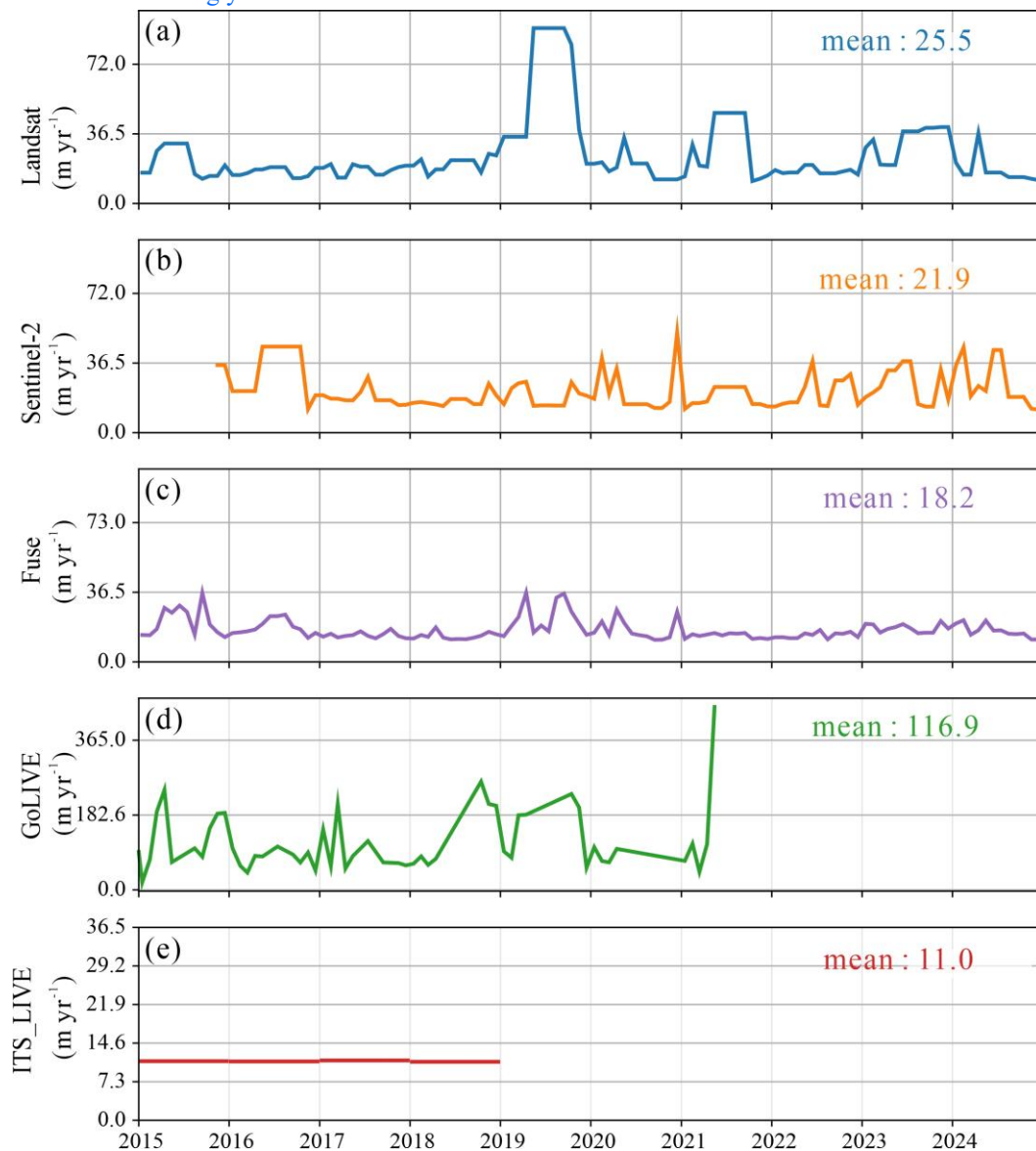


Fig 6. Same what's the unit of the axis ?

Modified accordingly.



1. 345 could you precise the null hypothesis that you have tested here ? is it the one for scipy ? <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.pearsonr.html>

Thank you for this helpful comment. Yes, the significance test reported here follows the default hypothesis test implemented in `scipy.stats.pearsonr`. Specifically, the null hypothesis is that the two variables are uncorrelated (i.e., the population Pearson correlation coefficient is $r=0$), and we used the default two-sided test.

Fig 12 : here there is no way of deciding if the trend is significant or not. For that, you need to perform a null hypothesis test, as done in Halas et al :

<https://doi.org/10.1016/j.rse.2022.113419>

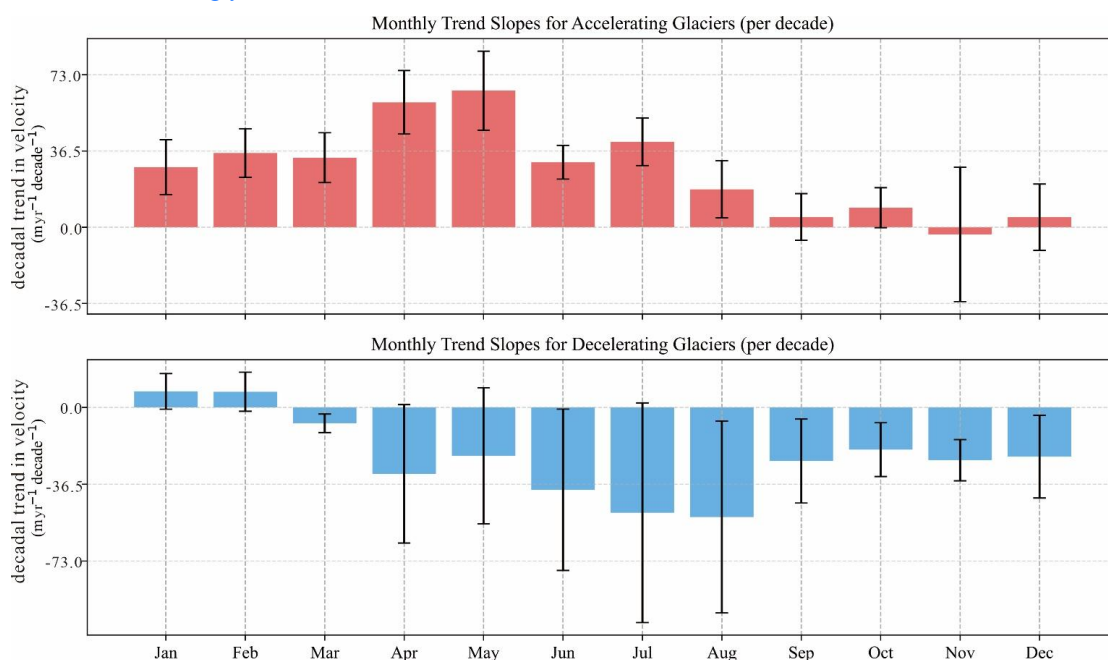
In the section just after you are providing a p value but it's not clear : what test did you used ?

Thank you for this important comment. We would like to clarify that Figure 12 does not represent a significance-tested trend map. Instead, it shows the direction and magnitude of the pixel-wise linear trends in glacier surface velocity during 2015–2024. The statistical significance test and the reported p values apply only to the glacier-level trend analysis presented in Figure 13, where linear regression was performed for each glacier (> 1 km²) and the resulting trends were classified as significantly accelerating, significantly decelerating, or non-significant.

We agree that the original wording could lead to confusion, especially where we used the term “significant” in the discussion of Figure 12. In the revised manuscript, we have therefore clarified this distinction explicitly and replaced the relevant wording in the Figure 12 discussion so that it refers only to the spatial pattern and magnitude of trends, not to statistical significance. The explanation of the p-value test has also been made explicit in the Figure 13 section.

Fig 14 : why is it written slope for the y-axis and not decadal trend in velocity ?

Modified accordingly.



In the conclusion « Pearson correlations indicate significant relationships between velocity and glacier area, slope, and aspect: » is a strong statement recording the pearson value you got.

Thank you for this helpful comment. We agree that the original wording in the Conclusion was too strong given the Pearson correlation coefficients obtained in this study. In the revised manuscript, we have softened this statement and now refer more cautiously to statistically significant correlations/associations between glacier surface velocity and glacier area, slope, and aspect, rather than implying a stronger relationship than is supported by the results.

“In spatial terms, glacier surface velocity during 2015–2024 exhibits the canonical “fast center, slow margins” pattern: multi-year mean maxima exceed $700 \text{ m}\cdot\text{yr}^{-1}$, whereas values in lower reaches and most tributaries are generally $<100 \text{ m}\cdot\text{yr}^{-1}$. Regarding attribute controls, Pearson correlation analysis **indicates statistically significant but moderate associations of** velocity with glacier area, slope, and aspect: larger glaciers tend to flow faster; within individual glaciers, velocity is more strongly correlated with slope; and after accounting for area and slope, south-facing glaciers are slightly faster than glaciers of other aspects.”