Dear Editor:

Thank you for inviting me as one reviewer of this manuscript. This article presents a 30 m, nominally monthly glacier surface‐velocity product for the Kangri Karpo region spanning 2015–2024, derived by fusing multi-source remote sensing velocity estimates (Landsat-8, Sentinel-2 and Sentinel-1) and applying an additional Sentinel-1–guided enhancement-coefficient procedure to mitigate data gaps. It analyses the spatial distribution of the decadal mean velocity field, characterizes intra-annual seasonality, and quantifies interannual trends, with targeted evaluations on representative glaciers.

The manuscript is generally well organized and mostly straightforward to read. I believe it would become a useful contribution to the community who are interested in southeastern of Tibet Plateau, if it could be considered for publication after a round of major revision. Several major concerns and specific comments are listed below.

Major comments:

1) In Sect. 3.2, the UAV coverage appears to be limited to a small terminal portion of a single glacier within the study area. Please clarify to what extent this training domain is representative in both space and time (seasonal/full-year conditions), and justify the applicability of UAV-calibrated parameters/weights when extrapolated to the entire region and the full annual cycle.

2) Although the manuscript states that velocities from the three data sources are harmonized to a 30 m spatial resolution, the underlying native grids of the different sensors/products are not necessarily co-registered. For pixel-wise WLS fitting and subsequent fusion to be valid and reproducible, the pre-fusion co-registration/harmonization steps should be described explicitly, including the choice of the reference grid, reprojection and resampling methods (and interpolation scheme), grid alignment strategy, NoData/mask propagation rules, and whether these operations are performed prior to the WLS fitting.

3) For the WLS-based fusion weights, please specify the exact objective function, the interpretation of the weights within the WLS formulation, and any constraints imposed during optimization. In particular, clarify whether non-negativity and/or a normalization constraint are enforced. If constraints are used, briefly describe the corresponding solution strategy/implementation to ensure reproducibility.

4) In Sect. 3.4, $(i,j)$is defined as the pixel location in the image. However, in Eq. (1) the weight estimation would normally iterate over all pixels of the 2-D grid, whereas the summation appears to run only over i. Sentinel-1–derived velocities are used as the information source for filling gaps in the fused product (via the enhancement-coefficient field), but the rationale for selecting Sentinel-1 as the gap-filling reference is not sufficiently articulated. Also, the first two

paragraph of Section 3.4 is a bit repetitive with Section 3.1.

5) In Eqs. (3) and (4), the enhancement coefficient is denoted as a_{Ω}, which naturally reads as a single constant/parameter associated with a domain Ω. However, based on the stated computation and definition (window-based estimation followed by Gaussian smoothing to form a spatial field), the enhancement coefficient should vary spatially and thus be a gridded quantity.

6) In Sect. 4.1, the manuscript reports an "~70% / ~30% improvement" in valid pixels, but it is unclear whether this refers to an absolute increase (percentage points) or a relative increase with respect to a baseline. To avoid ambiguity, please provide the explicit definition/formula used to compute the improvement.

7) In Sect. 3.5, "image smoothness" is defined as the pixelwise standard deviation within the mask for each month ("…we compute the pixelwise standard deviation…"). However, later (Fig. 5 and associated text) the manuscript refers to using variance to assess image smoothness.

8) In Sect. 4.4, the manuscript states that the observed bimodal intra-annual velocity pattern "accords with" the subglacial hydrology-dynamics evolution framework for maritime glaciers. While this interpretation is plausible, the current discussion remains largely qualitative and lacks a clear evidential link to the data presented in this study.

9) Based on the velocity formulation described in Sect. 3.3, the primary unit of the derived glacier surface velocity should be m d^{-1}. While, in several places (including Sect. 4.3) the manuscript reports velocities in m yr^{-1} or m/year. Please standardize the unit system throughout the manuscript.

Specific comments

Title: "High resolution" could mean high spatial resolution and temporal resolution. As you indicated "monthly glacier surface velocity", how about using "30-meter" to replace "high-resolution"?

Line 17: This sentence confused me the fusion method is for whether high spatial resolution or high spatial and temporal resolution?

Line 25: It would be better to state as "with similar area and slope, south-facing glaciers are slightly faster than northern ones".

Line 96: Please add detail precipitation of this area to illustrate this area is "among the wettest sectors". You also mentioned in Line 99 "high annual precipitation and humidity", which may be repetition in meaning.

Figure 1: "KM" or "km", but not "Km".

Figure 2: "Weight" not "Weigth". Also, the use of blue and white in the figure is a bit misleading.

UAV velocity image and Landsat velocity image are in white, which means they are processing and analysis steps as you indicate in the caption. What are the input datasets of them? Why the final outputs of the velocity results (2015-2024) are not in blue?

Line 159: Please provide the detail acquisition time of the six UAV images, rather than listing a time span of June to November in 2023.

Line 166: “32×32 pixels” not “32*32 pixels”. The results of “2 pixels for Landsat OLI” and “3 pixels for Sentinel-2 MSI” have the same spatial resolution. Why did you not set the sept size as 1 pixel for both Landsat and Sentinel-2, and resample the displacement result to 30m afterwards?

Line 170: I suggest to introduce more about the stabilization of time series in this paragraph, including the description of the use of α-trimmed mean filter, the threshold standard of selecting 0.33, and the use of “*”, which may not be a formal use in the main text.

Line 311: “0.10, lower than” not “0.10-lower than”

Line 321: “product” not “produc”

Line 396: use en dash. “January–March” not “January-March”