We wish to thank Referee 1 for the careful reading and the valuable comments.

We provide below a summary of the key changes implemented and a point-by-point response to all the raised queries (in italic).

General comments

a. Section 2.2.1 Geomatic survey: *The section does not provide a detailed explanation of the photogrammetric processing, and it is not supported by the relevant literature. In my opinion, this section needs to be improved. For instance, the name "SfM-MVS photogrammetry" is never mentioned throughout the article and no effort is made to explain the steps needed in a rigorous photogrammetric study. A few examples: the authors do not mention how the images were collected (e.g., drone flight geometries), how many ground control points they used, or which software or freeware they used to process the images. Finally, the author do not provide any information on the quality of the DSMs, therefore questioning if their results are reliable or not. This important weakness needs to be addressed.*

We are grateful to the reviewer for the specific feedback. Section 2.2.1 will be integrated with all required details about the photogrammetric workflow supported by the relevant literature.

Additional details on the Drone and Crewed Aerial Photogrammetric surveys will be provided and an updated Table 2 will be integrated in the article, detailing the number of Ground Control Points (GCP) and Check Points (CP). Line 138 will be updated adding a remark about the difficulty of retrieving or positioning/measuring control point coordinates in glacier environments, as in the following link: https://zenodo.org/record/8089499 .

Photogrammetric flight	Date of acquisition	Covered area	Extent (km^2)	Average flight height (m)	GSD (m)	Number of images	Number of GCPs	Number of CPs
Aerial	30/09/2020	Glacier and a portion of the proglacial area	25.2	818	0.07	867	18	7
Drone	9/07/2021	L1, L2 and L4	2.6	126	0.03	1480	8	6
Drone	20/07/2021	L1 and L4	0.4	89.2	0.02	369	5	2
Drone	20/07/2021	Glacier front and lower part	1.1	159	0.04	623	Direct georeferencing	
Aerial	13/09/2021	Glacier and proglacial area	34.5	877	0.06	1100	9	4

b. Section 2.2.4 Bedload monitoring: *As per the geomatics, I believe there is the need of providing a more detailed explanation of the bedload monitoring since the use of seismometers in bedload* *studies is relatively recent. It would be very useful to know in more detail how the data were processed and the steps required to go from the raw signal to the results presented here.*

We are thankful to the reviewer for the general and detailed comments. As requested, section 2.2.4 will be modify as follows:

- the sentence on line 252 "Data are recorded with a DATA-CUBE3 (solar power supply, 24-bit converter, GPS-based time synchronization) with a sampling frequency of 200 Hz and stored on site." will be changed to "Data are recorded with a DATA-CUBE3 (solar power supply, 24-bit converter, GPS-based time synchronization) configured with an amplifier gain of 16, with a sampling frequency of 200 Hz and stored on site."
- in line 261 we will add "The counts exported by the DATA-CUBE[3] are converted to vertical ground velocity considering logger and geophone sensitivities according to the manufacturer's specifications. The power spectral density is determined as the ratio of the square of the absolute value of the Fourier transform to the time window (Bakker et al., 2020)."

In this way, we believe that the procedure is clearer and the added reference should provide all details explaining the steps in data processing.

c. Data availability: *The 2020 orthophoto and DSM are not available on Zenodo (https://zenodo.org/record/7713299). Are you going to include them in future?*

Thanks for pointing out that also the availability of 2020 metric products could be valuable for the readers. They have now been uploaded to Zenodo at the following link: https://zenodo.org/record/8089499

Detailed comments:

1. Lines 26 – 28: *"Alpine glacier retreat is leading to increased exposure of formerly glaciated terrain, entailing the colonization of plants and animals, and changes in morphodynamics and sediment transfer."*

Consider adding one or more citations here.

We will add the following 3 citations:

- Mainetti A, D'Amico M, Probo M, Quaglia E, Ravetto Enri S, Celi L and Lonati M (2021) Successional Herbaceous Species Affect Soil Processes in a High-Elevation Alpine Proglacial Chronosequence. Front. Environ. Sci. 8:615499. doi: 10.3389/fenvs.2020.615499
- Brambilla, M. and Gobbi, M.: A century of chasing the ice: delayed colonisation of ice-free sites by ground beetles along glacier forelands in the Alps, Ecography, 37, 33–42, https://doi.org/https://doi.org/10.1111/j.1600-0587.2013.00263.x, 2014

2. Lines 29 – 30: *"Little Ice Age (LIA)"*

You already defined the acronym; perhaps just use LIA instead of "Little Ice Age (LIA)".

We will update the manuscript and use LIA instead of “Little Ice Age (LIA)”.

3. Lines 35 – 36: *“On the one hand, plant colonization stabilizes glacial sediment and reduces sediment fluxes; on the other hand, geomorphic processes disturb and limit vegetation succession.”*

Consider adding one or more citations here.

We will add the following 3 citations:

- Eichel, J.: Vegetation Succession and Biogeomorphic Interactions in Glacier Forelands, pp. 327–349, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-319-94184-4_19, 201
- Moreau, M., Mercier, D., Laffly, D., and Roussel, E.: Impacts of recent paraglacial dynamics on plant colonization: A case study on midtre Lovénbreen foreland, Spitsbergen (79°N), Geomorphology, 95, 1-2: 48-60., https://doi.org/10.1016/j.geomorph.2006.07.031, 2008
- Curry, A. M., Cleasby, V., and Zukowskyj, P.: Paraglacial response of steep, sediment-mantled slopes to post-‘Little Ice Age’ glacier recession in the central Swiss Alps, J. Quat. Sci., 21, 211–225, https://doi.org/10.1002/jqs.954, 2006

4. Lines 54 – 55: *“Sediment yield depends on water discharge and sediment availability which are both highly variable in space and time.”*

Consider adding one or more citations here.

We will add the following 2 citations:

- Heckmann, T. and Schwanghart, W.: Geomorphic coupling and sediment connectivity in an alpine catchment — Exploring sediment cascades using graph theory, Geomorphology, 182, 89–103, https://doi.org/https://doi.org/10.1016/j.geomorph.2012.10.033, 2013
- Hooke, R. L.: Toward a uniform theory of clastic sediment yield in fluvial systems, GSA Bulletin, 112, 1778–1786, https://doi.org/10.1130/0016-7606(2000)112

5. Line 148: *“manned photogrammetric flights”*

I would move away from “manned” and describe those as crewed or airborne.

We are going to replace the term Manned/Unmanned with Crewed/Uncrewed. Additionally, it will be detailed at line 150 that "The Crewed aerial flight was carried out with a P90e light aircraft. Its handling allows easy flight altitude changes to maintain a constant GSD"

6. Lines 154 – 159:

How many targets did you use in total? Did you deploy the targets only in 2021? Why did you not consider collecting independent checkpoints for quality assessment?

The artificial targets were deployed and measured in 2021 and independent Check Points have been used for quality assessment. The derived metric products (orthoimagery and DSM) generated using as GCP/CP the targets described above are used as reference data to extract GCPs and CPs for all the available datasets.

Section 2.2.1 (line 159) will be integrated with the following sentences.

"The 1 m x 1 m markers deployed in 2021 have been used as both GCPs and independent CPs (details in Table 2) for the 2021 crewed aerial survey. Considering that the focus is on relative displacements rather than on absolute values, natural GCPs and CPs have been then identified on the 2021 orthomosaic and DSM to orient the 2020 crewed aerial imagery and assess its 3D positional accuracy (considering that the artificial markers were not yet available in 2020).

The same approach has been used for drone surveys (except for one drone survey where, exploiting the RTK capabilities of the drone GNSS receiver, a direct georeferencing approach has been adopted)

Section 3.1 will be integrated with the following sentence (at line 268) and table.

"Table **XX** shows the planimetric and altimetric errors calculated on both GCPs and CPs. The "reference" dataset is characterised by a planimetric accuracy (CPs) of 8 cm and a vertical accuracy of 11 cm. The 3D accuracies of the other datasets are calculated considering the 2021 dataset as ground truth".

Table XX - planimetric and altimetric errors calculated on both GCPs and CPs.

Flight	n° GCP	n° CP	Residuals GCP [cm]		Residuals CP [cm]	
			RMS plan	RMS H	RMS plan	RMS H
aerial 2021	9	4	4,7	6,1	7,7	11,1
aerial 2020	18	7	24.3	10.6	9.5	16.7
drone 9th July 2021	8	6	2,6	1,9	5,4	7,8
drone 20th July 2021	5	2	22,7	8,8	1,8	1,5

7. Line 160: *"Unlike drone flights which were oriented exploiting a direct georeferencing approach"*

What do you mean with "direct georeferencing"? Did you use the camera positions alone? If yes, why? The GPS onboard of the DJI P4 is of poor quality for high-precision photogrammetric surveys, and it is a standard practice to use ground targets in SfM-MVS studies (particularly to reduce the occurrence of systematic deformations in DEMs).

Firstly, we would like to point out that only one of the drone surveys has been oriented by means of a Direct Georeferencing approach (having used a DJI Phantom 4 with RTK capabilities and a GNSS base station positioned on a point of known coordinates and sending RTK corrections through RTCM radio transmissions), and the highlighted sentence will be amended accordingly.

A few details about the direct georeferencing approach, including relevant references, will be added at line 160, namely:

"Direct Georeferencing refers to the orientation of remotely sensed imagery without the use of GCP, exploiting Real Time Kinematic (RTK) or Post Processing Kinematic (PPK) approaches. The RTK- or PPK-based approach enables the generation of metric products with 3D positional precision and accuracy in the range of few centimeters (Teppati Losé et al., 2020; Chiabrando et al., 2019)"

We will add the following 3 references:

- Teppati Losè, L., Chiabrando, F., and Giulio Tonolo, F.: Boosting the Timeliness of UAV Large Scale Mapping. Direct Georeferencing Approaches: Operational Strategies and Best Practices, ISPRS International Journal of Geo-Information, 9, https://doi.org/10.3390/ijgi9100578, 2020
- Chiabrando, F., Giulio Tonolo, F., and Lingua, A.: UAV DIRECT GEOREFERENCING APPROACH IN AN EMERGENCY MAPPING CONTEXT. THE 2016 CENTRAL ITALY EARTHQUAKE CASE STUDY, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2/W13, 247–253, https://doi.org/10.5194/isprs-archives-XLII-2-W13-247-2019, 2019
- Teppati Losè, L., Chiabrando, F., and Giulio Tonolo, F.: ARE MEASURED GROUND CONTROL POINTS STILL REQUIRED IN UAV BASED LARGE SCALE MAPPING? ASSESSING THE POSITIONAL ACCURACY OF AN RTK MULTI-ROTOR PLATFORM, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B1-2020, 507–514, https://doi.org/10.5194/isprs-archives-XLIII-B1-2020-507-2020, 2020.

The flight over the front was oriented with direct georeferencing and its positional accuracy was estimated based on the above-mentioned references (considering that no stable CP can be identified over the glacier area).

8. Lines 163 – 166: *"Due to a large number of well-distributed ground control points, the 2021 aerial survey was considered the reference model (referred to as 'Model" Zero') to be used for multitemporal analyses. The 2020 survey was, therefore, co-registered (i.e., georeferenced in the same reference system, enabling the overlap of all the derivative products) with the 2021 survey."*

This suggests that you did not use any target in 2020 (see comment 6), am I right? How did you co-register the 2020 survey? Could you please explain the co-registration procedure? Could you provide statistics on the quality of the co-registration?

As detailed in a previous reply, for 2020 survey 18 GCPs e 7 CPs have been used. The co-registration is granted by the fact that the GCPs have been identified on the 2021 final products. The relevant positional accuracy statistics will be reported in a new table (the Table X mentioned above) in the revised version of the manuscript.

9. Lines 282 – 284: *"The aerial DSMs were preliminary compared to the LiDAR DSM as of 2008 available on Valle d'Aosta Geoportal to verify the consistency of the produced model, checking the stability of the periglacial rocky areas. Subsequently, 2021 and 2020 DSMs were subtracted to quantify glacier ablation and displacement"*

I would move this section into the methods, and explain how you compared the DSMs. In the results, it would be more informative to provide the statistics of such a comparison (e.g., mean error, std of error – not the RMSE) in order to demonstrate that your DSMs were free of systematic (mean error close to 0) and random (std of error close to 0) errors.

We'll insert the following sentence in the methods section (2.2.1) at line 167. Additionally, the reference to the 2008 dataset will be removed, not being one of the main datasets described in the manuscript (following the advice of the reviewer in comment 11).

"Reiterating that the main focus is on evaluating relative displacements, the DSMs have been compared using a pixel-by-pixel approach. Specifically, the height from the 2021 DSM (i.e. the pixel value) was subtracted from the 2020 DSM one. The overall comparison enables the evaluation of the changes of the glacier surfaces, while the comparison limited to stable areas enables a further validation of the elevation products. To define the stable areas, we considered areas not covered by ice, snow or water: in order to retrieve a statistically relevant comparison dataset, we included in the stable areas also outwash plains, areas that may have been affected by geomorphological changes mainly due to water erosion and deposits. Since the stable areas are used as relative validation of the DSM, this choice is conservative since it may worsen the statistics." The outline of the stable areas will be inserted in Figure 8a.

In the results section (3), in addition to the result related to the glacier monitoring, a dedicated table related to the difference of DSMs over stable areas will be included, showing the statistics requested by the reviewer, with and without the exclusion of outliers in the DSMs. The median of differences is -0.098 m, the mean differences and standard deviation -0.082 m ± 0.788 m and the mean differences and standard deviation at 95% confidence level is -0.072 m ± 0.141 m. The Table XXX will be as follows.

Table XXX - Elevation differences (DSM 2020 - DSM 2021) on stable areas.

Elevation differences (DSM 2020 - DSM 2021) on stable areas	
Median	-0.098 m
Mean	-0.082 m ± 0.788 m
Standard deviation (95% confidence level)	0.072 m ± 0.141 m

10. Figure 8a:

I am a little concerned about the way you presented the DSM of difference. First, why did you not use a bivariate scale (from –X to +X)? A bivariate scale would help a lot in my opinion. Second, why did you not apply a Limit of Detection? The use of a Limit of Detection is a common practice in DoDs, and allows showing changes that are statistically significant (e.g., at 68 or 95% confidence limits). Lastly, from the DoD presented in Fig. 8a it seems that the whole study area experienced at least some movements in Z, is that really possible? Did you check for systematic deformations (e.g., doming, datum shift) in your DSMs? There is the need of providing statistics that illustrate the quality of your DSMs, e.g. the mean error in Z (i.e., systematic errors) and the std of error in Z (i.e., random errors) in respect to reference point altitudes or independent check points.

• bivariate scale (from –X to +X): We appreciate the advice and we have modified the image accordingly, introducing a bivariate scale and the Limit of Detection (95%) thresholds in figure 8a.

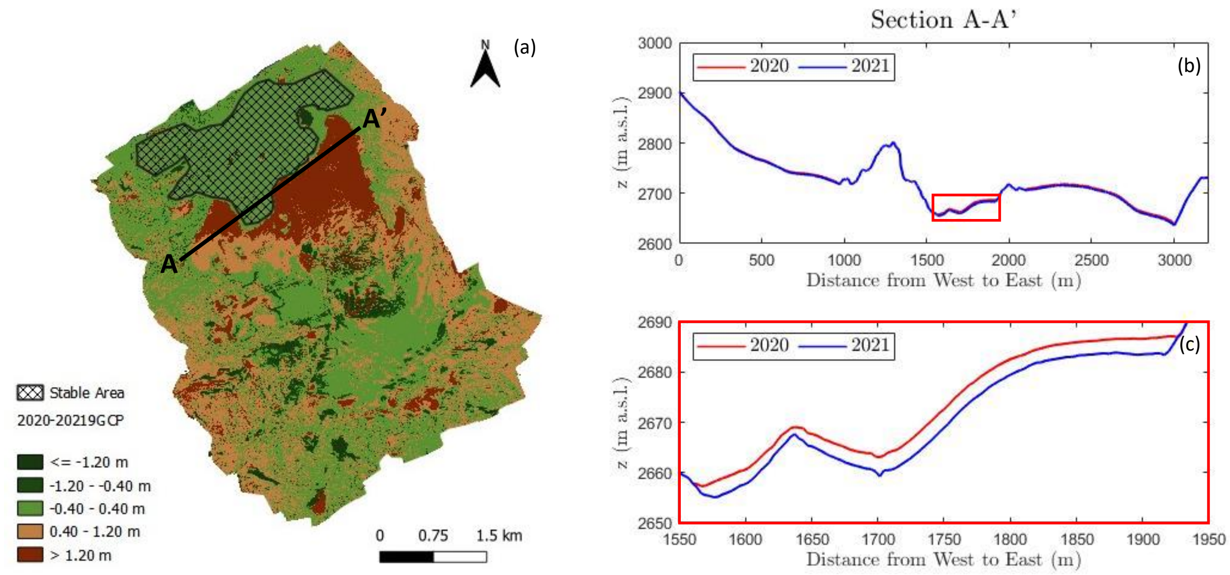

Figure 8. (a) Difference between the DSM of 2020 and 2021. The black line refers to the cross-section A-A', whose 2020 (red) and 2021 (blue) elevation profiles are shown in panel **(b)**, with a zoom-in on the central tongue of the glacier in panel **(c)**.

• Limit of Detection: Thank you for your suggestion. The text will be integrated with a discussion on the suggested LoD approach at line 285: "When comparing two DSMs (i.e. Difference of DSMs, DoD) it's crucial to distinguish the information (actual vertical displacement) from the noise. To this purpose, the Limit of Detection approach has been adopted. The vertical error of the two DSMs propagates when calculating their difference. From the standard deviation of the DSM, it is possible to calculate the standard deviation of the difference exploiting the error propagation theory. The vertical precision (based on CPs) of the DSMs is 11 cm for 2021 and 16.7 cm 2020. The DoD LoD at 68% confidence level is 20 cm. Figure 8a shows the differences between the 2021 and 2020 DSMs starting from a LoD threshold of 95% = 40 cm." - (ref. Azmoon, B., Biniyaz, A., and Liu, Z.: Use of High-Resolution Multi-Temporal DEM Data for Landslide Detection, Geosciences, 12, https://doi.org/10.3390/geosciences12100378 , 2022.)

• systematic deformations: Considering that 2020 aerial surveys were georeferenced using control points extracted from the 2021 one and the CPs horizontal residuals are comparable with the DSM resolution, we do not consider Datum Shift significant. Additionally, using the ETRF2000 reference system, the related displacements (about 1 mm/year) do not lead to a significant datum shift.

• in Fig. 8a: figure 8a will be replaced considering the Limit of Detection at 95% confidence level, as suggested.

11. Line 289: *"Additionally, a comparison with the 2008 DSM shows a lowering of glacier surface up to 50 meters in glacial front areas"*

Where is this result presented?

As previously mentioned (reply to comment 9), we'll remove the reference to the 2008 DSM (lines 282-283 and 289), not being the focus of the manuscript (we'll include a reference to such comparison in the conclusion section)

12. Lines 290 – 293: *"As far as very high-resolution satellite stereo pairs are concerned, they enable the extraction of 3D information with a lower vertical accuracy (metric level) with respect to aerial and drone data. Nevertheless, the coverage of a much larger area (in the range of hundreds of square kilometres) enables a multiscale and multiplatform approach to identify the most critical areas where to focus the monitoring activities in the field"*

You do not present these results nor discuss them later, what is the point of including such a thing?

As per the 2008 DSM, we do agree and we will remove line 167-170 and 290-293 related to satellite imagery.

13. Lines 300 – 303: *"The x-y-z locations of the first interface, representing the lake bottom, detected in all the GPR sections, were interpolated to produce a bathymetry map (Figure 10, which also displays the sediment thickness distribution and the electrical conductivity measurements). The perimeter of the lake, retrieved from the 6-cm-resolution orthophoto, was useful to fix the 0-depth in the interpolation process."*

This section reads like methods. I would move it to the methods.

We will re-arrange and move the paragraph under consideration to the methods section as follows.

- we will delete the sentence in lines 195-196: "The analysis of the GPR travel times provided the sections of the water depth and sediment thickness, which were interpolated into a bathymetric model.";
- we will add in line 195: "The GPR sections acquired were processed according to a set of standard processing steps, as detailed in Vergnano et al. (2023). The x-y-z locations of the first interface, representing the lake bottom, detected in all the GPR sections, were interpolated to produce a bathymetry map (Figure 11, which also displays the sediment thickness distribution and the electrical conductivity measurements). The perimeter of the lake, retrieved from the 6-cm-resolution orthophoto, was useful to fix the 0-depth in the interpolation process."

Nevertheless, the resulting interpolated lake bathymetry, being one of the research outputs, will be presented in the Result section.

14. Lines 326 – 334: *"The ecoLog1000 and CTDs instruments were first installed in July 2021 and June 2022, respectively. The measuring periods of each sensor are shown in a time: measured-quantity diagram in Table 1. At the L4 gauging station, a set of velocity-based discharge measurements (Q) taken in the summer of 2021 and 2022 were related to the corresponding water depth measured at the gauge (h), in order to plot the stage-discharge diagram (Fig. 11(a); details of the procedure followed to determine the stage-discharge relationship are given in Appendix A). Discharge measurements were also used to calibrate the lake outflow curve, i.e., the relationship between the hydraulic head (H) in the lake and the flowing discharge (see Fig. 11(c)). For this purpose, a linear fitting between the water depth at the gauge (h) and the Hydraulic head in the lake (H) was also calibrated (Fig. 11(b), R2 $\sim$ 0.98), since the water levels in the lake and in the control cross-section in the stream are strictly related but not equal, due to the head-dependent outflow process and water speed."*

This section reads like methods, I would therefore re-arrange this part.

As suggested, we will re-arrange and move the paragraph under consideration to the methods section as follows:

- we will add in line 219 “The ecoLog1000 and CTDs instruments were first installed in July 2021 and June 2022, respectively; the measuring periods of each sensor are shown in a time: measured-quantity diagram in Table 1.”;
- in line 231 we will change the sentence “Flow velocity measurements were taken with an Acoustic Doppler Velocimeter (ADV) current meter in the cross-section of gauging station L4 for a total of 9 surveys.” To “In the summer of 2021 and 2022, a set of nine flow velocity measurements were taken with an Acoustic Doppler Velocimeter (ADV) current meter in the cross-section of gauging station L4.”
- we will add before the paragraph on line 234: “The velocity-based discharge measurements *Q* were related to the corresponding water depth *h* measured at the gauge, in order to plot the stage-discharge diagram (Fig.4 (a)), further details on the procedure are given in Appendix A).”
- in line 236 we will change the sentence and add further details: “To monitor the water level in the lake, the relationship between the water level recorded continuously in the L4 gauging station and the water level in L4 was determined.” to “Due to backwater effects at the outflow, the water levels in the lake and in the control cross-section are not identical, but strictly related. Therefore, in order to monitor the lake's stage, a relationship between the continuously recorded water level at the gauging station and the water level in the lake far from the gauging station, was determined.”
- we will add in line 245 “The best fitting of relation between the water depth measurements at the gauge *h* and the Hydraulic head in the lake *H* was found to be linear (H=1.3h-0.1279, R^2 ~0.98; (Fig.4(b)). The stage-discharge diagram (h-Q) and the linear fitting (h-H) were used to calibrate the lake outflow curve, i.e., the relationship between the hydraulic head (*H*) in the lake and the flowing discharge (*Q*) (Fig.4(c)).”.

Accordingly, we will amend the first paragraph of subsection "3.3 Hydrometric Monitoring" as follows: “The investigation at the L4 gauging station involved: i) a set of 9 velocity-based discharge measurements which allowed the stage-discharge diagram (h-Q) to be assessed; ii) a set of 15 elevation difference measurements which led to the linear fitting (h-H) and the lake outflow curve (H-Q) (Fig. 4).”

14. Line 357: *“permits the identification of time intervals characterized by intense transport”*

Although bed load would easily occur at high discharges, is the signal you see necessarily related to intense transport events? Or, could the peaks be related to flow turbulence instead?

To better clarify this, we have added in line 363 the following sentences: ''It is assumed that the geophone signal (Figure 12b) permits the identification of time intervals characterised by intense transport since, in correspondence with the peaks of the envelope, the power increases for high values of frequency (Figure 14). Indeed, the power in the lower bands is attributed to turbulent fluid flow (Schmandt et al., 2013) while that in the higher bands to bedload (Schmandt et al., 2013; Bakker et al., 2020).”.

15. Lines 358 – 359: *"Raw seismic signals were filtered in the band 5-95 Hz and then the envelope was calculated as the average of the absolute value of the filtered signal over a time window of 1 min"*

The sentence reads like methods, consider moving it in the appropriate section.

We will move the sentence under consideration to the methods section below line 260, after the integration in the answer to general comment b.

16. Lines 363 – 364: *"In 2021, we directly observed the absence of bedload transport in three days (10 July, 20 July and 13 September)."*

What does "directly observed" mean here? Could you be more precise?

To clarify this concept, the sentence will be modified as follows: "In 2021, we observed, through direct inspection of the flow field, the absence of bedload transport in three days (10 July, 20 July and 13 September)."

17. Lines 364 – 371*: "During the 2022 season, we performed direct measurements of bedload transport at the glacier mouth by means of portable samplers on the occasion of one day of intense glacier melt (14 July) and at the end of the monitoring season (16 September). Bedload traps (4 mm mesh size, 20 × 30 cm opening, (Bunte et al., 2004)) were deployed simultaneously at 2 positions. Measured unit bedload rates feature a large variability ranging from 0.02 to 16.2 kg/m/min in a few hours, as already observed in glacierized basins (Coviello et al., 2022). Bedload samples were sieved and weighed to obtain the grain size distribution. The total bedload transport rate Qs (kg/min above 4 mm) for each sampling period (ranging 370 from 2 to 30 min) was estimated as width-weighted averages based on the available positions sampled."*

This section really is about methods, I would therefore move into the methods.

We will move this paragraph to the end of subsubsection "2.2.4 Bedload monitoring".

18. Lines 416 – 417: *"It is important to stress that the accurate georeferencing of all the acquired data with respect to the same Datum plays a crucial role in the data integration phase and in enabling the multitemporal analyses."*

This is right, but what about systematic deformations that could well lead to erroneous multitemporal analysis? See comment 10.

This comment has been addressed in a previous reply to comment 10.