We want to thank the anonymous referee for this technical review and comments that will help to improve some aspects of the manuscript. We agree with most of the comments and suggestions made by the anonymous referee. Nevertheless, very similar observations were also raised by the first reviewer, and therefore, already explained in the first author’s response (AR1). Where this is the case, we indicated that similar responses are also in the first author's response (AR1).

The referee comments are in italic font, and our responses and revisions are in normal font.

Review of the manuscript “Validation and application of sequential unmanned aerial vehicle surveys to monitor the kinematics of a rapid rock glacier” from Vivero et al. Submitted to The Cryosphere earth surface Dynamics.

This paper attempt to show the importance of using sequential UAV-surveys and Structure for Motion photogrammetry workflows for monitoring and quantifying changes in rock glaciers. This study is based on the analysis and comparison of very high-resolution photogrammetric products (i.e. DEMs and orthomosaics) acquired between 2016-2019.

The manuscript has as the main goal to propose the benefits of using UAV surveys through a “rigorous” methodological approach to estimate surface velocities and volume changes over the Tsarmine rock glacier.

Finally, the authors analyse the UAV monitoring strategies as well as the kinematical behavior and surface changes of this rock glacier.
Due to the general arguments and some detailed comments unfortunately I think that there are some major drawbacks in this manuscript that from my point of view could question the acceptance of the manuscript. In addition, even though I am not native English speaker, I think the text requires some corrections. I strongly recommend having the manuscript reviewed by a English native speaker.

Thank you for the overall summary of our work. We will review the English of this manuscript accordingly.

General and specific comments are described as follows:

General comments:

SfM methodology

First, the main contribution of the study is focused on the establishment of a “rigorous methodology” for monitoring rock glaciers using UAV. However, this aspect is crucial because in this study I could not find significant differences (in the methodology) concerning other studies.

We propose UAV surveys as a complementary method to monitor rock glacier kinematics. For achieving this, we exploited the capabilities of the image correlation software CIAS to derived landform-wide kinematic fields from sequential UAV-SfM-derived orthomosaics. A fundamental step before explaining the geomorphological changes was to validate the results via independent ground information during five consecutive periods. The good agreement with data collected by repeated dGNSS demonstrated the robustness and reliability of the monitoring protocol (Figure 3), which similar studies have not demonstrated. Please, see also our similar response to reviewer 1 (AR1).

No discussion is presented about the “implementation of this rigorous protocol” methodology. At this point, I wonder, what is the main difference between the traditional UAV surveys and the new one you propose? More details are needed about specific steps on the protocol, e.g. orientation parameters, the quality of the point cloud, and the differences between GCPs and external data.

We acknowledge that the UAV monitoring concept and validation was not well presented. We changed the emphasis to the evaluation of the UAV-SfM-derived products rather than the SfM processing. In doing so, we included a new section (3.4 Comparison with the TGS) where we better explained the validation protocol. We also expanded the section dealing with the coregistration assessment and added a new table. Please, see also our similar response to reviewer 1 (AR1).
Second, you mentioned in the text the implementation of the “co-registration” stage as the main difference between other studies. However, I think there is a misinterpretation of the term “co-registration”. Initially, you speak of multi-temporal co-registration (Figure 2) to the fact of always using the same 4 control points for the aerotriangulation stage. However, co-registration is more related to comparing one product with another, (i.e. master DEM to another, co-registration on SAR images that follows the same principle). The fact that you are always using the 4 points for the aerotriangulation stage is part of the "rigorous protocol", nevertheless, this is not necessarily a co-registration stage. In your case, the co-registration step should be implemented after ortho-images and DEMs generation and with independent references (points, DEMs, etc) that were not included in photogrammetric analysis.

We agreed that the term co-registration was incorrectly assigned to the four multitemporal GCP. We have changed accordingly and corrected the caption in Figure 2.

The GCPs and CPs are used to determine the absolute accuracy of the UAV-SfM-derived products. However, the relative accuracy was evaluated for each orthomosaic pair using a Helmert similarity transform based on 69 points sampled on stable terrain features without significant vegetation. In doing this, we assessed both the coregistration and image matching accuracies. See our similar response to review 1 (AR1).

Then you talk about that the co-registration based on the Helmert methodology you do not find rotations or changes of scale. This issue is obvious since you always used the same 4 points for aerotriangulation. Then, you mention shifts between orthomosaics (computed again using Helmert methodology) but these values are not reported anywhere. Is this important? Or not?

The mention of the lack of rotation or scale differences among the sequential orthomosaics may sound trivial, but we include it as initial quality control. Additional details are included on the new Table and supplementary material.

Later, only the ortho-images were "co-registered" and not the DEMs. In the case of x-y shifts in the ortho-images were important, this also implies strong variations in the elevation changes calculated using the non-coregistered DEMs.

We thank the reviewer for this comment, but we did not observe such “strong variations in the elevation changes” caused by the subpixel x-y-shifts. Therefore, we evaluated the uncertainty of each DEM according to their absolute accuracy (35 points), and we propagated the errors during the DoD analysis following Wheaton et al. (2010). We included these results in the supplementary material.

Finally, you mention that you use small areas to perform this co-registration, however,
these areas represent a very small portion of the overall DEMs (which is not presented indeed). The fact of selecting small areas can lead to misinterpretations in the uncertainties.

The co-registration was performed on 69 points distributed on well-evaluated areas (see Kummert and Delaloye, 2018) and without significant vegetation (as the NCC misses the coherence on this kind of surfaces). This selection was not trivial, as we have detected some landslide activity in the vicinity (see Barboux et al., 2014). Still, the “small areas” represent about 1/4 of the study area (Tsarmine rock glacier) and are distributed on five disconnected regions. We are not aware of this level of assessment of the stable regions at this high resolution by similar studies.

**Surface velocities**

Another methodological question arises on the presentation of rock glacier surface velocities. More details are needed to discuss the validation of the feature tracking products. Also, the current manuscript does not allow the correct interpretation about technical details like grid size, correlation windows size which is fundamental. For example, given the current displacements of Tsarmine rock glacier in the frontal part, adaptable window correlation size are maybe needed. Independent comparisons between products (correlation results) should be mentioned and results should be provided as supplementary material.

The decision to use a particular combination of reference and search window sizes was pragmatic because the expected maximum displacement can be estimated beforehand from the visual inspection of sequential orthomosaics. In this regard, keeping the exact configuration of the search and window sizes provided only a theoretical maximum displacement limit of about 13 m (more than twice the maximum value observed) for each pair of consecutive orthomosaics.

We included a new section (3.4 Comparison with the TGS) to explain the validation protocol better. We also expanded the section dealing with the coregistration assessment and added a new table. Please, see our similar response to reviewer 1 (AR1).

*Given the recent importance of these in geosciences, I consider that it is necessary to expand a little more literature on the methodology used in other studies and to be able to propose a new one.*

We expanded the references.
As showed in Table 1, your latest survey was made early in the morning. This aspect is not even mentioned, we don’t have any idea if the presence of shadows in the morning or light conditions could harm the feature tracking process.

We thank the reviewer for pointing on this valid concern; however, It is well known that the normalised cross-correlation (NCC) can account for the differences in contrast and brightness (see Heid and Kääb, 2012). We included this explanation.

The fact that you are using very high-resolution data leads to fine changes in image texture and quality. There is no discussion about the influence of using a very high-resolution image to compute surface velocities fields.

Agreed, we have put more emphasis on the very high-resolution datasets.

**Thickness changes**

This study leads with very high-resolution data, therefore the validation of these must be rigorous and independent. Although the manuscript presents a validation step (independent using GPS data), robust statistical analysis is necessary to support these results.

Additional details are included in the new supplementary material (DEM assessment).

Additionally, as you are computing differences between many DEMs, uncertainties between those should be presented at least as supplementary material.

Agreed, we included additional details in the new supplementary material (DEM assessment).

**Specific comments:**

P1L10-11: as details about UAV surveys were not detailed, the “rigorous procedure” is ambiguous and it is no possible to understand why it is rigorous.
Agreed, we removed “rigorous” from the abstract.

P1L19: On the phrase “we almost archived the same accuracy as the GNSS-derived velocities”, numerical comparations should be provided for this comparison.

Changed to: “We achieved uncertainties down to 0.08 m using the adequate UAV survey acquisition, processing and validation steps.”

P1L26-27; The phrase “Rock glacier analogues […] (Cliford et al., 2013)” does not contribute to the discussion of the document and is out of context.

Agreed, we remove this sentence from our manuscript.

P2L33: the word “top” is too ambiguous. I suggest changing this by “surface”.

Changed to “surface”.

P2L37-39: The phrase “The most extending […] Schneider, 2001)” should be revised because it is no clear.

Changed to “One of the first monitoring efforts goes back to the summer of 1938 at the outer Hochebenkar in the Ötztal Alps of Western Austria”.

P2L44: Annual surveys derived from what?

Annual surveys derived “from in-situ measurements”.

P2L62-63: on the sentence “[…] more recently using differential GPS (Berthling et al., 1998”) the reference is too old.

We removed “and more recently using”.
P3L64-68: In this paragraph, you mentioned remote sensing techniques to measure rock glacier displacements and all the references are at the end. However, we do not know which corresponds to which one. I suggest changing this paragraph putting the correct references after each technique.

Changed.

P3L70: There is no only Kaab that has been working on photogrammetry and rock glacier. For example, you forgot Kaufman publications and one of your Swiss colleagues Michelleti et al 2015, Australian and French colleagues.

Andreas Kääb did a lot of work in this field, and in our view, he produced some of the most inspiring articles dealing with photogrammetry applications on rock glaciers. Furthermore, we are employing the CIAS software which he developed. We expanded, however, the list of authors. Also, we are not aware of the “Australian colleagues” working on photogrammetry and rock glacier.

P3L80-81: Ok, we are gree that UAVs are changing and improving, but it will affect the way to make photogrammetry? I think this sentence is too ambiguous.

We have changed the sentence: “how rock glacier kinematics and geomorphic changes are evaluated still presents a challenge, and there is still scope for the assessment of monitoring protocols using different UAV configurations high altitude and challenging terrain”.

P3L85: What is necessary?

We modified this sentence to: “In the context of rapid rock glacier acceleration, a robust validation of the increasing velocities from particular remote sensing techniques is necessary”.

P3L87: the benefits of UAV have been already proved for rock glaciers? Why are you interested in proving these again?

Previous research has partly described these benefits, but as we had argued, a
destabilised rock glacier entails a specific UAV monitoring strategy and particular analyses. Topics such as substantial velocity variations, scarp development and frontal erosion processes have been described and discussed in the present manuscript. See our similar responses to reviewer 1 (AR1).

P4L119: Mounty peak instead that "peak monthly".

Changed.

P5L127: 58 points is a specific number, it cannot be "around".

Corrected.

P5L128: [...] six fixed points [...]. In this sentence, you write six? However, on the L127 you mentioned "58 points..." with a number. Please harmonise.

According to the “house standards” in ESurf: “For items other than units of time or measure, use words for cardinal numbers less than 10; use numerals for 10 and above (e.g. three flasks, seven trees, 6 m, 9 d, 10 desks). Therefore, we have to follow the “house standards”.

P5L141: Table 2 is mentioned before Table 1.

Revised.

P5L151: Add focal length units.

Added.

P5L152: Add focal length units.

Added.
P5L155- L152: The sentence “However […]” is too ambiguous. Poor information is provided about the independent test.

Clarified.

P6L164-165: The sentence “Moreover […] ground […] is repetitive. You already mentioned this on L146. Please check the redundancy.”

Thanks for point this redundancy issue. We modified this sentence.

P7L208: are you sure that the way to present “average density of points” is the correct way? Please verify this.

Density is usually quantified as a quantity per unit of volume (pts m$^{-3}$). We will verify this with the Copernicus typesetting and language copy-editing services.

P8L223: I am confused. On the L210 you said that you generate ortho-mosaics at 0.1 m resolution. But here on the L223, you calculate surface displacements using ortho-mosaics at “one-tenth of the original pixel size (i.e. 0.01m)” which is not coherent. Please verify.

We added the following sentence: “To achieve displacements at sub-pixel precisions (Debella-Gilo and Kääb, 2011) the original ortho-mosaics were resampled to one-tenth of the original pixel size (i.e. 0.01 m).”

P8L229: Dis you check the influence of shadows on very high-resolution ortho-mosaics? This aspect is not even mentioned.

The investigated area inside the ortho-mosaics is normally free from deep shadows caused by high mountain relief. Therefore, we did not observe such influence in the CIAS results. We included this explanation.

P8L234-237: Disagree. Please see my general comment. No detailed information is presented about this issue (i.e. differences in stable areas, x-y shifts, etc.).
Please, see our previous response to this comment.

P8L240: Again, 69 is a specific number, it cannot be "around". Also, these 69 "rock surfaces" are not displayed anywhere.

We removed “around”. Please, see our response to reviewer 1 (AR1).

P8L241: x-y-shift translation vectors are not mentioned anywhere. Please provide a detailed description of the specific values obtained from the Helmert similarity.

The x-y-shift translation vectors are included in a new supplementary material (Table) related to this manuscript.

P9L255: What means ca? Please check this. Moreover, no information on this comparison was presented. Please provide Z differences of these 35 points.

We removed “ca”. Also, the Z differences of these 35 points are included in a new supplementary material (Table) related to this manuscript.

P9L261-262: Uncertainties of DEMs, as well as uncertainties propagated, thought DoD calculation should be shown.

We included the DEM evaluation in a new supplementary material (Table) related to this manuscript. As the DoD calculation is not the main facet in our contribution, we refer to the original work of Wheaton et al. (2010) for further details.

P9L265: how much is this area examination? Please presents values on square meters.

Added. 50 064 m².

P9L272: Did you measure the displacement of these 35 points? Additionally, these are the same used in section 3.4 (DEM assessment)?
Yes, they are the same. We made this more evident.

**P11L335-339**: Disagree. *I think this paragraph is too ambiguous because it depends on your goals. Recent developments of satellite photogrammetry (i.e. world-view 2 and 3, Pleiades, etc) showed the high quality of their products even on rock glaciers.*

In the framework of remote sensing approaches, our sentence makes a discussion about platform operation between classical aerial surveys or satellite programs (notably companies) and new UAV surveys (primarily researchers). We argued that UAV devices could help to close the gap between expensive commercial providers and small research initiatives (Eltner et al., 2016). We improved the paragraph accordingly.

**P11L242**: Disagree. Again, too ambiguous. *The UAV survey can early take more than 30 minutes because it is weather-dependent. In this case. In case the conditions are not optimal, the acquisition time can easily exceed 30 minutes. I suggest changing this part.*

To our experience, the typical UAV setup and pre-flight checks take less than 5 minutes. Since the geodetic measurements are fixed and already provided, this step is not needed for subsequent UAV surveys. In any case, we now indicated that the actual data capture from the first to the last images usually takes no more than 30 minutes (for the size of Tsarmine rock glacier), which is considerably less time than the regular GNSS data capture. Please, check Table 1, where the individual UAV survey times are indicated.

**P12L344**: Disagree. *The limits of detection depend on the size of the window used for correlation and it should be tested on fixed areas. Some information about these values is partially reported in Table 3, however, more detailed information needs to be provided.*

We improved the evaluation of sequential orthomosaics in stable areas and included the results in a new table. We also indicated that the estimated LoD takes into account the orthorectification and NCC uncertainties.

**P12L350-356**: This paragraph does not contribute anything to the study. *It is maybe a general discussion about “optimal conditions” but this topic is already known.*

Thanks for point out this. We included more precise details about the pros and cons. We made clear that the investigated area size, weather conditions and specific UAV regulations are all potential limiting factors for our study area.
P12L258-269: Disagree. This paragraph is redundant and a repetition of what the others already do. However, no technical comparisons are given about your protocol against the others (i.e. planning, processing, etc). I suggest revising or excluding this paragraph.

We agree that this section needs revision. However, from previous studies, we highlighted two main protocols issues that we have accounted for in the present contribution. The assessment of the stable areas (uncertainty estimation) and surface velocity validation by independent GNSS data were explained for five consecutive periods. We have indicated the most relevant survey and processing details, particularly in sections 2 and 3. Additional details are included on the new supplementary material (UAV surveys settings, DEM assessment and Helmert results).

P12L367-369: Ok, you are using stable terrain to validate your products but detailed information about these areas is not presented anywhere. Also, these areas are very small compared to the total extension of your DEMs. Please see my remark on P8L234-237.

Please see our previous comments about the stable terrain.

P12L375. The fact that you perform all working steps does not guarantee high-quality results. It should be proved by comparing independent methods. Additionally, a detailed report is needed to support this argument. I suggest revising or exclude this paragraph.

Agreed, we removed this sentence.

P13L404: "[...] net frontal retreats (Fig. 6)" reference is missing.

We added new references (Kääb and Reichmuth, 2005; Kummert et al., 2018).

P14L411: Surface velocities are showed instead of kinematic points. Please change this and also on Figure 9.

Changed “kinematic points” to “velocities values derived from...”

P14L428-430. Agree, but what is the contribution of this paragraph to your work? This
paragraph takes place at the introduction instead than in the discussion section. Please revise this or exclude this paragraph.

We excluded this sentence.

**P14L430-432:** I could not find an uncertainty analysis. You are relating this to GNSS comparison? The fact that you are using very high-resolution products needs a robust statistical analysis concerning the uncertainties between photogrammetric products notable on stable areas (for DEMs) and surface velocities.

Changed to “The quality of the high-resolution UAV-derived datasets leads us to stress that regular and spatially distributed surface velocity measurements would permit landform-wide kinematics to be better monitored.”

**P14L434:** What are the differences between your UAV protocol against others? More details are needed to be related to “customisable data acquisition”.

Our protocol is tested and validated on a rock glacier during a destabilised period. See our similar responses to reviewer 1 (AR1).

**P14L436-437:** Agree. With your data, you could explain more extensively this process regarding even into sub-periods.

Thanks for acknowledging this. We included this in the new supplementary material.

**P15L449-450:** Even this works shows a detailed description of UAV surveys on rock glaciers, I difficult to me what’s is the main difference against a normal UAV survey. The current status of the manuscript does not show in detail the relevant differences of SfM photogrammetric studies.

We showed a detailed description of a destabilised rock glacier with a suitable UAV monitoring strategy. We changed the sentence: “This work detailed the use of UAV surveys to monitor rapid kinematic and geomorphological changes on destabilised rock glacier with enhanced temporal and spatial details.” Please, see our similar responses to reviewer 1 (AR1).
P18L564: the list of references needs revision. The reference Fugazza [...] is not cited in the text.

Thanks for spotting! We added this reference to the manuscript.

**Figures and tables**

**Figure 1.** Panel b: add a north arrow to make clear the image.

Added

**Figure 2 Panel b.** Disagree with multitemporal co-registration. Please see my remark in general comments.

Changed to “Sequential SfM processing using four permanent GCP outside the rock glacier.”

**Figure 4:** I think that the community has much more interest to know which are the values of these limits of detection (LoD) instead than only the text. Please show specific values of LoD.

Agreed. Information about the mean LoD is indicated on each panel. Information about the LoD for each period is shown in Table 3.

**Figure 5.** This figure is nice. However, there is not always clear to relate a 3D view to a plane view as shown in Figure 4. I suggest changing the 3D image by a more explicit figure that shows the localisation of the longitudinal profile.

Thank you for the feedback. We included the localisation of the profile in Figure 1.

**Figure 6.** It is difficult to see front evolutions because the lines overlay. You said the mean “front line displacements” did you calculate this displacement over the entire line? If yes, precise which method did you use?
As suggested by the first reviewer, we modified the hillshade of this figure. We indicated the method on the figure caption.

Figure 7. These results could be showed better. Is too difficult to interpret this figure because even if the intervals are regular, there is not obvious to see ±0.8. The intervals should be divided by entire numbers (-1.0 - -2.0 as an example) instead of 0.8. Also, on the color scale, you say that you found values greater (less) than +6.8m (-6.8m), I suggest that you present the min and max values as well as the north arrow. Finally, please use a color scale adapted for colorblind people.

Thanks for the suggestions to improve our figure. As suggested by the first reviewer, we modified the categories accordingly and added the north arrow. Regarding the colour scale and the min/max values, we have followed the journal style for the representation of similar results. Please, see the relevant figures published recently on ESurf (Ulrich et al., 2021; Van Woerkom et al., 2019).

Figure 8. On panel a), please indicate the photo was acquired.

Agreed. Photograph acquired during the same day of the UAV survey (24 September 2019).

Table 1. Please provide the “GSD” acronym.

Thanks. We included “ground sample distance (GSD)” directly in the text.

Table 1. if it is a rigorous protocol, all flights were made between 11h-13h. Why was your last flight in the morning? I found no justification in the text. Please specify this. Did this early flight have any consequences on the velocity calculation? (important presence of shadows).

The UAV flight is typically performed simultaneously as the TGS or vice versa, but we must adjust the survey accordingly due to weather and personnel constraints. Please, see our previous response about the influence of “shadows”.

Table 2. Change “errors” by uncertainties.
We retain “errors” as this is the case for this type of statistics.

*Table 2.* Why are you expressing uncertainties in pixels? It should be better expressed in meters.

Changed to “meters.”

*Table 3.* Additional statistical analysis is required.

Thanks for this suggestion. However, we are not aware of such “additional statistical analysis” for this kind of data. We included the LoD values for the GNSS technique.

REFERENCES


