Dear Editor and Reviewers,

Thank you for your detailed comments and suggestions about our manuscript entitled "A high frequency and high resolution image time series of the Gornergletscher – Swiss Alps – derived from repeated UAV surveys".

We identified two main caveats noted by both reviewers. We answer hereafter to these concerns in a general manner with propositions of improvement, and then give detailed responses (with additional tests and quality checks when needed) to all comments in an itemized manner.

Response to the main concerns:

Absence of actual GCPs.

The first concern is related to the absence of actual GCPs to geo-reference our dataset, and to avoid internal deformations within the dataset.

We agree that the absence of GCPs can lead to some deformations in the master bundle adjustment (9^{th} June), in particular a doming effect due to an imperfect lens calibration. However, while these deformations can lead to significant absolute positioning errors, we do not think that they translate in large errors for relative measurements, and in particular for the computed ice velocities. Indeed, velocities depend on the relative positioning errors, which we argue are small here (see response to reviewer#2 for details). This is because: (1) the internal deformations are relatively small (max. few meters) thanks to a careful photogrammetric processing, (2) these deformations are smoothly distributed in space thanks to the bundle adjustment of each campaign, and (3) the deformations are almost constant in time thanks to the co-registration procedure. It follows that for most of the expected uses of the present dataset, which focus on change detection, the errors imputable to the internal deformations of the master bundle adjustment can be regarded as negligible compared to other sources of uncertainty. As an illustration, the error in horizontal velocities over a two-weeks period that can be imputable to the internal deformations is less than 1mm/day (see response to reviewer#2 for details), which we believe is acceptable.

Having said this, we acknowledge that these points were not discussed enough in the manuscript. Furthermore, internal deformations must be carefully described, assessed and documented in the description of the dataset. We therefore propose the following modifications of our manuscript to address the reviewers' comments:

- Deformation errors will be described in detail in a new paragraph at the end of section 3.1: generation of co-registered orthomosaics and DEMs. We will include the references suggested by the reviewers, and inform the reader about the impact of internal deformations on our dataset.

- In addition, we adopted solution 3 proposed by Reviewer#2 to assess the internal deformations that may exist in the master bundle adjustment (June 9th). We therefore co-registered our master bundle adjustment on an orthomosaic provided by the Swiss mapping agency, SwissTopo. This orthomosaic has a 50 cm resolution and has been acquired in 2009. This led to two interesting products, which will be added to the dataset and discussed in the revised manuscript: (1) a transform function allowing to geo-reference our dataset in the Swiss national reference frame (CH1903-LV03), and (2) a map of residuals between adjusted and observed tie-point coordinates allowing to assess and visualize the internal deformations of the master bundle adjustment.

Non-standard procedure to compute the Matching Maps.

The second concern addressed by both reviewers is the application of a non-standard procedure to compute the matching maps.

We argue that we did follow a best-practice procedure but adapted it to match the requirements of the present application (see response to reviewer#1 for details). We tracked within-scene motions by block matching, which is the basis of the vast majority of change detection software used for ice velocity tracking (e.g. CIAS, ImCorr, COSI-Corr, GoLive). Compared to these tools, our feature tracking strategy only differs by some technical details, mostly driven by: (1) the need to handle high resolution data acquired by UAV (10 cm resolution) in contrast to the relatively coarse satellite images (few meters resolution) for which the aforementioned software have been designed; and (2) the wish to match every pixel of the input ortho-mosaics in order to generate dense matching maps and not sparse velocity fields. To ensure that the outputs of our Matching Map Maker utility are robust, we carried out some additional tests, which show that the velocities computed by our utility are in line with the ones obtained by two well-established glacier surface tracking algorithms (see response to revewer#1 for the detailed results).

This said, we fully understand the reviewers when they ask for better assessment and description of our feature tracking approach. We therefore propose the following modifications of our manuscript to answer these requests:
- We will describe in more detail the procedure we used to compute the matching maps, with particular attention paid to put our approach in context with existing methods (using the references suggested by both reviewers). To motivate these choices, we will add a paragraph focusing on the specific requirements of UAV acquisitions, the differences compared to satellite applications, and their implications for feature tracking.
- Building on a comment of Reviewer#1, we will discuss the sensitivity of the estimated velocity fields in relation to the similarity score used in the matching procedure.
- We will add the results of the benchmarking tests mentioned above to a sub-repository of the Matching Map Maker utility, and mention the results of these tests in the revised version of the manuscript. We hope this will help the users to evaluate the performance of our utility with respect to other tools.

Responses to the comments of Reviewer #1:

In the following point-by-point reply, RC denotes a reviewer comment and AR denotes our response to the comment.

RC: The first concern is based on the relative geo-referencing, the authors have chosen not to use real ground control points. While they have chosen one common bundle-adjustment as master, to "stitch" other control against it. It is common in photogrammetry to distribute GCP’s and have them placed especially at the outer ends. This reduces "banana" bending, caused by imperfections in the lens model. However, this is not done in this study, thus one can assume such effects are here at hand as well. I am well aware of the logistics within such terrain, thus I am not asking to do this procedure. Nevertheless, I propose the authors put a bit more emphasis into describing the potential errors associated with this effect/shortcoming. I hope if this is done rightfully, it will reduce the mistake of over interpretation by other users of this data.

AR: We agree that internal deformations (e.g. banana bending) are possible, and we share the vision of Reviewer#1 that it is difficult to remediate to such deformations by measuring GCPs due to logistical constraints in high mountains.
We therefore completely approve the proposition to better describe potential errors, their amplitude, and their effect on the final products (i.e. orthomosaics, DEMs and matching maps) in order to inform the reader about the limitations of our dataset, and avoid over-interpretation of the possible uses.
To implement this idea, we propose to first discuss the internal deformations based on the literature in a new paragraph added at the end of section 'section 3.1: generation of co-registered orthomosaics and DEMs'. We will also discuss the impact of these internal deformations on the possible uses of our dataset.
To complement this theoretical error analysis, we also assessed the amplitude of the internal deformations in the master bundle adjustment of 9th June. To this end, we followed a proposition of Reviewer#2 and co-registered our master ortho-image on an orthomosaic produced by the Swiss mapping agency SwissTopo (and therefore properly referenced in the Swiss national reference frame). This orthomosaic has a 50 cm resolution and has been acquired in 2009. The residuals after co-registration provide information on the potential deformations within our master bundle adjustment (maximum few meters, see results of the co-registration in the answer to Reviewer#2). Thanks to this additional quality check, we will also be able to discuss the impact of the actual deformations of our dataset. We are aware that this procedure only assesses the horizontal deformations and that the vertical component can be more affected by the doming effects than the horizontal one, but for now we do not have access to a precise enough DEM to carry out a similar assessment on the vertical component. This will also be discussed in the new paragraph about internal deformations.

RC: My second concern is focused around the matching maps, which are less standard products, and therefore implementation details need to be discussed more. Although not standard, there do exist best-practices, and because the authors deviate from this I highlight some steps which may need more clarification, or adjustments, in order to improve the resulting matching map product or get a better understanding why certain steps are taken.

AR: We agree that matching maps must be better explained, with a particular focus on the procedure adopted to compute them. To this end, we will rewrite and improve the section '3.2 Surface displacement tracking: generation of Matching Maps' following three guidelines:
- Clearly state that Matching Maps are nothing but displacement maps.
- Contextualize our procedure with respect to well established software, highlighting similarities and differences.
- Benchmark our Matching Map Maker utility with respect to two well-established tools dedicated to displacement calculation by image correlation (we chose CIAS - https://www.mn.uio.no/geo/english/research/projects/icemass/cias/ and Imcorr - https://nsidc.org/data/velmap/imcorr.html as benchmarks). The results of the tests will be added to a sub-repository of the MMM source code, and will be mentioned in the main manuscript.

RC: The similarity score is "maximum absolute error" on grayscale images. This is a peculiar choice as throughout the season and throughout the day the sun has changing illumination directions. Though this similarity measure is very sensitive towards such effects. Commonly, the normalized cross correlation is used […].

AR: The similarity score we use is the minimum (not maximum) of the mean absolute difference between two patches. This similarity score is indeed uncommon for glacier image matching, but is

widely used in other domains and in particular for video tracking (e.g. Liu and Zaccarin, 1993; Chuang et al, 2014) because it is fast to compute, especially on large images using convolution. We therefore choose this score to accelerate the processing of our very large matching maps.
This said, we agree that this similarity score is in theory sensitive to illumination differences. However, in practice, we did not notice much adverse effects in shadow areas. This is mostly because the images were acquired roughly at the same time of the day (between 11:30 and 16:00), and because the orthoimages used to generate the matching maps are always separated by less than one month, which should mitigate the illumination differences.
To be sure that our choice of similarity score does not degrade the resulting matching maps, we tested two other scores in addition to the mean absolute differences, namely: (1) the normalized cross-correlation and (2) a high-pass filter (cutting wavelength=25m) coupled with a cross correlation (similar to (Fahnestock et al, 2016)). The test is carried out for the period July 13th – July 26th and focuses on a small area to limit the computing time. The results below (Fig. R1.1) show that the straight use of the normalized cross-correlation produces the less reliable results with many outliers (in white in the figure R1.1). In contrast, the cross correlation tuned for feature tracking in glacier context (i.e. with a preliminary filtering) and the mean absolute differences (the selected option for this manuscript) perform almost perfectly in the test area. In addition, one can notice that these two methods generate almost indistinguishable displacement maps, which tends to show that using mean absolute differences as similarity score results in robust displacement estimates.

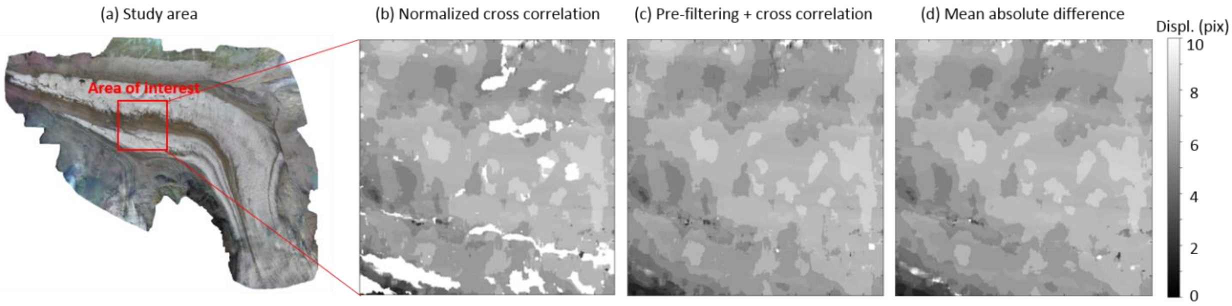

Figure R1.1: Surface displacements (July 13th – July 26th) estimated by feature tracking using different similarity scores. (a) Situation map. (b) Displacements estimated by maximum normalized cross-correlation. (c) Displacements estimated by high pass filtering (cutting wavelength = 25 m) and maximum cross-correlation. (d) Displacements estimated by minimum of the mean absolute differences. In (b-d), all parameters of the tracking algorithm besides the similarity score are the same for the three scenarios (search template 201 x 201 pixels, search window 400 x 400 pixels).

To make the reader aware of our choice of similarity score, we will expand the section '3.2 Surface displacement tracking: generation of Matching Maps' and mention the potential sensitivity of the Matching Map Maker utility to illumination differences. In addition, we will add the results of the tests shown above in a sub-repository of the Matching Map Maker utility and mention it in the manuscript. Finally, the MMM code in the repository has been updated to allow users to choose the similarity score used in feature tracking (the scores available are now: (1) normalized cross-correlation, (2) high pass filter + cross-correlation, and (3) mean absolute differences).

References:

Chuang M-C., Hwang J-N., Williams K., Towler R.: Tracking Live Fish from Low-Contrast and Low-Frame-Rate Stereo Videos, IEEE Transactions on Circuits and Systems for Video Technology, 25(1), 167-179, 2015.

Liu B., Zaccarin, A.: New Fast Algorithms for the Estimation of Block Motion Vectors, IEEE Transactions on Circuits and Systems for Video Technology, 3(2), 148-157, 1993.
Fahnestock M., Scambos T., Moon T., Gardner A., Haran T., Klinger M., Remote Sensing of Environment, 185, 84-94, 2016.

RC: […] the images are either pre-filtered with a high-pass-filter [Fahnestock 2016] or a Wallis-filter [Dehecq 2018]. I expect this will improve the results considerable.

AR: A first reason to pre-filter the images is the improvement of the normalization factor when the cross-correlation is used as similarity score. This is clearly visible in the results of the previous test assessing different similarity scores (Fig R1.1). But when mean absolute differences are used (as is the case in our manuscript), no normalization is required, and there is therefore no need for pre-filtering to help normalizing the similarity score.
Another argument for using pre-filtering is to focus on features with a characteristic scale close to the resolution of the images. We think that such use of pre-filtering is mostly relevant in case of matching satellite images with meter-scale resolution or coarser. In our case (10cm resolution), a high-pass-filter with a cutting wavelength close to the resolution of the images will extract very small features such as small superficial cracks, shadows or snow patches, which are probably not very persistent in time and thus not relevant to track. Instead, we prefer to use the full information of the patch of interest (5m x 5m) that also includes larger and more persistent features.

RC: Also the use of orientation-correlation [Heid 2012] or COSI-CORR [leprince 2007] might be a more robust procedure.

AR: To ensure that our Matching Map Maker utility is reliable, we benchmarked it by comparison with velocity maps calculated using the well-established CIAS and Imcorr programs. The results below (Fig. R1.2 and R1.3) show that our method does not introduce any bias or inconsistencies.

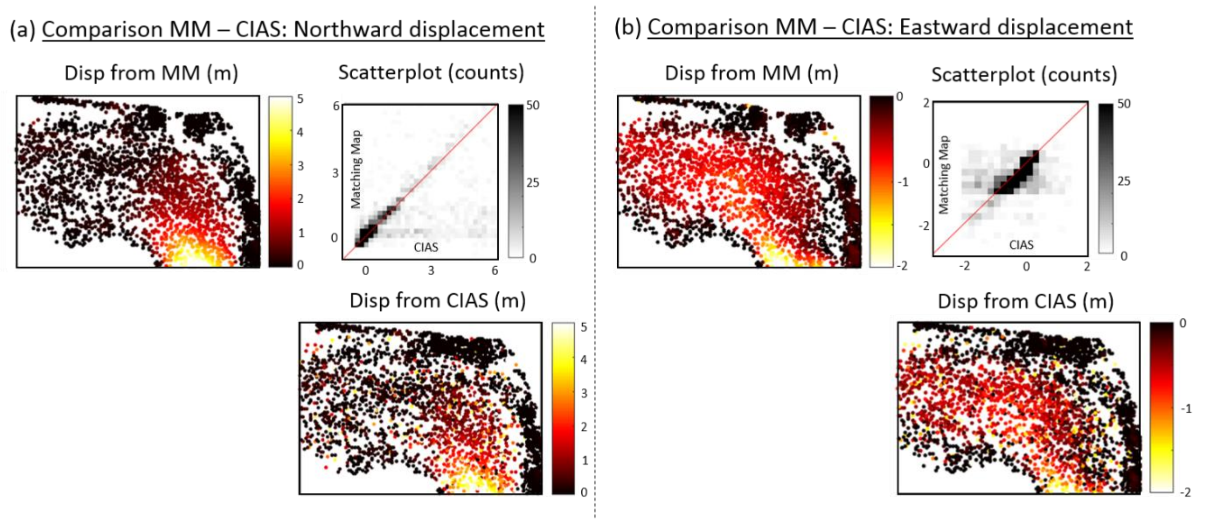

Figure R1.3: Comparison of estimated displacements derived from the Matching map (top plots) and from the CIAS software (bottom plots). Period of interest: July 13th – July 26th. (a) Northward displacement, (b) Eastward displacement. The only noticeable differences are due to residual noise in the CIAS results due to the simple set up we applied.

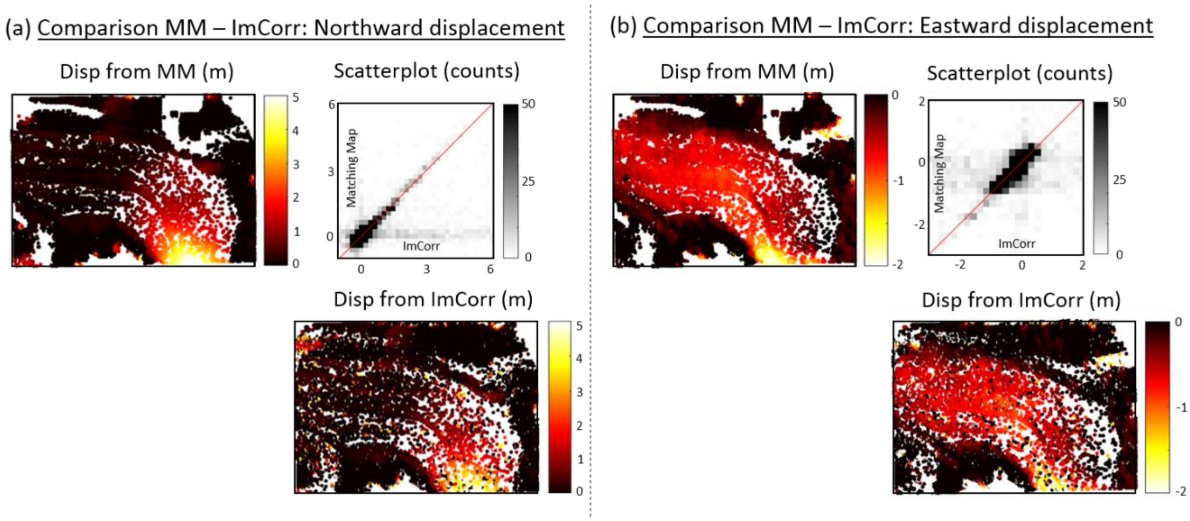

Figure R1.2: Comparison of estimated displacements derived from the Matching map (top plots) and from the ImCorr software (bottom plots). Period of interest: July 13th – July 26th. (a) Northward displacement, (b) Eastward displacement. The only noticeable differences are due to residual noise in the ImCorr results due to the simple set up we applied.

To reassure the user about the reliability of our utility, we will add the results above in the repository of the Matching Map Maker utility. In addition, we will mention the existence of this test (and the associated results) in the manuscript.

RC: The matching maps seem to have an integer displacement, and no sub pixel localization is applied. While these procedures are able to increase the precision considerably.

AR: We think that sub-pixel localization is mostly useful when the amplitude of the displacement between two dates is small with regard to the resolution of the images. This is the case for instance when glacier velocity is derived from satellite images, or when satellite images are used to measure ground deformations induced by tectonics (typical application of COSI-corr). However, in our case, the resolution of the ortho-images used to compute the matching maps is high (10 cm) compared to the expected displacement of the glacier (20-50 cm/day => 2.8-7 m displacement in two weeks). We therefore think that the absence of sub-pixel localization in our procedure is not a major issue.

RC: The spacing of velocity product is at pixel level, though in products like GoLIVE for example, the spacing of the grid is as large as the template size (in this case 300 meters). Similar processing in SAR speckle tracking is done with only 50% overlap [van wychen 2018]. This is done to have independent measurements, but now there will be a large smearing effect. It is possible to get to the resolution of a pixel, when pyramidal matching is applied, up to a point where optical flow can to be implemented.

AR: Yes the spacing of the Matching Maps is at the pixel level, and we think this is an advantage of this product rather than a drawback. Indeed, it allows the user to relate directly any pixel of a given orthomosaic to its counterpart in the following orthomosaic, without resorting to an interpolation of the velocity map. This makes the Matching Maps a convenient tool to navigate between the different acquisition dates.
But it is true that this also generates two drawbacks during the processing of the Matching Maps: (1) the matching procedure has to be applied to each pixel of the orthomosaic, which requires high computation. That is why we choose to boost the matching procedure by using the minimum of the

mean absolute difference as similarity score. (2) The results of the matching procedure are correlated for neighboring pixels. But a similar effect will appear if we compute sparse matching maps and then interpolate the results. And the advantage of the dense correlation is that the smoothness we introduce is not uniform. If there are large velocity contrasts, e.g. in a shear zone, this will be visible in our Matching Maps. In contrast, a sparse correlation coupled with an interpolation will probably miss it.
However, we agree that it is worth mentioning these aspects in the manuscript. We will therefore discuss these points in more details in section 3.2.

RC – minor comments.
AR: All minor comments will be implemented in the revised version of the manuscript. Hereafter we respond in details to three minor comments which we consider require a detailed answer.

RC: I assume all flights are nadir, or did the UAV also took oblique imagery. This is of interest, as it enhance the separation between internal parameters [james 2014].

AR: Due to the UAV platform we used (ebee from SenseFly), all flights are (unfortunately for lens calibration) nadir flights. We will specify it more clearly in section '2.2 Data acquisition' and discuss the implication on lens calibration in section '3.1 Generation of co-registered orthomosaics and DEMs'.

RC: "GCPs" is maybe not the correct term, as they are not real ground control, hence (manual) tie-points might be more correct.

AR: We fully agree with this terminology and we will adopt it in the revised version of the manuscript. We will also emphasize more the relative nature of our referencing.

RC: here seem to be multiple flights per campaign. Because a fixed wing is used, the landing must have had an impact. Hence, internal camera parameters might have been different between different flights. Thus, are these groups of data also separated in the camera optimization?

AR: Indeed, there are multiple flights per campaign, and only one set of internal parameters per campaign. Hence no separation of each group of data in the camera optimization. Since the SODA camera we used is specially designed for ebee flights (and associated possible rough landings) we supposed that the camera is robust to landing conditions, and therefore we expected that the camera parameters do not change much between flights.
To verify it, we plotted the camera parameters estimated for each date. In case of changing camera parameters induced by landing, different dates should have rather different estimated parameters. This is fortunately not the case. The estimated camera parameters are pretty stable along time (see figure R1.4 below), except one big 'jump' between July 13th and July 26th. This jump is due to the change of the camera, the device used for the first five acquisition dates having been destroyed in a UAV crash between these two dates during another (independent) project. It does not affect our data because the camera was never changed between the flights used for a single orthomosaic. In conclusion, we believe that the camera parameters are pretty stable along time, and are not much affected by the intra-campaign landings within a single day.

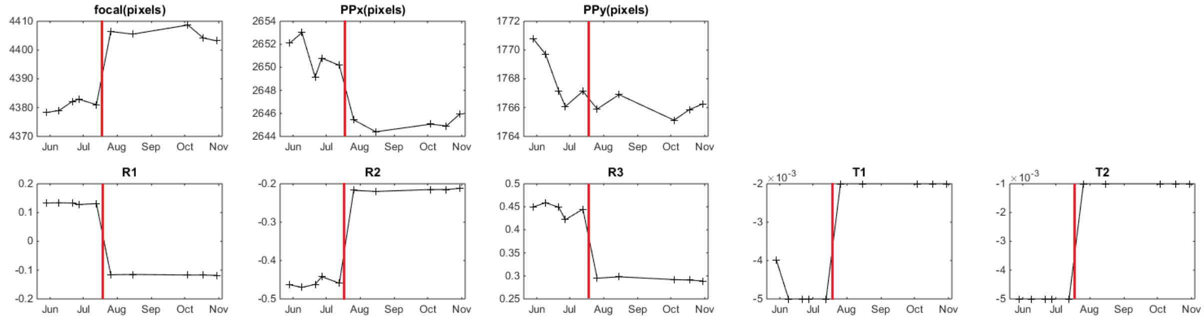

Figure R1.4: temporal variability of estimated internal camera parameters. Vertical red lines denote a change of camera (between 07/13 and 07/26).

Responses to the comments of the Editor:

In the following point-by-point reply, EC denotes an editor comment and AR denotes our response to the comment.

EC: I understand that you have used a commercial software package for the bulk part of the processing. This is ok, however, I believe the underlying algorithms should be explained/mentioned explicitly so that your study remains reproducible. For example, for people without access to pix4D Mapper, it is not helpful to state that the "default values" (p5/l5) were chosen.

AR: We agree with this comment, and we will therefore better explain the structure from motion (SfM) algorithms implemented in Pix4D, with references not only to the use of this software but also more general references explaining the use of SfM for glacier mapping.

EC: Fig. 5 has peculiar gray box behind that image which should be removed (maybed this is PDF viewer dependent? I have used Preview on MacOS.

AR : We will improve this figure.

EC: Consider changing the chosen colormaps to a sequential colormap. The rainbow colormaps (and their relatives) are prone to introduce false boundaries (e.g. https://www.climate-lab-book.ac.uk/2016/why-rainbow-colour-scales-can-be-misleading/ by Ed Hawkins).

AR: We will use a sequential colormap in our figures.

EC: Fig. 1c: Label Zwilings-GL is hard to read. Maybe move it to the heaven-part of the picture and use arrows? Not sure.

AR: We will improve this figure.