Reply to RC2

In this paper, the authors have provided a thorough overview of the technological and survey design considerations that impact the usability and accuracy of photogrammetric surveys and derived products. The authors then compare these considerations to the results of two prior surveys made on remote glaciers in Arctic Canada. I found the first "half" of the paper to be very well put-together and easy to read, which is no small feat for such a technical subject!

Because the surveys were (presumably) made before the start of the PhD that this work comes out of, it feels as though there is a slight disconnect between the review of the first half and the discussion in the second half - there isn't much indication that the lessons learned from the review informed the design of the field surveys, or that these lessons were "tested" by comparing the results of different approaches. This is entirely understandable, but I think that some additional critical reflection about how the surveys might have been improved would help to bridge this gap.
Thank you for the useful feedback. It's correct that the two parts of the study were undertaken at somewhat different times during the PhD, but we hope that the combination of a theoretical review combined with measurements under real-world conditions help to provide insights for how to optimize air photo surveys with field conditions that are often not as planned. To do this, we outline the weaknesses and discuss aspects that were improved or adjusted between the first and second survey, and those that should be improved in the future, but testing every factor, its impact on image quality and on the results of the bundle adjustment, is beyond the scope of this study. With multiple interacting factors contributing to the accuracy of topographic data, separating the impact of individual aspects is not straightforward, but we agree that it would be interesting to investigate in future work.

In response to the reviewer's comments, we have added a paragraph to the conclusion stating the main way the surveys could have been improved (lns 757–67):

> "Beyond any challenges inherent to the survey areas, logistical time constrains, and weather delays, the outcomes of the two surveys presented in this study were adversely impacted by unforeseen complications from system malfunctions (GNSS receiver defaulting to low logging rate) and human error (forgetting the intervalometer, accidental loss of focus). However, the main weakness was in the positioning performance being limited by the GNSS system which, despite consisting of dual-frequency, survey-grade components, was only able to record GPS observations (i.e., from a single satellite constellation). With continuous developments in GNSS technology, and multi-constellation/frequency equipment becoming increasingly accessible, upgrading the positioning system used here will likely result in fewer data gaps and significantly improve direct georeferencing accuracy in future surveys. Additional INS data could further help derive more precise control measurements taking into account variations in aircraft attitude. When coupled with precise synchronisation with the camera sensor, INS-aided GNSS navigation could help reduce uncertainties in camera position estimates and ensure more accurate topographic reconstructions."

Obviously, additional surveys are not feasible, but "simulation" of different settings after the fact (using ETTR images by shifting the histogram, using JPG instead of RAW images, not correcting chromatic aberrations, etc.) might be possible and give some indication of what impact these different choices have on the final product. This would help to solidify the connection between some of your proposed recommendations and the final results, though I only suggest doing this if it is easy to do - if not, some additional comparison to literature that has reviewed the impact of these different design choices is more than enough.
Unfortunately it's not possible to simulate the different settings as suggested (and access to our field sites is very difficult to repeat any surveys), so instead we have followed the reviewer's suggestions and included additional references to studies demonstrating the impact extreme lens distortions (Thoeni et al., 2014; Girod et al., 2017), rolling shutter effects (Vautherin et al., 2016; Zhou et al. 2020, Stark 2021), survey configuration and image capture geometry (Sanz-Ablanedo et al., 2020).

We also include an example from out dataset to illustrate the impact of image resolution on feature detection and tiepoint density which studies have shown to be directly related to result accuracy (lns 630–40):

> “Maximising image information content enhances feature detection and matching performance, resulting in higher point density, and significantly improving both reconstruction accuracy and precision (Gienko and Terry, 2014; Mosbrucker et al., 2017). In this study, feature detection was performed after upsampling all images by a factor of four. Increasing image resolution involves interpolating pixel values which can introduce artefacts (Rowlands, 2017), and is therefore only advantageous for high resolution and high quality images where it can be helpful with identifying more features and matching them with sub-pixel precision. Tested on a subset of 20 images from the EF dataset, the average number of features detected on upsampled images was 7–8 times greater than with the original resolution data (1.6 vs 0.2 million points per image), and the average number of valid matches increased by a factor of four (310,000 vs 75,000 points per image). The total point count in the corresponding clouds was five times greater based on the upsampled data (2.3 vs 0.46 million points each). This underlines the importance of selecting high performance imaging hardware maximising resolving power and image quality to ensure high point density, and in turn more accurate topographic reconstructions.”

Another consideration that I think could be better explored is the minimum feasible design. Obviously, a number of studies obtain usable results, even with less ideal camera setups (e.g., Welty et al., 2013; Girod et al., 2017), so at least some part of survey design depends on the needs or application. Given how rare the "ideal" situation is, some additional thought/reflection on what the "minimum" settings (or the non-negotiable choices) should be in order to achieve a useful output would be useful here.

This is a good point, and in response, we have added a paragraph in the discussion referencing other studies that have demonstrated adequate results using what could be considered the minimum non-negotiable camera setup, as well their key considerations for improving the system beyond the bare-minimum (lns 618–40):

> “The camera (Nikon D850) and lens (NIKKOR AF-S 24mm f/1.8G ED) used in this study were selected based on the key aspects discussed in Sect. 2.2, prioritising high resolving power, high SNR, and low geometric distortions. Low-cost compact cameras, including smartphones, action cameras, and cameras commonly mounted on consumer-grade RPAS, have been shown to provide adequate results for some applications, but comparisons with higher performance imaging systems show significant improvements in final outputs (e.g., Thoeni et al., 2014; Eltner and Schneider, 2015; Micheletti et al., 2015; Girod et al., 2017; Stark et al., 2021). With image resolution (GSD) being the main contributing factor to result accuracy, studies have also demonstrated the importance of maximising the level of detail captured (Mosbrucker et al., 2017), avoiding extreme distortions from very wide-angle (fisheye) lenses (Thoeni et al., 2014; Girod et al., 2017), minimising rolling-shutter effects with fast sensor readout speeds or global shutters (Vautherin et al., 2016; Zhou et al., 2020; Stark et al., 2021), and ensuring precise synchronisation between the camera and positioning system (Welty et al., 2013).”

We have also added a paragraph summarising the main priorities for survey planning and image capture settings (lns 663–72):

> “Image capture and survey planning are tightly related to the size, topography, and surface characteristics of the study area. For a given camera and lens combination, flying height will determine the GSD (Eq. 1), which should be selected based on the project and required level of detail, while keeping in mind the extent of the survey area and the flight time and number of images needed to provide full coverage with sufficient overlap. Other aspects should be configured to maximise the GRD while ensuring adequate exposure, keeping diffraction effects (Eq. 2) and motion blur (Eq. 4) to a minimum, ideally at or below 1.5 pixels. Maintaining image sharpness is crucial and, in low light conditions, decreasing flight velocity will allow faster shutter speeds while also reducing distortions from rolling shutter effects (Eq. 3). With large survey areas or limited flight time, where decreasing flight velocity is not possible, or with variable wind conditions causing sudden jumps in aircraft motion, it is preferable to increase ISO at the cost of introducing some noise, rather than risking blurry images.”

Specific comments

----

line 16: add "e.g." to Mölg and Bolch reference OK

line 67: why are helicopters preferred here? Can you explain the motivation/reasoning why this would be so? I would think that this will depend on the particular study area (size, remoteness), the available camera setup, and perhaps the prevailing weather conditions, but it's not clear to me why helicopters would be preferred in all cases.
The main point is that remotely piloted aircraft have several limitations which make them unsuitable to survey entire glaciers in certain conditions. At the latitudes of our study sites, and given the size of the glaciers that we want to measure, those unfavourable conditions are the norm, making manned aircraft the only option in almost all cases. Examples of issues that we have encountered in the field when trying (and failing) to use UAVs are frequent crashes due to proximity to the magnetic north pole, cold temperatures that severely limit range, and glaciers that are orders of magnitude larger than be covered by a drone.

As to why we use helicopters over fixed-wing aircraft, it mainly has to do with the logistics of chartering a fixed wing aircraft specifically for a scheduled survey 100s of km away from their home base, versus the flexibility of fitting in a survey flight in an existing field schedule with the helicopter stationed in camp. Fixed wing aircraft in the Canadian High Arctic also typically don't have a camera hatch or any way of taking photos out of an open window while airborne, and they can't perform the sharp turns at low elevations that we undertook in our surveys.

Table 1: can you provide, in either the caption or the text, an explanation of "high" and "low"?
Since the distinction is in the pixel count, and unrelated to sensor format, we removed "high" and "low" from the table column and added the following text in the caption:

> "The Nikon D850 fits in the FF category of newer sensors with relatively higher pixel count than the more standard 20–24 Mpx."

line 153: do you mean "stepping" instead of "stopping" here?
Here "stopping down" refers to decreasing the aperture by two stops, that is, increasing the f-number from f/4 to f/8. "stepping" up/down has another meaning and refers to fitting an adapter ring in front of the lens to use filters with a different thread size, but that isn't what we're referring to here.

line 165: perhaps "a combination" is a better word choice than "the product" to avoid confusion with the mathematical meaning of "product" - unless this is the meaning that is intended here.
"Product" is the correct word here, as multiplying the MTF of all separate components gives the system MTF.

line 206-7: I think this again depends on the needs of the study and equipment availability; it is possible to get usable results from "unsuitable" cameras.
Agreed. We added a paragraph in the discussion (lns 619–27) to underline this point (see the response to the last general comment above).

Figure 4: it might be useful to indicate the outline of (a2, b2) in (a1, b1). Added the outlines of a2/b2 in a1/b1.

line 359: insert space between "a.s.l." and "The" OK
line 530: remove stray "s" OK

References

Girod, L., Nuth, C., Kääb, A.M., Etzelmüller, B., Kohler, J., 2017. Terrain changes from images acquired on opportunistic flights by SfM photogrammetry. The Cryosphere 11, 827–840. https://doi.org/10.5194/tc-11-827-2017

Welty, E.Z., Bartholomaus, T.C., O'Neel, S., Pfeffer, W.T., 2013. Cameras as clocks. Journal of Glaciology 59, 275–286. https://doi.org/10.3189/2013JoG12J126