Answer to Michael Smith (Referee Comment #4)

One of your recurring comment seems to be that better gear and better accessory data (GCPs for instance) would make this method "not needed". This is absolutely true, however, such things are not always available (see other comments bellow) and the aim of this method is to circumvent the shortcomings of what is indeed available at the time of acquisition. We have tried to make this aim clearer in the paper.

This paper presents the idea of "opportunistic flights" where photos are acquired by allowing a camera to "hitch" a ride on a pre-booked research flight and using a low-tech approach to collect non-metric photos. The paper is well written and illustrated, with appropriate quantitative approaches. There are several points the authors make that are worth highlighting:

- *Errors are associated with the level of accuracy inherent in the method, the poor quality camera and a low contrast scene with high dynamic range*
- *The novelty lies in linking the timed GNSS points to the most appropriate photos. It's a good use of "extra" imagery and low cost acquisition of data that can bolster research is to be lauded. It is not too dissimilar to the original aims of SfM which had volunteered geographic information (VGI) in the form of photos scraped off websites as a source for point-cloud reconstruction. So, yes, there is NO novelty in using SfM to acquire (poor) quality photos of glaciated terrain and generate DEMs from them. The question is, is there enough novelty in synchronizing the camera and GNSS clocks to warrant a full paper?*

Section 2.1:

Covers the hardware and would benefit from examples of appropriate cameras (not in the conclusions), the estimated accuracy of the GNSS and optimal camera network design.

- **We abstained from recommending cameras at that stage because we did not test other cameras in this context and other issues could arise from them (stability issues, difficulty to mount to an aircraft...).**
- **The absolute accuracy of a code-based GNSS tracker in movement inside of a helicopter is not so easy to evaluate, and is fairly bad (few meters at best).**
- **The camera network design is not necessarily that controllable in the cases envisioned, our flight plans were only vaguely followed by the non-photogrammetric-flight-specialist pilot for instance. But as pointed out in the Referee Comment 3 (RC3), a straight line would be an issue. We added a comment on the necessity of at least some form of turn or zigzag in the flight trajectory.**

Section 2.3:

The authors have identified a lag in the EXIF time stamps of the photos on the GoPro used. Crucially, were any other cameras tested? Is this a problem on a Ricoh GR, Nikon Coolpix A or Sony A7? If it isn't, then this is a non-issue. Don't use the GoPro. This is a serious weakness as it simply demonstrates a numerical technique to overcome a limitation in a cheap camera.

Time stamps are never perfect, but the regularity of a time-lapse is indeed camera and memory card dependent. On "better" equipment, part of the method might indeed not be necessary. As pointed out in RC1, the SubSecTimeOriginal EXIF tag is present in some cameras (including the Nikon Coolpix A and Sony A7 cited here) and would attenuate the issue (if the tag is accurate and precise enough, which is not the case for all cameras that embed it). The advantage of our method is its robustness against cheap cameras, and such cameras might be the only ones available to an underfunded research project, or the only ones that someone would be willing to strap more or less precariously to an aircraft (I (LG) would personally refrain from ductaping a Sony A7 to the landing gear of an helicopter). Our method could also be very useful as a backup if the "good" camera/GNSS system fails (corrupted memory card, failing battery or other issues).

*The technique itself is well described, well illustrated and well implemented. There is actually no experimental design outlined – yes, you provide detail to the method used to correct the GNSS points, but at the start of section 2, give a brief paragraph in outline form detailing *exactly* how the data will be analysed and how it will be assessed.*

The introduction of the paper now contains a better overview of the outline in the following form :

> **We first present our generic method, then our test area, the specific survey (data acquired, specific equipment used, data processing, and additional datasets for comparison). We use this method to generate two DEMs from our low-cost equipment. The accuracy of the generated DEMs are analyzed based upon comparisons with professional surveyed DEMs. DEMs are compared in regions considered stable and statistics are used for quantifying their accuracy. In summary, we aim to show the improvements gained by including a few additional steps within the common SFM pipeline. We close with a brief initial geomorphological analysis and interpretation conducted from our products.**

A standard technique would be to test immobile points on the image with known coordinates. Can these not be extracted from the UltraCam imagery in sufficient detail? Or, use GCPs extracted from the UltraCam imagery to perform the geocorrection? It's surprising that this hasn't been undertaken to see how it compares to the use of the Garmin GNSS data. Again, if this is sufficient then the correction isn't needed.

By design, we decided not to use GCP (either collected on the ground of from trustworthy data) in our process, as they may not be available when this method would be applied in the future. We used a model obtained from a much higher end system (the ICE cam) to estimate the quality of our results.

Section 4.1:

Provide estimates of the pixel size for the flight heights, along with estimates of motion blur using standard photogrammetric methods.

The original images should be made available and they should be summarized (statistically) to highlight the range of aperture, shutter speed and ISO settings. Also note the effective aperture and focal length.

A “Data availability” section is now added to the end of the paper, stating that the data is available on request. Given that the images add up to a data volume of over 4GB, they cannot be made available as supplementary data to the paper and we do not have a distribution system setup for this kind of data.

Estimate of pixel size, motion blur amplitude (reasonable and not noticeably affecting the results) as well as image statistics and comments on saturation are now added to that section in the form of the following table:

Parameter	2014 Survey	2015 Survey
Focal	2.8mm (35mm eq : 15mm) - Fish-Eye	
Aperture	f/2.8	
ISO	100	
Exposure Time	$1/1774 \pm 1/426$s	$1/596 \pm 1/289$s
Mean GSD	0.56m.pix^{-1}	0.40m.pix^{-1}
Blur	0.07 ± 0.03pix	0.63 ± 0.27pix

The camera is poor and ill-suited to the work you have used it for. You ideally need to use one of the cameras noted above or in your conclusions.

We are aware (and agree in the conclusion) that GoPros are ill suited to this kind of work but it is the camera that we had at the time of the (indeed opportunistic) survey in 2014. We however still managed to get acceptable results out of it, and we would expect that it is the type of equipment that a number of teams have available when in the field. A comment was added in the conclusion.

Did the sensor saturate on any of the photos?

No saturation was detected, but the absence of contrast on large area of fresh snow is obvious.

What is the dynamic range of the GoPro at the highest ISO settings you used?

ISO was always at 100, GoPro does not provide a Dynamic Range estimate (their marketing style is more composed of statement like “AMAZING” than numbers).

It would be useful to see a full list of all the estimated errors and where they come from.

See next comment.

Section 5:

Well presented but it focuses on DoD which both have errors associated with them. One of the big problems here is that you are dealing with “whole system” error, not just the contribution from the GNSS, which is the novel part of the paper. What is the experimental design to test for this?

We indeed deal with "whole system" errors (even multi-system errors), and the error sources are mixing themselves. Purely synthetic tests would show more clearly the error sources, but would lack "real life" conclusiveness. This real life conclusiveness is what is provided by the glacial/pro-glacial analysis.

Why not flip this around and use GCPs from the Ultracam imagery and assess the difference with the GoPro DEMs (in the same way as TS check points)?

This could indeed be done, but we preferred using independent data for the validation (Ultracam ICEcam DEMs). The figure bellow shows the improvement offered by the method.

The glacial examples are interesting in and of themselves but they don't add to the technical aspects of the paper and can be removed.

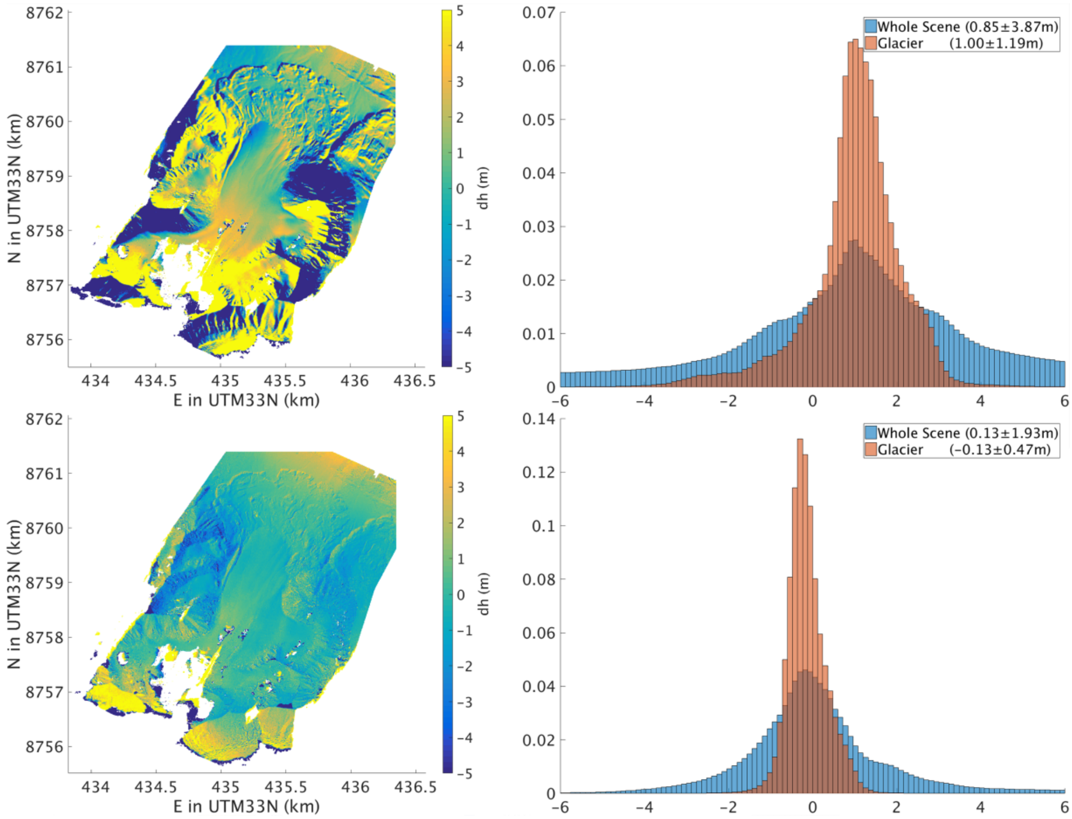
Again, what is provided by the glacial/pro-glacial analysis is a real life proof that actual geoscience data can be gathered with this method, not just accurate products to be archived. Moreover, one of the motivations for this study is to generate DEMs over mass balance glaciers for continuous estimates of the geodetic mass balance. Many glacier groups measuring mass balance require flight transport to their glaciers. In these cases, DEMs can be generated on almost an annual basis in order to constrain field measurements and biases that commonly accumulate over time (i.e. Zemp et al., 2013).

A table outlining all the DoDs generated would help to see what was compared with what.

All kind of comparisons were made; the ones that seemed of interest are used to illustrate diverse points along the paper.

*Overall the method used to correct the time lag introduced by the camera is elegant and well described, but the *effect* on the accuracy of the DEM is not demonstrated and the poor camera and large elements of error involved in various stages means that the conclusions that can be drawn are limited. And the method is possibly not needed if GCPs or a different camera are used.*

This was a great point, Thanks! We will add a comparison of the DEM produced using only a linear time delay to the EXIF times estimated from a picture of the GNSS tracker displaying just before the flight and with our method. We can see in the figure bellow that the un-refined data shows a clear distortion caused by the strain imposed by the bad geotagging. The application of a single refined delay as suggested by (Welty et al. 2013) improves the result dramatically already and our method only improves it slightly. However, the reduction of local strains in the bundle adjustment offered by our method is not insignificant and the end-of-line improvements on the DEM are not inexistent. They would also scale with aircraft speed (our survey was conducted at relatively slow speed). In any case, we believe that, as this paper is the first application of (Welty et al. 2013)'s method (in an improved form), the presented method is still novel and would still be if the method was not improved.



N in UTM33N (km)
E in UTM33N (km)
dh (m)
Whole Scene (0.85±3.87m)
Glacier (1.00±1.19m)
Whole Scene (0.13±1.93m)
Glacier (−0.13±0.47m)