Point-by-point response
Evaluating the Potential of PPK Direct Georeferencing for UAV-SfM Photogrammetry and Precise Topographic Mapping

He Zhang, Emilien Aldana-Jague, François Clapuyt, Florian Wilken, Veerle Vanacker, Kristof Van Oost

We really appreciate the suggestions that can make our work more valuable. We gratefully thank our reviewers and the editor for their careful advices and changed the manuscript accordingly. Please see the detailed answers (in italics) to the comments below:

Referee #1

General comments:

The paper presents analyses of topographic data acquired by UAV and SfM-MVS photogrammetry using a PPK-GNSS direct georeferencing approach. This is a technique with broad relevance to a wide range of disciplines because the method will become increasingly widespread.

We welcome this assessment.

Scientific comments:

1) Nevertheless, novel findings within the work are difficult to identify clearly and I haven't found the methods section sufficiently detailed to fully understand what has been done. The contribution of the work would be much clearer if existing similar work was evaluated more critically to provide a detailed context, and the aims and outcomes more concisely defined. Drawing more deeply on published work should allow statements of well-established principles (such as "camera properties have an impact on the accuracy") to be removed from the key sections such as abstract, discussion and conclusions so that the new findings can be more clearly communicated. The work is of interest but insufficiently described and, currently, the paper is somewhat challenging to assimilate. Overall, my suggestions below are aimed at highlighting the most transferrable new results from the work, by downplaying areas that have been previously covered and extending discussion to explore the underlying concepts further.

We agree that a substantial rewrite of the paper following the suggestions made here, will improve its focus, readability and will also more clearly identify the novel aspects. We do think that an assessment of the repeatability, reproducibility and efficiency of a PPK-SfM workflow in the context of 4D earth surface monitoring with time-lapse SfM photogrammetry is timely and highly relevant for geomorphological research. In our answer to the detailed comments below, we discuss how these improvements can be implemented.

2) With the paper focussed on PPK direct georeferencing for UAV surveys, the introduction would be well served by focussing on this. With UAV-SfM approaches not being so new, substantial regions of text (e.g. up to P 3), which introduce the broader aspects and uses of UAV photogrammetry could be condensed into a few sentences or a single paragraph. The introduction would be strengthened by incorporating Table A1 into the main text and critically evaluating the progress of PPK controlled UAV surveys so far. Consideration of established use of this approach for crewed aircraft could be covered briefly. Inclusion of the recent PPP work by Grayson et al. (2018; DOI: 10.1111/phor.12259) – and references included within it – will also strengthen this section.

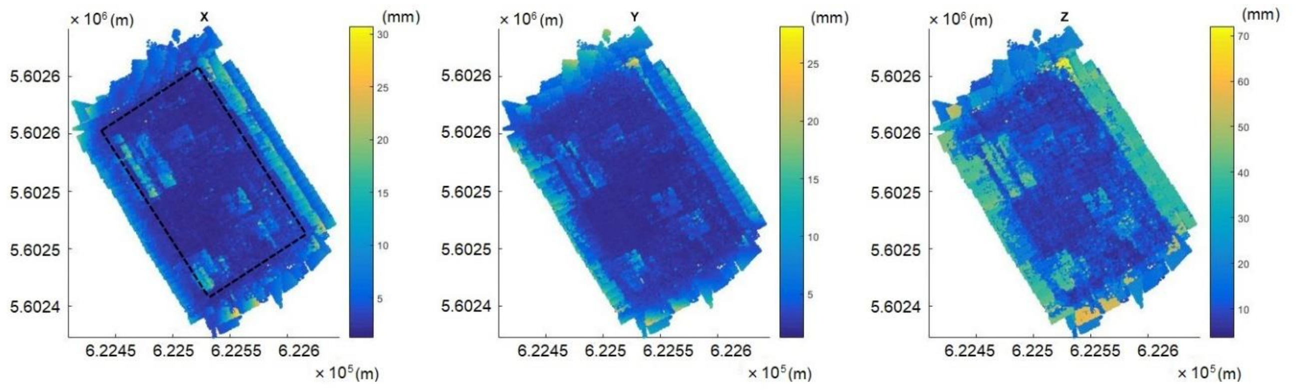
We will follow these suggestions. We will focus more on PPK direct georeferencing and remove the sections describing the UAV-SfM approach. Table A1 will be presented and discussed in the introduction to improve the description of the state-of-the-art in relation to PPK and the problem statement. The PPP work by Grayson et al. (2018; DOI: 10.1111/phor.12259) will be included in Table A1 as well as a PPK work by Padró et al. (10.1016/j.jag.2018.10.018).

3) One aspect of the work is an exploration of predictors for survey repeatability. The rationale for some of the selections could be strengthened here, and the utility of tie point density (which is shown to explain <50% of the variance for one camera) more critically considered. How useful is this, given that the analysis only appears to work for one camera and requires deployment of GCPs to determine the relationship? The number of tie points retained per image is usually a software setting that can be varied. Consequently, any parameterisation of repeatability would be software and UAV system dependent. Furthermore, other more important parameters are not considered. Within the bundle adjustment, measurement precision for a tie point is related to the number and angles of observations – how do these vary? The authors cite James et al. (2017) who show how point coordinate precision varies spatially and can be linked directly to these photogrammetric factors and other georeferencing factors within the adjustment. Consequently, maps of 3D point precision can be determined without any GCPs. The work here would be strengthened by discussing the authors 'spatialised error' approach in context with the 3D precision maps of James et al. (2017). The authors could also consider the findings of Mosbrucker et al (2017; DOI: 10.1002/esp.4066) within the discussion (or introduction).

The section were we presented a relation between tie point density and precision simply showed that it is a good indicator which can also be used to estimate the error patterns. We are aware of the fact that different sensors/flight patterns/surfaces/weather/illumination conditions could lead to differences in tie point magnitude. We agree with the author that the relationship requires deployment of GCPs and is empirical. However, we show below that the relationship is robust (see figure below) and could be used in temporal monitoring.

We suggest to substantially expand this section on tie point uncertainties by implementing the approach presented by James et al 2017 (based on monte carlo simulation) and by comparing this with our independent quantification of precision using GCP-based precision assessments. Following the workflow by James et al., 2017, we analysed DSLR camera (EOS) and action camera (GoPro) datasets, respectively. During the alignment process, georeferencing was achieved by PPK coordinates *without GCP reference. The Monte Carlo processing comprised 4000 iterations for each survey, and tie point uncertainties were calculated in X, Y and Z directions. The results are as follows:*

DSLR camera:



Action camera:

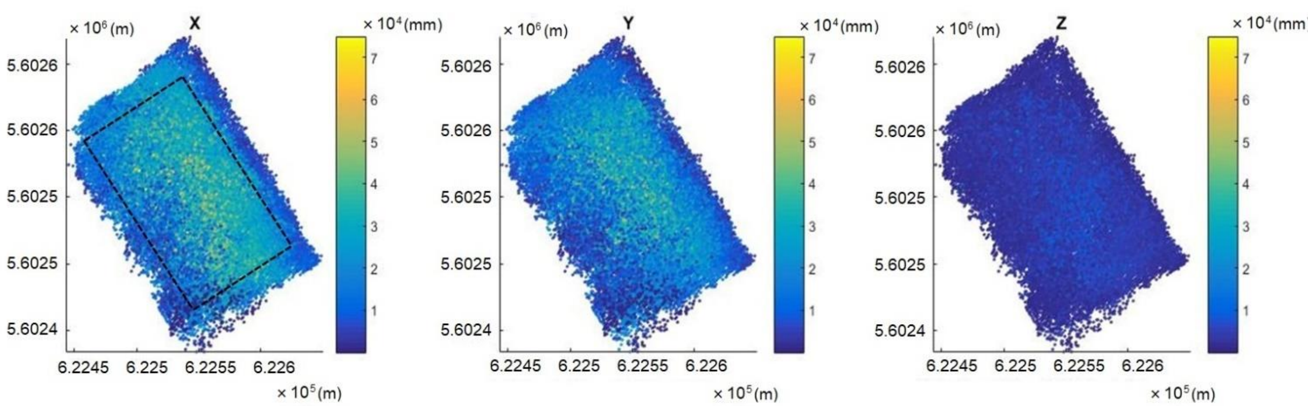
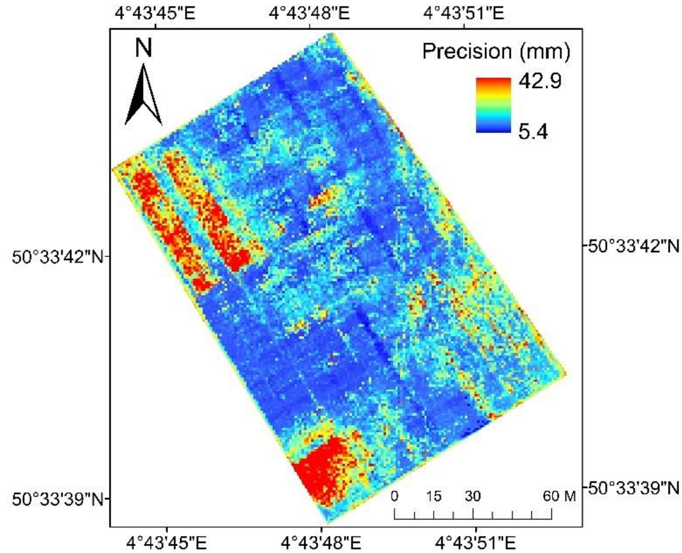


Figure. Precision and error maps for simulated UAV surveys. The region encompassed by the dashed line is the same processing area of the propagated error maps based on tie point density.

By comparing the two error maps derived from different approach, we observed highly consistent pattern of higher uncertainty/error area that can be distinguished from the DSLR camera dataset. Instead, the result of action camera showed no spatialized error pattern, which is also agree with the result that GoPro dataset was not spatially structured. To deeper interpret the Monte Carlo estimation, we calculated the XYZ precision (sXYZ) by points and then converted the points into raster (assigned by mean, resolution=0.92 m).

Precision estimated by Monte Carlo:



Precision estimated by tie point density:

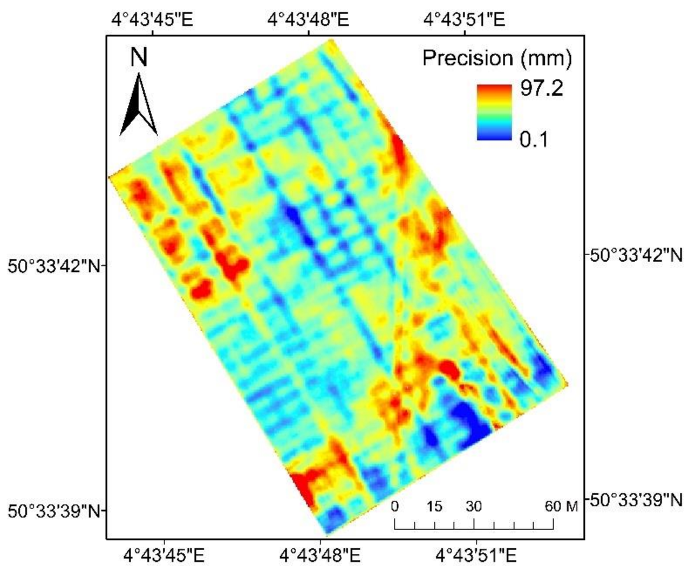
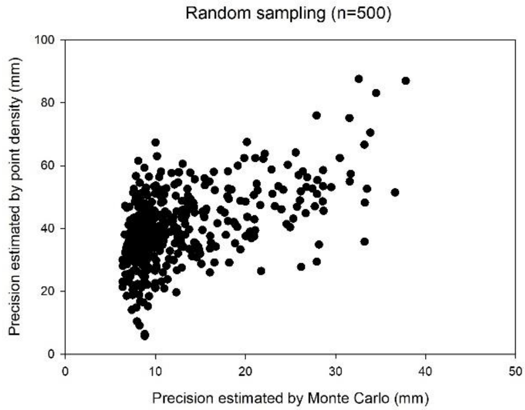
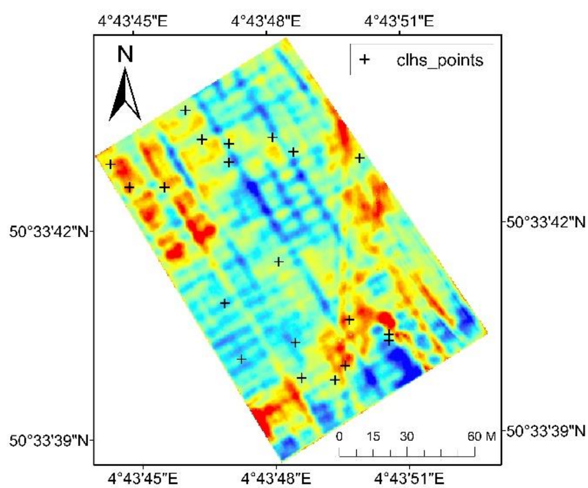


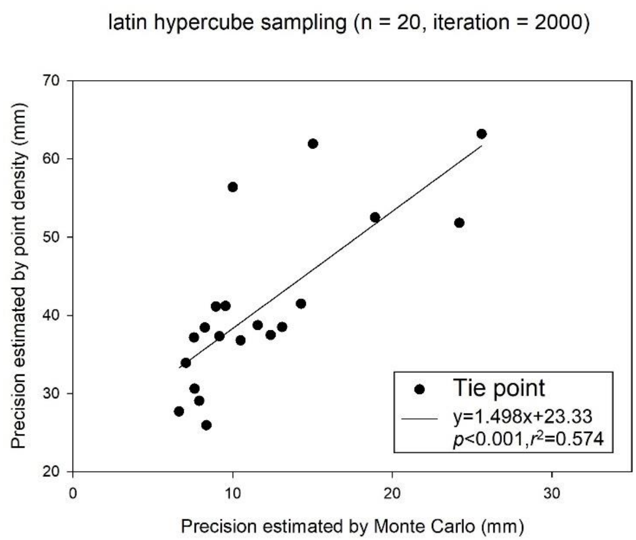
Figure. Precision map estimated based on Monte Carlo and tie point density. The XYZ error was estimated by the density-error regression model (y = -0.00041x + 0.125)

We observed similar shapes in vegetation area with higher error and the "artificial quadratic pattern" with less error in overlapping area between adjacent images. To test the consistency of the two estimation approaches, we randomly sampled 500 points within the region of interest and extracted values from the two maps, the results were as follows:



To better represent the real variability, we used Latin Hypercube sampling and selected 20 points to compare the two estimated datasets. The results were as follows:





Therefore, the estimation based on tie point density could be a quick alternative approach of evaluating tie point precision and spatializing the error.

4) Comparison of results from different cameras (S4.2; particularly the last paragraph) dominantly reflects established relationships between camera/flight parameters and conventional survey design principles. In my view, this material should form the rationale behind the survey design, and be given within the introduction or methods sections. Placing this within the discussion detracts from the newer aspects of the work (the PPK).

We will adjust the manuscript accordingly.

5) Throughout, when discussing results from different cameras, I suggest that a dimensionless approach based on pixels (or ground sampling distance) is also used. This could be used to assess the quality of the photogrammetric networks achieved, and to generate insight – again, see previous work, including that of Mosbrucker et al. (2017). I would actually see a much more detailed assessment of the PPK performance as providing the most useful (i.e. transferrable) insight.

Thank you! We will add the calculation of relative error based on GSD and the distance between surface and camera. The relative error is defined as the ratio between measured RMSE and surface to camera distance: $e_r = \frac{\sigma_m}{D}$*, where* e_r *is the relative error,* σ_m *the measured RMSE and D the mean distance between the camera and surface.*

Table 2 will be extended accordingly:

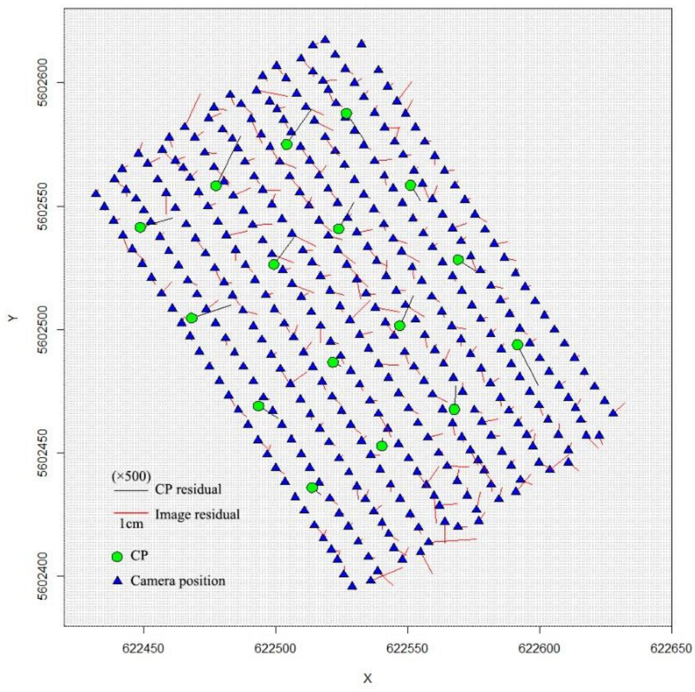
Camera	Flight Mission Date	Flight Height (m)	Mean(m)				RMSE(m)				Relative Error (px/m)	
			X	Y	XY	Z	X	Y	XY	Z	XY	Z
DSLR camera (EOS)	21, March	45	0.011	-0.011	0.016	0.012	0.016	0.015	0.022	0.017	0.071	0.055
	30, March	45	0.009	0.013	0.015	0.016	0.028	0.021	0.035	0.027	0.113	0.087
	05, April	45	0.005	0.014	0.014	0.003	0.014	0.02	0.024	0.022	0.077	0.071
	06, April	35	0.006	0	0.006	0.019	0.008	0.004	0.008	0.023	0.042	0.122
Action camera (GoPro)	21, March	45	0.018	0.021	0.027	0.049	0.028	0.031	0.042	0.076	0.039	0.071
	30, March	45	0.008	0.008	0.011	-0.042	0.016	0.015	0.022	0.051	0.021	0.048
	05, April	45	-0.01	0.008	0.013	0.009	0.014	0.013	0.019	0.024	0.018	0.022
	06, April	20	-0.013	0.02	0.024	-0.015	0.018	0.024	0.03	0.02	0.129	0.086

6) It would be good to see some more details to support the photogrammetric processing though – e.g. what were the rms image residual magnitudes? Did they vary image-to-image in any way that would help understanding of the repeatability? The clear image overlap outlines shown in Fig 6a suggest that camera positions may have been over-constrained in at least one survey (e.g. see a similar effect in Fig 1 of James et al. 2017a - http://dx.doi.org/10.1016/j.geomorph.2016.11.021, resulting from overweighting the GCPs in that case). Details of the a priori assigned camera position precisions used in the adjustments need to be provided and, given that they are often optimistically estimated, the effects of diluting the estimates could be explored.

We will add these analyses, including the effect of diluting the camera position precisions.

As for the image residuals, we selected two datasets from DSLR and action camera and depicted the planimetric residuals between calibrated position and original position by Pix4D SfM processing. The results are as follows:

DSLR camera:



Action camera:

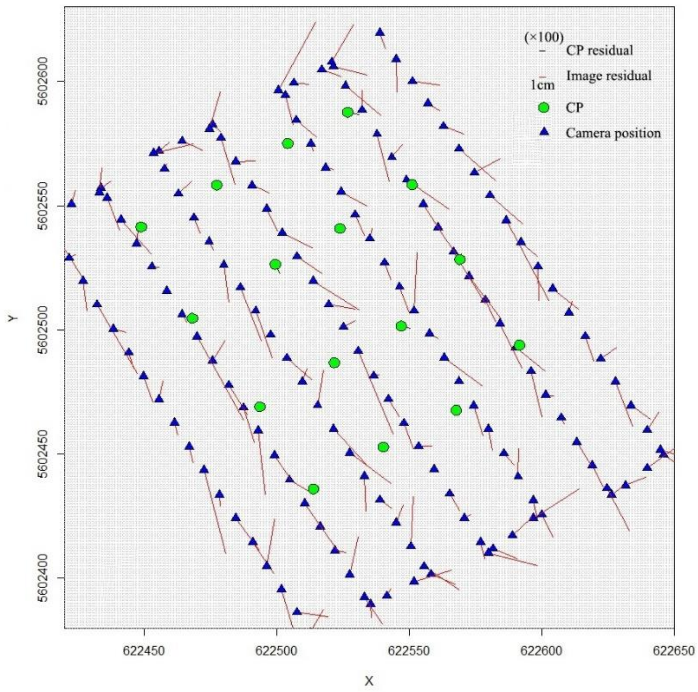


Figure. Residuals on the images and CPs in planimetric view (vectors give the horizontal residual component magnified by ×500 and ×100 for DSLR and action camera, respectively).

For the DSLR camera residuals, the outliers were mainly related to the points where the drone turned, while the rest were mostly within 1 cm (which is consistent with the RTK GPS precision). For the action camera dataset, the image residuals were larger while the CP residuals were not higher than those for the DSLR dataset. We ascribe this difference to the compensation of more convergent images from wider angle as discussed in the manuscript.

The clear overlap outlines (Fig 6a) were observed for the low FOV (field of view) camera. With a diagonal FOV of ca. 43º, one tie point can be observed simultaneously by 7-12 images for EOS camera in our case, while this is 60-80 images for the GoPro camera. This had an effect on the BBA process, where the GoPro output had no such overlap outline effects. Though the image overlap outlines shift in our case is ca. 2 cm, which can be eliminated by applying DoD threshold, we will investigate how image constraint rigidity impacts on the output accuracy as well as the quadratic outline pattern.

camera	*Vertical FOV*	*Horizontal FOV*	*Diagonal FOV*
EOS *Canon APS-C, lens 35 mm*	*24º*	*36º*	*43º*
GoPro *4×3 Wide (Zoom = 0%)*	*94.4º*	*122.6º*	*149.2º*

Fig. 2 Are photographs of the GPS system etc. really needed (c, d)? Much more valuable would be examples of the imagery processed (i.e. the underpinning data on which the work relies), with enlarged excerpts to illustrate the image quality and show how the GCPs have been imaged.

We will modify this figure and add the image-vision of GCPs.

All figures need to be checked for readability of the text labels. In particular, all map figures have scale and other labels which are far too small to be readable, and font sizes vary substantially across the figures. Labels must be readable: more consistent font sizes, of at least 9 point, will help.

Thank you for this comment, we will follow the suggestion.

Fig 4 Rephrase ‘error of detection’ for clarity.

DoDs – represent image overlaps etc for the DSLR but not for the action camera.

We will rephrase this concept.

Fig 7b LoD before/after lines indistinguishable – needs more careful visualisation.

We will follow the suggestion.

Fig. 8 Colour scale given to four decimal places could be tidied up.

We will follow the suggestion.

Table 1 The caption mentions three flights but I can only see data from two (i.e. one with DLSR, one with action camera). Which flights are these? Where are the results from the others? Table 2 The link to Table 1 is unclear. 05 April DLSR results are the same as in Table 1, but no similar repetition for Action camera. Maybe I haven't understood what Table 1 is?

We will add the information from all the flights conducted (in correspondence with data presented in Table 2).

Fig A1. I am not convinced how useful these visualisations are – it is difficult to extract much from them. I would suggest that a more informative plot would be as an XY map of discrepancy vectors, with symbols to indicate the check point position and Z-discrepancy showed by symbol colour. This way, any spatial systematics (which would be concealed in the current plots) would be clear.

We will follow the suggestion.

Table A2 This information is critical to a reader's understanding; it needs to be early in the main manuscript, not in the appendices. Why were some flights repeated? Where are these repeated data, and what did they show?

We will follow the suggestion.

Referee #2

General comments:

The paper "Evaluating the Potential of PPK direct Georeferencing for UAV-SfM Photogrammetry and Precise Topographic Mapping" fits the scope of the journal and I consider that the paper is very interesting for the Earth Surface Dynamics' readership. Moreover, it is a well-written paper, with very interesting results and rigorous validations. However, some minor revisions and comments must be fixed before the final publication:

Below, we show how we have revised the manuscript in light of these comments and recommendations.

Scientific comments:

1) Introduction (section 1) and Discussion (section 4.2): There is a very recent publication where it is compared the accuracy of different PPK approaches and other positioning alternatives, using DLSR cameras (10.1016/j.jag.2018.10.018). This could be in the introduction and in the discussion, since this research follows a similar workflow.

We will follow the suggestion and add this work by Padró et al. (10.1016/j.jag.2018.10.018). We will also include the PPP work by Grayson et al. (2018; DOI: 10.1111/phor.12259) in Table A1 to strengthen the advance in direct georeferencing.

2) P6 (section 2.3.2): Why did you not post-processed the static GNSS measurements?

The positioning measurement of GCPs was conducted using Reach RS in RTK mode, for which the differential correction data was transmitted via mobile IP network (the same approach for determining the base station coordinate). The Reach RS + GPS pole setup could provide a better control of antenna placement and reduce disturbance. We also tested the precision of the RTK solution by repeatedly measuring a fixed point near our department (12 km from the study field and 20 km from the BRUS station), and the error was ca. 0.010 m in XYZ. We thereby determined GCPs by RTK solution.

3) P7 (section 2.4.2): What was the interpolation method used in the DSM generation (TIN, bilinear, bicubic)?

It is the mean altitude from the point cloud data which is assigned to the DSM raster values. We will add this information in the revised manuscript.

4) P8 (section 2.5.2): How did you extracted the image coordinates? Could you detail the process (visually, number of iterations, . . .)?

After initial processing of pix4d, the software will generate a file with optimized position information based on the distortion model. The external camera parameters are given by:

T = (Tx, Ty, Tz) the position of the camera projection center in world coordinate system.

R the rotation matrix that defines the camera orientation with angles ω, φ, κ (PATB convention.)

If X = (X, Y, Z) is a 3D point in world coordinate system, its position X' = (X', Y', Z') in camera coordinate system is given by:

$$X' = R^T(X - T)$$

This was automatically computed in the initial processing, and the external camera parameters can be derived from the output "txt." file.

5) P10 (section 3.3) and Discussion (section 4.1): The authors explain and numerically detail the accuracy of several positioning procedures, but it would be interesting to compare them with a standard (e.g. ASPRS http://www.asprs.org/a/society/divisions/pad/Accuracy/Comments_NGTOC_Rev5_V1.docx), especially regarding the vegetated and non-vegetated terrain.

Thank you! We will strengthen this section.

6) P5 (section 2.3.1): Finally, the authors set the trigger interval in seconds, but they do not detail the rover velocity. Then, if the v is specified the reader could know how many meters lag between image captions and, if the GNSS rate is given,the distance between GNSS records.

The rover velocity is determined when the front-overlap, flight height, trigger interval and the camera parameter (especially horizontal FOV) are preset. Given that the trigger interval for DSLR camera (EOS) was 2 s and action camera (GoPro) was 4 s, the velocity was ca. 4.5 m/s and 2.8 m/s, respectively. The corresponding distance between each capture were 9 m for EOS and 11 m for GoPro. The image information can also be derived from the GPS logging or Pix4D output log after initial post-processing.

Referee #3

General comments:

The manuscript evaluated the repeatability of PPK UAV flight missions for precise topographic mapping. It is well structured and well written providing sufficient literature background and state-of-the-art methods. Results are presented from different perspectives and discussed broadly. The manuscript provides a contribution to the current debate of emerging PPK UAV data acquisition workflows that can be of interest to the readership of Earth Surface Dynamics. However, I encourage the authors to revise the manuscript based on some minor comments.

Scientific comments:

1) Compared to the entire manuscript, the introduction is very long, and some paragraphs could be more concise as most of the text builds on the general knowledgebase (e.g. general camera parameters, exterior orientation). Even though the authors stress the aims of the research, the novelty of this contribution is a bit fuzzy, as much research in this field has been done already. Efficiency is mentioned as one of the main objectives; however, there is little evidence on this in the results/discussion section as those parts mainly focus on repeatability/reproducibility.

We will compact the introduction and improve its focus, readability and will also identify the novel aspects more clearly (see also above).

2) Comparing metrical horizontal/vertical residuals of datasets with different GSD might not be the best approach, and normalized residuals could be more appropriate.

We will follow the suggestion and Referee #1 has similar comment to use a dimensionless approach. We will add the calculation of relative error based on GSD and the distance between surface and camera.

"The relative error is defined as the ratio between measured RMSE and surface to camera distance: $e_r = \frac{\sigma_m}{D}$*, where* e_r *is the relative error,* σ_m *the measured RMSE and D the mean distance between the camera and surface."*

Table 2 will be extended accordingly:

Camera	Flight Mission Date	Flight Height (m)	Mean(m)				RMSE(m)				Relative Error (px/m)	
			X	Y	XY	Z	X	Y	XY	Z	XY	Z
DSLR camera (EOS)	21, March	45	0.011	-0.011	0.016	0.012	0.016	0.015	0.022	0.017	0.071	0.055
	30, March	45	0.009	0.013	0.015	0.016	0.028	0.021	0.035	0.027	0.113	0.087
	05, April	45	0.005	0.014	0.014	0.003	0.014	0.02	0.024	0.022	0.077	0.071
	06, April	35	0.006	0	0.006	0.019	0.008	0.004	0.008	0.023	0.042	0.122

Action camera (GoPro)	21, March	45	0.018	0.021	0.027	0.049	0.028	0.031	0.042	0.076	0.039	0.071
	30, March	45	0.008	0.008	0.011	-0.042	0.016	0.015	0.022	0.051	0.021	0.048
	05, April	45	-0.01	0.008	0.013	0.009	0.014	0.013	0.019	0.024	0.018	0.022
	06, April	20	-0.013	0.02	0.024	-0.015	0.018	0.024	0.03	0.02	0.129	0.086

3) Line 3 page 6: Can you provide more information on the decision not to use a crossflight pattern or a single perpendicular strip as recommended by various authors?

One key aspect of our study is to provide practical tips for UAV survey, thus the autonomy (flight duration) is an important factor to take into account. A crosshatch pattern could improve accuracy by providing multi-angle images and increased overlap. However, it may double the survey time and battery consumption, and the promotion for accuracy is limited since the PPK enabled precise image georeferencing. To this end, we considered the tradeoff between survey accuracy and efficiency and opted the non-cross pattern throughout the experiment. It is also our motivation to investigate the large-FOV (field of view) action camera, so as to find a better solution for accuracy/efficiency ratio.

4) Did your UAVs also record attitude parameters? In the manuscript, you should make clear why you used some observations of the external orientation parameters and others not.

Sorry we didn't make this clear. Our cameras used a separated system without connecting to IMU, so the images only contained positional information without attitude parameters. We will rephrase the method section to make it clear.

5) You used Pix4D as a kind of black-box program. Which settings did you choose for camera calibration and the accuracy for geotagging information?

We will add details in 2.4.2 Point Cloud and DSM Generation. "The horizontal and vertical accuracy were both set to 0.05 m. We kept the remaining settings as default as 3D maps template, i.e., full keypoints image scale, automatic targeted number of keypoints and standard calibration method."

6) What are the reasons for the artificial quadratic pattern in the DoD of the DSLR in Figure 6a (upper left picture) – this almost looks like a kind of systematic error. This pattern needs to be explained in the text.

The artificial quadratic pattern of the DSLR camera were usually observed especially for the low-FOV camera. With a diagonal FOV ca. 43º, one tie point can be observed simultaneously by 7-12 images for EOS camera in our case. While we compared the output with the large-FOV camera GoPro and one single point can be captured by 60-80 images from wider angle. This had different effects on the BBA process, e.g., GoPro output had no such outline effects. Though the image overlap outlines shift in our case is ca. 2 cm, which can be eliminated by applying DoD threshold, we will investigate how image constraint rigidity impacts on the output accuracy as well as the quadratic outline pattern.

7) Some results are not very clear to the reader. I recommend extending the results section with some more explanation to enhance readability.

We will follow this suggestion.

Technical corrections:

1) Figures: Check consistency of font and readability of legends

We will follow the suggestion.

2) There are some typos and grammar mistakes – I recommend English proofreading prior to publication.

We will carefully correct these.