

Author response regarding re-review of: "Accuracy of snow depth estimation in mountain and prairie environments by an unmanned aerial vehicle"

By: Phillip Harder, Michael Schirmer, John Pomeroy, and Warren Helgason

The technical errors noted in the re-review by reviewer #2 have been addressed and additional typographical changes were made to the revised manuscript to remove errors and improve readability.

In light of reviewer 1 not re-reviewing our revised manuscript we are providing a more detailed response to their suggested edits as requested by the editor. The reviewer comments are in black, our general response to the comments are in red with the specific changes between versions highlighted in blue. The comments of reviewer 1 are difficult to identify specifically in the text as the comments were more general in nature and led to subtle changes throughout the paper as well as addition and rewriting of specific sections. To the best of our ability we will highlight what changes were performed with respect to reviewer 1's suggestions and comments.

Regarding General Comments Paragraph 1:

"More information, context, and discussion regarding the UAV system would help frame the results and conclusions presented in the study. The world of UAVs and their payloads is broad and quickly expanding. Given the diversity of aircraft, cameras, and processing techniques available, the authors should refrain from representing the results from one UAV system (theirs) as indicative of UAV / SfM snow estimation techniques as a whole. Quantitative results may be particular to the UAV system of choice. More discussion of how the choice of aircraft, payload, and processing software may have influenced results is needed."

We agree that articulating our results more clearly in terms of the platform that we use (fixed wing Ebee RTK) will help us to frame our results in the context of recent works that have used multirotor platforms and differentiate more clearly the results that come from unmanned and manned platforms.

In the introduction the results between manned and unmanned systems are more clearly communicated:

Previously (lines 107-114):

These examples have reported vertical accuracies (root mean square errors) from the manned platforms of 30 cm with horizontal resolution between 5-20 cm (Nolan et al., 2015) and 2 m (Bühler et al., 2015) and from the UAV 10 cm with a horizontal of resolution between 50 cm (Vander Jagt et al., 2015) and 10 cm (Bühler et al., 2016). The accuracy of assessment of the De Michele et al. (2015), Vander Jagt et al. (2015), and Bühler et al. (2016) studies were limited to a small number of snow depth maps, Bühler et al. (2016) had the most with four maps, and more

are needed to get a complete perspective on the performance of this technique and its repeatability.

Now (lines 131-139):

The manned aircraft examples have reported vertical accuracies of 10cm (Nolan et al., 2015) and 30 cm (Bühler et al., 2015) with horizontal resolutions of 5-20 cm (Nolan et al., 2015) and 2 m (Bühler et al., 2015). Unmanned aircraft examples have shown similar accuracies and resolution with vertical errors of reported to be ~10 cm with horizontal resolutions between 50 cm (Vander Jagt et al., 2015) and 10 cm (Bühler et al., 2016). The accuracy assessments of the De Michele et al. (2016), Vander Jagt et al. (2015), and Bühler et al. (2016) studies were limited to a small number of snow depth maps. Bühler et al. (2016) had the most with four maps, but more are needed to get a complete perspective on the performance of this technique and its repeatability under variable conditions.

An additional section was added to the discussion to place our results in the context of fixed wing vs. multirotor and manned versus un-manned systems

Added to end of section 3.3.1 (lines 330-344)

It is suggested that multirotor UAVs may be more stable and return better data products in windy conditions (Bühler, et al., 2016). There have not been any direct comparison studies that the authors are aware of that validate such assertions. A general statement regarding the use of fixed wing versus multirotor is also impossible with the broad spectrum of UAVs and their respective capabilities on the market. The only clear benefit of using a multirotor platform is that larger, potentially more sophisticated, sensors can be carried and landing accuracy is greater. That being said, the Ebee RTK returns data at resolutions that are more than sufficient for the purposes of this study (3cm pixel⁻¹), can cover much larger areas and has a higher wind resistance (>14 m s⁻¹ than many multirotor UAVs. Landing accuracy (±5 m) was also sufficient to locate a landing location in the complex topography of the alpine site. The more important issue relative to any comparison between platform types is that all UAVs will have limited flight times and results are compromised if conditions are windy and light is inconsistent. Until a direct platform comparison study is conducted this experience, and results of other recent studies (Vander Jagt et al., 2015; Bühler et al., 2016; De Michele et al., 2016), suggests that fixed wing platforms, relative to multi-rotor platforms, have similar accuracy and deployment constraints but a clear range advantage.

We also articulate that platform arguments are secondary to conditions at the time of deployment

Previously (lines 370-381)

Despite the limitations and deployment considerations discussed, UAVs are capable of providing data at unprecedented spatial and temporal resolutions that can advance understanding of snow processes. The most important consideration is whether the anticipated

signal-to-noise ratio will allow for direct estimates of snow depth or snow depth change. This limits the use of this technique to areas with snow depths or observable changes sufficiently larger than the SD of the error. This analysis established this threshold, at a minimum, to be ~30 cm. This threshold is equal to four times the mean observed SD (Rose criterion), but will vary with the application, site and user's error tolerance. Regardless of the accuracy of the absolute surface values, the relative variability within the DSM may offer fresh insights into the spatial variability of snow depth and snow surface roughness. Previous work on the statistical properties of snow depth (Deems et al., 2006; Shook and Gray, 1996) and snow surface roughness (Fassnacht et al., 2009; Manes et al., 2008) could be extended to consider even finer, centimetre-scale, variability over large areas.

Now (lines 384-395)

Despite the limitations and deployment considerations discussed, the Ebee RTK was capable of providing accurate data at very high spatial and temporal resolutions. A direct comparison between fixed wing and multirotor platforms is necessary to determine how snow depth errors may respond to variations in wind speed and lighting conditions. Until then, based on this experience and results of other recent studies (Vander Jagt et al., 2015; Bühler et al., 2016; De Michele et al., 2016), we do not expect there to be large differences in errors between platform type. Rather, the most important consideration when planning to map snow depth with a UAV should be whether the anticipated SNR will allow for direct estimates of snow depth or snow depth change. The SNR issue limits the use of this technique to areas with snow depths or observable changes sufficiently larger than the SD of the error. We propose a mean snow depth threshold of 30 cm is necessary to obtain meaningful information on snow depth distribution with current technology. This threshold is equal to four times the mean observed SD (Rose criterion), but will vary with the application, site and user's error tolerance.

More information is also included on the UAV platform and camera.

Previously (lines 152-167):

A Sensefly Ebee Real Time Kinematic (RTK) UAV (Fig. 2a) was used to collect imagery over both sites. It is marketed as a complete system, including the UAV platform and flight control and image processing software, capable of survey grade accuracy without the use of GCPs (Roze et al., 2014). The Ebee is a hand launched, fully autonomous, battery powered delta wing UAV with a wingspan of 96 cm and a weight of ~0.73 kg including payload. Maximum flight time is up to 45 minutes with cruising speeds between 40-90 km h⁻¹. A consumer grade camera, a Canon IXUS, captured imagery that was tagged with location and camera orientation information supplied by RTK corrected Global Navigation Satellite System (GNSS) positioning and IMU, respectively. A Leica GS15 base station supplied the RTK corrections to the UAV that resolve image locations to an accuracy of ± 2.5 cm. Bühler et al. (2015) found that snow depth mapping

improved with the use of near-infrared (NIR) imagery as the NIR spectrum is sensitive to variations in snow grain size and water content (Dozier and Painter, 2004), which increases the contrast and complexity of the snow surface. A NIR camera, a customized Canon S110, was also flown repeatedly during this campaign (three times at alpine site and 16 times at prairie site) and captured imagery in three bands; green, red and NIR (850 nm) bands. The Ebee was able to fly in all wind conditions attempted but image quality, location and orientation became inconsistent and/or was missed when wind speed at flight altitude approached or exceeded 14 m s⁻¹.

Now (lines 174-190):

A Sensefly Ebee Real Time Kinematic (RTK) UAV (version 01) was used to collect imagery over both sites (Fig. 2a). The platform is bundled with flight control and image processing software to provide a complete system capable of survey grade accuracy without the use of ground control points (GCPs) (Roze et al., 2014). The Ebee RTK is a hand launched, fully autonomous, battery powered, fixed wing UAV with a wingspan of 96 cm and a weight of ~0.73 kg including payload. Maximum flight time is up to 45 minutes with cruising speeds between 40-90 km h⁻¹. A modified consumer grade camera, a Canon PowerShot ELPH 110 HS, captures red, green and blue band imagery as triggered by the autopilot. The camera, fixed in the UAV body, lacks a stabilizing gimbal as often seen on multirotor UAVs, and upon image capture levels the entire platform and shuts off motor, to minimize vibration, resulting in consistent nadir image orientation. The camera has a 16.1 MP 1/2.3-inch CMOS sensor and stores images as JPEGs, resulting in images with 8-bit depth for the three color channels. Exposure settings are automatically adjusted based on a center weighted light metering. Images are geotagged with location and camera orientation information supplied by RTK corrected Global Navigation Satellite System (GNSS) positioning and IMU, respectively. A Leica GS15 base station supplied the RTK corrections to the Ebee to resolve image locations to an accuracy of ±2.5 cm. The Ebee was able to fly in all wind conditions attempted but image quality, location and orientation became inconsistent when wind speed at the flight altitude (as observe by an on-board pitot tube) approached 14 m s⁻¹.

Regarding General Comments Paragraph 2:

“For example, the Sensefly Ebee Real Time Kinematic aircraft was shown to be sensitive to wind speeds greater than 6 ms⁻¹. While this conclusion may be useful to future surveyors (i.e. it may not be worth their time to collect data on windy days), other platforms, such as rotary aircraft or even delta wings with more sophisticated autopilots, may be able to compensate and collect consistent data at higher wind speeds.”

A minor edit has been made to the wind sensitivity value. Initially a wind speed greater than 6 m s⁻¹ was reported to lead to an increase in DSM errors. Re-examination shows that any differences in DSM error with respect to wind speed were not larger for wind speeds up to 10 m s⁻¹ and this value is now used in the paper. This value is obviously platform specific. No changes to text were made with respect to this comment

“Do the authors recommend future surveys use rotary platforms?”

It has been suggested that multirotor UAV's may be more stable and return better data products in windy conditions (Bühler, et al., 2016). However, there have not been any direct comparison studies that the authors are aware of that validate such assertions. A general statement regarding the use of fixed wing vs. multirotor is challenged by the broad range of UAV designs and capabilities on the market. We see that the only clear benefit of using a multirotor platform is that larger, heavier, potentially more sophisticated, sensors can be carried (which may improve DSM accuracy as our camera's exposure settings were found to generate erroneous points) and landing accuracy is higher. Disadvantages of multirotor UAVs are that flight speeds and areal coverage are more limited than for fixed wing UAVs. We now note in the manuscript that the Ebee RTK returns data at resolutions that are more than sufficient for our purposes (3cm pixel⁻¹), can cover much larger areas and has a higher wind resistance (>14 m/s) than many multirotors – this seems to be a clear overall advantage. Landing accuracy (+/- 5 m) was also sufficient to locate a landing location in the complex topography of the alpine site. The more important issue relative to any comparison between platform types is that all UAVs will have limited flight times and results will be compromised if conditions are windy. A direct comparison between fixed wing and multirotor platforms is necessary to determine exactly how snow depth errors of various platforms may respond to variations in wind speed and lighting conditions. Until then, based on this experience and results of other recent studies (Vander Jagt et al., 2015; Bühler et al., 2016; De Michele et al., 2016), the sufficient image quality, reasonably good high-wind stability, suitable launching and landing procedures for alpine and prairie environments that are noted in the revised manuscript, in conjunction with the clear advantages in fixed wing range, may make fixed wing platforms preferable to the multi-rotor UAVs that have been described in the snow literature to date.

This is addressed as detailed in response to previous comments in lines 330-344 and 384-395

“Or does the decreased flight range / endurance of rotary aircraft compared to fixed wings outweigh the increased stability?”

This is platform specific but comparing this experience and results of other recent studies (Vander Jagt et al., 2015; Bühler et al., 2016; De Michele et al., 2016) would suggest that if the reported errors are similar than the increased range/ endurance of fixed wing platforms hold an advantage. That being said one cannot say anything with certainty without a direct side-by-side comparison. The manuscript has been amended with this discussion as noted above. (lines 330-344)

“How much of an operational concern is wind sensitivity, given that snow precipitation events and wind events frequently coincide?”

The reviewer does raise the concern that snow precipitation and wind events do sometimes coincide but those events should not be of concern as any UAV should not be flying in a snow

event and certainly not in a blowing snow storm because limited visible range (Pomeroy and Male, 1988) would make such operations illegal. Regulatory constraints (in Canada and other regions) restrict operations to visual line of sight, which is significantly hampered by snow in the atmosphere. Practically, airborne snow would significantly obscure surface features as seen from the UAV, reducing its ability to resolve the surface with SfM – there is no point in flying.

The most important consideration when planning to map snow depth with any UAV should be whether the anticipated signal to noise ratio will allow for direct estimates of snow depth or snow depth change. A discussion of platform type and its role in data quality that reflects these points is now in the revised manuscript. Lines 384-395

Regarding General Comments Paragraph 3:

“A similar discussion of the camera payload would be useful to readers as well. What is the specific model of the Canon IXUS used in this study? A quick Google search yields at least a dozen different models. What are the specifications of the camera? In particular, what is the bit depth? The point about the camera automatically adjusting exposure based on center-weighted values and overexposing some scenes, causing erroneous points, is important. More discussion of this type is useful – for example, that those planning a UAV snow survey should avoid cameras with automatic light metering. Also, the authors mention their system is not equipped with a stabilizing gimbal, which clearly increased wind sensitivity and decreased vertical accuracy. A 3-axis gimbal capable of maintaining an ideal camera orientation is a common feature of many consumer or “prosumer” level UAVs. A gimbal would certainly increase the quality of the SfM inputs, and therefore perhaps the snow depth resolution. Readers interested in snow, but perhaps UAV/SfM novices, would benefit from a more detailed discussion of the camera system used in the study.”

We concur that more details on the camera system would be beneficial. The camera a Canon PowerShot ELPH 110 HS, (which is the same as a Canon IXUS 125 HS) is used to capture red, green and blue band imagery and is modified to be triggered by the autopilot. Exposure settings are automatically adjusted based on a centre-weighted light metering and results may be improved in the future if one could manually adjust exposure settings (not possible with Canon ELPH). Most small fixed wing UAV’s do not employ a gimbal due to the space and weight requirements for such arrangements and in the case of the Ebee RTK the camera is fixed in the UAV body. To stabilize the camera when taking photos the UAV cuts power to the motor to minimize vibrations and levels the entire UAV resulting in consistent nadir image orientation. The camera has a 16.1 Mp 1/2.3-inch CMOS sensor and stores images as JPEGs, resulting in images with 8-bit depth for the three color channels. These details are now in the revised manuscript (lines 174-190) and addressed in the discussion of errors.

Previously (lines 295-303)

An attractive attribute of UAVs, relative to manned aerial or satellite platforms, is that they allow “on-demand” responsive data collection. While deployable under most conditions encountered, the significant variability in the DSM RMSEs is likely due to the environmental

factors at time of flight including wind conditions, sun angle, flight duration, cloud cover and cloud cover variability. In high wind conditions ($>14 \text{ m s}^{-1}$) the UAV struggled to maintain its preprogrammed flight path. This resulted in missed photos and inconsistent density in the generated point clouds. This UAV does not employ a gimbal to stabilize camera orientation and thus windy conditions also resulted in blurry images from the unstable platform that deviate from the ideal vertical orientation. The flights for the DSMs with the greatest RMSEs had the highest wind speeds as measured by the UAV.

Now (lines 309-319)

An attractive attribute of UAVs, versus manned aerial or satellite platforms, is that they allow “on-demand” responsive data collection. While deployable under most conditions encountered, the variability in the DSM RMSEs is likely due to the environmental factors at time of flight including wind conditions, sun angle, flight duration, cloud cover and cloud cover variability. In high wind conditions ($>14 \text{ m s}^{-1}$) the UAV struggled to maintain its preprogrammed flight path as it was blown off course when cutting power to take photos. This resulted in missed photos and inconsistent density in the generated point clouds. Without a gimballed camera, windy conditions also resulted in images that deviated from the ideal nadir orientation. The flights for the DSMs with the greatest RMSEs had the highest wind speeds as measured by the UAV. Four of the five problematic flights were due to high winds ($>10 \text{ m s}^{-1}$) and were identified by relatively low-density point clouds with significant gaps which rendered DSMs that did not reflect the snow surface characteristics.

Regarding General Comments Paragraph 5:

“Whether or not the erroneous points caused by overexposure are included in the authors’ results is unclear upon first reading. For example, section 3.1 (256 - 257) reads “These results exclude areas affected by erroneous points, as described in section 3.3.2, which was small compared to the total snow-covered area.” Which results are the authors referencing? Are the authors speaking generally about every single treatment? Or just the alpine-bare? For example, the authors should consider replacing “These results” with “The alpine-bare results” or “All results.” In general, an instance of the word “this” or “these” which lacks a referent can be confusing to the reader because they are unsure as to what precisely the writer is referring. After reading section 3.1 it seems the authors did not include the erroneous points for some or all of the results - but upon referencing section 3.3.2 (322 - 324) the reader finds conflicting information: “Erroneous points could be eliminated with the removal of overexposed images. However, reducing the number of images in such a large amount caused a larger bias and gaps in the point cloud, which made this method inappropriate.” Are the overexposed erroneous points included in the results or not? If the erroneous points are included, specify which results are impacted.”

We appreciate the reviewer’s identification of a confusing discussion on the identification and removal (or not) of erroneous points. This discussion has been simplified and limited to section

3.3.2. Some of the erroneous points encountered in early processing, only on alpine snow, coincide with snow surface measurement locations. On certain days, these errors limited the number of useful surface measurements. Incidentally, the erroneous points are located several metres above the surrounding surface, and thus are obvious and simple to exclude and so it does not make sense to include these in the error statistics.

The areas removed for each flight (as a percentage of the total snow covered area (SCA)) varied between 2% at the beginning of melt when the surface was predominantly snow-covered and 22% near the end of melt when a small number of snow patches persisted. The values of the removed SCA are now noted in the revised manuscript (added Table 3-see below). The point of this discussion was to note how we approached the errors in the hope of helping others who may encounter this issue in the future.

Previously (lines 314-340)

Erroneous points over snow were generated by post-processing with the default settings at the alpine sites. These points were up to several metres above the actual snow surface and were mainly located at the edge of snow patches, but also on irregular and steep snow surfaces in the middle of a snow patch. The worst cases occurred during clear sunny days over south-facing snow patches, where the whole snow patch was interspersed with these erroneous points. These points are related to the overexposure of snow pixels in the raw images, which typically occurred during direct sunlight over a small snow-covered area. A typical image with overexposed snow pixels had bare ground in the centre and small snow patches on the edges. The Canon IXUS camera automatically adjusts exposure based on centre-weighted light metering and is not adjustable. Erroneous points could be eliminated with the removal of overexposed images. However, reducing the number of images in such a large amount caused a larger bias and gaps in the point cloud, which made this method inappropriate.

The semi-global matching (SGM) option with optimization for 2.5D point clouds proved to be the best parameter setting within the post-processing software Postflight Terra 3D. Semi-global matching was employed to improve results on projects with low or uniform texture images, while the optimization for 2.5D removes points from the densified point cloud (SenseFly, 2015). The SGM option removed most of the erroneous points with best results if processing was limited to individual flights. Including images from additional perpendicular flights or merging subareas with overlapping images resulted in a rougher surface with more erroneous points. This is likely due to changes in the surface lighting conditions between flights, which challenges SfM. However, there was no additional bias introduced by the use of SGM and linear artefacts were visible when compared to default settings. These linear artefacts caused the standard deviation of the error to increase from 1 cm to 3 cm on bare ground. Areas with remaining erroneous points were identified and excluded from the presented analysis. The ability to reduce these erroneous points with SGM depended on the version of Postflight Terra 3D. Results achieved with version 3.4.46 were much better than results from the later version 4.0.81. This suggests that future users should test different versions to achieve optimal results.

The “black box” nature of this proprietary software and small number of adjustable parameters clearly limits the applications of this post-processing tool for scientific applications.

Now (lines 346-370)

Erroneous points over snow were generated in post-processing with the default software settings at the alpine site. These points were up to several metres above the actual snow surface and were mainly located at the edge of snow patches, but also on irregular and steep snow surfaces in the middle of a snow patch. The worst cases occurred during clear sunny days over south-facing snow patches, which were interspersed with these erroneous points. These points are related to the overexposure of snow pixels in the images which had bare ground in the centre and small snow patches on the edges. This is a consequence of the automatically adjusted exposure based on centre-weighted light metering of the Canon ELPH camera. It is recommended that erroneous points could be minimized with the removal of overexposed images; however, this increased the bias and led to gaps in the point cloud, which made this approach inappropriate.

The semi-global matching (SGM) option with optimization for 2.5D point clouds (point clouds with no overlapping points) proved to be the best parameter setting within the post-processing software Postflight Terra 3D. Semi-global matching was employed to improve results on projects with low or uniform texture images, while the optimization for 2.5D removes points from the densified point cloud (SenseFly, 2015). The SGM option removed most of the erroneous points with best results if processing was limited to individual flights. Including images from additional flights resulted in a rougher surface with more erroneous points. This may be caused by changes in the surface lighting conditions between flights. Biases did not change when using SGM though some linear artefacts were visible when compared to default settings. These linear artefacts caused the SD to increase from 1 cm to 3 cm on bare ground. Areas with remaining erroneous points were identified and excluded from the presented analysis. Table 3 summarises the extent of the areas removed with respect to the snow covered area at the alpine site. The fifth problematic flight identified (June 1, 2015flight over north area of alpine site) had a much larger bias with the inclusion of GCPs and the reason for this cannot be determined. The “black box” nature of this proprietary software and small number of adjustable parameters clearly limits the application of this post-processing tool for scientific purposes.

Table 3: Summary of areas excluded due to erroneous points with respect to snow covered area at Alpine site.

Flight ^a	Snow covered area (%)	Percentage of snow covered area excluded (%)
5-19_N	45.9	0.0
5-20_S	32.6	2.0
5-22_N	39.8	0.0
6-01_N	24.0	0.0

6-08_N	12.5	3.2
6-18_N	5.3	19.3
6-24_N	3.1	21.9
6-24_S	3.7	18.9

^amonth-day_portion of study area

Regarding General Comments Paragraph 6:

“Although the literature is sparse regarding SfM estimates of snow, the authors must be wary of comparing results derived from much different methods. For example, in the discussion section (286 - 292) the results are contrasted against the findings of Nolan et al. 2015, despite their methods using a manned aircraft. Similarly, Buhler et al. 2015 is a reference to a manned aircraft experiment. Given the topic sentence of this section begins “Differencing of UAV derived DSMs...” (emphasis mine) some readers may find the contrast of the authors’ results with that of a manned aircraft campaign misleading. Also, the 30 cm mean error reported by Nolan et al. is a geolocation error rather than a snow depth error. Snow depth errors were reported as 10 cm, and rigorously documented. Mean snow depths are not reported by Nolan et al. and thus as readers we cannot calculate or assess the SNR of his results, but it does seem like this study is suggesting a higher snow depth threshold for measurements than Nolan et al. That needs to be addressed.”

We agree it is important to differentiate that the imagery in this study was collected with a small fixed wing UAV rather than a multicopter or manned aircraft. The main difference between these studies is the collection platform, as application of SfM is fundamentally the same. Different processing software: Agisoft versus Pix4D Mapper versus Postflight Terra (and even between versions of Postflight Terra as we noticed) will give different results but the SfM principles are all the same. The differences in platform will lead to differences in the accuracy of image geotags and orientation, image resolution, bit depth and image overlaps. Regardless, very similar errors are being reported from the many recent studies applying SfM to snow despite the range of platforms and software being employed- this suggests to us that the greatest sources of uncertainty is the SfM procedure, followed by the differences in platform characteristics. The revised manuscript differentiates more clearly between the sources of uncertainty and the platforms used in the referenced studies (as noted in changes to introduction discussed earlier). The 30cm mean error that we attribute to snow depth error from Nolan et al. (2015) was an error and are grateful that the reviewer brought it to our attention. It is corrected in the revised manuscript.

Previously (lines 107-109)

These examples have reported vertical accuracies (root mean square errors) from the manned platforms of 30 cm with horizontal resolution between 5-20 cm (Nolan et al., 2015)

Now (lines 131-132)

The manned aircraft examples have reported vertical accuracies of 10cm (Nolan et al., 2015)....

Regarding Technical Corrections:

Both errors were typos, we thank the reviewer for noticing them, and they are corrected in the revised manuscript.

Regarding Overall recommendation:

"I recommend the authors make revisions to the paper based on the comments above. In general, the authors need to discuss the results appropriately with respect to the referenced work and use more precise language. Also, given the limited scope of this study, readers will prefer a considerably shorter paper. Striving for concision may improve the clarity of the paper as well. The paper could easily be shortened by up about 30%."

More precise language was implemented in the revised manuscript and we adjusted how we reference similar studies to more appropriately reflect their results and the platforms they used in contrast to this study (changes to introduction and discussion as already noted). Efforts were made to be more concise (deleting or substantially summarising lines 27-32, 94-95, 100-103, 115-118, 160-165, 176-181, 238-239, 266-269, 283-287, 343-345, 376-381, 395-396). The revised manuscript is a similar length than before and as there is additional information on the UAV platform, camera and more explicit discussion of this work with respect to platform type as requested as well as additions due to reviewer 2 comments.