The Cryosphere Discuss.,

doi:10.5194/tc-2016-11-RC1, 2016

Interactive comment on:

"Mapping snow depth in open alpine terrain from stereo satellite imagery"

by R. Marti et al.

Reviewer #2 : General comments

E. Thibert

Grenoble - 04/03/2016.

General comment

R. Marti and co-authors employ advanced methods of tri-stereo high resolution satellite imagery to retrieve DEMs and snow-cover thickness by DEMs differentiation in a mountainous open terrain. In a rigorous comparison with ground-based manual snow-probing, and UAV-derived DEMs differentiation, the authors find generally good agreement in these comparisons, the discrepancy been explainable by the natural scattering of the data, but also by some residual biases that remained unexplained. These favourable results should confer significant advances and improvements in mapping the snow mantle thickness over large open areas. The paper is almost clear, well organized, and properly focuses the scope of the journal.

We are very grateful to E. Thibert for his attentive revision of our manuscript. We agreed with most of its suggestions as detailed below in the point-by-point response.

(1) comments from Referees, (2) author's response, ***(3) author's changes in manuscript.***

The paper would be much more valuable were it also to provide a much clear presentation of the bias corrections in both xy and z-vertical directions for the DEMs derived from Pléiades and UAV flights. The resulting improvements in the residuals before/after adjustments are not clearly displayed (but dispersed along the text). A summary would be welcome with the numerical values of adjustments as an additional table or a supplementary column in Figure 4.

As suggested by the referee, we added the following table:

Table 3. Summary of the differents co-registrations and the bias corrections performed to produde the Pléiades and the UAV DEMs and dDEMS maps. SD means Standard Deviation.

Input data	Reference data	Type of coregistration	Values of adjustments	Comments
4 m-Pléiades winter DEM	4 m-Pléiades summer DEM	xy relative coregistration	-5.2 m North +2.8 m East	Workflow data Same shifts applied to the 1 m and 2 m-Pléiades winter DEMs
1 m-Pléiades winter ortho-image	1 m-Pléiades summer ortho-image	xy relative coregistration	-5.2 m North +3.2 m East	Verification data
1-2-4 m-Pléiades dDEMs	dDEM-snow free footbal field	z relative coregistration	$b_{1\,m} = -0.46\,m$ (SD=0.25 m) $b_{2\,m} = -0.48\,m$ (SD=0.20 m) $b_{4\,m} = -0.44\,m$ (SD=0.15 m)	Workflow data
2 m-Pléiades dDEMs	78 wide-spread points over snow-free areas	z relative coregistration	Median $b = -0.70\,m$ Mean $b = -0.74\,m$ SD $b = 0.26\,m$	Verification data
1 m-Pléiades summer ortho-image	6 wide-spread points on the 0.50 m-IGN ortho-image	xy absolute coregistration	+3 m North (SD=0.38 m) -0.8 m East (SD=0.35 m)	Workflow data Same shifts applied to the dDEMS
0.1 m-UAV-dDEM	353 wide-spread points over snow-free areas	ΔZ-correction based on a trend surface of order 3	RMSE: 0.34 m	Post-treatment correction. Same correction applied on the 1 m and 2 m-UAV dDEMs

Regarding the results of the comparisons between measurement methods, some questions remain:

- Is the remaining bias (-0.12 — -0.16 m) between the Pléiades estimation of snow thickness relative to manual probing significantly different from 0 from your error analysis?

If the reviewer agree with approach, we calculated a one-sample t-test on the residuals values to test the null hypothesis and it was rejected.

- Same question between Pleiades and UAV DEMs (-0.14 m) ?
A physical explanation should be attempted if you conclude theses biases to be significant.

As stated above, we calculated a one-sample t-test on the residuals values to test the null hypothesis and it was rejected.

Regarding specifically the UAV image acquisition and process from automatic correlation structure from motion, much more needs to be said about the image orientation, whether the camera model was estimated from a self-calibration in the process, or fixed from standard values or an independent prior calibration.

We agree that the method related to the UAV data was not sufficiently described. However, the UAV-images acquisition, the UAV-image processing, and the UAV-DEMs generation

were performed by a private company, and the statistics relative to the photogrammetric processes are not available. Thus, we are limited to comment the orientation residuals in line with the bias identified in the UAV dDEM. However, we now provide additional information, provided by the company, on the acquisition and processing of UAV data: number of images, camera model, focal length and a more detailed description of the method.
The main goal of our study is to evaluate if satellite data can play a role in snow depth mapping. To investigate this question, the 451 snow probe measurements is the foremost validation dataset. We did not use the UAV dataset to compute the Pléiades snow depth accuracy.

A minor point that arises throughout the paper is the indistinct use terms of error in systematic and random meanings, which can sometimes be confusing for the reader. To avoid confusion, I would as much as possible use systematic error, bias, and discrepancy to quantify the incorrectness, and random error and scattering to denote inaccuracy. Moreover, the random error should be defined with respect to the standard deviation (one or 2 times the standard deviation for example). This will help you to decide about the differences between the results to be notified as significant or not.

Thank you for these suggestion. As mentioned below (second substantive comment), maybe a part of that confusion comes from the use of precision and accuracy with a different meaning. As explained below, a strong bias in the residuals may be closely related to a bad accuracy, while a strong scattering in the residuals may be closely related to a bad precision. We agree with the reviewer to refer to the systematic and random errors terms to avoid confusion. We have checked systematically the use of the term "error" in line with this consideration throughout the manuscript and its supplement (also modified accordingly although not shown here):
P6-L21: "circular error": we refer here to the statistic as employed by Lebegue et al.2010 and by Gleyzes et al. 2013, and therefore we kept it as it.
P8-L8: "random error": use consistent with the reviewer commentary.
P8-L8: "systematic error": use consistent with the reviewer commentary.
P9- L7: "The **random** error on the DGPS measurements..."
P9-L8: "This term has a **random** error…"
P9-L9: "..the DGPS error": we deleted the term "error" here as it is redundant.
P9-L9: "Hence, two **random** error terms exist…"
P9-L10: "more details on the **random** error calculation".
P11-L9: "systematic error": use consistent with the reviewer commentary.
P11-L11:"random error": use consistent with the reviewer commentary.
P14-L28: "mean error": we replace the term here as follows "a mean of the residuals"

Random errors are combined assuming they are uncorrelated (page 5 of the supplement). This should be stated in the core of the text.

Thank you, this was added in the text:

P9-L7: **We assume all the randoms errors to be uncorrelated.**

An additional figure would be helpful, setting z-vertical direction and the different almost-horizontal surfaces Zs Zw from Pléiades and UAV and GPS, to define the notations of the variables, their differences, and better understand how standard deviations combine in the equations of the supplement.

To avoid to multiplicate the number of figure, we propose to improve the information available in the figure 4 (Pléiades worflow) in line with the new table 3 and the equation 1,2, and 3 to help to better understand the different notations in the text.

Here follow some detailed questions, comments, suggestions, and indications of minor typos in the paper.

Substantive comments

P1-L2. Specify ground resolution in meters for consistency with the rest of the paper (all lengths in meter elsewhere)

Thank you for that comment, we corrected the unit accordingly:
P1-L2: Two triplets of **0.70 m**-resolution images [...]

P1-L8. To me synonyms are accuracy and precision. Is it meant accuracy (random error) and correctness (bias or systematic error)?

The reviewer arises an important point: we do consider that accuracy and precision have different meaning. Accuracy is how close a measured value is to the actual (true) value. Precision is how close the measured values are to each other. As a bias may be defined as a systematic error which makes all measurements wrong by a certain amount, the precision is unaffected by a bias while accuracy does. In our case, we consider that the accuracy of the snow height (HS) measured by our method is characterized by the median of the residuals distribution. The precision may be assessed through the observed deviation in the residuals distribution characterized by the SD and NMAD values. In that statement, we consider all the HS values as a unique target.

P1-L16-18. I would set at first that snow cover is important in those areas for live and ecosystems, and second for anthropogenic needs.

We modify the beginning of the introduction in line with the reviewer's comment:
P1-L16-18: The seasonal snow cover in mountainous areas sustains mountain glaciers, alters frozen ground through its insulating effect, and plays a major role in mountainous ecosystems and plant survival (Keller et al., 2005). Snow cover is important for hydropower

production, irrigation, urban supply, risk assessment and recreation (Barnett et al., 2005).

P1-L20-21. The time for the seasonal snow thickness peak is very dependent on the elevation, varying from December at less than 1000 m a.s.l., to March in the 2000—2700 m range and April-May above 3000 m.

We agree with this comment. In the aforementioned sentence (P1-L20-21), we refer to the snowpack as a “water resource”, which implies a persistent snowpack during the favorable accumulation period (November--April). In the Pyrenees, the snow cover is persistent above 1 600 m-- 1 700 m, which correspond to the location of the 0°C isotherm between November--April (e.g. Lopez-Moreno, 2006). Only a very minor part of the Pyrenees are above the 3 000 m. Thus, the time for the seasonal snow thickness peak associated to a persistent snowpack in the Pyrenees is between March-April, although it is strongly affected by the interannual variability (e.g. Lopez-Moreno, 2013). To clarify this point and to take into account this comment of the reviewer, we propose to add the following sentence:
P1-L21: **In the Pyrenees, the accumulation peak associated to the persistent snow pack is generally between March--April (Lopez-Moreno, 2004, 2013)**.

P1-L23. The natural spatial variability of the snow cover thickness is preliminary due to the variability in precipitations, and post-deposition processes as wind drift, avalanches, snow densification.

To take into account this remark, and the remark of the other reviewer, we modified this sentence as follows:
P1L21--23: Even for small mountain catchments with areas of a few square kilometres, the spatial variability of the snow height and water equivalent is high because **of the elevation gradient of snow fall that is modified by the interaction of snow cover and topography, which leads to a large range of processes: preferential deposition of precipitation, redistribution of snow by wind, sloughing and avalanching (Grunewald,2014).**

P4-L20. use units in metres for consistency with the rest of the paper

Thank you for that remark, the units were corrected accordingly.

P4-L20. As far as I know, specify that this oversampling is carried out before image delivery

We modified the sentence accordingly:
P4-L20: The Pléiades’s pixel depth at acquisition is 12 bits, and the panchromatic images have an initial resolution of 0.70 m, but are oversampled at 0.50 m **before image delivery** by a post-processing algorithm that was implemented by the French Space Agency (CNES).

P4-L23. specify the local time.

We modified the sentences accordingly:

P4-L23: The snow-free acquisition was programmed on 26 October 2014 **(10:53:10, 10:53:31, and 10:53:52 LT).**

P4-L30: The second triplet was acquired on 11 March 2015 **(10:56:42, 10:57:03, and 10:57:27 LT).**

P5-L11—14. It should be clarified between winter and autumn surveys, what set up is used among RTK, ground control points.

In winter there was not GCP apart from the GPS base station, while in autumn there were GCPs. The RTK was activated in both cases.

P5-L23. The z-vertical correctness and accuracy is generally less than that of the xy plane, especially a vertical bias is unavoidable without a geoid model or an independent altimetry adjustment on levelling points.

The accuracy associated to the differential GPS positioning is much probably lower in the z-vertical estimation than in the xy-horizontal estimation, since z-positions are estimated only from positive z-satellite values. However, both accuracies are estimated being lower than 0.1 m after the post-treatment step. We exported the z-positions estimations as "heights relatives to an ellipsoid" and not relative to the sea level surface, thus a geoid model is unnecessary in that case.

P5-L28. Make it explicit whether your refer to influences in radiometry or ground surface roughness.

Although it could be an interesting possibility, we did not consider here the radiometry. The idea was to interpret the comparison between the Pléiades dDEM and the snow probe measurements according to the land cover classes. To clarify this point, we propose to modify the sentence as follows:

P5-L28: The vegetation types were aggregated into seven classes to reflect the type of land cover that may influence **the comparison between the Pléiades dDEM and the snow probe measurements.**

P6-L8. Figure 4 introduced before Figure 3 (in page 9 L14)?

Thanks for that remark, we swapped figures 3 and figure 4 accordingly.

P6-L8—11. Define RCP acronym just after "…Earth imagery that uses the RPC…" instead of the next sentence.

Thanks for that remark, we modified the text accordingly.

P6-L13. It should be clarified that each of the three resolution DEMs is retrieved from a rasterization of the 3 point cloud corpus (otherwise it may be understood that you derive one raster DEM per point cloud).

The description of our method is indeed something ambiguous as we omitted to mention that the three point clouds generated from the three stereoscopic pairs were merged before rasterization. Thus, we modified the text as follows:

P6-L13-14: We generated three point clouds from the three stereoscopic pairs from the *stereo command*, **and merged them**. The DEMs were rasterized at 1-m, 2-m and 4-m cell sizes from the **merged** raw point cloud through the *point2dem* command.

P6-L10. It should be clarified by a few words that RCP sets the image-to-ground geometry.

We propose to add the following sentence adapted from the Pléiades imagery user guide (Astrium, 2012):

P6-L10: **The RPC model is an analytical model, provided here as meta-data by Airbus Defense and Space (ADS), which gives a relationship between the image coordinates and the ground coordinates with z as the height above an ellipsoid, and which includes both a direct model (image to ground) and an indirect model (ground to image) (Astrium, 2012).**

We also modified the method description to be more consistent with the statements of the discussion part (P12 L26-27):

P6-L9: Spatio-triangulation was based on the RPC **model which was refined from an automated tie points generation without including ground control points (GCPs).**

P6-L15. I would rewrite as: "…grid point, with the Gaussian curve as weighting function…"

Thank you for the suggestion, we modified the text accordingly.

P6-L17. Which is the RMS residual in z after co-registration?

We present in the following table the statistic associated to the z-values before and after the co-registration process.

Statistics (N=5 944 407)	Before co-registration	After co-registration
Mean of the difference (m)	-3.43	-3.3
Median of the difference (m)	-1.51	-0.91
Standard deviation (m)	7.33	6.91
Normalized Median Absolute Deviation (NMAD)	5.21	3.75

We propose to include this table in the supplement if the reviewers or the editor consider it as relevant.

P6-L23. Is it meant that the same xy shift is applied to the higher resolution DEMs without proceeding to a new minimization? Why? And why the 4 m-resolution as the resolution to proceed for all DEMs?

Thank you for this comment. Yes, the same shift was applied to all the DEM (4m-, 2-m and 1-m) without proceeding to a new minimization at finer scale. Since all DEMs were generated from the same raw point cloud (see above), we estimated that the xy-shift should be the same. However, to check this assumption we have run again the minimization algorithm at 2 m and 1 m.

We obtained the following results:

2m DEM:
shift in N/S = 5.23 m
shift in E/W = -2.50 m
SD = 4.97 m

1m DEM:
shift in N/S = 5.44 m
shift in E/W = -2.45 m
SD = 4.99 m

As we expected, the best shifts are very similar at 1m, 2m and 4m (we obtained a shift in E/W of -2.83 m and 5.19 m in N/S at 4m). This was noted in the manuscript.

P5-L7—14. Please give more details about the UAV-on-board camera.
It is presumably a non-metric camera. Is it a fixed focal length lens camera? What is the focal length?

We added the following informations to the manuscript:

_in winter, 785 images during four parallels flights were acquired by a Canon IXUS 127 HS (4608 x 3465 pixels, sensor dimension: 6.170 mm x 4.628 mm). The focal length is 4.380 mm;

_during the summer campaign, 964 images during four parallels flights were acquired by a Sony DSC-WX220 (4896 x 3672 pixels, sensor dimension: 6.170 mm x 4.628 mm). The focal length is 4.572 mm.

The focal length is adjusted for each flight during the images processing, as well as the parameters used to modelize the lens distortion.

How on board RTK corrections to tag image centre coordinates and ground control points are used jointly for orientation, and maybe the camera self calibration? I suspect 5 ground control points insufficient for a calibration.

The referee is right that five control points were not enough, especially in this area with steep slopes. If we had to it again we would probably improve the setup of the survey.

P7-L3. Are you really sure that the IGN ortho has a more correct xy referencing than your DGPS ?

We associate an absolute (xy) georeferencing error of 2 m to the IGN orthophoto. Concerning the DGPS position, we consider an absolute xy localization error of 0.1 m. The problem is to interpret the DGPS position in the satellite images. As we did not use reference ground targets easy to photo interpret before the satellite image acquisitions, we are not able to interpret DGPS positions in the satellites images. Therefore, the IGN ortho-image is our unique source to provide an absolute georeferencing before the comparison between the Pléiades dDEM and the geolocalized snow measurements or the DGPS (x,y,z) coordinates.

P7-L6. Does that bias mean than the co-registration was not optimal? How compares this value with the co-registration residual?

We did not use ground control points (GCPs) during the spatio-triangulation step. Therefore both the summer and the winter DEMs contain a vertical bias as the DEMs are “floating” in (z) and have not been corrected yet. The (xy) co-registration step based on the DEM optimalshift method (P6-L17-20) do not aim at correcting such vertical bias. This is not a 3D-correction method. Therefore, the z-bias correction step is necessary in our workflow, after the xy-co-registration step to provide a final complete 3D co-registration. If the xy-co-registration step present large residuals, the search of a constant vertical bias over the entire image is complicated because it could vary significantly with the slope. However, it should not be affected over large flat areas.

P7-L8. Which photo? Does this refer to the satellite winter and autumn images?

The reviewer is right that the term "photo" is something ambiguous here. We replaced it as suggested:
P7-L8: where b is a constant vertical bias, which is determined from a unique, stable, and flat area of **the satellite winter and autumn images** that is easy to interpret.

P7-L11. I am not sure it is correct to remove negative snow depths as these values may not be significantly different from zero considering random errors in both you snow probing and dDEM calculations. They may well be acceptable in terms of confidence interval.
Removing some values is nevertheless conceivable considering they might be abnormal (irregular) if they discard negatively from zero at 2 times sigma (or more), according the way you define "aberrant" values.

We removed the negative values because we know that the snow height cannot be negative. This is the last step of the workflow to produce the snow height map. Then we evaluate the error on the snow map. Our primary objective is to evaluate the final product, not to provide a thorough error assessment of Pléiades DEMs or dDEMs. We asses the product that we would generate in an operational framework, or that we would distribute to hydrologists for example.

As answered to the reviewer Y. Bühler, the reviewer E. Thibert is right that it could potentially affect significantly the statistics of the residuals of the comparison between the Pléiades dDEM and the snow probes measurements. However as in indicated in the text (P7-L11) and in the table 3, it concerns only 8 to 10 occurrences (pixels) of the whole 451 snow probe measurements, therefore 2% or less of the number N of the whole validation dataset. We calculated here below the statistics considering the whole dataset (**in bold**) to assess the influence of removing the negative snow depths (in normal font) during the Pléiades dDEM assessment:

dDEM pixel size	Number of snow probing	Median (m)	Standard deviation	NMAD (m)
1 m	443 **/ 451**	-0.15 **/ -0.16**	0.62 / **0.61**	0.47 / **0.46**
2 m	442 / **451**	-0.16 / **-0.17**	0.58 / **0.58**	0.45 / **0.45**
4m	441 / **451**	-0.12 / **-0.13**	0.69 / **0.69**	0.51 / **0.50**

P7-L15--19. What is the result for the vertical bias calculated from these 78 points. Even if unused later in the paper, how does it compare to the bias calculated from the football field?

This important point is tackled in the section result "5.1. Pléiades and dDEM assessments":
P10-L25: The bias assessment which was performed over the entire Pléiades dDEM

(110 km²) and was based on 78 wide-spread values (see section 4.1) indicates a median of - 0.70 m, a mean of - 0.74 m and an SD of 0.26 m. The low SD value and the median difference confirm the possibility to remove a constant bias from a unique area, with 5 small random and systematic errors:
median(football field) - median(entire dDEM) = - 0.22m.

P7-L21—23. A bit more needs to be said about the image orientation (calibration?) process from then UAV acquisition.
Particularly, calibration for non-metric cameras is known to be critical and can generate significant orientation error when processed through automatic correlation Structure-from-Motion'based software as used here.
Which camera model is used for the orientation process? Did you used a simultaneous selfcalibration or a prior calibration? If calibrated, which camera parameters are estimated (decentration, radial distortion and associated polynomial coefficients—how many? , focal length)?

The UAV-data were treated by a private company, excepted the polynomial-based trend correction (cf. table 3). The photogrammetric software used was PIX4D, which uses a prior calibration of the camera model. The focal length and the lens distortion modelling parameters are adjusted for each flight during the photogrammetric treatment by the software.

We agree that this is a bit frustrating in the mark of a research approach, but the UAV-DEMs are not the main focus of our study, and were used only as comparison data. Recent works focus with a much more complete approach on this technics (Bühler, 2016; Harder, 2016).

Can you give the orientation residuals? This will help you to discuss about the discrepancy between Pléiades and UAV results to be significant or not.

Unfortunately, we do not dispose of such information. The UAV-treatment were performed by a private company (GeoFalco), which did not provide us the orientation residuals. According to the company, the software used (PIX4D) is a kind of "black box" with respect to the orientation residuals values.

P8-L9. It should be explained how probe [random error] is estimated as it is surprisingly high.

Snow probing in mountain area may be challenging and depend in part on the level of experience of the operator: i) an error of 0.05 m may be directly assigned to a reading error related to the graduations ii) it is difficult to maintain the snow probe perfectly vertical (e.g. a 15° inclination of the probe leads to a 0.01 m error at 3.2 m) iii) the vegetation by the snowpack introduce an extra and random error according to the fact that sometimes the bare ground is reached or not. Therefore, we propose to maintain this relative high random error associated to the snow probe sampling.

P8-L16. Instead of error, I would write "…is the median of the dDEM/probe discrepancies.".

The term "median of the errors" is commonly used in the literature to define the second term of the NMAD expression (e.g. (Berthier,2014), (Bühler, 2015)). We propose to use the term "median of the residuals".

P9-L11. Section 4.4.3. From the section 3.4, I expected here an analysis of the ground surface roughness effect, both on manual snow probing and in Zs uncertainty from the 2 DEMs.

A surface roughness index could indeed provide valuable information on the summer DEMs assessment. However, we do not focus on the seasonal DEMs evaluation. We aim at evaluate the DEM difference (dDEM) as it potentially contains the snow height information. We calculate statistics relative to the comparison between the 2m-Pléiades dDEM and the snow probe measurements for each of the five land cover classes. The idea here is to identify qualitatively a class that might introduce a systematic error (e.g. shrub compression by the snowpack) or a high random error that could be associated to a photogrammetric process issue.

P10-L31. How can you interpret physically this remaining systematic bias for the Pléiades snow thickness estimation? Is it significantly different from zero from your error analysis?

As stated above, we calculated a one-sample t-test on the residuals values to test the null hypothesis and it was rejected.

The z-bias correction of the Pléiades dDEM is characterized by a SD of 0.25 m considering the whole dDEM (see new table 3 above). Therefore, even on a large flat areas as the snow-off football field, it remains difficult to determine a unique b vertical offset and to coregister the summer and winter Pléaides DEMs in z. In that part of the image (football field), the resulting Pléaiades dDEM might be characterized by this systematic offset, which can lead to the observed underestimation of the HS values derived from the Pléiades dDEM.

P11-L5. Why did you force the intercept to be zero and did not fit to Y=aX+b in search of a systematic difference?

We assessed the systematic error through the residuals analysis (table 3). Here, the idea was to estimate how far the Pléiades and the UAV dDEMs are from the HS "signal" including both systematic and random errors. The linear model (y=ax) impedes the compensation of a systematic error instead of an affine model (y=ax+b) that introduces an extra b coefficient. In case of a strong bias, the correlation coefficient is thus affected applying the linear model.

P11-L11. I would expect a much lower value for the residual from the UAV DEM. Can you comment this in relation to the orientation residual of the images and the overall quality of the DEM geometry?

As stated above, we do not dispose of the orientation residuals and we are not able to comment the relation between the residual from the UAV **d**DEM and the UAV-images orientation residuals.

P11-L12.The subscript for the residual R is inconsistent with notations from equations 3, 5, 6 ?

Thank you for that remark, we corrected the residual expression to be consistent with the definition given by equation 3 in P10-L31 and in P11-L12.

P14-L8—17. I would also question the bias you identified here in relation to the quality of the geometry of the UAV DEM to originate from the orientation of your images.
More needs to be said about cameras, camera models, and the statistics of orientation results.
Instability is frequently associated to inaccurate or residual correlated camera model parameters after least-square adjustment (none unicity of solutions) which can result in poor quality of the geometry of the 3-D model (such as doming or bowl effects) after image orientation.

As stated in the manuscript and above, we are limited to provide valuable informations on the origin of the bias observed in the UAV dDEM, and our discussion remains merely speculative.

In this section, we refer now to recent works on HS mapping by rotor-UAV (Buhler et al. 2016) and winged UAV (Harder et al., 2016).

P14-L27. Is “dynamic” much more appropriated than “resolution” to denote the 12-bits depth?

The term “dynamic” seems indeed more appropriate, and we modified the sentence accordingly. However the expression “radiometric resolution” seems also correct and commonly employed in the literature (eg. Lee et al. 2008 (p.832) cited in the manuscript).

P14-L28. To make a better distinction between accuracy/correctness concepts, use bias instead of “error” when you refer to a systematic error. Or systematically qualify errors as random or systematic to avoid confusion.

Please consider the general answer above. Here, we cite the results of a publication (Lee et

al. 2008).
P14-L28: “mean error”: we replace the term here as follows “a mean of the residuals”

P15-L18. Inflect somewhat writing “…for clear-sky/limited cloud cover conditions…”

We modified the sentence as suggested:
P15-L18: As for all optical sensors, the main drawback of the Pléiades constellation is the need for clear-sky **or with limited cloud cover** conditions to obtain suitable images

P15-L34. “In hydrology and water resource applications, there remains…”

We completed the sentence as suggested.

P16-L8. This result is not trivial. The snow cover thickness you can estimate is not significantly different from that of the snowpack model, considering the overall uncertainty in both estimations.

As the snow cover thickness we estimated from Pléiades data and the snow cover thickness simulated by the snowpack model do not correspond to the same hydrological cycle, 2014-2015 and 2011-2012 respectively, we did not go further in that comparison.

P16-L26—29. I would mitigate/inflect your conclusions here mentioning that you nevertheless need an altimetry control/adjustment on a snow-free flat surface —as your z-vertical bias corrections demonstrate— that you have to infer on each satellite imagery. But this control surface can be located kilometres apart and at lower elevations.

The reviewer is right that several aspects of the conclusion should be mitigated here. We modified the text as follows:
P16-L26--29: Indeed, **the processing of the Pléiades data does not require mandatory field data like ground control points, although such reference measurements are always highly desirable**. **An adjustment on a snow-free flat surface, which can be located kilometres apart and at lower elevations, is needed to correct a vertical bias in the Pléiades DEMs difference.**

Supplement

It is not clear to me why probe appears in equations 3, 4 and 5. I would only expect this term in the uncertainty associated to the comparison of dDEMs with HS.

In equation 4 and 5, we propagate the error sources identified in equation 3. Snow probe error is present in the equation 3 because we evaluate the Pléiade and the UAV-snow free DEM from a Z_summer estimation based on the DGPS Z_winter value minus the snow probe height.

Figures

P23-Figure 1 — caption. “Pyrénées mountains. Bottom: Bassiès…”
Identically, in the legends of top right and main maps : “Bassiès”

Thank you for that remark, we corrected “Bassies” by “Bassiès” accordingly. “Pyrénées” is generally written “Pyrenees” in the english literature, so we would prefer keeping its english orthography.

P24-Figure 2 — caption. “Comparison of terrestrial oblique pictures taken by automatic cameras…”
The local time is mentioned here but not in the text.

Thank you for that remark, now the local time has been also mentioned in the text (please see above). We also added the word “terrestrial” as proposed by the reviewer.

P30-Figure 8 —left map. Add a title “Snow depth” at bottom right corner as you did for the 2 other maps at top left.

Thank you for that remark, the title “Snow depth” was added to the left map.

P32-Figure 10 — caption. Add plot colours “…and the 2m-Pléiades dDEM (black bars) according to the probe Id ranked in the ascending HS (red line) order (see equation section 4).”

Thank you for that remark, we added the “black bars” and “red line” in the caption.

P32-Figure 10 — labels. Point of notation: define the units for the residual errors as (m), and not (in m). Same for HS in the right-hand axis.

Thank you for that remark, we now define the units by (m) and not anymore by (in m).

Stylistic comments

P1-L1. At present…

As both expression seem equivalent in english, we would prefer to keep “To date”.

P1-L19. Snow Covered Area…. Snow Height

Thank you for that remark, we added uppercases in the acronym definition.

P1-L20. Snow Water Equivalent (SWE)

Thank you for that remark, we added uppercases in the acronym definition.

P4-L1. Bassiès and Pyrénées

As Bassiès has no english equivalent word in the english language, we would keep it in its french orthography. “Pyrénées” may be written “Pyrenees” in the english literature, so we would prefer keeping its english orthography.

P4-L7 “ 6.6°C and the mean annual precipitation is 1640 mm

Thank you for that remark, we completed the sentence accordingly.

P4-L9 “…and 25% by vegetation-free rock and bare soils.”

Thank you for that remark, we modified the sentence accordingly.

P4-L30. erase “which was”

We erased “which was” from the sentence.

P4-L30. “…115km², and centred on the Bassiès catchments, as achieved for snow-free images.”

Thank you for that remark, we completed the sentence accordingly.

P4-L32. “…-14° along track direction…”

Thank you for that remark, we corrected this grammatical aspect.

P5-L1. “…-6.4° across track direction.”

Thank you for that remark, we corrected this grammatical aspect.

P5-L9 “…mean Ground Sampling Distance (GSD)…”.

Thank you, we added uppercases in the acronym definition.

P5-L13. “…installed in a nearby mountain refuge…”

We modified this sentence by providing some more information:
P5-L13 [...] GPS-base, which was installed **on the flat dropping zone of the** mountain refuge during the survey.

P5-L16. “We collected up to 501 hand-probed…”

Thank you for that remark, we modified the text accordingly.

P5-L16. “…10 March 2015, at the time of the UAV survey…”

Thank you for that remark, we modified the text accordingly.

P5-L19. “…snow probes with lengths of 2.2m and 3.2m,…”

Thank you for that remark, we corrected this grammatical aspect.

P6-L3. “…to i) limit the areas potentially masked by the rugged topography…”

Thank you for that remark, we modified the text accordingly.

P6-L24. “…1-m resolution from their respective DEMs…”

Thank you for that remark, we corrected this typographical error.

P7-L24. “…snow probe”

Thank you for that remark, we corrected this typographical error.

P8-L1. “…the flat dropping zone of the mountain hut…”

Thank you for that remark, we modified “heliport” by “flat dropping zone”.

P8-L15. Here and in lines 5 and 18 on page 11, don’t know the correct sign to use for multiplication between dot and cross signs following the journal style ? my preference is cross…
It seems that several convention coexist to indicate the multiplication operator: dot (e.g. (Berthier, 2014)), or no sign at all (e.g. (Grünewald, 2010) or (Buhler,2015)). We propose to put no sign between variables to indicate a multiplication operator.

P9-L14. space to add between 3.2 m and (Fig.3).

Thank you for that remark, we corrected this typographical error.

P12-L27. “tie-point measurements, whose effect is equivalent…”

Thank you for that remark, we corrected this grammatical aspect.

P14-L4. space to erase after “safe”

Thank you for that remark, we corrected this typographical error.

P14-L28. In upper case letters “Motion Unit”

Thank you for that remark, we corrected the sentence accordingly.

P15-L3. Unclear sentence to correct, it seems that “by” is missing before Jagt et al. (2015)?

Thank you for that remark, we corrected the sentence accordingly:
P15-L3: A DSLR camera that was mounted on a UAV platform was used over a small mountainous terrain (0.07 km^2) with thick vegetation cover **by** Jagt et al. (2015) to map the snow depth [...]

P35-Table 3 — caption. Remind the reader what the acronym NMAD denotes.

Thank you for that remark, we added the NMAD acronym definition in table 3 (p.35) and table 4 (p.36).

P36-Table 4 — caption. It would be helpful to remind the reader that * denotes significant correlations as mentioned in Table 3 caption.

Thank you for that remark, we omitted to mention it in that table. We added the same sentence as in table 3:
P36 Tab 4: **Significant correlations (p values <0.05) are marked with asterisks.**

P36-Table 4. What does reversed brackets denote in the interval bin column?

Reverse brackets are used to exclude the limit of the interval bin, which allow to consider the values of the given variable (col.1 of table 4) only one time in the statistics calculation (col. 4-6).

Bibliographie

Astrium (Airbus Defense and Space), Pléiades Imagery User Guide, 2012.

Bühler, Y., Marty, M., Egli, L., Veitinger, J., Jonas, T., Thee, P., & Ginzler, C. (2015). Snow depth mapping in high-alpine catchments using digital photogrammetry. *The Cryosphere*, *9*(1), 229-243.

Grunewald, T., Schirmer, M., Mott, R., & Lehning, M. (2010). Spatial and temporal variability of snow depth and ablation rates in a small mountain catchment. *Cryosphere*, *4*(2), 215-225.

Lee, C. and Jones, S.: DEM creation of a snow covered surface using digital aerial photography, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37, 2008.

López-Moreno, J. I., & García-Ruiz, J. M. (2004). Influence of snow accumulation and snowmelt on streamflow in the central Spanish Pyrenees/Influence de l'accumulation et de la fonte de la neige sur les écoulements dans les Pyrénées centrales espagnoles. *Hydrological Sciences Journal*, *49*(5).

López-Moreno, J. I., Fassnacht, S. R., Heath, J. T., Musselman, K. N., Revuelto, J., Latron, J.,Moran-Tejeda E. & Jonas, T. (2013). Small scale spatial variability of snow density and depth over complex alpine terrain: Implications for estimating snow water equivalent. *Advances in Water Resources*, *55*, 40-52.

Poli, D., Remondino, F., Angiuli, E., & Agugiaro, G. (2015). Radiometric and geometric evaluation of GeoEye-1, WorldView-2 and Pléiades-1A stereo images for 3D information extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, *100*, 35-47.