Article

Centimetric Accuracy in Snow Depth Using Unmanned Aerial System Photogrammetry and a MultiStation

Francesco Avanzi 1,2,†, Alberto Bianchi 1,†, Alberto Cina 3,†, Carlo De Michele 1,5,†, Paolo Maschio 3,†, Diana Pagliari 1,†, Daniele Passoni 4,†, Livio Pinto 1,†, Marco Piras 3,†, and Lorenzo Rossi 1,†

1 Politecnico di Milano, Department of Civil and Environmental Engineering, Piazza Leonardo da Vinci 32, 20133 Milano, Italy; francesco.avanzi@polimi.it (F.A.); alberto.bianchi@polimi.it (A.B.); diana.pagliari@polimi.it (D.P.); livio.pinto@polimi.it (L.P.); lorenzo1.rossi@polimi.it (L.R.)
2 Department of Civil and Environmental Engineering, University of California, Berkeley, Berkeley, 94720 CA, USA
3 Politecnico di Torino, Department of Environment, Land and Infrastructure Engineering, Corso Duca degli Abruzzi 24, 10129 Torino, Italy; alberto.cina@polito.it (A.C.); paolo.maschio@polito.it (P.M.); marco.piras@polito.it (M.P.)
4 Università degli Studi di Genova, Laboratory of Geodesy, Geomatics and GIS, Via Montallegro 1, 16145 Genova, Italy; daniele.passoni@dicca.unige.it
* Correspondence: carlo.demichele@polimi.it; Tel.: +39-02-23996233
† Authors in alphabetical order.

Received: 28 March 2018; Accepted: 14 May 2018; Published: 16 May 2018

Abstract: Performing two independent surveys in 2016 and 2017 over a flat sample plot (6700 m²), we compare snow-depth measurements from Unmanned-Aerial-System (UAS) photogrammetry and from a new high-resolution laser-scanning device (MultiStation) with manual probing, the standard technique used by operational services around the world. While previous comparisons already used laser scanners, we tested for the first time a MultiStation, which has a different measurement principle and is thus capable of millimetric accuracy. Both remote-sensing techniques measured point clouds with centimetric resolution, while we manually collected a relatively dense amount of manual data (135 pt in 2016 and 115 pt in 2017). UAS photogrammetry and the MultiStation showed repeatable, centimetric agreement in measuring the spatial distribution of seasonal, dense snowpack under optimal illumination and topographic conditions (maximum RMSE of 0.036 m between point clouds on snow). A large fraction of this difference could be due to simultaneous snowmelt, as the RMSE between UAS photogrammetry and the MultiStation on bare soil is equal to 0.02 m. The RMSE between UAS data and manual probing is in the order of 0.20–0.30 m, but decreases to 0.06–0.17 m when areas of potential outliers like vegetation or river beds are excluded. Compact and portable remote-sensing devices like UASs or a MultiStation can thus be successfully deployed during operational manual snow courses to capture spatial snapshots of snow-depth distribution with a repeatable, vertical centimetric accuracy.

Keywords: UAS; laser scanning; snow; snow courses; MultiStation

1. Introduction

Monitoring snow distribution has important implications for both water resources management and risk prevention [1]. The amount of snow can be quantified indirectly as snow depth (HS, in m), or directly as snow water equivalent (SWE, in mm w.e. or kg/m², see [2]). Both variables are often measured with snow pits and manual probing [2], which are both time consuming and risky in
avalanche-prone, remote areas. HS can be also measured using ultrasonic [3] or laser [4] sensors, while SWE can be monitored using snow pillows [5] or cosmic rays [4]. The significance of local measurements has been often debated [6–10], especially in view of the marked spatial variability of snow processes [11–14]. To partially take this variability into account, snow manual measurements are often performed along snow courses and then averaged to provide a more representative estimation of available SWE and snow depth [15].

Remote sensing has recently emerged as a non-invasive alternative for monitoring snow water resources. Remote sensing captures the spatial and temporal patterns of snow and thus overcomes the potential undersampling of point measurements and the long surveys needed for snow courses [16,17]. Existing methods include Terrestrial Laser Scanner (TLS, [7,13,18–22]), digital photogrammetry [23,24], tachymetry [20], Ground Penetrating Radar (GPR) [25], time-lapse photography [26,27], or satellite-based sensors [16]. Among these alternatives, TLS is the commonest choice in most applications [28].

Several attempts of using Unmanned Aerial Systems (UASs) on snow have been recently carried out. These systems are commonly used for high-resolution surveys [29–38] as they allow flights to be performed in an automatic way [39,40]. The miniaturization of imaging and positioning sensors also reduces the payload and thus enables flights up to about one hour long [41]. The increasing use of UAS is also related to the improvement of Structure from Motion (SfM) and its combination with algorithms for automatic tie points extraction such as Scale Invariant Feature Transform (SIFT) [42] and Speed-Up Robust Feature (SURF) [43], which automatically reconstruct three-dimensional models from sequences of two-dimensional images [44,45]. These feature-based algorithms give several reliable matchings even in case of bad-textured surfaces [46,47].

Existing works employing UAS on snow or glaciers report an expected Root Mean Square Error (RMSE) for HS below 30 cm (e.g., see [28,48–55]). Larger errors are attributed to vegetated areas [28,50,52]. However, the performances of UAS on snow have mostly been quantified using datasets at low density [48–52] or mixed fresh-old snow and bare-ice surface textures [55], whereas only [54] and [28] present comparisons with a TLS under different illumination conditions. Snow tends to form homogeneous surfaces and therefore the identification of homologous points on different images of the photogrammetric block can be highly uncertain [55–58], especially in case of high-resolution images where each frame covers only a small area. This ambiguity represents one of the biggest challenges for UAS-based photogrammetry compared to a TLS. UAS flights may also suffer from strong wind, which is a frequent, yet variable condition in mountain areas, while a TLS needs several stationing points to map snow depth over an irregular terrain, hence a longer survey. These shortcomings are likely among the reasons why most operational services around the world prefer traditional, manual sampling over remote-sensing techniques. Further efforts are needed to establish UASs or laser-scanning devices as low-cost, precise, and portable tools for monitoring snow at slope-catchment scale.

By comparing UAS-based photogrammetric maps of snow depth with point clouds acquired with a Leica Nova MultiStation (MS) [59,60], we show that these instruments return highly consistent and repeatable results in measuring snow depth over a flat sample plot, a well-established scenario for operational services. While a laser scanner uses a mirror to reflect the laser beam on the target scene, a MS acquires a point cloud by moving the telescope collimation axis and scanning the target scene at a very high frequency, i.e., up to 1000 points/s for distances smaller than 300 m (accuracy of 2 mm + 2 ppm). A TLS enables to acquire a denser point cloud with an accuracy of the order of centimeters, whereas a MS can potentially reach the order of millimeters, as recently quantified by [60] on standard construction materials in laboratory and field conditions. This device has been here never tested on snow before, to our knowledge.

Surveys were performed at peak accumulation during the 2016 and 2017 snow seasons. We employed as ground truth for both techniques a relatively dense network of co-located manual measurements (135 and 115 snow-depth observations in 2016 and 2017, respectively). The high density of these measurements (~2 pt/100 m²) aims to assess the role that snow variability plays in ruling sensor performances. The snow-depth distribution is computed by comparing two different Digital
Surface Models (DSM), i.e., a DSM of snow distribution (April 2016 and March 2017) and the DSM of bare soil (September 2016).

Section 2 introduces the case study and the instrumentation used. Section 3 presents the processing methods and the results. Sections 4 and 5 report discussions and conclusions, respectively.

2. The Case Study

The study area is located nearby the Belvedere glacier (Piedmont region, Italy, 45°57′10.85″, 7°55′5.22″, 2070 m a.s.l) and extends for about 6700 m² (Figure 1). The site is characterized by sparse rocks and grass with no trees. The area is also crossed by two streams. While the topography is quite homogeneous (maximum variation of ∼7 m in correspondence of the highest rocks), the bare-ground coverage is variable and this enables to investigate the variability of sensor performances with different topographic features, representative of Alpine headwater catchments.

Figure 1. Topography of the surveyed area during summer 2016 ((a), top-left), winter 2016 ((b), top-right) and winter 2017 ((c), bottom-right). The boundaries of the study area are in black. Red dots denote the station points of the MultiStation, whereas the blue triangle represents the point used as angular reference. No bare-soil survey was performed during summer 2017.

Two different surveys are necessary to measure the depth and the distribution of snow. A first survey during summer is used to define the site topography and the reference, bare surface (henceforth, summer survey). A second survey is then carried out during peak accumulation (henceforth, winter survey). The vertical differences between the multi-temporal estimated DSMs yield a map of snow-depth distribution. The summer survey was executed on 27th September 2016, while the two winter surveys were performed on 14th April 2016 and on 31st March 2017 (Figure 1). Table 1 reports a schedule of the two field surveys on
snow. The schedule for the summer survey is not reported as sub-daily temporal differences between UAS and MS scans on bare soil are irrelevant for our scopes.

Table 1. Schedule of the two field surveys on snow (all times are local).

<table>
<thead>
<tr>
<th>Task</th>
<th>Winter 2016</th>
<th>Winter 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>UAS flights</td>
<td>12 PM</td>
<td>12 PM</td>
</tr>
<tr>
<td>MS surveys</td>
<td>12 PM to 1 PM</td>
<td>12 PM to 1 PM</td>
</tr>
<tr>
<td>Manual probing</td>
<td>1 PM to 4 PM</td>
<td>1 PM to 4 PM</td>
</tr>
</tbody>
</table>

2.1. UAS Flights

UAS surveys were performed using a hexacopter. This system is equipped with an on-board low-cost GNSS receiver (Ublox LEA GH), a triaxial magnetometer, and a MicroElectro-Mechanical System Inertial Measurement Unit (MEMS IMU). It has a maximum payload of 0.500 kg and is characterized by a flight autonomy of about 15 min. Images were acquired using a Canon EOS M camera with a fixed focal length of 22 mm, equivalent to a focal length of 35 mm in full-frame format. Both flights were planned following the guidelines by [36]. The average flight height for the summer and winter acquisitions was about 60 m, which corresponds to an average Ground Sample Distance (GSD) of about 0.02 m. The summer images were acquired at about 10 AM (local time), while the winter surveys were performed around midday (local time, Table 1).

Flights were arranged in 4 strips and planned to guarantee high overlaps, which were adjusted across campaigns basing on a trial-and-error procedure. The first survey was in winter 2016 (on snow): considering possible problems connected to image matching in case of homogeneous surfaces, we precautionarily used a very high overlapping (i.e., 90% along flight direction and 70% along crossed-flight direction). For the summer 2016 flight, the number of images was reduced by using overlaps equal to 80% along flight direction and 60% along crossed-flight direction, respectively. This decision was made to speed up the bundle block adjustment procedure, guaranteeing at the same time the same level of accuracy. In fact, the presence of different, recognizable texture on the ground surface (e.g., mixture of bare soil, grass, rocks etc.) guaranteed good results during the image matching stage. These overlaps have been planned considering the recommendation given in [35] and [61]. Finally, the winter 2017 flight was designed using the same overlaps used for the summer flight, thus allowing us to reduce computational efforts. Several tests conducted with winter 2016 data showed indeed that the same results in terms of photogrammetric block accuracy on snow can be achieved by reducing the number of images, if sufficient overlaps are guaranteed [35].

The photogrammetric blocks were georeferenced using Ground Control Points (GCPs), represented by black-and-white square targets (0.30 × 0.30 m). The position of both the summer and winter GCPs is given in the same reference frame, which is necessary to compute HS by means of a differentiation of photogrammetric DSMs. Therefore, three geodetic networks were realized and measured in the field, combining MS and GNSS measurements. GNSS measurements were only used to georeference the three surveys, while the MS (used in its Total Station mode) was used to measure all the GCPs. To correctly georeference the surveys, each of them was referred to some permanent GNSS stations, known in the ETRF2000(08) reference frame. The final adjustment leads to a centimetric accuracy for the position of GCPs for both cases in the global ETRF2000(08) reference frame.

2.2. MultiStation Scans

The MS used was a Leica Nova MultiStation MS60. This device acquired point clouds on a regular angular grid with a horizontal spacing between 0.01–0.10 m depending on the distance from the station point. During the summer survey, three station points were considered (Figure 1): this results in ~3 × 10^6 points, which cover the entire study area. On the other hand, only one station point (see again Figure 1) was used during the winter surveys due to time and accessibility restrictions. In 2016, the station point was located at
one edge of the investigated area, which resulted in $\sim 1 \times 10^6$ points and a relatively small scanned area due to increased obstructions and depressions in the middle of the study area. In 2017, we chose a barycentric station point to obtain a more complete point cloud of about $\sim 2.3 \times 10^6$. The scanned area on snow always obeys to a trade-off between maximizing survey extension and minimizing snow-melt due variations of the snow surface. Scans on snow were performed immediately after UAS flights, approximately between 12 PM and 1 PM (Table 1).

In both winter cases, the station points were materialized over a high rock to guarantee acquisitions from a raised point of view. The height of the scanning device was thus about 3 m above the snow surface (including tripod height). The MS was previously tested considering different combinations of materials, distances and incidence angles [60]. Results show that the accuracy is very well preserved when changing the incidence angle between 0 and 80 degrees. During the winter campaigns discussed in this paper, the zenith angles measured with the MS were in the range between 121 and 92 degrees (2016) and between 117 and 88 degrees (2017). The lowest value corresponds to some rocks emerging from the snowpack. The zenith angle coincides with the supplementary of the incidence angle when the terrain can be considered horizontal. Because of the flat topography and the morphology of the investigated site, small horizontal variation produced insignificant variation in the final DSM.

2.3. Manual Probing

Point measurements of snow depth were performed using portable stakes (aluminum, diameter $\sim 1$ cm, resolution 1 cm). A regular grid of points was defined and materialized in the field using ropes. The grid was composed by 12 (10) courses in 2016 (2017); the average spacing between measurement points on the same course was $\sim 5$ m, which aimed at reasonably capturing the variability of snow depth at plot scale. Each measurement took a few minutes and about two hours were needed to complete the manual survey. Both surveys were performed between 1 and 4 PM local time, meaning that snow surface was undisturbed during UAS and MS measurements (Table 1). This time span, however, introduced a slight decrease in the measured snow depth with time due to snowmelt (see Section 3). The position of each probing point was measured using the MS (centimetric precision), thus guaranteeing the co-registration with the GCPs used for the photogrammetric processing and with the MS point clouds.

3. Results

3.1. UAS Photogrammetric Blocks: Processing

The photogrammetric blocks of the three surveys were processed using Agisoft Photoscan (version 1.2.6) (www.agisoft.com). The summer block is composed by 84 images, while the winter ones are composed by 144 images in 2016 and 74 images in 2017. As stated before, searching correspondences (tie points) on snow may introduce large uncertainty, and for this reason an higher overlapping was used for the 2016 winter flight. However, we verified that the automatic algorithm performed very well also on quite homogeneous surfaces such as snow; for this reason the overlapping for 2017 flight was similar to that of the summer 2016 survey. Choosing the best overlapping is again the result of a trade off between high tie point multiplicity and wide baseline. While the first is necessary for image matching algorithms, the latter guarantees a satisfactory intersection between homologous rays. Both configurations (winter 2016 and winter 2017) allowed us to reach an accuracy of the order of one GSD, which means that the two solutions are characterized by the same level of accuracy. Of course, processing time will be lower if the photogrammetric block is composed by less images.

Each block was processed separately, following the standard photogrammetric procedure. First, tie points were extracted from multiple images and the External Orientation (EO) parameters were computed, constraining the block with the GCPs previously measured (bundle block adjustment). The accuracy of each GCP was specified for each coordinate, in agreement with the precision obtained from the geodetic network adjustment. Following this procedure, we ensured that the different
observations were used together with their correspondent weight during the bundle block adjustment. During the bundle block adjustment, camera Internal Orientation (IO) parameters were also optimized by means of a self-calibration tool embedded within Agisoft Photoscan. This step is fundamental in case of a UAS survey because of the disturbances to the system induced by take-off and landing, which can change these parameters compared to those obtained prior to the flight [35]. The empirical accuracies of GCPs for the two surveys are shown in Table 2.

Table 2. Root Mean Square Error between the measured and estimated coordinates of Ground Control Points (GCPs).

<table>
<thead>
<tr>
<th>Flight Season</th>
<th>East (m)</th>
<th>North (m)</th>
<th>Height (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer 2016</td>
<td>0.010</td>
<td>0.007</td>
<td>0.005</td>
</tr>
<tr>
<td>Winter 2016</td>
<td>0.017</td>
<td>0.010</td>
<td>0.004</td>
</tr>
<tr>
<td>Winter 2017</td>
<td>0.006</td>
<td>0.007</td>
<td>0.009</td>
</tr>
</tbody>
</table>

For each survey, the photogrammetric dense point cloud was computed considering an image downscaling factor equal to 4, namely, two times for each image size. The generated point clouds are composed by about $50 \times 10^6$ points for the summer survey, $55 \times 10^6$ points for the 2016 winter set and $37 \times 10^6$ points for the 2017 winter set. Then, the DSM was derived considering a pixel size equal to 0.03 m, as shown in Figure 2. This value was selected to have a cell size larger than the GSD, thus guaranteeing a sufficient number of observations for each cell.

![Figure 2](https://example.com/figure2.png)

Figure 2. Digital Surface Models from the UAS surveys during summer 2016 ((a), top-left), winter 2016 ((b), top-right) and winter 2017 ((c), bottom-right). The color map represents surface height (ASL). Contour lines (grey) are reported with an equidistant interval of 2 m. No bare-soil survey was performed during summer 2017.
3.2. UAS vs. MultiStation

For each survey, we compared the MS scans with both photogrammetric products, i.e., dense point clouds and DSMs. We first performed a cloud to cloud comparison (C1), meaning that the clouds of points from UAS and MS were directly compared. Second, we compared the MS cloud with the UAS-based computed DSM (C2) as DSMs are largely used as final products in snow applications [62]. We avoided a DSM to DSM comparison between MS and UAS as this would have been affected by uncertainties connected to each DSM generation. The MS scans were surveyed using the same geodetic network used for measuring the GCP, which guarantees the co-registration between the datasets.

C1 was performed by computing the height difference between each point of the MS scan and its nearest neighbour in the photogrammetric point cloud, which was found basing on the shortest 3D distance. Because the photogrammetric point cloud has some gaps due to shadows or poor matching over some areas of the snow surface, any resulting couple of points with a distance greater than 0.03 m in the horizontal plane was not included in the statistics. This corresponds to setting a maximum horizontal search radius of 0.03 m, which is assumed as the maximum acceptable distance between points on different datasets that make them physically correspondent. C2 was performed by interpolating the UAS DSM in correspondence of the horizontal coordinates of the MS point clouds and then comparing this interpolated height with that measured by the MS.

Table 3 reports the mean and the standard deviation of the residuals for each comparison along with the corresponding RMSE. In all cases, standard deviations are smaller than 6.8 cm, which means that the correspondence between the photogrammetric surveys and the MS point clouds is high. Moreover, the standard deviations of C1 (both summer and winter cases) are always smaller than the corresponding values of C2. This is because the DSMs used in C2 were obtained interpolating (hence, smoothing) the photogrammetric dense point clouds, which results in larger differences especially at discontinuities. As a result, height residuals increase in correspondence of abrupt terrain variations, as shown in Figure 3.

Table 3. Statistics of the winter and summer comparisons between the photogrammetric products and data from the MultiStation.

<table>
<thead>
<tr>
<th>Survey</th>
<th>Point Cloud (C1)</th>
<th>DSM (C2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (m)</td>
<td>St. Dev. (m)</td>
</tr>
<tr>
<td>Summer 2016</td>
<td>0.004</td>
<td>0.020</td>
</tr>
<tr>
<td>Winter 2016</td>
<td>0.026</td>
<td>0.025</td>
</tr>
<tr>
<td>Winter 2017</td>
<td>−0.003</td>
<td>0.015</td>
</tr>
</tbody>
</table>

The mean of the residuals for both summer 2016 and winter 2017 is not significantly different from zero compared to the corresponding standard deviations. However, both C1 and C2 show a bias for the winter 2016 case compared to other surveys. By projecting the GCPs on the MS point cloud, a vertical offset of about 2.6 cm, with a standard deviation of about 1 mm, is indeed found. During the winter 2016 survey, GCPs were acquired about 3 h after the measurement of the point cloud with the MS. This delay introduced a height variation due to progressive snowmelt. A nearby weather-snow station measured an average daily temperature on the survey day equal to +2.6 °C and an average decrease in snow depth of about 4–5 cm between 13 and 15 April (14 April being the survey day). We considered an average decrease across three days due to spurious fluctuations in data from these ultrasonic depth sensors at hourly to daily temporal resolutions [3]. The elevation and the slope of this station are comparable to that of the study area. This decrease in snow depth at the weather station is consistent with the observed bias, which refers to the three hours of peak snowmelt. During the winter 2017 survey the GCP’s acquisition and the MS scan were performed almost simultaneously, which avoided similar issues.
Figure 3. Height differences between the MultiStation (MS) dataset and the photogrammetric products for all surveys. C1 compares the MultiStation and the photogrammetric point clouds, whereas C2 compares the MultiStation point cloud with the DSM from UAS data.

Figure 4 reports the relative frequencies of the residuals for C2 and the corresponding scatter plots between UAS and MS heights (ASL). Results confirm a bias for the 2016 winter survey, even though all histograms are symmetric. The $R^2$ of a linear regression between the data of the MS and those of the UAS reads 0.97, 1.00, and 0.99 for the 2016 winter survey, the 2017 winter survey, and the summer survey, respectively. These results show that UAS and MS can provide a map of snow depth distribution with similar, competitive accuracy. In light of this high agreement, only UAS (for which we produced a DSM) was benchmarked against manual probing.

The horizontal range in Figure 4a excludes few locations with a significantly larger residual (see Figure 4b) because these points are characterized by a very small relative frequency. These larger
biases are expected in shaded areas or depressions (see Figure 3) and are mostly evident in C2, which means that they are generally the result of interpolation, i.e., DSM creation.

Figure 4. Histograms of the residuals between the MultiStation scan and the DSM from UAS data (C2) for each of the three surveys (a). Scatter plot of UAS-based vs. MS-based heights (ASL) for both winter and summer cases (b).

3.3. UAS vs. Manual Probing

Figure 5a shows the scatter plot between UAS-based (henceforth, HS$_{UAS}$) and manual (henceforth, HS$_{M}$) measurements of HS during the 2016 winter survey. A linear regression between these two datasets reads $\text{HS}_{UAS} = 0.72 \times \text{HS}_M + 0.51$ ($R^2 = 0.53$). While these statistics show that the two datasets are correlated (73%), the difference between the data increases with a decreasing HS$_M$. This plot also shows that UAS-based HS is generally greater than the manual one, i.e., $\text{HS}_{UAS} - \text{HS}_M > 0$ in $\sim 89\%$ of the data. Figure 5b reports the same statistics using the 2017 snow survey: the linear regression between HS$_{UAS}$ and HS$_M$ reads $\text{HS}_{UAS} = 0.93 \times \text{HS}_M + 0.069$ ($R^2 = 0.97$). In this case, the two datasets shows a much higher correlation (98%).

The mean difference between the UAS-based DSM and the 135 manual measurements (henceforth, HS$_{UAS} - \text{HS}_M$) in 2016 reads 0.20 m, whereas the standard deviation is equal to 0.24 m. This translates into a RMSE equal to 0.31 m. About 39% of the measurements are between zero and 0.12 m (the median), with maximum and minimum differences equal to $\sim 1$ m and $-0.23$ m, respectively (Figure 5e).

Figure 5c shows the spatial distribution of the differences between the UAS-based DSM and manual measurements for this first snow survey. Along courses A, B, E, F, J, L, M and N (Group 1), differences are smaller and more spatially homogeneous than along courses C, D, G, and H (Group 2, see Table 4). Because the measurement protocol as well as the manual probe were the same for all the courses, this discrepancy cannot be easily explained by a measurement error. While the differences are clustered along these courses, no evident spatial pattern in vegetation or soil coverage emerges that could clearly explain this mismatch.

The 2016 snow season at the study site was marked by an early snowpack (October 2015) that persisted at the site up to January 2016 and created a shallow, dense base layer (up to $\sim 10–15$ cm thick). Between January and April 2016, the snowpack increased up to $\sim 150$ cm, but both melt-freeze and rain-on-snow events occurred over the study area, especially in early April. Both processes, together with water retention at layer transitions due to capillary barriers [63], favor the development of ice layers [64]. While no pit was excavated during this first field survey due to time constraints, ice layers were observed close to the two streams crossing the study area, where the snow cover was patchy. The random occurrence of ice layers, coupled with slope redistribution of water in snow [65], may have impeded the full penetration of manual probes into snow, hence a systematic underestimation of snow depth [66].
Figure 5. Evaluation of UAS data using manual probing. (a,b): Scatter plots between UAS-based and manual measurements of HS for 2016 and 2017, respectively (regression line in red). (c,d): Spatial distribution of the differences between UAS-based and manual measurements of snow depth. The color legend represents the differences between UAS and manual probing. (e): Histograms of the differences between UAS-based and manual measurements of snow depth for both surveys.

During the second field survey in 2017, the mean difference between $\text{HS}_{\text{UAS}}$ and $\text{HS}_{\text{M}}$ was equal to 0.01 m, with a standard deviation equal to 0.20 m. This translates into a RMSE equal to 0.20 m (see Table 5). The maximum difference is equal to 0.52 m, while the minimum is equal to $-1.08$ m. During this second survey, the locations of maximum and minimum differences were more clearly correlated with topography, corresponding either to bushes or to the stream bed, which are both conditions that are prone to noise in manual probing as well as in DSMs. Both bushes and streams present complex elevation differences between winter and summer DSMs that may not be entirely due to snowpack, including potential micro-topographic variations in the surface of stream beds due to seasonal transport of sediments or larger rocks. This complicates the comparison at these
specific locations as probing is expected to reach the ground surface, while HS_{UAS} is the result of a differentiation between DSMs (e.g., photogrammetry reconstructs the surface of bushes, whereas probes are supposed to reach the underlying ground surface). If these outliers are removed from the sample, RMSE, minimum difference, and maximum difference decrease to 0.06 m, −0.72 m, and 0.19 m, respectively. One pit was excavated close to the study area during this second snow survey and no ice layer was found at any depth.

Table 4. Statistics of the differences between UAS and manual measurements in 2016. Results are separated between areas with significantly smaller errors (Group 1, courses A, B, E, F, I, L, M and N—84 samples) and areas with larger errors (Group 2, courses C, D, G, and H—51 samples).

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean (m)</th>
<th>St. Dev. (m)</th>
<th>RMSE (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.11</td>
<td>0.14</td>
<td>0.17</td>
</tr>
<tr>
<td>2</td>
<td>0.36</td>
<td>0.27</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Table 5. Statistics of the differences between UAS and manual measurements in 2017. Results are reported with or without outliers (8 samples out of 115).

<table>
<thead>
<tr>
<th></th>
<th>Mean (m)</th>
<th>St. Dev. (m)</th>
<th>RMSE (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>With outliers</td>
<td>0.01</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>Without outliers</td>
<td>0.04</td>
<td>0.05</td>
<td>0.06</td>
</tr>
</tbody>
</table>

4. Discussion

Our results show that both UAS photogrammetric flights and a new, state-of-the-art scanning device capable of millimetric accuracy (MultiStation, [60]) can measure snowpack distribution at cm scale with mutually consistent results (maximum RMSE of 0.036 between point clouds). The comparison between UAS flights and manual sampling show higher RMSEs, in agreement with what reported before by e.g., [49] (0.14 m), [50] (less than 0.15 on rocks and less than 0.3 m on grass), [52] (0.22 m in open areas and 0.42 m in forests), or [51] (between 0.085 m and 0.137 m) [53] reported a median difference of −0.11 m and a standard deviation of 0.62 m, whereas [55] found accuracies between 0.10 m and 0.25 m on three Swiss glaciers—including both snow and bare ice.

This finding expands existing comparisons between UAS flights and other techniques on snow, which have focused on either manual sampling (see previous paragraph) or terrestrial laser scanners, capable of centrimetric accuracy [28,54]. A MS was used here on snow for the first time, to our knowledge. Such a high precision suggests that SfM techniques can now correctly solve the potentially more complex matching of images on a homogeneous surface, such as snow. This statement is supported by the standard deviations of the residuals between the UAS and the MS point clouds, which are comparable for both the summer and all winter cases.

The obtained RMSE between UAS flights and the MS is significantly smaller than those reported by [54], who compared four UAS flights at different wavelengths and shadow conditions with a TLS (RMSE between 0.18 m and 0.77 m). [28] expanded on this comparison involving 12 flights with visible and near-infrared cameras (accuracy below 0.29 m in full sunlight). On the one hand, our results show the expected increased accuracy of a MultiStation compared to a TLS and further support the high agreement between scanning and photogrammetric techniques on snow. On the other hand, results were obtained under optimal illumination conditions, whereas [28] show that the accuracy of UAS flights decreases under suboptimal illumination (accuracy of 0.49 and 0.37 m for DSMs and snow-depth maps at visible wavelength). Comparing UAS flights with a MultiStation under such conditions is an important direction of future work. It is worth noting that the error distribution over the reconstructed surface is homogeneous in case of a flight planned and realized considering appropriate overlapping and a constant flight height (as for the case here presented). On the contrary, the precision of a MS or a TLS depends on the distance between the station point and the investigated
object (i.e., the distance traveled by the laser beam), on the incidence angle and on the material reflective properties.

Compared with most previous comparisons between UAS flights and manual probing, we focused here on a relatively dense grid. For example, [48] used 37 measurements over 6900 m$^2$ (0.5 pt/100 m$^2$), [23] took between 0.04 pt/100 m$^2$ and 0.11 pt/100 m$^2$, and [52] took between 0.04 pt/100 m$^2$ and 0.2 pt/100 m$^2$. [51] considered wide areas (0.65 km$^2$ and 0.32 km$^2$), but one of the two study sites included intensive plots, which is comparable to our survey. [28] also considered an intensive study plot with an average spacing of 18 m over 0.12 km$^2$. Because the spatial variability of snow may be very large even over short distances [8], the outcome of a comparison between UAS and manual probing may change with the spatial resolution of manual data. However, these surveys revealed that increasing the number of points does not significantly affect the overall precision of the survey (i.e., the RMSE). UAS represents a competitive choice among existing techniques for high-precision remote sensing of snow at high resolution.

These surveys focused on snow depth, one of the two key variables that are routinely monitored by operational services all around the world—the second being Snow Water Equivalent. In California, for example, the California Snow Cooperative Survey performs snow-depth and SWE manual courses at monthly time scale over the entire Sierra Nevada and southern Cascade (http://cdec.water.ca.gov/snow/). Over the same mountain range, the Airborne Snow Observatory measures snow depth every two weeks during the snowmelt season on selected catchments using LiDAR scans [62]. Over European Alps, various agencies routinely manually measure snow depth for avalanche and hydrologic forecasting (http://www.avalanches.org/eaws/en/main.php). While SWE is the primary variable of interest for hydrologic applications, snow depth is a key variable for avalanche forecasting [20]. Because snow density is generally less variable in space than snow depth [67,68], knowledge of snow depth distribution can also potentially be converted to SWE via empirical regressions [69] or dynamic models [62]. In this context, portable, low-cost techniques, such as UAS and a MS, can fill an important gap between laborious, manual measurements and large-scale surveys at lower resolution using satellites or manned aircrafts. Yet, snow-density estimates using models currently represent the most important source of uncertainty when converting LiDAR-based snow depth to SWE [70]: coupling these devices with co-located measurements of density is therefore an important direction of future development.

5. Conclusions

We compared UAS-based photogrammetry and a MultiStation (MS) in measuring snow-depth distribution over a sample plot (6700 m$^2$). Two different surveys at peak accumulation were performed in spring 2016 and 2017, whereas a survey in summer 2016 with both instruments defined the reference, bare ground. We collected a relatively dense network of manual measurements of snow depth (135 pt in 2016 and 115 pt in 2017) to compare both techniques with the standard, traditional probing method. While several comparisons between UAS photogrammetry and manual probing are available in the literature, a MS was tested on snow for the first time.

UAS-based photogrammetry and the MS show highly consistent results: the mean of the residuals between the MS and UAS point clouds is not significantly different from zero for the summer survey (0.004 m vs. a standard deviation of 0.020 m). For the winter cases, a bias equal to 0.026 m was found in 2016, which is mainly due to ongoing snowmelt (time difference of three hours). The mean of the residuals for the 2017 snow survey is much smaller, i.e., $-0.003$m when comparing clouds. The standard deviations of the residuals are of the order of the GSD for all the cases. These small variations are related to different soil coverage and illumination conditions, which cause different tie-point matching quality. The Root Mean Square Error (RMSE) between UAS and MS point clouds on snow are less or equal than 0.036 m. When comparing the photogrammetric DSMs with the MS point clouds, an increase in the mean and the standard deviations of the residuals was found. This is because a DSM interpolates a point cloud and introduces a smoothing effect, especially in correspondence of abrupt terrain variations.
The comparison between the UAS map and manual measurements shows RMSEs of the order of 0.30–0.20 m, which agrees with previous works. The mean difference between the two datasets is equal to 0.20 m and 0.01 m in 2016 and 2017, respectively. Larger differences between UAS and manual data of snow depth could be explained by canopy, topographic features, such as rocks, ice layers, and proximity to the river bed. When excluding areas of potential outliers, the observed RMSE decreases to 0.17 m in 2016 and 0.06 m in 2017. These results support the use of UAS and MS laser scanning as a support to traditional manual surveys for high-precision high-resolution monitoring of snow depth.

Author Contributions: C.D.M. conceived the study. A.B., A.C., C.D.M., P.M., L.P., and M.P. designed the field surveys. F.A., A.B., A.C., C.D.M., P.M., D.P., D.P., L.P., and L.R. performed the field surveys (realization of the geodetic networks, MultiStation acquisitions and snow height measurements). D.P. realized the photogrammetric flights. D.P. performed the photogrammetric processing. L.R. computed the geodetic network adjustment. F.A., C.D.M., D.P., L.P. and L.R. contributed to overall data analysis. The manuscript has been prepared with the contribution of all co-authors.

Acknowledgments: The authors would like to thank the students who participated to the Summer School “Design and realization of topographic surveys” (ed. 2015 and 2016) and to the 11th and 12th cycles of Alta Scuola Politecnica. We also thank Arpa Piemonte for sharing the daily weather data (season 2015–2016) and the observed profile of snow properties on 24th March 2016 at Rifugio Zamboni. Fruitful discussions and field logistical support by Mauro Spanò is acknowledged. We also would like to acknowledge discussions with the Harmosnow Consortium, COST ES 1404 action (www.harmosnow.eu) “A European network for a harmonised monitoring of snow for the benefit of climate change scenarios, hydrology and numerical weather prediction”. The project was supported by educational funds from Alta Scuola Politecnica (11th and 12th cycles—project DREAM and DREAM2).

Conflicts of Interest: The authors declare no conflict of interest.

References

5. Johnson, J.B.; Marks, D. The detection and correction of snow water equivalent pressure sensor errors. Hydrol. Process. 2004, 18, 3513–3525. [CrossRef]


24. Nolan, M.; Larsen, C.; Sturm, M. Mapping snow depth from manned aircraft on landscape scales at centimeter resolution using structure-from-motion photogrammetry. *Cryosphere* 2015, 9, 1445–1463. [CrossRef]


64. Wever, N.; Würzer, S.; Fierz, C.; Lehning, M. Simulating ice layer formation under the presence of preferential flow in layered snowpacks. Cryosphere 2016, 10, 2731–2744. [CrossRef]


67. Mizukami, N.; Perica, S. Spatiotemporal characteristics of snowpack density in the mountainous regions of the Western United States. J. Hydrometeorol. 2008, 9, 1416–1426. [CrossRef]

68. De Michele, C.; Avanzi, F.; Ghezzi, A.; Jommi, C. Investigating the dynamics of bulk snow density in dry and wet conditions using a one-dimensional model. Cryosphere 2013, 7, 433–444. [CrossRef]


© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).