Article

Accuracy Assessment of Digital Surface Models from Unmanned Aerial Vehicles' Imagery on Glaciers

Saskia Gindraux [1,*], Ruedi Boesch [1] and Daniel Farinotti [1,2]

[1]   Swiss Federal Institute for Forest, Snow and Landscape Research WSL, CH-8903 Birmensdorf, Switzerland; ruedi.boesch@wsl.ch (R.B.); daniel.farinotti@ethz.ch (D.F.)
[2]   Laboratory of Hydraulics, Hydrology and Glaciology (VAW), ETH Zurich, CH-8093 Zurich, Switzerland
[*]   Correspondence: saskia.gindraux@wsl.ch; Tel.: +41-44-739-2327

Abstract: The use of Unmanned Aerial Vehicles (UAV) for photogrammetric surveying has recently gained enormous popularity. Images taken from UAVs are used for generating Digital Surface Models (DSMs) and orthorectified images. In the glaciological context, these can serve for quantifying ice volume change or glacier motion. This study focuses on the accuracy of UAV-derived DSMs. In particular, we analyze the influence of the number and disposition of Ground Control Points (GCPs) needed for georeferencing the derived products. A total of 1321 different DSMs were generated from eight surveys distributed on three glaciers in the Swiss Alps during winter, summer and autumn. The vertical and horizontal accuracy was assessed by cross-validation with thousands of validation points measured with a Global Positioning System. Our results show that the accuracy increases asymptotically with increasing number of GCPs until a certain density of GCPs is reached. We call this the optimal GCP density. The results indicate that DSMs built with this optimal GCP density have a vertical (horizontal) accuracy ranging between 0.10 and 0.25 m (0.03 and 0.09 m) across all datasets. In addition, the impact of the GCP distribution on the DSM accuracy was investigated. The local accuracy of a DSM decreases when increasing the distance to the closest GCP, typically at a rate of 0.09 m per 100-m distance. The impact of the glacier's surface texture (ice or snow) was also addressed. The results show that besides cases with a surface covered by fresh snow, the surface texture does not significantly influence the DSM accuracy.

Keywords: UAV photogrammetry; glacier; unmanned aerial vehicle; accuracy assessment; digital surface models

1. Introduction

In recent years, Unmanned Aerial Vehicle (UAV) photogrammetry has emerged as an on-demand method to generate high-resolution datasets including Digital Surface Models (DSMs) and orthorectified images (orthophotos). This method offered a new range of application in many different research areas including forestry and agriculture (e.g., [1,2]), archeology (e.g., [3,4]) biology [5,6] and hydrology (e.g., [7,8]). An increase of studies using UAV photogrammetry is observed in geosciences, as well (e.g., [9,10]), where the DSMs and orthophotos are typically used for mapping and monitoring (e.g., [11,12]), object detection (e.g., [13]) or quantification of topographic changes, for instance (e.g., [14,15]). Classically, DSMs are produced through photogrammetric analysis of either terrestrial surveys or dedicated air- and space-borne campaigns (e.g., [16,17]). Light detection and ranging has more recently emerged as an alternative for generating high resolution DSMs (e.g., [18,19]). Compared to these methods, UAV photogrammetry can be appealing due to the (i) portability of the required UAV platforms; (ii) the possibility of self-designing and modifying the integrated sensors; (iii) the

availability of user-friendly software for data evaluation; (iv) the possibility to reach areas inaccessible with ground-based surveys and; (v) the relatively low cost.

The accuracy of DSMs derived from UAV photogrammetry is influenced by different factors that have been investigated separately, such as the camera's focal length (e.g., [20]), the flight altitude and camera orientation (e.g., [21,22]), as well as the image quality (e.g., [23,24]). Additionally, the number and spatial distribution of Ground Control Points (GCPs) used for georeferencing the acquired images have been cited as one of the most important factors (e.g., [25]). Several studies investigated the effect of varying number and position of GCPs on DSM accuracy. Shahbazi et al. [26], for instance, did several tests varying the number and locations of GCPs, as well as the number of images where the GCPs are visible. They found that a DSM georeferenced with a larger number of GCPs is more accurate that a DSM georeferenced with few and that an evenly distributed GCP-network generates a DSM of higher accuracy than a network in which GCPs are clustered. They also showed that DSM accuracy is higher when the GCPs are installed in a way that they are visible on many images. The studies of Tahar et al. [27], Tahar [28], as well as Rosnell and Honkavaara [29] also concluded that DSM accuracy increases with increasing number of GCPs. A similar conclusion was made in the study of Tonkin and Midgley [30], who additionally showed that when a certain number of GCPs is reached, the DSM accuracy does not increases further. All of these studies, however, performed their accuracy assessments on flat or undulated terrain and on well-structured surfaces, which facilitate DSM construction.

Glaciers surface texture is mainly constituted of ice and snow that can have high reflectance, a lack of structures and low contrasts, which make photogrammetry challenging. Although several studies already used UAV photogrammetry on glaciers (e.g., [31,32]) and snow (e.g., [33,34]), only a few of them thoroughly assessed the DSM accuracy. In particular, available assessments are based on a small set of validation points only, thus providing limited confidence in the obtained results. A rigorous assessment of the impact of GCP number and disposition on DSMs generated on ice or snow is missing to date.
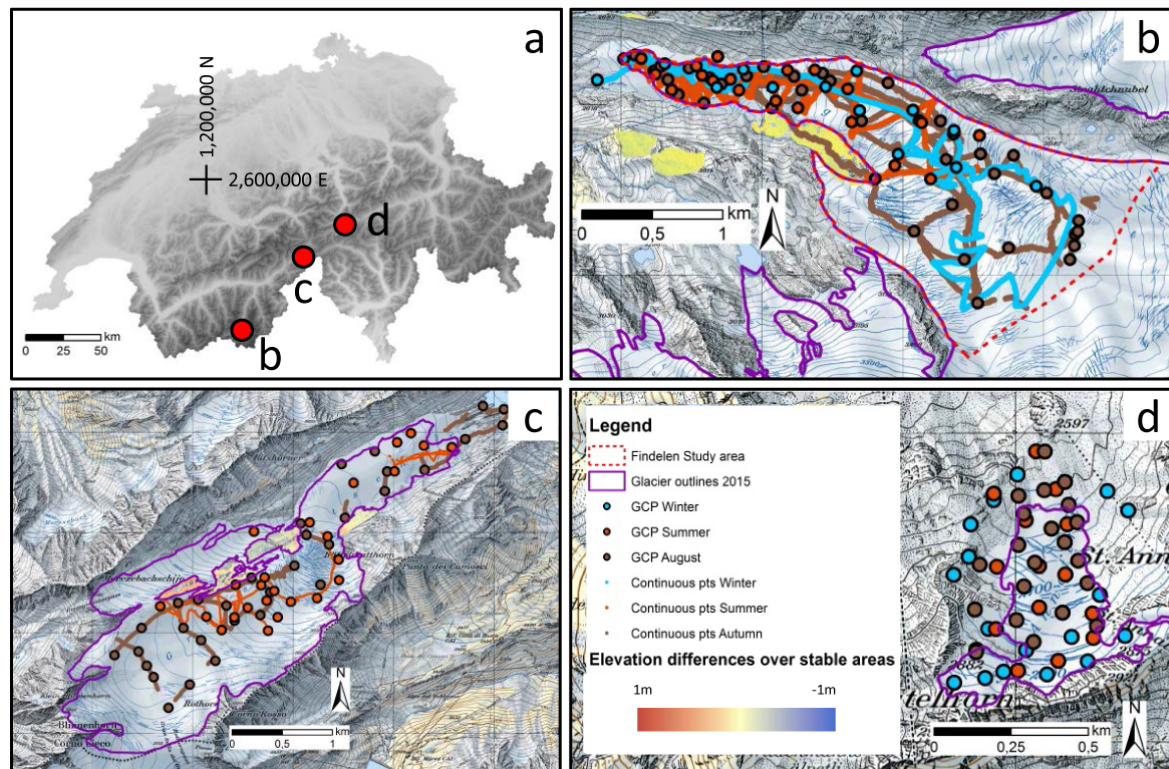
Here, we investigate the vertical and horizontal accuracy of 1321 DSMs generated from eight surveys. Each DSM is built with a different number of GCPs and validated with thousands of validation points independently acquired with differential Global Positioning Systems (dGPS). We focus on three different glaciers located in the Swiss Alps, and derive DSMs for different seasons (summer, autumn and winter). In addition, we test the DSM accuracy according to: (i) the distance to the closest GCP; (ii) the varying surface texture; and (iii) the parameter settings in the software Agisoft Photoscan that we use for DSM generation. Our results provide indications for the design of field surveys on glaciers and snow and for the accuracy that can be expected when deploying a given number of GCPs.

2. Study Areas and Data Acquisition

Three glaciers of different size, located in the Swiss Alps, were surveyed in this study (Figure 1). Findelengletscher, Sankt Annafirn and Griesgletscher have an area of $13 \, km^2$, $0.2 \, km^2$ and $4 \, km^2$ (as of 2015), respectively. Because of the large extent of Findelengletscher, only the main part of the tongue ($5 \, km^2$) was monitored.

GCPs were deployed on the glacier surface as red dots (55 cm in diameter) printed on thick paper sheets. These were fixed to the surface with wood screws in summer and with wooden sticks in winter (Figure 2a,b). The GCP center was measured prior to the UAV flights with a dGPS (Trimble R7 or Leica GPS 1200). From repeated measurements of fixed locations, we estimated the mean accuracy of the measurements to be in the order of 2 cm. The dGPS survey was executed in real-time kinematic mode using virtual reference stations from the permanent GNSS station network of Switzerland (AGNES: Automated GNSS Network for Switzerland). The same dGPS devices were used to measure transects of surface elevations between the individual GCPs (Figure 1). For these transects, points were recorded at regular intervals of about 0.50 m with an accuracy of 5 cm. We call the so-acquired points the

"continuous dGPS points". The number of GCPs and continuous dGPS points measured during each campaign can be found in Table 1.

Figure 1. (**a**) Overview of the study areas. Coordinates are given in CH1903+/LV95 [35]. The outlines of Findelengletscher (**b**), Griesgletscher (**c**) and Sankt Annafirn (**d**) are given in violet for the year 2015. The locations of the GCPs and the validation points measured with a differential Global Positioning Systems (dGPS) (continuous dGPS points) are depicted for each season. Because of the high spatial density, continuous dGPS points are not resolved individually, but appear as a line. The elevation differences between two DSMs over stable areas (Section 4.2) are shown for Findelen- and Griesgletscher. (Background maps: Swiss Federal Office of Topography, Swisstopo.)

Figure 2. Example of a GCP set on (**a**) ice and (**b**) snow. The GCP center is measured with a Leica GPS 1200 (**b**). (**c**) SenseFly eBee (**left**) and computer (**right**) linked to the transmitter sending the flight coordinates to the UAV.

For the UAV surveys, we used a SenseFly eBee system (Figure 2c; [36]). This ready-to-fly platform has a fixed wing span of 96 cm and weights 700 g including sensors. Its cruising speed is between 40 and 90 km/h, and the maximum flight duration at about 2500 m a.s.l. is of 30–35 min. The images were acquired with a customized Canon S110 camera, with a near-infrared filter (apart for the winter campaign on Sankt Annafirn, where a red-edge camera was used). The camera has a resolution of 12 megapixels and a sensor size of 7.44 mm × 5.58 mm. During every image acquisition, an on-board GPS and an inertial measurement unit provide information about the approximate 3D position, roll, pitch and yaw of the UAV. A compact UAV system including characteristics such as hand-starting (as opposed to catapult-starting), autonomous landing and small (ca. 50 m) radius turns is required for deployment in rough topography.

Flights were planned with the software eMotion 2.4 provided by SenseFly. The number of flights performed for each field campaign is shown in Table 1. An 80% lateral and 75% longitudinal ground overlap was ensured between adjacent images. The mean flight altitude was 115 m above ground, resulting in a mean ground sampling distance (GSD) of 6 cm.

Table 1. Overview of the conducted field campaigns. The season and date of each survey is given together with the total number of (a) UAV flights (N_{flight}), (b) images acquired (N_{img}), (c) GCPs deployed (N_{GCP}), (d) collected continuous dGPS points ($N_{cont.pts}$) and (e) produced DSMs (N_{DSM}). In addition, the area surveyed during each campaign (Area) is provided. Note that no continuous dGPS points were collected for Sankt Annafirn.

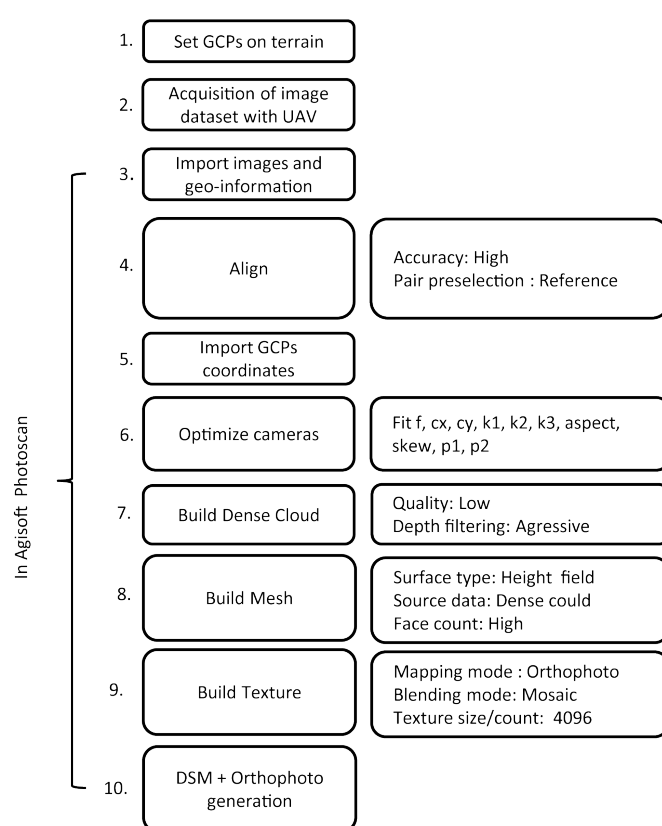| Glacier	Season	Date	N_{flight}	N_{img}	N_{GCP}	$N_{cont.pts}$	N_{DSM}	Area (km^2)
Sankt Annafirn	Summer	15 July 2015	4	750	14	-	224	1.4
Sankt Annafirn	Autumn	28 September 2015	5	856	21	-	378	1.6
Sankt Annafirn	Winter	15 April 2015	5	748	18	-	311	1.5
Findelengletscher	Summer	22–23 July 2015	9	2044	20	10,665	72	4.0
Findelengletscher	Autumn	21–22 September 2015	13	3002	36	38,813	90	5.2
Findelengletscher	Winter	19–20 April 2016	12	2901	21	12,914	66	4.2
Griesgletscher	Summer	3–04 August 2015	15	3307	25	15,795	96	6.2
Griesgletscher	Autumn	27–28 September 2015	14	3789	32	10,790	84	6.9

3. Methods

3.1. DSM and Orthophoto Generation

The images taken from the UAV were processed with a Structure-from-Motion (SfM) and multi-view stereo approach. These approaches allow the geometric constraints of camera position, orientation and GCPs from many overlapping images to be solved simultaneously through an automatic workflow [24]. In this study, the image datasets were processed with the software Agisoft Photoscan Pro 1.1.6. [37]. Computations were performed on a working station with two Intel Xeon CPU E5-2670 v3 processors (48 cores) and 128 GB random access memory. Ten steps were required to generate the DSMs and orthophotos (Figure 3):

1. In the field, the GCPs are deployed on the glacier surface and their positions recorded.
2. The image dataset is acquired through several (between four and fifteen; Table 1) UAV flights.
3. The images are imported in Agisoft Photoscan, together with the information about acquisition location (coordinates) and the roll, pitch and yaw of the UAV platform. The information is used to preliminarily orientate the images.
4. The "Align" step in Agisoft Photoscan comprises three phases. First, a feature-detection algorithm is applied to detect features (or "keypoints") on every image [38]. The number of detected keypoints depends on the image resolution, image texture and illumination [39]. Second, matching keypoints are identified and inconsistent matches removed. Third, a bundle-adjustment algorithm is used to simultaneously solve the 3D geometry of the scene, the different camera positions and the camera parameters (focal length, coordinates of the principal point and radial lens distortions).

The output of this step is a sparse point cloud. The preliminary orientation of the images (Step 3) reduces the processing time of the "align" operation, as only neighboring images (and not the entire image dataset) are searched for matching keypoints. Step 4, often referred to as SfM workflow (e.g., [40]), accounts for around 35% of the total Agisoft Photoscan processing time (Steps 3–10).

5. The GCPs are manually identified on the images and their coordinates imported.

6. The GCPs coordinates are used to refine the camera calibration parameters and to optimize the geometry of the output point cloud [24].

7. Multi-view stereo image matching algorithms (e.g., [40]) are applied to increase the density of the sparse point cloud. The density of the final georeferenced dense point cloud is strongly related to the number of matching keypoints. Different cloud quality parameters ("low", "medium" and "high") are available for the "build dense cloud" step. In this study, the parameter "low" was chosen. As shown later (Section 4.4), this reduces the DSM accuracy by 3 cm (6 cm) compared to a DSM built with the "medium" ("high") parameter, but reduces the processing time by a factor of six (eighteen). The "build dense cloud" step with the "low" parameter accounts for around 45% of the total processing time.

8. A polygon mesh is created from the dense point cloud.

9. A texture map derived from all images is applied to the polygon mesh and used to create an orthophoto.

10. A DSM with a cell size of 0.50 m is generated from the mesh and exported from Agisoft Photoscan. This cell size reflects a compromise between processing time and DSM accuracy (cf. Section 4.4).
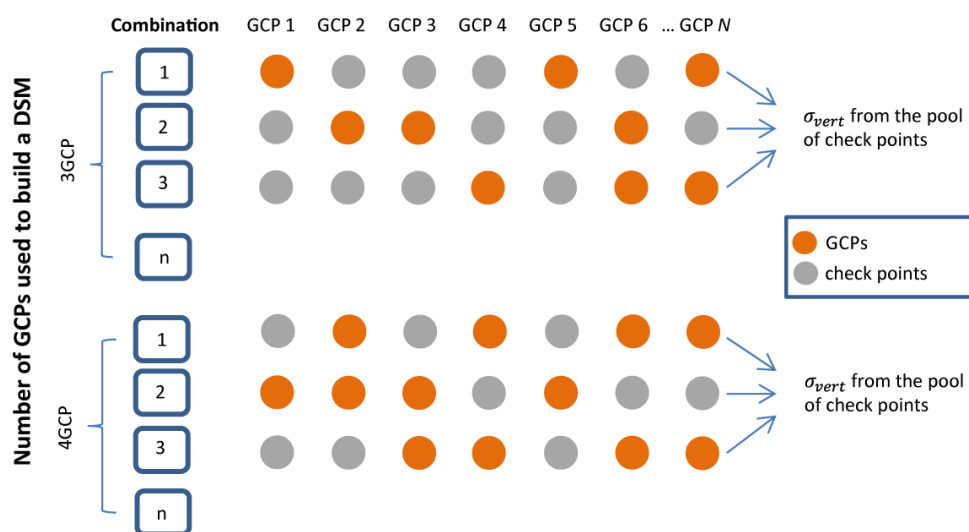

Figure 3. Workflow used for DSM and orthophoto generation (**left column**) and parameters selected in Agisoft Photoscan (**right column**).

3.2. Impact of the Number of GCPs on DSMs' Accuracy

The accuracy of the DSMs was assessed by comparing the DSMs to the acquired dGPS points. For this, 1321 different DSMs were built using different numbers and combinations of GCPs. The detailed number of DSMs built for each dataset is shown in Table 1. For every campaign, DSMs were generated by using between 3 and the maximum number of GCPs (N_{GCP}; see Table 1) for georeferencing. The remaining GCPs, which we call Check Points (CPs), were used for cross-validation (Figure 4). Analyzing all possible combinations (hundreds to thousands for every GCP number) was unfeasible. Therefore, for any given number of GCPs to be used for DSM generation, we only considered a subset of maximally $n = 6$ (Findelen- and Gries- gletscher) or $n = 21$ (Sankt Annafirn) combinations. The combinations were generated automatically, and an approximately homogeneous spatial distribution of GCPs was ensured by specifying a minimal distance between two GCPs. The standard deviation σ_{vert} of the elevation differences between the DSMs and the CPs was used to quantify the vertical accuracy. σ_{vert} was computed from the pool of all CPs obtained from the n combinations. For Findelen- and Gries- gletscher, the continuous dGPS points (Table 1) were additionally used for validation (cf. Section 2).

In order to assess the horizontal accuracy of the DSMs, a similar cross-validation analysis was performed using the orthophotos. The known horizontal coordinates of the CPs were compared with the positions of the GCPs that were manually extracted from the 1321 orthophotos. By doing so, it is the orthophoto's horizontal accuracy that is assessed. However, as the re-projection error of the final point cloud is small in our datasets, the DSM's horizontal accuracy is assumed to be very similar. As for the vertical accuracy, the standard deviation of the differences between the coordinates measured in the field and the coordinates extracted from the orthophotos (σ_{horiz}) was used as a measure for horizontal accuracy.

Figure 4. Methodology to assess the vertical accuracy for DSMs built with different numbers of GCPs. Here, an example is given for one dataset with *N* GCPs. Given a number of GCPs to be used for DSM generation (e.g., 3 and 4 in the top and bottom part of the figure, respectively; orange points), *n* different GCP-combinations (rows) are randomly selected. The remaining GCPs (check points; grey) are used for DSM validation. The standard deviation of the elevation difference between the DSM and the check points is then computed by pooling the *n* combinations (σ_{vert}).

3.3. Factors Impacting Local DSM Accuracy

The local accuracy of a DSM is governed by a number of factors, including the distance to the closest GCP, the surface structure, the topography, the varying contrast on the images, as well as the

degree of image overlap, for example. In this study, we investigated the first two factors. The influence of the other factors is not further analyzed, but is included in the discussion.

A relation between GCP density and DSM accuracy is expected to exist also at the local level, i.e., the accuracy is expected to decrease with increasing distance to the closest GCP. In order to quantify this effect, we computed the absolute elevation differences between the DSMs and the CPs, similarly to the methodology described in Section 3.2, and compared them against the distance to the nearest GCP used during DSM generation. The effect of the surface structure on the local DSM accuracy was assessed by computing the correlation between (a) the absolute elevation differences between DSMs and CPs and (b) the mean surface roughness calculated for a given area (varied between 2 and 50 m^2) around the CPs locations.

3.4. Impact of Surface Characteristics on DSM Accuracy

Snow has been cited as a potentially problematic surface for pattern-matching algorithms (e.g., [41]), although recent results have shown the applicability of SfM techniques for snow-covered surfaces, as well (e.g., [33,42]). For addressing the effect of the snow cover on the efficiency with which image matching keypoints are generated, we separately processed 20 images with fresh snow and 20 images of one-day-old snow, as well as 20 images of bare glacier ice. Images with snow-covered and bare-ice surfaces were taken from the Findelengletscher winter and summer campaigns, respectively. We processed these images up to Step 7 of the workflow presented in Section 3.1 and kept track of the number of oriented images, matching keypoints and points generated for the dense point cloud.

3.5. Effect of the Parameters Setting in Agisoft Photoscan on DSM Accuracy

The accuracy of the DSMs can depend on the SfM and multi-view stereo algorithms. The studies of Jaud et al. [43] and Vallet et al. [44] showed that depending on the software, accuracy differences in the order of 2–5 GSDs can emerge. We assess differences for Agisoft Photoscan parameters by generating a set of different DSMs for the autumn campaign on Findelengletscher's tongue. In particular, we built nine different DSMs with all available GCPs, varying the parameters controlling the "depth filter" and "cloud quality" in the image matching step (Step 7 in Figure 3). The different depth filters ("aggressive", "mild" or "moderate") control the degree to which outliers are removed from the point cloud, whilst the cloud quality parameters ("low", "medium" or "high") affects the level of detail in the reconstructed scenes. In addition, we investigated the impact of the DSM cell size by generating DSMs of 0.1 m, 0.5 m and 1.0 m for each parameter set. All generated DSMs were validated with the continuous dGPS points, and the resulting elevation differences were compared to each other.
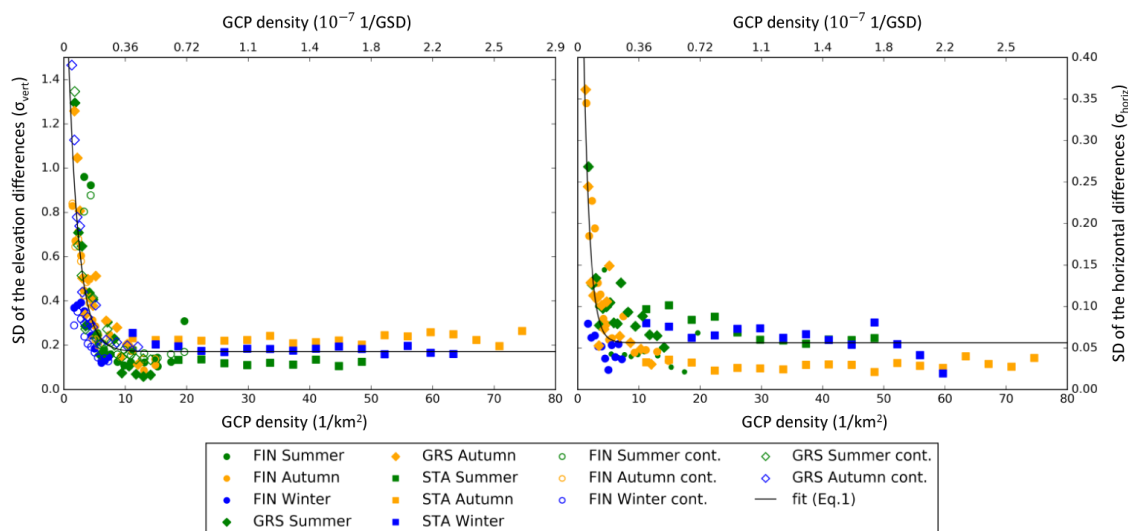
4. Results and Discussion

4.1. Influence of GCP Number on Vertical and Horizontal DSM Accuracy

The results of the accuracy assessment based on the elevation differences between DSMs and the CPs for all datasets are shown in Figure 5. In order to allow for the results of all campaigns to be compared, the number of GCPs used for DSM georeferencing was divided by the surveyed glacier area. We call the so-obtained values the GCP density. For Findelen- and Gries- gletscher, the vertical and horizontal accuracies increase (smaller σ_{vert} and σ_{horiz}) when increasing the number of GCPs used for DSM generation. This is expected, as more GCPs allow a reduction of the bundle adjustment error and therefore a more robust georeferencing (e.g., [45]). This increase in accuracy is asymptotic and can be characterized with a relation of the form:

$$\sigma = a \times exp(-b \times \rho_{\mathrm{GCP}}) + c \tag{1}$$

which we derived empirically. In the equation, σ (m) is either σ_{vert} or σ_{horiz}; ρ_{GCP} (1/km^2) is the density of GCPs per unit area; and a, b and c are three parameters to be estimated. Averaged over the three

glaciers we considered and using a non-linear least squares fit for parameter estimation, we found $a = 2.08$, $b = 0.59$, $c = 0.17$ ($r^2 = 0.70$) for σ_{vert} and $a = 0.98$, $b = 1.02$, $c = 0.06$ ($r^2 = 0.67$) for σ_{horiz}. This means, for example, that the σ_{vert} (σ_{horiz}) drops from 0.58 m (0.13 m) when 3 GCPs/km^2 are used, to 0.19 m (0.06 m) when 12 GCP/km^2 are considered. Most importantly, however, the relation indicates that σ_{vert} does not further decreases significantly for $\rho_{\text{GCP}} > 17$ GCP/km^2 (less than 5% variation compared to $\rho_{\text{GCP}} = \infty$) and that the same is true for $\rho_{\text{GCP}} > 7$ GCP/km^2 when considering σ_{horiz}. In our case, with GSD = 6 cm, the above means that the vertical (horizontal) accuracy does not increase significantly after a GCP density of 6.12×10^{-8} GCP/GSD (2.52×10^{-8} GCP/GSD). Assuming that the values of σ_{vert} and σ_{horiz} scale linearly with the GSD, we suggest that the so-expressed GCP density can be used as a rule of thumb for the planning of any UAV campaign on alpine glaciers with similar flight characteristics (e.g., image overlap, image block shape).

Figure 5. Vertical (**a**) and horizontal (**b**) DSM accuracy against GCP density. The accuracy is defined as the standard deviation (SD) of the elevation differences between DSMs and check points. The GCP density is calculated from the number of GCPs used to build a DSM, divided by the area investigated. Glaciers (seasons) are represented by different symbols (colors). The exponential fit (in black) follows Equation (1).

Our results show that the DSM accuracy increases with increasing the number of GCPs until a certain GCP density is reached, similarly as reported by Rock et al. [21] and Tonkin and Midgley [30]. The latter compared DSMs covering a glacier moraine of 0.145 km^2 with a mean GSD of 2 cm, built with varying numbers of spatially distributed GCPs (from 3–101 GCPs). They found that the vertical accuracy of the DSM built with 3GCPs (21 GCP/km^2 or 5.70×10^{-8} GCP/GSD) was 2.5-times lower than DSMs built with a larger GCP number (4–101 GCPs). In their study, the vertical accuracy did not decrease further for $\rho_{\text{GCP}} > 7.6 \times 10^{-8}$ GCP/GSD, which is similar to our results.

The parameter c in Equation (1) can be interpreted as the maximal accuracy that can be expected for a given DSM, independently of the number of GCPs used for DSM generation. In our datasets, this estimated maximal accuracy ranges between 0.10 and 0.25 m for the vertical accuracy and between 0.03 and 0.09 m for the horizontal accuracy (Table 2). For the vertical accuracy, our results are similar to three other studies that assessed the accuracy of DSMs on glaciers or snow ([31,46,47]; Table 3). For the horizontal accuracy instead, our results are comparable to the ones of Nolan et al. [47], who generated repeated DSMs for one site in the Alaskan tundra. We suspect that the comparatively low horizontal accuracies obtained by Whitehead et al. [46] (cf. Table 3) might be related to their use of natural features (instead of artificial targets) for georeferencing.

Our results for the horizontal and vertical accuracy are influenced by the process of manually setting the coordinates in the center of the GCPs (Step 6 in the workflow presented in Section 3.2). The accuracy of this manual processing step is estimated to be 0.3 pixels [43], which in our case corresponds to 1.8 cm on average. A similar uncertainty is expected when determining the coordinates of the GCP centers on the orthophotos, a determination that was used within the horizontal accuracy assessment (cf. Section 3.2). An additional source of error might affect the datasets generated over two days. On the one hand, surface lowering due to ice or snow melt can occur; on the other, ice flow can cause horizontal surface displacement. Both effects are strongest during summer, but repeated readings of stakes drilled into the ice indicate that these effects are below 0.03 and 0.02 m, respectively, on average.

Table 2. Vertical (σ_{vert}) and horizontal (σ_{horiz}) accuracies derived from parameter c in Equation (1). For Findelen- and Gries- gletscher, σ_{vert} as estimated from the continuous dGPS points (cont. pts) is given additionally.

Dataset	σ_{vert} from Check Points (m)	σ_{vert} from the Cont. Pts (m)	σ_{horiz} from Check Points (m)
Sankt Annafirn Summer	0.12	-	0.06
Sankt Annafirn Autumn	0.22	-	0.03
Sankt Annafirn Winter	0.18	-	0.06
Findelengletscher Summer	0.12	0.13	0.05
Findelengletscher Autumn	0.10	0.13	0.05
Findelengletscher Winter	0.04	0.10	0.03
Griesgletscher Summer	0.12	0.20	0.09
Griesgletscher Autumn	0.25	0.39	0.07

Table 3. Comparison between studies assessing DSM accuracy on snow or glaciers. The survey date, ground sampling distance (GSD), vertical (σ_{vert}) and horizontal (σ_{horiz}) accuracies, as well as the density of GCPs are given.

Study	Surface	Date	GSD	σ_{vert}	σ_{horiz}	GCP Density	
			(m)	(m)	(m)	(GCP/km^2)	(GCP/GSD)
Immerzeel et al. (2014)	Glacier	May 2013	0.05	0.30	0.20	6.3	1.55×10^{-8}
Immerzeel et al. (2014)	Glacier	October 2013	0.04	0.15	0.15	6.0	9.40×10^{-9}
Nolan et al. (2015)	Snow	April 2014	0.15	0.10	0.05	2.1	4.73×10^{-9}
Whitehead et al. (2013)	Glacier	July 2010	0.12	0.42	0.20	4.8	7.00×10^{-8}
Whitehead et al. (2013)	Glacier	July 2011	0.12	0.19	0.55	6.9	1.00×10^{-7}

4.2. Factors Impacting Local DSM Accuracy

The results for the relation between local DSM accuracy and the distance to the closest GCP show that, on average, the vertical accuracy decrease by 0.09 m when increasing the distance to the next GCP by 100 m (Figure 6a, light blue line). This was determined by binning the distance to the closest GCP into 10-m bins and performing a linear least squares fit of the 50th percentile of the elevation differences between the DSMs and the CPs for each of these bins. The large scatter in the relation (cf. empirical 95% confidence interval in Figure 6a) shows that the local DSM accuracy is not governed by the distance to the closest GCP alone. Tonkin and Midgley [30] performed a similar analysis. Instead of calculating elevation differences between DSM and CPs, however, they used the elevation differences between two DSMs built with 10 and 28 GCPs, respectively. They observed an ∼10-cm increase in elevation difference for every 100 m increase in distance to the closest GCP, which is very similar to our result.
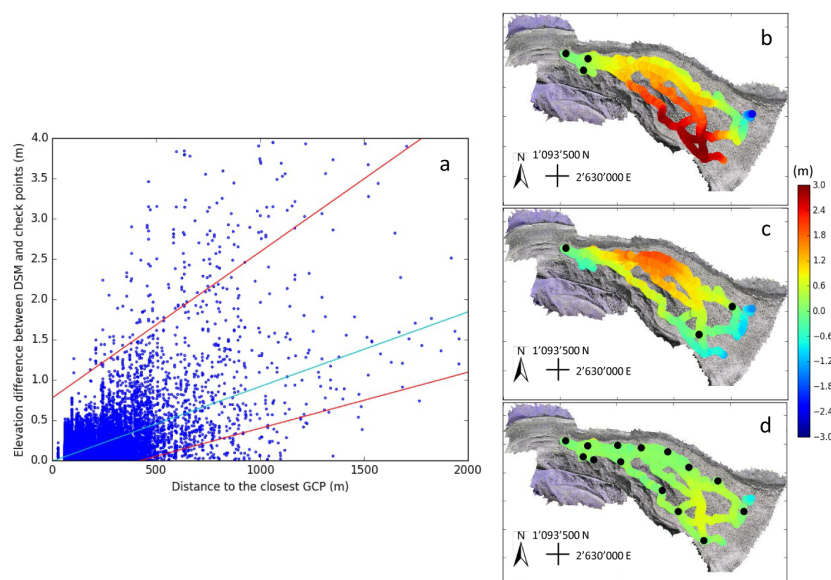
The analysis of surface roughness versus DSM accuracy provided inconclusive results. In particular, no clear relation was found between the two quantities. The DSM resolution of the

0.5-m cell might be too coarse to capture the characteristic roughness length of the analyzed surfaces. Similarly, no relation between accuracy and topographical parameters, such as aspect or slope, was found. We attribute this latter finding to the fact that our analyses focused on rather flat surfaces, i.e., areas for which such parameters are similar.
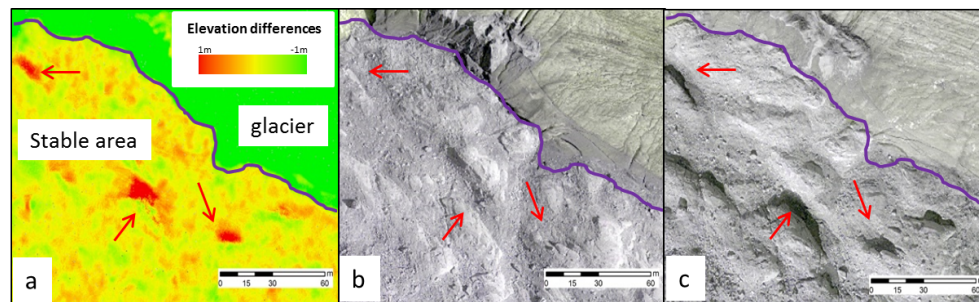
Varying contrast on the images was however found to potentially generate local DSM inaccuracies. Important elevation differences occurred, for example, when subtracting summer and autumn DSMs of Findelen- or Gries- gletscher over stable areas (cf. Figure 1) that strongly differed in terms of illumination between the two campaigns (Figure 7). As known from classical photogrammetry (e.g., [48]), sharp shadows were particularly prone to generate such differences.

The percentage of image overlap can also impact the DSM accuracy. This was shown in the study of Rosnell and Honkavaara [29], for instance, who performed several UAV flights over the same area, varying the longitudinal ground overlap from 60% to 90%. They found that increasing the overlap clearly improves the DSM accuracy, due to a better estimation of the image block orientation parameters.

The relation between vertical accuracy and distance to the closest GCP highlights the importance of a spatially-distributed GCP network for optimizing the DSM accuracy. An example comparing two DSMs built with the same number of GCPs, but with a different spatial distribution (distributed vs. clustered) is shown in Figure 6b,c. Here, it is visible that the difference between the generated DSMs and the continuous dGPS points increases with distance and that this difference does not increase linearly, i.e., both positive and negative differences can occur depending on location. Figure 6d depicts a network of GCPs located at the edge of the survey area, which is a common configuration in traditional photogrammetry. This figure shows that in an area where the distance between GCPs is large, the DSMs present a "dome" or "bowl" effect (e.g., [49,50]). A dome effect can appear if: (1) the GCP density is not sufficient; (2) the images are taken in near-parallel (nadir) directions; and (3) inexact camera calibration models are used (e.g., [51]). Based on our results, we recommend to place GCPs in a spatially distributed network for UAV surveys.

Figure 6. (**a**) Distance to the closest GCP used for DSM generation against absolute elevation differences between DSM and CPs. The data refer to all campaigns (cf. Table 1). The light blue line is a linear fit of the 50th percentiles of the elevation differences between the DSMs and the check points, calculated for 10-m distance bins. The red lines are linear fits of the 2.5 and 97.5 percentiles. (**b–d**) Elevation differences between continuous dGPS points and a DSM built with (**b**) three clustered GCPs on the glacier tongue, (**c**) three spatially distributed GCPs and (**d**) GCPs located at the edge of the survey area. The example is from the Findelengletscher autumn campaign.

Figure 7. (**a**) Elevation differences between summer and autumn DSMs. Orthophoto of the same area in (**b**) summer and (**c**) autumn. The glacier outline (violet line) and features changing appearance because of different illumination (red arrows) are highlighted.

4.3. Impact of Surface Characteristics on DSM Accuracy

The test performed with images from different surface characteristics shows contrasting results (Table 4). When processing the fresh snow images, only two of them oriented (Section 3.1, Step 4), providing 330 matching keypoints in total. For one-day-old snow, 19 out of the 20 images oriented, yielding a total of 34,800 matching keypoints. For bare ice, all 20 images oriented, providing over 380,000 matching keypoints. This indicates that the lack of structure of fresh snow makes it impossible to generate DSMs or orthophotos. However, even after being exposed to sunlight for one day only, snow seems to develop enough structure and contrast to allow for image orientation. For each surface type, Table 4 shows the number of points generated for the sparse and dense point cloud (Section 3.1, Steps 4 and 7, respectively). Although the number of matching keypoints for one-day-old snow is 10-times lower than for ice, the number of points in the dense point cloud is only 1.5-times smaller. This shows that the number of matching keypoints does not significantly affect the performance of the surface reconstruction algorithm, which can explain the similar accuracies obtained for our datasets in summer and winter (cf. Table 2). Furthermore, Piermattei et al. [52] showed that images do not orient on fresh snow surfaces, which can yield incomplete DSM coverage. They also highlighted that surface texture has an impact on the density of the final dense point cloud. In their case, the point density for fresh-snow surfaces was three- to four-times lower than for debris, ice or firn. We thus conclude that in the case of fresh snow, too few matching keypoints can be found to allow for image orientation, whilst in the case of other surface textures, the number of matching points has a minor influence on DSM accuracy.

Table 4. Number of oriented images, matching keypoints and points in the sparse point cloud after the "align" step (Section 3.1, Step 4) in Agisoft Photoscan. The number of point in the dense point cloud is determined after the "build dense cloud" step (Section 3.1, Step 7).
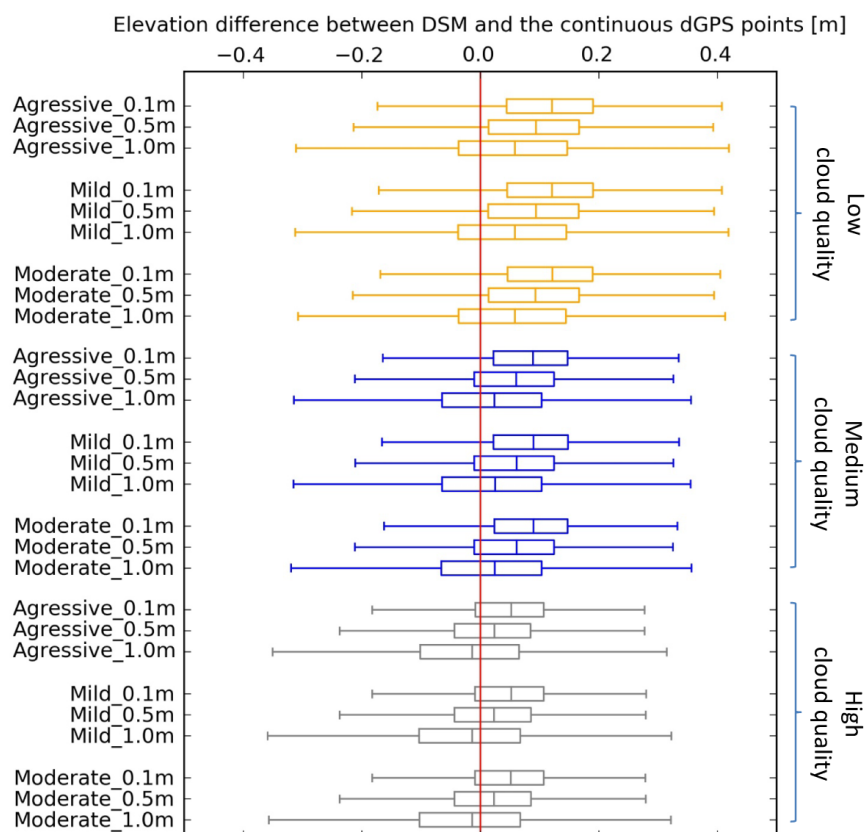
Surface Texture	Oriented Images	Matching Keypoints	Sparse Point Cloud	Dense Point Cloud
Fresh snow	2 of 20	330	143	115,250
One-day-old snow	19 of 20	34,800	15,400	1,032,700
Ice	20 of 20	380,000	118,700	1,540,000

4.4. Effect of the Parameters Settings in Agisoft Photoscan on DSM Accuracy

Our results indicate that there is virtually no difference between different depth filters on the final DSM accuracy and that the influence of switching the cloud quality parameter from "low" to "high" is in the order of a decimeter only (Figure 8). Note, however, that this assessment was performed in a rather flat area, in which even a low density point cloud might capture the mean surface elevation well. The influence of the cloud quality parameter might thus be larger in steeper terrain. The small

accuracy difference between DSMs generated with different cloud quality parameters is very much in contrast to the required computational time. In fact, changing the parameter from "low" to "high", for example, increased the required computational time by a factor of eighteen.

The DSM cell size was found to affect the mean elevation of the DSM itself. The mean elevation difference between DSMs and continuous dGPS points for DSMs with a 0.1-m cell size is 0.08 m. This difference decreases to 0.05 m for DSMs with a cell size of 0.5 m and to 0.02 m for DSMs with a cell size of 1.0 m. A dependence between the mean DSM elevation and cell size might be surprising at first. It has to be noted, however, that the elevation distribution of small-scale topographic features is not symmetric. Consider a 1-m^2-portion of a given surface for example. The probability of observing small scale features at a given elevation (say 1 m) above the mean elevation of this surface portion is obviously smaller than the probability of observing a feature that is the same distance below the mean. In the example and loosely speaking: at 1 m above-average elevation, the observer is very likely to be in free air, whilst at 1 m below-average elevation it is likely to be completely in ice. As the DSM cell size grows and small-scale features are less and less captured, the mean elevation is thus expected to decrease. This reasoning is obviously not true for any given cell size, but holds true as long as the cell size is in the same order of magnitude as the typical surface roughness.

Figure 8. Elevation differences between a DSM generated on the tongue of the Findelengletescher (autumn campaign) and the continuous dGPS points. The DSM was built with different parameters in Agisoft Photoscan for the step "build dense cloud": "aggressive", "mild" or "moderate" depth filters (**left**) and "low", "medium" or "high" cloud quality (**right**). The different DSM cell size is displayed on the left side of the plot. The boxplots show the median, interquartile range (box) and 95% confidence interval (whiskers).

5. Conclusions

Our study assessed the accuracy of DSMs derived over glaciers through UAV photogrammetry with a focus on the influence of the number and disposition of GCPs. We generated over a thousand DSMs with different GCP combinations and cross-validated them with precise validation points collected on the ground. Our results show that the vertical and horizontal accuracy is asymptotically related to the increasing number of GCPs. Combining all datasets, we found the maximal vertical (horizontal) accuracy to be between 0.10 and 0.25 m (0.03 and 0.09 m). To achieve this, a density of 6.12×10^{-8} GCP/GSD (2.52×10^{-8} GCP/GSD) was necessary for DSMs' georeferencing. We showed that on average, the local accuracy decreases with the distance to the closest GCP at a rate of about 0.09 m per 100-m distance. The large scatter of this relation, however, shows that other factors additionally impact the local accuracy. The surface type (i.e., ice or snow) was found not to significantly affect the DSM accuracy, as long as the surface is not covered with fresh snow. In the latter case, in fact, the used image-matching algorithm did not detect enough matching keypoints to orient the images. Different parameter choices in Agisoft Photoscan and variations in DSM cell size can additionally influence the DSM accuracy. In particular, the mean elevation can be affected by up to a decimeter. Our work highlights that DSMs with similar accuracy can be obtained for all seasons over glaciers and helps the design of efficient UAV field surveys.

Author Contributions: All authors conceived of and designed the experiments and analyzed the results. Saskia Gindraux analyzed the data and wrote the manuscript with contributions from Daniel Farinotti.

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

References

1. Grenzdörffer, G.J.; Engel, A.; Teichert, B. The photogrammetric potential of low-cost UAVs in forestry and agriculture. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *31*, 1207–1214.
2. Costa, F.G.; Ueyama, J.; Braun, T.; Pessin, G.; Osório, F.S.; Vargas, P.A. The use of unmanned aerial vehicles and wireless sensor network in agricultural applications. In Proceedings of 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 5045–5048.
3. Sauerbier, M.; Eisenbeiss, H. UAVs for the documentation of archaeological excavations. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2010**, *38*, 526–531.
4. Chiabrando, F.; Nex, F.; Piatti, D.; Rinaudo, F. UAV and RPV systems for photogrammetric surveys in archaelogical areas: Two tests in the Piedmont region (Italy). *J. Archaeol. Sci.* **2011**, *38*, 697–710.
5. Lucieer, A.; Robinson, S.; Turner, D.; Harwin, S.; Kelcey, J. Using a micro-UAV for ultra-high resolution multi-sensor observations of Antarctic moss beds. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *39*, 429–433.
6. Anderson, K.; Gaston, K.J. Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Front. Ecol. Environ.* **2013**, *11*, 138–146.
7. DeBell, L.; Anderson, K.; Brazier, R.; King, N.; Jones, L. Water resource management at catchment scales using lightweight UAVs: Current capabilities and future perspectives. *J. Unmanned Veh. Syst.* **2016**, *4*, 7–30.
8. Hoffmann, H.; Nieto, H.; Jensen, R.; Guzinski, R.; Zarco-Tejada, P.; Friborg, T. Estimating evaporation with thermal UAV data and two-source energy balance models. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 697–713.
9. Colomina, I.; Molina, P. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 79–97.
10. Whitehead, K.; Hugenholtz, C.H. Remote sensing of the environment with small unmanned aircraft systems (UASs), part 1: A review of progress and challenges 1. *J. Unmanned Veh. Syst.* **2014**, *2*, 69–85.
11. Nex, F.; Remondino, F. UAV for 3D mapping applications: A review. *Appl. Geomat.* **2014**, *6*, 1–15.

12. Fugazza, D.; Senese, A.; Azzoni, R.S.; Smiraglia, C.; Cernuschi, M.; Severi, D.; Diolaiuti, G.A. High resolution mapping of glacier surface features. The UAV survey of the Forni Glacier (Stelvio National Park, Italy). *Geogr. Fisica E Din. Quat.* **2015**, *38*, 25–33.

13. Kraaijenbrink, P.; Shea, J.; Pellicciotti, F.; Jong, S.d.; Immerzeel, W. Object-based analysis of unmanned aerial vehicle imagery to map and characterise surface features on a debris-covered glacier. *Remote Sens. Environ.* **2016**, *186*, 581–595.

14. d'Oleire Oltmanns, S.; Marzolff, I.; Peter, K.; Ries, J. Unmanned Aerial Vehicle (UAV) for Monitoring Soil Erosion in Morocco. *Remote Sens.* **2012**, *4*, 3390–3416.

15. Mancini, F.; Dubbini, M.; Gattelli, M.; Stecchi, F.; Fabbri, S.; Gabbianelli, G. Using Unmanned Aerial Vehicles (UAV) for High-Resolution Reconstruction of Topography: The Structure from Motion Approach on Coastal Environments. *Remote Sens.* **2013**, *5*, 6880–6898.

16. Baltsavias, E.P.; Favey, E.; Bauder, A.; Bösch, H.; Pateraki, M. Digital surface modelling by airborne laser scanning and digital photogrammetry for glacier monitoring. *Photogramm. Rec.* **2001**, *17*, 243–273.

17. Barrand, N.E.; Murray, T.; James, T.D.; Barr, S.L.; Mills, J.P. Optimizing photogrammetric DEMs for glacier volume change assessment using laser-scanning derived ground-control points. *J. Glaciol.* **2009**, *55*, 106–116.

18. Kenner, R.; Bühler, Y.; Delaloye, R.; Ginzler, C.; Phillips, M. Monitoring of high alpine mass movements combining laser scanning with digital airborne photogrammetry. *Geomorphology* **2014**, *206*, 492–504.

19. Fischer, M.; Huss, M.; Kummert, M.; Hoelzle, M. Use of an ultra-long-range terrestrial laser scanner to monitor the mass balance of very small glaciers in the Swiss Alps. *Cryosphere* **2016**, *10*, 1279–1295.

20. Clapuyt, F.; Vanacker, V.; Van Oost, K. Reproducibility of UAV-based earth topography reconstructions based on Structure-from-Motion algorithms. *Geomorphology* **2016**, *260*, 4–15.

21. Rock, G.; Ries, J.B.; Udelhoven, T. Sensitivity Analysis of UAV-Photogrammetry for Creating Digital Elevation Models (DEM). In Proceedings of the Conference on Unmanned Aerial Vehicle in Geomatics, Zurich, Switzerland, 14–16 September 2011.

22. Leitão, J.P.; Moy de Vitry, M.; Scheidegger, A.; Rieckermann, J. Assessing the quality of digital elevation models obtained from mini unmanned aerial vehicles for overland flow modelling in urban areas. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 1637–1653.

23. Nouwakpo, S.K.; Weltz, M.A.; McGwire, K. Assessing the performance of structure-from-motion photogrammetry and terrestrial LiDAR for reconstructing soil surface microtopography of naturally vegetated plots: SfM and LiDAR Performance on Vegetated Plots. *Earth Surf. Process. Landf.* **2015**, *41*, 308–322.

24. Smith, M.W.; Carrivick, J.L.; Quincey, D.J. Structure from motion photogrammetry in physical geography. *Prog. Phys. Geogr.* **2015**, *40*, 1–29.

25. Kääb, A.; Girod, L.; Berthling, I. Surface kinematics of periglacial sorted circles using structure-from-motion technology. *Cryosphere* **2014**, *8*, 1041–1056.

26. Shahbazi, M.; Sohn, G.; Théau, J.; Menard, P. Development and Evaluation of a UAV-Photogrammetry System for Precise 3D Environmental Modeling. *Sensors* **2015**, *15*, 27493–27524.

27. Tahar, K.N.; Ahmad, A.; Akib, W.A.A.W.M.; Mohd, W.M.N.W. Assessment on Ground Control Points in Unmanned Aerial System Image Processing for Slope Mapping Studies. *Int. J. Sci. Eng. Res.* **2012**, *3*, 1–10.

28. Tahar, K.N. An evaluation on different number of ground control points in unmanned aerial vehicle photogrammetric block. In Proceedings of the ISPRS, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Istanbul, Turkey, 27–29 November 2013; Volume XL-2/W2.

29. Rosnell, T.; Honkavaara, E. Point Cloud Generation from Aerial Image Data Acquired by a Quadrocopter Type Micro Unmanned Aerial Vehicle and a Digital Still Camera. *Sensors* **2012**, *12*, 453–480.

30. Tonkin, T.N.; Midgley, N.G. Ground-Control Networks for Image Based Surface Reconstruction: An Investigation of Optimum Survey Designs Using UAV Derived Imagery and Structure-from-Motion Photogrammetry. *Remote Sens.* **2016**, *8*, 1–8.

31. Immerzeel, W.; Kraaijenbrink, P.; Shea, J.; Shrestha, A.; Pellicciotti, F.; Bierkens, M.; de Jong, S. High-resolution monitoring of Himalayan glacier dynamics using unmanned aerial vehicles. *Remote Sens. Environ.* **2014**, *150*, 93–103.

32. Ryan, J.C.; Hubbard, A.L.; Box, J.E.; Todd, J.; Christoffersen, P.; Carr, J.R.; Holt, T.O.; Snooke, N. UAV photogrammetry and structure from motion to assess calving dynamics at Store Glacier, a large outlet draining the Greenland ice sheet. *Cryosphere* **2015**, *9*, 1–11.

33. Bühler, Y.; Adams, M.S.; Bösch, R.; Stoffel, A. Mapping snow depth in alpine terrain with unmanned aerial systems (UASs): Potential and limitations. *Cryosphere* **2016**, *10*, 1075–1088.

34. Harder, P.; Schirmer, M.; Pomeroy, J.; Helgason, W. Accuracy of snow depth estimation in mountain and prairie environments by an unmanned aerial vehicle. *Cryosphere* **2016**, *10*, 2559–2571.

35. Klokan Technologies GmbH, EPSG.io: Coordinate Systems Worldwide. Available online: https://epsg.io/2056 (accessed on 20 February 2017).

36. SenseFly SA, a Parrot Company. Available online: http://www.sensefly.com (accessed on 20 February 2017).

37. Agisoft LCC, Agisoft PhotoScan. Available online: http://www.agisoft.com (accessed on 20 February 2017).

38. Fonstad, M.A.; Dietrich, J.T.; Courville, B.C.; Jensen, J.L.; Carbonneau, P.E. Topographic structure from motion: A new development in photogrammetric measurement. *Earth Surf. Process. Landf.* **2013**, *38*, 421–430.

39. Westoby, M.; Brasington, J.; Glasser, N.; Hambrey, M.; Reynolds, J. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* **2012**, *179*, 300–314.

40. Seitz, S.M.; Curless, B.; Diebel, J.; Scharstein, D.; Szeliski, R. A comparison and evaluation of multi-view stereo reconstruction algorithms. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006; Volume 1.

41. Cline, D. Measuring Alpine Snow Depths by Digital Photogrammetry: Part 1. Conjugate Point Identification. In Proceedings of the Eastern Snow Conference, Quebec City, QC, Canada, 8–10 June 1993;

42. Jagt, B.; Lucieer, A.; Wallace, L.; Turner, D.; Durand, M. Snow Depth Retrieval with UAS Using Photogrammetric Techniques. *Geosciences* **2015**, *5*, 264–285.

43. Jaud, M.; Passot, S.; Le Bivic, R.; Delacourt, C.; Grandjean, P.; Le Dantec, N. Assessing the Accuracy of High Resolution Digital Surface Models Computed by PhotoScan® and MicMac® in Sub-Optimal Survey Conditions. *Remote Sens.* **2016**, *8*, 1–18.

44. Vallet, J.; Panissod, F.; Strecha, C.; Tracol, M. Photogrammetric performance of an ultra light weight swinglet UAV. In Proceedings of the International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g) (IAPRS), Zürich, Switzerland, 14–16 September 2011; Volume 38.

45. James, M.R.; Robson, S. Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. *J. Geophys. Res.* **2012**, *117*, 1–17.

46. Whitehead, K.; Moorman, B.J.; Hugenholtz, C.H. Brief Communication: Low-cost, on-demand aerial photogrammetry for glaciological measurement. *Cryosphere* **2013**, *7*, 1879–1884.

47. 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.

48. Luhmann, T.; Robson, S.; Kyle, S.; Boehm, J. *Close-Range Photogrammetry and 3d Imaging*, 2nd ed.; Walter de Gruyter: Berlin, Germany, 2014.

49. Nouwakpo, S.K.; James, M.R.; Weltz, M.A.; Huang, C.H.; Chagas, I.; Lima, L. Evaluation of structure from motion for soil microtopography measurement. *Photogramm. Rec.* **2014**, *29*, 297–316.

50. Ouédraogo, M.M.; Degré, A.; Debouche, C.; Lisein, J. The evaluation of unmanned aerial system-based photogrammetry and terrestrial laser scanning to generate DEMs of agricultural watersheds. *Geomorphology* **2014**, *214*, 339–355.

51. James, M.R.; Robson, S. Mitigating systematic error in topographic models derived from UAV and ground-based image networks. *Earth Surf. Process. Landf.* **2014**, *39*, 1413–1420.

52. Piermattei, L.; Carturan, L.; de Blasi, F.; Tarolli, P.; Dalla Fontana, G.; Vettore, A.; Pfeifer, N. Suitability of ground-based SfM–MVS for monitoring glacial and periglacial processes. *Earth Surf. Dyn.* **2016**, *4*, 425–443.