Accuracy of snow depth estimation in mountain and prairie environments by an unmanned aerial vehicle

Phillip Harder[1], Michael Schirmer[1, 3], John Pomeroy[1], Warren Helgason[1, 2]

[1]Centre for Hydrology,
University of Saskatchewan,
Saskatoon, Saskatchewan
Canada

[2]Department of Civil, Geological, and Environmental Engineering,
University of Saskatchewan,
Saskatoon, Saskatchewan
Canada

[3]now at WSL Institute for Snow and Avalanche Research SLF
Davos, Switzerland

All correspondence to Phillip Harder: phillip.harder@usask.ca

Abstract

Quantifying the spatial distribution of snow is crucial to predict and assess its water resource potential and understand land-atmosphere interactions. High-resolution remote sensing of snow depth has been limited to terrestrial and airborne laser scanning and more recently with application of Structure from Motion (SfM) techniques to airborne (manned and unmanned) imagery. In this study, photography from a small unmanned aerial vehicle (UAV) was used to generate digital surface models (DSMs) and orthomosaics for snowcovers at a cultivated agricultural Canadian Prairie and a sparsely-vegetated Rocky Mountain alpine ridgetop site using SfM. The accuracy and repeatability of this method to quantify snow depth, changes in depth and its spatial variability was assessed for different terrain types over time. Root mean square errors in snow depth estimation from differencing snow covered and non-snow covered DSMs were 8.8 cm for a short prairie grain stubble surface, 13.7 cm for a tall prairie grain stubble surface and 8.5 cm for an alpine mountain surface. This technique provided useful information on maximum snow accumulation and snow-covered area depletion at all sites, while temporal changes in snow depth could also be quantified at the alpine site due to the deeper snowpack and consequent higher signal-to-noise ratio. The application of SfM to UAV photographs returns meaningful information in areas with mean snow depth > 30 cm, however the direct observation of snow depth depletion of shallow snowpacks with this method is not feasible. Accuracy varied with surface characteristics, sunlight and wind speed during the flight, with the most consistent performance found for wind speeds < 10 ms^{-1}, clear skies, high sun angles and surfaces with negligible vegetation cover.

1. Introduction

Accumulation, redistribution, sublimation and melt of seasonal or perennial snowcovers are defining features of cold region environments. The dynamics of snow have incredibly important impacts on land-atmosphere interactions and can constitute significant proportions of the water resources necessary for socioeconomic and ecological functions (Armstrong and Brun, 2008; Gray and Male, 1981; Jones et al., 2001). Snow is generally quantified in terms of its snow water equivalent (SWE) through measurements of its depth and density. Since density varies less than depth (López-Moreno et al., 2013; Shook and Gray, 1996) much of the spatial variability of SWE can be described by the spatial variability of snow depth. Thus, the ability to measure snow depth and its spatial distribution is crucial to assess and predict how the snow water resource responds to meteorological variability and landscape heterogeneity. Observation and prediction of the spatial distribution of snow depth is even more relevant with the anticipated and observed changes occurring due to a changing climate and land use (Dumanski et al., 2015; Harder et al., 2015; Milly et al., 2008; Mote et al., 2005; Stewart et al., 2004).

The many techniques and sampling strategies employed to quantify snow depth all have strengths and limitations (Pomeroy and Gray, 1995). Traditionally, manual snow surveys have been used to quantify snow depth and density along a transect. The main benefit of manual snow surveying is that the observations are a direct measurement of the SWE; however, it requires significant labour, is a destructive sampling method and can be impractical in complex, remote or hazardous terrain (DeBeer and Pomeroy, 2009; Dingman, 2002). Many sensors exist that can measure detailed snow properties non-destructively, with a comprehensive review found in Kinar and Pomeroy (2015), but non-destructive automated sensors, such as acoustic snow depth rangers (Campbell Scientific SR50) or SWE analyzers (Campbell Scientific CS275 Snow Water Equivalent Sensor), typically only provide point scale information and may require significant additional infrastructure or maintenance to operate properly. Remote sensing of snow from satellite and aerial platforms quantify snow extent at large scales. Satellite platforms can successfully estimate snow-covered area but problems remain in quantifying snow depth, largely due to the heterogeneity of terrain complexity and vegetation cover. To date, Light Detection And Ranging (LiDAR) techniques have provided the highest resolution estimates of snow depth spatial distribution from both terrestrial (Grünewald et al., 2010) and airborne platforms (Hopkinson et al., 2012). The main limitations encountered are easily observable areas (sensor viewshed) for the terrestrial scanner and the prohibitive expense and long lead time needed for planning repeat flights for the aerial scanner (Deems et al., 2013). Typically, airborne LiDAR provides data with a ground sampling of nearly 1 m and a vertical accuracy of 15 cm (Deems and Painter, 2006; Deems et al., 2013). While detailed, this resolution still does not provide observations of the spatial variability of snow distributions that can address microscale processes such as snow-vegetation interactions or wind redistribution in areas of shallow snowcover, and the frequency of airborne LiDAR observations are typically low, except for NASA's Airborne Snow Observatory applications in California (Mattmann et al., 2014).

An early deployment of a high resolution digital camera on a remote controlled gasoline powered model helicopter in 2004 permitted unmanned digital aerial photography to support studies of shrub emergence and snowcovered area depletion in a Yukon mountain shrub tundra environment (Bewley et al., 2007). Since then, Unmanned Aerial Vehicles (UAVs) have become increasingly popular for small-scale high-resolution remote sensing applications in the earth sciences. The current state of the technology is due to advances in the capabilities and miniaturization of the hardware comprising UAV platforms (avionics/autopilots, Global-Positioning Systems (GPS), Inertial Momentum Units (IMUs) and cameras) and the increases in computational power for processing imagery. The conversion of raw images to orthomosaics and Digital Surface Models (DSMs) takes advantage of Structure from Motion (SfM) algorithms (Westoby et al., 2012). These computationally intensive algorithms simultaneously resolve camera pose and scene geometry through automatic identification and matching of common features in multiple images. With the addition of information on the respective camera location, or if feature locations are known, then georeferenced point clouds, orthomosaics and DSMs can be generated (Westoby et al., 2012). Snow is a challenging surface for SfM techniques due to its relatively uniform surface and high reflectance relative to snow-free areas, which limit identifiable features (Nolan et al., 2015). The resolution of the data products produced by UAVs depends largely on flight elevation and sensor characteristics but can promise accuracies of 2.6 cm in the horizontal and 3.1 cm in the vertical (Roze et al., 2014). The unprecedented spatial resolution of these products may be less important than the fact that these platforms are deployable at a high user-defined frequencies below cloud cover, which can be problematic for airborne or satellite platforms. Manned aerial platforms have the advantage of covering much larger areas (Nolan et al., 2015) with a more mature and clear regulatory framework (Marris, 2013; Rango and Laliberte, 2010) than small UAVs. However, the greater expenses associated with acquisition, maintenance, operation and training required for manned platforms (Marris, 2013), relative to small UAVs, are significant (Westoby et al., 2012). Many snow scientists have expressed great enthusiasm in the opportunities UAVs present and speculate that they may drastically change the quantification of snow accumulation and ablation (Sturm, 2015).

The roots of SfM are found in stereoscopic photogrammetry, which has a long history in topographic mapping (Collier, 2002). Relative to traditional photogrammetry, major advances in the 1990's in computer vision (Boufama et al., 1993; Spetsakis and Aloimonost, 1991; Szeliski and Kang, 1994) has automated and simplified the data requirements to go from a collection of overlapping 2D images to 3D point clouds. Significant work by the geomorphology community has pushed the relevance, application and further development of this technique into the earth sciences (Westoby et al., 2012). Recent application of this technique to snow depth estimation has used imagery captured by manned aerial platforms (Bühler et al., 2015; Nolan et al., 2015) and increasingly with small UAVs (Vander Jagt et al., 2015; Bühler et al., 2016; De Michele et al., 2016). The manned aircraft examples have reported vertical accuracies of 10cm (Nolan et al., 2015) and 30 cm (Bühler et al., 2015) with horizontal resolutions of 5-20 cm (Nolan et al., 2015) and 2 m (Bühler et al., 2015). Unmanned aircraft examples have shown similar accuracies and resolution with vertical errors of reported to be ~10 cm with horizontal resolutions between 50 cm (Vander Jagt et al., 2015) and 10 cm (Bühler et al., 2016). The accuracy assessments of the De Michele et al. (2016), Vander Jagt et al. (2015), and Bühler et al. (2016) studies were limited to a small number of snow depth maps. Bühler et al. (2016) had the most with four maps, but more are needed to get a complete perspective on the performance of this technique and its repeatability under variable conditions.

The overall objective of this paper is to assess the accuracy of snow depth as estimated by imagery collected by small UAVs and processed with SfM techniques. Specifically, this paper will: 1) assess the accuracy of UAV-derived snow depths with respect to the deployment conditions and heterogeneity of the earth surface, specifically variability in terrain relief, vegetation characteristics and snow depth; and 2) identify and assess opportunities for UAV generated data to advance understanding and prediction of snowcover and snow depth dynamics.

2. Sites and Methodology

2.1 Sites

The prairie field site (Fig. 1a) is representative of agricultural regions on the cold, windswept Canadian Prairies, where agriculture management practices control the physical characteristics of the vegetation which, in turn, influence snow accumulation (Pomeroy and Gray, 1995). There is little elevation relief and the landscape is interspersed with wooded bluffs and wetlands. Snowcover is typically shallow (maximum depth < 50 cm) with development of a patchy and dynamic snow-covered area during melt. Data collection occurred at a field site near Rosthern, Saskatchewan, Canada Mountains (52° 42’ N, 106° 27’ W) in spring 2015 as part of a larger project studying the influence of grain stubble exposure on snowmelt processes. The 0.65 km^2 study site was divided into areas of tall stubble (35 cm) and short stubble (15 cm). The wheat stubble (Fig. 1c), clumped in rows ~30 cm apart, remained erect throughout the snow season, which has implications for blowing snow accumulation, melt energetics and snow cover depletion. Pomeroy et al. (1993, 1998) describes the snow accumulation dynamics and snowmelt energetics of similar environments.

The alpine site, located in Fortress Mountain Snow Laboratory in the Canadian Rocky Mountains (50° 50’ N, 115° 13’ W), is characterized by a ridge oriented in SW-NE direction (Fig. 1b, d) at an elevation of approximately 2300 m. The average slope at the alpine site is ~15 degrees with some slopes > 35 degrees. Large areas of the ridge were kept bare by wind erosion during the winter of 2014/2015 and wind redistribution caused the formation of deep snowdrifts on the leeward (SE) side of the ridge, in surface depressions and downwind of krummholz. Vegetation is limited to short grasses on the ridgetop while shrubs and coniferous trees become more prevalent in gullies on the shoulders of the ridge. Mean snow depth of the snow-covered area at the start of the observation period (May 13, 2015) was 2 m (excluding snow-free areas) with maximum depths over 5 m. The 0.32 km^2 study area was divided between a North and a South area (red polygons in Fig. 1b) due to UAV battery and hence flight area limitations. DeBeer and Pomeroy (2010, 2009) and MacDonald et al. (2010) describe the snow accumulation dynamics and snowmelt energetics of the area.

2.2 Methodology

2.2.1 Unmanned Aerial Vehicle - flight planning - operation - data processing

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

At the prairie site, the UAV was flown 22 times over the course of the melt period (March 6 to 30, 2015) with three flights over the snow free surface between April 2 and 9, 2015. A loaner Ebee, from Spatial Technologies, the Ebee distributor, performed the first 11 flights at the prairie site due to technical issues with the Ebee RTK. The geotag errors of the non-RTK loaner Ebee were ±5 m (error of GPS Standard Positioning Service) and therefore required GCPs to generate georeferenced data products. At the alpine site, to reduce variations in the height of the UAV above the surface in complex terrain, flight plans were adjusted using a 1 m resolution DEM, derived from a LiDAR DEM. The UAV was flown 18 times over melt from 15 May to 24 June 2015 with four flights over bare ground on 24 July 2015. Table 1 summarises flight plan attributives of the respective sites. Figure 2b shows a typical flight plan generated by the eMotion flight control software for the prairie site.

Postflight Terra 3D 3 (version 3.4.46) processed the imagery to generate DSMs and orthomosaics. Though the manufacturer suggested that they are unnecessary with RTK corrected geotags (error of ±2.5 cm), all processing included GCPs. At the prairie site, 10 GCPs comprised of five tarps and five utility poles were distributed throughout the study area (blue points in Fig. 1a). At the alpine site, the north and south areas had five and six GCPs (blue points in Fig. 1b), respectively comprised of tarps (Fig. 3a) and easily identifiable rocks (Fig. 3b) spread over the study area.

Processing involved three steps. First, initial processing extracted features common to multiple images, optimized external and internal camera parameters for each image, and generated a sparse point cloud. The second step densified the point cloud and the third step generated a georeferenced orthomosaic and a DSM. Preferred processing options varied between the sites, with the semi-global matching algorithm in the point densification used to minimize erroneous points encountered at the alpine site (see Sect 3.3).

Generated orthomosaics and DSMs had a horizontal resolution of 3.5 cm at the prairie site and between 3.5 cm and 4.2 cm at the alpine site.

2.2.2 Ground truth and snow depth data collection

To assess the accuracy of the generated DSMs and their ability to measure snow depth, detailed observations of the land surface elevation and snow depth were collected. At the prairie site a GNSS survey, utilizing a Leica GS15 as a base station and another GS15 acting as a RTK corrected rover, measured the location (x, y and z) of 17 snow stakes on each stubble treatment to an accuracy of less than ±2.5 cm. This gives 34 observation points at the prairie site (locations identified as red dots in Fig. 1a). Over the melt period, the snow depth was measured with a ruler at each point (error of ±1 cm). Adding the manually measured snow depths to the corresponding land surface elevations from the GNSS survey gives snow surface elevations at each observation point directly comparable to the UAV derived DSM. At the alpine site, 100 land surface elevations were measured at points with negligible vegetation (bare soil or rock outcrops) with a GNSS survey to determine the general quality of the DSMs. For eight flights a GNSS survey was also performed on the snowcover (all measurement locations over the course of campaign are highlighted in Fig. 1b). To account for the substantial terrain roughness and to avoid measurement errors in deep alpine snowpacks, snow surface elevation was measured via GNSS survey and snow depth estimated from the average of five snow depth measurements in a 0.4 m x 0.4 m square at that point. Time constraints and inaccessible steep snow patches limited the number of snow depth measurements to between three and 19 measurements per flight. While the number of accuracy assessment points over snow is limited for each flight the cumulative number of points over the course of the campaigns used to assess accuracy over all flights is not; at the alpine site there were 101 GNSS surface measurements and 83 averaged snow depth measurements available, and at the prairie site 323 measurements on each stubble treatment.

At both the prairie and alpine site, the same GNSS RTK surveying method established GCP locations. Snow surveys (maximum one per day) and DSMs (multiple per day) are only compared if from the same days.

2.2.3 Snow depth estimation

Subtracting a DSM of a snow free surface from a DSM of a snow covered surface estimates snow depth assuming snow ablation is the only process changing the surface elevations between observation times. Vegetation is limited over the areas of interest at the alpine site and any spring up of grasses or shrubs is insignificant, based upon local observations, with respect to the large snow depths observed (up to 5 m). The wheat stubble at the prairie site is unaffected by snow accumulation or ablation. The snow-free DSMs corresponded to imagery collected on for the prairie site and July 24, 2015 for the alpine site.

2.2.4 Accuracy assessment

The accuracy of the UAV-derived DSM and snow depth was estimated by calculating the root mean square error (RMSE), mean error (bias) and standard deviation of the error (SD) with respect to the manual measurements. The RMSE quantifies the overall difference between manually measured and UAV derived values, bias quantifies the mean magnitude of the over (positive values) or under (negative values) prediction of the DSM with respect to manual measurements, and SD quantifies the variability of the error.

2.2.5 Signal-to-Noise Calculation

The signal-to-noise ratio (SNR) compares the level of the snow depth signal with respect to the measurement error to inform when meaningful information is available. The SNR is calculated as the mean measured snow depth value divided by the standard deviation of the error between the observed and estimated snow depths. The Rose criterion (Rose 1973), commonly used in the image processing literature, is used to define the threshold SNR where the UAV returns meaningful snow depth information. The Rose criterion proposes a SNR ≥ 4 for the condition at which the signal is sufficiently large to avoid mistaking it for a fluctuation in noise. Ultimately, the acceptable signal to noise ratio depends upon the user's error tolerance (Rose, 1973).

3. Results and Discussion

3.1 Absolute surface accuracy

The accuracy of the DSMs relative to the measured surface points varies with respect to light conditions at time of photography and differences in snow surface characteristics and extent. This is seen in the RMSE for individual flights varying from 4 cm to 19 cm (Fig. 4). Only a few problematic flights, which will be discussed in section 3.3.1, showed larger RMSEs, which are marked in blue in Figure 4. In general, the accuracy of the DSMs as represented by the mean RMSEs in Table 2, were comparable between the prairie short stubble (8.1 cm), alpine-bare (8.7 cm) and alpine-snow (7.5 cm) sites and were greater over the prairie tall stubble (11.5 cm). Besides the five (out of a total of 43) problematic flights (out of a total of 43 flights), accuracy was relatively consistent over time at all sites. More specifically, the prairie flights simultaneously sampled the short and tall stubble areas, thus there were only three problematic flights at the prairie site in addition to the two at the alpine site (Fig. 4). The larger error at the tall stubble is due to snow and vegetation surface interactions. Over the course of melt, the DSM gradually became more representative of the stubble surface rather than the snow surface. More points are matched on the high contrast stubble than the low contrast snow leading to the DSM being biased to reflect the stubble surface. This is apparent in the increasing tall stubble bias as the snow surface drops below the stubble height. By comparing the many alpine-bare points to the limited number of alpine-snow points (3 to 19) the relative difference in errors between the snow and non-snow surfaces was assessed. The benefit of the large amount of alpine-bare points (100) revealed the general errors, offsets and tilts in the DSM. It was concluded that the snow surface errors are not appreciably different from the non-snow surface errors.

The RTK level accuracy of the camera geotags should produce products with similar accuracy, without the use of GCPs, as those generated with standard GPS positioning and the use of GCPs (Roze et al., 2014). DSMs created with and without GCPs for flights where the Ebee's camera geotags had RTK-corrected positions with an accuracy of ±2.5 cm tested this claim. Nine flights from the prairie site and 22 flights from the alpine site met the requirements for this test. Inclusion of GCPs had little effect on the standard deviation of error with respect to surface observations, but resulted in a reduction of the mean absolute error of the bias from 27 cm to 10 cm and from 14 cm to 6 cm at the prairie and alpine sites, respectively.

3.2 Snow depth accuracy

The snow depth errors were similar to that of the surface errors with the alpine and short stubble sites having very similar errors, with mean RMSEs of 8.5 cm and 8.8 cm, but much larger errors over tall stubble, with a mean RMSE of 13.7 cm (Fig. 5 and Table 3). Snow depth errors were larger than the surface errors as the errors from the snow-free and snow-covered DSMs are additive in the DSM differencing. The usability of snow depth determined from DSM differencing requires comparison of signal-to-noise. Signal-to-noise, in Fig. 5, clearly demonstrates that the deep alpine snowpacks have a large signal relative to noise and provide useable information on snow depth both at maximum accumulation and during most of the snowmelt period (SNR >7). In contrast, the shallow snowpack at the prairie site, despite a similar absolute error to the alpine site, demonstrates decreased ability to retrieve meaningful snow depth information over the course of snowmelt; the signal became smaller than the noise. Applying the Rose criterion of a SNR ≥4, it is apparent that only the first flight at the short stubble and the first two flights at the tall stubble provided useful information on the snow depth signal. This is relevant when applying this technique to other areas with shallow, wind redistributed seasonal snowcovers such as those that cover prairie, steppe and tundra in North and South America, Europe and Asia. This is in contrast to other studies which do not limit where this technique can be reasonably applied (Bühler et al., 2016; Nolan et al., 2015).

3.3 Challenges

3.3.1 UAV Deployment Challenges

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

As the system relies on a single camera traversing the areas of interest, anything that may cause a change in the reflectance properties of the surface will complicate post-processing and influence the overall accuracy. Consistent lightning is important with a preference for clear skies and high solar angles to minimize changes in shadows. Diffuse lighting during cloudy conditions results in little contrast over the snow surface and large gaps in the point cloud over snow, especially when the snow cover was homogeneous. Three flights under these conditions could not be used and were not included in the previously shown statistics. Clear conditions and patchy snowcover led to large numbers of overexposed pixels (see Sect 3.3.2). Low sun angles should be avoided as orthomosaics from these times are difficult to classify due to the large and dynamic surface shadows present and the relatively limited reflectance range.

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

3.3.2 Challenges applying Structure from Motion over snow

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

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

3.4 Applications of UAVs and Structure from Motion over snow

The distributed snow depth maps generated from UAV imagery are of great utility for understanding snow processes at previously unrealized resolutions, spatial coverages and frequencies. Figure 6 provides examples of UAV derived distributed snow depth maps. The identification of snow dune structures, which correspond to in-field observations, is a qualitative validation that UAV derived DSM differencing does indeed provide reasonable information on the spatial variability of snow depth. Actual applications will depend upon the surface, snow depth and other deployment considerations as discussed.

Applications at the alpine site also include the ability to estimate the spatial distribution of snow depth change due to ablation (Fig. 7). To obtain ablation rates, the spatial distribution of snow density is still needed but it may be estimated with a few point measurements or with parameterizations dependent upon snow depth (Jonas et al., 2009; Pomeroy and Gray, 1995). In Fig. 7 the mean difference in snow depth between the two flights was 0.9 m; this gives a SNR of ~11 which is more than sufficient to confidently assess the spatial variability of melt.

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

The use of SfM in shallow snow environments, such as on the Canadian Prairies, is therefore limited to measuring near-maximum snow depths. Besides providing an estimate of the total snow volume, this information can also inform snow cover depletion curve estimation and description (Pomeroy et al., 1998). Simple snow cover depletion models can be parameterized with estimates of snow depth mean and coefficient of variation (Essery and Pomeroy, 2004), which otherwise need to be obtained from snow surveying. For 2015, coefficients of variation from the peak snow depth maps were 0.255 and 0.173, at the short and tall stubble sites respectively, which are similar to previous observations from corresponding landforms/surfaces (Pomeroy et al., 1998).

In addition to parameterising snow cover depletion models, UAV data could also be used to test the performance of these same models as Structure from Motion processing of UAV images produces orthomosaics in addition to DSMs. Sequences of orthomosaics are especially useful to quantify the spatio-temporal dynamics of snow covered area (SCA) depletion processes. Orthomosaics are complementary products to DSMs and their quality is subject to the same deployment conditions as DSMs. Orthomosaics have the same horizontal accuracy and resolution as the DSMs, but without a vertical component; any DSM vertical errors are irrelevant. Interpretation of SCA from orthomosaics is therefore possible regardless of surface characteristics or snow depth. The classification of orthomosaics to quantify surface properties will introduce error, and can be challenging in changing light conditions, which changes the spectral response of snow or non-snow covered areas across the surface. Typical supervised and unsupervised pixel based classification procedures can be readily applied. Since UAV imagery is at a much higher resolution than satellite or airborne imagery, classification differences in spectral response due to varying light conditions can be compensated for by using object oriented classification which also takes into account shape, size, texture, pattern and context (Harayama and Jaquet, 2004).

An example of a snow-covered depletion curve for the prairie site is presented in Fig. 8. A simple unsupervised classification of the orthomosaic into snow and non-snow classes quantifies the earlier exposure of the tall wheat stubble relative to the short wheat stubble. The tall stubble surface is an illustrative example of the advantages UAVs offer for SCA quantification. Tall stubble is a challenging surface on which to quantify SCA as snow is prevalent below the exposed stubble surface rendering other remote sensing approaches inappropriate. From an oblique perspective, the exposed stubble obscures the underlying snow and prevents the classification of SCA from georectification of terrestrial photography (Fig. 9). Due to the surface heterogeneity on small scales (stubble, soil and snow all regularly occurring within 30 cm) satellite, and most aerial, imagery struggles with clearly identifying SCA. To identify features accurately, in this case exposed stubble versus snow, multiple pixels are needed per feature (Horning and DuBroff, 2004). The 3.5 cm resolution of the orthomosaic corresponds to approximately three pixels to span the 10 cm stubble row which is sufficient for accurate SCA mapping over a tall stubble surface. The advantages of high-resolution UAV orthomosaics are obviously not limited to SCA mapping of snow between wheat stubble and can be readily applied to other challenging heterogeneous surfaces where SCA quantification was previously problematic. Snow cover data at this resolution can quantify the role of vegetation on melt processes at a micro-scale, which can in turn inform and validate snowmelt process understanding.

4. Conclusions

The accuracy of DSMs and orthomosaics, generated through application of SfM techniques to imagery captured by a small fixed-wing UAV, was evaluated in two different environments, mountain and prairie, to verify its ability to quantify snow depth and its spatial variability over the ablation period. The introduction of functional UAVs to the scientific community requires a critical assessment of what can reasonably be expected from these devices over seasonal snowcovers. Snow represents one of the more challenging surfaces for UAVs and SfM techniques to resolve due to the lack of contrast and high surface reflectance. Field campaigns assessed the accuracy of the Ebee RTK system over flat prairie and complex terrain alpine sites subject to wind redistribution and spatially variable ablation associated with varying surface vegetation and terrain characteristics. The mean accuracies of the DSMs were 8.1 cm for the short stubble surface, 11.5 cm for the tall surface and 8.7 cm for the alpine site. These DSM errors translate into mean snow depth errors of 8.8 cm, 13.7 cm and 8.5 cm over the short, tall and alpine sites respectively. Ground control points were needed to achieve this level of accuracy. The SfM technique provided meaningful information on maximum snow depth at all sites, and snow depth depletion could also be quantified at the alpine site due to the deeper snowpack and consequent higher signal-to-noise ratio. These findings demonstrate that SfM can be applied to accurately estimate snow depth and its spatial variability only in areas with snow depth >30 cm. This restricts SfM applications with shallow, windblown snowcovers. Snow depth estimation accuracy varied with wind speed, surface characteristics and sunlight; the most consistent performance was found for wind speeds <10 ms^{-1}, surfaces with insignificant vegetation cover, clear skies and high sun angles. The ability to generate good results declined over especially homogenous snow surfaces and southerly slope aspects in mountain terrain. Clear sky conditions were favourable for high snow-covered fractions with limited snow surface brightness contrast. During snowmelt with reduced snow-covered fraction, clear sky conditions caused overexposure of snow pixels and erroneous points in the point clouds.

The challenges of applying SfM to imagery collected by a small UAV over snow complicate the generation of DSMs and orthomosaics relative to other surfaces with greater contrast and identifiable features. Regardless, the unprecedented spatial resolution of the DSMs and orthomosaics, low costs and "on-demand" deployment provide exciting opportunities to quantify previously unobservable small-scale variability in snow depth that will only improve the ability to quantify snow properties and processes.

Acknowledgements

The authors wish to acknowledge the reliable assistance of Spatial Geomatics, Ltd of Calgary, Alberta, who provided strong technical UAV support, and Dr. Cherie Westbrook, for access to a differential GPS unit that made this research possible. Funding was provided by NSERC Research Tools and Instruments and Discovery grants, the NSERC Changing Cold Regions Network, the NSERC Postgraduate Scholarships-Doctoral Program, the Global Institute for Water Security and the Canada Research Chairs programme. Logistical support from Fortress Mountain Ski Resort, the University of Calgary Biogeoscience Institute and field assistance from May Guan, Angus Duncan, Kevin Shook, Sebastian Krogh and Chris Marsh of the Centre for Hydrology and post-processing support from Chris Marsh are gratefully noted.

References

Armstrong, R. and Brun, E.: Snow and Climate: Physical Processes, Surface Energy Exchange and Modeling, Cambridge University Press, Cambridge, UK., 222 pp., 2008.

Bewley, D., Pomeroy, J. W. and Essery, R.: Solar Radiation Transfer Through a Subarctic Shrub Canopy, Arctic, Antarct. Alp. Res., 39(3), 365–374, 2007.

Boufama, B., Mohr, R. and Veillon, F.: Euclidean Constraints for Uncalibrated Reconstruction, in 4th International Conference on Computer Vision (ICCV ’93), 466–470, IEEE Computer Society, Berlin, Germany, 1993.

Bühler, Y., Marty, M., Egli, L., Veitinger, J., Jonas, T., Thee, P. and Ginzler, C.: Snow depth mapping in high-alpine catchments using digital photogrammetry, Cryosph., 9, 229–243, doi:10.5194/tc-9-229-2015, 2015.

Bühler, Y., Adams, M. S., Bösch, R., and Stoffel, A.: Mapping snow depth in alpine terrain with unmanned aerial systems (UASs): potential and limitations, Cryosph., 10, 1075-1088, doi:10.5194/tc-10-1075-2016, 2016.

Collier, O.: The Impact on Topographic Mapping of Developments in Land and Air Survey: 1900-1939, Cartogr. Geogr. Inf. Sci., 29, 155–174, 2002.

DeBeer, C. M. and Pomeroy, J. W.: Modelling snow melt and snowcover depletion in a small alpine cirque, Canadian Rocky Mountains, Hydrol. Process., 23(18), 2584– 2599, 2009.

DeBeer, C. M. and Pomeroy, J. W.: Simulation of the snowmelt runoff contributing area in a small alpine basin, Hydrol. Earth Syst. Sci., 14(7), 1205–1219, 2010.

Deems, J., Painter, T. and Finnegan, D.: Lidar measurement of snow depth: a review, J. Glaciol., 59(215), 467–479, doi:10.3189/2013JoG12J154, 2013.

Deems, J. S. and Painter, T. H.: LiDAR measurement of snow depth: accuracy and error sources, in Proceedings of the 2006 International Snow Science Workshop, 330-338, Telluride, Colorado, 2006.

De Michele, C., Avanzi, F., Passoni, D., Barzaghi, R., Pinto, L., Dosso, P., Ghezzi, A., Gianatti, R., and Della Vedova, G.: Microscale variability of snow depth using U.A.S. technology, Cryosph., 10, 511-522, doi:10.5194/tc-10-511-2016, 2016.

Dingman, S. L.: Snow and Snowmelt, in: Physical Hydrology, 2nd ed., Prentice Hall, Upper Saddle River, New Jersey, 166-219, 2002.

Dumanski, S., Pomeroy, J. W. and Westbrook, C. J.: Hydrological regime changes in a Canadian Prairie basin, Hydrol. Process., 2015.

Essery, R. and Pomeroy, J. W.: Implications of spatial distributions of snow mass and melt rate for snow-cover depletion: theoretical considerations, Ann. Glaciol., 38(1), 261–265, doi:10.3189/172756404781815275, 2004.

Gray, D. M. and Male, D. H.: Handbook of Snow: Principles, Processes, Management, and Use, Pergamon Press, Toronto, Canada., 776 pp., 1981.

Grünewald, T., Schirmer, M., Mott, R. and Lehning, M.: Spatial and temporal variability of snow depth and ablation rates in a small mountain catchment, Cryosph., 4(2), 215–225, doi:10.5194/tc-4-215-2010, 2010.

Harayama, A., Jaquet, J.: Multi-source object-oriented classification of landcover using very high resolution imagery and digital elevation model. UNEP, Switzerland, 2004.

Harder, P., Pomeroy, J. W. and Westbrook, C. J.: Hydrological resilience of a Canadian Rockies headwaters basin subject to changing climate, extreme weather, and forest management, Hydrol. Process., 20, doi:10.1002/hyp.10596, 2015.

Hopkinson, C., Pomeroy, J. W., Debeer, C., Ellis, C. and Anderson, A.: Relationships between snowpack depth and primary LiDAR point cloud derivatives in a mountainous environment, Remote Sens. Hydrol., 352, 1–5, 2012.

Horning, N. DuBroff, N.: Myths and misconceptions about remote sensing. Version 1.0. American Museum of Natural History, Center for Biodiversity and Conservation. 4/29/2016 [online] Available from: http://biodiversityinformatics.amnh.org, 2004.

Vander Jagt, B., A, L., L, W. and M, T. D. D.: Snow Depth Retrieval with UAS Using Photogrammetric Techniques, Geosciences, 5, 264–285, doi:10.3390/geosciences5030264, 2015.

Jonas, T., Marty, C. and Magnusson, J.: Estimating the snow water equivalent from snow depth measurements in the Swiss Alps, J. Hydrol., 378, 161–167, doi:10.1016/j.jhydrol.2009.09.021, 2009.

Jones, H., Pomeroy, J. W., Walker, D. and Hoham, R.: Snow Ecology: An Interdisciplinary Examination of Snow-Covered Ecosystems, Cambridge University Press, Cambridge, UK., 384 pp., 2001.

Kinar, N. and Pomeroy, J. W.: Measurement of the physical properties of the snowpack, Rev. Geophys., 53, doi:10.1002/2015RG000481, 2015.

López-Moreno, J. I., Fassnacht, S. R., Heath, J. T., Musselman, K. N., Revuelto, J., Latron, J., Morán-tejeda, E. and Jonas, T.: Small scale spatial variability of snow density and depth over complex alpine terrain : Implications for estimating snow water equivalent, Adv. Water Resour., 55, 40–52, doi:10.1016/j.advwatres.2012.08.010, 2013.

MacDonald, M. K., Pomeroy, J. W. and Pietroniro, A.: On the importance of sublimation to an alpine snow mass balance in the Canadian Rocky Mountains, Hydrol. Earth Syst. Sci., 14(7), 1401–1415, 2010.

Marris, E.: Fly, and bring me data, Nature, 498, 156–158, 2013.

Mattmann, C. A., Painter, T., Ramirez, P. M., Goodale, C., Hart, A. F., Zimdars, P., Boustani, M., Khudikyan, S., Verma, R., Caprez, F. S., Deems, J., Trangsrud, A., Boardman, J. and Ave, A.: 24 hour near real time processing and computation for the JPL Airborne Snow Observatory, in Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International, 5222–5225., 2014.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundezewicz, Z. W., Lettenmaier, D. P. and Stouffer, R. J.: Stationarity Is Dead: Whither Water Management?, Science, 319, 573–574, 2008.

Mote, P. W., Hamlet, A. F., Clark, M. P. and Lettenmaier, D. P.: Declining mountain snowpack in western North America, Bull. Am. Meteorol. Soc., 86(1), 39–49, doi:10.1175/BAMS-86-1-39, 2005.

Nolan, M., Larsen, C. and Sturm, M.: Mapping snow depth from manned aircraft on landscape scales at centimeter resolution using structure-from-motion photogrammetry, Cryosph., 9, 1445–1463, doi:10.5194/tc-9-1445-2015, 2015.

Pomeroy, J. W. and Gray, D. M.: Snow accumulation, relocation and management, National H., Environment Canada, Saskatoon, SK., 144 pp., 1995.

Pomeroy, J. W., Gray, D. M. and Landine, P. G.: The Prairie Blowing Snow Model: characteristics, validation, operation, J. Hydrol., 144(1-4), 165–192, 1993.

Pomeroy, J. W., Gray, D. M., Shook, K., Toth, B., Essery, R. L. H., Pietroniro, A. and Hedstrom, N.: An evaluation of snow accumulation and ablation processes for land surface modelling, Hydrol. Process., 12, 2339–2367, 1998.

Rango, A. and Laliberte, A.: Impact of flight regulations on effective use of unmanned aircraft systems for natural resources applications, J. Appl. Remote Sens., 4, doi:10.1117/1.3474649, 2010.

Rose, A.: The Visual Process, in: Vision: Human and Electronic, Plenum Press, New York, 1-28, 1973.

Roze, A., Zufferey, J.-C., Beyeler, A. and Mcclellan, A.: eBee RTK Accuracy Assessment, Lausanne, Switzerland, 7 pp., 2014.

SenseFly: Menu Process Options Point Cloud Densification, 1/10/2015 [online] Available from: https://sensefly.zendesk.com/hc/en-us/articles/204542828-Menu-Process-Options-Point-Cloud-Densification, 2015.

Shook, K. and Gray, D. M.: Small-Scale Spatial Structure of Shallow Snowcovers, Hydrol. Process., 10, 1283–1292, 1996.

Spetsakis, M. and Aloimonost, J.: A Multi-frame Approach to Visual Motion Perception, Int. J. Comput. Vis., 6, 245–255, 1991.

Stewart, I. T., Cayan, D. R. and Dettinger, M. D.: Changes in Snowmelt Runoff Timing in Western North America under a `Business as Usual’ Climate Change Scenario, Clim. Change, 62(1-3), 217–232, doi:10.1023/B:CLIM.0000013702.22656.e8, 2004.

Sturm, M.: White water: Fifty years of snow research in WRR and the outlook for the future, Water Resour. Res., 51, 4948–4965, doi:10.1002/2015WR017242, 2015.

Szeliski, R. and Kang, S.: Recovering 3D Shape and Motion from Image Streams Using Nonlinear Least Squares, J. Vis. Commun. Image Represent., 5, 10–28, 1994.

Westoby, M., Brasington, J., Glasser, N., Hambrey, M. and Reynolds, J.: “Structure-from-Motion” photogrammetry: A low-cost, effective tool for geoscience applications, Geomorphology, 179, 300–314, doi:10.1016/j.geomorph.2012.08.021, 2012.

Table 1: Flight plan specifications

Variable	Prairie Site	Alpine Site
Flight altitude	90 m	90 m
Lateral overlap	70 %	85 %
Longitudinal overlap	70 %	75 %
Ground resolution	3 cm pixel^{-1}	3 cm pixel^{-1}
Number of flights (over snow/over non-snow)	22/3	18/4
Approximate area surveyed per flight	1 km^2	0.32 km^2

Table 1: Absolute surface accuracy summary[a]

Area	Variable	Mean (cm)	Maximum (cm)	Minimum (cm)	Total Points [c]
Apine-bare	RMSE	8.7	15	4	1120
Alpine-bare	Bias [b]	5.6	11	1	1120
Alpine-bare	SD	6.2	12	3	1120
Alpine-snow	RMSE	7.5	14	3	101
Alpine-snow	Bias [b]	4.4	13	1	101
Alpine-snow	SD	5.4	13	3	101
Short	RMSE	8.1	12.5	4.4	357
Short	Bias [b]	4.4	11.2	0	357
Short	SD	6.3	9.5	3.2	357
Tall	RMSE	11.5	18.4	4.9	357
Tall	Bias [b]	6.6	17.5	0.3	357
Tall	SD	8.4	14.2	3.1	357

[a] summary excludes five flights identified to be problematic
[b] mean of absolute bias values
[c] cumulative points used to assess accuracy over all assessed flights

Table 2: Absolute snow depth accuracy summary [a]

Area	Variable	Mean (cm)	Maximum (cm)	Minimum (cm)	Total Points [c]
Alpine	RMSE	8.5	14.0	3	83
Alpine	Bias [b]	4.1	11.0	0	83
Alpine	SD	7.1	12.0	3	83
Short	RMSE	8.8	15.8	0	323
Short	Bias [b]	5.4	15.2	0	323
Short	SD	6.1	10.3	0	323
Tall	RMSE	13.7	27.2	0	323
Tall	Bias [b]	9.8	26.4	0	323
Tall	SD	8.3	13.9	0	323

[a] summary excludes two flights identified to be problematic
[b] mean of absolute bias values
[c] cumulative points used to assess accuracy over all assessed flights

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

Flight [a]	Snow covered area (%)	Percentage of snow covered area excluded (%)
5-19_N	45.9	0.0
5-20_S	32.6	2.0
5-22_N	39.8	0.0
6-01_N	24.0	0.0
6-08_N	12.5	3.2
6-18_N	5.3	19.3
6-24_N	3.1	21.9
6-24_S	3.7	18.9

[a]month-day_portion of study area

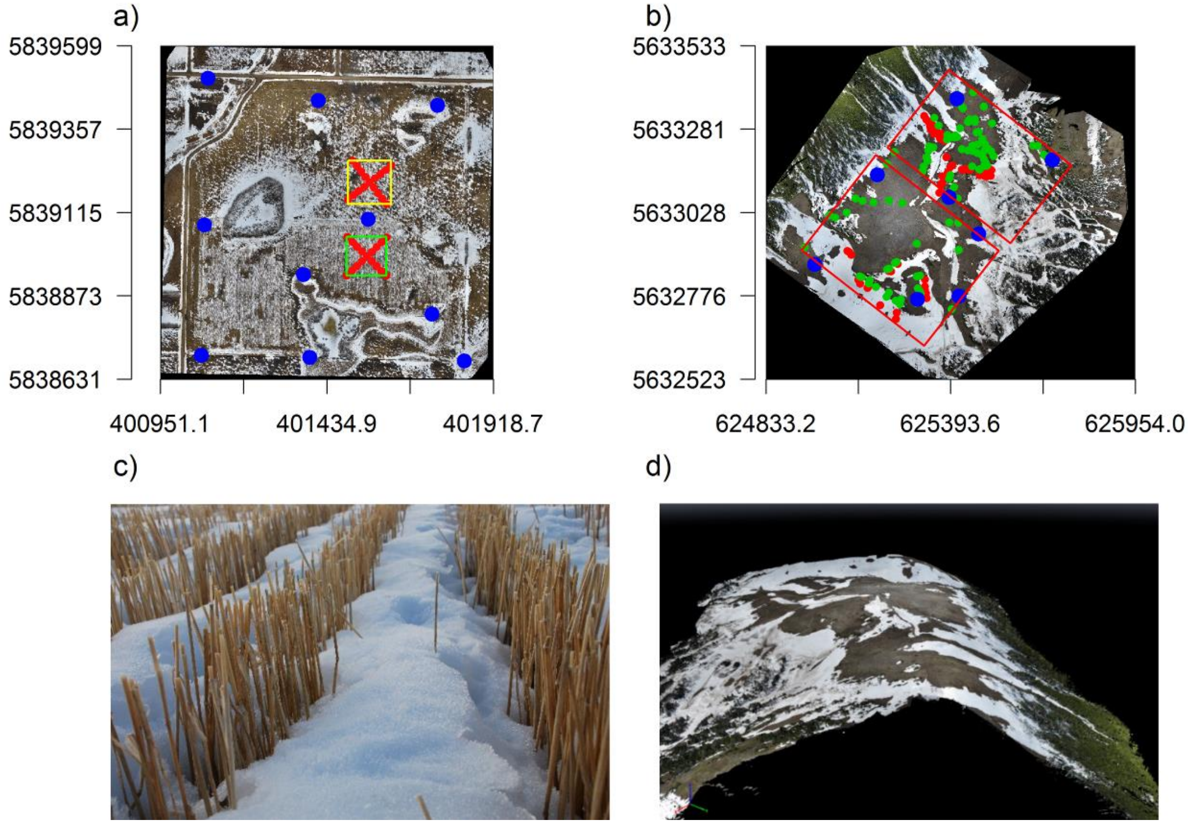

Figure 1: Orthomosaics of a) the prairie site located near Rosthern, Saskatchewan and b) the alpine site at Fortress Mountain Snow Laboratory, Kananaskis, Alberta. The prairie site image (March 19, 2015) has polygons depicting areas used for peak snow depth estimation over short (yellow) and tall (green) stubble. The alpine site image (May 22, 2015) was split into two separately processed subareas (red polygons). Red points in a) and b) are locations of manual snow depth measurements while green points at the alpine site b) were used to test the accuracy of the DSM over the bare surface. Ground control point (GCP) locations are identified as blue points. Axes are UTM coordinates for the prairie site (UTM zone 13N) and alpine site (UTM zone 11N). The defining feature of the prairie site was the c) wheat stubble (tall) exposed above the snow surface and at the alpine site was the d) complex terrain as depicted by the generated point cloud (view from NE to SW).

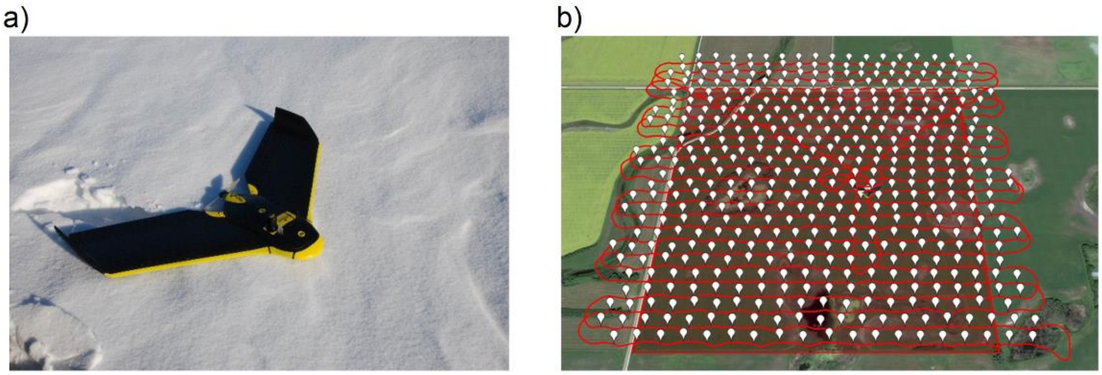

Figure 2: a) Sensefly Ebee RTK, b) a typical flight over the prairie site where red lines represent the flight path of UAV and the white placemarks represent photo locations.

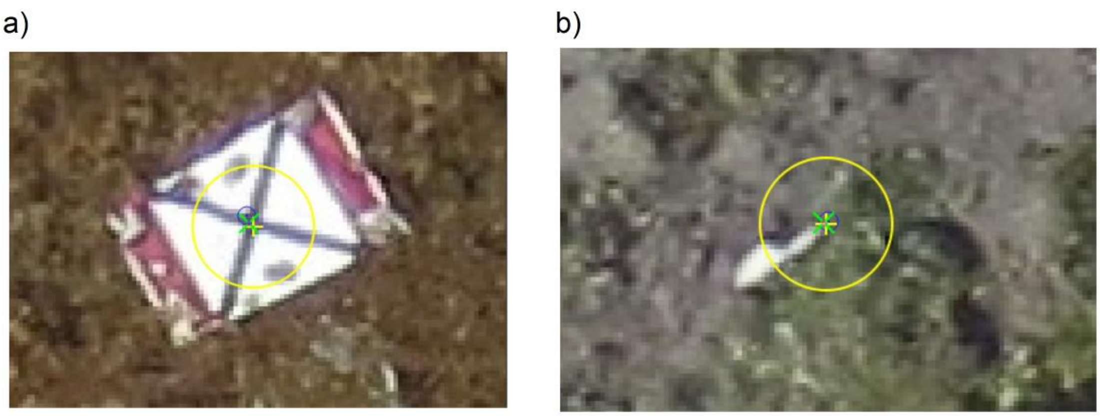

Figure 3: Examples of ground control points that included a) tarps (2.2 m x 1.3 m) and b) identifiable rocks at the same magnification as the tarp.

Figure 4: Root mean square error (RMSE, top row), Bias (middle row) and standard deviation (SD, bottom row) of DSMs with respect to surface over alpine-bare, alpine-snow, and short and tall stubble at prairie site, respectively. Blue bars highlight problematic flights and are excluded from summarization in Table 2. X-axis labels represent month-date-flight number of the day (to separate flights that occurred on the same day). Alpine-bare accuracies are separated into north or south areas, reflected with a _N or _S suffix. The last number in the alpine-snow x-axis label is the number of observations used to assess accuracy as the number of surface observations varied between 3 and 20.

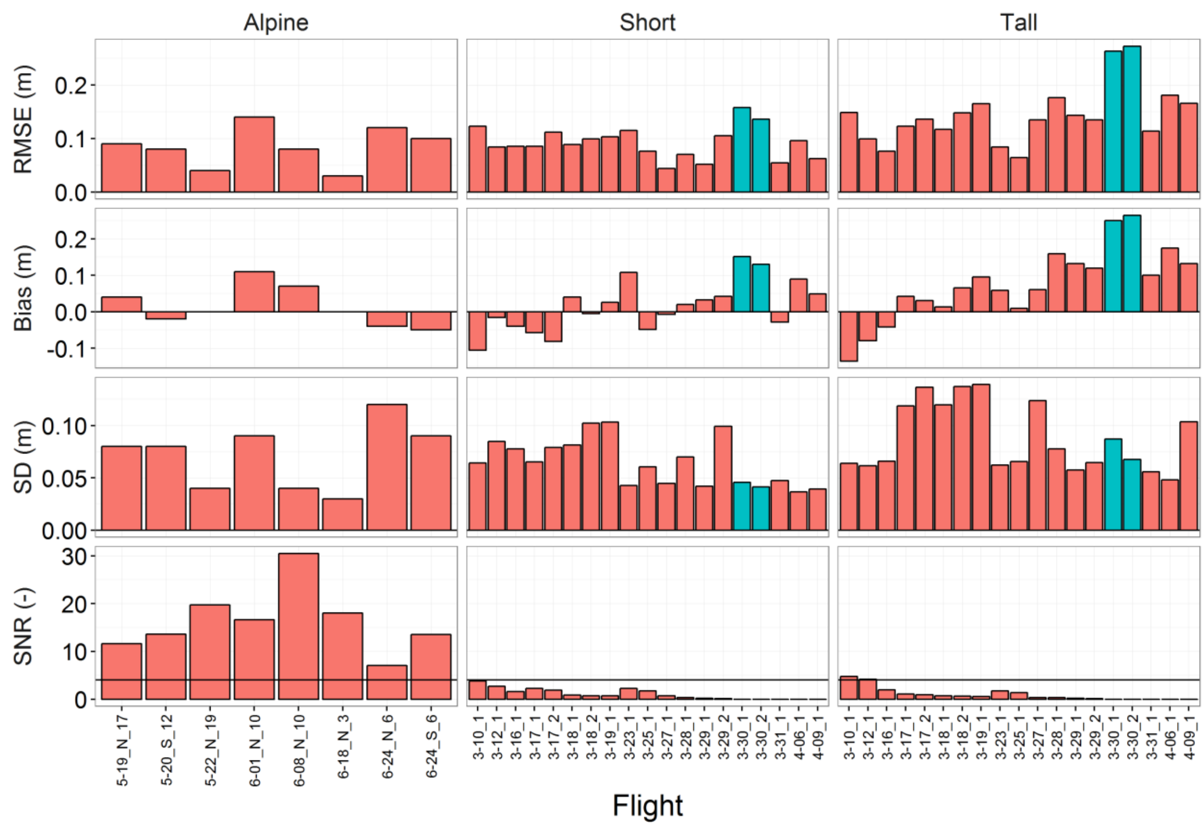

Figure 5: Estimated UAV snow depth error with respect to observed snow depth for the alpine site and the short and tall stubble treatments at prairie site. Blue bars highlight problematic flights and are excluded from summarization in Table 3. X-axis labels represent month-date. The last number in prairie labels is the flight of the day (to separate flights that occurred on the same day). Alpine labels separate the north or south flight areas suffixed as _N or _S respectively, and the last value is the number of observations used to assess accuracy as they vary between 3 and 19. Horizontal line in the SNR plots is the Rose criterion (SNR ≥4) that is used to identify flights with a meaningful snow depth signal.

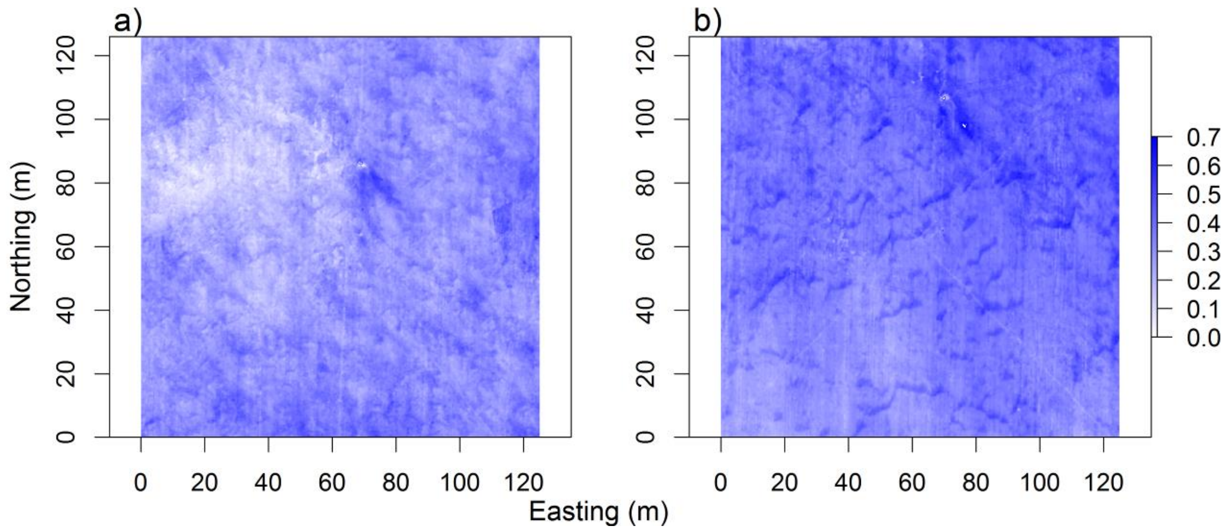

Figure 6: Bias corrected distributed snow depth (m) for a) short and b) tall stubble treatments at peak snow depth (March 10, 2015) at the prairie site.

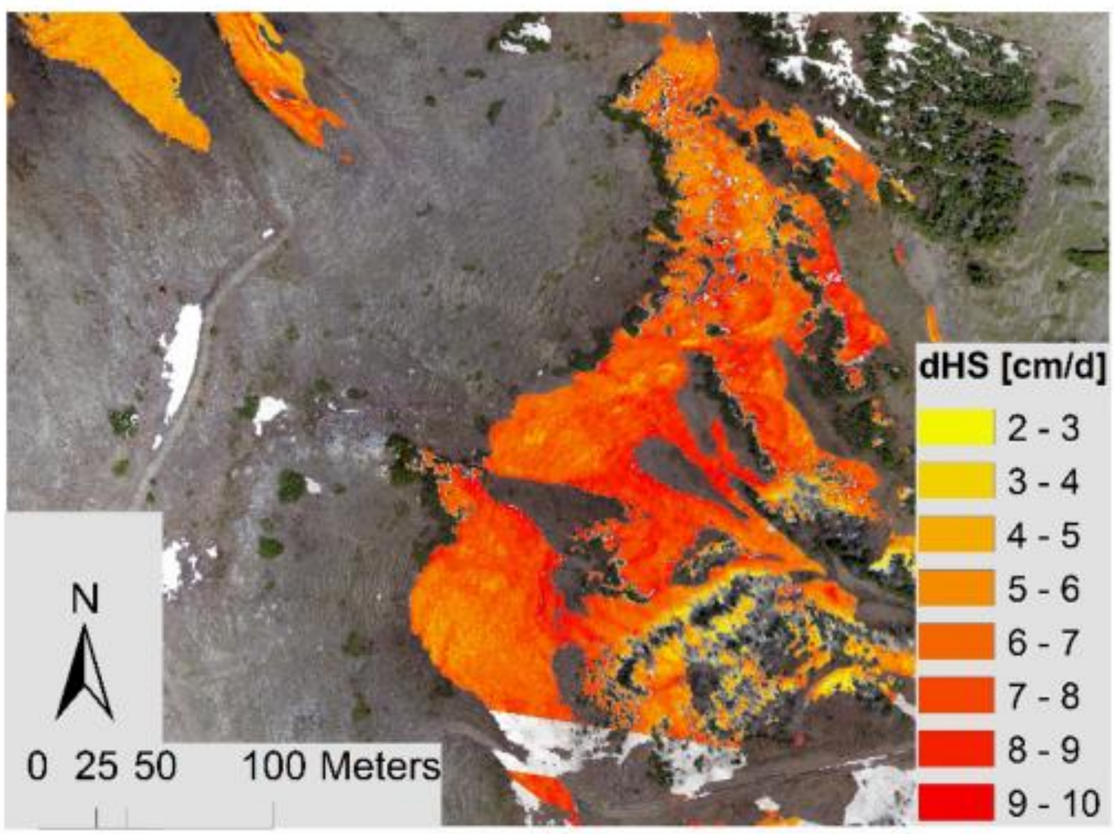

Figure 7: Rate of snow depth change (dHS day^{-1}) between May 19 and June 1, 2015 in the northern portion of the alpine site.

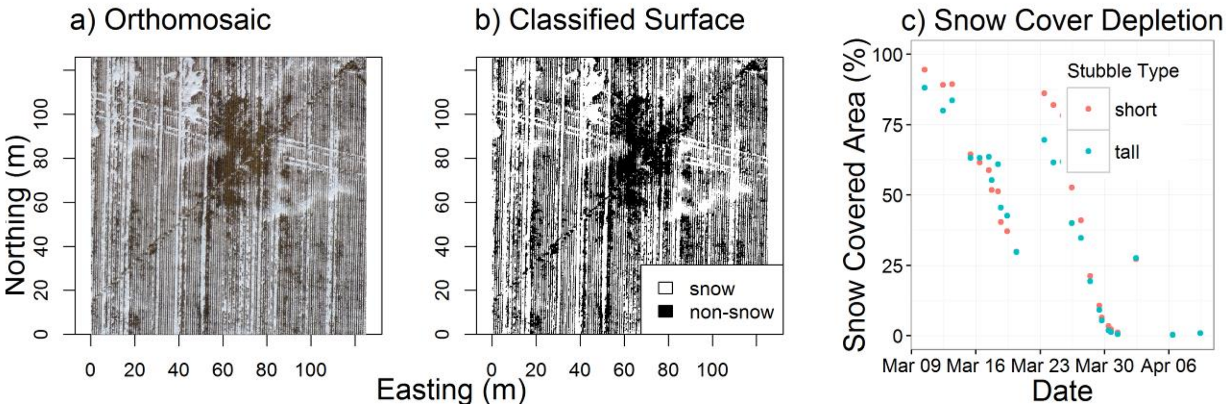

Figure 8: Estimation of snow covered area requires an a) orthomosaic which is then b) classified into snow and non-snow covered area. This produces a c) snow cover depletion curve when a sequence of orthomosaics are available. The short and tall stubble surface snow covered areas at the prairie site are contrasted, with a snowfall event evident on March 23, 2015.

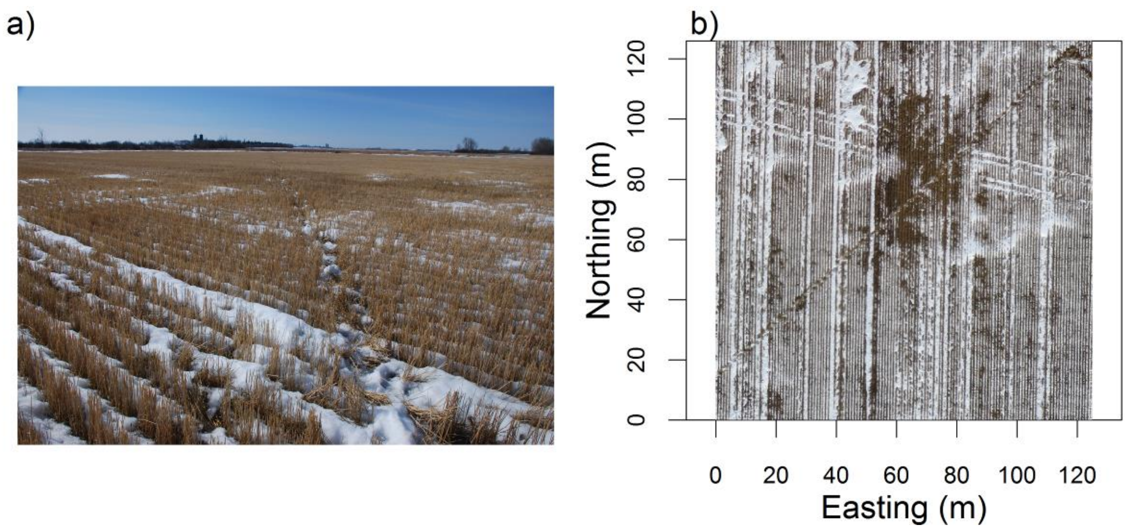

Figure 9: a) An oblique photograph demonstrates the issue of tall stubble obscuring underlying snowcover when considered in contrast to b) a UAV orthomosaic of the same area on the same date that clearly shows widespread snowcover.