Mapping Snow Depth from Manned Aircraft on Landscape-scales at Centimeter-Resolution using Structure-from-Motion Photogrammetry

Matt Nolan[1], Chris Larsen[2], and Matthew Sturm[2]

1. Institute of Northern Engineering, University of Alaska Fairbanks, 306 Tanana Loop, Fairbanks, AK 99775, matt2013@drmattnolan.org
2. Geophysical Institute, University of Alaska Fairbanks, 903 Koyukuk Drive, Fairbanks, AK 99775

Abstract

Airborne photogrammetry is undergoing a renaissance: lower-cost equipment, more powerful software, and simplified methods have significantly lowered the barriers-to-entry and now allow repeat-mapping of cryospheric dynamics at spatial resolutions and temporal frequencies that were previously too expensive to consider. Here we apply these advancements to the measurement of snow depth from manned aircraft . Our main airborne hardware consists of a consumer-grade digital camera directly-coupled to a dual-frequency GPS -- no Intertial Motion Unit (IMU) or on-board computer is required, such that system hardware and software costs less than $30,000, exclusive of aircraft,. The photogrammetric processing is done using a commercially-available implementation of the Structure from Motion (SfM) algorithm. The system is simple enough that it can be operated by the pilot without additional assistance and the technique creates directly-georeferenced maps without ground control, further reducing overall costs. To map snow depth, we made digital elevation models (DEMs) during snow-free and snow-covered conditions, then subtracted these to create difference DEMs (dDEMs). We assessed the accuracy (real-world geolocation) and precision (repeatability) of our DEMs through comparisons to ground control points and to time-series of our own DEMs. We validated these assessments through comparisons to DEMs made by airborne lidar and by a similar photogrammetric system. We empirically determined that our DEMs have a geolocation accuracy of ± 30 cm and a repeatability of ± 8 cm (both 95% confidence). We then validated our dDEMs against more than 6000 hand-probed snow depth measurements at 3 separate test areas in Alaska covering a wide-variety of terrain and snow types. These areas ranged from 5 to 40 km^2 and had ground sample distances of 6 to 20 cm. We found that depths produced from the dDEMs matched probe depths with a 10 cm standard deviation, and were statistically identical at 95% confidence. Due to the precision of this technique, other real changes on the ground such as frost heave, vegetative compaction by snow, and even footprints become sources of error in the measurement of thin snow packs (<20 cm). The ability to directly measure such small changes over entire landscapes eliminates the need to extrapolate limited field measurements. The fact that this mapping can be done at substantially lower costs than current methods may transform the way we approach studying change in the cryosphere.

1. Introduction

There are many reasons why being able to map snow depth over a landscape is desirable. In the Northern Hemisphere alone over 40 million km^2, almost half the land surface, becomes covered by snow each winter, making seasonal snow the largest annual topographic change on the planet (Déry and Brown, 2007; Lemke et al., 2007; Robinson et al., 1993). Billions of people rely on snow in some capacity, whether for drinking water, crop irrigation, or electricity (Barnett et al., 2005). Snow can also be a hazard, producing avalanches or floods (Castebrunet et al., 2014; Jamieson and Stethem, 2002). Snow plays a key role in the surface energy balance of the planet, thermally insulating the soil while efficiently reflecting sunlight because of its high albedo (Goodrich, 1982; Warren, 1982). The depth of the snow affects how much work grazing animals such as caribou will need to do in order to feed and it controls the quality of the habitat for sub-nivean animals like voles and weasels (Pauli et al., 2013; Pruitt, 1959; Russell et al., 1993).

Despite its importance, our current abilities to measure snow depth are limited. The simplest and oldest technique is to probe or core the snow by hand, but this technique has severe limitations with respect to areal coverage, and can be risky in avalanche country (Conway and Abrahamson, 1984; McKay, 1968; Sturm, 2009; Sturm and Benson, 2004). Automated point measurements such as snow pillows and sonic rangers have also been employed successfully for many years, but like hand probe measurements, require modeling to move from discrete point data to the landscape-scale (Liston et al., 2007; Liston and Sturm, 2002; Serreze et al., 1999; Slater and Clark, 2006). Remote sensing of snow coverage using optical sensors is fairly routine, but remote sensing of snow depth or snow water equivalent based on the microwave emissivity or radar scattering properties of the snow requires complex and problematic inversions in order to infer the depth and has kilometer-scale resolution (Clifford, 2010; Rittger et al., 2013; Rott et al., 2008). Similarly, it is possible to measure the SWE using an airborne gamma detector, but again the accuracy and spatial resolution of the method is low (Offenbacher and Colbeck, 1991). A technique that has received considerable attention in recent years is to measure the elevation of the snow surface by airborne or ground-based lidar and subtract from this the snow-free surface elevation, with the difference interpreted as snow depth (Deems et al., 2013; Fassnacht and Deems, 2006; Hopkinson et al., 2004; Prokop, 2008). Operating on the similar principles of repeat or overlapping coverage, but pre-dating lidar studies by 30 years, photogrammetry has also been used to produce snow depth maps (Cline, 1994; König and Sturm, 1998; Lee et al., 2008; McKay, 1968; Najibi and Arabsheibani, 2013; Otake, 1980; Rawls et al., 1980; Yan and Cheng, 2008), including using stereo-imagery from opto-electronic linescanners incorporating near-IR wavelengths in addition to RGB (Bühler et al., 2014; Buhler et al., 2015).

Airborne and terrestrial photogrammetry for determining snow depth were seriously investigated starting in the 1960s, though little published information is available (McKay, 1968). At that time, lacking any other method of mapping snow depth at the landscape scale, it was an obvious technique to consider as it was already being used for the study of glaciers (Brandenberger, 1959; Hamilton, 1965; Hitchcock and Miller, 1960; Post, 1995, 1969). However several issues hampered applying classical photogrammetry to snow cover. The low dynamic range of film combined with the difficulties of changing exposures mid-flight often produced over-exposed images of the snowfields, making it impossible for the photogrammetrist to determine elevation. Even when the snow images had suitable contrast, it took an extraordinary amount of time and skill to produce a map of sufficient vertical accuracy to measure snow depth (McCurdy et al., 1944), as the errors incurred produced uncertainty beyond the thickness of typical snowpacks.

These maps required identifying control points on the ground and establishing their elevation and position, and the process of subtracting one elevation field from another using paper or mylar maps was challenging. The overall complication and expense of this method in the pre-digital era was enough to cause the technique to largely be abandoned in the study of seasonal snow, though it has continued to be used for glacier volume change detection and for other large-scale deformation processes such as landslides (Bauder et al., 2007; Bitelli et al., 2004; Cox and March, 2003; Krimmel, 1989; Miller et al., 2009).

As we report here, recent advances in digital photogrammetric technology have now made it possible to not only produce accurate snow depth maps through airborne photogrammetry, but to do so at larger spatial-scales, at lower cost, and without loss of accuracy compared to most other techniques. These advances include improvements in consumer camera sensors, GPS processing techniques, desktop computational power, and especially, photogrammetric software. This software largely eliminates the need for purpose-built photogrammetric cameras and inertial motion units (IMUs), saving hundreds of thousands of dollars. These techniques are gaining popularity across all of earth sciences, being primarily deployed on low-cost unmanned aerial vehicles (UAVs). These systems are being used to map glaciers, river beds, coastlines, archeological sites, forest canopies, urban development, and more (d'Oleire-Oltmanns et al., 2012; Eisenbeiß, 2009; Fonstad et al., 2013; Gauthier et al., 2014; Hugenholtz et al., 2013; Irschara et al., 2010; Lucieer et al., 2013; Nex and Remondino, 2014; Rinaudo et al., 2012; Ryan et al., 2014; Vanderjagt et al., 2013; Westoby et al., 2012; Whitehead et al., 2013; Woodget et al., 2014). Our techniques were designed for manned aircraft, which can measure larger spatial scales with better accuracy and without the regulatory restrictions currently imposed on UAVs. Using an airborne equipment package costing less than $30,000 (excluding the aircraft), we demonstrate here that we can produce maps of snow depth accurate to ±10 cm with ground sampling distances (GSD) as low as 6 cm. We present results from 3 field sites in Alaska to show that the results produced using this technique (Figure 1) reveal details of snow depth distribution heretofore rarely available for study. The technique takes advantage of many of the technological developments of the past ten years, but in principle builds on the pioneering efforts of photogrammetrists and snow scientists beginning in the 1940s.

2. Recent Enhancements to Airborne Photogrammetric Methods

In this section we address the question "Why wasn't this method possible until now?" Our approach relies on three components that have undergone much improvement in recent years. These are the photogrammetric software used to create the maps, the digital cameras used to take the aerial photographs, and the airborne GPS techniques that geolocate the maps within the real world. We were not involved with these developments, our chief contribution here has been to integrate these components into a simplified and low-cost system. Below we describe the improvements to these components, as well as our choices for specific hardware/software. Evaluating whether our choices were optimal, and how other components might improve or degrade the results is beyond the scope of this paper, but it is likely to be an active topic of future research.

<u>2.1. Photogrammetric Software.</u> We used Agisoft's Photoscan software for processing, which uses a Structure from Motion (SfM) algorithm at its core (Koenderink and Van Doorn, 1991;

Westoby et al., 2012); at least 7 other software packages are currently available utilizing this algorithm. Both SfM and traditional photogrammetric-processing software triangulate the positions of points on the ground that have been imaged multiple times in overlapping photographs to create a 'point cloud' – a collection of X,Y,Z values defining the measured surface. This point cloud can then be gridded into a digital elevation model (DEM) or an orthometrically-corrected image mosaic (Maune, 2001); here we use the term *map* interchangeably with DEM. As part of this process, two types of unknowns must be determined before the maps can be made. Exterior orientations refer to the position and tilt of the photos and include 6 unknowns: X, Y, Z, yaw, pitch, and roll (that is, position and tilt of the camera). Interior orientations refer to the specifics of the camera and lens: focal length, sensor dimensions, pixel pitch of the sensor, lens distortions, and principle point. These result in about 10 unknowns, depending on the lens distortion model. Where the modern software has an advantage is that it requires no ground control points, no tilt information, and no a priori lens calibrations, as these can be calculated if the remaining variables are provided with adequate accuracy. Because tilts are not required as input, there is no need for an inertial measurement unit (IMU) on the aircraft. Because the software performs a camera/lens calibration on the fly, the need for a purpose-built aerial photography camera with strong camera-lens stability is also removed, allowing use of consumer-grade cameras. To create the point cloud, the software is able to access the full computational resources available, including the GPU of the graphics card.

2.2. Camera and Image Procesing. For this work we used a digital single lens reflex camera (DSLR), the Nikon D800E, which was the highest ranking DSLR (www.dxomark.com) when it was released. It costs about $3300 USD; in contrast, a modern, high-end photogrammetric-camera such as the Vexcel Ultracam might cost between $300,000 and $1,000,000. A primary attribute of photogrammetric cameras is their stable lens mount, but as we show, the SfM software adequately accounts for the less stable mounts on DSLRs. Photogrammetric cameras also have a greater number of pixels in the cross-track direction in comparison with a DSLR. For example, the D800E sensor has 7,360 x 4,912 pixels (36Mpix), compared to the Vexcel Ultracam with 11,704 x 7,920 (92 Mpix), resulting in flight lines that need to be about 60% closer for the same amount of overlap. In our applications the increased cost of extra flight time due to using a DSLR is more than offset by the reduced purchase price, high image quality, and ease of use of the consumer camera, all driven by relatively enormous consumer demand and competition. Similar advantages exist in consumer lens selection. The wide dynamic range and low noise of the D800E are largely responsible for our ability to capture texture in both bright snow and shadowed rock in the same image, problems that plagued film-based photogrammetry of snow in the past. Similar improvements in image processing now allow us to easily maximize local contrast (eg., sastrugi or suncups) while constraining global contrast to ensure the entire dynamic range is persevered. We used Adobe Camera Raw for this, though there are literally dozens of software packages with similar features. While the specifics for each data set varied, in general our approach consists of shooting in raw mode (with separate R, G, B channels), pushing the exposure as far as possible to the bright side of the histogram during acquisition where more bits are available for recording, then pulling the exposure down in post-processing (essentially turning the snow greyer) to enhance its visible contrast, while keeping the shadows from clipping. Despite these improvements in hardware and software, the quality of the photogrammetric results still depends on the skill of the photographer, especially in challenging lighting conditions, thus there is no simple prescription for camera settings or post-processing that can ensure success. However, as our results demonstrate it is possible to achieve accurate results, even in flat light.

2.3. GPS. While the GPS techniques we used have been available for some time, advances in processing software and hardware integration have streamlined the user-experience substantially. When maps are directly georeferenced (that is, without using ground control), the accuracy of the georeferencing is dependent on the accuracy of photo positions. To achieve our results, a modern multi-frequency GPS system must be used that can track aircraft position to within centimeters. We used a Trimble 5700 receiver, a discontinued model which measures only 12 GPS satellites at a time; modern receivers are capable of recording hundreds of channels from a variety of international constellations, which would likely improve position accuracy. The three dimensional offsets of the GPS antenna relative to the camera image plane, often referred to as "lever arms", must also be determined for each aircraft installation. In processing the GPS data, the lever arms are used in a coordinate transformation from the antenna position to the camera position. Without an IMU, this transformation relies upon the assumption that the aircraft frame of reference is aligned with the tangent of its trajectory. This assumption is often violated in the presence of crosswinds, but such errors associated with aircraft yaw can be mitigated by placing the GPS antenna directly above the camera. Finally, the exact time that the photo was taken must be used to determine its position within the post-processed GPS record. An aircraft traveling at 50 m s^{-1} (about 100 knots) will travel 5 cm in a millisecond. Thus to achieve a 5 cm accuracy in camera position requires a timing connection between camera and GPS with signal latencies reduced to below the millisecond level. There are a variety of ways this can be done. Our method converts the flash output from the camera into a TTL pulse for the event marker in the GPS; the camera and GPS receiver are thus directly coupled through this device without use of a computer.

3. Methods

3.1. Photo Acquisition and Processing. We pre-planned flight lines and shutter intervals to provide 60% sidelap and 80% endlap, such that most of the ground coverage within the map was imaged more than 9 times. Flight lines were uploaded into a Garmin aircraft-GPS for pilot display and navigation. The survey-GPS was set to record at 5 Hz. The Nikon D800E with Nikkor 24 mm lens was mounted vertically in the aircraft's camera port. The shooting interval rate (typically 2 to 5 s) was controlled by an intervalometer (contact www.fairbanksfodar.com for details), which also provided precise shutter-timing to the survey GPS as described in Section 2.3. Photos were acquired as raw NEF files, post-processed to maximize available contrast, and saved as JPGs for photogrammetric processing. A Cessna 170 flown by the first author was used to acquire the photos.

3.2. Airborne GPS Processing. GPS data were processed with GrafNav GNSS Post-Processing Software using their Differential GNSS method for projects near a CORS base station and using the PPP (Precise Point Positioning) method in remote areas (Gao and Shen, 2002; Snay and Soler, 2008). Positions were automatically interpolated within GrafNav from the 5 Hz GPS solution using the event markers created by the camera flash port to TTL pulse converter. Each photo position was exported and manually associated with image filenames to create an exterior orientation file that was imported into Photoscan Pro along with the photos themselves. The true accuracy of photo positions is difficult to assess, but most of the software's metrics (such as comparison of a forward and reverse solution) indicate that 95% of the points are within ±10 cm on most projects.

3.3. Photogrammetric Processing. We used Photoscan running on a dual Xeon eight-core computer with 192 GB RAM and a high end GPU for map construction. To make individual maps, a batch file was typically initiated within Photoscan to align the photos, optimize the bundle adjustment, construct the geometry, build a mesh, and export a DEM and orthophoto product. Total processing times ranged from 2-24 hours, depending on size of the project and processing resolution. As described in Section 2.1, processing time is dependent strongly on processing power, as well as having adequate RAM to prevent disk caching. Thus nearly any computer would work in this application, but processing times are dependent on computer resources.

3.4. DEM Differencing. To measure snow depth, we created a difference DEM (dDEM) by subtracting a snow-free DEM from a snow-covered DEM to determine the vertical change between them for each pixel (James et al., 2012; Maune, 2001; Nuth and Kääb, 2011; Wheaton et al., 2010). To optimize the differencing, the two maps were first co-registered horizontally to minimize errors in geolocation using simple 2D offsets determined with standard sub-pixel image correlation techniques using Matlab. Vertical alignment was done at snow-free locations in both maps (e.g., a wind-blown outcrop or a plowed runway). As described later, we found that we did not need to employ sophisticated techniques to determine misfits or non-affine co-registrations (Nuth and Kääb, 2011).

3.5. Snow probing. We tested the resulting snow depth maps by collecting about 6000 hand-probed depth measurements. We used several GPS-enabled depth probes to do this (Sturm and Holmgren, 1999). In most cases these depth data were collected along traverse lines that cut through obvious snow features (drifts, shallow areas, etc.), but in some cases we probed on a grid or on a spiral in a way that would allow the production of a snow depth map. Probe spacing varied depending on the length of the traverse line and the time available for the work, but was typically about 1 m. The GPS used on the probes is not a differential GPS and has a nominal accuracy of about 5 m. The probes have an inherent error due to penetration of the probe tip into the snow substrate of about ± 2 cm. In our remote field areas the substrate of tussocks and ice wedges usually had a surface roughness on a wavelength shorter than the probe spacing, which can introduce spatial aliasing when compared to airborne maps that have 6-20 cm resolution.

3.6. Validation DEMs. On the same day we acquired a photogrammetric DEM at the Minto Flats study area (3 April 14, described below), we also acquired a lidar DEM and a photogrammetric DEM from a system of slightly different design to validate our accuracy and precision assessments. This lidar and second photogrammetric system were operated simultaneously, carried in a Cessna 180 flown by the second author. This lidar system is based upon a Riegl Q240i and is the principal system used for NASA’s Operation IceBridge flights in Alaska. The system has been in extensive use since 2009 and is particularly well characterized with dozens of calibration flights and a careful program of boresight angle determination and monitoring (Johnson et al., 2013). At 95% confidence it has an accuracy of ±30 cm and precision of ±16 cm. The photogrammetric system differs from the one described above in that it used a 28 mm lens and routed its photo event markers through the IMU associated with the lidar system. With the GPS/IMU data, the software is able to directly calculate the full lever arm solution between the GPS antenna and camera. Thus image positions from this aircraft were derived from the fully coupled GPS/IMU processing, and there were other minor differences in processing workflow as well. This photogrammetric DEM was processed to a 12 cm ground sample distance (GSD).

3.7. Ground Control Points. We acquired ground control points for this project using the same Trimble 5700 receiver and Grafnav software used in airborne processing. Here we placed the antenna on a rod over photo-identifiable targets, as described later. We processed these measurements using the same Differential GNSS methods, which indicated a resulting accuracy of better than 3 cm in vertical and horizontal direction.

4. Study Areas and Measurements

We collected data from three study areas in Alaska: the Fairbanks International Airport, Minto Flats, and the Hulahula River watershed (location map in Supplemental Materials). As this was a technique-development project, these sites were chosen opportunistically to minimize our development costs, as described below.

The Fairbanks International Airport was selected due to its convenience and snow characteristics. It is located only a few miles from the University of Alaska Fairbanks and the plane we used for this work is located there. During the winter of 2013-14, about 43 cm of snow fell and remained undisturbed in the infields between runways. Near the runways and taxiways the snow gets extensively reworked to accommodate aircraft operations. The runways are kept clear of snow, which requires snow blowing, grading, and removal, all of which create berms adjacent to the runways of different thickness, and which change shape and depth frequently. Due to security and other issues, snow probing at the airport was limited to collection of a few hundred points and we do not statistically analyze these data. We made six airborne acquisitions over the airport (Table 1) mostly for assessments of accuracy and precision, using the snow-free runway as control. The maps made were roughly 5 km x 1 km and processed to 6 or 12 cm GSD. We used a GPS to measure 29 taxiway markings as photo-identifiable ground control points (GCPs); all GCPs used in this paper have an accuracy of about ±3 cm. The airborne imagery was acquired in a variety of lighting conditions, including low-angle mid-winter sun and beneath a thick overcast.

The Minto Flats site was selected because of its undisturbed snow cover and heterogeneous terrain. It is located about 50 km from Fairbanks and can be accessed using a ski-plane to land on its many frozen lakes. The area is characterized by tundra, swamps, areas of shrubs, spruce and birch forests, and taiga snow cover (Sturm et al., 1995). The airborne study area was about 2 km x 5 km and encompasses the full range of these terrain elements. Our snow-probe measurements were made at the edge of the largest lake in the area and cover about 9 hectares (about 1% of the area mapped by air). Using three separate GPS-enabled probes, 2,432 snow depth measurements were made on 2 April 2014, largely in a grid pattern with along-track separation of about 1 m and cross-track separation of about 6 m. Measured snow depths largely ranged from 0.1 – 0.6 m. We made six airborne maps of this area processed to about 15 cm GSD (Table 2); we also made two other maps on April 3rd using lidar and a 2nd photogrammetric system for validation, as described above. We also measured 21 GCPs on April 2nd using spray paint to create markers; these remained visible in the April 3rd orthoimagery as there was no intervening snow fall or melt.

The Hulahula River valley was selected for our snow research due to its history of hydrological studies, its relationship to the nearby, long-term McCall Glacier research project, its relevance to ecological research in the Arctic National Wildlife Refuge (Nolan et al., 2005; Nolan et al., 2011; Weller et al., 2007), and the availability of snow-probing conducted to support related snow research there (Sturm et al., in prep; Sturm et al., 1995). Located 330 miles north-east of Fairbanks, the valley extends from the continental divide of the Brooks Range to the Arctic Ocean, with a watershed of about 1800 km^2, about 6% of which is covered by glaciers (Nolan et al., 2011). Unlike most watersheds in the Alaskan Arctic, the snowmelt pulse is not the major hydrological event of the year due to the influence of glaciers and to a lesser extent aufeis. As the climate warms, however, these ice reservoirs are likely to disappear and allow snowmelt to dominate the run-off. A longer term project seeks to understand current rates and volumes of snowmelt, glacier melt, and aufeis melt through the photogrammetric techniques we describe here; these environmental questions will be addressed in subsequent papers. The probe data in the Hulahula River valley were collected in three terrain types on 18 March 2014: 1) a flat river terrace with a thin (15 - 20 cm), uniform snow cover, 2) a set of islands in the river with snow depths varying from 0.2 – 0.6 m, and 3) a series of drifted-in gullies cutting into a 40 m bluff with snow depth from 0 – 3 m. Airborne mapping was done on 20 April 14 (snow-covered) and 15 June 14 (mostly snow-free except in drifts). Though the snow-covered map was made 31 days after the probing, our results indicate that little change had occurred in snow depths over this period. The DEMs were processed to about 20 cm GSD and covered an area 14 km x 2.5 km. No GCPs were acquired.

5. Assessment and Validation of Map Accuracy and Precision

Our goal in this section is to answer two questions "How well do our airborne maps align with the real-world without using ground control?" and "After correcting for geolocation errors, how identical are our maps assuming no changes to the surface have occurred?" These questions address map ***accuracy*** and ***precision***, respectively. Because both the photogrammetric and GPS software we used to make our maps is proprietary and essentially black-box, we could not conduct a first-principle error analysis so we empirically assessed map errors, largely following Maune (2001). In all of our assessments we use the ± range to indicate the level of accuracy or precision at the 95% confidence interval for normal distributions (following Maune, 2001) and we simply cite the values of points ±47.5% about the mean for non-normal distributions; with 5 or less data points, we use +/- 50% of the full range.

We used two methods to assess accuracy. In the first, we assessed the difference between the maps and GCPs, calling the results *geolocation offsets*. The GCPs are accurate to about 3 cm, but the most we have for any one site is 29 and they are not well-distributed throughout the study area, making this a weak test spatially. In the second method, we applied these geolocation offsets to one of our maps, which we defined as a *reference map,* and then compared this map to the other maps (Maune, 2001); we term these map differences *co-registration offsets*. Using this method, the millions of pixels of the entire reference map become pseudo-GCPs, with their accuracy largely controlled by the precision of reference map itself (about ±8 cm, as we described below) rather than the GPS-GCPs (± 3 cm). We determined horizontal co-registration offsets using standard image correlation. We calculated vertical co-registration offsets at snow-free areas. The plowed runway in the airport data was the only location where we could do this statistically; at other sites we used the orthoimages to locate snow-free pixels for spot measurements only.

We report our precision as ±95% of the RMSE elevation difference between two DEMs after they have been optimally co-registered. Using this method, the magnitude of spatially correlated and uncorrelated errors are captured in the same precision metric. Given that our precision is on the centimeter-level and that we later show that this was sufficient to produce maps with excellent agreement to our snow probing data, we did not distinguish the amount of spatial-correlation within this ±95% RMSE further. Technically this RMSE measures the precision of a dDEM, not an individual DEM, but when computed from two maps where no changes in the surface have occurred and no gridding artifacts are present (both described later), the metric defines how identical the maps are and therefore the level of change-detection possible in the dDEMs.

Our overall assessment is that our maps (at 6 to 15 cm GSD) have accuracy better than ± 30 cm and precision better than ±8 cm, as described in sections 5.1-5.3. In this paper we do not address whether accuracy or precision vary with larger GSDs, but note that this remains to be explored. To validate these accuracy and precision assessments, in section 5.4 we compared one of our reference DEMs to two DEMs made on the same day using different systems and found that they confirmed our results.

5.1 Accuracy based on geolocation offsets from GCPs

We measured 29 GCPs at the airport. These were made at taxiway markings, all located within 300 m of each other. We compared these to the October snow-free acquisition and found a mean horizontal geolocation offset of 30 cm and a vertical offset of 13 cm (Table 1). Applying the offsets in Table 1, we define this October map as the reference map to determine co-registration offsets of the other maps made at the airport.

We measured 21 GCPs at the Minto Flats site. These targets were circles on the snow surface made with orange spray paint. They were too small for sub-pixel alignment within the orthomosaic, but they were suitable for determining that the horizontal geolocation offset was less than 15 cm (one pixel). The mean vertical offset was 23 cm (Table 2). This vertical offset was applied to our April 3rd photogrammetric DEM to create the reference map; no horizontal offset was applied given that a subpixel offset could not be reliably determined.

The results of these two GCP tests indicate a geolocation accuracy of ±30 cm.

5.2 Accuracy from co-registration offsets

We assessed the co-registration offsets of the other 5 maps from the airport time-series relative to October reference map. We calculated the horizontal offsets through image correlation of the snow-free runway markings, rounding to the nearest centimeter (Table 1, Columns 1-2). We calculated mean vertical offsets (Table 1, Column 3) using a block of pixels (roughly 20 m x 2000 m) surrounding the centerline of the runway, which was largely snow-free throughout the winter (Figure 2). The range of offset (highest minus lowest, last row Table 1) about the mean

(2nd to last row, Table 1) is a better indicator of accuracy than the mean itself, as the mean could be due to a systematic issue with the reference DEM. As discussed in more depth in Section 5.3, this “snow-free” area was not completely snow-free, so the range of vertical error has been impacted by real changes to the surface. Nonetheless, both the mean and the range indicate ±30 cm as a reasonable co-registration accuracy.

We repeated this same analysis for the Minto Flats time-series (Table 2). As shown in Table 2, the full range of horizontal co-registration offset is about ±0.05 m. Because there was no large snow-free surface like the runway, we determined vertical offsets by making spot measurements of the dDEMs in snow-free areas located using the orthoimage. These show a scatter of only ±0.07 m, with 5 of the 7 maps clustered within half that.

Overall the Minto Flats data showed better co-registration accuracy than the airport data, about ±15 cm compared to ±30 cm. The difference may relate to differences in relief of the terrain – the airport is nearly flat and thus perhaps making the solution geometry weaker due to fewer differences in scale. In any case, overall we conclude that our accuracy was ±30 cm, noting that is likely conservative. The underlying causes for why map geolocation accuracy is ±30 cm when photo position accuracy is ±10 cm remains unclear.

5.3 Precision

The primary challenge in determining map precision is that many real changes occur on the ground at the centimeter level that confound the precision assessment. For example, surface change at this level or higher can be caused by frost heave and thaw consolidation of the ground, or by compression of vegetation under the weight of snow (Esch, 1995; Ménard et al., 2014; Sturm et al., 2005; Taber, 1929). Thus the designs of our tests are largely about controlling for such confounding influences and we assessed the precision at the airport differently than we did at Minto Flats. At the airport, we used the same time-series of the snow-free runway sections that we used for accuracy assessments. At Minto Flats, we compared the November 6th and 8th maps as intervening changes were negligible.

5.3.1 Airport precision assessment

We tried to assess vertical precision in several ways using the runway time-series. Real changes in the surface elevation were present in these tests (but of unknown magnitude), yet the precision was still excellent.

First, we examined the data graphically as is shown in Figure 2A-C. This demonstrated that in the absence of confounding changes, our DEMs had a precision of about ± 3 cm. Figure 2A shows an example of a difference DEM, with Figure 2B showing the corresponding snow-covered scene for reference. Figure 2C shows transects from all 6 maps that extend across the snow-free runway. Over the crest of the centerline where plowing is best, we found that the elevations compared to within ± 3 cm (95% confidence).

Next we examined the scatter about the mean co-registration offsets described in Section 5.2. We did this over a block of the runway that was kept largely snow-free through winter. Column 4 of Table 1 indicates that once co-registered using the offsets in Table 1 (Columns 1-3), 95% of the vertical difference between the runway blocks were less than ±10 cm (about twice the standard deviation shown in Column 4). Visual inspection of the orthophotos (e.g., Figure 2B) shows that this block of pixels was not completely clear of snow and changed between maps. Further, our inspection of the difference maps indicates that spatially-correlated variations of 5-10 cm in elevation occur over segments separated by expansion joints across all of the tarmac, suggesting differential frost heave and settling. Despite these confounding influences (real changes in surface elevation), we still found only a range of ±10 cm, which is excellent.

Finally, we extracted elevation profiles down the centerline of that block where plowing is best to further eliminate the influence of snow (green line in Figure 2A). Figure 2D shows that each of these transects captured the same decimeter variations in runway topography, though each differs slightly. We measured the scatter of these centerline transects as function of distance along the runway. Here the maximum range between transect points was 21 cm, the mean range was 9 cm, and over 95% of the transect length these differences had a range less than 12 cm (±6 cm). Whether these differences are due to frost heave or spatially-coherent noise (perhaps caused by photo misalignments) is not known, but the fact that 95% of the variation is within ±6 cm is an outstanding result and, as we describe in Section 6, more than sufficient to measure snow depth variations at centimeter resolution.

To assess the horizontal precision, we used custom feature tracking software (Mark Fahnestock, pers. comm.., 2014) using a python version of the feature-tracking software Imcorr (Scambos et al., 1992). Such software is commonly used to measure velocity fields of glaciers from optical and radar satellite imagery (Berthier et al., 2005; Huang and Li, 2011). In our case, because we know that the position of runway markings and many other surface features are not moving, any relative motion between them detected by this software indicates a lack of horizontal precision within the maps. Using the two snow-free orthoimages (6 Oct 13 and 30 Sept 13) and search chips of 100 x 100 pixels (6 m x 6 m), we found that 95% of the RMSE pixel displacement about the mean was within ±6 cm (all subpixel). The mean value of displacement was also within a few centimeters of the co-registration offset we found through whole-image correlation (Table 1), as expected.

Thus our overall assessment of the airport time-series is that is that both vertical and horizontal map precision is ±6 cm or better when the confounding influence of real surface changes is removed.

5.3.2 Minto Flats precision assessment

Here we compare two DEMs of the Minto Flats area made two days apart with no intervening snow fall or snow melt (November 6^{th} and 8^{th}). Once co-registered we created the dDEM of the entire area at 15 cm GSD (~15 km^2, $n>6x10^8$) and found 95% of the vertical variation to be within ±44 cm. This distribution was non-gaussian, with tails extending to +/- 15 m. We cropped the dDEM to include only a large lake ($n>10^6$) and found the variation dropped to ±8 cm. These distributions are shown graphically in Figure 3A. The difference in scatter between the lake and entire area is largely caused by spatial aliasing of trees. Minto Flat trees are skinny black spruce and leaf-free birch, up to 20 m tall, typically separated from each other by a tree length or more like a forest of widely scattered flag poles. Even at 15 cm GSD, our DEMs are not able to resolve these spike-shape targets adequately and thus most trees are represented by several pixels that each average some fraction of tree height with surrounding ground height. The result is that trees appear as cones in the DEM, with cone height dependent on how the

DEM mesh happened to lie over that tree. Because these cones are so narrow, slight errors in horizontal co-registration or origin coordinates can cause dDEM errors approaching the heights of the trees; one of these maps was made when winds at ground level were over 15 m s-1, which could also cause similar aliasing at this resolution. Visual inspection of the dDEM confirms that within clearings between the trees that precision is the same as on the lakes. Thus any mapping system creating a DEM at this GSD would have these same spatial aliasing issues, and our precision is therefore represented better where gridding artifacts such as the spatial aliasing of trees are not present.

Based on our results at the airport and Minto Flats, we believe ± 8 cm is a reasonable value for the precision of our method. If any warps, tilts, or other spatially-correlated errors exist in our data, they are largely confined to within this level. Thus our DEMs should be repeatable to ±8 cm, exclusive of any spatial aliasing or other gridding artifacts.

5.4 Comparison to Validation DEMs

Here we seek to validate our accuracy and precision numbers by answering the question "How well do our DEMs compare to those made by other systems?" We do this by comparing our reference DEM for Minto Flats (April 3rd) to DEMs on the same day using lidar and a 2nd photogrammetric system (Section 3.5).

We co-registered the validation photogrammetry with our reference DEM using the same methods previously described and found a vertical co-registration offset of 21 cm, with variation of ±8 cm (95%) over the largest lake in the area. While we don't have any formal accuracy or precision specifications for the validation system, given its similarity to the system that created the reference DEM it seems reasonable that they should have similar specs.

Comparisons with the lidar DEM similarly validated our results. We created a 100 cm GSD DEM from the lidar point cloud, which had a point density of 2 points m^{-2} and a footprint of about 100 cm. We then resampled the reference DEM to this GSD. Because we have no orthoimage for the lidar, we created shaded relief images of the DEMs and then used these for sub-pixel image correlation to calculate horizontal offsets. Once co-registered, over the entire domain the vertical offset from our reference DEM was only 2 cm. Visual inspection of the dDEM showed no spatially-correlated errors, such as warps or tilts, greater than the lidar's precision level of 16 cm. Nearly all differences observed above that precision level were due to trees, likely caused by the different imaging physics between lidar and photogrammetry and by aliasing artifacts caused by the 100 cm GSD, as described in Section 5.3.2. Over the entire domain we found a variation of ±51 cm (95%), but over just the largest lake in the area the variation was only ±10 cm, with the latter being a better test in terms of validation; these distributions look nearly identical to those in Figure 3A. Statistically the lidar DEM is essentially identical to our reference DEM. We performed a Kolmogorov–Smirnov test and determined that statistically the two samples are from the same continuous distribution at the 95% confidence level. That is, our photogrammetric maps are essentially identical to the validation data. This is shown graphically in Figure 3B, which shows the similarity between the hypsometries of the lidar and the reference DEM.

6. Snow Depth Mapping Accuracy

Here we address the question "How well do our photogrammetric techniques measure snow depths?" To do this we compared our maps to over 6000 snow probe measurements. The mean of these differences is directly related to how well we can co-register the two DEMs used to produce the dDEM. This co-registration error, in turn, is related to finding snow-free areas that are not confounded by real changes to the surface such as vegetative compression, frost heave, aufeis melt, or erosion. Without suitable snow-free ground control points, the accuracy of our snow depth maps is limited to our geolocation accuracy, or about ± 30 cm. But when suitable ground control points can be found, this accuracy is effectively improved to the level of the precision of our maps, or about ±8 cm. Here we describe the accuracy our photogrammetric snow depth measurements by the standard deviation of the difference between probe and map values, as the mean is a function of ground control and co-registration, which have accuracies independent of system precision. As before, our assessment is confounded by real changes occurring on the ground, as we describe below. We conducted this map-probe analysis at three sites: the Fairbanks International Airport, Minto Flats near Fairbanks, and the Hulahula River valley, as described in Section 4.

6.1 Airport Snow Depth Analysis

Due to security and other issues we were only able to collect a few spot measurements of snow depth. We found the undisturbed snow depth to be about 43 cm, the packed and groomed ramp area snow depth to be 10-15 cm, and the plowed drifts to be greater than 1 m. Comparison of these values to Figure 2A shows close agreement, as described in the caption of Figure 2.

6.2 Minto Flats Snow Depth Analysis

Before statistically comparing our probe measurements to the dDEM (03 April 14 minus 28 Sept 13), we assessed whether the probe measurements were optimally co-registered to the maps using our footprints in the snow. These were clearly resolved in the DEM and orthophoto (Figure 4A-B). We each wore different footwear (ski, snowshoe, or boots), and the resolution of the map was such that we could differentiate these individual tracks based on their indentations (Figure 4C), which ranged from 6 cm to 10 cm deep and about 10 times as wide. The GPS units embedded into the probes each have an independent nominal accuracy of about 5 meters, thus the ground data has better vertical precision than the maps but a coarser horizontal precision. Analysis of all of the probe measurements together suggested there was no single shift that aligned them properly relative to the footprints, likely because each probe's GPS accuracy was independently varying. Short of manually shifting each of the 2432 measurements independently to the corresponding footprints, there was no simple spatial alignment possible. This meant that footprints' disturbance to the snow depth was included in the aerial mapping of snow depth, but not in the ground probe data. Nevertheless, even without exact co-registration the depth comparisons were 10-26 cm (on the order of footprints) and thus our results conservative, as we show next.

Figure 4D presents a comparison of about 500 probe measurements typical of the data set. The standard deviation of offset for those measurements was 10 cm. For the full 2432 measurements, including those made within the forests (with aliasing errors), the standard deviation was 26 cm, but careful visual examination of imagery reveals that nearly all of the offsets greater than 15 cm were located in areas where the vegetation was compressible, such as in the tall grasses near the edge of the lake or shrubs at the edge of the forest. The mapped summer surface in these areas is the top of the vegetative canopy. In winter, this canopy becomes compressed to the point where it can even produce 'negative' snow depths in the difference maps. Here we found such snow-vegetation dynamics were causing up to 30 cm of error. That is, the maps we produced here were no less precise than described in Section 5 (±8 cm), but the fundamental assumption that the differences between maps were caused only by snow accumulation has been violated where there is compressible vegetation.

6.3 Hulahula River Snow Depths

Similar to the other sites, we began this analysis by co-registering the DEMs. Using the same image correlation technique we used in Minto Flats, we found no horizontal offset. Using several snow free areas identified using the orthoimages, we determined there was a vertical offset of 55 cm. Subsequent analysis of the probe data indicated that 20 cm of that vertical offset needed to be removed to reduce the map-probe mean offset to zero over the snow-covered points that had the least likelihood of there being vegetation compression. Considering the surface amplitude of the tussock tundra here is about 15 cm, these shifts are small and within the noise of other confounding factors. Nevertheless, this process highlights that the primary errors in snow depth accuracy are co-registration in the absence of ground control points. Once the maps were co-registered, we created a dDEM and compared it to the probe values in the gullies, on the islands, and on a large river terrace (Figure 5).

Figure 6 highlights some results from the gullies. Here a series of ice wedges have thermally eroded to form a connected drainage system. In winter, this drainage network is completely drifted over by snow, as can be seen by comparison of Figures 6A-C with 6D, with snow depths of 100 to 200 cm. To the right of the gully a polygonal network can be seen in both the summer image and difference map with snow depths of only 10 to 20 cm. Figure 6D reveals a snow depth of near zero to the right of the gully and about 20 cm to the left of it. These values can be qualitatively confirmed by the winter image in Figure 6B, where exposed tussocks can be seen to the right but not to the left. Comparison of about 200 probe points in Figure 6E reveals that the maps match the probe depths and the features delineated by probing, including those parts of the gully that exceed the 120 cm range of the probes. The standard deviation of offset here was 20 cm, not including points where the probes did not reach the bottom. The bulk of this offset beyond 10 cm is likely attributed to 1) uncorrected probe positions resulting in misalignment between probes and maps, which matters more in steeper terrain where spatial depth heterogeneity is larger, 2) a spatial sample bias caused by the tussock terrain's surface roughness of 15 cm on spatial wavelengths below GSD and below probe spacing, and 3) real surface changes such as vegetative compressibility or frost heave. Considering these potential sources of error, the agreement makes clear that we are measuring snow depth at the centimeter to decimeter level.

The island transects (Figure 7) revealed a similarly strong correspondence between map and probe data as well as new sources of confounding error in interpreting the difference map as a change in snow depth. In winter, the river bed surrounding the island was completely snow covered and the transects extended over the edge of the island's summer boundaries (Figure 7A). In most of these edge locations, the map indicates changes up to a meter larger than revealed by the probe (Figure 7B). Interpretation of our difference maps in the active river bed is complicated by the fact that our photogrammetric technique does not work as accurately over water, for a variety of reasons outside the scope of this paper. Further, our stream gaging measurements (Nolan, unpub. data) show that the water height in spring can be over a meter higher than in fall here. Thus extra care in interpretation needs to be taken of differences over liquid water bodies. Given our map precision, it is therefore likely that remaining edge-offsets were caused by either the probe being stopped by river ice obscured by the snow or that the edges of the island were eroded, or both. On the island itself, numerous shrubs also influenced the correspondence, yet the agreement remains in the 10 – 20 cm range.

Map values along the terrace (orthogonal transects in Figure 5) showed even better correspondence with probe values than they did at gully and island sites. Here, the offset of *all* 1111 sample points spanning a transect of 1.6 km had a standard deviation of only 10 cm. This low variance could be explained by the relatively homogenous terrain of wide, shallow slopes characterized by a low shrub cover where sprigs and branches poked through the consistently 18 cm deep snow. However, despite the better standard deviation, the mean offset was 10 cm, as opposed to zero at the other sites. This mean offset could be eliminated using a different co-registration offset for the terrace points than used at the islands or gullies, but compression of the relatively uniform vegetative canopy, differential ablation or drifting of the prober's snow machine track over the intervening month, or the imprecise geolocation of the snow probe data could easily explain the offset as being real.

The offset between map and probe for all 3382 points measure at the Hulahula site had a standard deviation of 16 cm, without filtering for any of the sources of error noted above. We briefly explored the influence of different GSDs on results by using a 40 cm GSD compared to a 20 cm GSD; this did not appreciably change the standard deviation of offset, but it did change the individual pointwise comparisons. That is, comparing map data to map data (20 cm to 40 cm GSD) at the probe locations led to a 7 cm standard deviation, which is on the order of the precision we found in Section 4. Thus perhaps half of the 16 cm variation we found between map and probe may be attributable to real change on the ground. The similarity between map and probe data sets is further confirmed by a Kolmogorov–Smirnov test, which gives a value of 0.06; this is well below the critical D-value of 0.35, indicating that the two sample distributions are the same at the 95% confidence level. That is, to the best of our ability to determine, the photogrammetric maps are just as accurate as the probe data for characterizing snow depth, despite the many confounding influences besides depth that are incorporated into the maps.

7. Discussion

The photogrammetric method described here is sufficiently accurate to measure snow packs of nearly any thickness, and future software and hardware improvements are likely. The primary technological challenge for the future is improving geolocation accuracy, which relates to GPS data and how it is used within the photogrammetric bundle adjustment. Given the wealth of airborne-GPS research from lidar studies, it is likely that a map accuracy of 10 cm is currently a hard limit and one that will be difficult to overcome in the future. However, as we have demonstrated, geolocation (accuracy) is not as important as repeatability (precision). As long as stable, snow-free points within the mapped domain can be found such that the map differences there can be reduced to zero, a single affine translation appears to be enough to co-register an entire map and create excellent difference maps. A key lesson learned here is that it is not enough that these points are snow-free, but also that they be free of confounding real changes such as frost heave (as at the airport) or vegetative compression (as at Minto Flats). Similarly, the primary non-photogrammetric challenge for mapping of thin snow packs relates to the interpretation that changes in the difference map are being caused by snow depth. Because our technique can measure change at the centimeter to decimeter level, any real change at that level becomes noise when interpreting the results as purely changes in snow depth. These confounding changes in surface elevation are all site dependent and often a function of snow cover itself, such as the amount of vegetative compression or the rate of thermally-driven frost heave. However, given that our map-probe comparisons were still in the 10 – 20 cm range without accounting for these errors, it seems our technique is sufficient for many types of studies without further modification.

The issues of contrast and lighting that plagued the early pioneers of film photogrammetry to map snow depth can largely be overcome using modern technology applied with skill. With the advent of digital cameras and in-flight exposure evaluation, flat lighting conditions are still challenging but they do not prevent measurement. Such flat lighting conditions are typically caused by a thick overcast over fresh snow. Two types of map errors are produced by lack of contrast in deep shadows or flat lighting. In the worst of these cases, the spatial density of contrast features are reduced, resulting in the point cloud density also being reduced. In this case, either the resolution of the DEM must be reduced or a void of no data will result. This does occur, but rarely. Depending on camera settings (and camera) in such areas, the sensor noise itself can be misinterpreted by the photogrammetric software as real contrast features. Because the location of this sensor noise changes from image to image, topographic noise results. This noise is typically on the 1-2 m level, but in steep mountainous terrain can reach 10-20 m. We did not formally address such errors in this paper because none of the study areas used in this paper suffered from them due to suitable photographic technique. The most challenging contrast issues can also be avoided completely by waiting for better lighting. In any case, when these noise errors do occur they are easily identifiable in the DEM and confirmed by the orthoimage.

While there is currently a lot interest in using low-cost UAVs as platforms for SfM photogrammetry (also known as small Umanned Aerial Systems, or sUASs), our research requires manned aircraft for several reasons. Though it may be possible in the future to adapt our methods onto a UAV platform, we could not achieve the precision our needs required without use of multi-frequency GPS and high-quality optics, which both increase cost and payload outside the limits a low-cost sUAS. Our goal is also to measure snow depth of entire watersheds, covering hundreds to thousands of square kilometers, and this simply is not feasible with sUASs. Fundamentally, an sUAS is a field tool requiring the same logistics as ground-based measurements. For example, we flew our Hulahula missions as day trips from Fairbanks, over 500 kilometers away – to do similar work with an sUAS would require a multi-day field expedition with attendant logistical support and costs; even our work at Minto Flats, 30 miles from Fairbanks, would require overcoming similar challenges. Thus for use off the road system, an expeditionary field effort cannot be avoided without using a UAV that can truly replace a manned-aircraft, such as a Predator, Global Hawk, or Sierra (Fladeland et al., 2011; Schreiber et al., 2002; Whitlock, 2014). Such UAVs are considerably more expensive than the manned aircraft we used, are considerably more complicated to fly than small UAVs, and have a regulatory component that is currently undefined in the US. Thus manned-aircraft are the only choice throughout most of Alaska, where our research is based, when other ground-based field work is not required.

While lidar is also typically flown from manned-aircraft, photogrammetry offers several advantages for mapping snow depth. Both offer the advantage of mapping large spatial-scales, but the photogrammetric method allows creation of a color orthoimage that is perfectly co-registered with the DEM. For snow studies, this image allows us to unambiguously identify what is snow and what is not, especially useful in thin snow-packs or those covering aufeis, as well as useful for recognizing structures in the snow like barchans and sastrugi. When interpreting the difference maps, these summer and winter images allow us to investigate changes that seem suspect, such as those we described related to vegetation or sediment erosion. We found that our photogrammetric system had about twice the precision as the lidar system we compared to (8 vs. 16 cm respectively) and about the same accuracy, and thus the photogrammetric system can measure thinner snowpacks more accurately. The photogrammetric system is also substantially less expensive than most lidar units, reducing the cost of ownership for research groups wanting to operate their own systems.

Photogrammetry from manned-aircraft thus fills an important gap between ground-based and satellite methods, not just for snow depth but for measuring nearly any change in topography. No satellite methods can produce DEMs of our resolution and quality, though they operate on larger spatial-scales where such resolution and quality may not be required, such as ice sheets dynamics. Those satellite techniques that can detect change at the centimeter level, such as InSAR and its Persistent Scatter techniques, require substantial expertise to implement, have a variety of limitations (look-angles, shadowing/layover, phase decorrelation, scatterer permanence, etc), and have high data costs (Delacourt et al., 2007; Ferretti et al., 2001; Nolan and Fatland, 2003). Given the cost of repeat lidar from manned aircraft, most cryospheric scientists studying landscape change resort to extrapolation of ground-based measurements using GPS and increasingly sUASs, with the essentially unverifiable assumption that their measurements are representative of the broader area. Our study of snow-depths has demonstrated that using photogrammetry from manned-aircraft fills a niche that approaches the spatial-scales of satellites with the accuracy of ground-based measurements, for about the price of either. Glacier melt, coastal erosion, thermokarst, aufeis dynamics, and landslides are all examples of topographic changes in the cryosphere that we have also measured without resorting to extrapolation, and done so at lower cost than field measurements that generate only point measurements. Given that nearly all experimental field designs are attempts to minimize errors due to extrapolation of point measurements, this method has the potential to transform our study designs and thereby remove many of the impediments to our understanding of the cryosphere and the changes occurring within it.

8. Conclusions

This paper presents a method for measuring topographic change from manned aircraft that is accurate enough to measure the snow depth of most of the snow packs found worldwide. It can be used to map snow-depth of entire watersheds, with system costs that are much lower than lidar and operational costs on par with ground measurements that only yield transect measurements within those watersheds. This airborne method allowed us to measure topography with a geolocation accuracy of ± 30 cm and a precision of ±8 cm at a spatial resolution of centimeters to decimeters. We used these maps to measure snow depth by subtracting a snow-free map from a snow-covered map, and found these difference maps have a snow depth accuracy of ± 10 cm when confounding influences of other real changes could be minimized. The mapping technique is based on digital photogrammetry that uses consumer-grade cameras, multi-frequency GPS, and Structure from Motion algorithms, but requires no IMU, on-board computer, or ground control. The airborne methods are straightforward and the processing is done by off-the-shelf software that is reasonably user-friendly. All of the components of our system are under intense consumer pressure to improve, thus future improvements to our results are likely. The main conclusion of this paper is that centimeter-scale change-detection is now within reach of many earth scientists who previously could not afford it, and that this technology is already being used to measure snow depth as well as other cryospheric changes at unprecedented accuracy and cost.

Acknowledgements

We would like to thank members of the SnowStar 2014 team for data collection in the Hulahula River watershed, Turner Nolan for assistance in photo acquisitions, John Arvesen and Ted Hildum for developing our intervalometer, the US Fish and Wildlife Agency's Arctic Refuge staff for field support, and Mark Fahnestock for feature tracking assistance. We thank the two reviewers who provide many useful comments on the paper. This research was supported in part by the Arctic Landscape Conservation Cooperative and the USGS Alaska Climate Science Center (PI Nolan, Cooperative Agreements F10AC00755 and F11AC00607), by NASA (PI Larsen, Grant NNX13AD52A), and by NSF (PI Sturm, Grant OPP-1023052), and by Fairbanks Fodar (www.fairbanksfodar.com).

Author Contributions

All authors contributed substantially to the writing of the manuscript, data analysis, and the overall project. Nolan was primarily responsible for photogrammetric technique development and for acquiring and processing the mapping data used for snow depth measurements. Larsen was primarily responsible for acquiring and processing the lidar and SfM validation data used at Minto Flats. Sturm was primarily responsible for snow probe data collection and processing at all locations.

References

Barnett, T. P., Adam, J. C., and Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow-dominated regions, Nature, 438, 303-309, 2005.
Bauder, A., Funk, M., and Huss, M.: Ice-volume changes of selected glaciers in the Swiss Alps since the end of the 19th century, Annals of Glaciology, 46, 145-149, 2007.
Berthier, E., Vadon, H., Baratoux, D., Arnaud, Y., Vincent, C., Feigl, K., Remy, F., and Legresy, B.: Surface motion of mountain glaciers derived from satellite optical imagery, Remote Sensing of Environment, 95, 14-28, 2005.
Bitelli, G., Dubbini, M., and Zanutta, A.: Terrestrial laser scanning and digital photogrammetry techniques to monitor landslide bodies, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 35, 246-251, 2004.
Brandenberger, A. J.: Map of the McCall Glacier, Brooks Range, Alaska, American Geographical Society, New York, AGS Report 11 pp., 1959.
Bühler, Y., M. Marty, L. Egli, J. Veitinger, T. Jonas, P. Thee, and C. Ginzler (2015), Snow depth mapping in high-alpine catchments using digital photogrammetry, The Cryosphere, 9(1), 229-243, doi:10.5194/tc-9-229-2015
Buhler, Y., Meier, L., and Ginzler, C.: Potential of Operational High Spatial Resolution Near-Infrared Remote Sensing Instruments for Snow Surface Type Mapping, 2015. 2015.
Castebrunet, H., Eckert, N., Giraud, G., Durand, Y., and Morin, S.: Projected changes of snow conditions and avalanche activity in a warming climate: the French Alps over the 2020–2050 and 2070–2100 periods, The Cryosphere, 8, 1673-1697, 2014.
Clifford, D.: Global estimates of snow water equivalent from passive microwave instruments: history, challenges and future developments, International Journal of Remote Sensing, 31, 3707-3726, 2010.
Cline, D. W.: Digital photogrammetric determination of alpine snowpack distribution for hydrological modeling, Proceedings of the Western Snow Conference, 1994. 115, 1994.
Conway, H. and Abrahamson, J.: Snow stability index, J. Glaciology, 30 (106): 321, 327, 1984.
Cox, L. and March, R.: Comparison of geodetic and glaciological mass balance techniques, Gulkana Glacier, Alaska, draft, 2003. 2003.
d'Oleire-Oltmanns, S., Marzolff, I., Peter, K. D., and Ries, J. B.: Unmanned aerial vehicle (UAV) for monitoring soil erosion in Morocco, Remote Sensing, 4, 3390-3416, 2012.
Deems, J. S., Painter, T. H., and Finnegan, D. C.: Lidar measurement of snow depth: a review, Journal of Glaciology, 59, 467-479, 2013.
Delacourt, C., Allemand, P., Berthier, E., Raucoules, D., Casson, B., Grandjean, P., Pambrun, C., and Varel, E.: Remote-sensing techniques for analysing landslide kinematics: a review, Bulletin de la Societe Geologique de France, 178, 89-100, 2007.
Déry, S. J. and Brown, R. D.: Recent Northern Hemisphere snow cover extent trends and implications for the snow ‐ albedo feedback, Geophysical Research Letters, 34, 2007.
Eisenbeiß, H.: UAV photogrammetry, 2009.Dipl.-Ing., University of Technology Dresden, Zürich, Eidgenössischen Technischen Hochschule, ETH, Zürich, 237 pp., 2009.
Esch, D. C.: Long-term evaluations of insulated roads and airfields in Alaska, Transportation research record, 1995. 56-62, 1995.
Fassnacht, S. and Deems, J.: Measurement sampling and scaling for deep montane snow depth data, Hydrological processes, 20, 829-838, 2006.

Ferretti, A., Prati, C., and Rocca, F.: Permanent scatterers in SAR interferometry, Geoscience and Remote Sensing, IEEE Transactions on, 39, 8-20, 2001.

Fonstad, M. A., Dietrich, J. T., Courville, B. C., Jensen, J. L., and Carbonneau, P. E.: Topographic structure from motion: a new development in photogrammetric measurement, Earth Surface Processes and Landforms, 38, 421-430, 2013.

Gao, Y. and Shen, X.: A new method for carrier-phase-based precise point positioning, Navigation, 49, 109-116, 2002.

Gauthier, D., Conlan, M., and Jamieson, B.: Photogrammetry of fracture lines and avalanche terrain: Potential applications to research and hazard mitigation projects, International Snow Science Workshop, Banff, Canada, 109-115, 2014.

Goodrich, L.: The influence of snow cover on the ground thermal regime, Canadian Geotechnical Journal, 19, 421-432, 1982.

Hamilton, T. D.: Comparative glacier photographs from northern Alaska, Journal of Glaciology, 5, 479-487, 1965.

Hitchcock, C. B. and Miller, O. M.: Nine glacier maps, northwestern North America, American Geographical Society, New York, AGS Special Publication 34 pp., 1960.

Hopkinson, C., Sitar, M., Chasmer, L., and Treitz, P.: Mapping snowpack depth beneath forest canopies using airborne lidar, Photogrammetric Engineering & Remote Sensing, 70, 323-330, 2004.

Huang, L. and Li, Z.: Comparison of SAR and optical data in deriving glacier velocity with feature tracking, International Journal of Remote Sensing, 32, 2681-2698, 2011.

Hugenholtz, C. H., Whitehead, K., Brown, O. W., Barchyn, T. E., Moorman, B. J., LeClair, A., Riddell, K., and Hamilton, T.: Geomorphological mapping with a small unmanned aircraft system (sUAS): Feature detection and accuracy assessment of a photogrammetrically-derived digital terrain model, Geomorphology, 194, 16-24, 2013.

Irschara, A., Kaufmann, V., Klopschitz, M., Bischof, H., and Leberl, F.: Towards fully automatic photogrammetric reconstruction using digital images taken from UAVs, na, 2010.

James, L. A., Hodgson, M. E., Ghoshal, S., and Latiolais, M. M.: Geomorphic change detection using historic maps and DEM differencing: The temporal dimension of geospatial analysis, Geomorphology, 137, 181-198, 2012.

Jamieson, B. and Stethem, C.: Snow avalanche hazards and management in Canada: challenges and progress, Natural hazards, 26, 35-53, 2002.

Johnson, A. J., Larsen, C. F., Murphy, N., Arendt, A. A., and Zirnheld, S. L.: Mass balance in the Glacier Bay area of Alaska, USA, and British Columbia, Canada, 1995–2011, using airborne laser altimetry, Journal of Glaciology, 59, 632-648, 2013.

Koenderink, J. J. and Van Doorn, A. J.: Affine structure from motion, JOSA A, 8, 377-385, 1991.

König, M. and Sturm, M.: Mapping snow distribution in the Alaskan Arctic using aerial photography and topographic relationships, Water Resources Research, 34, 3471-3483, 1998.

Krimmel, R. M.: Mass balance and volume of South Cascade Glacier, Washington 1958–1985. In: Glacier fluctuations and climatic change, Springer, 1989.

Lee, C., Jones, S., Bellman, C., and Buxton, L.: DEM creation of a snow covered surface using digital aerial photography, Remote Sensing and Spatial Information Sciences., Beijing 831-836, 2008.

Lemke, P., Ren, J., Alley, R. B., Allison, I., Carrasco, J., Flato, G., Fujii, Y., Kaser, G., Mote, P., and Thomas, R. H.: Observations: Changes in snow, ice and frozen ground, Part of the Working Group I contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change
Cambridge University Press, 337-383 pp., 2007.

Liston, G. E., Haehnel, R. B., Sturm, M., Hiemstra, C. A., Berezovskaya, S., and Tabler, R. D.: Instruments and methods simulating complex snow distributions in windy environments using SnowTran-3D, Journal of Glaciology, 53, 241-256, 2007.

Liston, G. E. and Sturm, M.: Winter precipitation patterns in arctic Alaska determined from a blowing-snow model and snow-depth observations, Journal of hydrometeorology, 3, 646-659, 2002.

Lucieer, A., de Jong, S., and Turner, D.: Mapping landslide displacements using Structure from Motion (SfM) and image correlation of multi-temporal UAV photography, Progress in Physical Geography, doi: 10.1177/0309133313515293, 2013. 97-116, 2013.

Maune, D.: Digital elevation model technologies and applications: the DEM users manual, Asprs Publications, 2001.

McCurdy, P., Woodward, L., Davidson, J., Wilson, R., and Ask, R.: Manual of photogrammetry. American Society of Photogrammetry. Pitman Publishing Corporation, 1944.

McKay, G.: Problems of measuring and evaluating snow cover, 1968, 49-62.

Ménard, C. B., Essery, R., Pomeroy, J., Marsh, P., and Clark, D. B.: A shrub bending model to calculate the albedo of shrub ‐ tundra, Hydrological Processes, 28, 341-351, 2014.

Miller, P. E., Kunz, M., Mills, J. P., King, M. A., Murray, T., James, T. D., and Marsh, S. H.: Assessment of glacier volume change using ASTER-based surface matching of historical photography, Geoscience and Remote Sensing, IEEE Transactions on, 47, 1971-1979, 2009.

Najibi, N. and Arabsheibani, R.: Snow-covered surface variability and DEM generation using aerial photogrammetry in Mount Odin, Canada, Geodesy and Cartography, 39, 113-120, 2013.

Nex, F. and Remondino, F.: UAV for 3D mapping applications: a review, Applied Geomatics, 6, 1-15, 2014.

Nolan, M., Arendt, A., and Rabus, B.: Volume change of McCall Glacier, Arctic Alaska, from 1956 to 2003, Annals of Glaciology, 42, 409-416, 2005.

Nolan, M., Churchwell, R., Adams, J., McClellands, J., Tape, K., Kendall, S., Powell, A., Dunton, K., Payer, D., and Martin, P.: Predicting the impact of glacier loss on fish, birds, floodplains, and estuaries in the Arctic National Wildlife Refuge, Fairbanks, AK2011, 49-54.

Nolan, M. and Fatland, D. R.: Penetration depth as a DInSAR observable and proxy for soil moisture, Geoscience and Remote Sensing, IEEE Transactions on, 41, 532-537, 2003.

Nuth, C. and Kääb, A.: Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change, The Cryosphere, 5, 271-290, 2011.

Offenbacher, E. L. and Colbeck, S. C.: Remote Sensing of Snow Covers Using the Gamma-Ray Technique, DTIC Document, 1991.

Otake, K.: Snow survey by aerial photographs, GeoJournal, 4, 367-369, 1980.

Pauli, J. N., Zuckerberg, B., Whiteman, J. P., and Porter, W.: The subnivium: a deteriorating seasonal refugium, Frontiers in Ecology and the Environment, 11, 260-267, 2013.

Post, A.: Annual aerial photography of glaciers in northwest North America: How it all began and its golden age, Physical Geography, 16, 15-26, 1995.

Post, A.: Distribution of surging glaciers in western North America, Journal of Glaciology, 8, 229-240, 1969.

Prokop, A.: Assessing the applicability of terrestrial laser scanning for spatial snow depth measurements, Cold Regions Science and Technology, 54, 155-163, 2008.

Pruitt, W. O.: Snow as a factor in the winter ecology of the barren ground caribou (Rangifer arcticus), Arctic, 1959. 158-179, 1959.

Rawls, W., Jackson, T., and Zuzel, J.: Comparison of areal snow storage sampling procedures for rangeland watersheds, Nordic Hydrology, 11, 71-82, 1980.

Rinaudo, F., Chiabrando, F., Lingua, A. M., and Spanò, A. T.: Archaeological site monitoring: UAV photogrammetry can be an answer, The International archives of the photogrammetry, Remote sensing and spatial information sciences, 39, 583-588, 2012.

Rittger, K., Painter, T. H., and Dozier, J.: Assessment of methods for mapping snow cover from MODIS, Advances in Water Resources, 51, 367-380, 2013.

Robinson, D. A., Dewey, K. F., and Heim Jr, R. R.: Global snow cover monitoring: An update, Bulletin of the American Meteorological Society, 74, 1689-1696, 1993.

Rott, H., Cline, D., Duguay, C., Essery, R., Haas, C., Macelloni, G., Malnes, E., Pulliainen, J., Rebhan, H., and Yueh, S.: CoReH2O Â¿ A Ku-and X-Band SAR Mission for Snow and Ice Monitoring, 2008, 1-4.

Russell, D. E., Martell, A. M., and Nixon, W. A.: Range ecology of the Porcupine caribou herd in Canada, Rangifer, 13, 1-168, 1993.

Ryan, J., Hubbard, A., Todd, J., Carr, J., Box, J., Christoffersen, P., Holt, T., and Snooke, N.: Repeat UAV photogrammetry to assess calving front dynamics at a large outlet glacier draining the Greenland Ice Sheet, The Cryosphere, 9, 1–11, 2015 doi:10.5194/tc-9-1-2015

Serreze, M. C., Clark, M. P., Armstrong, R. L., McGinnis, D. A., and Pulwarty, R. S.: Characteristics of the western United States snowpack from snowpack telemetry (SNOTEL) data, Water Resources Research, 35, 2145-2160, 1999.

Slater, A. G. and Clark, M. P.: Snow data assimilation via an ensemble Kalman filter, Journal of Hydrometeorology, 7, 478-493, 2006.

Snay, R. A. and Soler, T.: Continuously operating reference station (CORS): history, applications, and future enhancements, Journal of Surveying Engineering, 134, 95-104, 2008.

Sturm, M.: Field techniques for snow observations on sea ice, Field Techniques for Sea Ice Research, 2009. 25-47, 2009.

Sturm, M. and Benson, C.: Scales of spatial heterogeneity for perennial and seasonal snow layers, Annals of Glaciology, 38, 253-260, 2004.

Sturm, M., Douglas, T., Racine, C., and Liston, G. E.: Changing snow and shrub conditions affect albedo with global implications, Journal of Geophysical Research: Biogeosciences (2005–2012), 110, 2005.

Sturm, M., Hellig, H., Urban, F., and Liston, G.: The snow cover of the Arctic National Wildlife Refuge, Arctic, in prep. in prep.

Sturm, M. and Holmgren, J.: Self recording snow depth probe. Office, U. P. (Ed.), US Army Corps of Engineers, 1999.

Sturm, M., Holmgren, J., and Liston, G. E.: A seasonal snow cover classification system for local to global applications, Journal of Climate, 8, 1261-1283, 1995.

Taber, S.: Frost heaving, The Journal of Geology, 1929. 428-461, 1929.

Vanderjagt, B., Turner, D., Lucieer, A., and Durand, M.: Retrieval of Snow Depth Using Low Cost UAV-Based Lidar and Photogrammetry, 2013, 0593.

Warren, S. G.: Optical properties of snow, Reviews of Geophysics, 20, 67-89, 1982.

Weller, G., Nolan, M., Wendler, G., Benson, C., Echelmeyer, K., and Untersteiner, N.: Fifty years of McCall Glacier research: from the International Geophysical Year, 1957-1958, to the International Polar Year, 2007-2008, Arctic, 60, 101-110, 2007.

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, 2012.

Wheaton, J. M., Brasington, J., Darby, S. E., and Sear, D. A.: Accounting for uncertainty in DEMs from repeat topographic surveys: improved sediment budgets, Earth Surface Processes and Landforms, 35, 136-156, 2010.

Whitehead, K., Moorman, B., and Hugenholtz, C.: Brief Communication: Low-cost, on-demand aerial photogrammetry for glaciological measurement, The Cryosphere, 7, 1879-1884, 2013.
Woodget, A., Carbonneau, P., Visser, F., and Maddock, I.: Quantifying submerged fluvial topography using hyperspatial resolution UAS imagery and structure from motion photogrammetry, Earth Surface Processes and Landforms, 2014. 2014.
Yan, K. and Cheng, T.: Close Shot Photogrammetry for Measuring Wind-Drifted Snow Distribution on Stepped Flat Roofs, 2008, 332-335.

Tables

Table 1. Fairbanks International Airport accuracy and precision assessment. Values for the reference DEM (6 Oct 13) are geolocation offsets to 29 GCPs. All other offset are co-registration offsets of that DEM minus the reference DEM for the snow-free area of the runway. The group statistics at bottom do not include the reference DEM. The first 3 columns of numbers represent accuracy while the 4th represents precision.

Date	**Easting offset (m)**	**Northing offset (m)**	**Elevation offset (m)**	**Elev. St. Dev (cm)**	**GSD (cm)**	**Notes**
06 Oct 13	0	0.30	0.13	1.7	6	Reference, snow free
30 Sept 13	-0.15	-0.51	0.45	5.3	6	Snow free
21 Jan 14	-0.11	-0.48	0.24	5.8	6	Snow covered
18 Feb 14	0.02	-0.18	-0.29	5.2	6	Peak snow
03 April 14	-0.18	-0.09	-0.04	4.2	12	Snow covered
20 April 14	-0.25	-0.46	0.31	5.0	14	Mostly melted
Means:	**-0.13**	**-0.34**	**0.13**	**5.1**		
±(Range/2):	**±0.13**	**±0.21**	**±0.37**	**±0.08**		

Table 2. Minto Flats accuracy assessment. Values for the reference DEM are geolocation offsets to 21 GCPs. All other values are co-registration offsets of that DEM minus the reference DEM. Statistics at bottom do not include the reference DEM.

Date	**Easting offset (m)**	**Northing offset (m)**	**Elevation offset (m)**	**GSD (m)**	**Notes**
03 April 14	0	0	0.23	0.15	Reference Map
28 Sept 13	-0.01	0.25	0.03	0.15	snow free
27 Jan 14	0.02	0.26	0.03	0.15	snow covered
19 April 14	-0.07	0.23	-0.02	0.14	snow melting
06 Nov 14	0.01	0.15	0.02	0.15	Frozen, snow dusting
08 Nov 14	-0.06	0.22	0.30	0.15	Frozen, snow dusting
Means:	**-0.02**	**0.22**	**0.07**		
±(Range/2):	**±0.05**	**±0.05**	**±0.16**		