Multitemporal terrestrial laser scanning point clouds for thaw subsidence observation at Arctic permafrost monitoring sites

Katharina Anders,¹ ² Sabrina Marx,¹ Benoit Boi,³ ⁴ Benjamin Herfort,¹ Evan James Milcox,⁵ Moritz Langer,³ Julia Boike,³ ⁴ Philip Marsh⁶ and Bernhard Höfle¹ ² ⁶

¹ ³D Geospatial Data Processing Group (3DGeo), Institute of Geography, Heidelberg University, 69120 Heidelberg, Germany
² Interdisciplinary Centre for Scientific Computing (IWR), Heidelberg University, 69120 Heidelberg, Germany
³ Alfred Wegener Institute (AWI), Helmholtz Centre for Polar and Marine Research, 14473 Potsdam, Germany
⁴ Geography Department, Humboldt University, 10099 Berlin, Germany
⁵ Cold Regions Research Centre, Wilfrid Laurier University, Waterloo N2L 3C5, Canada
⁶ Heidelberg Centre for the Environment (HCE), Heidelberg University, 69120 Heidelberg, Germany

Received 24 April 2019; Revised 29 January 2020; Accepted 30 January 2020

*Correspondence to: Katharina Anders, 3DGeo at the Institute of Geography, Im Neuenheimer Feld 368, 69120 Heidelberg, Germany.
E-mail: katharina.anders@uni-heidelberg.de

© 2020 The Authors. Earth Surface Processes and Landforms published by John Wiley & Sons Ltd

ABSTRACT: This paper investigates different methods for quantifying thaw subsidence using terrestrial laser scanning (TLS) point clouds. Thaw subsidence is a slow (millimetre to centimetre per year) vertical displacement of the ground surface common in ice-rich permafrost-underlain landscapes. It is difficult to quantify thaw subsidence in tundra areas as they often lack stable reference frames. Also, there is no solid ground surface to serve as a basis for elevation measurements, due to a continuous moss–lichen cover. We investigate how an expert-driven method improves the accuracy of benchmark measurements at discrete locations within two sites using multitemporal TLS data of a 1-year period. Our method aggregates multiple experts’ determination of the ground surface in 3D point clouds, collected in a web-based tool. We then compare this to the performance of a fully automated ground surface determination method. Lastly, we quantify ground surface displacement by directly computing multitemporal point cloud distances, thereby extending thaw subsidence observation to an area-based assessment. Using the expert-driven quantification as reference, we validate the other methods, including in-situ benchmark measurements from a conventional field survey. This study demonstrates that quantifying the ground surface using 3D point clouds is more accurate than the field survey method. The expert-driven method achieves an accuracy of 0.1 ± 0.1 cm. Compared to this, in-situ benchmark measurements by single surveyors yield an accuracy of 0.4 ± 1.5 cm. This difference between the two methods is important, considering an observed displacement of 1.4 cm at the sites. Thaw subsidence quantification with the fully automatic benchmark-based method achieves an accuracy of 0.2 ± 0.5 cm and direct point cloud distance computation an accuracy of 0.2 ± 0.9 cm. The range in accuracy is largely influenced by properties of vegetation structure at locations within the sites. The developed methods enable a link of automated quantification and expert judgement for transparent long-term monitoring of permafrost subsidence. © 2020 The Authors. Earth Surface Processes and Landforms published by John Wiley & Sons Ltd

KEYWORDS: change analysis; 3D geoinformation; ground surface displacement; permafrost monitoring; multitemporal LiDAR

Introduction

Background

The Earth’s surface is subject to permanent topographic change. Terrestrial laser scanning (TLS) is a highly accurate sensor system increasingly used in Earth Sciences for observing such surface changes. From repeated data acquisitions, multitemporal 3D point clouds can be generated and used to quantify surface change over time (Eitel et al., 2016). While applications of TLS-based change analysis are increasing, it has not been used to quantify thaw subsidence of the ground surface in Arctic permafrost environments, where ground surface changes are particularly difficult to observe.

Thaw subsidence is a vertical displacement of the ground surface in landscapes underlain by permafrost (i.e. ground that remains at or below 0°C for at least two consecutive years; NRC Canada, 1988). Permafrost occurs in 13–18% of the exposed land area in the northern hemisphere (Heginbottom et al., 2012) and is highly sensitive to changes in climate (e.g. Kane et al., 1991). Where permafrost is ice-rich, permafrost thaw involves melting of ground ice, which leads to soil consolidation and subsequently subsidence of the ground surface (Shur et al., 2005; Anisimov et al., 2007; Allison et al., 2012).
This phenomenon occurs at a rate of a few centimetres per year across large areas of permafrost-underlain landscapes (e.g., Shiklomanov et al., 2013; Günther et al., 2015; Strelets’ky et al., 2017; Antonova et al., 2018). However, thaw subsidence is difficult to observe in permafrost-underlain tundra landscapes, as there are often no absolute reference frames to compare to the subtle, widespread topographic change in the landscape (Shur et al., 2011; Shiklomanov et al., 2013). Consequently, thaw subsidence and how it is affected by environmental factors is not well understood.

When no fixed reference frame is available, thaw subsidence can be measured by referencing benchmarks that are anchored in permafrost. The benchmarks are thereby isolated from thaw subsidence processes in the active layer, which is the ground above the permafrost that thaws and freezes annually (cf. Brown et al., 2000; CALM, 2018), as well as from permafrost thaw. Figure 1 schematically illustrates the phenomenon of thaw subsidence and a benchmark-based measurement setup.

Besides this benchmark-based approach, other studies have used global navigation satellite system (GNSS) observations or GNSS interferometric techniques (e.g., Shiklomanov et al., 2013; Strelets’ky et al., 2017) for repeat measurements at fixed locations. Spatially extensive data acquired from repeat airborne laser scanning or unmanned aerial vehicle (UAV) photogrammetry provide elevation data to quantify permafrost-related surface deformation (Fraser et al., 2016; van der Sluijs et al., 2018) but lack accuracy in spatial measurements and georeferencing to quantify small-scale surface changes of thaw subsidence. Satellite-based InSAR techniques have been used to quantify small-scale surface changes (e.g., Liu et al., 2014; Antonova et al., 2018), but are limited by interferometric phase decorrelation (Antonova et al., 2018). In this paper, we investigate the use of 3D point clouds captured by TLS to spatially quantify small-scale thaw subsidence.

Motivation and research gap

Surveying protocols exist for performing manual in-situ measurements in a consistent manner for thaw subsidence quantification, but as in every manual measurement, uncertainty is introduced by the subjectivity of the individual performing measurements. The potential for such errors increases as monitoring sites are generally designed for long-term surveying in permafrost research and it is unlikely that the same surveyor will perform all the measurements over several consecutive years. Using TLS data holds the advantage that multiple analysts can perform the same digital measurements and take their time to fulfill the task. The measurement process also becomes more transparent as the survey conditions represented by the 3D point cloud can be revisited numerous times to increase consistency and develop best practices. Moreover, the sample size of in-situ ground surface measurements through TLS is much larger than is possible by physical field measurements. Depending on the survey setup, the sample size increases by a factor of more than 100 000. It follows that the ground surface can be determined more accurately from the TLS data, given the high sample size of precise LiDAR measurements. Another important benefit of TLS-based surveying is that it allows the site to be unimpacted by humans, while manual measurements repeated over a long time will lead to ground disturbance around the site.

It was found in 3D point cloud analysis that, for some geoinformation extraction problems, manual editing via a volunteered geographic information (VGI) approach (Goodchild, 2007) can produce better results than automated methods (Herfort et al., 2018). The framework of Herfort et al. (2018) provides a new method to capture VGI using 3D point clouds, which was shown to reduce uncertainties in data. Manual editing is performed in a web-based tool that splits a geoprocessing task into multiple components called ‘micro tasks’, with multiple users solving individual micro tasks. Aggregating multiple contributions from several users increased the accuracy compared to solutions from a single user.

Measuring benchmarks at permafrost monitoring sites is a geoinformation extraction problem that can easily be split up into micro tasks of the area around each benchmark pole. The difficult aspect of conducting such measurements – both in the field and in TLS point clouds – is determining the ground surface as reference for measuring the benchmark pole length. In permafrost regions, often no solid surface is available to serve as basis for ground surface elevation measurements, due to continuous moss and lichen cover (cf. Figure 3 later).

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Schematic of a survey setup for observing thaw subsidence resulting from an increase in thaw depth between consecutive summers under warming conditions (not to scale). The vertical displacement of the ground surface can be quantified at a subsidence station (a pole anchored deep into the permafrost that is not affected by displacement processes in the active layer) from repeated measurements of the pole length above the ground surface. A reference plane is used in manual surveys to represent the ground surface in a more consistent manner. Alternatively, the position of the ground surface can be recorded as 3D coordinates by terrestrial laser scanning in combination with an absolute reference frame, which is provided by the subsidence stations. [Colour figure can be viewed at wileyonlinelibrary.com]
Typically it is assumed that the thickness of the moss and lichen layer does not change from one year to the next, so the visible moss and lichen surface layer can be treated as the ground surface. Determining the ground surface using TLS point clouds becomes a complex task as the ground surface is rough and partly penetrable, requiring experts rather than non-expert analysts. For the purpose of our study, we define experts as people who have experience working at Arctic permafrost sites and/or with 3D point cloud data. We hypothesize that aggregating multiple contributions for individual micro tasks could reduce subjective errors of the single expert analysts and increase overall measurement accuracy, as was found in Herfort et al. (2018). Hereinafter we will refer to the approach of using multiple expert measurements as ‘expert-driven’.

Besides the expert-driven approach that relies on manual information extraction, ground surface measurements in the 3D point cloud can be automated using algorithms. However, it is not known if an automated method using 3D point cloud data is able to achieve a comparable accuracy to the expert-driven method, in contrast to in-situ measurements at benchmark poles.

Direct multitemporal comparison of ground surface elevations captured by TLS at tundra sites is challenging given the extensive soft ground cover and absence of stable ground (cf. Barnhart and Crosby, 2013; Marx et al., 2017). In addition, dense shrubs often grow in warmer tundra regions, which may prevent the laser beam from reaching the ground surface. This problem is compounded by highly variable microtopography, such as mineral earth hummocks that impede the full coverage of an area when using a typical TLS survey setup. The depth of laser beam penetration into the vegetation layer is further influenced by the beam incidence angles of a survey setup (Marx et al., 2017). As a result, sampling is strongly influenced by the scan geometry of individual surveys (cf. Fan et al., 2014). Given such highly variable surface characteristics, ground surface representations in 3D point clouds can differ greatly between acquisition dates (Lague et al., 2013; Eltner et al., 2016).

For change analysis in complex topographic environments, direct point cloud distance calculation is a widely used method for 3D surface change quantification (Lague et al., 2013). The accuracy of point cloud alignment is critical in this method, since errors in the point cloud co-registration propagate to the generated results (Wujanz et al., 2018). An advantage of comparing benchmark pole lengths measured for each survey date, in contrast to direct point cloud comparison, is that alignment errors in the multitemporal data are excluded from the quantification.

Research objective

To solve the previously outlined challenges in quantifying thaw subsidence, we develop a novel approach using TLS point clouds captured in a tundra landscape underlain by continuous, ice-rich permafrost in the western Canadian Arctic. We investigate to what degree TLS point clouds can improve benchmark-based surveys by increasing the accuracy of the quantification of thaw subsidence. This method could improve observations throughout long time series and for varying geographic locations of permafrost-underlain landscapes when compared to in-situ measurements of benchmarks. To achieve this, we compare the accuracy of different methods which comprise manual measurements made by experts (i) on site and (ii) in 3D point clouds. Further, we (iii) automate the expert-based measurement in the 3D point clouds, and lastly investigate the feasibility of (iv) directly comparing the ground surface between multitemporal point clouds.

Through the development and investigation of these multiple methods, our research presents a novel approach for surface change analysis in 3D geospatial data, which incorporates human surveyor input in the information extraction. The accuracy assessment of direct multitemporal 3D point cloud comparison provides information about the suitability of TLS for future research on thaw subsidence and the linkage between surface...
displacement and the complex interplay of soil ice content, thaw, and melt water drainage or evapotranspiration.

**Study Area and Data**

The study area is located at the Trail Valley Creek (TVC) research watershed (68°44′25″N, 133°29′36″W; Marsh et al., 2010), approximately 50 km north of Inuvik in the Northwest Territories, Canada (Figure 2). This upland plateau east of the Mackenzie Delta is characterized by gently sloping topography, incised by the broad river valley of TVC. Due to cryoturbation associated with permafrost processes, mineral earth hummocks dominate the ground surface (Quinton and Marsh, 1999). The site is within the northern margin of the transition zone from forest to tundra, and vegetation is dominated by dwarf shrub and herbaceous tundra, with a mix of short willows, birch, lichen, moss, and grass (Marsh et al., 2010).

Temperatures range from over 20°C during short summers to below −40°C in winter. Annual average temperature for the period 1981 to 2010 at the Inuvik Airport meteorological station is −8.2°C, with daily mean temperature of 14.1°C in July and −26.9°C in January. Average annual precipitation over the same period is 115 mm rainfall and 159 cm snowfall (Environment Canada, 2017). Active layer thickness measured in the plot sites of this study ranged from 58.0 to 95.0 cm (median: 81.0 cm) at the end of the thawing season in August 2016.

**Field surveys**

Surveys were conducted in June 2015, August 2015, and August 2016, in order to incorporate the beginning of the thawing season in June and the end of the thawing season in August. From these acquisitions, thaw subsidence can be quantified for two periods: (1) a seasonal period from June to August 2015 and (2) an interannual period from August 2015 to 2016.

Data were acquired for two plot sites with extents of 42 × 53 and 29 × 48 m, respectively (Figure 2A). Each site was selected to represent a relatively homogenous surface structure, but different vegetation covers. Site 1 is predominantly covered by moss and lichen with sporadically distributed small shrubs (Figure 2C), while more consistent dwarf shrub cover characterizes site 2 (Figure 2D). All data used in this paper are openly available with detailed metadata (Anders et al., 2018a). In the following, we describe the important aspects of data acquisition and processing of relevance for this paper.

Since the study area lacks any stable landscape elements to reference ground surface change, 12 fixed glass-fibre poles of 1.0 cm diameter were installed as benchmarks at both sites by drilling holes roughly 1.5 m into the underlying permafrost, at the beginning of the first field campaign in June 2015. The benchmark poles, subsequently referred to as subsidence stations, were placed on the tops of hummocks to make them easily visible to the TLS from surrounding scan positions. Each pole sticks out of the ground about 25–35 cm (Figure 3).

Assuming that the poles are stable, subsidence and heave of the active layer can be determined around the poles by measuring their height above the ground surface at successive points in time. The length of the subsidence station pole above the ground surface was measured manually during each field campaign. To reduce the subjectivity of different surveyors’ interpretation of the ground surface, a reference plane was established for the measurement procedure. For every measurement, a conventional Compact Disc (CD, radius: 12.0 cm, thickness: 0.1 cm) was laid on the ground surface during surveying by putting the pole through its centre hole (Figure 3). The pole length from the reference plane to the pole tip was measured five times around each pole using a folding ruler. The median of these five measurements determined the length value of the pole, and the difference from the previous length value of the pole at each subsidence station defines the net subsidence at each location. We use positive values for thaw subsidence, negative values represent ground heave.

![Photo of a subsidence station installed in the study area of Trail Valley Creek, NWT, Canada with a reference plane (here: conventional Compact Disc) for measurement.](https://wileyonlinelibrary.com)
TLS point cloud data

TLS data for the two sites were acquired during all three field campaigns using a Riegl VZ-400 terrestrial laser scanner (Riegl LMS, 2017) with a horizontal and vertical point spacing of 3 mm at 10 m measurement range. Each site was recorded from seven scan positions, with six positions surrounding the site extent and one centre position. The single point clouds acquired from these positions during a field campaign were registered using four cylindrical reflectors placed around the edges of a site. A GNSS receiver (Leica Viva GS15; Leica Geosystems, 2012) was used with base station (Leica Viva GS10; Leica Geosystems, 2015) in real-time kinematic mode to obtain measurements for georeferencing the data. Registration and georeferencing of the data were conducted in the software RiSCAN PRO (Riegl LMS, 2016).

To increase the accuracy of point cloud alignment between survey dates, georeferenced point clouds were co-registered using 7 of the 12 subsidence station tips as fixed control points at each site. The remaining five subsidence station tips were used to assess the co-registration accuracy based on residual distances between the control points. The co-registration reduced the root-mean-square (RMS) of vertical control point distances between survey dates to 0.2 cm; the RMS of 3D distances amounts to 1.0 cm.

The residual distances of multitemporal pole tip coordinates after co-registration not only provide an estimate of manual digitization accuracy and geodetic point cloud alignment, but also indicate the stability of the benchmark poles. To test for movement of the poles relative to each other within the sites between survey dates, we compute the rigid 3D transformation matrix between the pole tips of multitemporal acquisitions. For this, we use the pole tips of all subsidence stations, not only the independent control points. The RMS of residual 3D distances of all pole tips is <0.1 cm. This confirms our assumption that the benchmark poles are fixed in their relative position and are not subject to movement caused by freezing and thawing processes or other external influences.

Preprocessing of the TLS datasets comprises filtering to a maximum range of 30 m around the respective scan position and removing spatially isolated points using a filter for statistical outlier removal (SOR; Rusu and Cousins, 2011). All man-made objects, including the subsidence stations, are manually removed from the point clouds.

As input for the methods, we further filter non-ground vegetation points from the 3D point cloud data. We identify and remove vegetation points using the echo ratio (ER), a geometry-based measure for local transparency. ER is calculated as the relation between the number of points in a search sphere with a defined radius and the number of points in a vertical search cylinder with the same radius and infinite height (Höfle et al., 2009). We calculate the ER for the 3D point cloud of each survey date in a search radius of 2 cm and set the filtering threshold to 50%. We use the modular program system OPALS (Orientiation and Processing of Airborne Laser Scanning Data; Pfeifer et al., 2014) to calculate the echo ratio value for each point. The ER-filtered point clouds provide the data basis for all analyses in this research.

Methods

We develop and investigate different methods for quantifying thaw subsidence using multitemporal 3D TLS point clouds, for which Figure 4 depicts the structure of the main analysis steps.

First, we investigate how multitemporal TLS point clouds improve thaw subsidence quantification at discrete locations by measuring fixed benchmark poles. For this feature-based approach, two methods are used to determine the ground surface. The first method is expert-driven and aggregates contributions from multiple expert analysts who specify the 3D position of the ground surface plane for each subsidence station in the 3D point cloud. The second feature-based method automatically determines the ground surface plane from the 3D point cloud at the subsidence stations. To assess the accuracy of both methods independently, we validate the expert-driven method by using additional artificially generated micro tasks with known parameters of the ground surface plane. Therefore, results of the expert-driven method serve as reference data to assess the performance of the other methods.

Subsequently, we investigate direct 3D point cloud comparison for area-based quantification of thaw subsidence. We quantify change as the distance in ground surface elevation between multitemporal point clouds and compare the result at
the subsidence stations to the feature-based method, which is based on pole length differencing between survey dates. The resulting accuracy indicates if thaw subsidence observation can be extended from discrete benchmark locations within the sites to an areal assessment using multitemporal TLS point clouds.

**Expert-driven ground surface determination**

The feature-based method of thaw subsidence quantification follows the in-situ measurement procedure of quantifying the difference between the lengths of the subsidence station poles between survey dates. The pole length is derived as the vertical distance between pole tips and the ground surface. Given that the 3D position of the pole tips remains stable, performing the length measurements relies on the determination of the changed ground surface elevation in each point cloud.

**3D Micro tasks for ground surface plane adjustment**

In the expert-driven approach, we aggregate the assessment of the ground surface by multiple expert analysts in order to reduce measurement bias through individualsubjectivity. Experts are presented with 3D point cloud subsets of the subsidence stations (including artificial poles for validation) and position a 3D plane on the perceived ground surface for each pole and every survey date. For better comparability with the manual in-situ measurements, the presented point cloud subsets and reference planes around the subsidence stations have a radius of 6 cm (similar to the CD used as reference plane in the field).

The task of determining the reference plane on the ground surface in 3D point cloud subsets is implemented using a web-based tool for 3D geoinformation extraction from micro tasks developed by Herfort et al. (2018). The source code of our micro task project is openly available, together with a video of the online tool (Herfort et al., 2020). Figure 5A shows an example micro task from the web-based tool. As non-expert contributions are not used in this study, the level of expertise was queried from every contributor before starting the task.

The experts had to position and orientate the plane on the ground surface by modifying the height and rotation of the plane around its centre at the pole. The perspective on the micro task can be adjusted freely by zooming, shifting, and rotating the view. By submitting a contribution, the adjusted 3D centre coordinate and rotation of the plane around the x- and y-axes are stored. Using these parameters, we can reconstruct the 3D representation of the ground surface plane via the general equation of a plane in 3D Cartesian coordinate systems.

The contributions of each micro task are stored individually. We obtain one aggregated result of the determined ground surface plane per subsidence station using the median of contributions.

**Accuracy assessment of expert-driven plane determination**

Our deployed technology and methods are presumably more accurate than available in-situ measurements to quantify thaw subsidence, and there is no higher-accuracy reference dataset for validating results. Therefore, we generate artificial data to assess the accuracy of our method with respect to the known values. Any deviation of produced results subsequently describes the accuracy range of the approach.

Data for validating the expert-driven method of plane positioning are generated by including an artificially constructed point cloud of a subsidence station among the micro tasks in eight randomly, rigidly transformed versions (Figure 5B). The artificial point clouds were manually constructed from real TLS points of a subsidence station pole merged with points sampled on a plane of known parameters. Noise was induced by shifting the individual plane points slightly and included from the original measurements of the subsidence station by TLS. Experts were unaware of the artificial micro tasks among the real point cloud subsets. Expert contributions are analysed with respect to the estimated height and orientation of the determined ground surface plane. We compare the deviation of these parameters from the true plane for both the individual and the aggregated contributions in order to confirm that aggregation improved the accuracy of results (cf. Herfort et al., 2018).

**Automatic determination of the ground surface plane**

In the second feature-based approach, we replace manual methods with an automatic method that determines the ground surface plane using a robust least-squares plane adjustment. The method fits a plane to the input data points

![Figure 5](image-url)
by minimizing the RMS of point distances to the plane. This is done iteratively by detecting outliers by weighting points based on their residual. After eliminating all outliers, the final adjustment is performed with equally weighted remaining points (Dorninger and Nothegger, 2007). Planes are adjusted to all ground surface points within a horizontal distance of 6 cm around each subsidence station. We use the normal estimation procedure in OPALS for plane adjustment (Pfeifer et al., 2014).

In order to exclude locations of subsidence stations where the ground surface is poorly represented in the TLS point clouds, we use the RMS of point distances to the adjusted plane as intrinsic error measure of plane determination. As input points represent the ground surface, the RMS of point distances must be in the range of ground surface roughness and noise in the TLS data. Therefore, we set the threshold of excluding outlier subsidence stations from the surface change quantification to an RMS > 5.0 cm.

Feature-based thaw subsidence quantification

Thaw subsidence at each subsidence station is quantified by differencing pole lengths of successive survey dates. Using the ground plane obtained from either the expert-driven or the automatic method, pole length is quantified using the median of five measurements on the plane around the pole at a distance of 2.4 cm. We choose this approach to maintain consistency with manual in-situ measurements using a CD as reference plane.

We quantify overall thaw subsidence in the study area over a period as the median of thaw subsidence values per subsidence station. To assess the transferability of our methods between sites of different surface characteristics, we additionally provide thaw subsidence values separated by the two TLS plot sites.

Area-based quantification by direct point cloud comparison

We also quantify the vertical displacement of the ground surface by directly comparing multitemporal point clouds. Direct point cloud comparison enables an area-based assessment of thaw subsidence within the sites. The comparison to the feature-based approach at the benchmarks determines if direct point cloud comparison is feasible given the alignment uncertainty between the multitemporal 3D point clouds caused by the lack of stable reference frames in the study area.

We use the Multiscale Model to Model Cloud Comparison (M3C2) algorithm for multitemporal point cloud distance calculations (Lague et al., 2013). The algorithm estimates local positions in two input point clouds by using the surface normal vectors to determine the median point within a cylinder of defined radius. The distance between the local position estimates within each point cloud gives the local distance between the multitemporal point clouds. As thaw subsidence is a vertical ground surface displacement process, we set the M3C2 to orient the normal vectors strictly vertically. Similar to the reference plane used in the field, we set the radius for normal vector estimation to 6.0 cm.

As reference data are available only at the benchmark poles, we assess the accuracy of point cloud comparison using the median of resulting point distances within the area of the reference planes around each subsidence station. From this, the resulting M3C2-based point cloud distances can be compared to the expert-driven feature-based thaw subsidence quantification.

Validation of results

We base the overall performance of each method on the accuracy of results. The accuracy describes the objectivity of the methods and is expressed in the mean and standard deviation (SD) of differences to the reference value. Difference values are calculated by subtracting the reference value from the respective result. Results of the expert-driven method are validated with artificial data and subsequently provide the reference dataset for assessing the accuracy of all other methods. The measures previously introduced for thaw subsidence quantification using the different methods are listed in Table I. Following this, we use the respective denominations to indicate to which measure we refer in the accuracy assessment.

Results

Ground surface plane adjustment using aggregated expert contributions

Eight experts contributed to the task of ground surface plane determination. The results of the experts’ plane determination for the artificial subsidence stations are summarized in Table II. We separate the results per micro task into individual expert contributions and the aggregated contributions of the expert group. In contrast to the vertical position of centroids, the elevation measurements integrate the angular deviation caused by the non-flat ground surface plane, as measurements are taken around the pole at a distance of 2.4 cm. This influences the obtained elevation value (Figure 6B). The accuracy of the expert-determined plane compared to the true plane shows a large increase when the contributions of multiple experts are aggregated (Figure 6A). The mean of differences does not differ much between individual (0.4 cm) and aggregated (0.1 cm) elevation measurements. The SD of differences improves from 1.1 to 0.1 cm.

| Table I. Types of measures and their derivation for quantifying thaw subsidence using the different methods. Denominations are used for indicating the types of measure in the results, which are presented in the subsequent section |
|-------------------------------------------------|----------------|----------------------------------|
| Type of measure                                | Denomination  | Derivation of measure            |
| Pole length of a subsidence station (cm)        | PL            | Median of five single pole length measurements around a subsidence station |
| Thaw subsidence at benchmark locations (cm)     | TS            | • Feature-based methods: Difference of pole lengths (PL) of two survey dates for one subsidence station |
|                                                |               | • Direct point cloud comparison: Median of point-wise distances within the area of the reference planes around one subsidence station |
Table II. Results of individual and aggregated expert contributions for artificial micro tasks compared to the true values as angular deviation of plane surface normals and vertical offset of the centroid from the true plane height

| Variable                                  | Contributions | RMS  | Mean  | SD   | Median |
|-------------------------------------------|---------------|------|-------|------|--------|
| Angular deviation (°)                     | Individual    | 3.9  | 2.5   | 3.0  | 1.5    |
|                                           | Aggregated    | 0.7  | 0.7   | 0.2  | 0.7    |
| Vertical offset of centroids (cm)         | Individual    | <0.1 | <0.1  | <0.1 | <0.1   |
|                                           | Aggregated    | <0.1 | <0.1  | <0.1 | <0.1   |
| Difference of elevation measurement (cm)   | Individual    | 1.1  | 0.4   | 1.1  | 0.1    |
|                                           | Aggregated    | 0.1  | 0.1   | 0.1  | 0.1    |

Expert-driven and in-situ thaw subsidence quantification

We compute the vertical displacement of the ground surface for every subsidence station according to the feature-based method and assess the differences of in-situ field measurements from the expert-driven measurements at subsidence stations for both the pole length measurements and the subsequently derived thaw subsidence values.

The mean and SD of differences of individual pole length compared to the expert-driven method are larger and more variable (0.7 and 1.7 cm) than the differences of quantified thaw subsidence (0.4 and 1.5 cm). Quantifying thaw subsidence with in-situ measurements thus achieves a similar average measurement result, with a mean of differences close to our method’s accuracy (0.4 cm).

Automatic thaw subsidence quantification

In the automatic feature-based method we replace the manual determination of the ground surface plane with a fully automatic method. This method is validated by comparing pole length measurements for every subsidence station and survey date to the expert-based results as reference, as the accuracy of this method is validated independently. One subsidence station is excluded as outlier by the automatic method because the RMS of point distances to the adjusted plane exceeds the predetermined threshold of 5.0 cm. This subsidence station is the only one placed in the immediate vicinity of multiple shrubs, and consequently has poor representation of the ground surface in the TLS point clouds from any scan position.

The results (Table III) demonstrate that the automatic method achieves a low mean of differences for both pole length measurements and derived thaw subsidence, when compared to the expert-driven method.

Lastly, we assess the accuracy of direct point cloud distance calculation compared to the expert-driven feature-based approach. The differences of results from the reference dataset are listed in Table III, with thaw subsidence values quantified as median of point-wise distances around each subsidence station.

Thaw subsidence in the study area

Overall thaw subsidence in our study sites in TVC is 1.4 cm from the end of the thawing season in August 2015 to August 2016 according to our expert-driven method. When considering the individual values at the subsidence stations, thaw subsidence variability can be described by the SD of values and amounts to 1.0 cm. Figure 7 visualizes the distribution of individual thaw subsidence values separated into the two plot sites. The shrub-dominated site (2) exhibits larger thaw subsidence (1.7 cm) with higher variability within the site (1.1 cm). The open tundra site (1) exhibits a thaw subsidence of 1.2 cm with a variability of 0.8 cm within the site.

The distribution of individual thaw subsidence values separated into the two plot sites illustrates that thaw subsidence at different locations within the sites is highly variable. Such spatial variation in thaw subsidence is visualized for a subset of site 1 in Figure 8.

Discussion

Expert-driven and in-situ thaw subsidence quantification

The adjustment of ground surface planes strongly improves from using individual (1.1 cm) to aggregated (0.1 cm) expert contributions. Our result indicates that subjective errors caused by individual surveyors’ bias have less of an impact on results when aggregating the individual measurements. Particularly, the orientation of the surface planes substantially improves when aggregating contributions.

When comparing results obtained from in-situ measurements to the expert-driven quantification, average thaw subsidence values correspond better (mean of differences: 0.4 cm) than measured pole lengths (mean of differences: 0.7 cm). This suggests that the ground surface was determined differently in the field, but in such a consistent way that thaw subsidence values correspond better than individual pole length measurements. The high SD of differences for in-situ measurements of pole length (1.7 cm) and also derived thaw subsidence (1.5 cm) indicate that these manual in-situ measurements are less accurate than the expert-driven point cloud-based method as there is a higher variability in the difference values of individual measurement results. This is particularly important as the SD of differences is within the range of the quantified ground surface.
Table III. Statistics of differences of pole lengths (PL) and thaw subsidence (TS) for the two automatic methods compared to the expert-driven method for subsidence stations of both sites (and separated by site 1/site 2). In the feature-based approach, pole lengths are measured based on the automatic plane determination. In the point cloud distance calculation, statistics of differences are derived for the median of point-wise distances around each subsidence station (in a radius of 6 cm)

| Approach                              | Measure | RMS (cm) | Mean (cm) | SD (cm)     | Median (cm) |
|---------------------------------------|---------|----------|-----------|------------|-------------|
| Feature-based, fully automatic         | PL      | 0.5 (0.5/0.4) | 0.1 (0.1/0.1) | 0.4 (0.5/0.4) | 0.1 (0.1/0.1) |
|                                       | TS      | 0.5 (0.5/0.4) | -0.2 (-0.1/-0.2) | 0.5 (0.5/0.4) | -0.2 (-0.1/-0.2) |
| Point cloud distance calculation       | TS      | 1.0 (0.5/1.2) | -0.2 (-0.2/-0.3) | 0.9 (0.5/1.2) | -0.2 (-0.2/-0.1) |

Automatic determination of the ground surface plane

The low mean of differences of the automatic method compared to the expert-driven method (0.1 cm) shows that the automatically adjusted plane closely corresponds to the ground surface plane determined by the expert group. Considering the SD of differences, the variability of difference values of the automatic method is close to the uncertainty of the expert-driven method – to which we compare the resulting values per subsidence station. By repeatedly capturing a site using TLS, the developed automatic feature-based method of thaw subsidence quantification hence allows automatic observation of ground surface displacement. For difficult cases, such as the subsidence station that was excluded in the automatic processing because of high residual distances in the plane adjustment, the automatic method can be combined with expert input. Such a combined approach reduces the human input while maintaining the accuracy achieved by the expert-driven method.

In addition to the advantages of high sample size and objectivity, using TLS offers two other possible advantages over a benchmark-based survey at selected point locations. Firstly, it is a non-invasive measurement, which is particularly important in the thermally sensitive permafrost environment. Though great effort is made to minimize the effect of the benchmark installation on the surrounding ground, any influence on natural processes may only be excluded if an accurate assessment is available independent of the benchmarks.

As the TLS sensor is not placed within the area being measured, potential for disturbance to influence observed change is reduced. Secondly, through high-resolution recording of discrete points the TLS point clouds enable an areal assessment of thaw subsidence.

A further benefit of TLS-based analysis is that the time series built up through repeated capturing of sites can be revisited in the future as new approaches of quantifying thaw subsidence emerge. To help enhance such future analyses, we have made our multitemporal dataset openly accessible in the data library PANGAEA (Anders et al., 2018a).

Thaw subsidence quantification by direct point cloud comparison

The mean of differences of direct point cloud comparison to the reference dataset (-0.2 cm) shows good agreement with the expert-based method, as the value is well within the accuracy of 3D point measurements in the TLS datasets. The high SD of differences in site 2 (1.2 cm) is mainly caused by the high error of the point cloud comparison at one subsidence station with differences from the reference of 4.3 and 3.2 cm for the two periods. The same subsidence station pole is excluded from the automatic feature-based quantification due to the high plane fit error. While point cloud distance computation performs well at locations where subsidence stations are easily visible, it performs poorly in localized areas of dense shrub cover. This also explains the overall difference in SD values of quantified thaw subsidence between sites 1 and 2. The residual point distances to a locally fit plane could hence provide a useful estimate for the confidence of point-wise distance computations in the direct point cloud comparison. Such a measure of assessing the confidence level of change quantification based on point distances in local plane fits is implemented in the M3C2 algorithm.

![Figure 7](https://example.com/figure7.png) Distribution of thaw subsidence values quantified at the subsidence stations with the expert-driven method using multitemporal TLS point clouds separated by site. [Colour figure can be viewed at wileyonlinelibrary.com]
(Lague et al., 2013). This can be integrated in future thaw subsidence quantification by direct point cloud comparison to identify problematic areas that need human surveyor input or must be excluded from the analysis.

An area-based assessment of thaw subsidence is therefore feasible through the direct comparison of multitemporal TLS point clouds. However, ground surface cover should be taken into account for assessing spatial variations in the accuracy of quantification. The feasible application of areal thaw subsidence quantification in open tundra allows us to analyse ground surface displacement at positions further away from the installed benchmark poles. Spatial patterns of thaw subsidence can now be linked to local characteristics at smaller horizontal scales than benchmarks could previously measure at sparse locations within a site. Spatially variable thaw subsidence likely occurs depending on local surface characteristics such as vegetation and/or microtopographic properties (e.g. Overduin and Kane, 2006), and subsurface differences in ground ice. These are typically highly variable across small horizontal distances in Arctic tundra landscapes. So far, such linkages are not understood comprehensively. Our approach of combining subsidence stations with multitemporal TLS point clouds therefore offers a method for high-resolution areal thaw subsidence quantification at scales of millimetres to a few centimetres.

Conclusions

Our results show that TLS-based methods increase the overall accuracy of thaw subsidence quantification compared to in-situ measurements in a conventional field survey.

The automatic method comparing pole lengths from measurements in individual point cloud scenes is particularly powerful, as it does not require multitemporal point clouds to be aligned to the accuracy required for direct multitemporal point cloud comparison. While our monitoring setup includes fixed benchmark poles, which are always required to substitute for the lack of stable landscape elements, no static observation system is needed, such as a permanently fixed TLS scan position.

Our analysis of extending thaw subsidence quantification to an area-based assessment confirms that thaw subsidence can be quantified based on absolute point cloud distances spatially beyond benchmark pole locations, by using fixed benchmarks as control points for multitemporal point cloud alignment. This provides the novel opportunity to investigate the permafrost system using spatially distributed high-accuracy observations, which are independent from the potential influence of ground disturbance caused during the installation of benchmark poles and in-situ measurements.

The web-based 3D geoinformation extraction tool provides an easy-to-use framework for collecting expert measurements. Subjective errors are reduced by using the judgement of multiple contributors. This is an important advantage compared to in-situ measurements at the remote sites, which are typically conducted by one person during each survey. A practical approach for TLS point cloud-based monitoring setups can include a combination of automatic measurements and human expert input in the quantification. Positions in the point cloud scene that exhibit high uncertainty in the automatic thaw subsidence quantification could be distributed to expert analysts who submit their contribution online without being present at a specific location and at particular points in time. In this way, the expert-driven approach can enhance the link between automatic 3D surface change quantification and in-situ judgement of surveyors at permafrost monitoring sites.

Generally, the point cloud-based methods are more transparent than conventional in-situ methods, as each measurement – as well as the site conditions at the time of the survey as represented in the 3D point cloud – can be revisited. Compared to in-situ surveying methods using single measurements at benchmark locations, TLS point clouds thereby enable large improvements in thaw subsidence quantification.

Acknowledgements—This work was core funded by the Federal Ministry of Economics and Technology (BMWi) and the German Aerospace Centre (DLR), Germany, in the framework of the project PemaSAR (FKZ: 50EE1418).

We appreciate the support from the Wilfrid Laurier University Cold Regions Research Centre (CRRC) to use the Laurier TVC field camp. We thank everyone who helped in the field campaigns.

Katharina Anders was supported in part by the Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS MathComp), founded by DFG grant GSC 220 in the German Universities Excellence Initiative.

The implementation of the web-based tool for 3D geoinformation extraction benefited from developments in the frame of the 3D-MAPP project funded by the Vector Foundation (Project ID: 2015-051).

We thank three anonymous reviewers for their comments, which helped to improve the manuscript.
Data Availability Statement
The terrestrial laser scanning point clouds and in-situ measurements used in this study are openly available in PANGAEA at https://doi.org/10.1594/PANGAEA.888566.

Conflict of Interest Statement
The authors declare that they have no conflict of interest.

References
Allison I, Carrasco J, Kaser G, Kwok R, Mote P, Murray T, Paul F, Ren J, The authors declare that they have no conflict of interest. 56
Conflict of Interest Statement
PANGAEA at https://doi.org/10.1594/PANGAEA.888566.

in-situ
Data Availability Statement
Allison I, Carrasco J, Kaser G, Kwok R, Mote P, Murray T, Paul F, Ren J, The authors declare that they have no conflict of interest. 56
Conflict of Interest Statement
PANGAEA at https://doi.org/10.1594/PANGAEA.888566.

in-situ
Data Availability Statement
Allison I, Carrasco J, Kaser G, Kwok R, Mote P, Murray T, Paul F, Ren J, The authors declare that they have no conflict of interest. 56
Conflict of Interest Statement
PANGAEA at https://doi.org/10.1594/PANGAEA.888566.
Riegler LMS. 2017. Datasheet: RIEGL VZ-400. Available at www.riegl.com/uploads/tx_pxpriegldownloads/10_DataSheet_VZ-400_2017-06-14.pdf (accessed 3 September 2018).

Rusu RB, Cousins S. 2011. 3D is here: Point Cloud Library (PCL). IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China.

Shiklomanov NI, Streletskiy DA, Little JD, Nelson FE. 2013. Isotropic thaw subsidence in undisturbed permafrost landscapes. Geophysical Research Letters 40(24): 6356–6361.

Shur Y, Hinkel KM, Nelson FE. 2005. The transient layer: implications for geocryology and climate-change science. Permafrost and Periglacial Processes 16(1): 5–17.

Shur Y, Jorgenson MT, Kanevskiy MZ. 2011. Permafrost. In Encyclopedia of Snow, Ice and Glaciers, Singh VP, Singh P, Haritashya UK (eds). Springer: Dordrecht; 841–848.

Streletskiy DA, Shiklomanov NI, Little JD, Nelson FE, Brown J, Nyland KE, Klene AE. 2017. Thaw subsidence in undisturbed tundra landscapes, Barrow, Alaska, 1962–2015. Permafrost and Periglacial Processes 28(3): 566–572.

van der Sluijs J, Kokelj SV, Fraser RH, Tunnicliffe J, Lacelle D. 2018. Permafrost terrain dynamics and infrastructure impacts revealed by UAV photogrammetry and thermal imaging. Remote Sensing 10(11): 1–30. https://doi.org/10.3390/rs10111734.

Wujanz D, Avian M, Krueger D, Neitzel F. 2018. Identification of stable areas in unregistered laser scans for automated geomorphometric monitoring. Earth Surface Dynamics 6(2): 303–317.