Referee Comments 1

Dear reviewer, we would like to thank you for your helpful suggestions and constructive comments on our manuscript. They will help us to further improve and develop our work. We provide a response to your major points below in italics.

General comments:

This paper addresses an important and evolving topic of using terrestrial laser scanning for studying Arctic permafrost. However, I have several questions and suggestions which need to be addressed. I am detailing them below. Nevertheless, the intent and objective behind this work presently are unclear to me and need to be addressed properly. If this work is about methodology development, then I must say that there is very less novelty involved and the advancements in the preexisting approaches are not significant. If this work is about reporting the state of changing permafrost in the Arctic then that part is feebly touched upon, understandably because the changes over a period of just one year would not be as drastic as to definitely comment on the state of the permafrost. It needs more observations. Presently, I feel that although this is a good initiative, this is also a premature reporting and the authors can make a better study by gathering data from more sampling sites and for different seasons over duration of 2-3 years to report on the seasonal dynamics. The highlights of this research need to be emphasized in the present version.

Our manuscript presents a methodological study which evaluates TLS for quantifying small-scale thaw subsidence. The scientific contributions are that, firstly, we assess the applicability of current state-of-the-art approaches for TLS-based subsidence monitoring, which is challenging in Arctic tundra-ecosystems due to a typically dense moss-lichen layer and micro-topographic characteristics. We show, for example, that standard DEM differencing in Arctic tundra-ecosystems is error-prone due to spatial sampling effects. Secondly, we introduce a new point-based filter strategy to overcome the described spatial sampling and signal occlusion effects. The presented method identifies TLS ground points suitable for multi-temporal deformation analyses and allows to deliver highly accurate ground-truth data for small-scale subsidence. Finally, recommendations for TLS subsidence monitoring are given.

Specific comments:

1. Entire paper needs a thorough language editing in the sentence structures and presentation styles. There is a lot of redundancy that can be avoided.

   Abstract: It needs modification, particularly in the conclusive lines. The relevance of the research has been mentioned through the four opening lines. It can be shortened. 2-3 more sentences can be added elaborating more on the results, particularly mentioning some quantitative assessments. The end is abrupt and it can be modified by adding a conclusive line that highlights the contribution of the study in filling the research gaps and the future prospects.

2. Introduction:

   P1 L23: Add reference for the first line of the introduction.

   Zhang et al., 2008 and van Everdingen, 2005 are the references for this line:

Zhang, T., Barry, R. G., Knowles, K., Heginbottom, J. A., and Brown, J.: Statistics and characteristics of permafrost and ground-ice distribution in the Northern Hemisphere, Polar Geography, 31, 47–68, doi:10.1080/10889370802175895, 2008.

van Everdingen, R. O.: Multi-language glossary of permafrost and related ground-ice terms: http://globalcryospherewatch.org/reference/glossary.php.

P2 L3: "...allow a detailed... " Replace "detailed" by "spatially continuous".

P2 L25-27: "Such ALS...cm". The resolutions have improved in recent years, particularly with the use of UAV-mounted ALS. So this sentence needs improvement.

New technologies like Unmanned Laser Scanning (ULS) have been recently introduced, showing potential for mapping topography and vegetation with higher point density and accuracy (e.g. Wieser et al., 2016). Nevertheless, the registration accuracy of kinematic acquisition systems is still lower (up to cm level) compared to static data acquisition (up to mm level). We therefore rely on TLS for monitoring small-scale subsidence.

3. Study area:

P4 L10-11: Is it the annual average temperature or an average from 1971-2000?

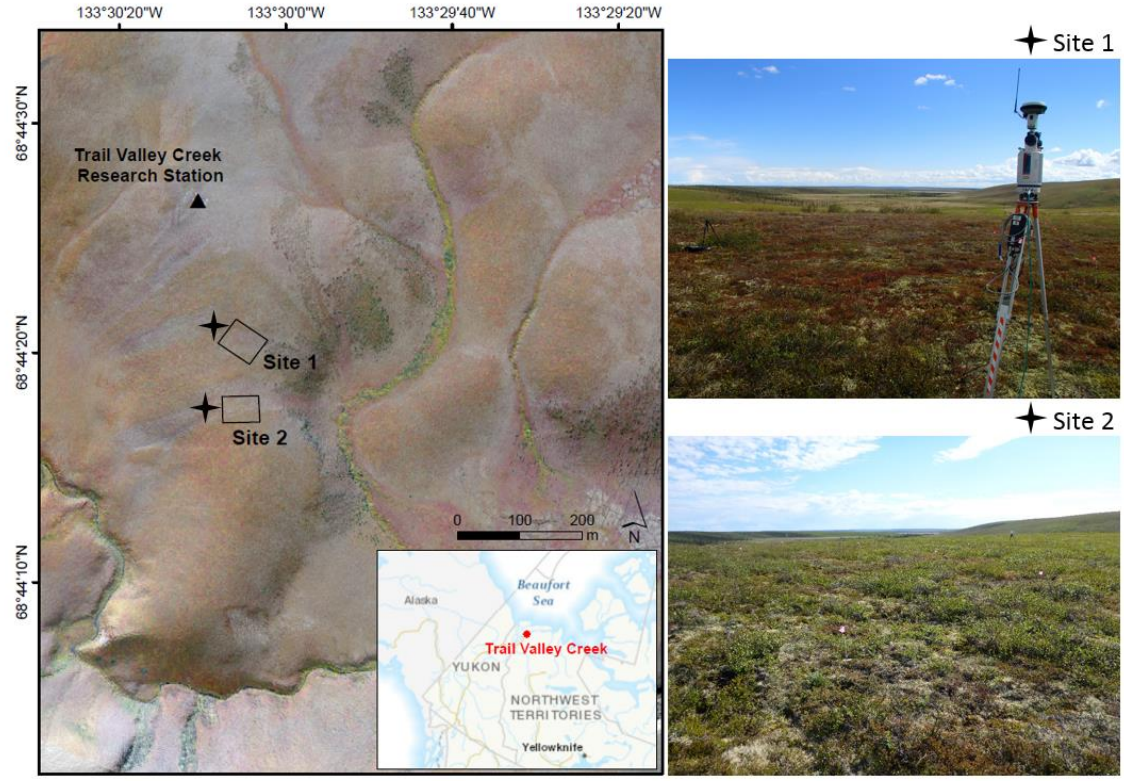
It is the mean annual air temperature (MAAT) between 1971 and 2000.

P4 L12-13: How far are the 2 sites from each other? In Figure 1, they appear to be ∼100 m apart. As mentioned "Site 1 is about 50x40 m, almost flat and covered by low tundra vegetation and Site 2 is equal in size but contains more shrubs", I was wondering that the sites are so close that they must be having overlapping areas (going by the 50X40 m dimensions and considering that the dots in Fig.1 represent the middle point of the plots). What was the need to keep the sites so close? Is there really any difference in the vegetation as it seems similar for both the sites in Figure 1? I would suggest displaying close-range photos to establish it. These photos should not be only for a single time period for each of the sites but must be consisting of repeat images of the same sites during various surveys so that the reader can visualize the changes in the vegetation.

The modified version of Figure 1 (see below) shows the actual ALS extents: The sites do not have an overlapping area. We agree that it is not easy to visually recognize differences in the vegetation based on the provided photo: Regarding the photo of site 1: The shrubby area next to the scanner is not part of the study site. We replaced this photo to provide a better impression how site 1 looks like. Furthermore, the photographer's position was marked in the map. Now it is visible that site 1 contains considerably less shrubs compared to site 2.

The legends from fig. 1 are missing. What do the colors in the map represent?

The orthophoto used as background map is a RGB image captured in August 2015.



I do not like the fact that although the paper is about a permafrost region, there is not enough description of that in the study area section. Mere citation of a few articles is not enough. The authors must talk about the MAAT and monthly temperatures through some graphs using the data from the nearest monitoring station for past several years to show that the region can still be considered as of having a continuous permafrost. Several close range photos of relevant surface features corroborating the permafrost (for e.g., palsa or hummocks) in the study area could have been added here. This section needs improvement and justifications.

We had a look at the meteorological data from a weather station at TVC (distance to the study sites: < 1km). The MAAT measured at TVC was -7 °C from 2013 to 2016. The regional permafrost is estimated to reach depths between 100 m and 150 m (Marsh et al. 2008) and maximum active layer depths ranging from 0.4 to 0.8 m (Qinton and Marsh 1999).

Marsh, P., Pomeroy, J., Pohl, S., Quinton, W., Onclin, C., Russell, M., Neumann, N., Pietroniro, A., Davison, B., & McCartney, S. (2008). Snowmelt Processes and Runoff at the Arctic Treeline: Ten Years of MAGS Research. In M.-k. Woo (Ed.), Cold Region Atmospheric and Hydrologic Studies: The Mackenzie GEWEX Experience (Vol. 2: Hydrologic Processes). Berlin, Heidelberg: Springer. pp. 97-123. doi: 10.1007/978-3-540-75136-6_6.

Quinton, W. L. & Marsh, P. (1999). A Conceptual Framework for Runoff Generation in a Permafrost Environment. Hydrological Processes, 13, pp. 2563-2581.

P4, L24-27: "Additionally, ... environment." This is an extremely important step for repeat surveys. I would have liked to see the pictures of the installed rods. Were these rods marked? How exactly did they serve the purpose of common reference points during the repeat surveys? Did the authors also check for the change in the inclination of the rods during repeat surveys? These points need detailing because it's just a matter of mm scale accuracy and any error in the methodology can compromise the entire results. Presently, I cannot comment on the accuracy standards unless I get the information on this step.

We installed the rods for two purposes. Firstly, to obtain reference data for the TLS-based subsidence rates. Secondly, we identified the top of each rod in the TLS point clouds and used those coordinates as fix points to co-register the TLS datasets. We did check for potential change in the inclination of the rods during repeat surveys by calculating the 2D-distances between all extracted coordinates. The 2D-distance is constant for all three survey dates (difference between 2D-distances < 1mm) meaning that the rods are stable.

4. Methods:

   P6, L9: Full form of OPALS?

   Orientation and Processing of Airborne Laser Scanning

   There is no description of interpolation and DEM raster generation algorithms. This cannot be avoided.

   The DEM raster generation is explained in section 3.4.1 DEM differencing: manuscript "DEMs are generated based on the lowest z-value within each raster cell […] different TLS raster cell sizes (from 1 cm up to 500 cm) are evaluated."

5. Results:

   P8, L19: NMAD and RMSE are quite high for a TLS-based DEM! Figure 4 represents the poor accuracy of the DEMs.

   In this section (4.1 Raster-based deformation analysis) we evaluate raster-based DEM differencing as one approach to derive subsidence rates. Yes, we found that the DEMs and the resulting DEMs difference maps are of poor accuracy. Therefore, we state "All in all, this reveals a significant limitation of TLS DEM differencing for detecting spatial patterns of small-scale subsidence" (P9, L8). Our conclusion is not to rely on raster-based approaches but to apply our proposed filter strategy and then to calculate changes directly in the point cloud e.g. using the M3C2 distance calculation algorithm.

   P8, L23: "change rate" is actually absolute change. How can it be called "rate"?

   Thanks for the hint. Correct is "The average TLS-based vertical change for site 1 was recorded at approximately -2 cm …".

6. Conclusions:

   P11, L20-21: Why is the seasonal subsidence (just 2 months gap) more than the yearly subsidence? If seasonal subsidence is in cm then the yearly subsidence should also be at least in cm and not in mm. The conclusion seems to be consisting of several general points such as multi-point scanning for increasing the accuracy. However, what is seriously lacking is a clear-cut advancement provided by the present research and precise future prospects for studying the Arctic permafrost. The levels of accuracy achieved during the entire field data collection (table 1, appendix 1) does not seem sufficient to me for commenting on the validity of the results

In Table 1 we present RMSE and mean error at five control points per study sites. RMSE and mean error (z-distance) after co-registration are of relevance to assess uncertainties for the TLS-derived vertical change maps (as suggested by Orem and Pelletier, 2015). As the co-registration error in z-direction is between 0.2–0.3 cm, our results support the observation that the TLS data are sufficient to derive small-scale subsidence maps. Appendix 1 shows the registration error between the single scan positions for each survey date which is of minor relevance to assess uncertainties of the derived results.