Dear Tobias,

We would like to thank both you and the reviewer for your suggestions to further improve the clarity and readability of this manuscript. We have addressed each point in turn below and in particular have substantially reduced the length of the paper. We hope the changes we have made here meet your requirements for publication.

Lindsey and co-authors

Comments from the editor

I feel that the paper can be shortened by ~20% without loss of relevant information.

The text has been reworked to reduce its overall length (excluding abstract) from 9020 words to 7693 words.

Abstract/General content:

You nicely write about the general importance of the study, e.g. with respect to understanding glacier changes and water resources. However, in the abstract no information about the results in this respect are given. Be more quantitative and include at least rough information about the mean measured surface lowering (and an estimate for the entire glacier, if possible).

Here we have added some specific values: *"Between November 2013 and January 2014 penitentes become fewer, wider, deeper, and the distribution of surface slope angles becomes more skewed to steep faces. Surface lowering during this core ablation season was in the order of 0.04m day^{-1}."*

In the absence of supporting data, we prefer not to provide an estimate for the whole glacier on the basis of measurements over <10m^2.

L. 14: You mention a potential error of up to 0.3 m due to insufficient overlap. However, it is unclear from this sentence whether this is a specific problem of your study and how the 0.3 m affects the significance of the result.

This was quoting the single worst alignment to an individual ground control point in our study. This has been clarified and now reads: *"The three-dimensional positional error of alignment between the digital surface and ground control points, were on average 0.08m, but in one case reached 0.3 m, due to poor overlap of individual scanned sections comprising the surface."*

The impact of errors of alignment to the GCPs is included in the calculations of volume change, and in the comparison to the manual measurements.

L. 25: “in line with the roughest values previously published for glacier ice”. This is also unclear.

This now reads: *“The computed roughness values are in the order of 0.01-0.10 m during the early part of the ablation season, increasing to 0.10-0.50 m after the end of December, in line with the largest previously published surface roughness values for glacier ice.”*

Methods:

L. 81ff (Section 2.1). The heading is misleading. You include here also information of measurements, but no information about the general climatic conditions (which I expected here).

The heading has been changed to: “Description of field area and measurement setup” to better accommodate the material we wish to cover here.

Additional information regarding the climatic conditions has been included: *“Interannual climate variability is controlled by the El Niño Southern Oscillation (ENSO), such that during El Niño events, higher precipitation and warmer conditions are experienced (Escobar and Aceituno, 1998). Most precipitation is received during the winter (Vuille and Ammann, 1997), however convective storms can cause small precipitation events in the period from December to March (Schotterer et al., 2003). Although the glacier mass balance in the area is highly sensitive to precipitation, warming at elevation over the last 40 years has produced a rise of the glacier equilibrium line al”titude of over 120 m (Carrasco et al., 2008). Annual mean temperature is below freezing and annual mean relative humidity is below 30% (Ginot et al., 1999). The glacier experiences year-round ablation by sublimation, however, melt is only produced during the summer (Sinclair and MacDonell, 2016).*

L. 85: “study areas”. This term might be misleading. I understand that you had two test sites, correct? Then write so.

Terminology changed to ‘test site’ and ‘glacier site’, and described more explicitly.

L. 116: Include short information about how you corrected for atmospheric conditions and where you got the information about the conditions (probably from the AWS but not clear from the sentence).

Corrections were made using data from a weather station in the glacier forefield, and this is now stated: *“Resulting point clouds were corrected for atmospheric pressure, temperature and humidity using data from a weather station in the glacier forefield, and then trimmed using ILRIS Parser software,…”*

L. 122: Unclear here why the measurements could not be done on the same day.

This was due to a logistical error, and subsequently limited availability of the TLS and Kinect at the same time (TLS *en route* to Antarctica), and is now stated as such.

Section 2.7.This section is hard to read and to understand. It is also too long. Write briefly about the experience from previous studies and then provide a rational for your approach based on these experiences and describe then clearly what you did. It might also be worth to think to move some info about previous studies into the introduction.

The section has been significantly shortened. Superfluous material has been removed, and relevant information moved to the introduction and discussion sections.

L. 270/271: Here you describe the test site and where TLS scans were taken. These sentences fit, hence, better into the methods sections.

The two sentences have been removed as the information was already presented in the method section.

Methods:

L. 284: The differences are not in figure 2b but in the data. Correct the sentence. What do you mean with "holes"? "Data gaps"?

This now reads: *"The difference between the two aligned meshes where overlapping data existed was always < 2 cm (Fig 2b), which is well within the uncertainty of the georeferenced TLS surface model. Larger differences of up to 5 cm, evident in Figure 2b, occur only where there are data gaps in one of the surfaces being compared."*

'Holes' replaced with 'data gaps'.

Section 3.2 About meteorological conditions.
This section is a bit misplaced here. The knowledge about the meteorological conditions and the methods of measurements should be presented earlier and not in the results section.

This section has been removed. Table 2 is referred to in the methods section instead, and further salient points are moved to the discussion.

Additional information regarding the general climatic conditions is presented in the field site description.

Section 3.3 "Aerial scans of penitents surfaces"
The heading fits not well as the first part of this section presents important information about surface lowering and mass loss. This is important and more information could be provided. Tell about the mass loss of the entire glacier based on these measurements here.

This heading has been changed to: *"Morphometric changes and surface lowering"*

L. 313: The sentence and the approach is not really clear to me. How did you calculate the volume change and the surface lowering? This should be described in the methods section and not the results section.

These methods are described in the methods section: *"As the surfaces contain overhanging parts, DSM differencing cannot be performed by simple subtraction. Instead surface lowering was calculated in two ways: Firstly by differencing area-weighted mean surface elevations, and secondly by computing the volume change between scan dates. For the latter approach, volumes for all surfaces were computed relative to a baselevel horizontal reference. Volumes relative to this horizontal reference for upward-facing triangles were computed column-wise, by projecting the area of each triangular face onto the reference surface and using the height coordinate of the triangle centroid as the height dimension for each column. These were summed and volumes for overhanging triangles, calculated in the same way, were subtracted to derive the total volume between the reference surface and each scanned penitente surface. Successive volumes were then subtracted to obtain the volume change over each measurement interval."*
We now refer to area-weighted mean surface height rather than hypsometric mean surface height in the hope that this is clearer to most readers.

L. 322 and elsewhere (e.g. but not limited to L. 345/346): Use uncertainty instead of error. Error refers to the deviation of the truth.

This has been done throughout the text.

L. 343: "compare well". Be more precise.

Now reads: *"This differs from the value calculated from volume change computed from surface meshes consisting of over 1.3 million points and covering an area of 7 m^2 by only 28 kg m^{-2}, which is within the uncertainty of the two measurement methods."*

L. 364ff and L. 995ff: These lines fit better in the discussion and could shortly be further elaborated.

These lines now open the discussion.

Discussion:

General: The discussion is quite lengthy, hard to follow and should more precise. Shorten and link the sections better to each other. Section 4.1 and 4.3: The surface roughness depends on the morphology. They can maybe be combined.

The discussion has been considerable shortened and re-ordered. However, we kept the morphology and roughness sections separate.

L. 402: "broadly meet" be more precise.

We now say 'are similar to'

L. 408/410: shortwave “radiation”.

Corrected

L. 420: Was the modelling done by Lhermitte et al. 2014? This study seems to be quite relevant as it addresses the same glacier. Present this study in the Introduction.

This paper was already referenced in the introduction, but we now highlight its findings as follows: *“Previous studies, based on radiative modelling within idealized penitente surfaces, have investigated the impact of penitentes on the shortwave radiative balance (Corripio and Purves, 2005; Cathles et al., 2014; Lhermitte et al., 2014). The results suggest that penitentes reduce effective albedo by up to 40% compared to flat surfaces and that both shape and penitente size impact the apparent albedo as measured by ground and satellite sensors (Lhermitte, et el., 2014).”*

L. 453ff: These are important results and do not fit in a discussion about methods. I suggest moving to the results section and discuss the implication in the discussion section.

Now in results section 3.3

L. 458: How can you state that these measurements would underestimate the mass loss of the entire glacier? Are mass balance measurements existing?

Accurate mass balance measurements are not available for the period of the study. We therefore remove reference to the glacier-wide mass balance and simply state: *“Assuming that this difference holds true for the whole ablation season of 120 days, point measurements underestimate the seasonal mass loss obtained from the Kinect digital surface models by 86 kg m^{-2}.”*

L. 488: What kind of glaciers do you refer to?

The highest values are for rough ice in the ablation zone and this now reads: *“These values are in line with values previously published for rough glacier ice (Smeets et al., 1999; Obleitner, 2000).”*

Conclusions:

The conclusions should be more specific. Where should and could the presented methods be used?

We added some specifics regarding potentially useful applications: *“This study demonstrates that the Microsoft Kinect sensor be used successfully at close range over rough snow and ice surfaces under low light conditions, to generate small-scale digital surface models useful for assessing morphometry and surface roughness properties of complex terrain, as well as detailed assessments of spatial variability of*

surface ablation. The data collected in this study offers the first detailed study of how the geometry of penitentes evolve through time, highlighting the rate of change of surface properties over an ablation season that can serve as a guideline for parameterizing surface properties required for energy and mass balance modelling of penitente surfaces. The method demonstrated here could be useful for investigating glacier surface features such as sastrugi, crevasses or meltwater streams and determining the patterns of surface change associated with such features."

L. 534f: Here you mention the first time a "number of potentially superior alternative set-up … and software is now available". This is important information. These alternatives should be shortly mentioned and discussed in the discussion section. Maybe some of your problems can be overcome with other alternatives.

Removed this from the conclusions and instead included the following in the discussion: *"The practical utility of the Kinect on glacier surfaces is limited to small study areas, but integrating local findings with glacier wide TLS or photogrammetric information of surface conditions may offer a means to usefully extrapolate small scale findings to the glacier scale. Surface scanning technology and software is an area of rapid development, and ongoing development of new sensors and airborne platforms may eliminate the challenges of producing high quality depth maps over larger areas using similar technology to the Kinect."*

Supplement:
The Supplementary files should be in a single pdf and only the ply scripts in one .zip fie.

This has been prepared as requested but note that it seems only one document can be uploaded as supplementary material. Hence the zip file containing the data will be provided on request.

Comments from reviewer

The purpose of this paper has become clear from the first sentence of Abstract "In this study, the first small-scale digital surface models (DSMs) of small natural penitentes on a glacier surface were produced using a Microsoft Xbox Kinect sensor on Tapado Glacier, Chile (30°08'S; 69°55'W)."

Then one of the major results, though sounding negative for the method developer, may be concluded as that the MXK method does not have a distinct advantage for studies of mass balance or changes of glaciers, judging from "surface lowering was comparable to that derived from manual measurements" L20-30 in Abstract, L600-601 in Discussion, and L733-734 in Conclusion.

This has been highlighted as a key finding in the abstract: *"Although these morphological changes cannot be captured by manual point measurements, a key finding is that mean surface lowering of the scanned areas was comparable to that derived from manual measurements of penitente surface height at a minimum*

density of 5 m^{-1} over a 5 m transverse profile, indicating that more limited manual measurements adequately capture the mean lowering of the complex surface."

However, comparisons of the results by MXK with those by TLS and manual measurements are not well explained in the text, especially Section 3.3 and 3.4 are not clear for what the authors intend to state. Also, Figure 5, which shows the result of comparisons, is not well illustrated with very poor captions.

The results and discussion sections comparing the manual and Kinect measurements have been simplified to clarify the comparison of total change over time.

The Figure 5 caption now reads: "Figure 5: Comparison of surface height through time from manual measurements (points) and extracted from the Kinect scans (solid lines ± vertical error) along the horizontal reference (site A, Figure 1). Triangles indicate original snow depth compared to the surface measured on 25/11/13 and solid black triangles indicate locations where snowdepth exceeded the length of the 3 m probe."

If the authors wish to stress much more 'surface properties', that may be surface roughness, presentations of comparisons of surface lowering, meteorological conditions and glaciological issues can be reduced.

The sections have been reduced as suggested by the reviewer, and as outlined in the response to the editor.

Although I recognize that the revised manuscript has been improved much, the paper is still lengthy and not easy to read in some parts. It is recommended to refine the paper further.

The manuscript has been substantially shortened and we hope this now also makes it clearer to read.