The authors would first like to thank Evan Miles for his instructive comments regarding the surface roughness techniques investigated in this study. We hope that the following response addresses all of the comments and believe they significantly strengthen the content of the paper as well.

1. Error assessment:

1a. From Sections 3.1 and 4.2, it is unclear how and when the 'DEM Error' presented in Table 2 is calculated. Specifically, are these values the mean GCP [ground control point] errors in the SFM-derived DEM [digital elevation model], or some other estimate? Agisoft calculates an error estimate, but this methodology is not transparent. Does this analysis occur before or after the x-y plane-fitting? The error assessment is critical as it determines the choice of final DEM resolution and therefore the scale of the analysis.

- The DEM error presented in Table 2 is the error that Agisoft reports after the GCPs (markers) have been identified in all of the images. As defined by Agisoft LLC (2013) this error is the "distance between the input (source) and estimated positions of the marker", which we interpret to be the root mean square error (RMSE) of the modelled and measured GCPs. The x-y plane-fitting was performed after these errors were calculated.

1b. The authors note that the error is dominated by human error, largely due to the choice of a type of cone (visible in Figure 2) which has no point. This certainly makes the total station survey and photogrammetric georeferencing very difficult, which is unfortunate as the dense point clouds from 40+ photos are probably very self-consistent (low internal error), although without a meaningful unit distance. What is the maximum resolution that could be achieved with the photo survey?

- Although the DEM error is dominated by GCP placement, we agree that the relative accuracy (or the internal consistency) of the DEM will be high. Our z_0 calculations should therefore be largely unaffected by the type of GCP marker we used.

  The maximum resolution that we were able to achieve was partly determined by the calculated errors in our data, which ranged from 0.008 m – 0.024 m (Table 2). Given that we wished to keep the DEM resolution on the same order as the errors we chose a pixel spacing of 0.01 m (errors were below this threshold for 3 of the 4 sites). The question of whether the chosen DEM resolution impacts the estimates of z_0 is interesting and is something we quantify in response to comment 2b.

2. Justification for the novel z_0 method and equation:

2a. As presented, the authors use an established (Lettau-Munro) method to calculate topographic roughness, but find that it doesn't agree with their expectations based on literature. They then opt to devise a new method with an arbitrary threshold, which then produces values in-line with their expectations based on literature. Unfortunately, the field instrumentation did not include meteorological equipment to determine the effective aerodynamic roughness for any of their plots, which could have established their method rigorously (that is a problem that has faced prior authors including, e.g., Rees and Arnold, 2006, and my own analyses). Instead, the modified values (which are used to constrain the optimization of z_0) may as well have been selected from reported values. On the other hand, the lower-than-expected roughness values are worth reporting, and should at least demonstrate the same pattern of inter-site variability seen by the modified method. It could have been useful for the authors to also consider the many microtopographic methods in literature to estimate z_0 (several are considered in Smith, 2014, also see Nield, 2013a). Alternatively, simply choosing surface roughness estimates from the literature will not change the bounds of the optimization or sensitivity test, and would provide a more consistent methodological approach, as both albedo and debris thermal conductivity are allowed to vary within literature bounds rather than according to in-situ observations.

- The results of the established Lettau-Munro method, referred to as the standard-deviation approach in this study (although "Lettau-Munro method" will be used from hereon), were reported in the original discussion paper (Section 4.2, p3516, line 20) showing average values of z_0 of 0.0037, 0.0091, 0.0022, and 0.0033 m for Sites A-D, respectively. As suggested, these values do demonstrate the same pattern of inter-site variability that is seen using the modified method. However, as the study discusses these values are towards the lower end of those previously reported in literature and the values appear to be underestimated based on comparison with estimates from other similar field sites.

  The improvement of the modified method is to use real measurements of obstacles over thousands of points and transects, rather than just a few. Unfortunately, meteorological equipment was not available to validate our estimates of z_0, which we agree is a limitation of this study. We feel that the use of other methods (such as those considered in Smith (2014) and Nield et al. (2013a)) is beyond the scope of this paper as the primary focus is the surface energy balance modelling, and we were limited by the lack of meteorological instrumentation and the small size of the plots compared to Nield et al. (2013a). The use of these other metrics and how they compare to the modified Lettau-Munro approach is an important area of future work that we intend to focus on.

  We can see that the selection of the threshold may appear to be arbitrary and thus address it fully in response to comment 2b.

2b. The new z_0 method is not clearly motivated in the manuscript. With a few basic assumptions it is numerically equivalent to the original Lettau (1969) method except that the authors here choose to use a different definition of an obstacle, by 1) initially looking only at profile changes greater than 0.01 m to determine an average obstacle height, and 2) then only considering obstacles larger than this average obstacle height. In other words, the authors are implicitly suggesting that aerodynamic roughness is best predicted by the largest obstacles, where Nield and others (2013b) found a non-linear increase in z_0 for an increase in obstacle size. The idea that a subset of obstacles (whether large or small) dominates the roughness effect is an interesting suggestion that has not been considered much in the literature, which I hope the authors include in their discussion and methodological justification for the z_0 derivation.

- The modified method aimed to i) use the high resolution DEM to derive the obstacle's average height, silhouette area, and specific area as opposed to estimating these parameters using the standard deviation and counting the number of continuous positive crossings, and ii) then apply this method to a high resolution DEM such that the number of samples (both points and transects) were greatly improved. In doing this we quickly realized a major limitation of Lettau (1969) is the lack of a clear definition of what is considered to be an obstacle. The initial approach in our study was to identify all obstacles that were greater than the error in the DEM (0.01 m) and assess the obstacles that were greater than the average of these "potential" obstacles. However, as noted, this causes the small obstacles to be removed thereby assuming that the large obstacles dominate the roughness effect. Figure D1 shows that as the obstacle threshold is increased, the surface roughness linearly increases. Therefore, we agree that the selection of the obstacle threshold or minimum obstacle size is crucial for properly estimating surface roughness.

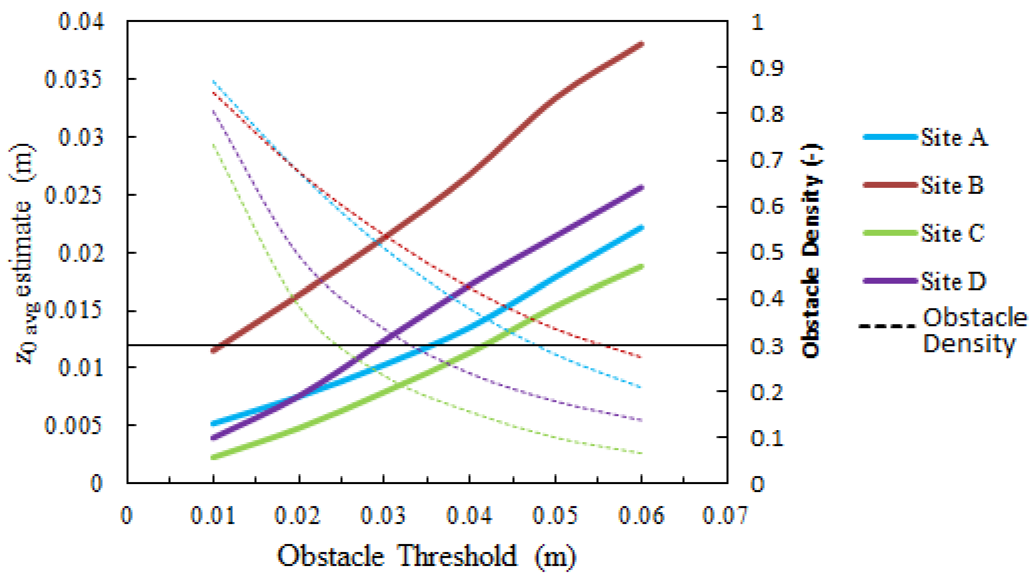

Figure D1. The effect of obstacle threshold (m) on both obstacle density and estimate of surface roughness using the modified method for Sites A-D using a DEM of 0.01 m resolution.

Nield et al. (2013a) found that the height of surface roughness is the best predictor of aerodynamic roughness, specifically for "surfaces with large elements, or [those] that exhibit mixed homogenous patches of large and small roughness elements, maximum height is the best predictor of z_0." This indicates that on a highly heterogeneous surface similar to those encountered at our study site, the larger elements are likely to be the best predictor of aerodynamic roughness. Furthermore, Nield et al. (2013a) found that k_{LM}, which is equivalent to the estimations of z_0 using the Lettau-Munro method, was a good predictor of aerodynamic roughness ($R^2 > 0.79$), even if the two values were not necessarily equal. As discussed in 2a, our study also found that the

Lettau-Munro method yielded the expected patterns of inter-site variability, indicating that they were likely a good predictor of the values, but were not necessarily accurate values of the aerodynamic roughness.

The difficulty of this analysis is to objectively select an obstacle threshold that yields reasonable results of surface roughness. For the purposes of this study, our estimates of surface roughness were compared to those previously reported in literature to assess their credibility. Furthermore, the selection of obstacle threshold should yield similar estimates of surface roughness regardless of the resolution of the DEM. Smith (2014) states that the relationship developed by Lettau (1969) holds at low roughness densities (< 20-30% of the surface area), beyond which the observed z_0 is less than that predicted by Lettau (1969). Above these roughness densities, the obstacles begin to aerodynamically interfere with one another.

Figure D1 shows the relationship between obstacle threshold and obstacle density, where the obstacle density is defined as the cumulative depth of all obstacles above the obstacle threshold divided by the length of the transect. As expected, as the obstacle threshold is increased, the obstacle density decreases. As the modified method is essentially a literal definition of Lettau (1969) using a high resolution DEM, the obstacle density should be less than 30% such that the estimates of z_0 are valid as discussed in Smith (2014). Therefore, the obstacle threshold can be objectively determined as the height at which the obstacle density is less than 30%.

Using this methodology, the obstacle thresholds were found to increase as the DEM resolution became coarser (Table D1). This is not surprising as small obstacles become more difficult to identify as the resolution of the DEM decreases. For the purposes of this study, the coarsest DEM resolution assessed was 0.04 m as this is well below the resolution that can be determined using SfM or acquired from terrestrial laser scanning (Nield et al. (2013a,b), yet still has a sufficient number (~50) of points for the transects. Furthermore, the obstacle thresholds showed the same inter-site variability that is expected, i.e., Site B had the largest threshold and Site C had the smallest. Interestingly, despite variations in obstacle thresholds, the estimates of surface roughness remained relatively constant (less than ±0.004 m). The values of z_0 were 0.017, 0.036, 0.006, and 0.014 m for Sites A-D, respectively. These values are reasonable compared to previous studies as discussed in the paper (Section 4.2) and capture the inter-site variability that is expected. This lends confidence to the use of the 30% obstacle density to determine the obstacle threshold and the modified method used to estimate the surface roughness.

Table D1. Surface roughness, z_0, estimates and obstacle thresholds as a function of DEM resolution (m) at Sites A-D.

Site	DEM Resolution (m)	Obstacle Treshold (m)	z_0 (m)
A	0.01	0.048	0.017
	0.02	0.052	0.016
	0.04	0.054	0.015
B	0.01	0.056	0.036
	0.02	0.073	0.040
	0.04	0.078	0.036
C	0.01	0.025	0.006
	0.02	0.027	0.006
	0.04	0.027	0.006
D	0.01	0.033	0.014
	0.02	0.037	0.015
	0.04	0.040	0.014

2c. One difficulty with this implicit suggestion is the scale-dependency of the method relative to the plot size. Filtering the candidate obstacles as in the modified method removes the small obstacles from influencing z_0, yet the ~2m by 2m plots are also unable to encompass very large boulders present at similar study sites (the authors note a boulder ~1m in diameter, but larger boulders of 3-5m diameter have been commonly observed on similarly debris-covered glaciers (e.g., Hambrey et al. (2008)). Thus, only middle-sized obstacles (over 1 cm based on the obstacle thresholding, and presumably under 50 cm based on the plot size) are considered in the analysis. It would be particularly useful to see if the authors' z_0 estimates change linearly with a different obstacle threshold or DEM resolution.

- The size of the plots (~2 m x 2 m) certainly limits the maximum obstacle size and is a limitation of this study. The use of TLS or an unmanned aerial vehicle (UAV) would be helpful in the future to broaden the size of the plots, while maintaining the high resolution necessary for the analysis. The DEMs generated in this study are of similar resolution to those used by Nield et al. (2013a,b), which appears to show that identifying obstacles at a smaller scale than 0.01 m is difficult and may require an alternative method to those presented in this study. However, for our study site, the surface is dominated by debris and obstacles that are greater than 0.01 m, which supports the use of the methods applied in this study.

  As is noted, there are obstacles that can range up to 5 m in size on the debris-covered glacier. Temperature sensors and ablation stakes were not deployed at these sites for obvious reasons; however, it would be important to consider these obstacles in a distributed model. Nonetheless, determining how z_0 varies across the glacier and at different scales is important for future work. For example, on a kilometer scale, the hummocky terrain of Imja-Lhotse Shar glacier is a potential obstacle of ~50-60 m. If meteorological data is measured such that the aerodynamic roughness can be compared to the surface roughness at various length-scales, then the length-scale that is suitable for debris-covered glaciers could be identified.

As to the variations of z_0 estimates as a function of obstacle threshold and DEM resolution, this has been addressed in response to comment 2b.

3. 'Topographic' vs 'aerodynamic' roughness:

3a. What is clear from literature (Nield et al., 2013a) is that topographic surface roughness estimates do not always match the effective aerodynamic roughness obtained by meteorological instrumentation. This is a difficult problem to solve, because the length-scale of analysis is important for determining aerodynamic processes (and therefore turbulent energy transfer; e.g., Smith (2014)) but this presents a direct conflict between microtopographic methods (which can only be performed for a few-square-meter area while free of instrumentation) and meteorological methods (which reflect the aerodynamic roughness over an unknown area and require instrumentation in place). The microtopographic methods developed to estimate aerodynamic roughness have been primarily developed for surfaces with low permeability relative to the surface of debris-covered glaciers, where airflow into the debris matrix could influence actual aerodynamic roughness, without being accounted for in transect approaches. Consequently, there is room for new microtopography-based approaches such as the method proposed by the authors; such approaches need to be shown to reproduce in-situ observations, however.

- This is an excellent point regarding the differences between aerodynamic roughness and surface roughness and the additional difficulties associated with debris-covered glaciers, which was discussed in 2c.

References:

Agisoft LLC.: Agisoft PhotScan User Manual Professional Edition, Version 1.0.0, accessed June 13 2015. <http://fieldofviewllc.com/wp-content/uploads/bsk-pdf-manager/27_AGISOFT%20PHOTOSCAN%20PRO%20USER%20GUIDE.PDF>, 2013.

Lettau, H.: Note on aerodynamic roughness-parameter estimation on the basis of roughness element description, J. Appl. Meteorol., 8, 828-832, 1969.

Nield, J.M., Chiverrell, R.C., Darby, S.E., Leyland, J., Vircavs, L.H., and Jacobs, B.: Complex spatial feedbacks of tephra redistribution, ice melt and surface roughness modulate ablation on tephra covered glaciers, Earth Surf. Process. Landforms, 38, 95-102, doi:10.1002/esp.3352, 2013.

Nield, J.M., King, J., Wiggs, G.F.S., Leyland, J., Bryant, R.G., Chiverrell, R.C., Darby, S.E., Eckardt, F.D., Thomas, D.S.G., Vircavs, L.H., and Washington, R.: Estimating aerodynamic roughness over complex surface terrain, Journal of Geophysical Research: Atmospheres, 118, 12, 948-12, 961, doi:10.1002/2013JD020632, 2013.

Smith, M.W.: Roughness in the Earth Sciences, Earth-Science Reviews, 136, 202-225, 2014.