

Response to Editor:

The discussion is about the methodology and the results is a bit short and hidden in the results section. Please either include an own discussion section or rename the “results” section to “results and discussion” and expand the discussion at least slightly.

- The editor makes a good point that a lot of the text in the results section regards the justification of certain decisions made, e.g., the selection of the upper bound for the thermal conductivity, the new method for the surface roughness, etc. Therefore, the paper has been restructured such that the results sections strictly focus on reporting the results and the justifications and validity of the decisions made to produce the results are reported in a separate discussion section. A significant amount of content has been removed from the discussion sections regarding the impact of thermal conductivity on model calibration, the assessment of surface roughness results, and the model performance. An entire discussion section has been added as follows:

“6 Discussion

6.1 Thermal Conductivity

One of the limitations with regards to the thermal conductivity measurements is that all the measurements were made near the surface. Therefore, it is possible that the estimates of the average thermal conductivity at each site are potentially underestimated because the deeper layers that may be more compact and humid are not considered. The lack of any trends with respect to depth appears to dispute this theory; however, this is based on a limited number of measurements near the surface. Interestingly, the thermal conductivities measured at Sites 4 and 11 are similar to those estimated by Nicholson and Benn (2012) for debris cover on Ngozumpa glacier with 10% and 20% of the void space being filled with water (1.42 and 1.55 W m⁻¹ K⁻¹, respectively).

The number of sites that reached the upper bound during the model calibration is concerning as it may indicate that the actual thermal conductivity throughout the debris is higher. To address this issue, an additional calibration was performed allowing the thermal conductivity to be unbounded. This calibration revealed that 3 or more out of the 10 sites for each method had thermal conductivities greater than 3.0 W m⁻¹ K⁻¹ with one thermal conductivity as high as 4.5 W m⁻¹ K⁻¹. The lithology of the debris cover in the Everest region is predominantly granite, gneiss, and pelite (Hambrey et al., 2008). Robertson (1988) found the thermal conductivity of solid granite gneiss to be 2.87 W m⁻¹ K⁻¹, so the unbounded thermal conductivities do not appear to make physical sense when one considers that the thermal conductivity of debris should be much lower than solid rock due to the pore spaces being filled with air and water. Furthermore, an optimization performed using the total sum of squares of all the surface sites reveals that increasing the thermal conductivity from 1.6 W m⁻¹ K⁻¹ to its minimum of 2.6 W

$\text{m}^{-1} \text{K}^{-1}$ only reduces the total sum of squares by 3%. These results and the similar values to Nicholson and Benn (2012) lend confidence to the use of $1.62 \text{ W m}^{-1} \text{K}^{-1}$ as the upper bound, but highlights the importance of understanding how the moisture varies within the debris and its influence on the thermal conductivity similar to the findings of Collier et al. (2014). Future work could improve measurements of the thermal conductivity by a) accurately measuring the depth of the temperature sensors during installation and retrieval, b) installing additional sensors (e.g. 5 cm spacing) that allow thermal conductivity within the debris to be computed at more depths, and c) measuring moisture in the debris at various depths.

6.2 Surface Roughness

The development of an alternative method for estimating z_0 was required, as the Lettau-Munro method appeared to greatly underestimate the values of z_0 . The alternative method applies the relationship developed by Lettau (1969) to a high resolution DEM using the selection of an obstacle height (threshold) based on an obstacle density of 30%. One of the main limitations of this study is the lack of aerodynamic roughness measurements to validate the developed methods. Previous work, e.g., Rees and Arnold (2006), has relied upon surface roughness estimates from other studies to assess the reasonableness of their results when aerodynamic data were not collected. This study relies upon the results of Inoue and Yoshida (1980), which estimated surface roughness using wind speed profiles at two sites on the Khumbu glacier. Specifically, Sites B and C in this study have similar debris cover to Areas III and IV from Inoue and Yoshida (1980), respectively.

Site B had the highest value of z_0 (0.043 m) of the four sites in this study and consisted of larger boulders up to 1 m in size. This value is similar to the higher value of 0.060 m for z_0 derived from a region on the Khumbu glacier that consisted of large granitic boulders of 1-2 m in size lying on top of schistose rocks with a grain size varying from a few centimeters to 0.5 m (Inoue and Yoshida, 1980). The larger boulders observed by Inoue and Yoshida (1980) may explain the slightly higher value of z_0 compared to Site B. Site C, which comprised the smallest grain sizes of the four sites in this study, agrees well with the smaller value of z_0 (0.0035 m) derived by Inoue and Yoshida (1980) for an area where the supraglacial debris comprised dispersed boulders ranging in size of 0.01 – 0.05 m. The few boulders ranging in size of up to 0.15 m may be the reason for Site C's slightly larger value of z_0 (0.006 m). Sites A and D were composed of boulders and grains that varied in size between those found in Sites B and C; therefore, we deem the value of z_0 of 0.016 and 0.014 m for Sites A and D, respectively, to be reasonable. Furthermore, these values agree fairly well with the z_0 of 0.016 m measured by Brock et al. (2010) on a debris-covered glacier in Italy that comprised a mixture of granites and schists of predominantly cobble size, with occasional boulders of < 1 m size.

Future work should seek to compare these estimates of surface roughness with aerodynamic roughness to determine the scale at which these two values agree. Brock et al. (2006) found there to be no significant difference between the use of a 3 m and 15 m transect; however, they did state that a shorter pole would be unlikely to capture a sufficient sample of roughness elements if the vertical changes are greater than 1 m. The use of hundreds of transects over a $\sim 4 \text{ m}^2$ grid has the benefit of expanding the number of surface roughness elements that can be captured compared to a single transect. However, Brock et al. (2006) was comparing microtopographic and aerodynamic roughness over snow, slush, and ice, which is significantly different from the hummocky and heterogeneous terrain on debris-covered glaciers. Therefore, it will be important to determine the scale or fetch length at which the surface roughness agrees with the aerodynamic roughness. Nonetheless, the method developed in this paper provides an objective approach to select an obstacle height and yields consistent and reasonable estimates of z_0 for various grain sizes independent of the resolution of the DEM.

6.3 Model Performance

One of the limitations of the calibration procedure is the LE_{Rain} , LE_{RH100} , and LE_{Dry} models all performed reasonably well. The lack of a single model clearly outperforming the others indicates that either a) the modeling of the latent heat flux is insignificant or b) the latent heat flux is significant, but the calibration procedure allows for changes in the latent heat flux to be compensated for via other model parameters. Brock et al. (2010) found that latent heat fluxes may be a significant energy sink when rain falls on warm debris indicating that the latent heat flux is important to include. They also assessed the importance of each component of the energy balance and found that inclusion of the latent heat flux improved the correlation coefficient of their model. The average latent heat flux for both the LE_{Rain} and LE_{RH100} models were comparable with values ranging from -53 to 10 W m^{-2} over the day. The peak instantaneous latent heat fluxes varied greatly between the two models with fluxes as high as -714 and -323 W m^{-2} , for the LE_{Rain} and LE_{RH100} models, respectively. These values are similar to those reported by Brock et al. (2010) and support the importance of including the latent heat flux term. However, they do not yield any insight into preference between the LE_{Rain} or LE_{RH100} models. These results suggest that the selection of the LE_{Rain} or LE_{RH100} model should be based on data availability. Future work should seek to measure the thermal conductivity, albedo, and surface roughness, which would allow the differences between models to be evaluated. Furthermore, detailed knowledge of the debris properties, including how the thermal conductivity and water content vary with depth, would allow the performance of these models to be compared to other debris-covered glacier melt models (Collier et al., 2014; Fujita and Sakai, 2014)."

Supra-glacial lakes and ice cliffs are hot spots of melt at many debris-covered glaciers. This should at least be mentioned in introduction.

- The following has been added to the introduction:

“These glacial lakes and their surrounding bare ice faces also play a crucial role in glacier melt as they typically have ablation rates that are orders of magnitude greater than those observed beneath debris cover (Benn et al., 2012).”

Include at least the most important information about the study region and do not only refer to your own previous study (section 2.1, L. 5/6).

- The following has been added in place of the sentence referring to our previous study:

“Imja-Lhotse Shar Glacier refers to both Imja Glacier and Lhotse Shar glacier, which are avalanche-fed debris-covered glaciers that converge and terminate into Imja Lake. The debris primarily consists of sandy boulder gravel (Hambrey et al., 2008) with the debris thickness increasing towards the terminal moraine. A more detailed description of the glacier may be found in Rounce and McKinney (2014).”

The results section contains also important information about the methodology, e.g. in section 5.3. This information should be moved to the methods section.

- This is a good point. The following section has been moved from Section 5.3 into the methods Section 2.2:

“Ablation rates were modeled over the same time period as the ablation stakes (18 May to 09 November). For days where no meteorological data was available, i.e., the data gaps, the ablation for that day was assumed to be equal to the daily ablation rate for that specific month. As the available meteorological data began on 31 May, the daily ablation rate for the month of May was assumed to be equal to the daily ablation rate of the first week of June. Temperature sensors revealed the debris was snow covered from 26 May to 01 June, so the melting during these days was assumed to be zero. Temperature profiles also show the debris was snow covered from 13-20 October and deeper thermistors revealed the temperature remained around freezing until the sensors were removed in November. Therefore, the melt rates after the 12 October were assumed to be zero.”

You should usually not refer to other studies in the conclusions as you should draw conclusions from your own work. I would rather prefer to include Collier et al. (2014) in the discussion. However, this is also a matter of style and I leave therefore the decision to you.

- The reference has been removed from the conclusions and the following has been added to section 5.1:

“Future work should continue to work on modeling the water content of the debris and its impact on the latent heat flux similar to models developed by Collier et al. (2014).”

P. 5, L.26 (Section 2.2): Regarding the met data: Did you do any kind of quality control with the “unvalidated” data, such as outlier detection or plausibility check? If yes, please include this information; if not you need to include a statement about the quality of the data.

- A limited amount of quality control was performed on the unvalidated data to ensure the data was plausible. The only notable adjustments were negative precipitation values were removed from the dataset. Furthermore, as snow has a large effect on how the model was treated, only snow depths that recorded snow on the ground for the entire half-hour were used. Furthermore, any average snow depth less than 1 mm was assumed to be zero to remove any impact that tiny amounts of snowfall may have on model performance. The quantity of snow was compared to previous years and found to be similar, which supported the limited quality control that was performed. The following has been added to the paper:

“The air temperature, wind speed, relative humidity, and global radiation data were reviewed and deemed plausible, so no adjustments were performed. The half-hour precipitation data was determined by summing the precipitation over each half-hour time step. A few of the minute measurements recorded negative precipitation, which were assumed to be zero as negative precipitation is not feasible. The half-hour snow depth data was processed to assume a snow depth of zero if snow was not recorded on the ground for the entire half-hour. The average snow depth over the half-hour was then computed and any average snow depth less than 0.001 m was considered to be zero.”

Additional revisions:

The standard deviation associated with the surface roughness measurements was not reported in the revised version of the paper. During the computation of the standard deviations it was found that Figure 4, regarding the obstacle density, z_0 estimates, and obstacle thresholds had accidentally been created using the x-axis based on 0.02 m resolution as opposed to 0.01 m resolution. This error was corrected, but it slightly altered that z_0 estimates of Sites A and B from 0.017 m to 0.016 m and 0.036 m to 0.043 m, respectively. Additionally, the model optimization was redone using 0.043 m as the upper bound for z_0 as opposed to 0.036 m, which was used for the previous optimization. This only impacted the optimization of Site 6; however, the changes made to Site 6 altered the average calibration parameters for each model.

The changes in z_0 estimates and the addition of their associated standard deviations were made throughout the entire paper. Furthermore, the updated optimization of Site 6 and changes associated with the average parameters for each model were made to the script. These were very minor changes that did not significantly effect the content of the script and are recorded in the tracked changes document.

Dear Dr. Lindsey Nicholson,

Thank you very much for your insightful and constructive comments. We hope that the following response will address all the comments and we will make appropriate changes in the revised paper to strengthen its content.

Points to be addressed:

1. SfM DEMs: more specific detail on the error assessment of the derived DEMs is required as detailed below in specific comments.

- Please see the response to the specific comments addressed below.

2. Newly proposed method for determining microtopographic z_0 : While I support what was trying to be achieved through this effort, I am not convinced of the value of introducing this new method without validation of it against an aerodynamic roughness length derived from meteorological instrumentation. The relationships of Lettau and others are usually validated against wind profile determinations of z_0 carried out in the field or in a wind tunnel. As stated above, I think the paper would be better and more focused without this section.

- The lack of aerodynamic measurements is a limitation of this as well as other similar studies and is a very important area of future work. Previous work, e.g., Rees and Arnold (2006), has relied upon surface roughness estimates from other studies to assess the reasonableness of their results when aerodynamic data were not collected. In this study, we originally intended to use the Lettau-Munro method, but found that estimations using this method greatly underestimated the values of surface roughness of debris-covered glaciers in the field (Section 4.2, p3516, L18-27). As aerodynamic data were not collected in this study, we rely upon the results of Inoue and Yoshida (1980), which estimated surface roughness using wind speed profiles at two sites on the Khumbu glacier. Based on the description of these two sites, these sites seemed to be similar to the debris cover on Imja-Lhotse Shar. Specifically, Sites B and C in our study have similar debris cover to Areas III and IV from Inoue and Yoshida (1980), respectively. This provides a means of assessing the reasonableness of the methods developed in this paper.

Additionally, the various sites assessed in this study provide a range of debris cover that is typical of debris-covered glaciers in this region. Therefore, it is important that the method developed in this paper captures this inter-site variability (Section 4.2, p3517, L1-24). We also believe that the robustness of the method developed in this paper was greatly improved in response to the discussion comment from Evan Miles. The use of an obstacle density of 30% provides an objective approach for the selection of the obstacle threshold height that is supported scientifically (Smith, 2014). This 30% obstacle density threshold was also shown to hold for various resolutions of the DEM further supporting its use.

Therefore, while we acknowledge its limitations, we believe that the method developed in this study presents a novel approach for estimating the surface roughness that yields reasonable results and is a significant step forward in the current state of knowledge of surface roughness techniques.

Changes to the manuscript: A discussion regarding the lack of aerodynamic roughness measurements has been added to Section 4.1:

“One of the main limitations of this study is the lack of aerodynamic roughness measurements to validate the developed methods. Previous work, e.g., Rees and Arnold (2006), has relied upon surface roughness estimates from other studies to assess the reasonableness of their results when aerodynamic data were not collected. This study relies upon the results of Inoue and Yoshida (1980), which estimated surface roughness using wind speed profiles at two sites on the Khumbu glacier. Specifically, Sites B and C in this study have similar debris cover to Areas III and IV from Inoue and Yoshida (1980), respectively. This provides a means of assessing the reasonableness of the methods developed in this paper. Additionally, the four sites selected in this study provide a range of debris cover that is typical of debris-covered glaciers in this region. Therefore, it is important that the method developed in this paper captures this inter-site variability.

Site B had the highest value of z_0 (0.036 m), which was expected since the debris cover includes larger boulders up to 1 m in size (Figure 2). Furthermore, this value of 0.036 m is similar to the higher value of 0.060 m for z_0 derived from a region on the Khumbu glacier that consisted of large granitic boulders of 1-2 m in size lying on top of schistose rocks with a grain size varying from a few centimeters to 0.5 m (Inoue and Yoshida, 1980). Site C, which comprised the smallest grain sizes of the four sites in this study, agrees well with the smaller value of z_0 (0.0035 m) derived by Inoue and Yoshida (1980) for an area where the supraglacial debris was deposited as dispersed boulders ranging in size of 0.01 – 0.05 m. The few boulders ranging in size of up to 0.15 m may be the reason for Site C’s slightly larger value of z_0 (0.006 m). Sites A and D were composed of boulders and grains that varied in size between those found in Sites B and C; therefore, we deem the value of z_0 of 0.017 and 0.014 m for Sites A and D, respectively, to be reasonable. These values also agree fairly well with the z_0 of 0.016 m measured by Brock et al. (2010) on a debris-covered glacier in Italy that comprised a mixture of granites and schists of predominantly cobble size, with occasional boulders of < 1 m size.”

3. Model calibration: It would be advantageous to additionally perform a multisite optimization to obtain single optimized values for albedo, k , and roughness, rather than a value for each stake. These values could then be applied to all ablation stake sites to give an idea of how useful the model will be when applied to sites for which no specific optimization is available.

- As the reviewer notes, the use of a single optimized parameter set for all sites would be beneficial in estimating the ablation at other sites. However, in order to equally weight all the sites, this set of parameters was derived as the mean value from the model optimization at the 10 sites. The resulting values were a z_0 of 0.014 m, an albedo of 0.32, and a thermal conductivity of $1.5 \text{ W m}^{-1} \text{ K}^{-1}$. These parameters will be applied to the other sites for which an individual optimization was not performed to estimate ablation rates.

Changes to the manuscript: The following has been added to Section 5.3 to explain the use of the average calibrated parameters: “For sites that only had an ablation stake, the average calibrated parameters for that particular latent heat flux model were used. Additionally, a comparison was performed for the LE_{Rain} model using both the average calibrated parameters for all the sites and the calibrated parameters for each individual site (Figure 7).”

4. Model results: (a) Including scatter plots as well as the line plots in Figure 4 could be more helpful for visualizing the prevalence and nature of model biases. As you observe a positive model bias for the nightly minimum temperature I would suggest color coding the scatter plot according to the hour of the day to show the timing of any biases. This approach might also be useful for discriminating more detail about the relative performance of differing methods of modeling LE, by highlighting scatter points for which LE was being modelled in a different color. (b) I became a bit confused as to exactly what data were being used for model validation, so this needs [to be] cleared up and state more explicitly regarding the use of R^2 values between modeled and measured total ablation. I'd also like a small table or explicit listing of values compared the 3 model time resolutions to available stake data.

- Figure R1-1 shows the positive bias of the nightly minimum, i.e., the nightly low was typically overestimated by the model. One possible explanation for this is the use of NCEP/NCAR data for the incoming longwave radiation, which dominates the energy flux at night.

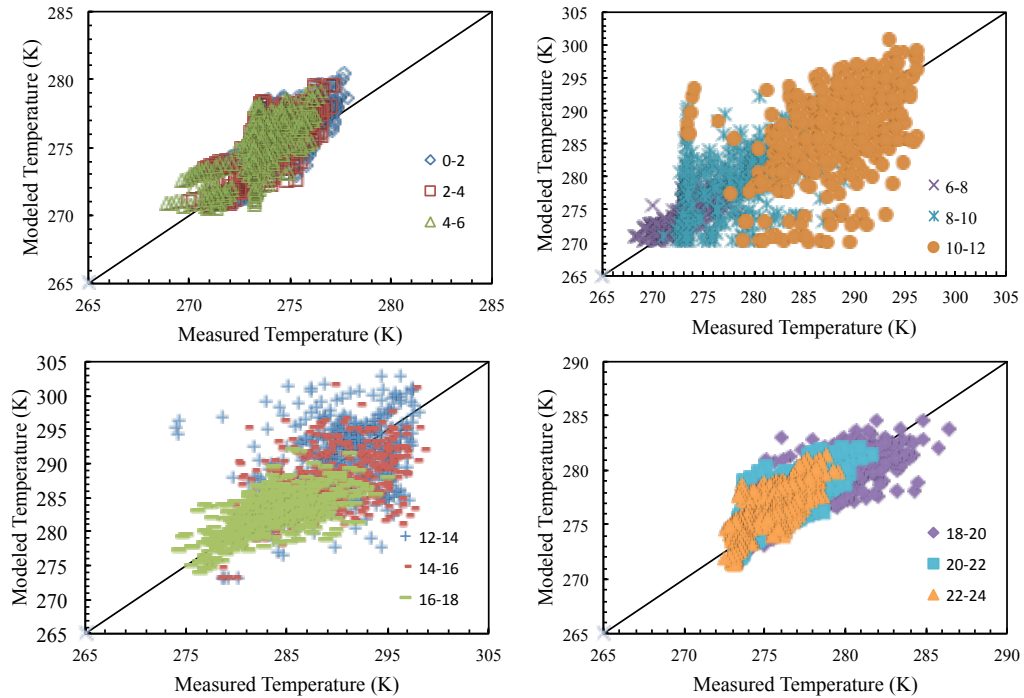


Figure R1-1. Scatterplots of measured and modeled temperature for Site 11 at the surface for the LERain model.

The surface temperature sensors for the months of June and July were used to calibrate the model, while the surface temperature sensors for the months of August, September, and October were used to validate the model. Ideally, two separate years of data would be available such that one year could be used for calibration and the latter for validation, but these were not available.

As requested by the reviewer, Table R1-1 shows a comparison of the total ablation as a function of the different temporal resolutions for the three sites where ablation measurements were made. As expected and detailed in the paper, the 6 hrly time step underestimates the modeled ablation with the 30 min time step. The differences between measured and modeled ablation rates are likely due to site-specific properties and not the temporal resolution.

Table R1-1. Modeled and measured ablation rates (m) using the LERain model for the three sites where the ablation stakes did not completely melt out of the ice.

	Site 8	Site 13	Site 15
Measured	0.92	0.85	0.89
30 min	1.76	0.76	1.22
6 hrly	1.47	0.69	0.98
Daily	1.76	0.74	0.97

Changes to the manuscript: The following figure has been added to the paper to show the nightly bias:

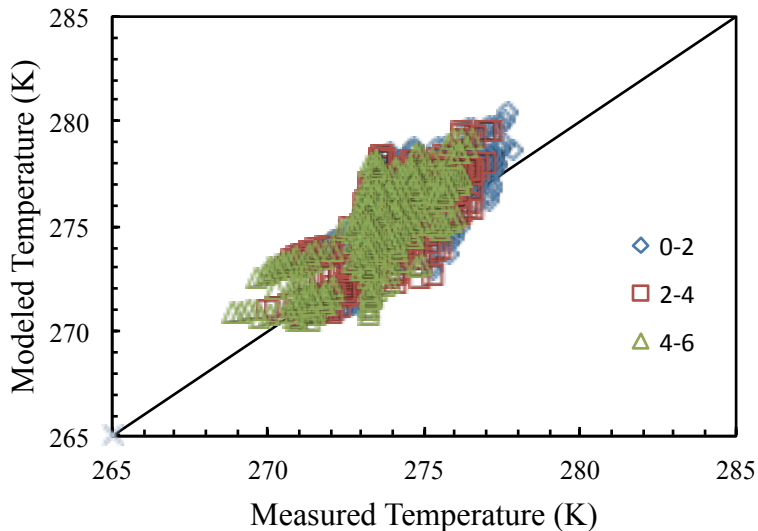


Figure 6. Scatterplot of measured and modeled temperature for Site 11 at the surface for the LE_{Rain} model showing the positive temperature bias overnight.

This has also been accompanied by the following sentence in Section 5.1: “The positive bias of the nightly minimum is apparent between the hours of 0:00 and 6:00 (Figure 6).”

Specific comments:

P3507/L23: Replace ‘laying’ with ‘overlying’

- The manuscript has been changed accordingly.

P3507/L26: What was the reason for the data loss? Failed loggers? Inaccessible? Sensors becoming exposed? Might be useful information for others.

- Large changes in the topography, i.e., some sites on a collapsed slope burying the sensors such that they could not be found, and a couple of the loggers on the sensors failed. For future work, we recommend tying the sensors to a string such that all the sensors remain together. A sentence has been added to the manuscript regarding the reason for data loss as follows: “... as sensors were lost due to large changes in the topography and some of the loggers failed.”

P 3508/L8: Why was site 4 treated differently?

- Site 4 had a debris thickness greater than 1 m, which was the limit of our manual excavation (p3508, line 11). The sentence has been revised to read as follows: “...

with the exception of Site 4, where the debris thickness was greater than 1.0 m and therefore was estimated assuming a linear temperature profile from the mean temperatures over the study period...”

P3508/23: How did you compute k from the temperatures? I assume you used the method of Conway and Rasmussen (2000), but you need to state the method and reference.

- Yes, Conway and Rasmussen (2000) was used and has been added to the manuscript.

P3509/L9: What do you mean by unvalidated here?

- Pyramid Station provided us with raw meteorological data prior to their quality control processing. This detail has been added to the paper as follows: “... unvalidated, i.e., prior to their quality control processing, ...”

P3509/L15: When density did you assume for your snowfall rate to get SWE? Was it a constant value?

- We assumed a density of 150 kg m^{-3} . This detail has been added to the paper as follows: “... to derive a snowfall rate assuming a density of snow of 150 kg m^{-3} .”

P3509/L16: Perhaps it’s useful to add a % of missing data?

- Missing 11.9% of data from May 31 to October 12. This has been added to the paper as follows: “... with a few short gaps (missing 11.9% data).”

P3509/L20: Was this comparison done on a month by month basis or on the average of all 4 months?

- The NCEP/NCAR downward longwave radiation data from 2003 to 2010 between the months of June and September were resampled using a linear interpolation such that the temporal resolution of the NCEP/NCAR data would agree with the hourly meteorological measurements from Pyramid Station. The comparison was performed on all of these data and it was found that the NCEP/NCAR data overestimated the incoming longwave radiation by an average of 29 W m^{-2} . It is important to note that the comparison was not done for any time period where data were missing from Pyramid Station.

Changes to the manuscript: The language has been adjusted as follows to clarify: “... the incoming longwave radiation flux at Pyramid Station from 2003 to 2010 (neglecting any data gaps) between the months of June and September revealed that NCEP/NCAR overestimated the incoming longwave radiation by an average of 29 W m^{-2} (results not shown).”

P3510/L26: How well did the linearly interpolated diurnal LWI cycle represent that measured at the Pyramid Station 2003 – 2010?

- Figure R1-2 and R1-3 show that the linearly interpolated NCEP/NCAR values capture the diurnal cycle of incoming longwave radiation, but consistently overestimates the values at Pyramid Station. Furthermore, R2 shows there is a low correlation ($R^2 = 0.21$) between Pyramid Station and NCEP/NCAR. This poor correlation resulting from the use of reanalysis data as opposed to in-situ meteorological data is likely a major source of error in the model.

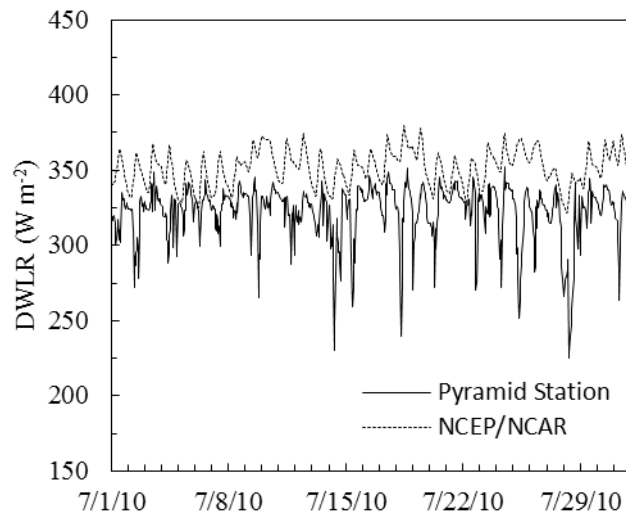


Figure R1-2. Sample time series of Pyramid Station and NCEP/NCAR incoming longwave radiation (W m^{-2}) for July 2010.

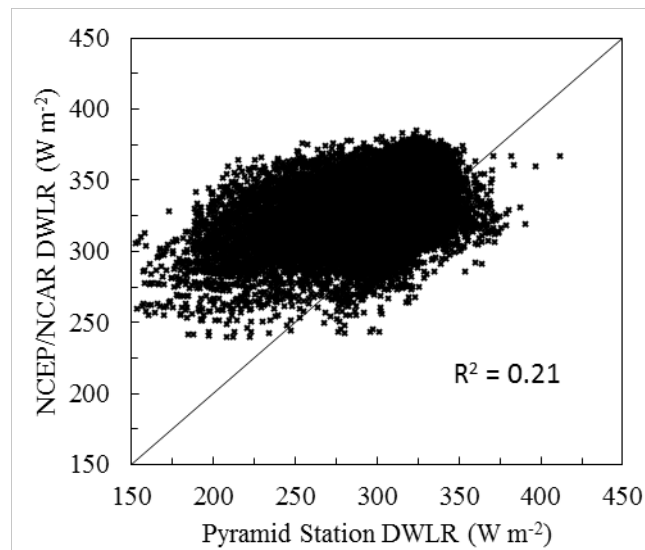


Figure R1-3. Comparison of incoming longwave radiation for Pyramid Station and NCEP/NCAR for all time steps between June and September for the years 2003 – 2010.

P3510/L21: Can you provide information on the accuracy of your GCPs as it affects the resultant scale of the SfM model as far as I understand it.

- The error associated with the total station is very small (0.4 mm) due to the inherent accuracy of the total station and the short distances between the total station and the measurement of the GCPs. For our analysis, the use of cones made it difficult to identify the exact location on the top of the cone in each photo, and this is a source of uncertainty. However, the absolute accuracy of the DEMs is not crucial for our work; it is the relative accuracy that is most important for our roughness calculations, and this will only be affected by GCP placement to a very small extent.

Changes to the manuscript: the sentence has been revised as follows: “Ground control points (GCPs), collected using a total station with an error less than 0.4 mm, are then used to...”

P3510/L12: State here that your GCPs were obtained using a total station with an accuracy better than 1 mm.

- This change was made in response to the comment above.

P3512/L7: I am not clear what you mean by unit width in this part of your method.

- The original relationship between surface roughness and obstacles derived by Lettau (1969) uses information regarding the obstacle’s height, depth, and width (Equation 1). However, the methods used in this paper are based on a transect approach. Therefore, the concept of a unit width (which ultimately cancels out with itself) is used to show the relationship between the original method by Lettau (1969) and the method developed in this paper using a transect approach.

P3513/L5: This is a little unclear to me, but I think you mean that $LE = 0$ unless the RH in the overlying air (at 2m height at the Pyramid AWS) is at 100%, at which point you also set the surface RH to be 100%? Can you express this more precisely in the text please.

- Yes, your understanding is correct. This has been edited in the text “(2) assuming it is dry unless the relative humidity is 100%, at which point the surface relative humidity is assumed to also be 100% based on the assumption that the water vapor above the surface is well mixed.”

P3514/L16: In addition to the reference, please add a sentence describing the nature of the simple snowmelt model used.

- Details of the snowmelt model have been added: “In the event of snow, a simple snowmelt model was used (Fujita and Sakai, 2014), which applies an energy balance over the snow surface that includes net radiation, turbulent heat fluxes, and conductive heat flux with the debris layer in addition to a variable surface albedo of the snow based on the number of days since fresh snow and the air temperature.”

P3515/L11: What was the reason for not computing k at site 14 for the 0.05 m depth?

- Site 14 had sensors at 1 cm, 5 cm, and 24 cm. Since the sensors surrounding the 0.05 m depth were not approximately equidistant, a thermal conductivity was not computed.

P3515/L22: You found previously that k varied with depth, or it was dependent on total debris thickness?

- Rounce and McKinney (2014) found the thermal conductivity of the upper 10 cm to be $0.60 \text{ W m}^{-1} \text{ K}^{-1}$ compared to those below 10 cm, which were $1.20 \text{ W m}^{-1} \text{ K}^{-1}$, i.e., the thermal conductivity varied with depth. No relationships between thermal conductivity and total debris thickness were developed.

P3515/L25: Reiterate here the duration of the measurements used in Rounce and McKinney (2014). Also, was there any observable trend in k over time – that would also indicate temperature dependency?

- The duration of measurements used in Rounce and McKinney (2014) was 13-24 Sept 2013.

Figure R1-4 shows the trends in monthly thermal conductivity (solid lines) for all three sites and the respective depths at which thermal conductivity was measured. Once again, there is no apparent trend between depth and thermal conductivity. On the other hand, there does appear to be a seasonal trend in thermal conductivity such that the thermal conductivity is typically higher in July and August (and June in a few cases) and is smaller in September and October. This appears to closely follow trends in air temperature, where higher air temperatures are observed in June, July, and August. In addition to air temperature, it is likely that rainfall is also contributing to differences in thermal conductivity as moisture in the debris will greatly impact the thermal conductivity. Developing an understanding of the moisture in the debris cover is an important area of future work as it likely will have a great influence on the thermal conductivity in addition to the latent heat fluxes.

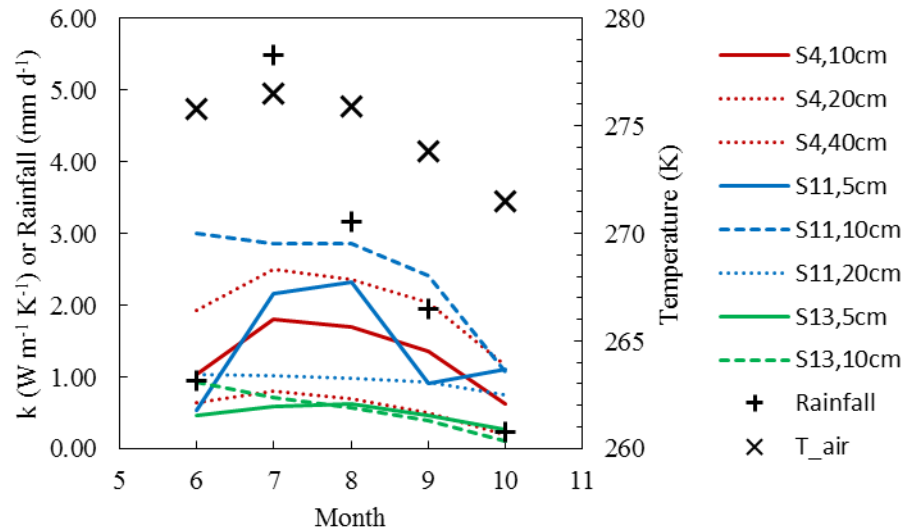


Figure R1-4. Trends in monthly thermal conductivity compared to rainfall and air temperature (T_{air}).

Changes to the manuscript: The duration of Rounce and McKinney (2014) was added to the text as follows: "... it is likely that the temporally-limited data (13-24 Sept 2013) presented in Rounce and McKinney (2014)...".

A sentence regarding the monthly trends associated with thermal conductivity was added to Section 4.1 as follows: "The thermal conductivities did appear to have a trend over the monsoon season where the highest thermal conductivities were typically observed in July and August, which coincides with higher average air temperature and increased precipitation compared to the other months."

P3516/L7: State explicitly how the error is computed – is it the RMSE computed between the total station location of the GCP and the SfM DEM location of all 4 corner markers?

Presumably before you remove the planar slope? Are the values in the table the average or the maximum of these 4 corner marker errors? What is the error on the DEM produced by Agisoft? Did you need to reject any images from the analysis?

- The error for each marker is calculated as the difference between the source (measured) coordinates and the coordinates estimated by PhotoScan Pro and are reported as the root mean square error. In other words, it is the root mean square error between the measured GCPs (all 4 corner markers) from the total station and their estimated position in the SfM DEM. We use this value of root mean square error as the error of the DEM. No images were rejected from the analysis.

Changes to the manuscript: The following sentence has been added to the manuscript: "The error of the DEM was computed as the root mean square error based on the differences between the measured GCPs from the total station and the modeled position of the GCPs from the software."

P3516/L15: Do you really think 4 GCPs per plot is too sparse? Why?

- Given the quality and distribution of the GCPs within our plots, four points are probably sufficient to generate an accurate transformation. We took care to ensure the points were not clustered, which has been shown in previous work to be a source of error (e.g., James and Robson, 2012; Javernick et al., 2014). The internal consistency of the model is therefore likely to be robust. Adding further GCPs would be a trade-off between improving the distribution of the ground control data and introducing additional sources of error into the process. On reflection, therefore, we suggest that a low number of well-surveyed points is sufficient for plot-scale analyses such as this provided they are evenly spaced throughout the plot.

Changes to the manuscript: “as well as the sparse coverage of GCPs in our plots” has been removed from the text.

P3516/L25: How was this value for z_0 determined. Why do you consider it more accurate than Munro’s Method?

- Inoue and Yoshida (1980) estimated z_0 from wind speed profiles of near neutral runs at two sites on the Khumbu glacier. The range of values that we estimated from the Lettau-Munro method was 0.0022 – 0.0091 m compared to those reported on the Khumbu glacier, which ranged from 0.0035 – 0.060 m (Inoue and Yoshida, 1980; Takeuchi et al., 2000). Based on the description of the debris cover from Inoue and Yoshida (1980) and our observations of debris cover on Imja-Lhotse Shar glacier, our Site B seemed comparable to their Area III, which had a larger value of z_0 . Furthermore, the smallest roughness estimation made by Inoue and Yoshida of 0.0035 m was performed on an area of small schist with bare ice, which is further described as the “uppermost part of the ablation zone where debris first appears on the glacier surface. The upper part of this area is the ogive zone which is essentially debris-free, and the lower part is the ice pinnacle zone where supraglacial debris deposits as a dispersed bouldery veneer around ice pinnacles.” One would expect this area of mixed debris and bare ice to have a smaller surface roughness than Sites A, C, and D from our study; however, they all have similar values. These comparisons help show that the Lettau-Munro appears to be underestimating surface roughness estimates. Hence, we believe the modified method developed in this paper is an important step forward as it provides reasonable estimates of debris-covered surface roughness that is based on the original methods developed by Lettau (1969), but utilizes the high accuracy of the SfM DEM. Furthermore, the use of the 30% threshold, which is also rooted in the methods of Lettau (1969), provides an objective approach for selecting the minimum obstacle height.

Changes to the manuscript: The method of measurement has been added to Section 4.2 in the sentence regarding the previous studies as follows: “These values are towards the lower end of those previously reported in literature, which were estimated

from wind speed profiles and range from 0.0035 to 0.060 m (Inoue and Yoshida, 1980; Takeuchi et al., 2000; Brock et al., 2010).”

P3518/L24: Specifically which field data? Just temperature?

- Yes, model calibration was performed only using debris temperatures. This has been revised in the text to read “... how well each method models the measured debris temperatures.”

P3519/L2: Remove this sentence. It is tautological as the optimization must achieve reasonable values for these parameters as you constrain their possible range according to values from the literature.

- This sentence has been removed.

P3519/L11: R^2 between what variables? It seems you performed this on all the available temperature records, correct? Perhaps you'd be better off just doing it for the surface temperatures as (a) you have few measurements at depth and one might be poorly located and (b) other things might be going on within the debris and affecting individual temperature readings at depth are in some way taken into account by using a single optimized k value for each site.

- The R^2 was calculated for all the available temperature records, i.e., those on the surface and at depth. The reviewer makes a strong point that optimizing the model using only the surface temperatures would be consistent across all sites and the lack of knowledge of the moisture and thermal conductivity within the debris make it difficult to effectively model the subsurface temperatures. The subsurface temperatures will be better used for estimating the thermal conductivity. Furthermore, the potentially poorly located sensor should not influence the thermal conductivity calculations as shown by Conway and Rasmussen (2000).

Additionally, Table R1-2 shows that the differences between an optimization performed using all the temperature sensors and only the surface sensors is fairly minimal with the only large discrepancy being Site 13 likely a result of compensating for the poorly located sensor. These results were also reported for the unbounded case to highlight the differences in thermal conductivity, which is the parameter that one would expect to change the most through the incorporation of sensors within the debris. However, these optimized thermal conductivities are once again higher than those previously measured in the field and close to the thermal conductivity of solid granite gneiss (Robertson, 1988). Therefore, with the exception of Site 13, the thermal conductivities would be equal for the bounded condition, which shows there is little difference between including all the sensors or only using the surface sensors. For consistency between the sites and the other reasons discussed above, all the optimizations in the revised paper only use the surface temperatures.

Table R1-2. LE_{Rain} model optimization with k unbounded using all the temperature sensors and only the surface sensors.

Site	LE_{Rain} - Surface Only			LE_{Rain} - All Sensors		
	α	k^1	z_0^2	α	k^1	z_0^2
4	0.13	2.51	0.012	0.13	2.55	0.016
11	0.20	3.40	0.006	0.19	3.47	0.009
13	0.10	0.92	0.025	0.10	1.88	0.020
14	0.31	2.95	0.009	0.37	2.41	0.010
Avg	0.19	2.45	0.013	0.20	2.58	0.014
Std	0.09	1.08	0.008	0.12	0.66	0.005

¹units of $W\ m^{-1}\ K^{-1}$; ²units of m

Changes to manuscript: The calibration was performed using only the surface temperatures as recommended. This has been made clear in Section 3.2: “The calibration was performed by minimizing the total sum of squares of the measured versus modeled surface temperature for each site and was done independently for the three methods used to estimate the latent heat flux.”

And also in Section 5.1: “The albedo, thermal conductivity, and surface roughness for each of the three methods were optimized by minimizing the sum of squares of the surface temperature for each site (Table 4).”

P3519/L13: What physical field evidence leads you to believe the sensor moved down over time? I’m not sure how it could do so? Could it just have been poorly located at the outset? If the sensor at 20 cm depth in site 13 was actually at a greater depth, this would also affect the calculation of k at that site and be a reason for your anomalously low k value for this site.

- Unfortunately, there is no field evidence that led us to believe the sensor moved down over time, since the sensor depths were unable to be re-measured prior to retrieval. It is very possible that the sensor was simply poorly located during installation. Another possibility is the estimate of thermal conductivity was too small. We have removed the sentence.

P3520/L5: Might it be clearer to use an alternate data format given that much of the English speaking world does not use the US month/day convention?

- Yes, the manuscript has been changed to (16-18 June and 25-27 July).

P3520/L4: Consider using ‘... There are a few days for which a positive bias in temperature can be seen during the daily high and nightly low’, as overestimating the ‘low’ might imply modeled temperatures lower than those measured during the nightly low.

- Good point. The manuscript has been changed to read, “there are a few days for which a positive bias in temperature can be seen during the daily high and nightly low (see for example 16-18 June and 25-27 July).”

P3520/L8: Typo – do you mean daily high here?

- Yes, it was a typo. It was supposed to be referring to the nightly low having a positive temperature bias. The manuscript has been changed to read: “one possible explanation for the positive bias in temperature in the nightly low is an overestimation of the incoming longwave radiation...”

P3521/L15: I’d suggest removing this last sentence, as it’s not really necessary.

- Agreed. The sentence has been deleted.

P3522/L1: Why not compute ablation for all 14 ablation stake sites? It might provide a more useful model test, as in reality researchers will likely be applying the model to sites for which optimized inputs are not available. See my point above about a single multi-site optimization.

- The reviewer makes a great point here. The sites that had ablation stakes, but lacked temperature sensors to calibrate the parameters, were assumed to have the mean values of thermal conductivity, albedo, and surface roughness as those from the LE_{Rain} model such that ablation was computed at all the sites that had a temperature sensor or an ablation stake.

Changes in the manuscript: The results are reported in Figure 5 and discussed in further detail in Section 5.3 as follows: “For sites that only had an ablation stake, the average calibrated parameters for that particular latent heat flux model were used.”

P3522/L21: The time periods of the ablation stake measurements and the modeled ablation do not match. Did you run the optimized model for the whole period of the ablation stake measurements to provide a comparison of these data? It becomes clear from later text that you did this, but make it clear here as well.

- The inconsistency between the time periods for the ablation stake measurements and the modeled ablation estimates is an important factor that needs to be addressed. Meteorological data was only available from 31 May to 12 October. The meteorological data was required to run the model; therefore, the model was only run during this time period. The temperature sensors and ablation stakes were installed on 17-18 May, but the first 48 hours of the temperature sensor data was discarded to allow the sensors to equilibrate with the debris. Figure R1-5A shows the debris temperature at various depths at Site 4 throughout the entire duration of the study. Figure R1-5B shows the time series from when the ablation stakes and temperature sensors were installed until meteorological data were available. The temperature profiles indicate that there was snow on top of the debris from 26 May to 1 June.

During this time, we assume that no melting was occurring within the debris. However, from 18 May to 26 May, the debris profiles indicate that melting was occurring. We assume that the daily melt rate over this time period is equivalent to the average daily melt rate that was modeled for the first week of June, i.e., 02 – 09 June. Figure R1-5C shows a heavy snow event, consistent with reports in the field, on 13 October, such that the debris remained snow covered until 20 October. The lower thermistors, located at a depth of 40 and 83 cm in the debris (S4-40 and S4-83, respectively), show that the temperature in the debris remained around freezing until the temperature sensors were removed from the debris on 09 November. Therefore, we assume that no melting occurred after 12 October. This is supported by field observations during the retrieval of sensors where the debris was completely frozen and an ice axe was needed to remove the temperature sensors at select sites. Furthermore, the melt during the transition seasons is much smaller than the melt during July and August, so this assumption should not significantly impact the ablation rates over the entire melt season. Nonetheless, these assumptions are required such that the modeled ablation rates and the measured rates from the ablation stakes are temporally consistent.

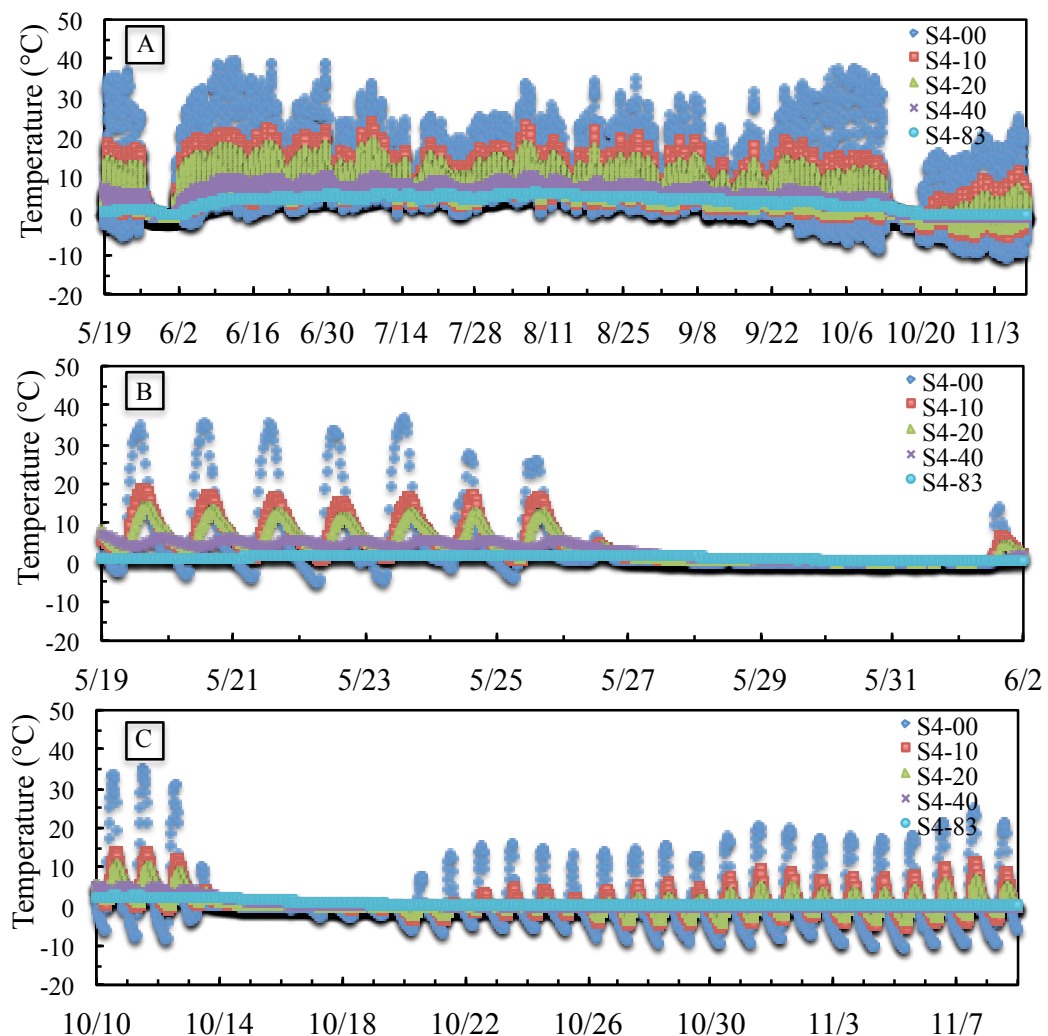


Figure R1-5. Debris temperature measurements from Site 4 for (A) entire study period, (B) prior to modeling time period, and (C) post modeling time period.

Changes in the manuscript: The time period and the assumptions used to have the modeled ablation be consistent with the measured ablation stakes are described in detail in Section 5.3 as follows:

“Ablation rates were computed for all 15 sites that had a temperature sensor or an ablation stake. For sites that only had an ablation stake, the average calibrated parameters for that particular latent heat flux model were used. Additionally, a comparison was performed for the LE_{Rain} model using both the average calibrated parameters for all the sites and the calibrated parameters for each individual site (Figure 7). Ablation rates were modeled over the same time period as the ablation stakes (18 May to 09 November). For days where no meteorological data was available, i.e., the data gaps, the ablation for that day was assumed to be equal to the daily ablation rate for that specific month. As the available meteorological data began on 31 May, the daily ablation rate for the month of May was assumed to be equal to the daily ablation rate of the first week of June. Temperature sensors reveal the debris was snow covered from 26 May to 01 June, so the melting during these days was assumed to be zero. Temperature profiles also show the debris was snow covered from 13-20 October and deeper thermistors reveal the temperature remained around freezing until the sensors were removed in November. Therefore, the melt rates after the 12 October were assumed to be zero.”

P3524/L5: Your sensitivities to k are similar to those found in Nicholson and Benn (2006) for debris cover of ca. 30 cm, but in that work we found that the sensitivity of modeled melt (as a % of melt) to k is dependent on the thickness of the debris cover (see Table 2). Your results show higher k sensitivity at sites with lower ablation, coinciding with thicker debris. I think this is worth mentioning in the context of our results where we explicitly explored the thickness dependency of the sensitivities. Nicholson and Benn (2006) found slightly lower sensitivity to albedo than you, and our sensitivity to z_0 was variable but similar to your findings. It might be worth comparing these explicitly as the model in Nicholson and Benn (2006) uses only daily averages, but the sensitivities as quoted appear to be similar.

- As the reviewer comments, Nicholson and Benn (2006) assessed the sensitivity of modeled melt compared to z_0 and also assessed the uncertainty associated with depth averaged thermal conductivity, k , for both wet and dry debris. Interestingly, the uncertainty associated with the wet debris at Larsbreen (1.67 ± 0.35) was very similar to the uncertainty approximated by our results (± 0.40). The uncertainty associated with the thermal conductivity from our results showed that thicker debris is more sensitive than thinner debris. Nicholson and Benn (2006) show the maximum absolute percent change in melt rate in response to a 1% change in the individual input parameters. For thermal conductivity, the maximum absolute percent change in

melt rate increases with increasing debris thickness, which is consistent with the results of our sensitivity analysis.

Similarly, the sensitivity analysis shows that the sensitivity of the total melt with respect to surface roughness is also more sensitive for thicker debris. This trend was also observed by Nicholson and Benn (2006) for the wet debris, but not for the dry debris. Interestingly, the initial sensitivity analysis was performed by varying z_0 by $\pm 10\%$, which resulted in an average $\pm 1.5\%$ change in total ablation. This result is highly consistent with the sensitivity found by Nicholson and Benn (2006) of 1-2% caused by a 10% change for surface roughness. However, the revised sensitivity analysis (in response to a comment from the other review) shows that the total ablation is quite sensitive to changes in surface roughness, especially for decreases in surface roughness.

Lastly, the thickness of the debris does not appear to have any effect with respect to variations in albedo, which is also consistent with Nicholson and Benn (2006).

Changes in the manuscript:

The discussion regarding the sensitivity analysis has been changed to incorporate these comparison to Nicholson and Benn (2012) and the uncertainties regarding each parameter (in response to a comment from the other reviewer). Section 5.4 has been completely revised as detailed in the response to the other reviewer.

P3525/L8: So here for the temporal resolution you model ablation for all 14 sites, but in the evaluation of your standard model only for 10 sites?

- The discussion paper only compared the 10 sites where ablation was modeled with their respective ablation stake measurements. However, the use of the average optimized parameters allows the other 5 sites to be modeled such that all 14 ablation stake measurements may be compared to the modeled ablation estimates. For the assessment of temporal resolution, the R^2 values of the temperature sensors were used for comparison and the modeled ablation estimates for all 15 sites were also used to compare differences in melt rates between the different time steps.

Changes to the manuscript: The text has been revised to read “The R^2 correlation coefficients for the sites with temperature sensors and the modeled total melt for all 15 sites were used to assess the effect of temporal resolution on model performance.”

P3525/L25: So here you are comparing the total ablation from your stake data – only available at sites 8, 13, and 15? Or are you comparing to your higher temporal resolution model? I am unclear as to which temporal resolution is performing best as compared to the stake data at these 3(2) sites. I say 2 sites because site 13 is clearly complicated, and by optimizing the model on a poorly located temperature sensor at depth it makes a poor test of the model.

- The ablation comparisons are simply assessing modeled total ablation for various temporal resolutions such that only the effects of the temporal resolution are being assessed. The comparison of the modeled ablation and the measured ablation rates reveals that 2 of the 3 sites overestimated ablation, while the other underestimated ablation. A comparison of the modeled ablations for the different temporal resolutions at these three sites was not included in the text, as it does not yield any useful information. The response to the previous comment should clarify what is being compared in the text to assess the effect of temporal resolution.

P3526/L4: In addition to your comment about snow, which is at least partly dependent on the manner in which snow is treated in the model, a daily average model is likely to perform poorly during seasonal transitions even if the surface is not snow-covered as temperature profiles through the surface debris that is either heating up or cooling down through the seasonally transition are often not linear, as shown by the measurements in Nicholson and Benn (2012).

- Good point. The sentence has been adjusted to include this as well: "... the model will perform poorly towards the transition seasons due to changes in the temperature profiles resulting from freeze/thaw conditions and higher amounts of snow fall. Additionally, caution should be taken with respect to ablation estimates as these will likely be slightly underestimated."

References:

- Rees, W.G. and Arnold, N.S.: Scale-dependent roughness of a glacier surface: implications for radar backscatter and aerodynamic roughness modelling, *J. Glaciol.*, 52, 214-222, 2006.
- James, M.R. and Robson, S.: Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application, *Journal of Geophysical Research*, 117, F03017, doi:10.1029/2011JF002289, 2012.
- Javernick, L., Brasington, J., and Caruso, B.: Modeling the topography of shallow braided rivers using Structure-from-Motion photogrammetry, *Geomorphology*, 213, 166-182, 2014.

Dear Reviewer,

Thank you very much for your insightful and constructive comments. We have addressed these below and have revised text in the manuscript accordingly.

Specific Comments

Sections 3.1 and 4.2: z_0 estimation using Structure from Motion

How was the accuracy of the DEM assessed? You analyze variations at the 0.01 m scale in estimation of z_0 ; is that resolution justified? The maximum area of the DEM according to the values given at line 20 on p. 3510, is 4 m², but how much of this area is useable for assessing microtopography?

- The accuracy of the DEM is assessed using the errors estimated by Agisoft in the SfM software. This is the root mean square error based on the differences between the measured GCPs from the total station and the modeled position of the GCPs from the software.

The total RMSE values associated with each DEM are reported in Table 2 and ranged from 0.008 m – 0.024 m. We chose a pixel spacing of 0.01 m such that the resolution of the DEM was on the same order as the DEM errors (3 of the 4 sites had errors less than 0.01 m), which justifies the choice of the resolution.

At each site, the area inside the cones was clipped such that the cones would not influence the surface roughness calculations. The actual area of these clipped plots that were used for analysis were 3.98, 3.23, 5.29, and 3.72 m² for Sites A-D, respectively.

Changes to the manuscript:

“The error of the DEM was computed as the root mean square error based on the differences between the measured GCPs from the total station and the modeled position of the GCPs from the software” has been added to the paragraph on the surface roughness methods (p3510).

“The DEMs were resampled to a resolution of 0.01 m such that their resolution was on the same order as their respective errors (3 of the 4 sites had a total RMSE less than 0.01 m)” has been added to the field results of surface roughness (Section 4.2, p3516).

SfM is not my area of expertise, so I cannot comment on the validity of the methodology, but I like the fact that your approach attempts to generalize the characteristic roughness over the area of the DEM and generate a probable range of z_0 rather than specific z_0 values on individual profiles. Something that is missing from the discussion in Section 4.2, however, is an appreciation that the atmospheric boundary layer and its corresponding z_0 value develops over

100s of m of fetch. It is therefore questionable whether 4 m² ‘samples’ of the glacier microtopography are sufficient to characterize the surface roughness over such large areas, and it is doubtful whether the wind profile would be able to adjust to order of magnitude roughness changes (Table 2) over the short distances in your study area. Consequently it is physically meaningless to have z_0 as a tuneable parameter at individual stakes. It would be more realistic if the 4 SfM sites were combined to generate a single z_0 value and its range.

- As the reviewer notes, the scale at which the atmospheric boundary layer develops is a critical parameter in determining how z_0 will vary over the course of the debris-covered glacier. Lettau (1969) conducted his experiments over fetches of 50 m in length. In regions of relatively homogenous terrain, this is not an issue; however, on debris-covered glaciers, the terrain is highly heterogeneous over a 50 m span as the terrain is both hummocky and the grain sizes vary greatly. Therefore, it is a crucial area of future work to understand the scale (fetch length) at which z_0 varies over the glacier.

Brock et al. (2006) compared microtopographic and aerodynamic roughness over snow, slush, and ice and found close agreement. Interestingly, they found no significant difference between the use of a 3 m and 15 m transect; however, they did state that a shorter pole would be unlikely to capture a sufficient sample of roughness elements and a longer pole should be used if vertical changes are greater than 1 m. The use of hundreds of transects over a ~4 m² grid has the benefit of expanding the number of surface roughness elements that can be captured compared to a single traditional cross section. In this regard, while the approach is likely limited to roughness elements < 1 m, the approach should be able to capture a sufficient number of elements. Additionally, the method developed in this paper using an obstacle density of 30% allows one to take advantage of the high resolution DEM that can help to capture the irregularly shaped elements as opposed to using the method of Munro (1989), which as Brock et al. (2006) states was “a necessary generalization... due to the difficulty of measuring and converting irregularly shaped elements in a z_0 value.”

Unfortunately, there is no comparison between microtopographic and aerodynamic z_0 values on debris-covered glaciers. The z_0 values reported on debris cover (Inoue and Yoshida, 1980; Takeuchi et al., 2000; Brock et al., 2010) only measured aerodynamic roughness. Therefore, the scale (or fetch length) at which the aerodynamic roughness may vary over the terrain is unknown. This is an important area of future work. Nonetheless, the techniques developed in this study show a range of z_0 values for various grain sizes found on Imja-Lhotse Shar glacier. The average value of these four sites is 0.018 m (standard deviation of 0.013), which is similar to the value of 0.016 m found by Brock et al. (2010) on a debris-covered glacier in Italy “comprising a mixture of granites and schists of predominantly cobble size, with occasional boulders of < 1m size.”

As requested by the reviewer, an optimization was performed using a constant value for z_0 of 0.018 m. Table R2-1 shows that the thermal conductivity is at the upper

bound for 8 of the 10 sites, which is similar to the optimization allowing z_0 to vary for each site. The noticeable difference between optimizations is the increase in average z_0 from 0.014 m to 0.018 m causing the average albedo to decrease from 0.32 to 0.27. This makes intuitive sense as a higher surface roughness removes more energy from the debris, which is compensated by a lower albedo that will cause the surface to absorb more energy. The performance of the model appears to be comparable with the average R^2 of the surface sites being 0.73 for the constant z_0 compared to 0.74 when z_0 is allowed to vary for each site. Furthermore, the total sum of squares only varied by 5%. Without detailed knowledge of albedo, surface roughness, and thermal conductivity, it is difficult to determine which model performs better.

Table R2-1. LE_{Rain} model optimization allowing z_0 to vary and setting it at a constant value of 0.018 m.

Site	LE_{Rain}			LE_{Rain}		
	α	k^1	z_0^2	α	k^1	z_0^2
4	0.26	1.62	0.006	0.10	1.62	0.018
5	0.40	1.62	0.014	0.40	1.62	0.018
6	0.40	1.09	0.036	0.40	1.39	0.018
11	0.37	1.62	0.006	0.26	1.62	0.018
13	0.10	0.92	0.025	0.20	0.61	0.018
14	0.39	1.61	0.015	0.38	1.62	0.018
15	0.38	1.62	0.006	0.29	1.62	0.018
17	0.30	1.62	0.006	0.19	1.62	0.018
19	0.33	1.62	0.019	0.34	1.62	0.018
20	0.28	1.62	0.006	0.18	1.62	0.018
Avg	0.32	1.50	0.014	0.27	1.50	0.018
Std	0.09	0.26	0.010	0.11	0.32	0.000

¹units of $W\ m^{-1}\ K^{-1}$; ²units of m

This analysis does stress the importance of accurately measuring the albedo, surface roughness, and thermal conductivity. For example, Sites 6 and 13 show how varying z_0 causes the thermal conductivity to change by $0.30\ W\ m^{-1}\ K^{-1}$, which would have a significant impact on ablation rates. Future work should strive to measure all three of these parameters. Specifically, it is important to determine the scale at which surface roughness is equal to aerodynamic roughness, understand how albedo varies throughout the day and throughout the melt season, and improve measurements of thermal conductivity. The latter will likely be hampered by the limited ability of the temperature sensors to capture the small temperature changes with respect to time at deeper depths in the debris and our understanding of soil moisture within the debris profile(s). One promising method is to measure the albedo and surface roughness in

conjunction with meteorological data and ablation stakes, such that the ablation stakes may be used to approximate the thermal conductivity.

Changes to the manuscript: A paragraph has been inserted into the manuscript in Section 4.2 as follows: “Future work should seek to compare these estimates of surface roughness with aerodynamic roughness to determine the scale at which these two values agree. Brock et al. (2006) found there to be no significant difference between the use of a 3 m and 15 m transect; however, they did state that a shorter pole would be unlikely to capture a sufficient sample of roughness elements if the vertical changes are greater than 1 m. The use of hundreds of transects over a $\sim 4 \text{ m}^2$ grid has the benefit of expanding the number of surface roughness elements that can be captured compared to a traditional single transect. However, Brock et al. (2006) was comparing microtopographic and aerodynamic roughness over snow, slush, and ice, which is significantly different from the hummocky and heterogeneous terrain on debris-covered glaciers. Therefore, it will be important to determine the scale or fetch length at which the surface roughness agrees with the aerodynamic roughness. Nonetheless, the method developed in this paper provides an objective approach to select an obstacle height and yields consistent and reasonable estimates of z_0 for various grain sizes independent of the resolution of the DEM.”

Section 4.1: Calculation of thermal conductivity:

Why were only the near-surface thermistors down to 20 cm used and not the deeper thermistors at sites 4 and 11? This implies that the k calculation is biased towards the openwork clast layers at the surface and not the more compact and humid lower debris layers. It could be reasoned that as void spaces of deeper layers are filled with water and fine rock material the thermal conductivity here would be higher than near the surface where void spaces are filled with low conductivity air. This implies an underestimation of the full depth k due to the use of temperature data only from the upper layers. This reasoning is supported by the fact that in the calibration of model parameters in Section 5.1 the optimal k value is at the maximum for most sites, implying the true value is greater than 1.62.

- The method from Conway and Rasmussen (2000) is based on approximations of T'' ($\delta^2 T / \delta Z^2$) and \dot{T} ($\delta T / \delta t$) using standard centered finite-difference expressions. While Conway and Rasmussen (2000) show in the Appendix that the analysis is not affected by the thermistor position and calibration errors, this is referring to errors in measured position and errors in temperature readings between different sensors. The use of a standard centered finite-differences means the ideal set up is to have three sensors that are all more or less equidistant from one another. For example, three temperature sensors at 10 cm, 20 cm, and 30 cm allow $\delta T / \delta Z$ at 15 cm and 25 cm to be approximated, which is then used to approximate $\delta^2 T / \delta Z^2$ at 20 cm. This coincides with the 20 cm temperature sensors, which records $\delta T / \delta t$. Nicholson and Benn (2012) note that deeper sensors are also problematic as temperature sensors may not be sensitive enough to capture the small temperature changes that occur, which was

found to be the case with the temperature sensor at a depth of 0.83 m at Site 4. Therefore, the combination of the lack of sensors spaced at a reasonable distance to use the standard centered finite-difference approach and the thermistors not being sensitive enough to record temperature changes at deeper depths caused the deepest thermal conductivity to be reported at 20 cm. This could be improved in the future by spacing sensors similar to Nicholson and Benn (2012) and by using more sensitive temperature sensors.

The reviewer brings up an important point that the lack of thermal conductivity calculations at depth may cause the actual thermal conductivity at the site to be underestimated, as the deeper layers that may be more compact and humid within the debris are not considered. Nicholson and Benn (2012) estimated the effects of 10% and 20% of the void space being filled with water and found the thermal conductivities increased to 1.42 and 1.55 $\text{W m}^{-1} \text{K}^{-1}$ compared to the summer dry debris value of 1.29 $\text{W m}^{-1} \text{K}^{-1}$. This lends confidence to the higher average values of k found in this study at Sites 4 and 11 of 1.44 and 1.62 $\text{W m}^{-1} \text{K}^{-1}$. However, adjusting the thermal conductivities based on the percent of void space filled with water requires detailed knowledge of the soil moisture within the debris and how the soil moisture varies with depths, which unfortunately was not measured in this study.

To address the comment that the thermal conductivities may be underestimated, an optimization was performed for all three models with k unbounded and only using the surface temperature sensors at each site (in response to a comment from the other reviewer). In this optimization, the thermal conductivities at Sites 4, 11, and 13 were also allowed to vary (as opposed to being held constant at their average measured value) to avoid the bias associated with only measuring thermal conductivities near the surface. The results of the unbounded thermal conductivity optimization are shown in Table R2-2.

Table R2-2. Model optimization with k unbounded

Site	LE _{Rain}			LE _{RH100}			LE _{Dry}		
	α	k^1	z_0^2	α	k^1	z_0^2	α	k^1	z_0^2
4	0.13	2.51	0.012	0.21	2.34	0.008	0.10	3.30	0.024
5	0.35	3.99	0.006	0.31	4.50	0.006	0.35	4.20	0.006
6	0.40	1.09	0.036	0.40	1.34	0.036	0.40	1.20	0.036
11	0.20	3.40	0.006	0.22	3.18	0.006	0.15	3.90	0.012
13	0.10	0.92	0.025	0.10	1.77	0.021	0.10	2.10	0.024
14	0.31	2.95	0.009	0.38	2.42	0.009	0.35	2.70	0.012
15	0.18	3.64	0.006	0.17	3.76	0.006	0.20	3.60	0.006
17	0.26	2.03	0.006	0.27	1.96	0.006	0.25	2.10	0.006
19	0.20	2.66	0.026	0.29	2.24	0.017	0.15	3.00	0.036
20	0.26	1.80	0.006	0.27	1.75	0.006	0.25	2.10	0.006
Avg	0.24	2.50	0.01	0.26	2.53	0.01	0.23	2.82	0.02
Std	0.10	1.04	0.01	0.09	0.99	0.01	0.11	0.95	0.01

¹units of W m⁻¹ K⁻¹; ²units of m

Table R2-2 shows the optimized values of k are as high as 4.5 W m⁻¹ K⁻¹. The lithology of the debris cover located in the Everest region is predominantly granite, gneiss, and pelite (Hambrey et al., 2008). Robertson (1988) reports a thermal conductivity of solid granite gneiss to be 2.87 W m⁻¹ K⁻¹. Therefore, many of the values reported in Table R2.1 appear to be unreasonable, especially when considering the fact that the pores of the debris are filled with air and water, which would cause the effective thermal conductivity to be smaller than that of solid rock. Interestingly, Site 13 has the lowest value of thermal conductivity for all three models, which is consistent with the estimations of thermal conductivity compared to Sites 4 and 11. However, its value (~1.20 W m⁻¹ K⁻¹) is still more than twice as high as the average k measured (0.47 W m⁻¹ K⁻¹). Therefore, while the unbounded analysis yields some interesting results, it appears to be an unreasonable method to optimize the thermal conductivity at each site as many of the sites have values of k that are unreasonable.

The unbounded analysis shows that it is important to have an upper bound for the optimization. Figure R2-1 shows the variations in total sum of squares for all the surface sites for various values of thermal conductivity. The total sum of squares appears to plateau around 2.0 W m⁻¹ K⁻¹ and does not appear to significantly change for values above 1.6 W m⁻¹ K⁻¹. In fact, the difference between the minimum sum of squares, which occurs at 2.6 W m⁻¹ K⁻¹, and the minimum associated with 1.6 W m⁻¹ K⁻¹ is only 3%. This study found the thermal conductivity to vary from 0.42 to 2.28 W m⁻¹ K⁻¹ with the highest average thermal conductivity being Site 11 with a value of 1.62 W m⁻¹ K⁻¹. As previously discussed, these measurements may not be representative of the entire debris as they were measured closer to the surface; however, the value of 1.62 W m⁻¹ K⁻¹ is higher than those previously reported in the Everest region (Conway and Rasmussen, 2000; Nicholson and Benn, 2012). This

value is also similar to the value reported by Nicholson and Benn (2012) of $1.55 \text{ W m}^{-1} \text{ K}^{-1}$, which assumed that 20% of the pore space was filled with water. Therefore, a value of $1.62 \text{ W m}^{-1} \text{ K}^{-1}$ appears to be a reasonable upper bound as higher values do not significantly reduce the total sum of squares and it is a reasonable value of effective thermal conductivity. The uncertainty associated with the thermal conductivity and the upper bound will be addressed via the sensitivity analysis, which will use an uncertainty of $\pm 0.4 \text{ W m}^{-1} \text{ K}^{-1}$ as this covers the range of the value where the sum of squares levels off (2.0) and the effective thermal conductivity reported by other studies (1.22).

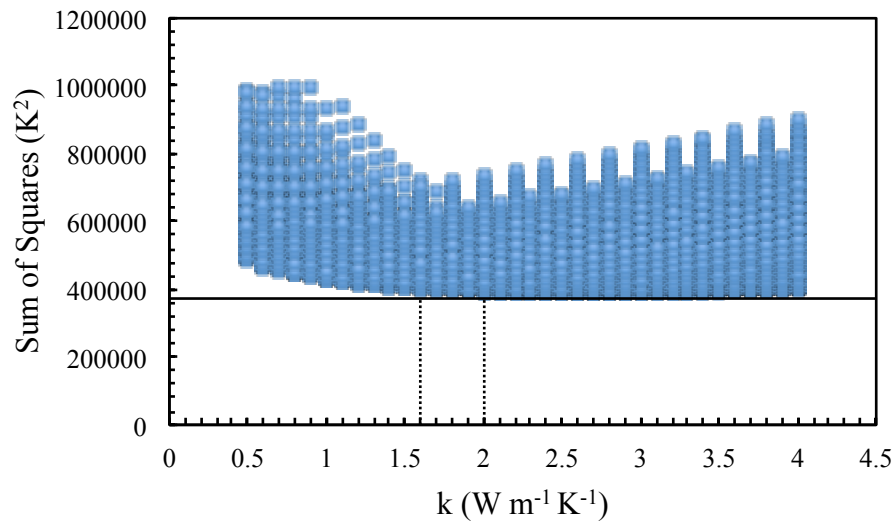


Figure R2-1. Total sum of squares for all surface sites as a function of k

Changes to the manuscript: A discussion regarding the upper bound and the calibration of k has been added to Section 5.1 as follows: “For the LE_{Rain} and LE_{RH100} model, 7 of the 10 sites had a thermal conductivity at the upper bound ($1.62 \text{ W m}^{-1} \text{ K}^{-1}$), while for the LE_{Dry} model 9 of the 10 sites were at the upper bound. An additional calibration was performed allowing the thermal conductivity to be unbounded and found 3 or more out of the 10 sites for each method had thermal conductivities greater than $3.0 \text{ W m}^{-1} \text{ K}^{-1}$ with one thermal conductivity as high as $4.5 \text{ W m}^{-1} \text{ K}^{-1}$. The lithology of the debris cover in the Everest region is predominantly granite, gneiss, and pelite (Hambrey et al., 2008). Robertson (1988) found the thermal conductivity of solid granite gneiss to be $2.87 \text{ W m}^{-1} \text{ K}^{-1}$, so the unbounded thermal conductivities do not appear to make physical sense when one considers that the thermal conductivity of debris should be much lower than solid rock due to the pore spaces being filled with air and water. Furthermore, an optimization performed using the total sum of squares of all the surface sites reveals that increasing the thermal conductivity from $1.6 \text{ W m}^{-1} \text{ K}^{-1}$ to its minimum of $2.6 \text{ W m}^{-1} \text{ K}^{-1}$ only reduces the total sum of squares by 3%. Therefore, the results reported in this study use an upper bound of $1.62 \text{ W m}^{-1} \text{ K}^{-1}$ and the importance of accurately measuring the thermal conductivity will be discussed in the sensitivity analysis.”

You use values from Nicholson and Benn (2012) for debris characteristics in the k calculation. Are these values representative for the sites you measured, and how sensitive is k value to e.g. a 10% change in porosity?

- Nicholson and Benn (2012) was performed on Ngozumpa glacier, which is located ~25 km away from Imja-Lhotse Shar glacier. The characteristics of the debris cover on Ngozumpa glacier appear to be very similar to those on Imja-Lhotse Shar glacier, which lends confidence to the use of the same values as their study. A 10% change in porosity would result in a 5% change in thermal conductivity.

Section 4.3

It probably doesn't need pointing out, but it is a shame that most of the stakes melted out. Why didn't you drill the stakes in deeper, e.g., 2 m?

- A mechanical drill was used to drill the ablation stakes and after drilling ~1 m into the ice, the drill bit would get stuck. A thermal drill would be preferable for future work.

Section 5.1

A problem with the method of tuning model parameters separately for each of the 3 different methods of estimating turbulent fluxes is that the parameter values will compensate for any errors in the turbulent flux calculation. For example, in the LE_{Dry} method, neglect of the energy used in evaporating water could lead to overestimation of the conductive heat flux, however this is offset by the relatively high albedo and z_0 optimized for this run (Table 3), which reduce net shortwave radiation and increase sensible heat transfer away from the surface during the daytime. This makes interpretation of the differences in performance of the model with three different LE flux formulations difficult. Maybe you should select one set of optimal parameters and look at the difference in performance between the three model formulations again. At least this issue needs some discussion.

- The optimization of each model certainly allows the LE_{Dry} model to compensate for the lack of a latent heat flux term through higher thermal conductivities and surface roughnesses, which allow more energy to be removed from the surface. The difficulty in assessing model performance, as the reviewer commented, was discussed in the text (p3520, line 20). As albedo, thermal conductivity, and surface roughness were not measured at each site, this study lacks the ability to assess which model most accurately estimates the debris properties. However, the latent heat fluxes have been found to be a significant energy sink after rain events as discussed in the paper,

so the instantaneous and daily average over the melt season latent heat fluxes may be used to show that the LE_{Dry} model is not physically accurate.

Section 5.4

I think you should redo the sensitivity analysis, varying each parameter in its range of uncertainty rather than a flat value of $\pm 10\%$. It is obvious a priori that varying z_0 by $\pm 10\%$ will have little effect on the magnitude of the turbulent fluxes when this parameter can vary across several orders of magnitude. In contrast, a variation in albedo of 10% is quite a significant change as its range of variation is much smaller. This would give a better assessment of model uncertainty which reflects the possible range of values of the input parameters.

- The range of uncertainty is difficult to assess as there are a limited number of measurements of the debris properties in this region. However, as the reviewer comments, it would be more beneficial to perform the sensitivity analysis by trying to account for the uncertainty within each parameter as opposed to using a set value of $\pm 10\%$. The uncertainty in thermal conductivity, as discussed above, will be $\pm 0.40 \text{ W m}^{-1} \text{ K}^{-1}$. The uncertainty associated with the surface roughness, will be $\pm 0.010 \text{ m}$, which is the approximate standard deviation associated with the z_0 values from the model calibration for each of the three models and similar to the standard deviation between the four sites where z_0 was measured ($\pm 0.013 \text{ m}$). Lastly, the uncertainty of the albedo will be estimated as ± 0.10 , which is the approximate standard deviation within the model calibration for each of the three models.

Table R2-3 shows the percent changes in the total melt (m) as a function of the uncertainty associated with each parameter. The LE_{Rain} model is used as the baseline case and the average value for each of the calibrated parameters (α, k, z_0) from the model optimization is used for each site. The use of the average calibrated parameters at every site was required because a -0.010 m adjustment to z_0 for sites with low z_0 values (0.006 m) would cause z_0 to be negative, which is impossible. Nonetheless, Table R2-3 shows how the uncertainty within each parameter effects the total ablation rates. The total ablation is most sensitive to changes in thermal conductivity, where a $\pm 0.40 \text{ W m}^{-1} \text{ K}^{-1}$ change in thermal conductivity causes on average a $\pm 20.6\%$ change in total melt. Total ablation is also moderately sensitive to changes in albedo, where a ± 0.10 change in albedo caused a $\pm 12.0\%$ change in total melt. The effect of surface roughness is quite interesting as the model is quite sensitive to decreases in surface roughness, i.e., a -0.010 m change in z_0 caused a $+15.9\%$ change in melt; while, a $+0.010 \text{ m}$ increase in z_0 caused only a -7.6% change in total melt on average. This reveals that the model is much more sensitive to decreases in surface roughness compared to higher values of surface roughness. surface roughness compared to increases.

Table R2-3. Sensitivity analysis showing percent changes relative to the total melt (m) as a function of the calibrated parameters (α, k, z_0) for all sites over the study

period using the LE_{Rain} model in conjunction with the average calibrated parameters for all the sites.

Parameter		α		k		z_0	
<i>Adjustment</i>		<i>+ 0.10</i>	<i>- 0.10</i>	<i>+ 0.40</i>	<i>- 0.40</i>	<i>+ 0.010</i>	<i>- 0.010</i>
Site	Total Melt (m)	% Change					
4	0.29	- 12.6	+ 12.8	+ 30.1	- 30.0	- 10.2	+ 22.8
5	0.9	- 12.7	+ 12.9	+ 22.5	- 24.3	- 9.6	+ 20.6
6	2.7	- 11.3	+ 11.2	+ 13.0	- 16.6	- 4.4	+ 8.2
7	0.92	- 12.6	+ 12.5	+ 22.1	- 23.9	- 9.2	+ 20.1
8	1.02	- 11.7	+ 11.7	+ 20.7	- 22.9	- 8.0	+ 17.1
10	1.61	- 12.2	+ 12.0	+ 19.0	- 21.5	- 7.9	+ 16.2
11	1.11	- 12.1	+ 12.4	+ 20.3	- 22.4	- 8.0	+ 17.1
12	1.3	- 11.9	+ 12.0	+ 19.9	- 22.2	- 8.1	+ 17.1
13	1.05	- 12.5	+ 12.6	+ 21.3	- 23.3	- 8.8	+ 18.9
14	1.72	- 12.0	+ 11.7	+ 18.1	- 21.0	- 7.3	+ 15.0
15	0.9	- 12.5	+ 12.8	+ 21.9	- 23.7	- 8.8	+ 18.7
16	1.76	- 11.7	+ 12.0	+ 18.2	- 20.5	- 7.0	+ 14.8
17	2.77	- 11.0	+ 10.9	+ 11.9	- 15.4	- 3.2	+ 5.9
19	2.04	- 11.7	+ 11.6	+ 16.5	- 19.7	- 6.6	+ 13.0
20	1.81	- 11.6	+ 11.2	+ 16.5	- 19.7	- 6.4	+ 12.8
Average		- 12.0	+ 12.0	+ 19.5	- 21.8	- 7.6	+ 15.9

Changes to the manuscript:

Table 4 has been replaced by Table R2-3, which has the updated values incorporated into the sensitivity analysis. Additionally, the text accompanying the sensitivity analysis (Section 5.4) has been revised accordingly to read:

“A sensitivity analysis was performed to assess how albedo, thermal conductivity, and surface roughness affect the total ablation (Table 4) based on the uncertainty with respect to each parameter. The uncertainty in thermal conductivity was $\pm 0.40 \text{ W m}^{-1} \text{ K}^{-1}$ as described above. The uncertainty associated with the surface roughness was $\pm 0.010 \text{ m}$, which is the approximate standard deviation associated with the z_0 values for each of the three models (Table 3) and similar to the standard deviation between the four sites where z_0 was measured ($\pm 0.013 \text{ m}$). Lastly, the uncertainty of the albedo was estimated as ± 0.10 , which is the approximate standard deviation within the model calibration for each of the three models and is the difference between the mean and median albedo measured by Nicholson and Benn (2012) on Ngozumpa glacier. The LE_{Rain} model was used as the baseline case and the average value for each of the calibrated parameters (α , k , z_0) from the model optimized is used for each site.

Table 4 shows the total ablation is most sensitive to changes in the thermal conductivity, where a $\pm 0.40 \text{ W m}^{-1} \text{ K}^{-1}$ change causes a $\pm 20.6\%$ change in total ablation on average. The uncertainty associated with the thermal conductivity is also more sensitive to thicker debris, which is consistent with the findings of Nicholson and Benn (2012). Total ablation is also moderately sensitive to changes in the albedo, where a ± 0.10 change causes a $\pm 12.0\%$ change in total ablation. Lastly, the total ablation is least sensitive to changes in increasing the surface roughness, as a $+0.010 \text{ m}$ increase in z_0 caused only a -7.6% change in total ablation. However, the model was quite sensitive to a reduction in the z_0 of -0.010 m , which caused an average change in total ablation of $+15.9\%$. The sensitivity associated with z_0 also appears to increase with an increase in debris thickness. These results highlight the importance of properly estimating the thermal conductivity, but also show the surface roughness and the albedo are important as well.”

Minor Corrections

Paper title. You need to insert glacier between ‘Debris-covered’ and ‘energy’. As it stands, the title literally means that the model is debris-covered, not the glacier. Make this change everywhere the phrase ‘Debris-covered energy balance model’ appears in the paper, e.g., section 3.2 title.

- Good point. The changes have been made to the title and throughout the entire text.

P3506/17: I think you mean partial density of water vapour, not water vapour pressure. If vapour pressure was the same at 2 m as at the surface, there would be no vapour pressure gradient and hence no latent heat flux.

- Correct. It has been changed to “water vapour partial pressure”.

P3507/19: ‘cobble and gravel’.

- Correct. It has been changed to “cobble and gravel”.

P3509: Section 2.2, make it clear that the Pyramid Station is an off-glacier station.

- The manuscript has been changed to “... located off-glacier, next to the Khumbu glacier, approximately 14 km northwest of Imja-Lhotse Shar glacier.”

P3519/23: ‘...this particular temperature sensor...’, which one? Be specific.

- Site 13 at 20 cm depth. The manuscript has been changed to include “(Site 13, 20 cm).”

P3521/14: I think you mean overestimate not underestimate.

- The other reviewer also commented on this. The language has been changed such that we refer to overestimating temperatures as a positive bias in temperature to avoid confusion with overestimating the nightly low, which could be thought of as lower temperatures.

P3522/12: You can't say that thin debris 'promotes ablation' as you have not measured bare ice melt rates you can't determine that the melt rate beneath thin debris is greater than that for bare ice. In all likelihood even the thinnest of your debris layers is reducing ablation through insulation.

- Good point. The sentence has been modified to read as "... as thin debris has higher rates of ablation compared to thicker debris, which insulates the ice to a greater extent thereby further retarding ablation."

P3519/23: These are not surface temperatures but temperatures at 0.01 m depth – an important distinction.

- Changed from "surface temperatures" to "temperatures close to the surface" and "near the surface".

P3523/20: You can't say the model agrees with the ablation stake data, when the stakes have melted out. All you can see is that the stake data do not contradict the model calculations.

- The sentence has been modified accordingly to read as "All the other model estimates of ablation were near to or greater than 1 m, which was also observed by their respective ablation stakes as they completely melted out of the ice."

Figure 1: What is the background image?

- "Landsat 8 panchromatic image from 14 Nov 2014" has been added to the figure's caption.

Figure 2: State the size of the target discs in the caption.

- The target discs are 19 cm in diameter. This has been added to the caption.

Figure 5: Why is the ablation for LE_{Dry} so much higher than the other two methods for the thinnest debris layer, compared with the other sites? There is an inconsistency in terminology with the text here: LE_{Dry} , LE_{Zero} .

- The ablation for that particular thin debris layer (Site 6) was much higher than any of the other sites as a result of the individual calibration process. The three models each have an albedo of 0.40 and a z_0 of 0.030, so the only way for that site to compensate for the lack of a turbulent heat flux was to increase its thermal conductivity, i.e., from $1.15 \text{ W m}^{-1} \text{ K}^{-1}$ for the LE_{Rain} and LE_{RH100} models to $1.34 \text{ W m}^{-1} \text{ K}^{-1}$ for the LE_{Dry}

model. This increase in thermal conductivity greatly increased the melting. Other sites adjusted the albedo and z_0 to compensate for the lack of a latent heat flux terms, which is why the differences in the melt for Site 6 were so much greater compared to those differences at other sites.

LE_{Zero} was a typo and has been changed to LE_{Dry} throughout the text.

References

- Hambrey, M.J., Quincey, D.J., Glasser, N.F., Reynolds, J.M., Richardson, S.J., and Clemmens, S.: Sedimentological, geomorphological and dynamic context of debris-mantled glaciers, Mount Everest (Sagarmatha) region, Nepal, *Quaternary Sci Rev*, 27, 2361-2389, 2008.
- Robertson, E.C.: Thermal Properties of Rocks. US Geological Survey (No. 88-441), 1988.

Debris-Covered Glacier Energy Balance Model for Imja-Lhotse Shar Glacier in the Everest Region of Nepal

D. R. Rounce¹, D. J. Quincey², and D. C. McKinney¹

[1]{Center for Research in Water Resources, University of Texas at Austin, Austin, Texas, USA}

[2]{School of Geography, University of Leeds, Leeds, LS2 9JT, UK}

Correspondence to: D. R. Rounce (david.rounce@utexas.edu)

Keywords: Debris-covered glaciers, Himalaya, Energy Balance, Structure from Motion, Ablation, Surface Roughness

Abstract

Debris thickness plays an important role in regulating ablation rates on debris-covered glaciers as well as controlling the likely size and location of supraglacial lakes. Despite its importance, lack of knowledge about debris properties and associated energy fluxes prevents the robust inclusion of the effects of a debris layer into most glacier surface energy balance models. This study combines fieldwork with a debris-covered glacier energy balance model to estimate debris temperatures and ablation rates on Imja-Lhotse Shar glacier located in the Everest region of Nepal. The debris properties that significantly influence the energy balance model are the thermal conductivity, albedo, and surface roughness. Fieldwork was conducted to measure thermal conductivity and a method was developed using Structure from Motion to estimate surface roughness. Debris temperatures measured during the 2014 melt season were used to calibrate and validate a debris-covered glacier energy balance model by optimizing the albedo, thermal conductivity, and surface roughness at 10 debris-covered sites. Furthermore, three methods for estimating the latent heat flux were investigated. Model calibration and validation found the three methods had similar performance; however, comparison of modeled and measured ablation rates revealed that assuming the latent heat flux is zero may overestimate ablation. Results also suggest that where debris moisture is unknown, measurements of the relative humidity or precipitation may be used to estimate wet debris periods, i.e., when the latent heat flux is non-zero. The effect of temporal resolution on the model was also assessed and results showed that both 6-hour data and daily average data slightly underestimate debris

temperatures and ablation rates, thus these should only be used to estimate rough ablation rates when no other data are available.

1 Introduction

Debris-covered glaciers are commonly found in the Everest region of Nepal and have important implications with regard to glacier melt and the development of glacial lakes. It is well understood that a thick layer of debris (i.e. > several cm) insulates the underlying ice, while a thin layer of debris (i.e. < several cm) may enhance ablation (Østrem, 1959; Nakawo and Young, 1981; Nicholson and Benn, 2006; Reid et al., 2012). Spatial variations in debris thickness, particularly where the debris layer thins up-glacier, can also lead to reverse topographic and ablation gradients, glacier stagnation and, ultimately, the development of lakes (Benn et al., 2012). [These glacial lakes and their surrounding bare ice faces also play a crucial role in glacier melt as they typically have ablation rates that are orders of magnitude greater than those observed beneath debris cover \(Benn et al., 2012\).](#) The importance of debris thickness has led many studies to develop models in conjunction with knowledge of the surface temperature to derive debris thickness (Zhang et al., 2011; Foster et al., 2012; Fujita and Sakai, 2014; Rounce and McKinney, 2014). With knowledge of debris thickness, energy balance models may be used to model debris surface temperature, sub-debris ablation rate, and/or runoff downstream (Nicholson and Benn, 2006; Reid et al., 2012; Collier et al., 2014; Fujita and Sakai, 2014). The main factors affecting the performance of these models are the amount of knowledge of the debris properties, the spatial and temporal resolution of the meteorological data, and the assumptions/complexity of the model.

The properties of the debris typically required in debris-covered glacier energy balance models are the albedo, thermal conductivity, and surface roughness. The albedo of debris on glaciers in the Everest region has been found to range from 0.1 – 0.6 (Inoue and Yoshida, 1980; Kayastha et al., 2000; Nicholson and Benn, 2012; LeJeune et al., 2013). Specifically, Nicholson and Benn (2012) reported that 62% of measured values ranged between 0.1 and 0.3. Similarly, Kayastha et al. (2000) showed that most values fall between 0.2 and 0.4. The thermal conductivity of debris in the Everest region has been found to range from 0.60 to 1.29 W m⁻¹ K⁻¹ (Conway and Rasmussen, 2000; Nicholson and Benn, 2012; Rounce and McKinney, 2014). The surface

roughness, z_0 , is arguably the most difficult parameter to measure as it requires an eddy covariance instrument, horizontal wind speed measurements at multiple heights above the surface, or detailed microtopographic measurements (Brock et al., 2006). In the Everest region, Inoue and Yoshida (1980) estimated z_0 to be 0.0035 m and 0.060 m for two sites, one consisting of small schist and bare ice and another comprising mainly large granite, respectively. Takeuchi et al. (2000) estimated a similar value of z_0 on the Khumbu glacier of 0.0063 m. On Miage glacier in the Italian Alps, Brock et al. (2010) measured z_0 to be 0.016 m on a debris-covered glacier.

In addition to the properties of the debris, the amount and source of meteorological data available may also greatly influence the model performance. In particular, knowledge related to the latent heat flux on debris-covered glaciers is very limited. This has led previous studies to assume the surface is dry (Foster et al., 2012; Lejeune et al., 2013; Rounce and McKinney, 2014), assume it is dry unless the surface relative humidity was 100% (Reid and Brock 2010; Reid et al., 2012; Fyffe et al., 2014), assume a relationship between debris thickness and wetness (Fujita and Sakai, 2014), or use a reservoir approach to model the moisture in the debris (Collier et al., 2014). Collier et al. (2014) suggested that if the atmospheric surface layer is well mixed, then the water vapor partial pressure between the surface and the air may be assumed to be constant, thereby resulting in a latent heat flux based on the vapor pressure gradient. Fyffe et al. (2014) also commented that the lower portion of the debris near the ice interface was observed to be saturated indicating that there may be evaporation and condensation occurring within the debris, albeit small, even when the surface relative humidity is less than 100%. The lack of knowledge of the moisture in the debris and at its surface makes it difficult to accurately model the latent heat flux term. These problems are further exacerbated in data scarce regions where automatic weather stations are not available. In these situations, reanalysis datasets must be used for all the required meteorological data (Fujita and Sakai, 2014).

This study develops a method to estimate z_0 using a microtopographic method in conjunction with Structure from Motion (SfM) photogrammetry techniques (Westoby et al., 2012). The z_0 values are used with measured values of thermal conductivity, and previously reported values of albedo to calibrate a debris-covered glacier energy balance model on Imja-Lhotse Shar glacier.

Temperature sensors installed at various depths at debris-covered sites were operated from May, to November 2014 on Imja-Lhotse Shar glacier and are used for calibration and validation of the

Daene McKinney 10/27/15 12:58 PM

Deleted: June

Daene McKinney 10/27/15 12:58 PM

Deleted: October

model. Various methods for estimating the latent heat flux are investigated. Furthermore, sub-debris ablation rates are compared to ablation stake measurements to assess model performance and the effects of temporal resolution are investigated.

2 Data

2.1 Field Data

Field research was conducted on the debris-covered portion of Imja-Lhotse Shar glacier (27.901°N, 86.938°E, ~5050 m a.s.l., Figure 1) from May to November 2014. [Imja-Lhotse Shar glacier refers to both Imja glacier and Lhotse Shar glacier, which are avalanche-fed debris-covered glaciers that converge and terminate into Imja Lake. The debris primarily consists of sandy boulder gravel \(Hambrey et al., 2008\) with the debris thickness increasing towards the terminal moraine. A more detailed description of the glacier may be found in Rounce and McKinney \(2014\).](#) The field expedition focused on 19 sites on the debris-covered portion of the glacier to determine how debris thickness and topography affect ablation rates. Four sites were used to analyze the surface roughness through the use of SfM and are referred to as Sites A-D (Figure 2). These sites were selected to represent various grain sizes and mixes of debris that were observed on Imja-Lhotse Shar glacier. Site A was relatively homogenous with the majority of debris being cobble and gravel ranging in size from 0.05 to 0.25 m. Site B comprised similar cobbles typically ranging in size from 0.15 to 0.25 m with larger boulders lying on top of the cobble of up to 1.0 m. Site C had the finest debris, which primarily consisted of fines and gravel with some cobbles on the surface up to 0.15 m in size. Lastly, Site D was the most heterogeneous site with boulders ranging up to 0.40 m overlying a surface of cobble of similar size to Site A mixed with the fine and gravel material found in Site C.

Temperature sensors and ablation stakes were installed at 20 other sites; however, data could only be retrieved from 15 of the sites (Table 1) as sensors were lost due to large changes in the topography and some of the loggers failed. Sites 4-14 were located in a single area that appeared to have developed from differential backwasting over the years. This was the same focus area as described in Rounce and McKinney (2014) and was selected because it appeared to be representative of the debris-covered terrain on Imja-Lhotse Shar glacier and was accessible. Sites 15-20 were located outside of the focus area in an adjacent melt basin to determine if the

Daene McKinney 10/26/15 10:29 PM

Deleted: Rounce and McKinney (2014) provide a detailed description of Imja-Lhotse Shar glacier so only study-specific details are given here.

1 focus area was representative of other debris-covered areas. At each site, the debris thickness
2 was determined following the methods described in Rounce and McKinney (2014) with the
3 exception of Site 4, where the debris thickness was greater than 1.0 m and therefore was
4 estimated assuming a linear temperature profile from the mean temperatures over the study
5 period similar to the extrapolation used in Nicholson and Benn (2012). The debris thickness of
6 these sites ranged from 0.07 m to greater than 1.0 m. A debris thickness of 1.0 m was considered
7 the maximum due to labor constraints. The slope was also approximated by measuring two
8 points, one 0.5 m uphill from the site and the other 0.5 m downhill, using a total station (Sokkia
9 SET520, $\pm 2.6\text{mm}/100\text{m}$). The slope at each site ranged from 17° to 37° . The aspect of each site
10 was measured using a compass (Table 1).

11 Temperature sensors (TR-42 ThermoRecorder, T&D Corporation) were installed and
12 successfully retrieved at 10 sites. These sensors recorded data every half hour from 19 May to
13 09 November 2014. Each of the 10 sites had a sensor at its surface, which was considered to be
14 installed 1 cm into the debris since debris was placed on top of the sensor. Sites 4, 11, 13, and
15 14 also had temperature sensors installed within the debris to capture the nonlinear temperature
16 variations in the debris and at three of the four sites the sensors were retrieved such that the
17 thermal conductivity could be estimated (Conway and Rasmussen, 2000).

18 Ablation stakes were also installed at 14 sites. One site had a debris thickness greater than 1.0 m,
19 so an ablation stake could not be installed. The ablation stakes were installed by excavating to
20 the debris-ice interface, at which point the debris thickness was measured, and then a 2 inch
21 diameter hole was drilled vertically approximately 1.0 m into the ice using a manual ice drill
22 (Kovacs Enterprise). A 2.0 m piece of 1 ½ inch PVC pipe was placed into the hole and the
23 height from the top of the ice to the top of the pipe was measured to determine the exact length
24 that the PVC pipe was inserted into the ice. A PVC end cap was then placed on top of the pole to
25 prevent anything from entering the hole through the pipe. The debris was then replaced in its
26 approximate original position.

27 **2.2 Meteorological Data**

28 The meteorological data used in the model calibration and validation was from an automatic
29 weather station (AWS), Pyramid Station (27.959° N, 86.813° E, 5035 m a.s.l., SHARE Network
30 operated by EV-K²-CNR), located off-glacier, next to the Khumbu glacier, approximately 14 km

northwest of Imja-Lhotse Shar glacier. The meteorological data provided by Pyramid Station were unvalidated, i.e., prior to their quality control processing, minute measurements of air temperature, wind speed, relative humidity, global radiation, precipitation, and snow depth. The data were processed to be consistent with the half-hour debris temperature measurements on Imja-Lhotse Shar glacier. The air temperature, wind speed, relative humidity, and global radiation data were reviewed and deemed plausible, so no adjustments were performed. The half-hour precipitation data was determined by summing the precipitation over each half-hour time step. A few of the minute measurements recorded negative precipitation, which were assumed to be zero as negative precipitation is not feasible. The half-hour snow depth data was processed to assume a snow depth of zero if snow was not recorded on the ground for the entire half-hour. The average snow depth over the half-hour was then computed and any average snow depth less than 0.001 m was considered to be zero. Wind speed data were collected at 5 m and adjusted to 2 m to be consistent with air temperature measurements for the turbulent heat fluxes assuming a logarithmic dependence (Fujita and Sakai, 2014). The snow depth data were used to derive a snowfall rate assuming a density of snow of 150 kg m^{-3} . The data were available from 31 May to 12 October 2014 with a few short gaps (missing 11.9% data). The first two days of meteorological data were used as start-up time for the model.

Longwave radiation was not measured at Pyramid Station during this period; therefore, the downward longwave radiation flux from NCEP/NCAR reanalysis data (Kalnay et al., 1996) was used with a minor modification. A comparison of the downward longwave radiation flux from NCEP/NCAR and the incoming longwave radiation flux at Pyramid Station from 2003 to 2010 (neglecting any data gaps) between the months of June and September revealed that NCEP/NCAR overestimated the incoming longwave radiation by an average of 29 W m^{-2} (results not shown). Therefore, the NCEP/NCAR downward longwave radiation flux was adjusted to account for this overestimation when being used in conjunction with the Pyramid Station data. This reanalysis dataset provides 6-hour meteorological data and was resampled using a linear interpolation such that the temporal resolution of the incoming longwave radiation agreed with the half-hour debris temperature measurements.

Ablation rates were modeled over the same time period as the ablation stakes (18 May to 09 November). For days where no meteorological data was available, i.e., the data gaps, the ablation for that day was assumed to be equal to the daily ablation rate for that specific month.

As the available meteorological data began on 31 May, the daily ablation rate for the month of May was assumed to be equal to the daily ablation rate of the first week of June. Temperature sensors revealed the debris was snow covered from 26 May to 01 June, so the melting during these days was assumed to be zero. Temperature profiles also show the debris was snow covered from 13-20 October and deeper thermistors revealed the temperature remained around freezing until the sensors were removed in November. Therefore, the melt rates after the 12 October were assumed to be zero.

3 Methods

3.1 Surface roughness (z_0)

Structure from Motion (SfM) was used to derive fine-resolution (i.e. centimetric) digital elevation models (DEMs) at four sites (Sites A-D) located on the debris-cover of Imja-Lhotse Shar glacier (Figure 2). In brief, SfM relies upon the acquisition of a series of overlapping images that capture the features of the terrain from a number of different vantage points. Computer vision techniques detect matching features between images using multiscale image brightness and colour gradients and a highly iterative bundle adjustment procedure is used to develop a three-dimensional structure of the surface (Snavely et al., 2008). Camera positions and orientations are solved simultaneously with surface geometry utilizing the high level of redundancy afforded by a large overlapping image set. Ground control points (GCPs), collected using a total station with an error less than 0.4 mm, are then used to transform the relative three-dimensional surface into an absolute coordinate system. The resulting point-cloud data are comparable in both density and accuracy to those generated by terrestrial laser scanning (Westoby et al., 2012) and can either be used as-is, or decimated (as in this study) to generate gridded elevation data. The use of SfM within geoscience is well reviewed by Westoby et al. (2012) and specific details of the mathematical operations involved can be found in Snavely (2008) and Szeliski (2011). Here, we therefore focus mostly on our field method and subsequent roughness analysis.

At each of our sites ~40 photos were taken around a roughly 2 m x 2 m grid. Cones were placed in the four corners of the grid as GCPs and their location was measured using a total station with

a local coordinate system. The GCPs and photos were processed using Agisoft PhotoScan Professional Edition Version 1.1.0 to create a DEM for each site. At each stage, the highest accuracy settings were chosen. No *a-priori* information about camera position or orientation was recorded, so these were estimated coincidentally as part of the adjustment. In each case the initial estimates of camera position and altitude were accepted and used to generate a sparse point cloud ($10^3 - 10^4$ points). A moderate depth filter was then used to derive a dense cloud ($10^6 - 10^7$ points), and subsequently a mesh was constructed using the height field as the surface type. The error of the DEM was computed as the root mean square error based on the differences between the measured GCPs from the total station and the modeled position of the GCPs from the software. The resulting DEMs were then resampled in ArcGIS 10.3 to a resolution of 0.01 m and were clipped to remove the cones from the subsequent analyses. The DEM was then fit with an x-y plane using a method of least squares such that the DEM was flattened with a mean elevation of zero.

These processed DEMs of the four sites were analyzed to determine the surface roughness, z_0 . Lettau (1969) developed an empirical relationship to estimate z_0 :

$$z_0 = 0.5h^* \frac{s}{S} \quad (1)$$

where h^* is the average vertical extent or effective obstacle height, s is the silhouette area or area of the upwind face of an average element, and S is the specific area or unit ground area occupied by each obstacle. Previous studies have estimated the variables in Equation 1 through a simplified standard deviation approach (which will be referred to as the Lettau-Munro method), based on the variations in elevations and the number of continuous positive groups above the mean elevation (Munro, 1989; Rees and Arnold, 2006; Brock et al., 2006). Initially, the Lettau-Munro method was applied to measure z_0 for every row and column transect of the four DEMs; however, the resulting values of z_0 did not capture the variations between sites and may have been slightly underestimated (see results).

Consequently, an alternative method was developed to estimate the effective height, silhouette area, and unit ground area of each obstacle using a similar transect approach and taking

advantage of the high resolution DEM. One problem with applying the method from Lettau (1969) is the lack of a clear definition of what constitutes an obstacle. The surface roughness will greatly vary depending on what is considered to be an obstacle, so a method must be developed that (i) objectively determines the obstacle height and (ii) yields reasonable 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) because the obstacles begin to aerodynamically interfere with one another. Therefore, a method was developed to select an obstacle height based on an obstacle density of 30%.

Initially, all the relative topographic highs and lows were identified. This was done for all of the transects in each of the four cardinal directions with respect to the DEM, i.e., every East-West, North-South, West-East, and South-North transect. Every elevation change between a relative low and high was considered a potential obstacle. The depth of each obstacle was defined as the distance between two low points surrounding the obstacle's high point. In the event that an obstacle was identified, but there was no low point following the high point, i.e., the low point was outside the extent of the transect, then the depth of the obstacle could not be determined. Figure 3 shows an example of a transect from Site B, which identifies the obstacle's height and depth based on the method developed in this study. The obstacle density was then defined as the cumulative depth of all the obstacles above the obstacle threshold divided by the length of the transect. An iterative approach was then used to determine the obstacle height that causes the obstacle density to reach the 30% threshold.

Once the obstacle height has been determined, the silhouette area and unit ground area were approximated from the height and depth of the obstacles. Specifically, the silhouette area was taken to be the height of the obstacle times a unit width and the unit ground area was estimated as the depth of the obstacle times a unit width. Based on these definitions, Equation 1 may be simplified to (Eqn 2):

$$z_0 = 0.5 \frac{h^2}{d_{obst}} \quad (2)$$

where d_{bst}^* is the average depth of the obstacle. The surface roughness, z_0 , was computed using the average effective obstacle height and average obstacle depth for each transect. In the event that an obstacle was identified, but did not have a depth, the obstacle's height was still used in the average.

3.2 Debris-Covered Glacier Energy Balance Model

The model used in this study was a steady-state surface energy balance model for a debris-covered glacier, where:

$$R_n(T_s) + H(T_s) + LE(T_s) + P(T_s) + Q_c(T_s) = 0 \quad (3)$$

where R_n is the net radiation flux, H is the sensible heat flux, LE is the latent heat flux, P is the heat flux supplied by rain, and Q_c is the ground heat flux (all in W m^{-2}). The net radiation and sensible heat fluxes are fully described in Rounce and McKinney (2014); however, in the current study the incoming shortwave radiation was only corrected for the effects of topography as shading could not be considered due to the lack of a high resolution DEM of the glacier.

The latent heat flux is difficult to determine without detailed knowledge of the moisture in the debris or the relative humidity at the surface. As the surface relative humidity was unknown, this study has analyzed three methods for estimating the latent heat flux: (1) assuming the debris is dry ($LE = 0$), (2) assuming it is dry unless the relative humidity is 100%, at which point the surface relative humidity is assumed to also be 100% based on the assumption that the water vapor above the surface is well mixed, and (3) assuming the surface is saturated when it is raining. These methods for modeling the latent heat flux will be referred to herein as LE_{Dry} , LE_{RH100} , and LE_{Rain} , respectively. The reservoir-approach detailed by Collier et al. (2014) and the empirical relationship between debris thickness and wetness (Fujita and Sakai, 2014) were not applied to this study due to the limited amount of knowledge of moisture within the debris and how the debris properties change with respect to depth. The latent heat flux is thus estimated according to Nicholson and Benn (2006):

$$LE = \left(\frac{0.622 \rho_{air}}{P_0} \right) L_e A u (e_z - e_s) \quad (4)$$

where

$$A = \frac{k_{vk}^2}{\ln\left(\frac{z}{z_0}\right) \ln\left(\frac{z}{z_0}\right)} \quad (5)$$

4

5 where ρ_{air} is the density of air at standard sea-level pressure (1.29 kg m^{-3}), P_0 is the standard air
6 pressure at sea level ($1.013 \times 10^5 \text{ Pa}$), L_e is the latent heat of evaporation of water ($2.49 \times 10^6 \text{ J}$
7 kg^{-1}), A is a dimensionless transfer coefficient, u is the wind speed collected at a height of 2 m
8 (m s^{-1}), e_z and e_s are the vapor pressures (Pa) at height z , 2 m, and on the surface of the debris,
9 respectively, k_{vk} is von Karman's constant (0.41), and z_0 is the surface roughness.

10 The heat flux due to precipitation was estimated following Reid and Brock (2010):

11

$$P = \rho_w c_w w (T_r - T_s) \quad (6)$$

13

14 where ρ_w is the density of water (999.97 kg m^{-3}), c_w is the specific heat capacity of water ($4.18 \times$
15 $10^3 \text{ J kg}^{-1} \text{ K}^{-1}$), w is the rainfall rate (m s^{-1}), and T_r is the temperature of rain (K), which was
16 assumed to be equal to the air temperature.

17 The debris layer was broken down into layers of 0.01 m such that the nonlinear temperature
18 profiles in the debris could be captured using a Crank-Nicholson Scheme (Reid and Brock, 2010).
19 The conductive heat flux at the surface and at the debris/ice interface were estimated following
20 Reid and Brock (2010):

21

$$Q_{c,s} = k_{eff} \frac{T_d(1) - T_s}{h} \quad (7)$$

$$Q_{c,ice} = k_{eff} \frac{T_d(N-1) - T_{ice}}{h} \quad (8)$$

24

1 where k_{eff} is the effective thermal conductivity ($\text{W m}^{-1} \text{K}^{-1}$), h is the height of each layer in the
2 debris set at 0.01 m, and $T_d(1)$, $T_d(N-1)$, T_{ice} are the temperatures (K) of the first layer in the
3 debris, the last layer before the debris/ice interface, and the temperature of the ice (273.15 K),
4 respectively.

5 The surface temperature was computed at half-hour time steps using an iterative Newton-
6 Raphson method approach as detailed in Reid and Brock (2010). In the event of snow, a simple
7 snowmelt model was used (Fujita and Sakai, 2014), which applies an energy balance over the
8 snow surface that includes net radiation, turbulent heat fluxes, and conductive heat flux with the
9 debris layer in addition to a variable surface albedo of the snow based on the number of days
10 since fresh snow and the air temperature. The thermal conductivity of snow was assumed to be
11 $0.10 \text{ W m}^{-1} \text{K}^{-1}$ (Sturm et al., 1997; Sturm et al., 2002; Rahimi and Konrad, 2012) and the surface
12 roughness of the snow was assumed to be 0.002 m (Brock et al., 2006). If snow was on the
13 surface, all the heat fluxes at the debris surface were assumed to be zero with the exception of
14 the conductive heat flux in the debris and at the debris/snow interface. If all the snow was
15 melted on the surface, then the next time step returned to the snow-free energy balance model.

16 As detailed knowledge of albedo, thermal conductivity, and surface roughness was not available
17 for the sites where temperature sensors were installed, the debris-covered glacier energy balance
18 model was calibrated at each site from 02 June to 30 July 2014. The calibration was performed
19 by minimizing the total sum of squares of the measured versus modeled surface temperature for
20 each site and was done independently for the three methods used to estimate the latent heat flux.
21 Bounds for the thermal conductivity and surface roughness were based on measured field data
22 (see results), while the bounds for the albedo were 0.1 – 0.4 (Inoue and Yoshida, 1980; Kayastha
23 et al., 2000; Nicholson and Benn, 2012; LeJeune et al., 2013). A validation was then conducted
24 at each site using data from 08 August to 12 October 2014 to assess how well the calibrated
25 model performed.

4 Field results

4.1 Thermal Conductivity (k)

The thermal conductivity, k , of the debris was computed using the temperature measurements from Sites 4, 11, and 13 over the time period of the study (02 June – 12 October 2014) following the methods of Conway and Rasmussen (2000). The calculations used standard values for the density of rock (2700 kg m^{-3}), volumetric heat capacity of rock ($750 \text{ J kg}^{-1} \text{ K}^{-1} \pm 10\%$), and effective porosity (0.33) based on Nicholson and Benn (2012). Depending on the vertical spacing of temperature sensors at a site, k was computed at depths of 0.05, 0.10, and 0.20 m. The values of thermal conductivity ranged from $0.42 (\pm 0.04)$ to $2.28 (\pm 0.23) \text{ W m}^{-1} \text{ K}^{-1}$. The average value of k for each site was $1.44 (\pm 0.14)$, $1.62 (\pm 0.16)$, and $0.47 (\pm 0.04) \text{ W m}^{-1} \text{ K}^{-1}$ for Sites 4, 11, and 13, respectively.

These values agree well with other studies in the Everest region that have found the thermal conductivity to vary between 0.60 to $1.29 \text{ W m}^{-1} \text{ K}^{-1}$ (Conway and Rasmussen, 2000; Nicholson and Benn, 2012; Rounce and McKinney, 2014). In September 2013, Rounce and McKinney (2014) found the thermal conductivity to be greatly influenced by depth; however, this trend was not apparent in our current data. We believe this disparity can be explained by the time period during which the data were collected. It is likely that the temporally-limited data (13-24 Sept 2013) presented in Rounce and McKinney (2014) represent a constantly dry surface, whereas here we observed an entire melt season, where the surface is exposed to precipitation and snow. The thermal conductivities appeared to have a trend over the monsoon season where the highest thermal conductivities were typically observed in July and August, which coincided with higher average air temperature and increased precipitation compared to the other months. [As \$k_{\text{eff}}\$ is one of the parameters that is used to calibrate the model, the range of average thermal conductivity \(\$0.47 - 1.62 \text{ W m}^{-1} \text{ K}^{-1}\$ \) will be used to bound \$k_{\text{eff}}\$.](#)

4.2 Surface roughness (z_0)

The DEMs generated using the SfM workflow had a total root mean square error of $0.008 - 0.024 \text{ m}$. Table 2 shows that the errors in elevation (i.e., z) were smaller than in planform (i.e., x and y) with a maximum error of 0.007 m . The contrast between elevation and planimetric errors is likely a result of the identification of the GCPs in each photo during the SfM workflow, since

Daene McKinney 10/27/15 8:11 PM

Deleted: -

Daene McKinney 10/27/15 9:57 AM

Moved down [2]: One of the limitations with regards to the thermal conductivity measurements is that all the measurements were made near the surface. Therefore, the estimates of the average thermal conductivity at each site are potentially underestimated because the deeper layers that may be more compact and humid are not considered. Interestingly, the thermal conductivities measured at Sites 4 and 11 are similar to those estimated by Nicholson and Benn (2012) for debris cover on Ngozumpa glacier with 10% and 20% of the void space being filled with water (1.42 and $1.55 \text{ W m}^{-1} \text{ K}^{-1}$, respectively). This highlights the importance of understanding how the moisture varies within the debris and its influence on the thermal conductivity. As k_{eff} is one of the parameters that is used to calibrate the model, the range of average thermal conductivity ($0.47 - 1.62 \text{ W m}^{-1} \text{ K}^{-1}$) will be used to bound k_{eff} . The influence of the selection of these bounds on the model calibration will be discussed in the model calibration section. -

1 it was easier to identify the top of the cone in each photo than it was to determine the exact point
2 on the rim of the cone. As the error with the total station is small (maximum of 0.4 mm), this
3 human error likely dominated the total error, although errors in estimates of both camera position
4 and orientation will also have contributed. The DEMs were resampled to a resolution of 0.01 m
5 such that their resolution was on the same order as their respective errors (3 of the 4 sites had a
6 total RMSE less than 0.01 m). The DEMs were then de-trended to account for variations in the
7 local topography.

8 Initially, z_0 was estimated from Equation 1 using the Lettau-Munro method. The average value
9 of z_0 was 0.0037, 0.0091, 0.0022, and 0.0033 m for Sites A, B, C, and D, respectively. These
10 values are towards the lower end of those previously reported in literature, which were estimated
11 from wind speed profiles and range from 0.0035 to 0.060 m (Inoue and Yoshida, 1980; Takeuchi
12 et al., 2000; Brock et al., 2010). In particular, the average value of z_0 for Sites A, C, and D was
13 comparable or smaller to Area IV on the Khumbu glacier (Inoue and Yoshida, 1980), which
14 comprised small schist with bare ice. Sites A, C, and D were all debris-covered with boulders
15 ranging up to 0.40 m, so these small estimations of z_0 are concerning. Site B also appears to be
16 underestimated as it has similar debris characteristics to Area III on the Khumbu glacier (Inoue
17 and Yoshida), yet its average value of z_0 was much smaller (0.0091 m compared to 0.060 m,
18 respectively). These apparent underestimations of z_0 led to the development of an alternative
19 method.

20 The alternative method relies upon the selection of the obstacle height (or threshold) such that
21 the obstacle density is 30% (Smith, 2014). Figure 4 shows that as the obstacle threshold is
22 increased, the obstacle density decreases, which makes intuitive sense as there will be fewer
23 obstacles in the transect. Table 3 shows the values of z_0 using an obstacle density of 30% and
24 the highest resolution DEM (0.01 m) were 0.016, 0.043, 0.006, and 0.014 m for Sites A-D,
25 respectively. These values agree well with the range of z_0 values previously reported (Inoue and
26 Yoshida, 1980; Takeuchi et al., 2000; Brock et al., 2010). Furthermore, these z_0 values appear to
27 capture the inter-site variability as Site B had the highest value of z_0 (0.043 m), which is expected
28 since the debris cover includes large boulders up to 1 m in size (Figure 2). Site C, which
29 comprised the smallest grain sizes of the four sites in this study had the lowest estimation of z_0
30 (0.006 m). Sites A and D had similar values of surface roughness and the standard deviations of

Daene McKinney 10/27/15 10:35 AM

Deleted:

Daene McKinney 10/27/15 10:36 AM

Deleted: Furthermore,

Daene McKinney 10/27/15 10:37 AM

Deleted: appeared to have

Daene McKinney 10/27/15 10:37 AM

Deleted: is further supports the notion that the values of z_0 from the Lettau-Munro method are likely underestimated, which led to the development of an alternative method.

Daene McKinney 10/27/15 11:31 AM

Deleted: less than

Site A and D (0.008 m and 0.012 m, respectively) appear to capture the more homogenous surface of Site A compared to the highly heterogeneous surface of Site D (Figure 2).

The impact of DEM resolution on obstacle threshold and z_0 was analyzed to determine the robustness of this alternative method. Since the terrain is not changing and only the sampling frequency is varying, the z_0 values should remain fairly constant. Table 3 shows that the obstacle threshold increases as the resolution of the DEM becomes coarser. This occurs because the coarser resolution cannot capture the subtler changes in surface height over the debris cover. On the other hand, the estimations of z_0 remain relatively constant (± 0.004 m) as the DEM resolution is reduced to 0.04 m. The consistency of this method despite variations in DEM resolution and obstacle thresholds, and the objective approach for deriving the obstacle threshold using a 30% obstacle density lends confidence to this method. Furthermore, Nield et al. (2013b) found that measures regarding surface heights are the best predictor of aerodynamic roughness, specifically for surfaces that comprise large elements or have patches of large and small elements. Therefore, it is expected that the obstacle threshold should vary for different sites as a function of their largest elements, which is consistent with the inter-site variability and the obstacle thresholds reported in this study.

4.3 Ablation Stakes

Ablation stakes were installed on 18-19 May 2014 approximately 1 m into the ice at 14 sites with debris thicknesses ranging from 0.07 to 0.54 m (Table 1). The ablation stakes were measured on 09 November 2014. At 11 of the 14 sites, the ablation stakes completely melted out of the ice indicating there was greater than 1 m of ablation. Sites 8, 13, and 15, had ablation measurements of 0.92, 0.85, and 0.89 m, respectively. These three sites had debris thicknesses of 0.20, 0.33, and 0.37 m and were oriented in the southern, northeast, and northwest directions, respectively. The lower ablation rates of Sites 13 and 15 compared to the other 12 sites is likely due to a combination of their debris thickness and aspect as they are oriented in a manner that receives less solar radiation throughout the day. Site 8 appears to be an anomaly as it has a smaller debris thickness than 8 of the sites with ablation stakes and a southerly aspect, which positions it in a manner to receive a greater amount of solar radiation throughout the day. It is possible that Site 8 had a higher albedo and/or a lower thermal conductivity, which would greatly reduce its ablation; unfortunately, these properties could not be measured in the field. Nevertheless, the

Daene McKinney 10/27/15 11:23 AM

Formatted: Not Highlight

Daene McKinney 10/27/15 11:23 AM

Formatted: Not Highlight

Daene McKinney 10/27/15 11:23 AM

Formatted: Not Highlight

Daene McKinney 10/27/15 11:23 AM

Formatted: Not Highlight

Daene McKinney 10/27/15 10:41 AM

Moved down [3]: The robustness of this method was assessed by varying the DEM resolution to assess its effect on the resulting obstacle thresholds and z_0 values. Since the terrain is not changing and only the sampling frequency is varying, the z_0 values should remain fairly constant. 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 of points (~50) for the transect analysis.

Daene McKinney 10/27/15 11:24 AM

Deleted: .

... [1]

1 ablation measurements indicate that understanding ablation rates on debris-covered glaciers is
2 greatly influenced by slope, aspect, and properties of the debris (albedo, thermal conductivity,
3 and surface roughness).

4

5 5 Modeled Results

6 5.1 Model Calibration

7 Three different methods were used to estimate the latent heat flux to determine how well each
8 method models the measured debris temperatures. These methods are referred to as LE_{Rain} ,
9 LE_{RH100} , and LE_{Dry} . The albedo, thermal conductivity, and surface roughness for each of the
10 three methods were optimized by minimizing the sum of squares of the surface temperature for
11 each site (Table 4). For the LE_{Rain} and LE_{RH100} model, 7 of the 10 sites had a thermal
12 conductivity at the upper bound ($1.62 \text{ W m}^{-1} \text{ K}^{-1}$), while for the LE_{Dry} model 9 of the 10 sites
13 were at the upper bound. These results indicate that the selection of the upper bound for the
14 thermal conductivity is important and its impact on model performance is detailed in the
15 discussion section. The albedo values ranged from 0.10 – 0.40 and had an average value around
16 0.32, which is consistent with albedos measured in the Khumbu (Inoue and Yoshida, 1980;
17 Kayastha et al., 2000; Nicholson and Benn, 2012; LeJeune et al., 2013). The values of z_0 had an
18 average around 0.014 m, which is consistent with z_0 measured in this study and those reported on
19 other debris-covered glaciers (Inoue and Yoshida, 1980; Takeuchi et al., 2000; Brock et al.,
20 2010). For the LE_{Rain} and LE_{RH100} models, 5 of the 10 sites had a value of z_0 at its lower bound
21 (0.006 m), which highlights the importance of measuring the surface roughness of the debris
22 cover and will be discussed in the sensitivity analysis.

23 The performance of each model was assessed using the total sum of squares and the R^2
24 correlation coefficients. The R^2 values ranged from 0.34 to 0.92 for all three models. The
25 average R^2 values over the calibration period for the LE_{Rain} , LE_{RH100} , and LE_{Dry} models were 0.72,
26 0.72, and 0.71, respectively. Figures 5C and 5D show the correlation between the modeled and
27 measured surface temperature at Site 11, which had an R^2 of 0.77 and 0.75 for the calibration and
28 validation periods, respectively. Figure 5C shows there is good agreement between the modeled
29 and measured temperature sensors. The modeled temperatures appear to capture the daily

Daene McKinney 10/27/15 9:51 AM

Moved down [1]: An additional calibration was performed allowing the thermal conductivity to be unbounded and found 3 or more out of the 10 sites for each method had thermal conductivities greater than $3.0 \text{ W m}^{-1} \text{ K}^{-1}$ with one thermal conductivity as high as $4.5 \text{ W m}^{-1} \text{ K}^{-1}$. The lithology of the debris cover in the Everest region is predominantly granite, gneiss, and pelite (Hambrey et al., 2008). Robertson (1988) found the thermal conductivity of solid granite gneiss to be $2.87 \text{ W m}^{-1} \text{ K}^{-1}$, so the unbounded thermal conductivities do not appear to make physical sense when one considers that the thermal conductivity of debris should be much lower than solid rock due to the pore spaces being filled with air and water. Furthermore, an optimization performed using the total sum of squares of all the surface sites reveals that increasing the thermal conductivity from $1.6 \text{ W m}^{-1} \text{ K}^{-1}$ to its minimum of $2.6 \text{ W m}^{-1} \text{ K}^{-1}$ only reduces the total sum of squares by 3%. Therefore, the results reported in this study use an upper bound of $1.62 \text{ W m}^{-1} \text{ K}^{-1}$ and the importance of accurately measuring the thermal conductivity will be discussed in the sensitivity analysis.

Daene McKinney 10/27/15 10:03 AM

Deleted:

Daene McKinney 10/27/15 10:05 AM

Deleted: -

Daene McKinney 10/27/15 10:27 AM

Deleted: .

Daene McKinney 10/27/15 10:27 AM

Deleted: This

1 variations in temperature well. However, there are a few days for which a positive bias in
2 temperature can be seen during the daily high and nightly low (e.g., Figure 5C from 16-18 June
3 and 25-27 July). Interestingly, the overestimation of the daily high typically occurs after the
4 nightly low has a positive bias in temperature during the previous night. The positive bias of the
5 nightly minimum is apparent between the hours of 0:00 and 6:00 (Figure 6). One possible
6 explanation for the positive bias in temperature in the nightly low is an overestimation of the
7 incoming longwave radiation due to the poor temporal and spatial resolution of the
8 NCEP/NCAR reanalysis dataset compared to the other meteorological data from Pyramid Station.
9 Typically, the wind speed during the night is relatively low thereby limiting the turbulent heat
10 fluxes, which causes the incoming longwave radiation to be a major source of energy during this
11 time.

12 Nonetheless, the model performs reasonably well for all of the temperature sensors.
13 Unfortunately, it is difficult to determine which latent heat flux model performs the best using
14 the total sum of squares and/or the R^2 values as there was not one particular model that
15 consistently had a lower total sum of squares and/or a higher R^2 at each site. The average R^2
16 value was fairly comparable for all three models. The total sum of squares of all the sites was
17 the lowest for the LE_{RH100} model followed by the LE_{Rain} model and then the LE_{Dry} model, but the
18 difference between models was less than 5%.

19 5.2 Model Validation

20 Model validation was assessed from 08 August to 12 October 2014 for all three models using the
21 R^2 values for each temperature sensor. The R^2 values for all the temperature sensors at the 10
22 sites ranged from 0.39 to 0.81 for all three methods. The average R^2 value for the LE_{Rain} ,
23 LE_{RH100} , and LE_{Dry} was 0.67, 0.67, and 0.68, respectively. Again, the similar performance
24 between the three models does not provide any insight into preference for one model and is
25 likely a result of the calibration procedure. Figure 5D shows that the LE_{Rain} model performs well
26 through the entire validation period. Similar to the calibration period, the LE_{Rain} model appears
27 to underestimate the nightly low, which causes the following daily high to be overestimated.

28 Reid and Brock (2010) found R^2 values of 0.94 and 0.52 for temperature sensors at the surface
29 and at a depth of 15 cm, respectively. While the R^2 value of 0.94 is higher than those found in
30 this study, the range of R^2 is comparable. In contrast to the findings of Reid and Brock (2010),

Daene McKinney 10/27/15 8:22 PM

Deleted: (see for example

Daene McKinney 10/27/15 12:10 PM

Deleted: The lack of a single model clearly outperforming the others indicates that either a) the modeling of the latent heat flux is insignificant or b) the latent heat flux is significant, but the calibration procedure allows for changes in the latent heat flux to be compensated for via other model param ... [2]

1 the average R^2 value for the surface temperature sensors (0.67-0.68 for all three models) was
2 very similar to the average R^2 value of those buried in the debris (0.66-0.68 for all three models).
3 The slightly lower R^2 values in this study may be a result of using meteorological data from an
4 AWS located 14 km away from the glacier. Furthermore, longwave radiation was estimated
5 from remotely sensed data, which may also influence model performance as previously
6 discussed.

7 5.3 Modeled Ablation Rates

8 Ablation rates were computed for all 15 sites that had a temperature sensor or an ablation stake.
9 For sites that only had an ablation stake, the average calibrated parameters for that particular
10 latent heat flux model were used. Additionally, ablation rates were estimated for the LE_{Rain}
11 model using the average calibrated parameters for all the sites to assess the differences between
12 using a single set of parameters compared to optimizing the parameters at each site. The
13 modeled ablation over the entire duration of the study period varied from 0.39 to 2.85 m among
14 the three methods (Figure 7). On average the LE_{Dry} model overestimated both the LE_{Rain} and
15 LE_{RH100} models by 7.9%. The slight variations in ablation between the models are directly
16 related to the differences in their calibrated parameters. The slightly higher ablation rates for the
17 LE_{Dry} model is likely attributed to the higher values of thermal conductivities and the lack of a
18 latent heat flux term to remove heat from the debris. Figure 7 shows there is a clear relationship
19 between debris thickness and ablation as thin debris has higher rates of ablation compared to
20 thicker debris, which insulates the ice to a greater extent thereby retarding ablation. The scatter
21 found throughout the curve, specifically between 0.25 and 0.50 m, is due to the site-specific
22 debris properties and the slope and aspect of each site. A comparison between the LE_{Rain} model
23 using the optimized parameters at each site and those using the average calibrated parameters at
24 each site highlights the effect that site-specific properties has on ablation. Site 6, with a debris
25 thickness of 0.08 m, is a good example as the use of average calibrated parameters increased the
26 melt from 2.03 m to 2.70 m due to an increase in thermal conductivity from 1.29 $W m^{-1} K^{-1}$ to
27 1.52 $W m^{-1} K^{-1}$. These differences in melt and the sensitivity to thermal conductivity highlight
28 the importance of properly estimating/measuring the thermal conductivity of the debris cover.

29 The modeled ablation rates may also be compared to the measured ablation rates. Specifically,
30 Sites 8, 13, and 15 had measured ablation rates of 0.92, 0.85 and 0.89 m compared to their

Daene McKinney 10/27/15 9:00 AM

Deleted: Ablation rates were modeled over the same time period as the ablation stakes (18 May to 09 November). For days where no meteorological data was available, i.e., the data gaps, the ablation for that day was assumed to be equal to the daily ablation rate for that specific month. As the available meteorological data began on 31 May, the daily ablation rate for the month of May was assumed to be equal to the daily ablation rate of the first week of June. Temperature sensors revealed the debris was snow covered from 26 May to 01 June, so the melting during these days was assumed to be zero. Temperature profiles also show the debris was snow covered from 13-20 October and deeper thermistors revealed the temperature remained around freezing until the sensors were removed in November. Therefore, the melt rates after the 12 October were assumed to be zero.

Daene McKinney 10/26/15 10:36 PM

Formatted: Highlight

Daene McKinney 10/2/15 3:17 PM

Deleted: 8.3

Daene McKinney 10/2/15 3:18 PM

Deleted: 1.85

Daene McKinney 10/2/15 3:18 PM

Deleted: 0

Daene McKinney 10/2/15 3:19 PM

Deleted: 0

Daene McKinney 10/27/15 12:17 PM

Moved down [4]: The thermal conductivity could be better resolved by a) accurately measuring the depth of the temperature sensors during installation and retrieval, b) installing additional sensors (e.g. 5 cm spacing) that allow thermal conductivity within the debris to be computed at more depths, and c) measuring moisture in the debris at various depths.

modeled ablation rates of 1.76, 0.76, and 1.22 m, respectively, for the LE_{Rain} model. The large discrepancy between the measured and modeled ablation rates at Site 8 may be due to the lack of knowledge of the debris properties at Site 8 as previously discussed. The difference between the modeled and measured ablation rates at Site 15 may also be a result of the thermal conductivity parameter ($1.61 \text{ W m}^{-1} \text{ K}^{-1}$), which is slightly higher than thermal conductivities previously reported in the Khumbu, which ranged from 0.60 to $1.29 \text{ W m}^{-1} \text{ K}^{-1}$. A comparison of the daily average temperatures for Site 15 reveals there was about an hour lag between the modeled and measured temperatures (results not shown). Lags between temperatures are typically a result of their depth (Conway and Rasmussen, 2000), which is apparent in Figure 5A as the 0.10 m sensor lags behind the 0.01 m sensor. It is possible that debris may have shifted over the melt season causing the measured temperature to be at a lower depth than 0.01 m, which would greatly influence the model calibration and potentially cause the thermal conductivity to vary. Site 5 was the only other site where a slight lag was observed between the measured and modeled temperatures. The modeled ablation at Site 5 was 0.87 m using the LE_{Rain} model, while the ablation stake melted completely out of the ice indicating greater than 1 m of ablation. All the other model estimates of ablation were near to or greater than 1 m, which was also observed by their respective ablation stakes as they completely melted out of the ice.

The ablation results also show strong seasonal trends with maximum melt rates occurring in June, July, and August. These ablation rates appear to taper off towards the transition seasons. Melt rates in July and August ranged from $0.3 - 2.5 \text{ cm day}^{-1}$ based on the debris thickness, which is consistent with empirical relationships between mean daily ablation rate and debris thickness found on other glaciers (Nicholson and Benn, 2006). The total ablation rates are also similar to those measured on Imja-Lhotse Shar glacier using multiple DEMs, which were $1.45 \pm 0.52 \text{ m yr}^{-1}$ (Bolch et al., 2011), thereby lending confidence to the results.

5.4 Sensitivity Analysis

A sensitivity analysis was performed to assess how albedo, thermal conductivity, and surface roughness affect the total ablation (Table 5) based on the uncertainty with respect to each parameter. The uncertainty in thermal conductivity was $\pm 0.40 \text{ W m}^{-1} \text{ K}^{-1}$ as described above. The uncertainty associated with the surface roughness was $\pm 0.010 \text{ m}$, which is the approximate standard deviation associated with the z_0 values for each of the three models (Table 4) and

1 | similar to the standard deviation between the four sites where z_0 was measured (± 0.016 m).
2 | Lastly, the uncertainty of the albedo was estimated as ± 0.10 , which is the approximate standard
3 | deviation within the model calibration for each of the three models and also the difference
4 | between the mean and median albedo measured by Nicholson and Benn (2012) on Ngozumpa
5 | glacier. The LE_{Rain} model was used as the baseline case and the average value for each of the
6 | calibrated parameters (α , k , z_0) from the model optimized is used for each site.

7 | Table 5 shows the total ablation is most sensitive to changes in the thermal conductivity, where a
8 | $\pm 0.40 \text{ W m}^{-1} \text{ K}^{-1}$ change causes a $\pm 20.5\%$ change in total ablation on average. The uncertainty
9 | associated with the thermal conductivity is also more sensitive to thicker debris, which is
10 | consistent with the findings of Nicholson and Benn (2012). Total ablation is also moderately
11 | sensitive to changes in the albedo, where a ± 0.10 change causes a $\pm 12.0\%$ change in total
12 | ablation. Lastly, the total ablation is least sensitive to changes in increasing the surface
13 | roughness, as a $+0.010$ m increase in z_0 only caused a -7.3% change in total ablation. However,
14 | the model was quite sensitive to a reduction in the z_0 of -0.010 m, which caused an average
15 | change in total ablation of $+15.0\%$. The sensitivity associated with z_0 also appears to increase
16 | with an increase in debris thickness. These results highlight the importance of properly
17 | estimating the thermal conductivity, but also show the surface roughness and the albedo are
18 | important as well.

19 | 5.5 Temporal Resolution

20 | Nicholson and Benn (2006) proposed that the temperature gradient in the debris may be assumed
21 | to be linear at a time step greater than a day, but is nonlinear for shorter time steps. This would
22 | have important implications for modeling melt on remote debris-covered glaciers where
23 | meteorological data is not available and reanalysis datasets could be used instead. The
24 | importance of temporal resolution was analyzed using 6-hour and daily average data from
25 | Pyramid Station, which are consistent with the temporal resolution of NCEP/NCAR reanalysis
26 | datasets. To be consistent with this reanalysis dataset such that only the effects of temporal
27 | resolution were analyzed, the wind speed and relative humidity used were instantaneous values
28 | from Pyramid Station, while all the other variables were 6-hour averages. For the daily time step,
29 | all the parameters were daily averages and the temperature profile in the debris is assumed to be
30 | linear. The LE_{Rain} model was used to model the latent heat flux.

Daene McKinney 10/15/15 8:17 AM
Deleted: 3

Daene McKinney 10/27/15 8:18 PM
Deleted: is

Daene McKinney 10/2/15 3:21 PM
Deleted: 6

Daene McKinney 10/2/15 3:21 PM
Deleted: 6

Daene McKinney 10/2/15 3:21 PM
Deleted: 9

1 The R^2 correlation coefficients for the sites with temperature sensors and the modeled total melt
2 for all 15 sites were used to assess the effect of temporal resolution on model performance. The
3 R^2 using the 6-hour data ranged from 0.30 to 0.80 with an average of 0.55 over the calibration
4 period and was significantly poorer during the validation period with R^2 values ranging from
5 0.15 to 0.65 with an average of 0.35. Figure 8A shows the surface temperature at Site 11 does
6 fairly well ($R^2 = 0.63$) at modeling the measured surface temperatures over the calibration period.
7 The lower R^2 values compared to the 30 minute time step appears to be a result of the 6-hour
8 model underestimating the daily high, which occurs around 15:00 each day. Furthermore, Figure
9 8B shows the 6-hour model poorly replicates the measured data towards the transition seasons
10 when snowfall becomes significant, which explains the poorer R^2 values for the validation period.
11 Snowfall is problematic in the model for large time steps because the model assumes the snow is
12 on the surface for the entire time step. Therefore, a small snow event that could melt quickly on
13 the debris and then allow the debris to warm up during the day is perceived to remain on the
14 snow for the 6-hour time step (e.g., Figure 8B from 27 Sept to 03 October). The same problem
15 arises at the daily time step, so a snow-free model was used instead. For the daily time step, the
16 R^2 values ranged from 0.18 to 0.63 with an average of 0.29. Figure 8C shows the daily time step
17 is able to capture some of the temperature fluctuations over the melt season, but does not
18 perform as well as the 30-minute or 6-hour models.

19 Since we are most interested in understanding the effects of temporal resolution, the 6-hour data
20 and daily averages from 02 June to 25 September 2014 were assessed, which is prior to the time
21 when snowfall was recorded each day. A comparison of all the modeled and measured
22 temperatures at the surface reveals the 6-hour model underestimates the measured temperatures
23 by an average of 1.0 (± 4.3) K over the entire time period. The modeled total ablation from 02
24 June to 25 September reveals the ablation is consistently underestimated at all sites by an
25 average of 11 (± 5) %. The lower estimates of ablation are likely a result of the underestimation
26 of the daily high as previous discussed. Similar to the 6-hour model, the daily time step model
27 slightly underestimates the measured temperatures at the surface on average by 0.3 (± 1.9) K.
28 The modeled total ablation is also underestimated by an average of 6 (± 10) %. However, it is
29 important to note that 5 of the 15 sites are actually slightly overestimated the melt. These results
30 suggest that if high-temporal meteorological data are not available, a first-order estimate of
31 ablation over a melt season could be obtained using the daily time step model. It is important

that the estimate is made over the entire melt season, as the daily model does not capture the daily temperature fluctuations well. Furthermore, caution should be used to avoid the transition seasons as both the daily model and the 6-hour model do not have a small enough time step to properly account for snow melt.

6 Discussion

6.1 Thermal Conductivity

One of the limitations with regards to the thermal conductivity measurements is that all the measurements were made near the surface. Therefore, the estimates of the average thermal conductivity at each site are potentially underestimated because the deeper layers that may be more compact and humid are not considered. The lack of any trends with respect to depth appears to dispute this theory; however, this is based on a limited number of measurements near the surface. Interestingly, the thermal conductivities measured at Sites 4 and 11 are similar to those estimated by Nicholson and Benn (2012) for debris cover on Ngozumpa glacier with 10% and 20% of the void space being filled with water (1.42 and $1.55 \text{ W m}^{-1} \text{ K}^{-1}$, respectively).

The number of sites that reached the upper bound during the model calibration is concerning as it may indicate that the actual thermal conductivity throughout the debris is higher. To address this issue, an additional calibration was performed allowing the thermal conductivity to be unbounded. This calibration revealed that 3 or more out of the 10 sites for each method had thermal conductivities greater than $3.0 \text{ W m}^{-1} \text{ K}^{-1}$ with one thermal conductivity as high as $4.5 \text{ W m}^{-1} \text{ K}^{-1}$. The lithology of the debris cover in the Everest region is predominantly granite, gneiss, and pelite (Hambrey et al., 2008). Robertson (1988) found the thermal conductivity of solid granite gneiss to be $2.87 \text{ W m}^{-1} \text{ K}^{-1}$, so the unbounded thermal conductivities do not appear to make physical sense when one considers that the thermal conductivity of debris should be much lower than solid rock due to the pore spaces being filled with air and water. Furthermore, an optimization performed using the total sum of squares of all the surface sites reveals that increasing the thermal conductivity from $1.6 \text{ W m}^{-1} \text{ K}^{-1}$ to its minimum of $2.6 \text{ W m}^{-1} \text{ K}^{-1}$ only reduces the total sum of squares by 3%. These results and the similar values to Nicholson and Benn (2012) lend confidence to the use of $1.62 \text{ W m}^{-1} \text{ K}^{-1}$ as the upper bound, but highlights the

Daene McKinney 10/27/15 9:57 AM
Moved (insertion) [2]

Daene McKinney 10/27/15 10:09 AM
Deleted:

Daene McKinney 10/27/15 10:10 AM
Deleted:

Daene McKinney 10/27/15 10:18 AM
Deleted: This highlights the importance of understanding how the moisture varies within the debris and its influence on the thermal conductivity. As k_{eff} is one of the parameters that is used to calibrate the model, the range of average thermal conductivity ($0.47 - 1.62 \text{ W m}^{-1} \text{ K}^{-1}$) will be used to bound k_{eff} . The influence of the selection of these bounds on the model calibration will be discussed in the model calibration section. .

Daene McKinney 10/27/15 9:51 AM
Moved (insertion) [1]

Daene McKinney 10/27/15 10:22 AM
Deleted: An

Daene McKinney 10/27/15 10:22 AM
Deleted: and

Daene McKinney 10/27/15 10:23 AM
Deleted: found

Daene McKinney 10/27/15 10:24 AM
Deleted:

1 importance of understanding how the moisture varies within the debris and its influence on the
2 thermal conductivity. Future work should improve measurements of the thermal conductivity by
3 a) accurately measuring the depth of the temperature sensors during installation and retrieval, b)
4 installing additional sensors (e.g. 5 cm spacing) that allow thermal conductivity within the debris
5 to be computed at more depths, and c) measuring moisture in the debris at various depths.

6 **6.2 Surface Roughness**

7 The development of an alternative method for estimating z_0 was required, as the Lettau-Munro
8 method appeared to greatly underestimate the values of z_0 . The alternative method applies the
9 relationship developed by Lettau (1969) to a high resolution DEM using the selection of an
10 obstacle height (threshold) based on an obstacle density of 30%. One of the main limitations of
11 this study is the lack of aerodynamic roughness measurements to validate the developed methods.
12 Previous work, e.g., Rees and Arnold (2006), has relied upon surface roughness estimates from
13 other studies to assess the reasonableness of their results when aerodynamic data were not
14 collected. This study relies upon the results of Inoue and Yoshida (1980), which estimated
15 surface roughness using wind speed profiles at two sites on the Khumbu glacier. Specifically,
16 Sites B and C in this study have similar debris cover to Areas III and IV from Inoue and Yoshida
17 (1980), respectively.

18 Site B had the highest value of z_0 (0.043 m) of the four sites in this study and consisted of larger
19 boulders up to 1 m in size. This value is similar to the higher value of 0.060 m for z_0 derived
20 from a region on the Khumbu glacier that consisted of large granitic boulders of 1-2 m in size
21 lying on top of schistose rocks with a grain size varying from a few centimeters to 0.5 m (Inoue
22 and Yoshida, 1980). The larger boulders observed by Inoue and Yoshida (1980) may explain the
23 slightly higher value of z_0 compared to Site B. Site C, which comprised the smallest grain sizes
24 of the four sites in this study, agrees well with the smaller value of z_0 (0.0035 m) derived by
25 Inoue and Yoshida (1980) for an area where the supraglacial debris comprised dispersed
26 boulders ranging in size of 0.01 – 0.05 m. The few boulders ranging in size of up to 0.15 m may
27 be the reason for Site C's slightly larger value of z_0 (0.006 m). Sites A and D were composed of
28 boulders and grains that varied in size between those found in Sites B and C; therefore, we deem
29 the value of z_0 of 0.016 and 0.014 m for Sites A and D, respectively, to be reasonable.
30 Furthermore, these values agree fairly well with the z_0 of 0.016 m measured by Brock et al.

Daene McKinney 10/27/15 10:25 AM

Deleted: Therefore, the results reported in this study use an upper bound of $1.62 \text{ W m}^{-1} \text{ K}^{-1}$ and the importance of accurately measuring the thermal conductivity will be discussed in the sensitivity analysis. .

(2010) on a debris-covered glacier in Italy that comprised a mixture of granites and schists of predominantly cobble size, with occasional boulders of < 1 m size.

Future work should seek to compare these estimates of surface roughness with aerodynamic roughness to determine the scale at which these two values agree. Brock et al. (2006) found there to be no significant difference between the use of a 3 m and 15 m transect; however, they did state that a shorter pole would be unlikely to capture a sufficient sample of roughness elements if the vertical changes are greater than 1 m. The use of hundreds of transects over a ~ 4 m² grid has the benefit of expanding the number of surface roughness elements that can be captured compared to a single transect. However, Brock et al. (2006) was comparing microtopographic and aerodynamic roughness over snow, slush, and ice, which is significantly different from the hummocky and heterogeneous terrain on debris-covered glaciers. Therefore, it will be important to determine the scale or fetch length at which the surface roughness agrees with the aerodynamic roughness. Nonetheless, the method developed in this paper provides an objective approach to select an obstacle height and yields consistent and reasonable estimates of z_0 for various grain sizes independent of the resolution of the DEM.

6.3 Modeled Results

One of the limitations of the calibration procedure is the LE_{Rain} , LE_{RH100} , and LE_{Dry} models all performed reasonably well. The lack of a single model clearly outperforming the others indicates that either a) the modeling of the latent heat flux is insignificant or b) the latent heat flux is significant, but the calibration procedure allows for changes in the latent heat flux to be compensated for via other model parameters. Brock et al. (2010) found that latent heat fluxes may be a significant energy sink when rain falls on warm debris indicating that the latent heat flux is important to include. They also assessed the importance of each component of the energy balance and found that inclusion of the latent heat flux improved the correlation coefficient of their model. The average latent heat flux for both the LE_{Rain} and LE_{RH100} models were comparable with values ranging from -53 to 10 W m⁻² over the day. The peak instantaneous latent heat fluxes varied greatly between the two models with fluxes as high as -714 and -323 W m⁻², for the LE_{Rain} and LE_{RH100} models, respectively. These values are similar to those reported by Brock et al. (2010) and support the importance of including the latent heat flux term. However, they do not yield any insight into preference between the LE_{Rain} or LE_{RH100} models.

Daene McKinney 10/27/15 10:41 AM
Formatted: Normal

Daene McKinney 10/27/15 9:31 AM
Formatted: Normal

These results suggest that the selection of the LE_{Rain} or LE_{RH100} model should be based on data availability. Future work should seek to measure the thermal conductivity, albedo, and surface roughness, which would allow the differences between models to be evaluated. Furthermore, detailed knowledge of the debris properties, including how the thermal conductivity and water content vary with depth, would allow the performance of these models to be compared to other debris-covered glacier energy balance models (Collier et al., 2014; Fujita and Sakai, 2014).

7 Conclusions

Debris thickness greatly impacts ablation rates on debris-covered glaciers; however, incorporating debris cover into energy balance models is still hampered by a lack of knowledge of the debris properties. Fieldwork performed on Imja-Lhotse Shar glacier over the 2014 melt season was used to develop new techniques to measure surface roughness, which yielded reasonable values for various grain sizes. Temperature sensors and ablation stakes installed in the debris were also used to assess the performance of a debris-covered glacier energy balance model using three different methods for estimating the latent heat flux. All three models performed well, as a result of the calibration procedure, which allowed variations in the lack of latent heat flux to be compensated for by adjusting the debris properties. However, the LE_{Rain} and LE_{RH100} models yielded more reasonable values of latent heat fluxes. This suggests that in a data-scarce region either the LE_{Rain} or LE_{RH100} model may be used if relative humidity or precipitation data are available.

A sensitivity analysis revealed ablation rates were most sensitive to variations in thermal conductivity, followed closely by albedo and surface roughness. This highlights the importance of measuring the thermal conductivity and the moisture content in the debris. The effect of temporal resolution on model performance was also explored using a 6-hour time step and a daily time step. The 6-hour time step was found to underestimate the daily high each day, which caused the ablation rates to also be slightly underestimated. The daily time step did not model the daily average temperature as well, but yielded better estimates of ablation over the entire melt season.

Future studies should continue to work on incorporating the water content in the debris into debris-covered glacier energy balance models and determine its effect on thermal conductivity

Daene McKinney 10/27/15 12:17 PM

Moved (insertion) [4]

Daene McKinney 10/27/15 12:18 PM

Deleted: The thermal conductivity could be better resolved by a) accurately measuring the depth of the temperature sensors during installation and retrieval, b) installing additional sensors (e.g. 5 cm spacing) that allow thermal conductivity within the debris to be computed at more depths, and c) measuring moisture in the debris at various depths.

Daene McKinney 10/27/15 10:41 AM

Moved (insertion) [3]

Daene McKinney 10/27/15 12:04 PM

Deleted: The robustness of this method was assessed by varying the DEM resolution to assess its effect on the resulting obstacle thresholds and z_0 values. Since the terrain is not changing and only the sampling frequency is varying, the z_0 values should remain fairly constant. 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 of points (~50) for the transect analysis.

... [3]

and the latent heat flux. Furthermore, an increased understanding of how the albedo may vary over the course of the day, the course of the melt season, and as a function of debris saturation, may significantly improve model performance. Lastly, the methods developed in this study have the potential to be scaled up such that maps of surface roughness on a whole glacier scale may be developed in the future, but it is imperative to determine the scale at which the surface roughness and aerodynamic roughness agree with one another.

Acknowledgements

The authors acknowledge the support of the USAID Climate Change Resilient Development (CCRD) project for the support of Rounce. The meteorological data from Pyramid Station used in this study was collected within the SHARE Project thanks to contributions from the Italian National Research Council and the Italian Ministry of Foreign Affairs. The NCEP Reanalysis data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado USA from their website <http://www.esrl.noaa.gov/psd/> was also very helpful. We also acknowledge the support of Dr. Dhananjay Regmi of Himalayan Research Expeditions for temperature sensors and logistical support during fieldwork.

References

- Benn, D.I., Bolch, T., Hands, K., Gulley, J., Luckman, A., Nicholson, L.I., Quincey, D., Thompson, S., Toumi, R., Wiseman, S.: Response of debris-covered glaciers in the Mount Everest region to recent warming, and implications for outburst flood hazards, *Earth-Science Reviews*, 114:156–174, 2012.
- Bolch, T., Pieczonka, T., and Benn, D.I.: Multi-decadal mass loss of glaciers in the Everest area (Nepal Himalaya) derived from stereo imagery, *The Cryosphere*, 5:349-358, 2011.
- Brock, B.W., Willis, I.C., and Sharp, M.J.: Measurement and parameterization of aerodynamic roughness length variations at Haut Glacier d’Arolla, Switzerland., *J. Glaciol.*, 52(177):281-297, 2006.
- Brock, B.W., Mihalcea, C., Kirkbride, M.P., Diolaiuti, G., Cutler, M.E.J. and Smiraglia, C.: Meteorology and surface energy fluxes in the 2005–2007 ablation seasons at the Miage debris-

Daene McKinney 10/27/15 6:45 AM

Deleted: (Collier et al., 2014)

1 covered glacier, Mont Blanc Massif, Italian Alps. *J. Geophys. Res.*, 115, D09106,
2 doi:10.1029/2009JD013224, 2010.

3 Collier, E., Nicholson, L.I., Brock, B.W., Maussion, F., Essery, R., and Bush, A.B.G.:
4 Representing moisture fluxes and phase changes in glacier debris cover using a reservoir
5 approach, *The Cryosphere*, 8:1429-1444, doi:10.5194/tc-8-1429-2014, 2014.

6 Conway, H. and Rasmussen, L.A.: Summer temperature profiles within supraglacial debris on
7 Khumbu Glacier, Nepal, *Debris-Covered Glaciers*, Proceedings of a workshop held at Seattle,
8 Washington, USA, September 2000.

9 Foster, L.A., Brock, B.W., Cutler, M.E.J., and Diotri, F.: Instruments and methods: A physically
10 based method for estimating supraglacial debris thickness from thermal band remote-sensing
11 data, *J. Glaciol.*, 58(210):677-691, doi:10.3189/2012JoG11J194, 2012.

12 Fujita, K. and Sakai, A.: Modelling runoff from a Himalayan debris-covered glacier, *Hydrol.*
13 *Earth Syst. Sci.*, 18:2679-2694, doi:10.5194/hess-18-2679-2014, 2014.

14 Fyffe, C.L., Reid, T.D., Brock, B.W., Kirkbride, M.P., Diolaiuti, G., Smiraglia, C., and Diotri,
15 F.: A distributed energy-balance melt model of an alpine debris-covered glacier, *J. Glaciol.*,
16 60(221):587-602, doi:10.3189/2014JoG13J148, 2014.

17 Hambrey, M.J., Quincey, D.J., Glasser, N.F., Reynolds, J.M., Richardson, S.J., and Clemmens,
18 S.: Sedimentological, geomorphological and dynamic context of debris-mantled glaciers, Mount
19 Everest (Sagarmatha) region, Nepal, *Quaternary Sci Rev*, 27, 2361-2389, 2008.

20 Inoue, J. and Yoshida, M.: Ablation and heat exchange over the Khumbu Glacier, Seppyo,
21 41:26-31, 1980.

22 Kalnay, E., Kanamitsu, M., Kistler, R. et al.: The NCEP/NCAR 40-year reanalysis project,
23 *Bulletin of the American Meteorological Society*, 77, 437-470, 1996.

24 Kayastha, R.B., Takeuchi, Y., Nakawo, M. and Ageta, Y.: Practical prediction of ice melting
25 beneath various thickness of debris cover on Khumbu Glacier, Nepal, using a positive degree-
26 day factor, *Int. Assoc. Hydrol. Sci. Publ.*, 264, 71–81, 2000.

1 Lejeune, Y., Bertrand, J., Wagnon, P., and Morin, S.: A physically based model of the year-
2 round surface energy and mass balance of debris-covered glaciers, *J. Glaciol.*, 49(214): 327-344,
3 doi:10.3189/2013JoG12J149, 2013.

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

6 Munro, D.S.: Surface roughness and bulk heat transfer on a glacier: comparison with eddy
7 correlation, *J. Glaciol.*, 35(121):343-348, 1989.

8 Nakawo, M. and Young, G.J.: Field experiments to determine the effect of a debris layer on
9 ablation of glacier ice, *Ann. Glaciol.*, 2:85-91, 1981.

10 Nicholson, L. and Benn, D.I.: Calculating ice melt beneath a debris layer using meteorological
11 data, *J. Glaciol.*, 52:463–470, 2006.

12 Nicholson, L. and Benn, D.I.: Properties of natural supraglacial debris in relation to modelling
13 sub-debris ice ablation, *Earth Surface Processes and Landforms*, 38(5):490-501,
14 doi:10.1002/esp.3299, 2012.

15 Nield, J.M., Chiverrell, R.C., Darby, S.E., Leyland, J., Vircavs, L.H., and Jacobs, B.: Complex
16 spatial feedbacks of tephra redistribution, ice melt and surface roughness modulate ablation on
17 tephra covered glaciers, *Earth Surf. Process. Landforms*, 38, 95-102, doi:10.1002/esp.3352,
18 2013a.

19 Nield, J.M., King, J., Wiggs, G.F.S., Leyland, J., Bryant, R.G., Chiverrell, R.C., Darby, S.E.,
20 Eckardt, F.D., Thomas, D.S.G., Vircavs, L.H., and Washington, R.: Estimating aerodynamic
21 roughness over complex surface terrain, *Journal of Geophysical Research: Atmospheres*, 118, 12,
22 948-12, 961, doi:10.1002/2013JD020632, 2013b.

23 Østrem, G.: Ice Melting under a thin layer of moraine, and the existence of ice cores in moraine
24 ridges, *Geografiska Annaler*, 41(4):228-230, 1959.

25 Rahimi, C.J. and Konrad, J-M.: Thermal conductivity of compacted snow. *Cold Regions*
26 *Engineering: Sustainable Infrastructure Development in a Changing Cold Climate*, ASCE, 833-
27 843, 2012.

1 Rees, W.G. and Arnold, N.S.: Scale-dependent roughness of a glacier surface: implications for
2 radar backscatter and aerodynamic roughness modelling, *J. Glaciol.*, 52(177):214-222, 2006.

3 Reid, T.D. and Brock, B.W.: An energy-balance model for debris-covered glaciers including heat
4 conduction through the debris layer, *J. Glaciol.*, 56(199):903-916, 2010.

5 Reid, T.D., Carenzo, M., Pellicciotti, F. and Brock, B.W.: Including debris cover effects in a
6 distributed model of glacier ablation, *J. Geophys. Res.*, 117(D18), D18105,
7 doi:10.1029/2012JD017795, 2012.

8 Robertson, E.C.: Thermal Properties of Rocks. US Geological Survey (No. 88-441), 1988.

9 Rounce, D.R., and McKinney, D.C.: Debris thickness of glaciers in the Everest area (Nepal
10 Himalaya) derived from satellite imagery using a nonlinear energy balance model, *The*
11 *Cryosphere*, 8:1317-1329, doi:10.5194/tc-8-1317-2014, 2014.

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

13 Snavely N.: Scene reconstruction and visualization from internet photo collections, Unpublished
14 PhD thesis, University of Washington, USA, 2008.

15 Snavely, N., Seitz, S.M., and Szeliski, R.: Modeling the world from internet photo collections,
16 *Int. J. Comput. Vis.*, 80:189-210, doi:10.1007/s11263-007-0107-3, 2008.

17 Sturm, M., Holmgren, J., König, M., and Morris, K.: The thermal conductivity of seasonal snow,
18 *J. Glaciol.*, 43, 26-41, 1997.

19 Sturm, M., Perovich, D.K., and Holmgren, J.: Thermal conductivity and heat transfer through the
20 snow on the ice of the Beaufort Sea, *J. Geophys. Res.*, 107(C21), 2002.

21 Szeliski, R.: Computer Vision: algorithms and applications. Springer, London, 2011.

22 Takeuchi, Y., Kayastha, R. B. and Nakawo, M.: Characteristics of ablation and heat balance in
23 debris-free and debris-covered areas on Khumbu Glacier, Nepal Himalayas, in the pre-monsoon
24 season, *Int. Assoc. Hydrol. Sci. Publ.*, 264, 53–61, 2000.

25 Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., and Reynolds, J.M.: ‘Structure-
26 from-Motion’ Photogrammetry: A low-cost, effective tool for geoscience applications,
27 *Geomorphology*, 179:300-314, 2012.

- 1 Zhang, Y., Fujita, K., Liu, S., Liu, Q., and Nuimura, T.: Distribution of debris thickness and its
- 2 effect on ice melt at Hailuoguo Glacier, Southeastern Tibetan Plateau, using in situ surveys and
- 3 ASTER imagery, *J. Glaciol.*, 57(206):1147-1157, 2011.

1 Table 1. Details of the debris thickness, topography, and monitoring equipment installed at each
2 site. *Italics* note an estimation of debris thickness. T_s denotes surface temperature.

Site	Debris Thickness (m)	Slope (°)	Aspect (°)	Temperature Sensor Depth (m)	Ablation Stake	z_0 photos
4	<i>1.50</i>	17	232	T_s , 0.10, 0.20, 0.40, 0.83	-	-
5	0.54	24	158	T_s	x	-
6	0.08	37	237	T_s	x	-
7	0.52	31	65	-	x	-
8	0.20	32	187	-	x	-
10	0.07	32	337	-	x	-
11	0.45	32	197	T_s , 0.05, 0.10, 0.20, 0.36	x	-
12	0.15	19	265	-	x	-
13	0.33	29	295	T_s , 0.05, 0.10, 0.20	x	-
14	0.26	23	148	T_s , 0.05, 0.24	x	-
15	0.37	29	40	T_s	x	-
16	0.15	32	264	-	x	-
17	0.27	29	228	T_s	x	-
19	0.37	33	198	T_s	x	-
20	0.20	29	200	T_s	x	-
A	-	-	-	-	-	x
B	-	-	-	-	-	x
C	-	-	-	-	-	x
D	-	-	-	-	-	x

3

1 Table 2. Errors associated with the DEM for each site

Site	DEM Error (m)			
	<i>x</i>	<i>y</i>	<i>z</i>	Total
A	0.015	0.018	0.007	0.024
B	0.004	0.007	0.001	0.008
C	0.010	0.007	0.002	0.012
D	0.006	0.007	0.001	0.009

2

1 Table 3. Surface roughness (z_0) estimates and obstacle thresholds as a function of DEM
 2 resolution (m) at Sites A-D.

Site	DEM Resolution	Obstacle Threshold	z_0 (m)	
	(m)	(m)	avg	std
A	0.01	0.048	0.016	0.008
	0.02	0.052	0.016	0.007
	0.04	0.054	0.015	0.007
B	0.01	0.067	0.043	0.037
	0.02	0.073	0.040	0.029
	0.04	0.078	0.036	0.022
C	0.01	0.024	0.006	0.005
	0.02	0.026	0.006	0.004
	0.04	0.027	0.006	0.006
D	0.01	0.033	0.014	0.012
	0.02	0.037	0.015	0.013
	0.04	0.040	0.014	0.011

Table 4. Optimized values of albedo, thermal conductivity, and surface roughness for three methods of estimating latent heat flux during calibration period.

Site	LE _{Rain}			LE _{RH100}			LE _{Dry}		
	α	k^1	z_0^2	α	k^1	z_0^2	α	k^1	z_0^2
4	0.26	1.62	0.006	0.24	1.62	0.006	0.23	1.62	0.011
5	0.40	1.62	0.014	0.40	1.62	0.013	0.40	1.62	0.017
6	0.40	1.29	0.043	0.40	1.35	0.043	0.40	1.31	0.043
11	0.37	1.62	0.006	0.37	1.62	0.006	0.39	1.62	0.006
13	0.10	0.92	0.025	0.16	1.03	0.015	0.10	1.62	0.021
14	0.39	1.61	0.015	0.40	1.62	0.012	0.39	1.62	0.017
15	0.38	1.62	0.006	0.37	1.62	0.006	0.38	1.62	0.006
17	0.30	1.62	0.006	0.30	1.55	0.006	0.30	1.62	0.006
19	0.33	1.62	0.019	0.37	1.62	0.013	0.30	1.62	0.028
20	0.28	1.62	0.006	0.29	1.62	0.006	0.30	1.62	0.006
Avg	0.32	1.52	0.015	0.33	1.53	0.013	0.32	1.59	0.016
Std	0.09	0.23	0.012	0.08	0.19	0.011	0.10	0.10	0.012

¹units of W m⁻¹ K⁻¹; ²units of m

Daene McKinney 10/2/15 3:33 PM

Deleted:

... [4]

1 Table 5. Sensitivity analysis showing percent changes relative to the total melt (m) as a function
 2 of the uncertainty associated with the calibrated parameters (α , k , z_0) for all sites over the study
 3 period using the LE_{Rain} model in conjunction with the average calibrated parameters for all sites.

Parameter		α		k		z_0	
Adjustment		+ 0.10	- 0.10	+ 0.40	- 0.40	+ 0.010	- 0.010
Site	Total Melt (m)	% Change					
4	0.30	-12.3	+13.0	+30.1	-29.4	-9.7	+21.7
5	0.90	-12.6	+13.0	+22.3	-24.0	-9.2	+19.4
6	2.70	-11.1	+11.3	+13.1	-16.3	-4.1	+7.9
7	0.92	-12.6	+12.6	+22.0	-23.5	-8.8	+18.8
8	1.03	-11.6	+11.9	+20.7	-22.5	-7.7	+16.1
10	1.61	-12.2	+12.1	+18.9	-21.2	-7.7	+15.3
11	1.12	-12.0	+12.5	+20.3	-22.1	-7.7	+16.2
12	1.30	-11.8	+12.0	+19.8	-21.9	-7.8	+16.1
13	1.05	-12.4	+12.7	+21.2	-22.9	-8.5	+17.8
14	1.72	-11.9	+11.8	+18.1	-20.7	-7.0	+14.1
15	0.90	-12.4	+12.9	+21.8	-23.4	-8.6	+17.7
16	1.76	-11.6	+12.1	+18.3	-20.2	-6.7	+14.0
17	2.77	-10.8	+11.1	+12.0	-15.0	-2.9	+5.7
19	2.04	-11.6	+11.7	+16.5	-19.4	-6.3	+12.3
20	1.81	-11.4	+11.4	+16.5	-19.4	-6.1	+12.1
Average		-11.9	+12.1	+19.4	-21.5	-7.3	+15.0

5

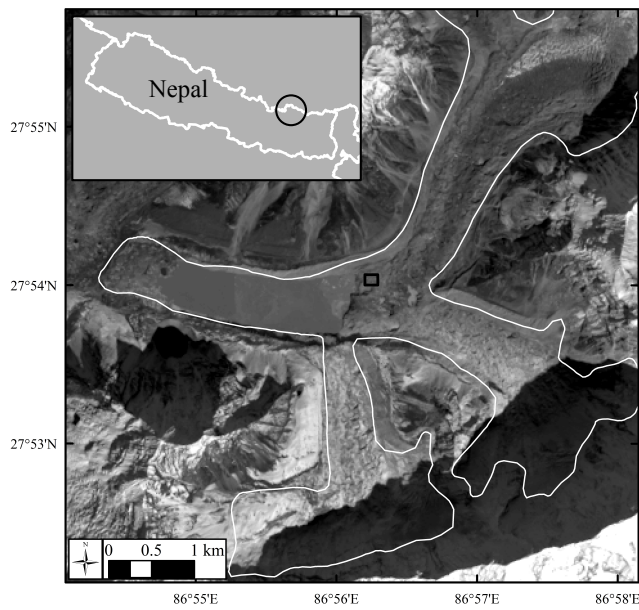
Daene McKinney 10/2/15 3:41 PM

Deleted:

... [S]

Daene McKinney 10/27/15 8:14 PM

Formatted Table



1
2 Figure 1. Landsat 8 panchromatic image from 14 Nov 2014 of Imja-Lhotse Shar glacier with the
3 focus area of this study highlighted by the rectangular box several kilometers up-glacier from the
4 terminus, and the site location within Nepal shown in the inset.

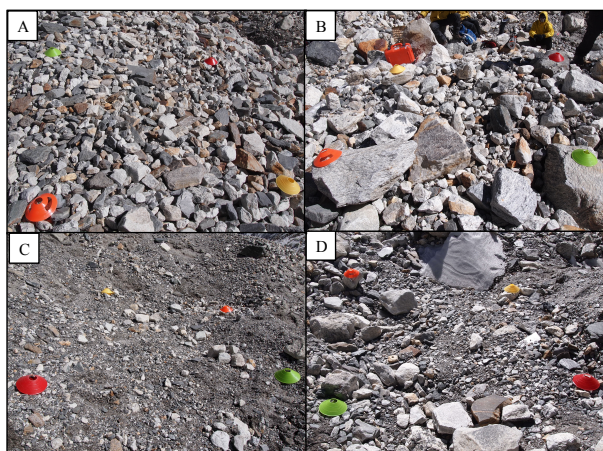
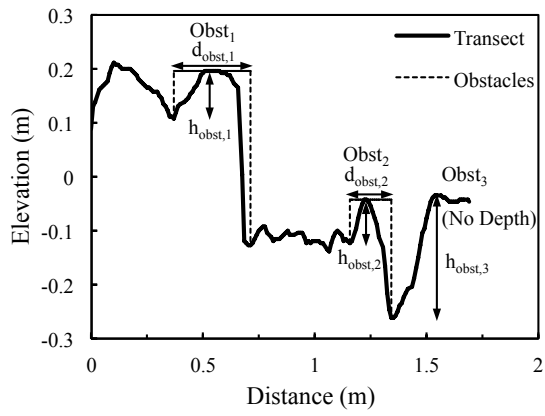


Figure 2. Sites A-D highlighting the variations in grain sizes that are found over the debris-covered portion of Imja-Lhotse Shar glacier (cones 0.19 m diameter).



1
2 Figure 3. Transect from Left-Right of Site B showing the identification of obstacles (Obst) and
3 their corresponding heights (h_{obst}) and depths (d_{obst}).
4

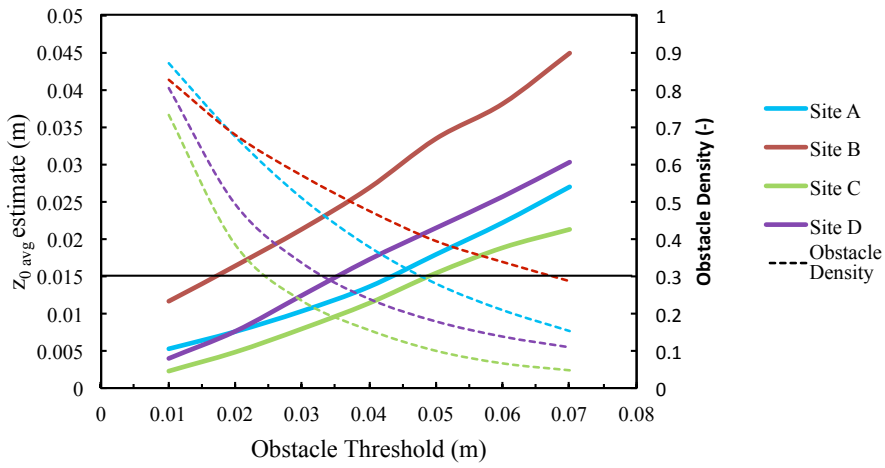
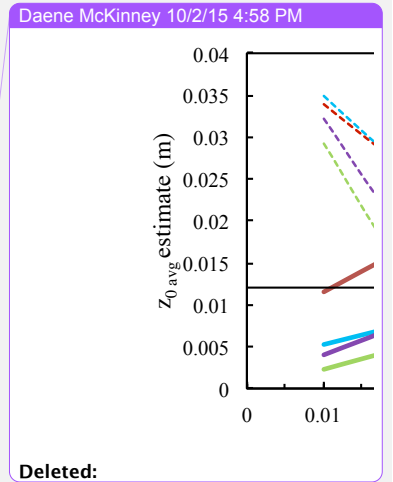
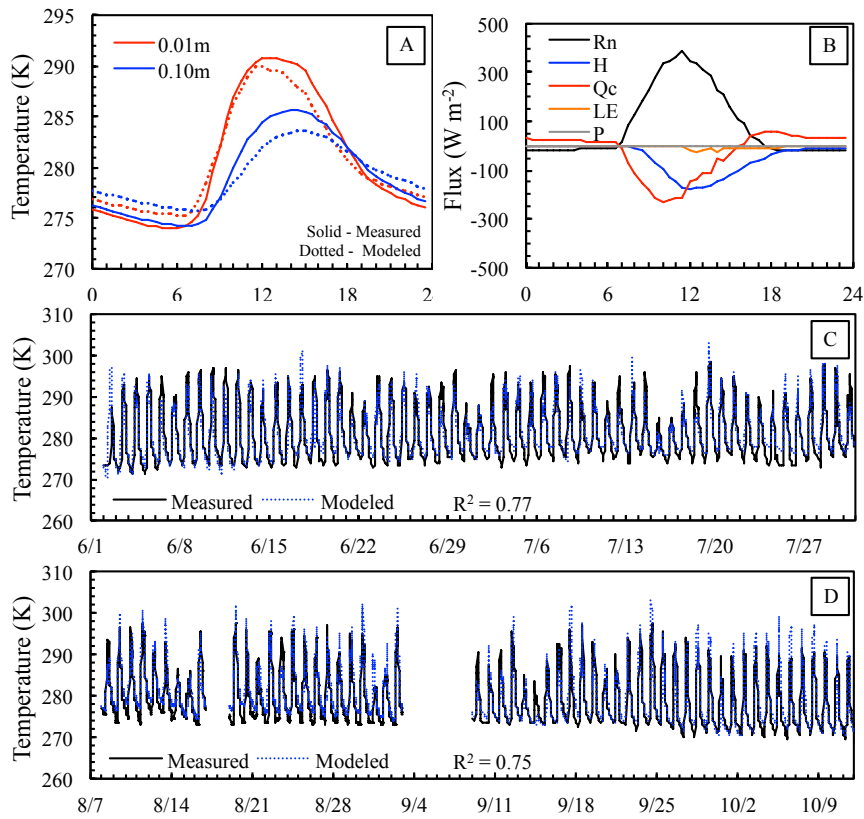
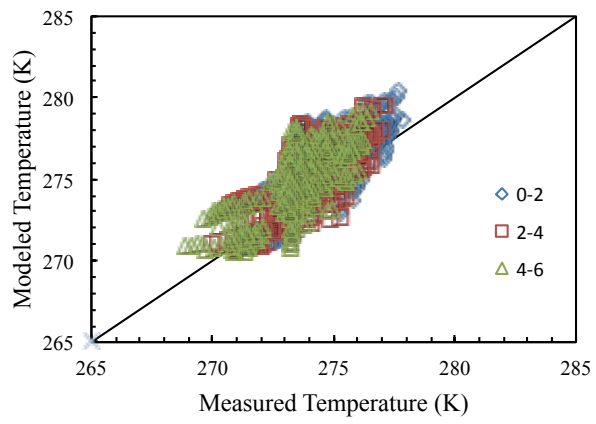


Figure 4. The effect of obstacle threshold (m) on both obstacle density and the estimate of surface roughness using the modified method for Sites A-D and the DEMs with a resolution of 0.01 m.





1
2 Figure 5. Various plots for Site 11 using the LE_{Rain} model showing (A) average daily
3 temperatures at two depths (solid and dashed lines indicate measured and modeled temperatures,
4 respectively), (B) average daily energy fluxes, (C) measured and modeled temperatures at a
5 depth of 0.01 m over the calibration period, and (D) measured and modeled temperatures over
6 the validation period.



- 1
- 2 Figure 6. Scatterplot of measured and modeled temperature for Site 11 at the surface for the
- 3 LE_{Rain} model showing the positive temperature bias overnight.

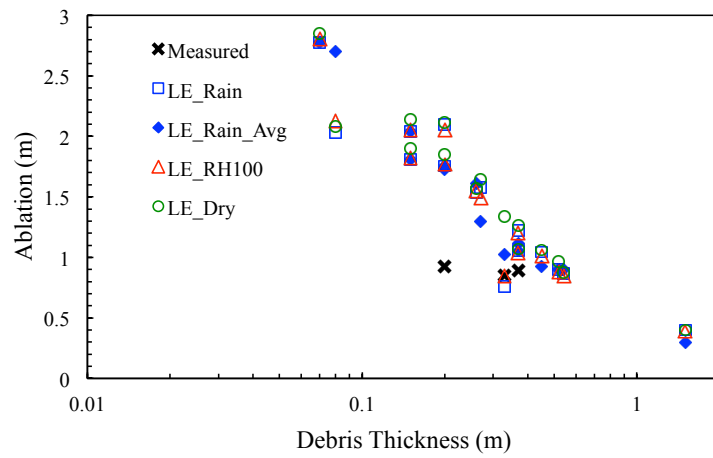
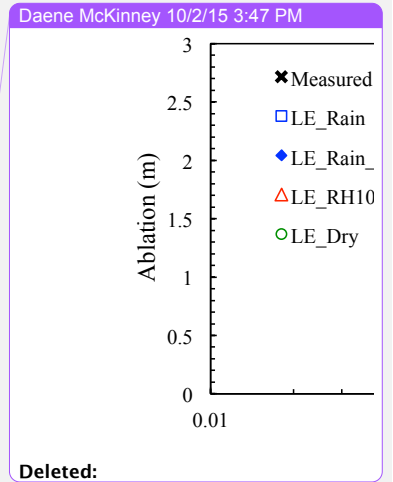
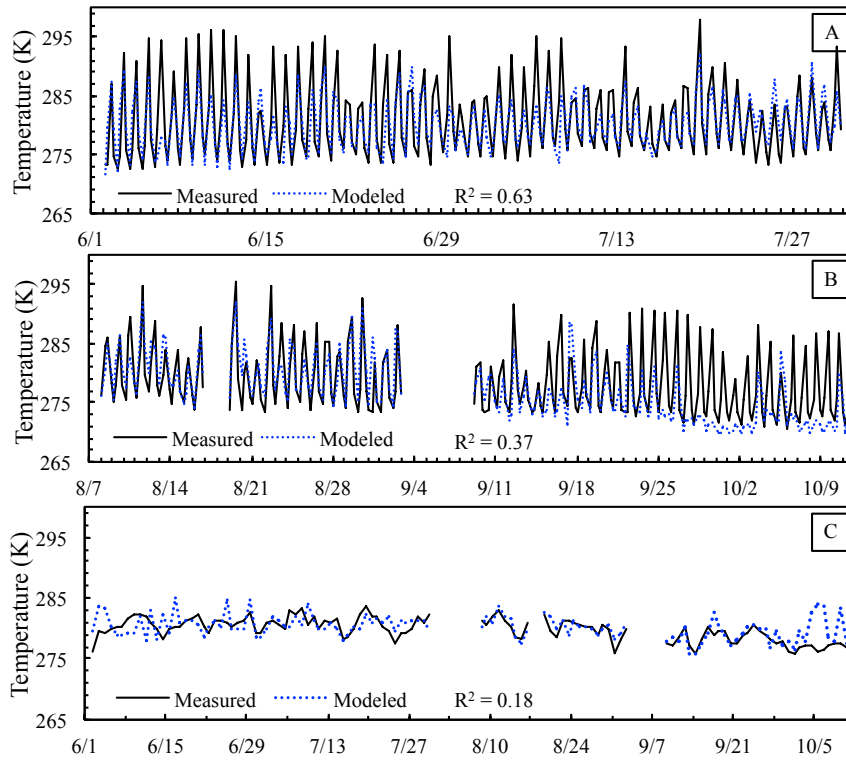


Figure 7. Modeled ablation with respect to debris thickness for all 15 sites from 18 May to 09 November 2014 for each of the three latent heat flux models including the LE_{Rain} model with average values and the three measured stakes that did not exceed 1 m.





1
2 Figure 8. Modeled and measured surface temperature at Site 11 over the (A) calibration period
3 and (B) validation period using 6-hour data, and (C) entire period using daily averages.