Journal of Glaciology



Article

Cite this article: Aubry-Wake C, Lamontagne-Hallé P, Baraër M, McKenzie JM, Pomeroy JW (2023). Using ground-based thermal imagery to estimate debris thickness over glacial ice: fieldwork considerations to improve the effectiveness. *Journal of Glaciology* **69**(274), 353–369. https://doi.org/10.1017/jog.2022.67

Received: 28 December 2021
Revised: 13 July 2022
Accepted: 14 July 2022
First published online: 8 August 2022

Keywords:
Debris-covered glaciers; glaciological instruments and methods; ground ice; remote sensing; supraglacial debris

Author for correspondence:
Caroline Aubry-Wake,
E-mail: caroline.aubrywake@gmail.com

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is included and the original work is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use.

cambridge.org/jog

Using ground-based thermal imagery to estimate debris thickness over glacial ice: fieldwork considerations to improve the effectiveness

Caroline Aubry-Wake[1], Pierrick Lamontagne-Hallé[2], Michel Baraër[3], Jeffrey M. McKenzie[2] and John W. Pomeroy[1]

[1]Centre for Hydrology, University of Saskatchewan, Canmore, Canada; [2]Department of Earth and Planetary Sciences, McGill University, Montréal, Canada and [3]Département de génie de la construction, École de Technologie Supérieure, Montréal, Canada

Abstract

Debris-covered glaciers are an important component of the mountain cryosphere and influence the hydrological contribution of glacierized basins to downstream rivers. This study examines the potential to make estimates of debris thickness, a critical variable to calculate the sub-debris melt, using ground-based thermal infrared radiometry (TIR) images. Over four days in August 2019, a ground-based, time-lapse TIR digital imaging radiometer recorded sequential thermal imagery of a debris-covered region of Peyto Glacier, Canadian Rockies, in conjunction with 44 manual excavations of debris thickness ranging from 10 to 110 cm, and concurrent meteorological observations. Inferring the correlation between measured debris thickness and TIR surface temperature as a base, the effectiveness of linear and exponential regression models for debris thickness estimation from surface temperature was explored. Optimal model performance (R^2 of 0.7, RMSE of 10.3 cm) was obtained with a linear model applied to measurements taken on clear nights just before sunrise, but strong model performances were also obtained under complete cloud cover during daytime or nighttime with an exponential model. This work presents insights into the use of surface temperature and TIR observations to estimate debris thickness and gain knowledge of the state of debris-covered glacial ice and its potential hydrological contribution.

Introduction

Rock debris is found on 44% of the world's glaciers (Herreid and Pellicciotti, 2020). Debris thickness is one of the key factors that modulate the sub-debris ice melt – a thin layer (less than several centimeters) enhances melt, while a thick layer of debris insulates the underlying ice and reduces melt (Østrem, 1959; Nicholson and Benn, 2006). Generally, debris thickness increases toward a glacier terminus (Anderson and Anderson, 2018), but exhibits a strong small-scale variability, caused by a variety of factors (Nicholson and others, 2018; Shah and others, 2019). Manual excavations (e.g., Reid and others, 2012), observations of debris thickness above exposed ice cliffs (e.g., Nicholson and Benn, 2013) or ground-penetrating radar surveys (e.g. McCarthy and others, 2017; Giese and others, 2021) at a high enough spatial resolution to capture this small-scale variability are both time and labor-intensive. As these methods either apply to limited zones or are highly labor intensive, there is a risk of bias in the results linked with the number and distribution of the measurements.

Due to the difficulties in obtaining field-based debris thicknesses, different methods have recently been developed to estimate debris thickness from remotely sensed observations, such as satellite-derived surface temperatures. Two main approaches are used to derive debris thickness from surface temperatures. Both are based on the assumption that surfaces above thinner debris have cooler temperatures than areas with thicker debris. The first approach is to derive debris thickness from an empirical relationship with surface temperature, but uncertainties remain regarding the best form of empirically-derived relationship to use, with linear (Mihalcea and others, 2008a) or exponential (Mihalcea, and others, 2008b; Juen and others, 2014; Minora and others, 2015; Kraaijenbrink and others, 2017; Tarca and Guglielmin, 2022) functions performing best in different studies. This empirical approach, appealing for its simplicity and low amount of required data, works when heat conduction between buried ice and the surface is the predominant surface energy balance term compared to others such as turbulent transfer and net radiation. The performance of these empirical relationships depends on the thermal conductivity of the surface material as well as microscale variations in energy balance terms. Other processes in the debris, such as heat advection in the subsurface, complicate the relationship between surface temperature and debris thickness and result in lower empirical model performance. Another key limitation of empirically-derived models to estimate debris thickness is that their performance is strongly dependent on the availability of well-distributed input data that represent the full range of debris thickness and surface temperature (Boxall and others, 2021).

The second common approach to derive debris thickness from surface temperature is through a physically based energy-balance model (Foster and others, 2012; Rounce and McKinney, 2013; Schauwecker and others, 2015; Rounce and others, 2018; Stewart and others, 2021). A major challenge for this approach is to obtain reliable meteorological data to accurately quantify the energy fluxes at the debris surface (Foster and others, 2012; Rounce and McKinney, 2013; Rounce and others, 2015). In addition to the difficulties in obtaining the energy fluxes, debris properties, such as the thermal resistance, surface roughness or surface albedo, often have to be estimated across the glacier area and add further uncertainty to the method.

Both the empirical relationships and the energy-balance approach only provide an accurate debris thickness estimate for debris shallower than approximately 0.5 m, after which a decoupling between surface temperature and debris thickness is observed (Mihalcea and others, 2008a; Foster and others, 2012; Tarca and Guglielmin, 2022). Rounce and others (2018, 2021) addressed this limitation and obtained robust estimates for debris thicker than 0.5 m by combining an inverted sub-debris melt model with calculations of elevation change and flux divergence. This method requires a series of digital elevation models and surface velocity data in addition to the meteorological data and debris parameters of the melt model. Therefore, it faces similar difficulties linked with data and parameter uncertainty.

As pointed out by Rounce and others (2013), a key limitation in using surface temperature from satellite to derive debris thickness, either using empirical models or surface-energy balance, is the poor resolution of the satellite thermal band (typically between 60–100 m). At this resolution, local slope and aspect variations, which play an important role in controlling the surface energy balance and therefore the surface temperature, are not captured. Due to the large pixel area of these thermal images, debris, ice cliffs and supraglacial ponds can all appear in a single pixel, resulting in an underestimation of the debris thickness (Rounce and others, 2018; Herreid, 2021). Nicholson and others (2018) showed that ignoring the small-scale variability in debris thickness and using spatially averaged values can have a strong impact on the melt predicted by glacier models, underestimating the ablation rate by 11–30%.

One option to capture surface temperature at a high spatial resolution in glacierized terrain is to use near-surface remote sensing, such as from a plane or uncrewed aerial vehicles (UAVs, Kraaijenbrink and others, 2018) or ground-based oblique imagery (Hopkinson and others, 2010; Aubry-Wake and others, 2015; Herreid, 2021; Tarca and Guglielmin, 2022). These instruments offer the possibility to measure surface temperatures from thermal infrared radiometry (TIR) both at a high spatial resolution and with flexible timing to account for different times of the day or variable weather conditions. This increased flexibility in measurement methods enables further investigation of the relationship between surface temperature and debris thickness (Herreid, 2021), but it also facilitates assessments of when and how surface temperature should be measured to optimize debris thickness estimates. For example, Mihalcea and others (2008a) suggested that early morning hours optimized the correlation between surface temperature and debris thickness, while afternoon provided the weakest correlation at the debris-covered Miage glacier in the Alps. Herreid (2021) further examined the optimal time of day at which surface temperature should be collected for the Canwell Glacier in Alaska and found that cold nights provided poor conditions to derive debris thickness from surface temperature, as the debris surface became isothermal. Herreid (2021) suggested that weather conditions play a key role in obtaining suitable thermal infrared imagery to derive debris thickness, with clear days being the least suitable and cloud cover providing optimal conditions. However, obtaining surface temperature from UAV or ground-based TIR imagery presents complications that need further assessment, such as issues related to the radiometer viewing angle, which can cause shifts in emissivity, the distance between the TIR radiometer and the study surface, which increases the atmospheric interference, and small-scale variations in surface emissivity linked with varying surface types (Aubry-Wake and others, 2015; Kraaijenbrink and others, 2018; Baker and others, 2019; Herreid, 2021). Considering the advances in measuring surface temperature using near-surface remote sensing and its widespread potential use in debris thickness reconstructions, there is a need for further experimentation and testing to explore best practices to measure debris surface temperature as a method to estimate debris thickness.

The objective of this paper is to examine how debris thickness over glacial ice can best be estimated using ground-based, oblique TIR imagery. More specifically, this paper explores how the simple empirical models of debris thickness based on TIR surface temperature and debris thickness are influenced by:

a) The empirical regression type
b) The time of day of TIR image acquisition
c) Weather conditions
d) The number and depth distribution of the manual debris thickness measurements
e) The spatial resolution of the images
f) The distance between the region of interest and the TIR camera.

Study area

The ice-cored lateral moraine of Peyto Glacier (51.676°N, −116.554°W), a well-studied outlet glacier of the Wapta Icefield in the Canadian Rockies and a World Glacier Monitoring Services reference glacier (Pradhananga and others, 2021; WGMS, 2021), was selected to investigate the relationship between debris surface temperature and debris thickness. This ice-cored moraine used to be connected to the glacier toe, but ongoing glacier retreat has disconnected it from the clean ice glacier. The study area is 100 by 50 m (Fig. 1) study area and is predominantly east-facing, with an average slope of 23.5̊, gradually increasing toward the upper ridge, and is located between elevations of 2138 and 2159 m a.s.l. (Fig. S1). During the TIR surveys, sunrise at the study site above the horizon created by surrounding mountains was ~ 0800 h MDT and sunset was at 1800 h, reaching full darkness by 2200 h. The mountains surrounding Peyto Glacier are composed of dolomitic limestone rock, which was representative of the debris composing the study area. Two weather stations located in the vicinity of the study area recorded the meteorological conditions during the experiment: $AWS_{moraine}$ located 1.4 km northeast of the study area on the moraine below the glacier toe, and AWS_{ice}, an on-ice station located 415 m east of the study area (Fig. 1).

Data collection and processing

Thermal infrared imagery survey

Two thermal infrared imaging radiometers were used at four locations to survey the study area and surroundings. A Jenoptik VarioCam HD thermal infrared imaging radiometer was installed on the glacier facing the study area at location TIR1 (Fig. 1a), and recorded images at 5-min intervals from 5 August 2019, 17:15 to 9 August 2019, 9:00 for a total of 1077 images over 90 consecutive hours. The radiometer's field of view is shown in Fig. 1b. The field of view of TIR1 is considered the reference frame and is the primary field of view for the analysis.

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

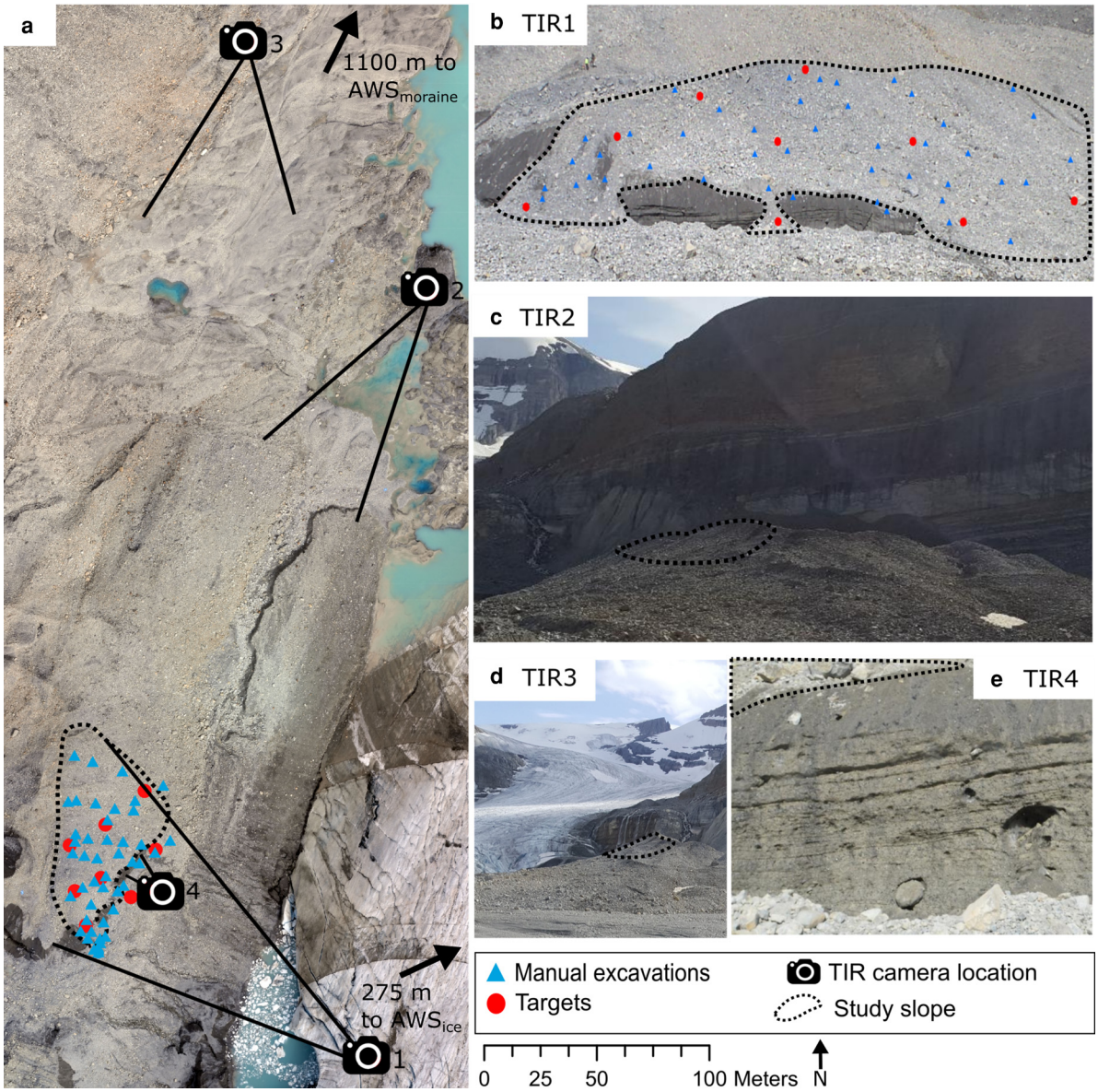


Fig. 1. Study area (a) showing the location and camera angle from the four thermal infrared imaging radiometer locations and the distance and direction to the AWS_{ice} and $AWS_{moraine}$. The field-of-view from the four locations is shown in (b–e). The blue triangles in (a) and (b) show the manual excavation locations and the red circle shows the control point. In all pictures, the dashed black line delimits the study area. Note that the scale bar and north arrow apply to (a) only.

A second Jenoptik VarioCam Basic thermal imaging radiometer was installed at three locations for intermittent measurements during the experiment. The only difference between the two cameras is the pixel resolution: 1024 × 768 for the VarioCam HD and 640 × 480 for the VarioCam Basic. Locations TIR2 and TIR3 (Fig. 1a) had a survey area larger than TIR1, with a field of view including a side-view of the study area, the toe of Peyto Glacier, mountain cliffs and the sky (Figs 1c and 1d). The fourth location was a close-up view of an ice cliff in the study area (Fig. 1e). These three locations also recorded thermal images at 5-min intervals. Details of the image collection timing can be found in Table 1.

Both thermal imaging radiometers use an uncooled microbolometer sensor with a spectral range of 7.5 to 14 µm, with a manufacturer-stated accuracy of ± 1.5°C. The radiometers were automatically corrected for an emissivity of 1.0, a distance of 10.0 m and an atmospheric temperature of 20.0°C. The emissivity of limestone debris is closer to 0.93 (Salisbury and D'Aria, 1994; Sobrino and Cuenca, 1999; Kraaijenbrink and others, 2018) which means that its reflectance of incoming longwave radiation is 0.07 and it is not a full black body for application of the Stephan-Boltzmann equation. An emissivity of 1.0 is a standard assumption of land surface models and other situations where it is difficult to explicitly account for near-black body emission and reflectance and permits direct application of the Stephan-Boltzmann equation. This assumption results in the total radiation measured by the radiometer and considered in calculating the surface temperature being a combination of the radiation emitted and reflected from the target object and atmospheric radiation emitted along the beam path. To calculate accurate absolute measured surface temperatures, a correction needs to be applied to the total radiation to isolate the radiation emitted from the target (Cardenas and others, 2008; Aubry-Wake and others, 2015). Surface emissivity and longwave irradiance are needed to solve for the surface radiant temperature in this manner, however as discussed in Baker and others (2019), the uncertainty associated with the small-scale variability in surface emissivity and longwave irradiance to the surface due to slope, aspect, sky view, wetness, ice exposure and surface mineral characteristics, coupled with the variability of the radiant temperature of the surface, results in a limited ability to apply these spatially

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

Table 1. Details of thermal infrared time-lapses

Location	Scene description	Number of pixels in study area	Start time – End time	Recorded time (min)	Distance to ice cliff (m)	Number of images
TIR1 *Fig. 1b*	Study area	91 408	5 Aug, 17:15 – 6 Aug, 9:00	89:45	140	1077
TIR2 *Fig. 1c*	Ice-cored moraine and cliff	18 966	5 Aug, 17:35 – 6 Aug, 2:20	8:10	290	98
			7 Aug, 13:55 – 7 Aug, 21:20	7:30		90
			8 Aug, 11:40 – 8 Aug, 23:35	11:55		143
TIR3 *Fig. 1d*	Ice-cored moraine and glacier toe	4052	6 Aug, 17:25 – 6 Aug, 20:45	3:25	370	41
TIR4 *Fig. 1e*	Ice cliff close-up	10 251	9 Aug, 8:35 – 9 Aug, 10:20	1:50	6	22

The picture from Fig. 1 corresponding to the field-of-view for each location is indicated in italic.

variable corrections accurately to field TIR images. In the field TIR imaging radiometer used, any emissivity correction or atmospheric correction is applied uniformly across the image and so modifies all surface temperatures equally. Such coarse corrections would not advance this study, which focusses on how observed correlations between TIR imagery observations and debris thickness are influenced by different TIR field data collection procedures. The debris emissivity was therefore assumed to be 1.0 and images were not corrected for atmospheric TIR emission.

The TIR images recorded at TIR1 were first co-registered to account for minor position shifts of the camera using an automated, intensity-based image registration algorithm from the MATLAB Image Processing Toolbox (MathWorks, 2017). For TIR1, three groups of images were considered, corresponding to the images recorded between battery changes. Then, the three image groups were registered using manual control-point image registration to account for the sudden shift in the field of view of the camera at battery changes. A similar procedure was followed for locations TIR2, TIR3 and TIR4. For each location, a visual image was also registered using a control-point approach (Figs 1b–e).

Manual excavations

Forty-four manual debris excavations were dug at an array of locations aiming to capture the local debris thickness variability of the study area. The debris layer thickness was measured at each excavation using a measuring tape along the vertical axis. Manual excavations were surveyed with differential GPS to obtain accurate locations.

Even though the aim was to select excavation spots randomly, some bias may have been introduced to their location. First, a debris depth of zero was not recorded, even though some areas showed exposed ice. The steepest slopes of the study area were not sampled for safety reasons. Sparse or thin debris sites, in a region of discontinuous debris on the southern end of the study area (left side of the TIR1 images), were unstable and multiple mass movements were observed during the surveys, meaning that the recorded depth might only be valid in temporal proximity to the measurements. For the rest of the study area, debris cover was continuous and stable throughout the observation period. Manual measurements were not performed under large boulders as it was impossible to displace them. These two factors led to systematic under-sampling of the extremes in debris thickness and bias samples to depths of 10 to 70 cm The manual excavations were refilled while attempting to recreate the stratigraphic distribution of the debris.

Geospatial localization

During the TIR surveys, nine targets were installed on the study area surface and located with a differential GPS (red circles, Figs 1a–b). All nine targets are visible in the registered visual image of the study area, but only six targets, which were partially covered with aluminum foil to create a significantly different emissivity, are visible in the TIR images.

A high-resolution orthomosaic and digital surface model was obtained from an uncrewed aerial vehicle (UAV) survey on 28 August 2019, 18 days after the TIR surveys. A comparison of visual images recorded during the TIR surveys and the UAV flight shows the ice cliff in the study area receded slightly, but only affecting the area immediately adjacent to the ice cliff and not affecting the rest of the study area. The orthomosaic visual imagery is used to select an additional six boulders as tie-in points between the TIR images and the UAV survey to provide 15 ground control points. The control points were used to geolocalize the TIR images, create a local spatial frame of reference and assign spatial coordinates to every pixel in the TIR images. These local spatial frames, in the form of matrices of latitude and longitude values of the same dimensions as the TIR images, were used to localize the manual excavations in the TIR images.

Results and discussion

Debris thickness from manual excavations

The manual excavations showed a debris thickness ranging from five cm to more than 110 cm (Fig. 2a). The most common measured depth was between 40–45 cm ($n = 6$), followed by 45–55 cm ($n = 10$) (Fig. 2b). Stratigraphy was recorded for every excavation more than 15 cm deep. Most stratigraphy showed a layer of coarse to very coarse debris (average thickness of 15 ± 9 cm) overlaying a layer of sand, typically moist up to approximately two cm below the sand/pebble interface, with interspersed boulders (Fig. S2). The bottom of the sand layer, at the ice-debris interface, was water-saturated and ponding was observed at the ice interface in four of the 45 excavated holes. This lack of ponding at the debris-ice interface might be linked to the slope of the study area, promoting quick evacuation of the meltwater, and might be less representative of other debris-covered areas with lower surface slopes (Giese and others, 2020).

The debris thickness measurements were interpolated following a natural neighbor algorithm to obtain distributed debris thickness over the study area (Fig. 2a). The debris thickness across the slope showed a similar distribution as the measured manual excavation, with debris thickness ranging from 0 to 120 cm with a mode of 35–40 cm. To simplify the image processing, the ice cliffs at the bottom of the study area were not included in the analysis. Since this study focuses on the correlation between debris thickness and surface temperature rather than ice cliff processes or absolute surface temperature analysis, including the ice cliffs in the analysis was considered out of the scope of this current study.

The interpolated debris thickness is used as the reference to which the simulated debris thicknesses are compared. Due to the high density of manual excavations performed to calculate the interpolated debris thickness, it is likely to represent the spatial patterns occurring on the slope. However, the interpolated debris thickness might miss small-scale spatial patterns that could be reflected by the surface temperature variations. Therefore, the

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

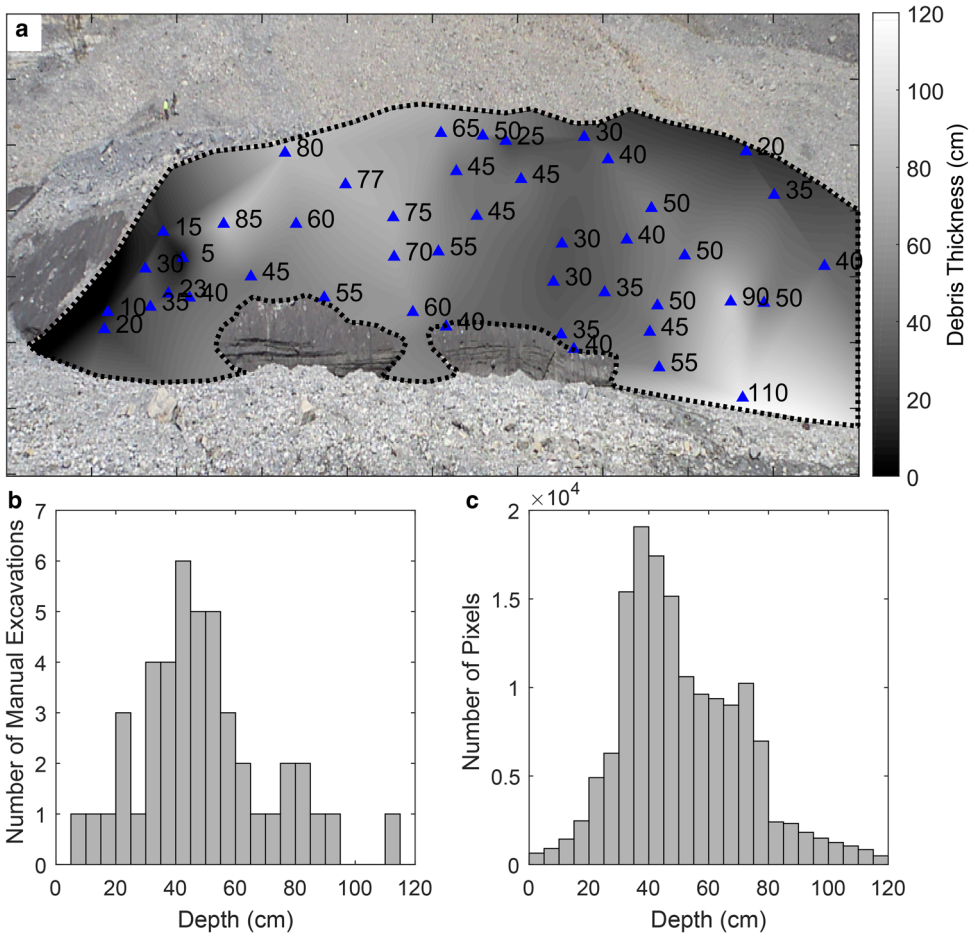


Fig. 2. Location and depth of manual excavations (blue triangles) and interpolated thickness from point excavation in (a). The dashed line indicates the study area. (b) The distribution of the manual excavations and (c) the distribution of the interpolated debris thickness across the study area.

interpolated debris thickness is used as a reference in the comparisons, but should not be considered the known debris thickness.

Measured surface temperature at manual excavation locations

The surface temperature of 2 × 2 pixels (corresponding to 0.12 × 0.4 m) at the location of each manual excavation was extracted from the TIR1 images (Fig. 3). The manual excavations did not cause detectable thermal anomalies in the surface temperature record. This was assessed by comparing the temperatures at the excavation locations, for the period when the manual excavation was performed, to the retrieved time series for those in the immediate vicinity (Fig. S3). Generally, thicker debris had higher surface temperatures than thinner debris, with the largest differences occurring at peak shortwave irradiance (~ 1400–1500 h MDT). The surface temperature pattern differs on the first day of the study period, 6 August 2019, from those on the subsequent two days, 7 August 2019 and 8 August 2019, due to a shift in weather, with the first day being predominantly cloudy, with rainfall occurring on the evening of 6 August 2019 and the following days being sunny. The influence of weather on the surface temperature and debris thickness correlation is further discussed in section 4.1.1. (Fig. 6). 6 August 2019 shows cooler peak surface temperatures, reaching a daily maximum of 21.8°C compared to 27. and 27.1°C on 7 August 2019 and 8 August 2019. Similarly, night temperatures on 5–6 August 2019 only reached lows of −0.11°C, compared to −2.75, −2.64 and −2.30°C for the following nights. The timing of minimum and maximum temperatures also differed between the first day of the study period and subsequent days. On the night of 5–6 August 2019, the minimum temperature occurred at 0110 h MDT, compared to 0555, 0620 and 0700 h MDT for subsequent nights. The variation in surface temperatures observed at the sites of manual excavations reached between 16 and 23°C in the early afternoon, but only 6°C at the end of the night.

Relationship between debris thickness and surface temperature

The 44 manual excavations and corresponding surface temperatures were used to calculate empirical relationships for each TIR1 image ($n = 1077$). Five regression types were tested with debris thickness as the only control of surface temperature (linear, quadratic, power, exponential, log). Multivariate linear regression models were also tested using slope, aspect and elevation in addition to surface temperature as possible controls on debris thickness. Slope and aspect are important controls on debris thickness in addition to surface temperature (Rounce and McKinney, 2013). For the simple regression models, the linear and logarithmic empirical regression performed very similarly, and the exponential and power empirical regression also showed comparable performance. The second quadratic fit provided a model performance between the two groups (Figs 4a–d). Out of the 1077 time steps analyzed, model performance was improved for 20, 238 and 25 time steps when including either slope, aspect and elevation, respectively, compared to the linear regression model considering surface temperature only, and no improvements were obtained when including a combination of slope, aspect and elevation (Fig. 4e). These improvements were minor, with an average increase of 0.01 in the adjusted R^2 values

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

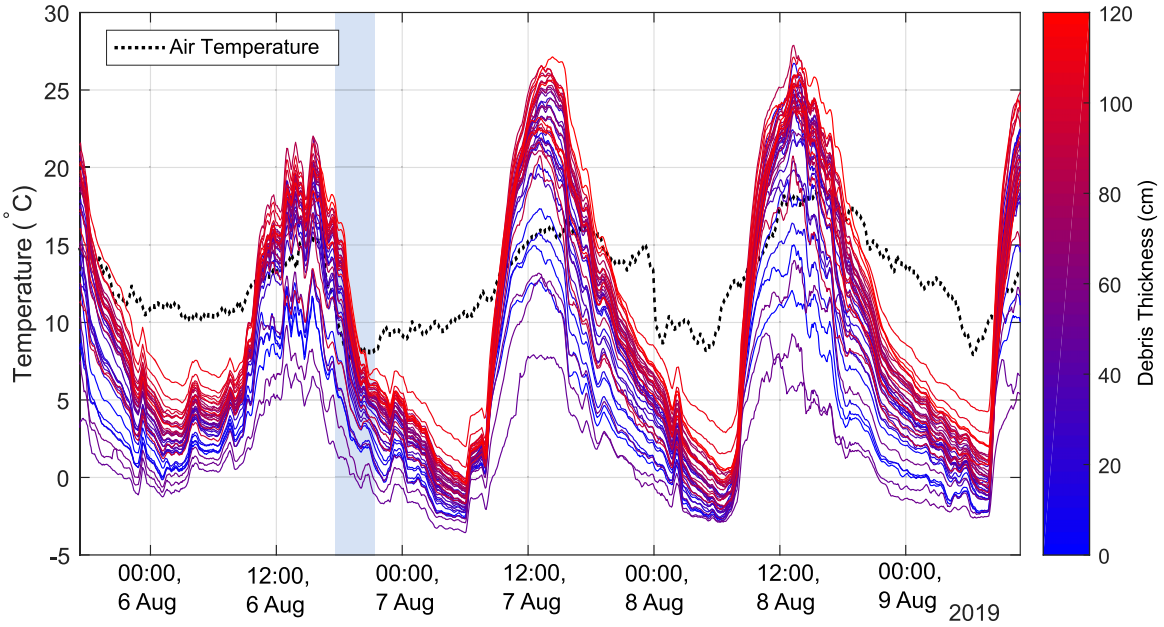


Fig. 3. TIR1-derived surface temperature at the location of each manual excavation (full line) and air temperature measured at the moraine $AWS_{moraine}$ (dotted line). Blue lines correspond to thin debris and red lines to thicker debris. Lines are smoothed using a 30-min moving frame for visual clarity. The shaded blue area corresponds to a period of intermitted rain.

compared to the linear regression model. The time steps that showed improvement when including either slope, aspect or elevation in the linear regression models still performed more poorly than the exponential regression model. This limited model improvement is likely caused by the largely uniform slope, aspect and elevation of the study area. As limited improvements were observed when including topographic variables in the linear regression model, and the comparable performance between the varied empirical model tested, only the exponential or linear were selected for further analysis and discussions.

The coefficient of determination, R^2 between surface temperature and debris thickness from both linear and exponential model fit is shown in Figs 4a–c for 8 August 2019 at 12:00 (noon), 8 August 2019 at 20:00 and 9 August 2019 at 6:45. These times were selected as reasonable times for fieldwork data collection in remote glacial environments, corresponding to just before sunrise, ~ sunset, and midday. Additionally, these images show examples of relatively poor, average and good model performances, with the good performances corresponding to R^2 values above the median (0.55 and 0.50 for the exponential and linear model), the average values near the median, and the poor model values corresponding to the bottom 10% of R^2 values.

For each regression model, the goodness-of-fit at each manual excavation was assessed by calculating the adjusted coefficient of determination adjR^2 and the root mean square error normalized to the range of observed data (nRMSE) between the measured debris thickness and the modeled debris thickness. The coefficient of determination ranged between 0.09 and 0.70 (Fig. 4d), with nearly all models showing a significant correlation (with a p-value lower than 0.05). Only three thermal images for the exponential fit and six images for the linear fit, out of 1077, did not have a significant correlation. The nRMSE values range between 0.13 and 0.21 (Fig. 4e) for the duration of the study. Both metrics show a strong variation in model performance based on time of day, which is further discussed in section 4.4.2. The coefficients obtained for each linear and exponential model varied with time and are included in Fig. 5.

Under clear conditions from 7 August 2019 to 9 August 2019, model performance was at the lowest for both models when direct solar radiation first reached the study area, ~ 8:00 AM. Model performance increased throughout the day, to reach a maximum at late night or very early morning (~ 2:00 for exponential fit and 6:00 for linear fit). The performance of both models increased throughout the night as the negative net longwave radiation gradually cooled the debris and the influence of warming from shortwave net radiation in the previous day decreased relative to heat conduction. This explains why both models perform better late at night or very early in the morning, long after the sun sets behind the mountains at 1800 h. When sunlit, the surface energy balance becomes dominated by net shortwave radiation, and the amount of irradiation absorbed by the surface fluctuates at a very small spatial scale due to varying rock albedo, slope, aspect and shading from both individual boulders and surface microtopography which further decouples the surface temperature from the debris thickness. The simple exponential and linear regression models tested herein could not reliably estimate debris thickness in shortwave radiation-dominated regimes since the surface temperature increase caused by absorption of shortwave radiation is decoupled from subsurface thermal regimes associated with debris thickness over ice, but rather is influenced by many other factors (e.g., partial cloud cover, local slope inclination, shading, multiple reflections, solar elevation), which is in agreement with finding from Herreid (2021). These results suggest that the ideal timing to measure surface temperature for estimating debris thickness over ice using empirical relationships as presented here is late at night or early morning before sunshine hits the study site. Simple relationships ignoring the surface energy-balance processes or not accounting for the strong influences of shortwave radiation on surface temperature, should not be used to estimate debris thickness from surface temperatures measured during daylight, as it is usually done with satellite imagery and UAV applications.

The variation in model performance between the exponential and linear fit also followed a pattern: an exponential curve provides a higher coefficient of determination and a lower RMSE during most of the day, but a linear fit performs better late at night, approximately from 3:00 AM to 8:00 AM. This diurnal pattern between the two models was broken on the first day of data collection (6 August 2019), when dense clouds, fog and rain occurred over the study area, causing similar high model

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

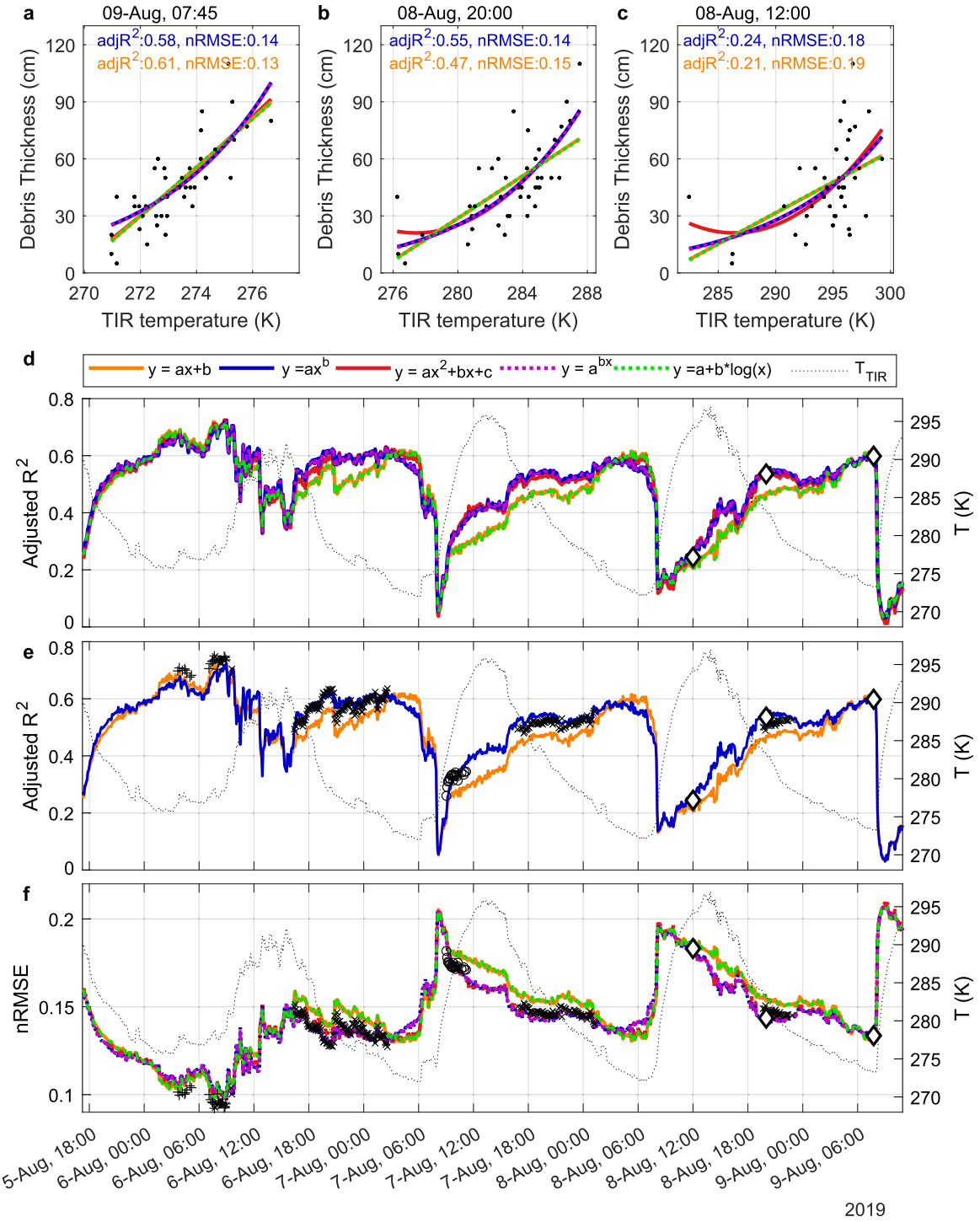


Fig. 4. Correlation of debris thickness and TIR surface temperature with empirical regression for linear, exponential, quadratic, power and logarithmic fit for three-time steps over the study period (a–c). Panel (d) shows the calculated adjusted coefficient of determination R^2 for the five types of regression tested and panel (e) shows the linear and exponential fit in combination with the multiple linear regression model including slope, aspect and elevation. Panel (f) shows normalized Root Mean Square Error (nRMSE) for each TIR image for 5-August to 9-August 2019. The average TIR surface temperature is shown in (d–f) on the right axis. The timing of the (a–c) scatter plots is indicated by the diamond on the (d–f) panels and the blue shading indicates the intermitted rainfall period.

performance for both the exponential and linear models. This pattern is likely linked to the dual controls on surface temperature occurring throughout the day. During daylight, the solar radiation heats the debris surface, but thin debris stays cool due to the close presence of the underlying ice, creating an exponential fit between the debris thickness and surface temperature. At night, the thermal inertia linked with the presence of the underlying ice dominates, and the correlation between surface temperature and debris thickness is better represented by a linear model.

Factors impacting the surface temperature-debris thickness correlation

Over this study period, the empirical relationships derived between surface temperature and debris thickness showed a high variability both in terms of model performance and parameters values (Figs 4, 5). In the next section, different factors that influence the surface temperature-debris thickness relationship, and the resulting modeled debris thickness, are investigated.

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

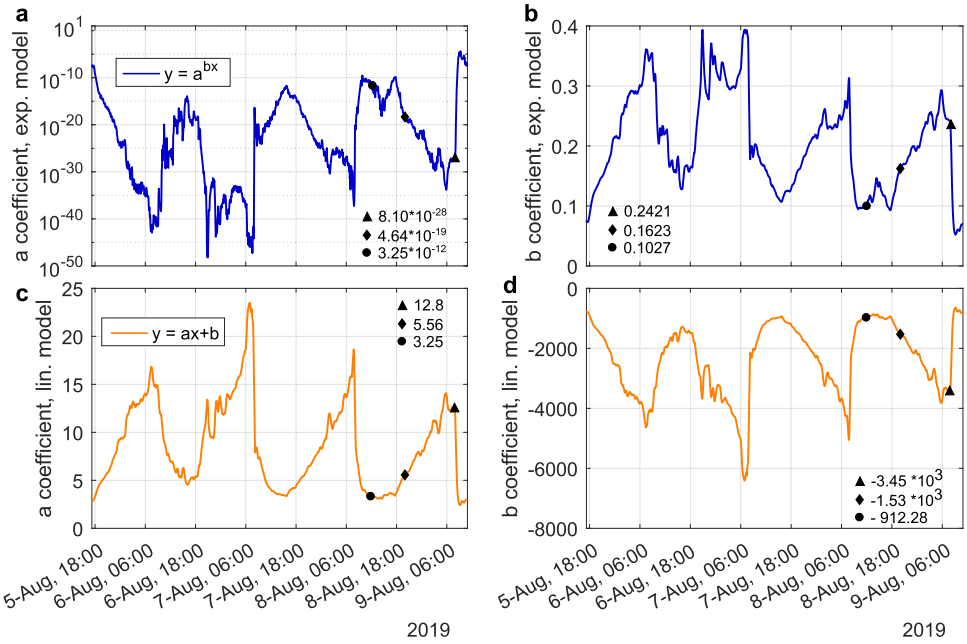


Fig. 5. Parameter values for both the exponential and linear model. The timing of the scatter plots shown in Fig. 4a–c is indicated by the triangle (08-Aug-2019 12:00), the diamond (08-Aug-2019 20:00) and the circle (08-Aug-2019 12:00), with the corresponding values for the parameters indicated on figure.

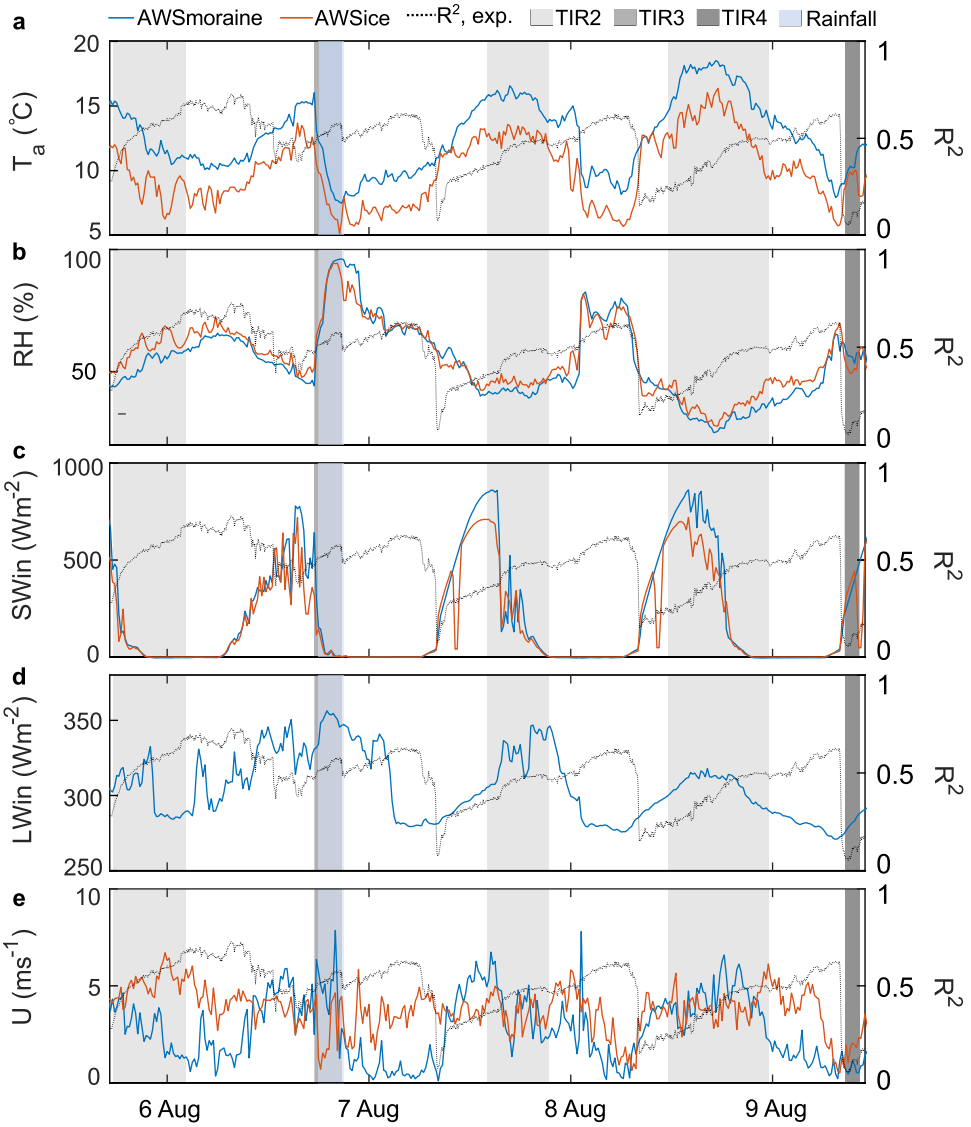


Fig. 6. Meteorological measurements at the Peyto Moraine and Ice weather stations. The coefficient of determination R^2 for the exponential model between surface temperature and debris thickness (dashed line) is shown on the right axis. The shading represents the periods at which TIR2 to TIR4 were taken while TIR1 was measured during the entirety of the plotted data. The intermittent rainfall period is shaded in blue and is overlapping with the TIR3 measurement period.

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

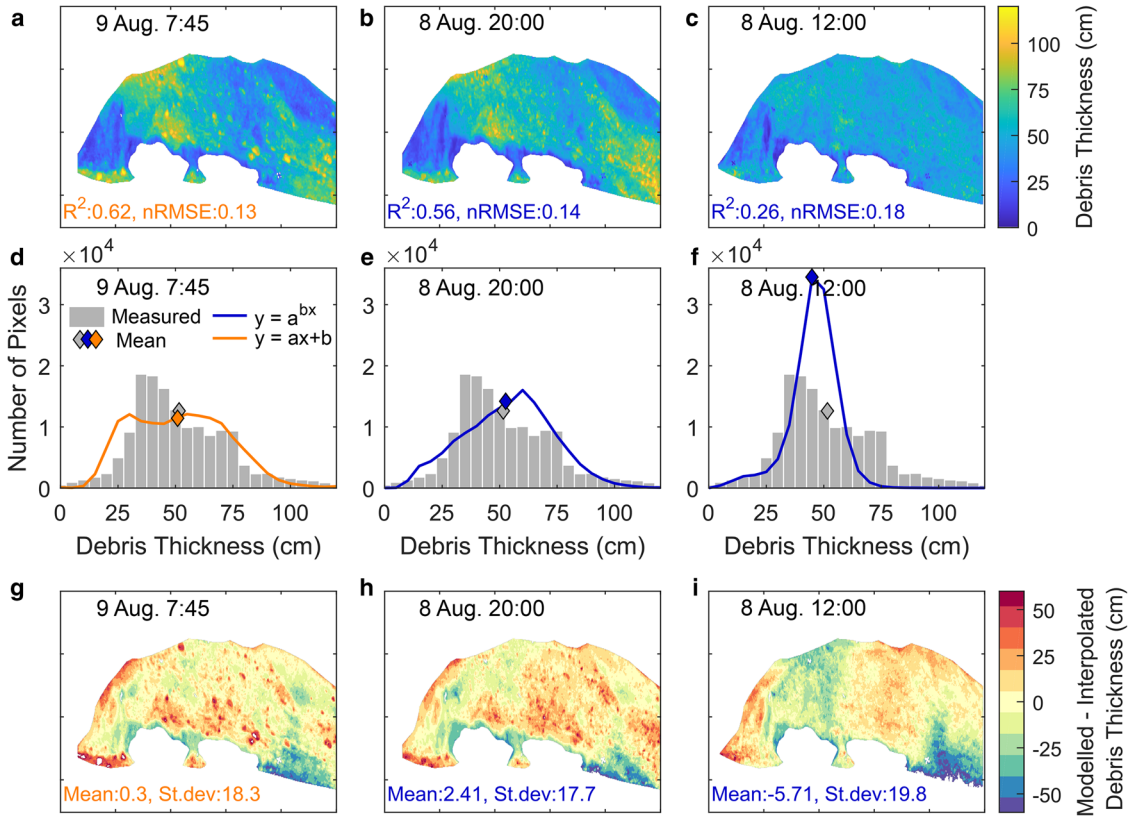


Fig. 7. Modeled debris-thickness based on the 'good' (a), 'average' (b) and 'poor' (c) model performance and associated debris thickness distribution (d–f). The difference between the modeled debris thickness and the interpolated debris thickness (in Fig. 2a) is shown in (g–h). Orange refers to the linear model and blue to the exponential model. Note that the measured debris distribution, in (d) – (f) is the same as shown in Fig. 2c.

Five environmental and data collection factors were tested: (1) weather conditions, (2) the time of the day at which the TIR images are collected, (3) the number and (4) distribution of the manual excavations, and (5) the spatial resolution of the TIR images.

Weather conditions

The meteorological conditions observed during the experiment explain some of the measured surface temperature as well as the modeled debris thicknesses. Over the study period, weather was overcast during the first day of the three-day study (6 August 2019), with intermittent rainfall occurring in the evening, with 10 mm rainfall recorded at the $AWS_{moraine}$ between 6:00 PM and 9:00 PM, followed by two predominantly sunny days on 7 August 2019 and 8 August 2019. This cloudiness is also discernible in the incoming shortwave and longwave radiation measurements from the two weather stations adjacent to the study area: the moraine and on-ice Peyto weather stations (Figs 6b–c). Both meteorological datasets are shown to showcase the variable conditions in the area. The AWS_{ice} is located only 425 m to the study area, and at only 25 m higher, but on a different surface, and the $AWS_{moraine}$ is located 1400 m away, and 90 m higher, but on similar moraine sediments.

Over the three nights analyzed, the best model performance occurs ~0745 h MDT during the morning of 6 August 2019 under cloudy conditions, which contrasts with the following predominantly clear nights, when the highest correlation between surface temperature and debris thickness occurred slightly earlier (before 0700 h). Additionally, on 6 August 2019, both the linear and exponential model performances stay relatively high throughout the day compared to the following two days. The high cloudiness on 6 August 2019 shifted the incoming radiation regime from one dominated by shortwave radiation with its high spatial variability to one dominated by longwave radiation, which has a lower dependency on slope and aspect of the surface. The cloudiness allowed the conductive heat flux to continue to cause the observed variability in surface temperature and resulted in better daytime model performance. This suggests that one can reliably obtain reasonable model performance toward the end of the night under clear conditions, but other times may be suitable if a continuous cloud cover is present, which is suitable to ground- or UAV-based surface temperature survey, but not applicable to satellite imagery analysis. These results are in agreement with Herreid (2021), who found optimal timing to use TIR imagery to estimate debris thickness during nighttime or during daytime hours but under cloudy weather, and Mihalcea and others (2008a), who found a stronger correlation between elevation-band averaged surface temperature obtained from thermistor and debris thickness during nighttime.

Low clouds and short intense precipitation occurred between 19:00 and 21:00 on 6 August 2019, likely causing interference between the TIR cameras and the study area and resulting in doubtful temperature measurements. Due to the short distance between TIR1 and the study area (140 m) and the short duration of the intermittent rainfall, it was not possible to clearly detect the rainfall events in the TIR1 visible or infrared images. No image appeared obviously blurred from the atmospheric interference, and therefore, all the images were kept in the analysis. TIR3 was also active during the rainfall event, and in this case, due to the wider field of view and longer distance between the study area and the TIR3 camera, the atmospheric interference due to clouds was noticeable in the above and behind the study area, but not at its exact location, and all images were kept in the analysis. However, the surface temperature and modeled debris thickness obtained during the rainfall event should be treated carefully. Examples of the TIR1 and TIR3 visible and TIR images are shown

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

in Fig. S3. Once the clouds lifted and the atmospheric path between the study area and the camera was clear, the possible presence of rainwater on the debris of the studied area did not seem to significantly affect the correlation between surface temperature and debris thickness, as shown by the relatively high correlation for the images of TIR1 on 6 August 2019 from 18:00 to 21:00 in Fig. 4d, during which the rainfall occurred. The good empirical performance model performance under wet conditions obtained here are contrasting with results from Herreid (2021), who observed the opposite.

Temporal variation

The distribution of calculated and interpolated debris thickness is compared for the same times of day as shown in Figs 4a–c, selected to represent relatively good (0745 h), average (2000 h) and poor (1200 h) model performance. For each image, the best-fit model was selected. Both examples for the good (0745 h, linear) and average (2000 h, exponential) model performances provide reasonable estimates of debris thickness (Figs 7a–b). Both obtain average debris thicknesses over the study area that are one cm or less from the mean debris thickness interpolated from the manual excavations. However, they show mixed successes in replicating the interpolated debris distribution (Figs 7d–e). They can capture the range of measured debris thickness, from 0 to 110 cm thick, and capture the distribution of debris thinner than 30 cm and thicker than 70 cm, but do not capture the mode of the measured debris thickness, which occurs at 35 cm for the interpolated debris thickness, but at 55–65 cm for the modeled debris. The 'poor' (noon, exponential) model was unable to replicate the measured debris thickness patterns, indicated by a debris thickness strongly centered on 50 cm, strongly missing the presence of debris thinner than 25 cm or thicker than 75 cm (Fig. 7f). The 'poor' model underestimates the average debris thickness across the study area by seven cm. This suggests that, when using TIR imagery obtained late at night, or under cloudy conditions, when the coupling between surface temperature and debris thickness is the strongest, it is possible to derive debris thickness for thick debris up to 100 cm, beyond the thermal decoupling suggested to occur ~ 50 cm when the surface temperature was obtained during daytime.

The difference between the modeled and interpolated debris thickness is shown in Figs 7g–i. All models show an underestimation of debris thickness in the lower right region of the study area, where the thickest debris measurement was located. The difference between the 'good' modeled and the interpolated debris thickness (Fig. 7g) is mainly linked to the presence of large boulders, which are present in the modeled debris thickness but were not well captured but the smooth interpolation of the manual excavations across the study area. The 'average' model and the 'poor' model both over and underestimate the debris thickness across wider areas of the study area (Fig. 7h–i). The distribution of the difference between modeled and debris thicknesses is shown in Fig. S5.

The cooling rates at the debris surface were also analyzed to see how they relate to measured debris thickness but using surface temperature change to estimate debris thickness did not provide stronger model performances than the instantaneous surface temperatures (Fig. 8). Cooling rates were tested from hourly intervals to daily intervals. For example, using the temperature change between 12:00 (noon) and 15:00 on 6 August 2019 as a predictor of debris thickness in the exponential empirical model did not provide a better estimate than only using the surface temperature measured at 15:00. Similarly, using the surface temperature change between 18:00 and 00:00 (midnight) on 6 August 2019 showed a similar model performance as using the surface temperature at midnight. These two examples showcase early and

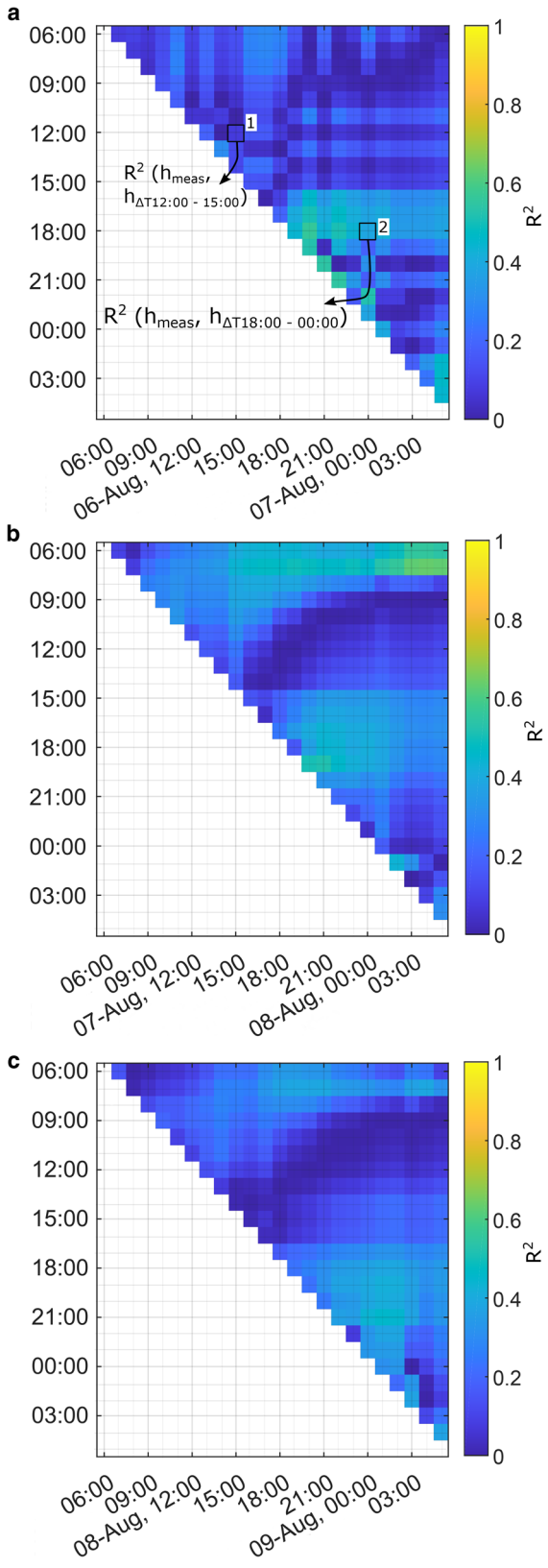


Fig. 8. Correlation of determination (R^2) between modeled ($h_{\Delta T}$) and measured debris thickness (h_{meas}) with the exponential model, when the empirical model is based on surface temperature change (d–f) for 6, 7, and 8 August 2019. The outlined cell (1) shows the R^2 value calculated from the measured debris thickness (h_{meas}) and modeled debris thickness obtained with an exponential empirical model based on a change in temperature between 6 August, 12:00 (noon) and 6 August, 15:00, ($h_{\Delta T\ 12:00-15:00}$). The outlined cell (2) shows the R^2 value calculated from the measured debris thickness (h_{meas}) and modeled debris thickness obtained with an empirical model based on the change in surface temperature between 6 August, 18:00 and 7 August, 00:00 ($h_{\Delta T\ 18:00-00:00}$).

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

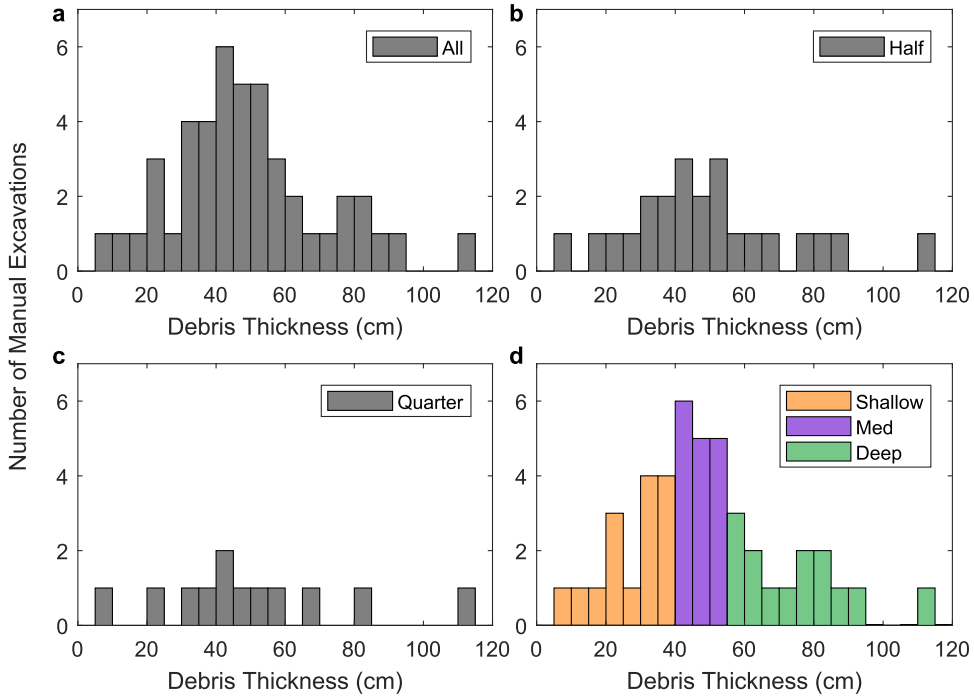


Fig. 9. Manual excavation distribution used to build the regression models for showing (a) all the validation points ($n = 44$), (b) half the points ($n = 22$), (c) a quarter for the points ($n = 11$) and the shallow, medium and deep validation points (d).

late-night cooling rates where a signal linked to the debris thickness might be expected to emerge. Hopkinson and others (2010) found that the active areas of the ice-cored moraine at Peyto Glacier peaked at a cooler temperature during the daytime and presented a more rapid cooling after sunset than the cooling experienced in the stable areas of the moraine and concluded that these cooler areas were indicators of buried ice but did not draw any conclusion on debris thickness. The results presented here suggest that, while the overnight temperature changes can be used to derive empirical models of debris thickness, these models do not provide better model performances than using instantaneous surface temperature. However, this analysis is limited by the lack of correction of the TIR images for atmospheric emissions along the path from the target to the imaging radiometer, which can change with time and influence the temperature change observed between the images and add noise to the measured temperature change.

Number and distribution of the manual excavations

The number of manual excavations used to determine the correlation of the regression models between surface temperature and debris thickness was varied following two approaches to estimate their effect on the correlation parameters and performances. First, the number of manual excavations was reduced to half and then to a quarter of the available measurements while attempting to conserve the range and distribution of the debris thickness (Figs 9a–c). This resulted in using 22 and then 11 measurement points instead of 44. Secondly, the manual excavations used were classified by depth: shallow (<35 cm), medium (35–50 cm) and deep (>50 cm) (Fig. 9d). The selection of this subset of data aimed to represent fieldwork scenarios when lower resources were available, causing a lower number of manual excavations to be performed and potentially resulting in a bias toward certain debris thicknesses in the empirical models that are developed.

Reducing the number of manual excavations used to derive the empirical model, while keeping a similar depth distribution of the validation points, shows an improvement in the model performance (Fig. 10a). When only using a quarter of the manual excavations to establish the exponential regression models, the maximum R^2 over the study period improved from 0.73 to 0.91 (Fig. 10a, Table 2). Comparing the three empirical models for 7 August 2019, at 16:50, the coefficient of determination improves from 0.54 to 0.66 and 0.73 when only half and a quarter of the available measurements were used, even though the models are highly similar (Figs 10b–d) and provided comparable spatial distribution for the modeled debris thickness (Figs 10e–f). However, the models using only one half, and one quarter of the manual excavations overestimated the mean thickness across the study area by 6.14 and 4.33 cm respectively. Those errors are substantially above the mean deviation reached when all the available measurements are used (less than one cm).

The modeled debris thickness distribution obtained with one quarter and one half of the manual excavations underestimated the presence of very thin debris (< 10 cm) and medium thickness debris (30–50 cm) but overestimated the presence of thick debris (> 75 cm). The TIR images used to derive these modeled debris thicknesses are shown in Figs S6a–c.

Using only manual excavations measurements corresponding to shallow, medium or thick debris to derive empirical regression models, shown in Fig. 9d, lead to very different patterns in model performance (Fig. 11). When only shallow or medium debris was used in the empirical models, the resulting modeled debris thickness was highly uniform and strongly underestimated the interpolated measurements across the study area (Figs 11e–f, h–i). When using only thick debris, the modeled debris thickness showed a higher range of debris with a spatial distribution closer to the interpolated debris measurements, but strongly overestimated the mean debris thickness over the study area (Figs 11g, j). The TIR images used to derive these modeled debris thicknesses are shown in Figs S6d–f. These results illustrate how an apparently strong model performance based on a limited number of ground observations can lead to inaccurate modeled debris thicknesses. For example, strong model performance was obtained when building a regression model on a few measurements, but the

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

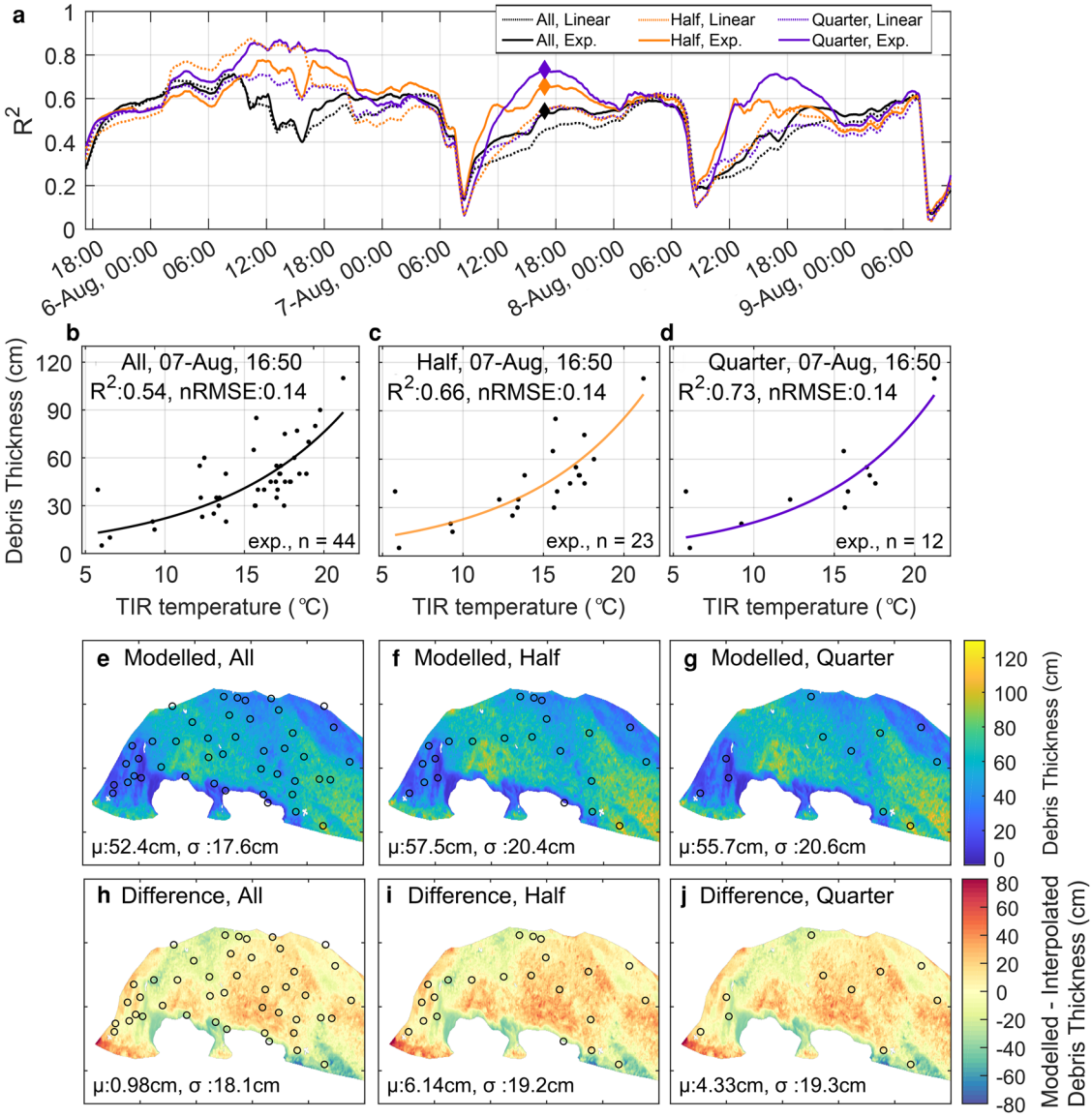


Fig. 10. Coefficient of determination (R^2) obtained when using all, half or a quarter of the manual excavations in the regression model (a), with the regression models the TIR image taken on 07-Aug, 16:50 showed in (b–d), along with the coefficient of determination and the normalized RMSE. The model type (exponential or linear) and the number of point manual excavations used are shown in the bottom left of the panel. The timing of the (d–e) scatter plots is indicated by the diamond on the (a) panel. The resulting modeled debris thickness for the study area is shown in (e–g) and the difference between the modeled and interpolated debris thickness from the manual excavations is shown in (h–j). For the panels (e–j), the mean and the standard deviation are shown, and the locations of manual excavation used in the regression model are shown as circles.

modeled variability in the debris thickness was wrong. Similarly, when a bias existed in the measurements used in the regression model, the average depth simulated over the study area was incorrect. This emphasizes the limitations of developing empirical models in highly heterogeneous environments with low numbers of observations and indicates that careful interpretation of statistical analysis is warranted.

Spatial resolution

Variation in the coefficient of determination was also explored by varying the spatial resolution of the surface temperature data. The spatial resolution of the images was artificially degraded to obtain images at resolutions 10, 20 and 30 times lower than the original images from TIR1. The original TIR image provided a pixel resolution of 1000×350 over the study area, which was degraded to 36×100, 18×50 and 12×34 pixels. These correspond to pixel sizes of 0.06×0.2 m (original), 0.6×2.0 m, 1.2×4 m and 1.8×6 m, respectively. Further degrading the spatial resolution was not possible due to the small area covered by the study area and would have led to not having enough points to develop the correlations. Even though these spatial resolutions are higher than the ones available from satellite-based thermal imagery, they are comparable to those from UAV or plane-based instruments. For each degraded resolution, a linear and exponential regression model was fit, and the coefficient of determination was calculated between modeled and measured debris thickness for each image (Fig. 12a). Changing the spatial resolution of the image caused a moderate decrease in the model performance, but a strong impact on the small-scale variability of the calculated debris thickness. Decreasing the spatial resolution produced a similar overall pattern of debris thickness compared to the high-resolution images, but the debris thickness variability linked to individual clusters of boulders, for example, was not captured (Figs 12e–j). This suggests that high-resolution images are not necessary to estimate the overall pattern of debris thickness and that high-resolution images only introduce noise due to microtopography. However, this comparison is limited by the use of the interpolated debris thickness as the reference, as it does not capture small-scale variability in debris thickness and microtopography. The TIR images used to derive these modeled

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

Table 2. Summary of model performance for the number and depth of manual excavations as well as spatial resolution scenarios

	R^2_{mean}	R^2_{min}	R^2_{max}	$nRMSE_{mean}$	$nRMSE_{min}$	$nRMSE_{max}$
Normal	0.51 (0.48)	0.05 (0.05)	0.73 (0.73)	0.15 (0.16)	0.13 (0.13)	0.21 (0.21)
Half	0.56 (0.51)	0.07 (0.06)	0.8 (0.73)	0.16 (0.17)	0.11 (0.14)	0.23 (0.23)
Quarter	0.58 (0.52)	0.03 (0.02)	0.91 (0.89)	0.18 (0.2)	0.1 (0.12)	0.28 (0.29)
Shallow	0.47 (0.52)	0 (0)	0.76 (0.79)	0.24 (0.23)	0.17 (0.16)	0.33 (0.33)
Medium	0.3 (0.3)	0 (0)	0.65 (0.65)	0.35 (0.36)	0.25 (0.24)	0.42 (0.42)
Deep	0.31 (0.3)	0 (0)	0.61 (0.56)	0.27 (0.28)	0.2 (0.21)	0.38 (0.38)
Resolution/10	0.45 (0.43)	0.05 (0.05)	0.64 (0.67)	0.16 (0.17)	0.13 (0.13)	0.24 (0.24)
Resolution/20	0.37 (0.32)	0 (0)	0.57 (0.6)	0.18 (0.18)	0.14 (0.13)	0.24 (0.24)
Resolution/30	0.33 (0.28)	0.02 (0.02)	0.59 (0.57)	0.18 (0.19)	0.15 (0.16)	0.24 (0.24)

Values refer to the exponential model, and values in parentheses refer to the linear model

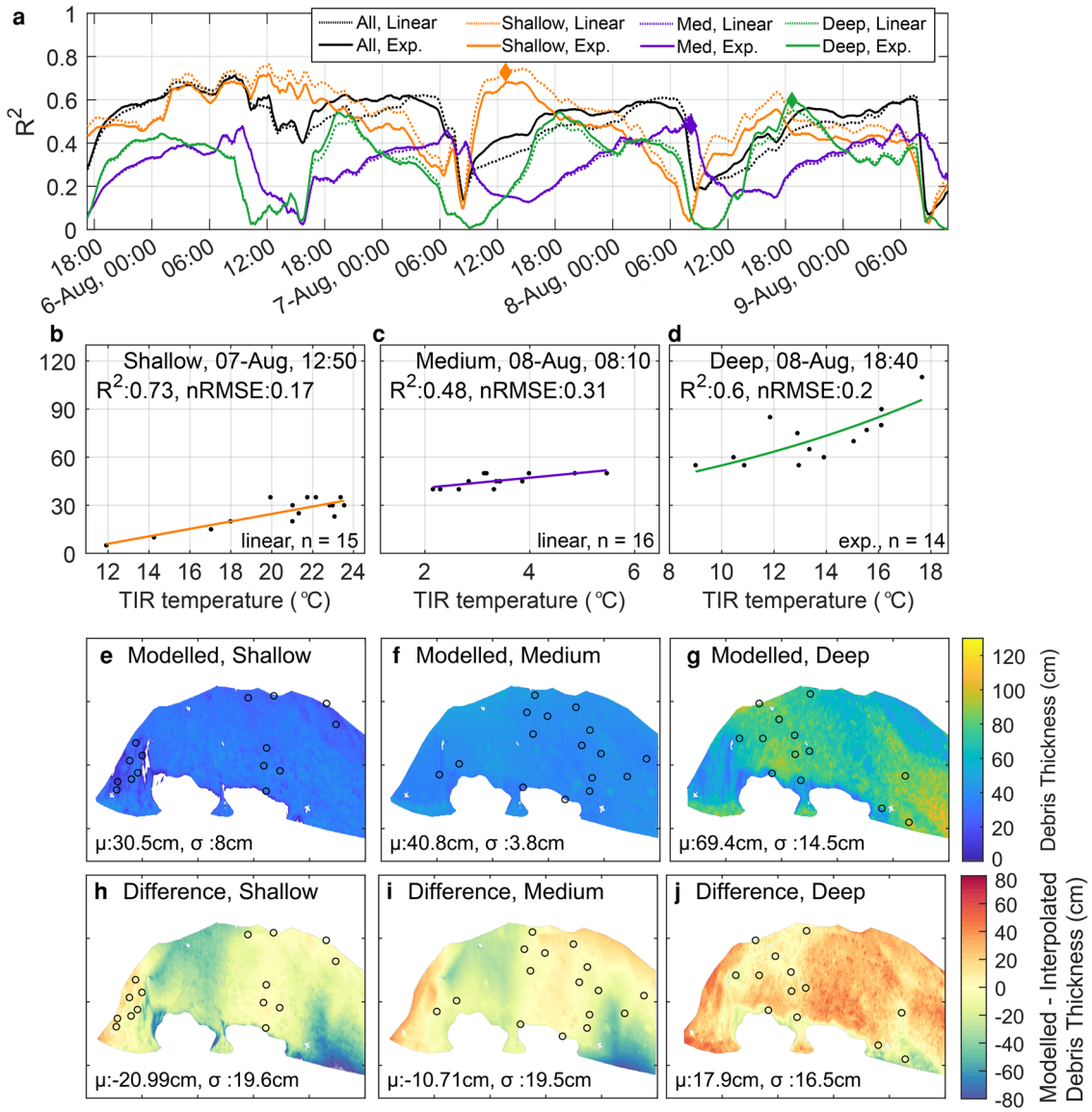


Fig. 11. Coefficient of determination (R^2) obtained when using only shallow, medium or deep manual excavations in the regression model (a), with an example of good performing models shown in (b–d), along with the coefficient of determination and the normalized RMSE. The model type (exponential or linear) and the number of point manual excavations used are shown in the bottom left of the panel. The timing of the (d–e) scatter plots is indicated by the diamond on the (a) panel. The resulting modeled debris thickness for the study area is shown in (e–g) and the difference between the modeled and interpolated debris thickness from the manual excavations is shown in (h–j). For the panels (e–j), the mean and the standard deviation are shown, and the locations of manual excavation used in the regression model are shown as circles.

debris thicknesses are shown in Fig. S6g–i. The summary statistics of model performance for the number and distribution of measured debris thickness used in the regression model as well as the spatial resolution of the TIR images are shown in Table 2 (Fig. 12).

Camera location and angle

Another factor influencing data collection is the distance and angle between the camera and the study area. A camera located further away will cause a decrease in spatial resolution, but

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

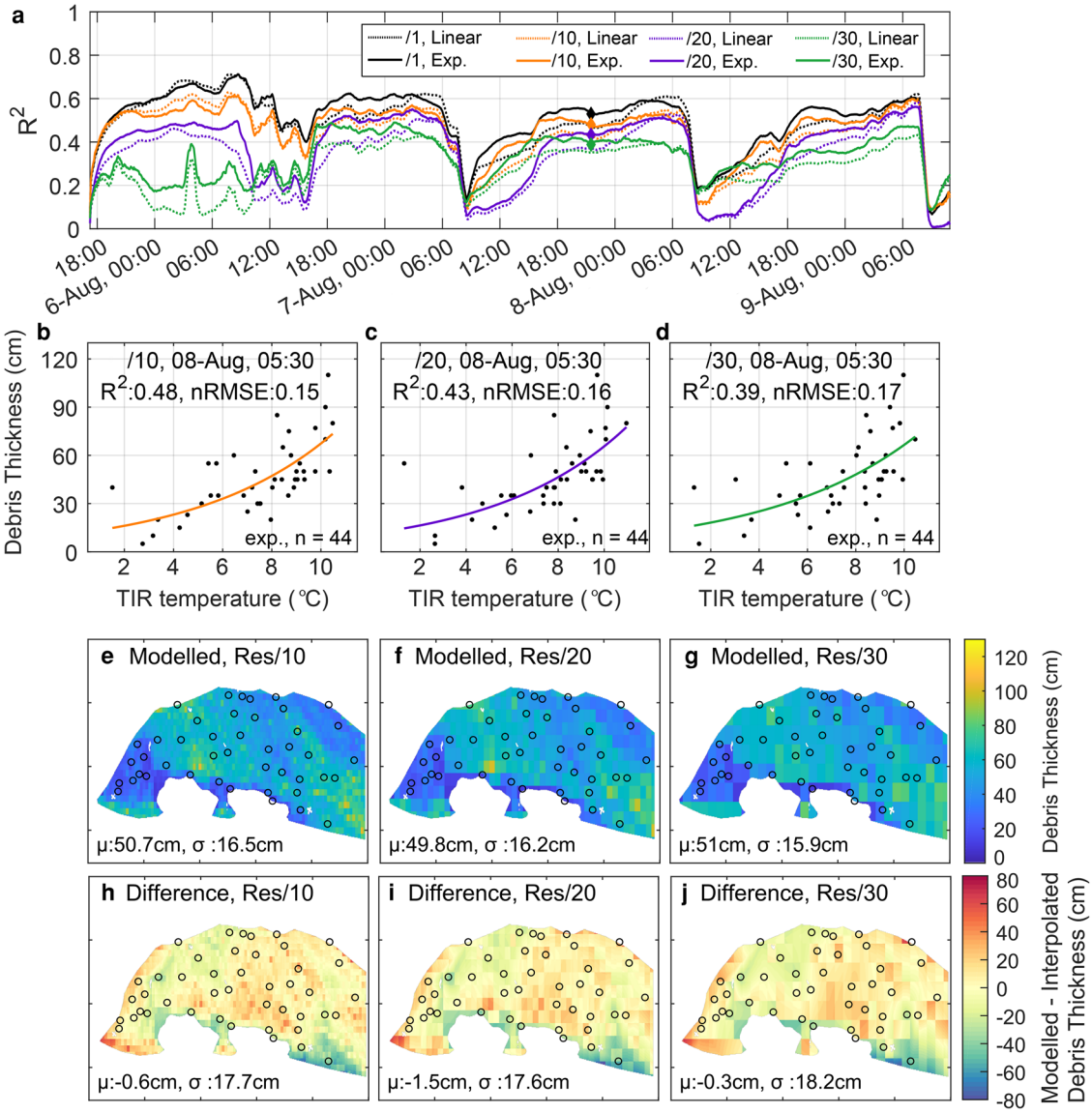


Fig. 12. Coefficients of determination (R^2) obtained when using all, half or a quarter of the manual excavations in the regression model (a), with the regression models the TIR image taken on 07-Aug, 16:50 showed in (b–d), along with the coefficient of determination and the normalized RMSE. The model type (exponential or linear) and the number of point manual excavations used are shown in the bottom left of the panel. The timing of the (d–e) scatter plots is indicated by the diamond on the (a) panel. The resulting modeled debris thickness for the study area is shown in (e–g) and the difference between the modeled and interpolated debris thickness from the manual excavations is shown in (h–j). For the panels (e–j), the mean and the standard deviation are shown, and the locations of manual excavation used in the regression model are shown as circles.

other sources of error can appear, such as atmospheric interference along the optical path between the camera and the study area. To assess how this might affect calculated debris thickness, the average measured surface temperature for the study area captured from the different TIR locations was compared. This showed that the surface temperatures obtained by the four camera locations followed a similar pattern with similar average measured temperatures, even though the distance between the study area and the camera, the camera view angle, and the number of pixels of the study area greatly differed amongst the shots (Fig. 13). TIR1 was located 140 m from the study area, while TIR2 and TIR3 were at 290 and 370 m, and TIR4 was only six m from the ice cliffs. These distances were relatively short compared to other oblique imagery applications (945–1145 m in Aubry-Wake and others, 2015; 10–1000 m in Herreid, 2021). Overall, the standard deviations of the surface temperature of continuous debris measured with TIR1, TIR2, and TIR3 were very similar as well. This suggests that TIR imagery is a robust approach to monitor debris surface temperature and does not strongly depend on the distance and view angle, as long as the distance does not allow for significant atmospheric disturbances between the target area and the camera (Aubry-Wake and others, 2015; Baker and others, 2019; Herreid, 2021). This conclusion might be due to the specific atmospheric conditions and short distances of this study site, and more efforts should be put forward to understand the environmental conditions that impact atmospheric emission of TIR along the path from target to imaging radiometer, and how these emissions affect TIR imagery in glacial environments.

Considering that the surface temperature between the different camera locations follows similar patterns of absolute temperature, one could assume that the resulting calculated debris thickness provides similar results. However, this is difficult to assess with the TIR images in this study. Due to the changing camera angle and the increased distances, most validation points and TIR targets were not discernable in the images from TIR2-4. Regression models could not be derived specifically for these images. When the linear and exponential regression equations derived from TIR1 are used on TIR2, TIR3 and TIR4 images,

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

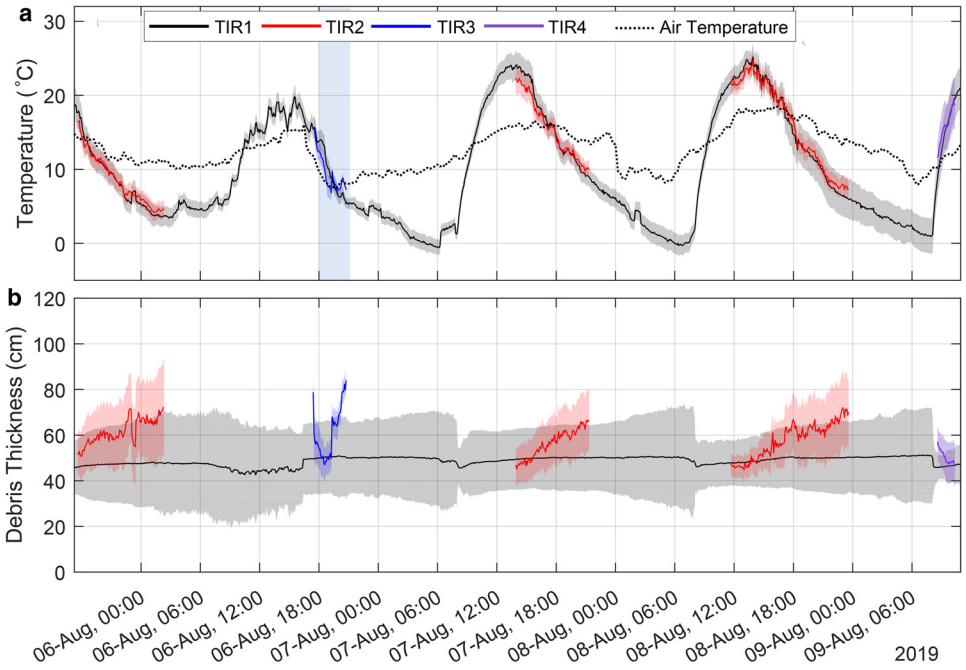


Fig. 13. TIR-derived average surface temperature for the study area (a) for location TIR1 (grey), TIR2 (red), TIR3 (blue) and TIR4 (purple). The mean temperature is represented by a solid line and the standard deviation is illustrated by the shading. Air temperature is shown as a dotted line for comparison. The modeled debris thickness using the exponential model from the TIR1 location is shown in (b), for the TIR1 (grey), TIR2 (red), TIR3 (blue) and TIR4 (purple) locations.

modeled average debris thicknesses over the study area are very different (Fig. 13b). This emphasizes the limitations of empirical model transferability and their applications outside the specific setting in which they were developed. A future experiment comparing surface temperature from TIR imagery from different perspectives and distances should ensure that the difference in camera location provides a strong overlap in the view of the study area to allow matching geospatial localization. Specifically, the localization targets should be visible from all the camera perspectives. Such an experiment could be used to assess not only average surface temperature but also assess the changes in longwave radiation transmission between the camera and the study area.

Debris thermal emissivity

The TIR-derived surface temperatures used in the empirical debris thickness models are radiant temperatures obtained under the assumption that the debris cover has a uniform emissivity of 1.0 and thus behaves as a perfect blackbody. This allows more straightforward processing of the images, and as the TIR-derived temperatures are not analyzed for accuracy of the absolute temperature, but only to derive the relationship with debris thickness, this possible bias in surface temperature due to the lack of emissivity correction cannot influence the results. However, some small-scale variability in surface exitance of longwave radiation linked to the varying lithology, sky view, reflectance, slope and aspect, and reflectance of thermal irradiance is likely, even though this small-scale variability is mitigated by the uniform rock debris cover type analyzed in the study area, as the ice cliffs were removed from the analysis. However, despite the cover type of the study area being uniformly rock debris, it presented varying clast size and specific lithology, which also affects the emissivity (Salisbury and D'Aria, 1994; Sobrino and Cuenca, 1999). Additionally, emissivity decreases when the view-angle moves further from nadir with the decrease following different curves for varying surface types (Sobrino and Cuenca, 1999). Individual boulders, therefore, present varying viewing slopes relative to the camera, shifting the emissivity. This small-scale variability in emissivity, reflectance, irradiance and hence exitance may have added noise to the correlation between radiant surface temperature and debris thickness. In this study, this was mitigated by having a uniform slope in the area with a small study area with consistent lithology.

It should be noted that given the short distance between the camera and the study area (140 m) and low absolute humidity in the cool, unsaturated air, atmospheric TIR emissions along the path from target to imaging radiometer are likely to be minimal. The distance from the TIR imager to the furthest and closest points in the study area is very close, and therefore, any atmospheric TIR emissions are likely to be relatively uniform for each image. Atmospheric TIR emission can vary over time, but as each image was assessed independently and so this has minimal effect on the analysis.

Suggestions for optimal TIR surveys

Modeled debris thickness was highly sensitive to biases in the range of measured debris thickness used in the empirical models, with a strongly weakened predictive capacity at the study area scale when the regression models were developed using only shallow, medium or thick debris thicknesses. This supports findings by Boxall and others (2021), who drew similar conclusions when using empirical models to estimate debris thickness at the regional scale in High Mountain Asia. When comparing different empirical debris thickness models, Boxall and others (2021) also suggested that optimal model type was related to debris thickness, with linear models performing better for thinner debris, which is observed in this study as well.

The number of debris thickness measurements used in the regression models, as well as the spatial resolution of the TIR images, had a relatively smaller impact on modeled debris thickness. Even when reducing spatial resolution by a factor of three, reasonable modeled debris thickness was obtained despite a decrease in model performance. This suggests spatial resolution, once high enough to capture patterns in slope and angle and resolve features such as ice cliffs, does not need to be a key priority when estimating

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press

debris thickness from surface temperature. In this study, a spatial resolution of 0.6×2.0 m or 1.2×4 m, instead of the original 0.06×0.2 m, would have been adequate to capture the debris cover pattern on the study area while smoothing over the microtopography that causes additional noise in the relationship. Similarly, based on this analysis, the temporal resolution used in this study, with image acquisition occurring every five minutes, was superfluous. Images acquired hourly under stable weather conditions, or every 15 to 30 min interval during quickly changing weather and ~ sunrise would have been sufficient, with the added benefit of preserving battery power and allowing longer data collection.

Due to the highly variable model performance that was due to a large range of the factors analyzed here, the empirical models presented in this study have no validity or known performance outside of the four days and the small study area used in their generation. Developing debris thickness empirical models that hold transferability in time or space approaches using the information in the infrared images to constrain the exponential scaling, such as in Kraaijenbrink and others (2017) and Herreid (2021) should be further explored.

Conclusions

This study explored how the correlation between oblique ground-based TIR-observed surface temperature and debris thickness over ice is influenced by data collection methods and weather conditions in the field. A total of 1478 thermal-infrared images were collected on 5–9 August 2019 using two TIR imaging radiometers at four locations ~ a small ice-cored moraine complex at the edge of Peyto Glacier, in the Canadian Rockies. These images were used to produce high spatial and temporal resolution surface temperature maps of the ice-cored moraine. Along with these maps, 44 excavations of the moraine were made to measure debris thickness above the ice and surface temperatures were correlated to debris thicknesses.

This study presents a thorough analysis of the different factors relating to weather conditions and data collection approaches that influence the empirical relationship between surface temperature and debris thickness. High-resolution temperature maps of debris-covered glaciers, such as those processed by Kraaijenbrik and others (2018) and Tarca and Guglielmin (2022), provide a valuable snapshot of surface temperature distribution but only allow limited investigation of how the relationship between surface temperature and debris thickness varies throughout the day. Similarly, previous studies linking surface temperature to debris thickness using empirical relationships were limited both in spatial and temporal resolution by the use of satellite imagery (Mihalcea and others, 2008a, 2008b; Minora and others, 2015). Therefore, this study provides a substantial advance in the understanding of surface temperature and debris-cover variability.

Using 44 debris thickness measurements ranging from five to 110 cm, a linear and exponential regression model between debris thickness and surface temperature was generated for each image. A pattern emerged, where an exponential model provided better performance than the linear model throughout the day, but the linear model performed better than the exponential model for a short period in late-night conditions, reaching R^2 values of 0.71 and nRMSE of 0.13 cm. Based on these results, the most reliable times to obtain surface temperatures that correlate well to debris thickness are from late at night to just before sunrise under clear conditions or any time of the day or night under cloudy conditions.

The findings of this study have several practical implications for future efforts involving the acquisition of TIR-derived surface temperature. This is of obvious use for debris thickness calculations as presented in this work but can also be of use to a broader range of geophysical investigations using surface temperatures, such as studies in the field of permafrost, soil moisture, surface water-groundwater interactions or volcanology. Beyond presenting practical implications for TIR radiometer data collection, this study provides an initial insight into the small-scale radiative and conductive controls of surface temperature in debris over glacial ice.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/jog.2022.67

Acknowledgements. The authors wish to acknowledge funding from the Natural Sciences and Engineering Research Council of Canada in the Discovery Grants and Vanier and Michael Smith Scholarship programs, Alberta Innovation, the Canada Research Chairs program and the Canada First Excellence Research Fund's Global Water Futures program. We thank two anonymous reviewers for their constructive comments that helped to improve the manuscript significantly.

Data availability. The scripts used to process and analyze the TIR imagery can be found at: https://github.com/caubrywake/PeytoIR and the thermal infrared images (raw and processed), meteorological data and excavations data can be found at: https://www.hydroshare.org/resource/d2768996e761412381c2051b99d02c87/.

References

Anderson LS and Anderson RS (2018) Debris thickness patterns on debris-covered glaciers. *Geomorphology* **311**, 1–12. doi: 10.1016/j.geomorph.2018.03.014

Aubry-Wake C and 7 others (2015) Measuring glacier surface temperatures with ground-based thermal infrared imaging. *Geophysical Research Letters* **42**(20), 8489–8497. doi: 10.1002/2015GL065321

Baker EA, Lautz LK, McKenzie JM and Aubry-Wake C (2019) Improving the accuracy of time-lapse thermal infrared imaging for hydrologic applications. *Journal of Hydrology* **571**, 60–70. doi: 10.1016/j.jhydrol.2019.01.053

Boxall K, Willis I, Giese A and Liu Q (2021) Quantifying patterns of supraglacial debris thickness and their glaciological controls in High Mountain Asia. *Frontiers in Earth Science* **9**, 504. doi: 10.3389/feart.2021.657440

Cardenas MB, Harvey JW, Packman AI and Scott DT (2008) Ground-based thermography of fluvial systems at low and high discharge reveals potential complex thermal heterogeneity driven by flow variation and bioroughness. *Hydrological Processes* **22**(7), 980–986. doi: 10.1002/hyp.6932

Foster LA, Brock BW, Cutler MEJ and Diotri F (2012) A physically based method for estimating supraglacial debris thickness from thermal band remote-sensing data. *Journal of Glaciology* **58**(210), 677–691. doi: 10.3189/2012JoG11J194

Giese A, Arcone S, Hawley R, Lewis G and Wagnon P (2021) Detecting supraglacial debris thickness with GPR under suboptimal conditions. *Journal of Glaciology* **67**(266), 1108–1120. doi: 10.1017/JOG.2021.59

Giese A, Boone A, Wagnon P and Hawley R (2020) Incorporating moisture content in surface energy balance modeling of a debris-covered glacier. *The Cryosphere* **14**(5), 1555–1577. doi: 10.5194/TC-14-1555-2020

Herreid S (2021) What can thermal imagery tell us about glacier melt below rock debris? *Frontiers in Earth Science* **9**, 738. doi: 10.3389/feart.2021.681059

Herreid S and Pellicciotti F (2020) The state of rock debris covering earth's glaciers. *Nature Geoscience* **13**(9), 621–627. doi: 10.1038/s41561-020-0615-0

Hopkinson C, Barlow J, Demuth MN and Pomeroy JW (2010) Mapping changing temperature patterns over a glacial moraine using oblique thermal imagery and lidar. *Canadian Journal of Remote Sensing* **36**, S257–S265. doi: 10.5589/m10-053

Juen M, Mayer C, Lambrecht A, Han H and Liu S (2014) Impact of varying debris cover thickness on ablation: a case study for Koxkar Glacier in the Tien Shan. *The Cryosphere* **8**(2), 377–386. doi: 10.5194/tc-8-377-2014

Kraaijenbrink PDA and 6 others (2018) Mapping surface temperatures on a debris-covered glacier with an unmanned aerial vehicle. *Frontiers in Earth Science* **6**, 64. doi: 10.3389/feart.2018.00064

Kraaijenbrink PDA, Bierkens MFP, Lutz AF and Immerzeel WW (2017) Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers. *Nature* **549**(7671), 257–260. doi: 10.1038/nature23878

MathWorks (2017) *MATLAB and the Image Processing Toolbox, Version 9.2.0 (R2017a)*. Natick, Massachusetts, USA: The MathWorks Inc.

McCarthy M, Pritchard H, Willis I and King E (2017) Ground-penetrating radar measurements of debris thickness on Lirung Glacier, Nepal. *Journal of Glaciology* **63**(239), 543–555. doi: 10.1017/jog.2017.18

Mihalcea C and 7 others (2008a) Using ASTER satellite and ground-based surface temperature measurements to derive supraglacial debris cover and thickness patterns on Miage Glacier (Mont Blanc Massif, Italy). *Cold Regions Science and Technology* **52**(3), 341–354. doi: 10.1016/j.coldregions.2007.03.004

Mihalcea C and 7 others (2008b) Spatial distribution of debris thickness and melting from remote-sensing and meteorological data, at debris-covered Baltoro glacier, Karakoram, Pakistan. *Annals of Glaciology* **48**, 49–57. doi: 10.3189/172756408784700680.

Minora U and 10 others (2015) A simple model to evaluate ice melt over the ablation area of glaciers in the Central Karakoram National Park, Pakistan. *Annals of Glaciology* **56**(70), 202–216. doi: 10.3189/2015AoG70A206

Nicholson L and Benn DI (2006) Calculating ice melt beneath a debris layer using meteorological data. *Journal of Glaciology* **52**(178), 463–470. doi: 10.3189/172756506781828584

Nicholson L and Benn D (2013) Properties of natural supraglacial debris in relation to modelling sub-debris ice ablation. *Earth Surface Processes and Landforms* **38**(5), 490–501. doi: 10.1002/esp.3299

Nicholson L, McCarthy M, Pritchard HD and Willis I (2018) Supraglacial debris thickness variability: impact on ablation and relation to terrain properties. *The Cryosphere* **12**(12), 3719–3734. doi: 10.5194/tc-12-3719-2018

Østrem G (1959) Ice melting under a thin layer of moraine, and the existence of ice cores in moraine ridges. *Geografiska Annaler* **41**(4), 228–230. doi: 10.1080/20014422.1959.11907953

Pradhananga D and 8 others (2021) Hydrometeorological, glaciological and geospatial research data from the Peyto Glacier Research Basin in the Canadian Rockies. *Earth System Science Data* **13**(6), 2875–2894. doi: 10.5194/essd-13-2875-2021

Reid T, Carenzo M, Pellicciotti F and Brock BW (2012) Including debris cover effects in a distributed model of glacier ablation. *Journal of Geophysical Research Atmospheres* **117**(17), 1–15. doi: 10.1029/2012JD017795

Rounce DR and 10 others (2021) Distributed global debris thickness estimates reveal debris significantly impacts glacier mass balance. *Geophysical Research Letters* **48**(8), e2020GL091311. doi: 10.1029/2020GL091311

Rounce DR, King O, McCarthy M, Shean DE and Salerno F (2018) Quantifying debris thickness of debris-covered glaciers in the everest region of Nepal through inversion of a subdebris melt model. *Journal of Geophysical Research: Earth Surface* **123**(5), 1094–1115. doi: 10.1029/2017JF004395

Rounce DR and McKinney DC (2013) Debris thickness of glaciers in the everest area (Nepal Himalaya) derived from satellite imagery using a nonlinear energy balance model. *The Cryosphere* **8**(4), 1317–1329. doi: 10.5194/tc-8-1317-2014

Rounce DR, Quincey DJ and McKinney DC (2015) Debris-covered glacier energy balance model for Imja-Lhotse Shar Glacier in the Everest region of Nepal. *The Cryosphere* **9**(6), 2295–2310. doi: 10.5194/tc-9-2295-2015

Salisbury JW and D'Aria DM (1994) Emissivity of terrestrial materials in the 3-5 µm atmospheric window. *Remote Sensing of Environment* **47**(3), 345–361. doi: 10.1016/0034-4257(94)90102-3

Schauwecker S and 7 others (2015) Remotely sensed debris thickness mapping of Bara shigri glacier, Indian Himalaya. *Journal of Glaciology* **61**(228), 675–688. doi: 10.3189/2015JoG14J102

Shah SS, Banerjee A, Nainwal HC and Shankar R (2019) Estimation of the total sub-debris ablation from point-scale ablation data on a debris-covered glacier. *Journal of Glaciology* **65**(253), 759–769. doi: 10.1017/jog.2019.48

Sobrino JA and Cuenca J (1999) Angular variation of thermal infrared emissivity for some natural surfaces from experimental measurements. *Applied Optics* **38**(18), 3931. doi: 10.1364/ao.38.003931

Stewart RL and 6 others (2021) Using climate reanalysis data in conjunction with multi-temporal satellite thermal imagery to derive supraglacial debris thickness changes from energy-balance modelling. *Journal of Glaciology* **67**(262), 366–384. doi: 10.1017/JOG.2020.111

Tarca G and Guglielmin M (2022) Using ground-based thermography to analyse surface temperature distribution and estimate debris thickness on Gran Zebrù glacier (Ortles-Cevedale, Italy). *Cold Regions Science and Technology* **196**, 103487. doi: 10.1016/j.coldregions.2022.103487

WGMS (2021) *Fluctuations of Glaciers Database*. Zurich, Switzerland: World Glacier Monitoring Service. doi: 10.5904/wgms-fog-2021-05

https://doi.org/10.1017/jog.2022.67 Published online by Cambridge University Press