



Article

*Present address: Department of Geography, University of British Columbia, Canada.

Cite this article: Winter-Billington A, Moore RD, Dadic R (2020). Evaluating the transferability of empirical models of debris-covered glacier melt. *Journal of Glaciology* 66 (260), 978–995. <https://doi.org/10.1017/jog.2020.57>

Received: 3 October 2019
Revised: 1 July 2020
Accepted: 2 July 2020
First published online: 8 September 2020

Keywords:

Debris-covered glaciers; glacier modeling; melt-surface; supraglacial debris

Author for correspondence:

A. Winter-Billington, E-mail: winter-billington@alumni.ubc.ca

Evaluating the transferability of empirical models of debris-covered glacier melt

A. Winter-Billington¹ , R. D. Moore¹ and R. Dadic²

¹Department of Geography, University of British Columbia, Canada and ²Antarctic Research Centre, Victoria University of Wellington, New Zealand

Supraglacial debris is significant in many regions and complicates modeling of glacier melt, which is required for predicting glacier change and its influences on hydrology and sea-level rise. Temperature-index models are a popular alternative to energy-balance models when forcing data are limited, but their transferability among glaciers and inherent uncertainty have not been documented in application to debris-covered glaciers. Here, melt factors were compiled directly from published studies or computed from reported melt and MERRA-2 air temperature for 27 debris-covered glaciers around the world. Linear mixed-effects models were fit to predict melt factors from debris thickness and variables including debris lithology and MERRA-2 radiative exchange. The models were tested by leave-one-site-out cross-validation based on predicted melt rates. The best model included debris thickness (fixed effect) and glacier and year (random effects). Predictions were more accurate using MERRA-2 than on-site air temperature data, and pooling MERRA-2-derived and reported melt factors improved cross-validation accuracy more than including additional predictors such as shortwave or longwave radiation. At one glacier where monthly ablation was measured over 4 years, seasonal variation of melt factors suggested that heat storage significantly affected the relation between melt and energy exchange at the debris surface.

Introduction

Glacier melt contributes significantly to streamflow and is important for hydro-electric power generation and irrigation in many places around the world (IPCC, 2014). Glacier melt is also contributing to the global sea-level response to climate change (Radić and Hock, 2014). To address the effects of ongoing and future glacier changes on water resources, glacier-related hazards and sea-level rise, managers and policy-makers depend on the output from melt models driven by climate data. Being able to quantify the uncertainties in estimated glacier melt provides decision-makers a basis for assessing risks associated with different management options.

The physical processes controlling glacier surface melt are well understood and energy-balance models are accurate at a range of scales (Hock, 2005). However, energy-balance models are computationally demanding and require extensive input data that may not be available or reliable over large regions. Temperature-index models (TIM) only require air temperature as an input, which can often be interpolated from ground-based weather stations or extracted from reanalysis products with reasonable accuracy at the basin scale (Stahl and others, 2006; Rienecker and others, 2011; Trubilowicz and others, 2016; Gelaro and others, 2017). TIM are adequate for a wide range of applications and can even exceed the performance of energy-balance models when forcing data are limited (Gabbi and others, 2014; Réveillet and others, 2018). However, the physical processes explicitly represented in energy-balance models are reduced to simplified, site-specific empirical parametrizations in TIM, which may limit their transferability.

Debris-covered glaciers are a relatively small but significant part of the global cryosphere. Estimates vary, but recent work found 4.39% (Scherler and others, 2018) to 16.8% (Sasaki and others, 2016) of the global glacierized area outside Greenland and Antarctica to be debris-covered – up to 47.4% by region. Theory, models and recent observations suggest the extent of supraglacial debris cover is increasing as glaciers recede (Bozhinskiy and others, 1986; Mayer and others, 2011; Scherler and others, 2018). The accuracy of models of debris-covered glacier melt is therefore an increasingly important area of research.

The surface exchanges of mass and energy that lead to ice melting under debris are different to those of snow, firn and ice, because debris temperatures can exceed 0°C. In most conditions, sublimation is expected to be negligible under debris so that surface ablation occurs entirely by melt (Bozhinskiy and others, 1986; Conway and Rasmussen, 2000). Like melt of bare ice, sub-debris melt has been well reproduced by energy-balance models applied at point and glacier scales (Nakawo and Young, 1982; Haidong and others, 2006; Reid and Brock, 2010; Reid and others, 2012; Lejeune and others, 2013; Brown and others, 2014; Collier and others, 2014). However, the observational data to develop, calibrate and drive energy-balance models are as rare in debris-covered environments as elsewhere. As an alternative, a number of empirical models based on the temperature index approach have been developed for sub-debris melt (Booker and Dunbar, 2008; Carezzo and others, 2016; Möller and others, 2016).

© The Author(s) 2020. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Table 1. Classical sub-debris melt factors, k ($\text{mm}^\circ\text{C}^{-1}\delta t^{-1}$), that have been applied to model sub-debris melt with time intervals (δt) of an hr = hour, d = day or m = month

Study area	k provenance	k	δt	References
Langtang Khola Basin, Central Himalaya	Modeled	2.4 to 6.5	d	Braun and others (1993)
Langtang Khola Basin, Central Himalaya	Modeled	2.1	d	Konz and others (2007)
Langtang Khola Basin, Central Himalaya	Modeled	$0.7 \cdot k_i^R$ or $0.15 \cdot k_i^R$	d	Immerzeel and others (2012)
Langtang Khola Basin, Central Himalaya	Modeled	$0.5 \cdot k_i$ or $0.58 \cdot k_i$, ($5.0 \leq k_i \leq 12.5$)	d	Pradhananga and others (2014)
Upper Indus, Hindu Kush, Karakoram and Western Himalaya	Modeled	Log-Gaussian distribution, $\mu = 2$, $\sigma = 2$	d	Immerzeel and others (2015)
Lirung Glacier, Central Himalaya	Modeled	$0.5 \cdot k_i^R$	hr	Parajuli and others (2015)
Indus Basin	Defined <i>a priori</i>	4, 6.5 or 9	d	Koppes and others (2015)
Dudh Koshi Basin, Central Himalaya	Literature-derived	3.34	d	Shea and others (2015)
Echaurren Norte, Chilean Andes	Empirically derived using observed glacier-wide mass balance	90.1	m	Masiokas and others (2016)

Subscript i is for bare ice, snow or both. h is debris thickness (m), and k^R denotes a restricted melt factor calculated with air temperature as one of a set of predictor variables.

Table 2. Restricted sub-debris melt factors, k^R , calculated with air temperature as one of a set of predictors, that have been applied to model sub-debris melt

Study area	k^R provenance	k^R	δt	References
Keqicar Baqi, Tien Shan	Unspecified	2	d	Zhang and others (2007)
Hunza Basin, Karakoram	Defined <i>a priori</i>	0.06	d	Ragettli and others (2013)
Langtang Khola Basin, Central Himalaya	Empirically derived using modeled melt	$0.03 \cdot h^{-0.8}$	hr	Ragettli and others (2015)
Shaune Garang, Western Himalaya	Empirically derived using observed melt	2.6 to 9.3	d	Kumar and others (2016)
Miage Glacier, Italian Alps and Haut Glacier d'Arolla, Swiss Alps	Empirically derived using modeled melt	0.636 to 2.362	hr	Carenzo and others (2016)
Rio del Yeso, Chilean Andes	Empirically derived using modeled melt	$0.017 \cdot h^{-0.4043}$	hr	Ayala and others (2016)
Inylchek glacier basin, Tien Shan	Empirically derived using observed melt	$0.0011 \cdot h^{-0.3625}$	d	Hagg and others (2018)

Units and notation as in Table 1.

Observations show that melt rate and its sensitivity to meteorological forcing decreases as debris thickness (h) increases from an effective thickness (h_c) of around 0.02–0.05 m (Østrem, 1959; Reznichenko and others, 2010; Singh and others, 2019). Many debris-covered glacier TIM include a dependence on debris thickness and some include additional variables such as incident solar radiation (e.g. Zhang and others, 2007; Ragettli and others, 2013; Ayala and others, 2016; Carenzo and others, 2016; Möller and others, 2016; Hagg and others, 2018). Few studies have quantified the transferability of empirical sub-debris melt models, or the associated uncertainty in melt estimates.

The key parameter in TIM is the melt factor (k) that relates melt to air temperature. Melt factors have been applied to predict melt of debris-covered glaciers in at least 16 reported studies. The melt factors applied in those 16 studies, listed in Tables 1 and 2, vary over two orders of magnitude. Only three of the 16 values were calculated using measurements of melt; the remaining 13 melt factors were calculated using melt modeled with an energy-balance scheme, modeled as a function of bare ice melt, modeled using an assumed statistical distribution, or defined *a priori*. There is currently no statistically robust method for selecting melt factors for debris-covered glaciers where data for calibration are unavailable, and the accuracy of predictions made with transferred, assumed or modeled melt factors remains difficult to quantify.

No studies appear to have examined the effect of the provenance of air temperature data on the accuracy of transferred sub-debris melt factors, which has been shown to affect the transferability of k for bare snow and ice (Lang and Braun, 1990; Hock and others, 2007; Wheler and others, 2014). Consequently, it is unclear whether sub-debris melt factors calculated using air temperature data measured at an on-glacier weather station, for example, can be used to accurately predict melt using input data interpolated from a remote weather station or gridded climate data.

Gridded climate reanalysis products are increasingly used in hydrological and glaciological modeling because they provide broader spatial and temporal coverage and continuity than

many traditional data sources (e.g. Koppes and others, 2015; Hagg and others, 2018). However, if data provenance is critical to the accuracy of TIM, it would be inappropriate to use climate reanalysis data with melt factors derived from ground-based temperature observations; instead, k derived from climate reanalysis air temperature would be required.

A popular reanalysis product is The Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) from the US National Aeronautics and Space Administration (NASA). It provides surface climate fields from 1980 to present, lagging a few weeks behind real time. It has six-hourly temporal resolution and $0.5^\circ \times 0.65^\circ$ spatial resolution. It has been applied successfully to model melt of debris-free glaciers over extended regions (Mernild and others, 2014). Application of MERRA-2 to model melt of debris-covered glaciers has not yet been tested.

The objectives of this study were to (1) quantify the accuracy of classical TIM and extended TIM for debris-covered glaciers, (2) test whether air temperature data provenance affects prediction accuracy, (3) derive a generalizable relation for sub-debris melt factors that can be transferred with a known margin of error, and (4) evaluate the usefulness of MERRA-2 data for modeling sub-debris glacier melt.

Definition and calculation of melt factors

Temperature index models relate surface ablation to air temperature as follows:

$$A = k \cdot PDD \quad (1)$$

where A is cumulative ablation (mm w.e.) over an observation period, k is the melt factor and PDD is the sum of positive degree-days ($^\circ\text{C} \Delta t$). Thorough reviews of temperature index modeling of debris-free glaciers were given by Ohmura (2001) and Hock (2003). Time steps (δt) vary between studies (e.g. hours, days or months), which makes it difficult to directly compare values of k . In practice, $\delta t = 1$ day is most common, so the units of k are

Table 3. Data and variables

Variable	Symbol	Units	Resolution	Source
Response variables				
Melt rate	a	mm w.e. d^{-1}	Point	Reviewed publications
Melt factor	k_o	mm w.e. $^{\circ}\text{C}^{-1} \text{d}^{-1}$	Point	Reviewed publications
	k_M	mm w.e. $^{\circ}\text{C}^{-1} \text{d}^{-1}$	Point	Calculated from a and D
Candidate fixed-effects predictor variables				
Observed debris thickness	h	m	Point	Reviewed publications
Observed air temperature	T_o	$^{\circ}\text{C}$	Point	Reviewed publications
2 m air temperature	T_M	$^{\circ}\text{C}$	$0.625^{\circ} \times 0.5^{\circ}$	inst1_2d_asm_Nx, MERRA-2
Daily mean positive degree-days	D	$^{\circ}\text{C}$	$0.625^{\circ} \times 0.5^{\circ}$	Calculated from T_M
Surface temperature	T_s	$^{\circ}\text{C}$	$0.625^{\circ} \times 0.5^{\circ}$	tavg1_2d_rad_Nx, MERRA-2
Surface temperature positive degree-days	D_s	$^{\circ}\text{C}$	$0.625^{\circ} \times 0.5^{\circ}$	Calculated from T_s
Incoming shortwave radiation	SW_{in}	W m^{-2}	$0.625^{\circ} \times 0.5^{\circ}$	tavg1_2d_rad_Nx, MERRA-2
Net shortwave radiation	SW_{net}	W m^{-2}	$0.625^{\circ} \times 0.5^{\circ}$	tavg1_2d_rad_Nx, MERRA-2
Incoming longwave radiation	LW_a	W m^{-2}	$0.625^{\circ} \times 0.5^{\circ}$	tavg1_2d_rad_Nx, MERRA-2
Outgoing longwave radiation	LW_e	W m^{-2}	$0.625^{\circ} \times 0.5^{\circ}$	tavg1_2d_rad_Nx, MERRA-2
Net longwave radiation	LW_{net}	W m^{-2}	$0.625^{\circ} \times 0.5^{\circ}$	tavg1_2d_rad_Nx, MERRA-2
Albedo	α_M		$0.625^{\circ} \times 0.5^{\circ}$	tavg1_2d_rad_Nx, MERRA-2
	α_{ML}		$1/60^{\circ}$	Terra/MODIS
	α_{ML}		$30 \text{ m} \times 30 \text{ m}$	Landsat 5 TM
Continentality	C	$^{\circ}\text{C}$	1 km^2	WorldClim V2
Bedrock geology	L		Various	Reviewed publications and geological maps
Region	R			RGI 6.0

Continentality, C , is computed as an annual mean value. All other numeric predictor variables are daily means.

mm w.e. $^{\circ}\text{C}^{-1} \text{d}^{-1}$. For simplicity and to permit direct comparison, only data and models for $\delta t = 1$ day are included in this analysis. The following relation is used to calculate PDD :

$$PDD = \delta d \sum_{i=1}^{n_d} (T_i - T_b) \cdot \delta(T_i) \quad (2)$$

with $\delta(T_i) = \begin{cases} 1, & \text{if } T_i \geq T_b \\ 0, & \text{if } T_i < T_b \end{cases}$

where T is mean air temperature for each day, subscript i is one day in the observation period of n_d days, T_b is a threshold air temperature below which ablation ceases and $\delta(T)$ is a binary variable that sets ablation to zero below the temperature threshold. T_b is often set to 0°C , but non-zero values have been found in some studies (Pellicciotti and others, 2012). In this work, ablation was standardized as a mean rate, a (mm w.e. d^{-1}):

$$a = \frac{A}{n_d} \quad (3)$$

and mean positive degree-days, D ($^{\circ}\text{C}$), were used to enable comparison, calculated as

$$D = \frac{PDD}{n_d} \quad (4)$$

In this paper, the term ‘classical’ is used to distinguish melt factors calculated with only air temperature and debris thickness as the predictors (k), as in Table 1. Extended models with one or more predictors in addition to air temperature, such as incident shortwave radiation or surface slope, have been developed with a range of different formulations (e.g. Hock, 1999; Pellicciotti and others, 2005; Carezzo and others, 2016; Möller and others, 2016). When air temperature is one of a larger set of predictor variables, as in Table 2, the coefficient associated with air temperature is called a ‘restricted’ melt factor (k^R) (Kustas and others, 1994). Restricted melt factors are not directly comparable to classical melt factors. Models with restricted melt factors, including predictors in addition to air temperature and debris thickness

(such as solar radiation or surface characteristics), are called ‘extended TIM’.

Data collection and preparation

Overview

Data for this study were gathered from free digital archives. Observed cumulative melt, A , melt rates, a , and melt factors, k_o , were compiled from existing publications. A literature search was conducted through the University of British Columbia library databases and Google Scholar using combinations of the keywords ‘glacier’, ‘ice’, ‘debris’, ‘rock’, ‘ablation’, ‘melt’, ‘melt factor’, ‘temperature index’, ‘empirical’ and ‘model’. Metadata in the publications were also collected, including geographic coordinates and site elevation, observation date and period, observational set-up, debris thickness, lithology and texture and observed air temperature. If the resolution of the geographic coordinates was low, Google Earth Pro was used to ground-truth and adjust them to a rough mid-point within the debris-covered area of each glacier (as it appears in the composite images in Google Earth Pro). Additional spatial data representing elements of the surface energy balance at the time and place of observation were collected to explore causes of variability in the observations. These were taken from MERRA-2 climate reanalysis and WorldClim V2, additional publications and geological maps. The data are included in the Supplementary Material. Procedures followed in data collection are detailed below.

The symbols, units and resolution of the data and variables used are listed with data sources in Table 3. Throughout this paper, data provenance is distinguished by a letter. The subscript ‘o’ stands for ‘observed’ and is used to identify variables, models and metrics associated with observational data copied directly from the literature. The letter ‘M’ is used to distinguish variables extracted from MERRA-2, or melt factors calculated using MERRA-2 air temperature data with observed melt rates. The letter ‘P’ stands for ‘pooled’, and is used for the observational and MERRA-2-derived melt factors combined as a single variable. The letter ‘d’ is used for melt factors calculated for Dokriani Glacier alone, using MERRA-2 air temperature data and observed melt rates.

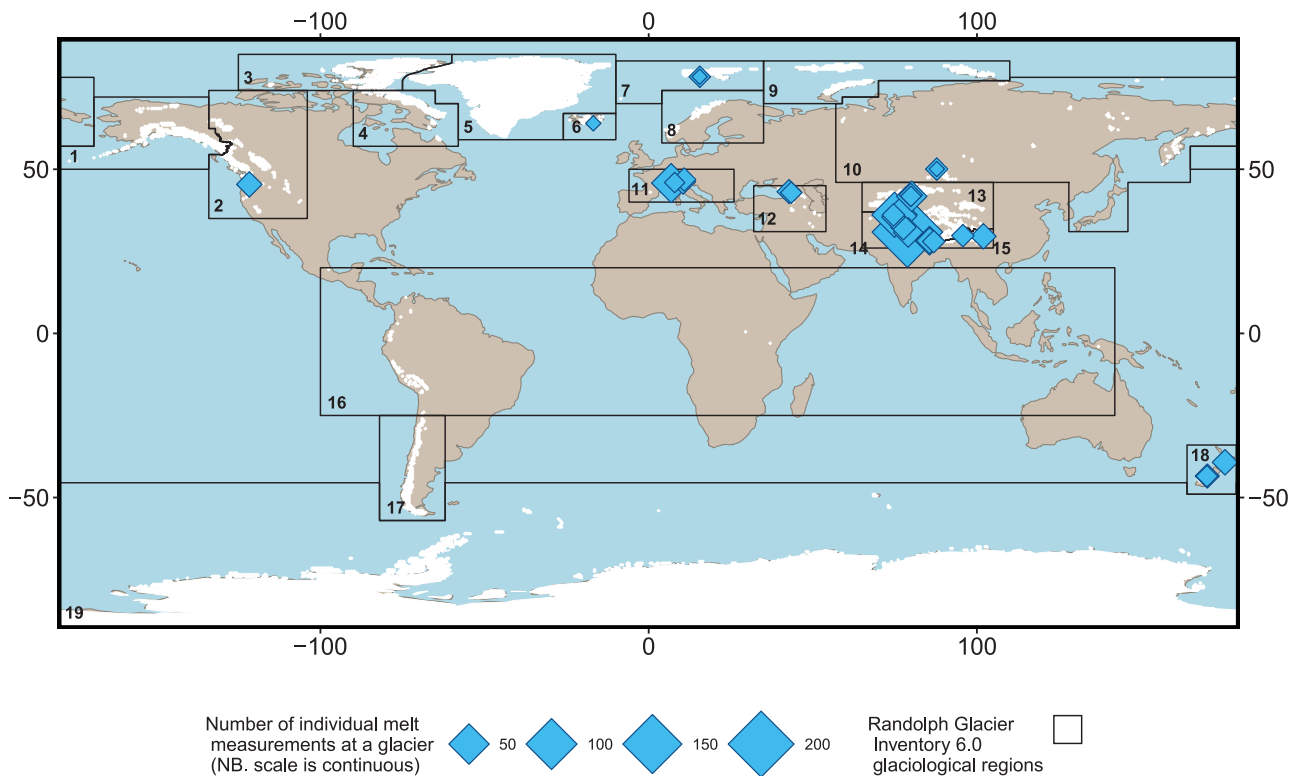


Fig. 1. The location of debris-covered glaciers in this study (blue diamonds). Symbol size indicates the number of individual observations from each glacier, on a continuous scale. The black polygons are the Randolph Glacier Inventory (RGI) 6.0 first-order glaciological regions (RGI Consortium, 2017), labeled with their numeric IDs. The white areas are the RGI 6.0 glacier outlines, Greenland and Antarctica. The background map is based on data from [naturalearthdata.com](https://www.naturalearthdata.com).

Melt rates and melt factors

The observational data are from 27 individual glaciers located in North America, Europe, Asia, Svalbard, Iceland and New Zealand, with observation points ranging from 140 to 5350 m a.s.l. and -43° to 78° latitude. The data are individual measurements of melt and were copied directly from text and tables or by digitizing plots. In the instances where corresponding values of a and k_o were published together, both were used, but separately. New melt factors, k_M , were computed using Eqn (1) with observed melt rates and mean positive degree-days calculated with MERRA-2 air temperature data, D . Four complete years of data from Dokriani Glacier, Central Himalaya, which were published in Pratap and others (2015), were generously contributed as a data file by those authors. As a subset of k_M , melt factors calculated for Dokriani Glacier are denoted k_d .

A total of seven observations from the period before 1980 were found in the literature: one from a 1956 study of Isfallsglaciären published in Østrem (1959), and six from a 1979 study of Peyto Glacier published in Nakawo and Young (1981). The MERRA-2 climate reanalysis project covers the period from 1980 to present. To avoid introducing uncertainty by using more than one source of climate reanalysis data, this study was limited to the period from 1980 to present and the observations from Isfallsglaciären and Peyto Glacier were excluded.

'Thick' debris is thought to cover more than twice the area of 'thin' debris globally (Sasaki and others, 2016). To isolate the 'insulation effect' of 'thick' debris, only data for debris thickness exceeding the critical thickness ($h > h_c$) were included, where $h_c = 0.05$ m for all glaciers except Summit Crater Glacier (Richardson and Brook, 2010), for which $h_c = 0.07$ m was used. To allow direct comparison among sites, only 'classical' values of observed melt factors (k_o) and studies focused on melting ice (rather than snow or firn) were included.

Following personal communication with the authors, the values of k_o reported in Richardson and Brook (2010) were corrected by applying a factor of 10^{-1} , and the units of k_o stated in Table 2 of Koppes and others (2015) were confirmed to be mm w.e. $^{\circ}\text{C}^{-1} \text{d}^{-1}$, not $\text{m } ^{\circ}\text{C}^{-1} \text{a}^{-1}$ as stated in the text of page 7. Of multiple units for k_o reported in Hagg and others (2008), mm w.e. $\text{K}^{-1} \text{d}^{-1}$ was assumed to be correct because $\text{cm } \text{K}^{-1} \text{d}^{-1}$ would be unusually large.

Data processing was conducted using the R programming language (R Core Team, 2017) in R Studio. To allow comparison with observed air temperature reported as daily mean values, and of observation periods ranging from days to years, melt sums were converted to melt rates using Eqn (3). In all cases where melt rates were reported in units of water equivalent, the value for ice density used in unit conversion was stated in the corresponding publications to be 890 kg m^{-3} . Therefore, both a and k_o that were not reported in water equivalent units were standardized to mm w.e. with an assumed ice density of 890 kg m^{-3} .

Two data subsets from Dokriani Glacier ($n = 24$) (Pratap and others, 2015) and one from Pichillancahue–Turbio Glacier ($n = 3$) (Brock and others, 2007) were omitted because they exhibited a positive correlation between melt rate and debris thickness, indicating the insulation signal was confounded by other processes in those instances. The final dataset included 561 individual observations, consisting of 86 melt factors (from 12 different glaciers) and 482 melt rates (from 21 different glaciers, 235 from Dokriani Glacier). The glacier locations are shown in Figure 1. Observed melt factors (k_o) ranged from 0.30 to $18.40 \text{ mm w.e. } ^{\circ}\text{C}^{-1} \text{d}^{-1}$, and debris thickness ranged from 0.05 to 0.65 m. Melt factors computed with MERRA-2 positive degree-days (k_M) ranged from 0.06 to $16.89 \text{ mm w.e. } ^{\circ}\text{C}^{-1} \text{d}^{-1}$ and h from 0.05 to 1.20 m. The number of data points from each glacier, summary statistics

for k_o and data references are provided in Tables 6 and 7, Appendix A.

Air temperature, shortwave and longwave radiation, surface temperature

On-site air temperatures were provided in 17 of the source publications, either as a multi-day mean or as a time series. Daily mean air temperature for each melt observation period, T_o , was assumed to be equal to the multi-day mean in the absence of a time series.

The NASA Earthdata subsetter tool (<https://earthdata.nasa.gov>) was used to extract meteorological data from MERRA-2 products for grid cells that contained the ground-truthed coordinates for each glacier. Four MERRA-2 data collections were used. The variable for 2 m air temperature (T_2) was extracted from the collection **inst1_2d_asm_Nx** (2-dimensional, 1-Hourly, Instantaneous, Single-Level, Assimilation, Single-Level Diagnostics, V5.12.4; Global Modelling and Assimilation Office, (GMAO, 2015b)). Surface skin temperature (T_s), incoming and net shortwave radiation (SW_{in} , SW_{net}), and incoming, outgoing and net longwave radiation (LW_o , LW_e , LW_{net}) were taken from the collection **avg1_2d_rad_Nx** (2-dimensional, 1-Hourly, Time-Averaged, Single-Level, Assimilation, Radiation Diagnostics, V5.12.4) (GMAO, 2015c). Grid cell elevation (E_M) was extracted from the constants file, **const_2d_asm_Nx** (2d, Constants, V5.12.4) (GMAO, 2015a).

Elevation lapse rates (Γ) were used to adjust T_2 to the elevation of observations (E_o , in m a.s.l.) that was reported with k_o and/or a , as follows:

$$T_{M,j} = T_2 + \Gamma_j(E_{o,j} - E_{M,j}) \quad (5)$$

for each observation site j . Two approaches were used to arrive at values for Γ_j . First, a constant value of $6.5^\circ\text{C km}^{-1}$ was assumed and applied uniformly for all j . Second, free-atmosphere lapse rates were calculated for each grid cell from the **avg3_3d_asm_Nv** collection (3-dimensional, 3-Hourly, Time-Averaged, Model-Level, Assimilation, Assimilated Meteorological Fields, V5.12.4) model-levels 72 and 71 (respectively, surface-adjacent and second from surface) (GMAO, 2015d) and applied site by site. The elevation-adjusted MERRA-2 air temperature variables were compared to T_o to evaluate which best represented on-glacier conditions. T_M with $\Gamma_C = 6.5^\circ\text{C km}^{-1}$ was found to be in closest agreement with T_o and was therefore used for the remainder of the study (see Appendix B 'Comparison of MERRA-2 and on-site air temperature data' for details). Positive degree-days were calculated from T_M following Eqn (2) and converted to daily mean positive degree-days, D , using Eqn (4), to match the period of a .

Albedo

Albedo was extracted from the **avg1_2d_rad_Nx** collection and averaged over each observation period. Given the spatial mismatch of observations and the MERRA-2 grid, and given the availability of albedo products from satellite imagery, higher resolution albedo from MODIS and Landsat 5 TM (α_{ML}) was also obtained, using Google Earth Engine. To avoid bias from patches of debris-free ice, ice cliffs and supraglacial ponds, α_{ML} was collected from nearby moraine rather than the glacier surface, and the moraine was identified manually. Cloud cover prevented the record of albedo during some of the observation periods. In those cases, values reported for the day nearest the date of observation were used. Albedo taken both from MERRA-2 and from MODIS and Landsat was found to vary with time, presumably

due to the change of the incidence angle of solar radiation. Correcting such effects was beyond the scope of this study.

Debris geology

Geological information about the debris was provided in few of the source publications. When not stated in the source publication, debris rock type was found in a related publication or digital geological map. The debris was assumed to have the same lithology as the dominant basin bedrock in the absence of more specific information. The distribution of k in bins of debris geology at the order of, for example, limestone, granite and pumice, was found to be significantly different, but those results are not included because the classes were not well resolved and the analysis was inconclusive (in some cases the highest resolution geological data were 1:1M). Instead, debris rock type was redefined as *metamorphic*, *igneous* or *sedimentary* and included in the analysis as the categorical variable L .

Continentalinity

Continentalinity (C , in $^\circ\text{C}$) was calculated as follows

$$C = T_{\max} - T_{\min} \quad (6)$$

where T_{\max} and T_{\min} are annual maximum and minimum monthly mean air temperatures ($^\circ\text{C}$), respectively, extracted from WorldClim V2. WorldClim V2 is a 1 km^2 resolution interpolated climate data product that is derived from ground-based and satellite observations (Fick and Hijmans, 2017), which is widely cited for a variety of applications and is freely available from worldclim.org. WorldClim is a purely observational data product with structure and resolution that is appropriate for climate-scale analyses with minimal data storage and processing. The relatively high temporal and spatial resolutions of MERRA-2 reanalysis products make it appropriate for synoptic scale analyses, but require data storage and processing that were impractical for calculating climatic continentalinity in the current application.

Statistical analysis

Model fitting

Model fitting and analysis were performed using the R programming language (R Core Team, 2017) within R Studio. R packages used in addition to base R are listed at the end of this document and the R code is included in the Supplementary Material. The statistical models presented here are specified in R syntax rather than matrix notation, a practice that is encouraged for clarity and reproducibility (e.g. Gandrud, 2015; Mair, 2016). Readers are directed to Bates and others (2015) for mathematical expression of the R syntax.

Models were fitted separately for the two different types of melt factor, k_o (observed) and k_M (calculated). The melt factors calculated for Dokriani Glacier, k_d , were included in the analysis of k_M and also analyzed separately. This resulted in three sets of models, one for each of k_o , k_M and k_d . Hereafter, these model sets are called KO, KM and KM_d , and each individual model within a set is designated by a number (i.e. KO1, KM1, KM_d1). An additional model set, KP, was generated by pooling k_o and k_M . The same data were not repeated in the pooled dataset. Melt factors in the KO dataset were prioritized over the KM dataset. If melt rates corresponding to the melt factors in the KO dataset were available, those melt rates were used to calculate values of

k_M that were included in the KM dataset but were not included in the pooled dataset.

A range of logarithmic and exponential transformations of the response and predictor variables were tested to meet the linear modeling assumption of normally distributed error terms. Models with base-10 logarithmic transformation of the response variable met that requirement and were the most accurate in cross-validation. The base-10 logarithms of the melt factors are denoted y .

An attempt was made to model ablation directly as a function of the predictor variables rather than from a predicted melt factor. However, those models were consistently less accurate than those based on a predicted melt factor and are not considered further.

Both fixed-effects and mixed-effects linear models for y were fit, the latter to account for the variability of model coefficients among glaciers and among years of observation. The motivation for using mixed-effects models in this study is that they explicitly address the hierarchical structure of the dataset when observations are grouped – e.g. by repeat sampling of the subjects in a study – and the dependency of the error terms on grouping levels (e.g. sampling sites) must be accounted for (Galecki and Burzykowski, 2013; Galwey, 2014).

Mixed-effects models include both fixed effects and random effects. Fixed effects represent the effects of predictors that are assumed to apply uniformly throughout the population of interest. When a model that only includes fixed effects is applied to data with a grouped structure (e.g. multiple observations within multiple sites), the residuals from the model often exhibit within-group correlations; for example, some sites may have dominantly positive residuals while other sites may have dominantly negative residuals. Such within-group correlation violates the assumption of independence that underlies fixed-effects models. The inclusion of site as a random effect in this case is one way to account for such within-group correlation.

The fundamental assumption underlying a mixed-effects model is that the model coefficients can vary randomly from group to group, and are drawn from a normal distribution. When predictions are made for a new group that was not in the dataset used to fit the model, the prediction is made using the fixed-effects coefficients, which are assumed to apply through the entire population. The calculated prediction limits include the uncertainty associated with the among-site variability of the random effects. Multiple random effects and fixed effects can be included in a single mixed-effects model.

Data with a grouped structure are common to many areas of earth science (e.g. Booker and Dunbar, 2008; Kasurak and others, 2011; Azócar and others, 2017), making mixed-effects modeling especially useful. To illustrate the application of mixed-effects models, consider the case of a linear relation between a response variable y and a single predictor variable x , with multiple sites and multiple observations at each site. In this case, the model could be represented as

$$y_{ik} = B_0 + b_{0k} + (B_1 + b_{1k}) \cdot x_{ik} + \epsilon_{ik} \quad (7)$$

where y_{ik} and x_{ik} are the i th observations of y and x at site k , B_0 and B_1 are fixed effects that are assumed to apply to the entire population of sites, b_{0k} and b_{1k} are deviations from the fixed-effects coefficients for site k , and ϵ_{ik} is a random error term. If the model is applied to new observations at one of the sites used to fit the model, then both the fixed effects and the random effects for that site are included. However, if the model is applied to observations at a new site, then only the fixed-effects coefficients are used. With a nested data structure, such as the present case with multiple sites, multiple years of observation at each site and multiple observations in each year, the model

would be

$$y_{ijk} = B_0 + b_{0k} + b_{0jk} + (B_1 + b_{1k} + b_{1jk}) \cdot x_{ijk} + \epsilon_{ijk} \quad (8)$$

where j is the year of observation at a given site, and the index jk indicates a deviation from the relation at site k in year j . The random effects b_{0k} and b_{1k} are included when the model is applied to new observations from a site that was used to fit the model; the random effects b_{0jk} and b_{1jk} are also included if the new observations were taken in one of the site-years used to fit the model. Otherwise, for a new site, only the fixed effects are applied, as for the simpler case of Eqn (7).

The prediction limits computed for a mixed-effects model include not only uncertainty in the fixed-effects coefficients and the random error term, as is the case for fixed-effects models, but also the uncertainty associated with the variability of coefficients among sites as represented by the variances of the mixed-effects coefficients. An important advantage of using mixed-effects models is that the values of B_0 and B_1 are not biased by individual sites with large numbers of observations, as would be the case for a fixed-effects model. Fixed-effects versions of all the models tested in this study were significantly outperformed by the mixed-effects models and are not considered further.

The mixed-effects models were fit using the *lmer* function from the **lme4** R package using maximum likelihood estimation (Bates and others, 2015). Model fitting was carried out in two stages, starting with a base model including debris thickness. In R syntax, the base mixed-effects model was defined as:

$$\hat{y} \sim h + (1|G) \quad (9)$$

where \hat{y} is the base-10 logarithm of the predicted melt factor (\hat{k}_o , \hat{k}_M or \hat{k}_d), h is debris thickness and G is the name of each glacier as a categorical variable. Equation (9) contains two fixed-effects coefficients, B_0 and B_1 , which are the intercept and a slope coefficient for h , and the $(1|G)$ term indicates a set of random-effects coefficients representing a random deviation from the fixed-effects intercept for each glacier (b_{0j}). Base models were also tested that allowed the slope coefficient for debris thickness to vary randomly among glaciers in addition to the intercept (b_{1j}):

$$\hat{y} \sim h + (h|G) \quad (10)$$

Because multiple years of data were available for a number of glaciers, the year of observation, yr , was included as an additional random effect, nested within each glacier. The following model, for example, allows both the intercept and the slope coefficient for debris thickness to vary randomly among glaciers and years (b_{0jk} and b_{1jk}):

$$\hat{y} \sim h + (h|G) + (h|G:yr) \quad (11)$$

The candidate mixed-effects base models in Eqns (9)–(11) were evaluated with the Akaike Information Criterion (AIC). The AIC is a metric that weighs model performance against parsimony, penalizing unnecessary complexity. The base model with the lowest AIC was advanced to the second stage of model development, where additional fixed-effects predictors and interaction terms were tested. Variables tested were T_s , D_s , SW_{net} , SW_{in} , LW_{net} , LW_a , LW_e as well as rock type (L), continentality (C), albedo (α_M and α_{ML}) and RGI 6.0 glaciological region (R). To minimize issues with collinearity, pairs of predictors with a Pearson's correlation coefficient > 0.5 were not

included in a model together. Otherwise, combinations of the fixed-effects variables and interaction terms, such as $h \cdot SW_{net}$, $h + SW_{net} + C$ and $h + SW_{net} \cdot C$, and so on, were tested exhaustively. Calculated melt factors for Dokriani Glacier varied through the ablation seasons of 2010–2014. In the models for Dokriani Glacier (model set KM_d), *month* was included as a fixed effect, and *yr* as a random effect.

The fitted coefficients and 95% prediction limits calculated by the *lmer* function are the average of the maximum likelihood estimates from a user-defined number of simulations. Here, the number of simulations per iteration was set to 10,000. Models were discarded if any individual predictor was not significant at the 0.95 confidence level or if any fitted coefficient was physically unrealistic. Valid models for each response variable were ranked by AIC. Up to five of the top-ranked models were selected for cross-validation. Only the fixed-effects coefficients were used in model validation, when fitted models were used to make predictions for glaciers not included in the model fitting.

Model validation

Leave-one-out cross-validation was used to evaluate the candidate models’ ability to reproduce observed melt rates. At each iteration, the input values for one glacier were withheld and the model fit using data from the other glaciers. The newly fitted model was then applied to predict the values for the withheld glacier. For model set KM_d , 150 m elevation bins were used in place of glaciers.

Smearing estimates (*s*) were calculated to avoid the introduction of bias during back-transformation of the predicted values from base-10 logarithms (i.e. $\hat{k} = 10^{(\hat{y}+s)}$) (Duan, 1983). Predicted melt rates, \hat{a} , were calculated using Eqn (1) and the back-transformed melt factors (i.e. $\hat{a} = \hat{k} \cdot D$). When discussing predicted values of \hat{y} and \hat{a} , subscripts are used to indicate which model was used (e.g. \hat{y}_{KO3} , \hat{a}_{KM1}).

Metrics used to evaluate the performance of the models in cross-validation were the root-mean-square error (RMSE), root-mean-square relative error (RMSRE), mean bias error (MBE) and relative mean bias error (RMBE), calculated as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (a_i - \hat{a}_i)^2} \tag{12}$$

$$RMSRE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{a_i - \hat{a}_i}{a_i} \right)^2} \tag{13}$$

$$MBE = \frac{1}{n} \sum_{i=1}^n (a_i - \hat{a}_i) \tag{14}$$

$$RMBE = \frac{1}{n} \sum_{i=1}^n \left(\frac{a_i - \hat{a}_i}{a_i} \right) \tag{15}$$

where a_i and \hat{a}_i are the observed and predicted values, respectively, for each observation i , and n is the total number of observations. The percentage of points within $\pm 25\%$ error and the graphical linearity of the predictions also served as indicators of the models’ goodness of fit. Models that performed relatively poorly in cross-validation were discarded and a maximum of three top-performing models for each model set are presented in the Results. A range of prediction uncertainty was estimated for the strongest model in each set, where the 95% prediction limits

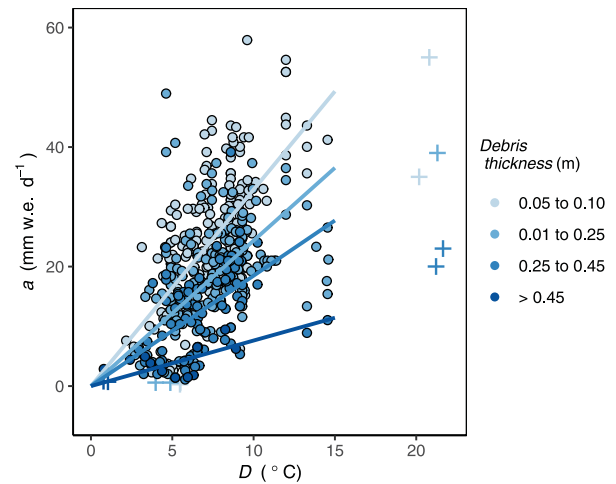


Fig. 2. Observed melt rate (*a*) and mean daily positive degree-days based on MERRA-2 air temperatures (*D*), with least-squares best-fit lines through the origin for bins of debris thickness. Crosses are extreme values that were removed for model fitting.

were converted to a percentage of the predicted values and averaged over the observations.

Air temperature data provenance

To test whether melt factors calculated with air temperature data of different provenances were significantly different, k_o and k_M were combined as a single response variable, k_P . Air temperature data provenance was included in the base model as a fixed-effects categorical predictor (i.e. observed or MERRA-2). Significance of the provenance variable was evaluated at the 0.90 confidence level. This liberal threshold was chosen to reduce the chance of incorrectly accepting the null hypothesis, that k_o and k_M came from the same distribution.

The importance of using air temperature data of the same provenance for both calculation and application of melt factors was also evaluated. Model set KO was tested twice, once using observed air temperature as input and again using MERRA-2 air temperature as input. If the provenance of air temperature data (i.e. MERRA-2 vs observed on-site) had a significant affect on predictive accuracy, then \hat{a} calculated with \hat{k}_o and observed positive degree-days should be more accurate than \hat{a} calculated with \hat{k}_o and MERRA-2 positive degree-days.

A complication in the second analysis is that values of k_o were calculated with observed positive degree-days, which were not reported in the source publications and were thus unavailable for this test. Therefore, the mean observed temperature, T_o , was used rather than positive degree-days, whereas the MERRA-2-based temperature statistic, D , was computed as positive degree-days. It is possible that using different statistics for the calculation and application of melt factors (i.e. positive degree-days vs daily means) could affect prediction accuracy. To evaluate that possibility, cross-validation of all the models was performed with D first, and then using daily mean MERRA-2 air temperature, T_M . Differences in the RMSE and MBE of the same model but different air temperature input statistic were consistently $< 0.001 \text{ mm w.e. d}^{-1}$. Therefore, applying T_o and D in the KO models was deemed to be a valid indicative test of the effect of input data provenance on prediction accuracy. Values of \hat{a}_{KO} predicted using T_o and D are distinguished with T or D appended to the model name (i.e. \hat{a}_{KO2-T} , \hat{a}_{KO2-D}).

Once the significance of air temperature data provenance had been assessed, the full model fitting and validation procedures were carried out with the pooled melt factors, k_P , as the response variable. That resulted in the additional set of models designated KP.

Table 4. Fitted coefficients with standard errors and AIC for models of \hat{y}_{KO} , \hat{y}_{KM} , \hat{y}_{KP} and \hat{y}_{KM_d} .

Model	R formula	B_0	B_1	B_2	AIC	n
KO1.	$h + (1 G)$	0.62 ± 0.08	-1.43 ± 0.11		-50.2	81
KO2.	$h + SW_{net} + (1 G)$	-0.68 ± 0.32	-1.44 ± 0.11	0.006 ± 0.001	-47.8	
KM1.	$h + (h G) + (h G : yr)$	0.58 ± 0.04	-1.23 ± 0.17		-189.1	476
KM2.	$h + \alpha_M + (h G) + (h G : yr)$	1.05 ± 0.10	-1.24 ± 0.17	-2.16 ± 0.42	-209.6	
KM3.	$h + LW_{net} + (h G) + (h G : yr)$	0.75 ± 0.05	-1.24 ± 0.17	$0.002 \pm 4 \times 10^{-4}$	-204.0	
KP1.	$h + (h G) + (h G : yr)$	0.62 ± 0.04	-1.46 ± 0.17		-225.0	550
KP2.	$h + LW_{net} + (h G) + (h G : yr)$	0.80 ± 0.05	-1.47 ± 0.17	$0.002 \pm 3 \times 10^{-4}$	-239.7	
KM _d 1.	$h + (h yr)$	0.44 ± 0.04	-0.50 ± 0.18		92.8	231
KM_d2.	$h + month + (h yr)$	0.52 ± 0.05	-0.52 ± 0.16	May: -0.367 ± 0.05 June: -0.127 ± 0.05 July: 0.010 ± 0.05 September: 0.067 ± 0.05	43.5	

G is a class variable representing individual glaciers. B_0 to B_2 are fitted fixed-effects coefficients. Also shown is the number of data points in each model set, n . Random-effects coefficients are not shown. The model in each set that was most accurate in cross-validation is highlighted with bold font

Results

Models were fitted for the KO, KM, KP and KM_d datasets, with a total of 557 individual observations from 27 glaciers around the world. From the original 568 observations, 11 values of a from Hinarche Glacier, Dokriani Glacier and Koxkar Glacier were unusually high or low given D and were removed as outliers prior to model fitting (Fig. 2, crosses). None of the predictor variables explained those values and the results presented below are for the reduced dataset ($n = 557$).

The model coefficients presented below are in log base-10 transformed units. For example, the coefficients for the intercept, B_0 , are in units of \log_{10} (mm w.e. °C⁻¹ d⁻¹), while the coefficients for debris thickness, B_1 , are in units of \log_{10} (mm w.e. °C⁻¹ d⁻¹)m⁻¹.

Dependence of ablation on degree-days and debris thickness: overview

Figure 2 illustrates the dependence of measured melt rate on positive degree-days and debris thickness. As expected, both the magnitude of melt and its sensitivity to air temperature decline as debris thickness increases. Also notable is the high degree of variability around the best-fit linear regression lines for each thickness bin.

Model set KO: melt factors for multiple glaciers extracted directly from the literature

In these models, the slope coefficient, B_1 , is the direct dependence of melt factors on debris thickness, while the intercept, B_0 , is the variation of melt factors that is independent of debris thickness. As seen in Table 4, the strongest models for predicted values of the log base-10 of k_o are the base model with random effects in the intercept of h by glacier (G), KO1, and the extended model including SW_{net} , KO2. Despite the inclusion of SW_{net} in KO2, which results in different estimates of the intercept, the models have consistent estimates of the coefficient for debris thickness, B_1 .

Although the AIC of model KO1 is 2.4 points lower than that for KO2, model KO2 was more accurate in cross-validation (Table 5). The models have similar biases, but the RMSE and RMSRE of \hat{a}_{KO2} are almost half the magnitude of those of \hat{a}_{KO1} . The percentage of points within $\pm 25\%$ error is not a strong metric of the performance of the KO models, because the influence of individual points in the small validation dataset was exaggerated. Nonetheless, those values are also favorable to model KO2, irrespective of which air temperature input was used. Differences

due to the air temperature input variable are discussed with the results for air temperature data provenance, below.

There is a negative bias in \hat{a}_{KO2} that increases as a increases (Fig. 3a), as a result of the uneven distribution of observations in the lower range of h (Fig. 3b). The 95% prediction limits for \hat{k}_{KO2} (calculated with SW_{net} set to its mean value, 225 W m²) range from -49 to 203% of the predicted values on average.

Model set KM: melt factors for multiple glaciers calculated with observed melt rates and MERRA-2 air temperature

The best KM models are the base model with random effects in the slope and intercept by glacier and by year (KM1), the model extended with α_M (KM2), and the model extended with LW_{net} (KM3). The estimated coefficients for h are consistent between models KM1, KM2 and KM3 and consistent with those of the KO models, within the standard errors of the estimates.

Model KM2 has the lowest AIC but performed poorly by all criteria in cross-validation, with RMSE and MBE an order of magnitude larger than those of models KM1 and KM3, as seen in Table 5. Models KM3 and KM1 performed similarly well to each other, with differences in the errors on the order of 0.1 mm w.e. d⁻¹ and one percentage point; on the basis of fractionally more favorable metrics and parsimony, the base model is deemed to be more robust.

Like \hat{a}_{KO2} , around half the values of \hat{a}_{KM1} fell within $\pm 25\%$ error, but the MBE of KM1 is an order of magnitude smaller than that of KO2. The RMSE is also smaller for \hat{a}_{KM1} than \hat{a}_{KO2} , but the RMSRE is around three times as large. The RMSRE was strongly influenced by the lack of observations of low melt rates in the k_o dataset with which to test KO2, as seen in Figures 3a and c. It can be seen in Figure 3d that the 95% prediction limits for \hat{k}_{KM1} are also wider than those for \hat{k}_{KO2} , ranging from approximately -38 to 270% of the predicted value on average. Again, that resulted from the larger number of low melt factors in the k_M dataset that were not present in the k_o dataset.

Of two observations of melt rate found in published literature for $h > 0.65$ m, one was unusually large given D and was removed as an outlier before model fitting. The value of k_M calculated for the sole remaining observation at $h = 1.2$ m is poorly reproduced by model KM1 and lies well above the prediction limits (Fig. 3d).

Model set KP: k_o and k_M pooled as a single response variable

Air temperature data provenance was not found to be significant in a model of k_o and k_M pooled as a single response variable, k_P

Table 5. Cross-validation metrics for predicted melt rates, \hat{a}

Model	Input variable (°C)	RMSE (mm w.e. d ⁻¹)	RMSRE	MBE (mm w.e. d ⁻¹)	RMBE	% of points within ± 25 % error	Number of glaciers in the model set
KO1.	T_o	14.8	0.52	-5.7	0.17	23	12
KO1.	D	15.1	0.59	-1.7	-0.01	11	
KO2.	T_o	7.9	0.24	-5.4	0.19	46	
KO2.	D	8.3	0.35	-2.5	0.02	56	
KM1.	D	7.7	0.99	-0.5	-0.23	51	20
KM2.	D	11.7	1.22	-1.6	-0.22	34	
KM3.	D	7.8	0.95	-0.6	-0.22	50	
KP1.	D	7.7	0.98	-0.2	-0.23	52	27
KP2.	D	7.9	0.98	-0.1	-0.23	50	
KM _d 1.	D	7.0	3.84	-1.0	-0.65	54	1
KM_d2.	D	7.0	3.50	0.2	-0.50	53	

The most accurate model in each set is highlighted with bold font.

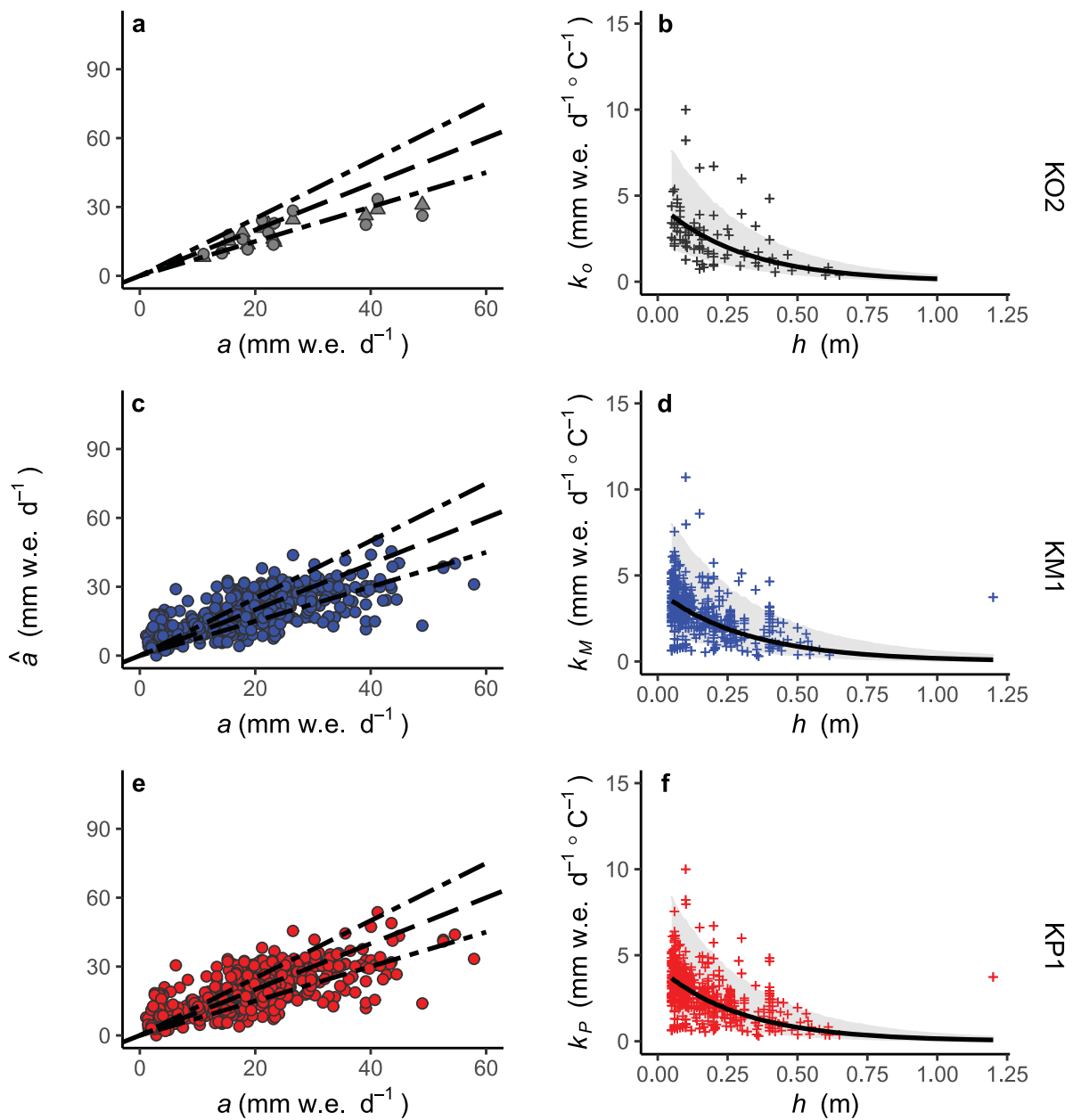


Fig. 3. Left column: observed melt rates, a , and modeled melt rates, \hat{a} , using the best models for k_o , k_M and k_P (rows 1–3, respectively). MERRA-2 positive degree-days were used as input to models KM1 and KP1. In plot a, \hat{a}_{KO2-T} and \hat{a}_{KO2-D} are represented with triangles and dots, respectively. The solid lines represent perfect agreement and the dashed lines are ± 25% error. Right column: observed melt factors, k (crosses), modeled melt factors, \hat{k} (curves) and 95% prediction limits (shaded areas), against debris thickness, h .

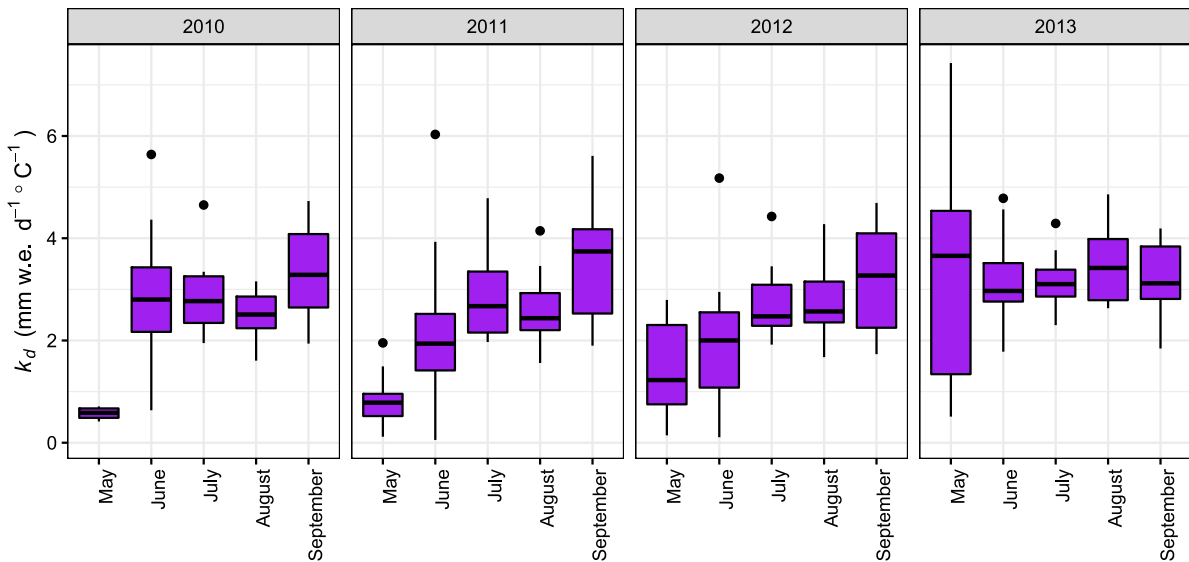


Fig. 4. The variation of k_d among ablation stakes for each month and year of the 2010–2013 ablation seasons.

($p = 0.23$). The cross-validation metrics in Table 5 vary between models KO given different air temperature input variables, T_o and D . The RMSE and RMSRE are smaller using T_o in both KO1 and KO2, but the MBE is around three times as large. The difference in the percentage of points within $\pm 25\%$ error varies inconsistently between inputs. The differences between \hat{a}_{KO2-T} and \hat{a}_{KO2-D} in Figure 3a are not striking.

In the absence of evidence that inconsistent air temperature data provenance reduced prediction accuracy, the model fitting and validation procedures were carried out with k_p as the response variable. The strongest models set in KP are the base model with random effects in the slope and intercept of debris thickness by glacier and by year (KP1), and the model extended with LW_{net} (KP2) (Table 4). As with model set KM, the AIC for KP2 is lower than that for the base model but the base model was more accurate in cross-validation. The RMSE of KP1 is fractionally smaller than that of KP2, and KP1 has 3% more points within $\pm 25\%$ error than KP2 (Table 5). The RMSRE and MBE for the two models are equal. On the basis of those metrics and parsimony, KP1 is deemed to be more robust.

Only around 50% of \hat{a}_{KP1} fall within $\pm 25\%$ error, comparable to models KO2 and KM1. The RMSRE of KP1 falls between KO2 and KM1, the RMSE is lowest for KP1 and equal with KM1, while the MBE is lowest for KP1 of all the models. The uncertainty in \hat{k}_p , ranging from -40 to 254% of the predicted value, is intermediate between KO2 and KM1.

Model set KM_d : melt factors calculated for Dokriani Glacier with observed melt rates and MERRA-2 air temperature

Melt factors for the 2010–2013 ablation seasons at Dokriani Glacier, k_d , were calculated using D and monthly melt measured at a network of 30 ablation stakes. In Figure 4 it can be seen that k_d increased over the ablation season in all years except 2013. Figure 4 also shows that inter-site variability of k_d within each month did not vary markedly from year to year. However, for May, the inter-site variability was low in 2010 and increased each year to the point that there was greater inter-site variability in May 2013 than in any other year-month. The difference between k_d for June and September was found to be statistically significant in all years except 2013, while the difference between k_d in July and August was not found to be significant in any year.

The best models for k_d are the base model with random effects in the slope and intercept by year, KM_{d1} , and the model extended with a categorical variable for *month*, KM_{d2} (Table 4). The estimated coefficients for the slope of debris thickness, B_1 , are consistent between KM_{d1} and KM_{d2} , but notably smaller than those estimated for KO, KM and KP. The estimated coefficients for *month*, B_2 , reflect the pattern of increase through the ablation season described above.

In cross-validation, the RMSE of the base model and KM_{d2} were equal, and KM_{d1} had 1% more predicted values within $\pm 25\%$ than KM_{d2} (Table 5). However, the MBE of KM_{d2} is an order of magnitude smaller than that of KM_{d1} , and the RMSRE is also fractionally smaller. Model KM_{d2} is deemed to be more robust primarily on the basis of the MBE.

Both KM_{d2} and KM_{d1} have smaller RMSE than the strongest multi-glacier models, but RMSRE more than three times as large. That is because the distribution of k_d is skewed by the low May and June values while the multi-glacier datasets are more normally distributed. The MBE of KM_{d2} and KP1 are equal in magnitude, but the RMBE of KM_{d2} is twice as large as that of KP1. The prediction limits for \hat{k}_d , calculated as the mean for each month, are also wider than any of the multi-glacier models at -32 to 312% of the predicted value.

In Figure 5a, it can be seen that May melt rates were least well reproduced. It is also of note that, while k_d were highest in September, a were relatively low. Figure 5b shows that the overall increase of k_d between May and September, muted in July and August, is reflected in $\hat{k}_{KM_{d2}}$.

Discussion

The data compiled and analyzed here are from debris-covered glaciers around the world, covering a wide range of geographic and climatic settings. The analysis confirms the generality of the dependence of melt rates on positive degree-days and debris thickness that has been well documented at individual sites.

Accuracy of modeled melt factors applied to new glaciers without calibration

The best-performing models generated predictions that were within $\pm 25\%$ error only around 50% of the time, while the RMSE and RMSRE of the highest performing model, $8.7 \text{ mm w.e. d}^{-1}$ and

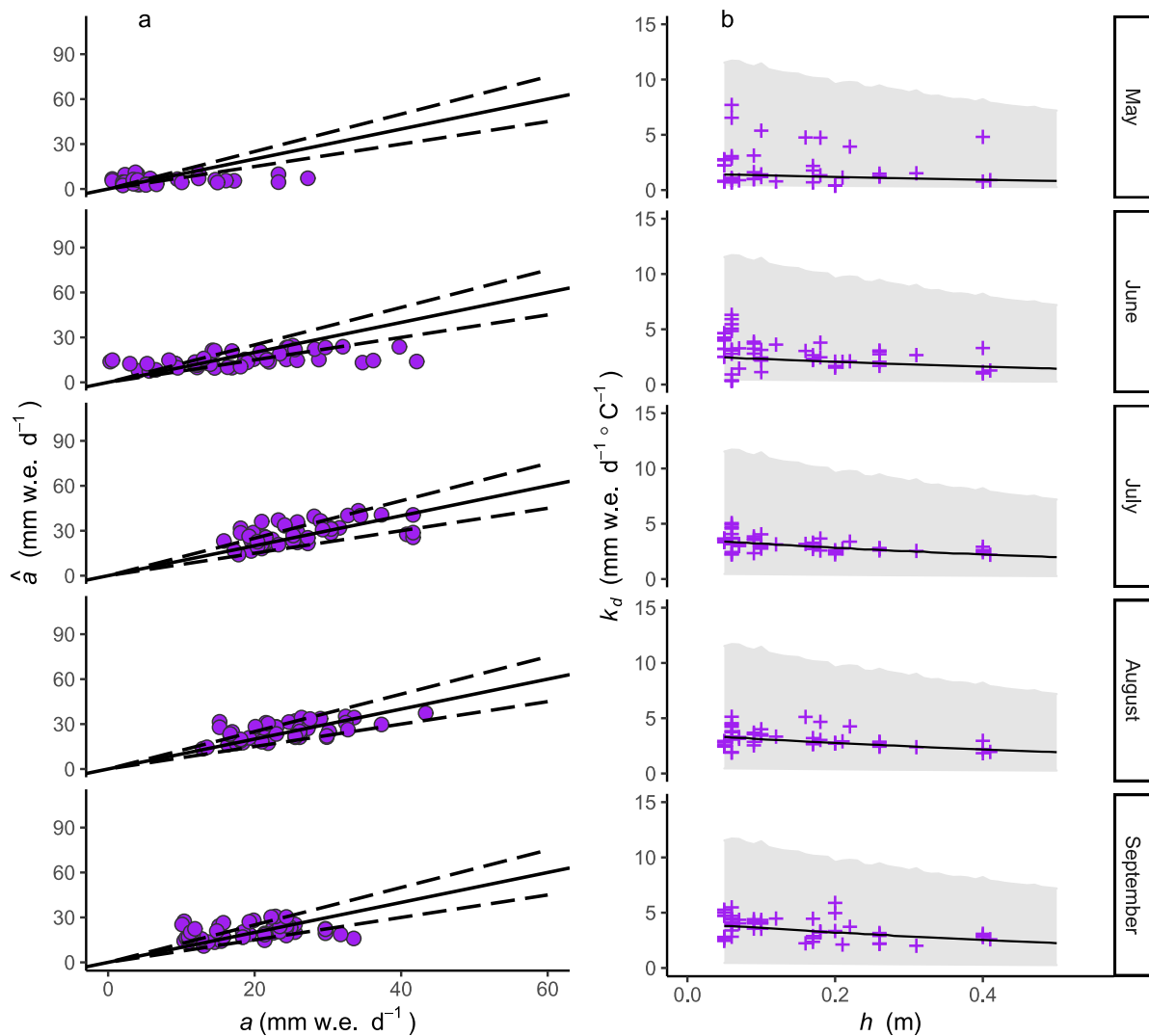


Fig. 5. Left column: observed melt rates, a , and predicted melt rates, \hat{a} , for Dokriani Glacier and each month of the ablation season, using model KM_d2 and MERRA-2 positive degree-days. The solid lines represent perfect agreement and the dashed lines are $\pm 25\%$ error. Right column: observed k_d (crosses), predicted values, \hat{k}_d (curves) and 95% prediction limits for each month (shaded areas), against debris thickness, h .

0.98 respectively, are substantial (model KP1, Table 5). This reflects the known strengths and weaknesses of empirical melt models, tending to be accurate on average but imprecise for specific times and places. The prediction limits of mixed-effects models are typically wider than those of equivalent fixed-effects models, because they explicitly consider the sample used to fit the model within the context of a broader population of glaciers to which the model might be applied (Shadish and others, 2002). It is notable, therefore, that the accuracy of model KP1 is comparable to the reported accuracy of TIM transferred between neighboring debris-free glaciers (e.g. Wheler, 2009).

Extending classical TIM with additional predictors did not improve prediction accuracy, but increasing the size of the calibration dataset did. In cross-validation, model KO2, extended with SW_{net} , out-performed the base model (KO1), but the base models KM1 and KP1 out-performed the extended models KM2, KM3 and KP2. Then, KP1 predicted k_o more accurately than either KO1 or KO2. This shows that increasing the size of the dataset let the underlying, common relation between k and h to be more accurately estimated, and that was more important than additional information about the surface energy exchanges.

The inability of other candidate predictors to contribute to the performance of the models suggests that positive degree-days and

debris thickness are strong predictors of sub-debris melt, which is in agreement with previous findings (e.g. Möller and others, 2016). It has been found that air temperature is reproduced more accurately than other variables in some climate reanalyses (e.g. Trubilowicz and others, 2016); here, MERRA-2 air temperature was found to be in good agreement with observations (root mean squared deviation = 2.56°C, mean bias < 0.1°C, $n = 32$, Appendix B). The relative performance of the predictors may also have been influenced by relative accuracy of the MERRA-2 variables.

The increased volume of data used to fit models may have improved the accuracy of the distribution of melt factors as a function of debris thickness, or, with more repeat observations within the hierarchical sampling structure of the data, improved accuracy of the random-effects coefficients. The models in sets KM and KP were improved with random effects in the intercept and slope of h by glacier (G) and by year (yr), having repeat observations at multiple glaciers over multiple years. In contrast, the models in set KO were only improved by random effects in the intercept of h by G , because, despite the same sampling structure, there were not enough repeat observations for each value of G and yr to accurately estimate coefficients for the slope of h by G , or either the intercept or slope of h by yr . With more data, the fixed- and random-effects coefficients for both k_M and k_o were more accurately approximated.

Empirical melt models are often fit to data from a single glacier and applied to neighboring glaciers. The coefficients for debris thickness estimated in the models of Dokriani Glacier were significantly different to those estimated in the multi-glacier models sets, KO, KM and KP, while the coefficients for debris thickness estimated in the multi-glacier model sets were consistent with each other. This finding suggests that empirical models fitted using data from a single glacier are location specific, and that models fitted using data from multiple glaciers more accurately reflect the direct relation between melt, air temperature and debris thickness. Models fitted using data from multiple glaciers may, therefore, be more accurate when they are transferred between glaciers than models fitted using data from a single glacier. Indeed, the accuracy of predictions made with model KP1 for glaciers that KP1 had not been calibrated for was only fractionally lower than that of predictions made for bins of elevation on Dokriani Glacier with model KM_{d2}, which was calibrated for Dokriani Glacier.

The fractionally higher accuracy of predictions made with model KM_{d2}, with RMSE = 7.0 mm w.e. d⁻¹ and MBE = 0.2 mm w.e. d⁻¹, is likely to relate to the properties of the debris at Dokriani Glacier (e.g. rock type, porosity), which were uniform by comparison with those in the multi-glacier datasets. However, the elevations of observation sites at Dokriani Glacier were also known precisely while only a single, average or nominal value was known for most of the source data. Therefore, the accuracy of KM_{d2} might simply have been due to more accurate elevation adjustment of *D*.

Mixed-effects models are based on the assumption that the observation sites and years are sampled randomly from broader populations. In the current case, the sampled population includes only eight of 18 RGI glaciological regions and elevations < 5350 m a.s.l.; therefore, the applicability of the models and prediction limits outside that population of sites is not known.

The models could not be validated for debris thicker than 0.65 m. Model KP1 predicts $k = 0.5 \text{ mm w.e. } ^\circ\text{C}^{-1} \text{ d}^{-1}$ for $h = 0.65 \text{ m}$, decaying to $0.1 \text{ mm w.e. } ^\circ\text{C}^{-1} \text{ d}^{-1}$ for h between 1.0 and 1.2 m, which is in general agreement with empirical models of sub-debris melt in the literature (e.g. Brook and Paine, 2012; Brook and others, 2013; Hagg and others, 2014, 2018; Juen and others, 2014; Möller and others, 2016). However, the single observation for $h > 0.65 \text{ m}$, at $h = 1.2 \text{ m}$ from Juen and others (2014), has a value of $k_M = 4 \text{ mm w.e. } ^\circ\text{C}^{-1} \text{ d}^{-1}$ that is considerably higher than the models predicted, suggesting the models might not be valid for $h > 0.65 \text{ m}$.

Importance of air temperature data provenance

The success of TIMs relies on positive degree-days having a consistent relation to the surface energy balance. A number of studies have shown that, because air temperature data collected from different locations and/or by different methods differ (e.g. recorded at an AWS on the glacier or on a lateral moraine, or recorded at a remote weather station and lapsed to the elevation of the glacier), the accuracy of predictions can suffer if air temperature data of different provenances are used for the calculation and application of melt factors (e.g. Lang and Braun, 1990; Shea and others, 2009; Shea and Moore, 2010; Wheler and others, 2014).

No clear evidence was found here that air temperature data provenance affected the magnitude of melt factors, given the considerable scatter inherent in the data. The slightly different distributions of k_o and k_M were not statistically significant, and differentiating between them did not improve the performance of the models. Nor did changing the provenance between model fitting and application of melt factors have a significant effect on prediction accuracy: the RMSE of the KO models was

< 1 mm w.e. d⁻¹ lower with observed air temperature input data than with MERRA-2 positive degree-days, but the MBE was more than 3 mm w.e. d⁻¹ larger. Model KP1 was derived from melt factors calculated using data with two different provenances, yet, with input data of a single provenance (*D*), it was the most accurate of the multi-glacier models in cross-validation. This indicates that enlarging the imprecise dataset by combining observations from different sources was more valuable than maintaining consistent air temperature data provenance.

It is plausible that melt factors calculated with climate reanalysis air temperature data are more highly transferable than melt factors calculated with point-scale data. Micro-climates can affect air temperature recorded at point scale, while MERRA-2 air temperature variables reflect the full, integrated system of mass and energy fluxes through the modeling and assimilation process. Melt factors calculated with air temperature data that reflects the average, regional scale energy balance may more accurately predict melt under debris at new sites than melt factors calculated with point-scale air temperature data.

Variability of melt factors within the ablation season

The melt factors calculated for Dokriani Glacier, k_d , were found to vary systematically through the 2010–2012 ablation seasons. Including *month* as a categorical variable in model KM_{d2} reduced the RMSRE of predictions by 0.32 and the MBE by an order of magnitude, from -1.1 to $0.2 \text{ mm w.e. d}^{-1}$.

In 2010–2012, the values of k_d increased significantly from June to a maximum in September, which is reflected in Figure 5b and the KM_{d2} coefficients for *month* in Table 4. Short-term observations have found that heat storage in supraglacial debris is close to zero on the order of days (Brock and others, 2007) but may be higher at seasonal scale (Nicholson and Benn, 2013). The increase of values of k_d between June and September at Dokriani Glacier suggests that heat accumulated in the debris layer after melting had begun, so that melt rates were increasingly sensitive to surface energy inputs. There was no significant difference between July and August values in any year. July is the first month of the monsoon in the Central Himalaya, characterized by thick cloud cover and intense rainfall. The rate of heat accumulation may have declined in response to reduced net radiation, or the increasing importance of heat advection by rainwater percolating through the debris.

The sensitivity of melt factors to debris thickness in model KM_{d2} was close to one-third that of the multi-glacier models, at $-0.52 \log_{10}(\text{mm w.e. } ^\circ\text{C}^{-1} \text{ d}^{-1})\text{m}^{-1}$. That also suggests that advection by rainwater percolation was an important mechanism of heat transfer in the debris layer. Continuous percolation of rain water at air temperature might reduce the thermal gradient in the near-surface, making the effect of debris thickness on melt energy supply less pronounced.

Very low values of k_d in May 2010–2012 suggest that much of the surface heat input early in the ablation season was used to warm the debris layer and thus not immediately available to melt ice. May 2013 melt factors were higher than they were in previous years and more variable among sites than in any other month or year. Air temperature in the region of Dokriani Glacier was not significantly different in 2013 by comparison with previous years, but there were precipitation deficits in March, April and May 2013 (Kaur and Purohit, 2013, 2014). Further, precipitation in June 2013 was 123–191% greater than the climatic normal (Kaur and Purohit, 2014, visualize.data.gov.in) indicating that macro-scale atmospheric circulation was unusual that year. With low spring snowfall, there may have been an unusual spatial distribution of heat in the debris, because warming could begin earlier at sites without snow, and that could

explain the high inter-site variation of k_d in May 2013. Studies of debris-free glaciers have shown that empirical models do not necessarily hold during seasonal transitions (e.g. Essery and others, 2013; Irvine-Fynn and others, 2014; Litt and others, 2019) and the same conclusion can be drawn from the poor prediction accuracy for May using model KM_d2 . This finding shows that antecedent conditions are sometimes important to seasonal melt at debris-covered glaciers.

More than half of the data in this analysis were from High Mountain Asia, under a climate similar to Dokriani Glacier (Fig. 1). However, no significant difference was found between melt factors or melt rates by region and there was no indication that the timing of observations systematically affected the values of k or a . Therefore, it is unclear whether the temporal variability of melt factors at Dokriani Glacier can be generalized.

Conclusions

This meta-analysis of previously published observations has demonstrated the generality of the dependence of sub-debris melt rates on positive degree-days and debris thickness. In cross-validation, empirical mixed-effects models describing melt factors as an exponential function of debris thickness have been found to predict melt rates with accuracy comparable to that reported for debris-free glacier TIMs transferred between neighboring glaciers. Predictions made for new glaciers with a model fitted with data from multiple glaciers were as accurate as those made for a single glacier with a model fitted with 4 years of data from the same glacier.

Overall, simple models were more accurate in cross-validation than models extended with additional predictors and interaction terms. There was no indication that air temperature data provenance significantly affected melt factors or prediction accuracy. Pooling melt factors calculated with observed melt rates and on-site air temperature and melt factors calculated with observed melt rates and MERRA-2 air temperature into a single large dataset improved prediction accuracy.

Melt factors calculated for Dokriani Glacier increased through the ablation season in 3 out of 4 years of observation. Low values for May in 2010–2012 suggested that energy was used to heat the debris after the spring snow pack melted, delaying the onset of the ablation season for a period on the order of days to weeks. Increasing values between June and September, as well as high inter-site variability in May 2013, suggest that heat storage in the debris layer was a significant part of the energy balance.

The models presented here can be used with MERRA-2 air temperature data to predict regional and glacier scale sub-debris melt rates with a known margin of error, which may be useful for estimating melt rates in the absence of more robust alternatives, or as a baseline for comparison with physically based models. However, the utility of the models for specific purposes may be limited by the relatively wide prediction limits. The 95% prediction limits around predicted melt factors is substantial at approximately –40 to 250% of the predicted values, which demonstrates that melt rates predicted with assumed melt factors without a margin of error may be misleading. An advantage of these models is that the uncertainty associated with predicted ablation is quantified, so this uncertainty can be considered when the output is used in further calculations (e.g. catchment water balances) or decision making.

To advance the science of sub-debris melt modeling, additional studies quantifying the transferability of empirical models would be valuable. It would be worthwhile for the models presented in this paper to be tested independently in a range of geographical settings. Only three observations of melt under debris > 0.65 m thick were found in the published literature and the

models were unable to reproduce those observations. Additional observations would be valuable for testing the accuracy of predictions of melt under debris > 0.65 m thick. The importance of air temperature data provenance to predictions of sub-debris melt at higher temporal and spatial resolution should be tested. Seasonal and multi-year observational studies of debris heat storage would be useful to constrain the significance of heat storage to ice melt in different places. Finally, both empirical and physically-based modeling would benefit if the characteristics of supraglacial debris, including quantities such as porosity and grain size distribution as well as qualities such as rock type, were reported with future field-based studies.

R packages used

beepR (Bååth and Dobbryn, 2018), **bookdown** (Xie, 2018a), **Cite** (Schmitt, 2016), **compiler** (R Core Team, 2017), **digitize** (Poisot, 2011), **dplyr** (Wickham and others, 2018), **ggplot2** (Wickham, 2010), **ggpubr** (Kassambara, 2018), **ggplotify** (Yu, 2019), **grid** (R Core Team, 2017), **gridExtra** (Auguie, 2017), **here** (Müller, 2017), **kableExtra** (Hao and others, 2018), **knitr** (Xie, 2018b), **lme4** (Bates and others, 2015), **lmerTest** (Kuznetsova and others, 2017), **magrittr** (Bache and Wickham, 2014), **merTools** (Knowles and others, 2018), **ncdf4** (Pierce, 2017), **ncdf4.helpers** (Bronaugh, 2014), **ncdf.tools** (Buttler, 2015), **raster** (Hijmans and others, 2017), **RColorBrewer** (Neuwirth, 2014), **rgdal** (Bivand and others, 2018), **rnaturalearth** (South, 2017a), **rnaturalearthdata** (South, 2017b), **tiff** (Urbanek, 2013).

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/jog.2020.57>.

Acknowledgments. This work was partly funded by a UBC Doctoral Fellowship to AWB. Production of the manuscript was supported by a Research NSERC Grant to RDM. We thank V. Radić (University of British Columbia), A. Banerjee (Indian Institute of Science and Education and Research) and R. Shankar (The Institute of Mathematical Sciences, India) for valuable discussions and comments on draft manuscripts that improved the final product beyond measure. We thank the authors of Pratap and others (2015) for contributing their original data from Dokriani Glacier as a data file. We also thank M. Bosilovich at NASA for prompt responses to questions and useful suggestions for working with MERRA-2, and thanks to the Scientific Editor, Joe Shea and an anonymous reviewer for comments leading to improvement of the final manuscript.

References

- Auguie B** (2017) *gridExtra: Miscellaneous Functions for 'Grid' Graphics*. R package version 2.3. Available at url: <https://CRAN.R-project.org/package=gridExtra>
- Ayala A** and 6 others (2016) Modelling the hydrological response of debris-free and debris-covered glaciers to present climatic conditions in the semi-arid Andes of central Chile. *Hydrological Processes* **30**, 4036–4058. doi: [10.1002/hyp.10971](https://doi.org/10.1002/hyp.10971).
- Azócar GF, Brenning A and Bodin X** (2017) Permafrost distribution modeling in the semi-arid Chilean Andes. *The Cryosphere* **11**, 877–890. doi: [10.5194/tc-11-877-2017](https://doi.org/10.5194/tc-11-877-2017).
- Bååth R and Dobbryn A** (2018) *beepR: easily play notification sounds on any platform*. R package version 1.3. Available at url: <https://CRAN.R-project.org/package=beepR>.
- Bache SM and Wickham H** (2014) *magrittr: a forward-pipe operator for R*. R package version 1.5. Available at url: <https://CRAN.R-project.org/package=magrittr>.
- Bates D, Mächler M, Bolker B and Walker S** (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**(1), 1–48. doi: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01).
- Bivand R and 10 others** (2018) *rgdal: bindings for the 'geospatial' data abstraction library*. R package version 1.5–10. Available at url: <https://CRAN.R-project.org/package=rgdal>.

- Bocchiola D and 6 others** (2015) An ablation model for debris-covered ice: the case study of Venerocolo Glacier (Italian Alps). *Geografia Fisica e Dinamica Quaternaria* **38**, 113–128. doi: [10.4461/GFDQ.2015.38.11](https://doi.org/10.4461/GFDQ.2015.38.11).
- Booker DJ and Dunbar MJ** (2008) Predicting river width, depth and velocity at ungauged sites in England and Wales using multilevel models. *Hydrological Processes* **22**, 4049–4057. doi: [10.1002/hyp.7007](https://doi.org/10.1002/hyp.7007).
- Bozhinskiy AN, Krass MS and Popovnin VV** (1986) Role of debris cover in the thermal physics of glaciers. *Journal of Glaciology* **32**(111), 255–266. doi: [10.3189/S0022143000015598](https://doi.org/10.3189/S0022143000015598).
- Braun LN, Grabs W and Rana B** (1993) Application of a conceptual precipitation–runoff model in the Langtang Khola Basin, Nepal Himalaya. *Snow and Glacier Hydrology (Proceedings of the Kathmandu Symposium, November 1992)*. IAHS Publication **218**, 221–237. http://hydrologie.org/red-books/a218/iahs_218_0221.pdf.
- Brock BW, Mihalcea C, Kirkbride MP, Diolaiuti G, Cutler MEJ and Smiraglia C** (2010) Meteorology and surface energy fluxes in the 2005–2007 ablation seasons at the Miage debris-covered glacier, Mont Blanc Massif, Italian Alps. *Journal of Geophysical Research* **115**(D09106), 1–16. doi: [10.1029/2009JD013224](https://doi.org/10.1029/2009JD013224).
- Brock B, Rivera A, Casassa G, Bown F and Acuña C** (2007) The surface energy balance of an active ice-covered volcano: Villarrica Volcano, Southern Chile. *Annals of Glaciology* **45**, 104–114. doi: [10.3189/172756407782282372](https://doi.org/10.3189/172756407782282372).
- Bronaugh D** (2014) *ncdf4.helpers: helper functions for use with the ncdf4 package*. R package version 0.3–5. Pacific Climate Impacts Consortium. Available at <https://CRAN.R-project.org/package=ncdf4.helpers>.
- Brook MS, Hagg W and Winkler S** (2013) Debris cover and surface melt at a temperate maritime alpine glacier: Franz Josef Glacier, New Zealand. *New Zealand Journal of Geology and Geophysics* **56**(1), 27–38. doi: [10.1080/00288306.2012.736391](https://doi.org/10.1080/00288306.2012.736391).
- Brook MS and Paine S** (2012) Ablation of ice-cored moraine in a humid, maritime climate: Fox Glacier, New Zealand. *Geografiska Annaler: Series A, Physical Geography* **94**, 339–349. doi: [10.1111/j.1468-0459.2011.00442.x](https://doi.org/10.1111/j.1468-0459.2011.00442.x).
- Brown ME and 16 others** (2014) An integrated modelling system for estimating glacier and snow melt driven streamflow from remote sensing and earth system data products in the Himalayas. *Journal of Hydrology* **519**, 1859–1869. doi: [10.1016/j.jhydrol.2014.09.050](https://doi.org/10.1016/j.jhydrol.2014.09.050).
- Buttler JV** (2015) *ncdf.tools: easier 'NetCDF' file handling*. R package version 0.7.1.295. Available at <https://CRAN.R-project.org/package=ncdf.tools>.
- Carenzo M, Pellicciotti F, Mabilard J, Reid TD and Brock BW** (2016) An enhanced temperature index model for debris-covered glaciers accounting for thickness effect. *Advances in Water Resources* **94**, 457–469. doi: [10.1016/j.advwatres.2016.05.001](https://doi.org/10.1016/j.advwatres.2016.05.001).
- Chand MB, Kayastha RB, Parajuli A and Mool PK** (2015) Seasonal variation of ice melting on varying layers of debris of Lirung Glacier, Langtang Valley, Nepal. *Remote Sensing and GIS for Hydrology and Water Resources (Proceedings RSHS14 and ICRHWE14, Guangzhou, China, August 2014)*. IAHS Publication **368**, 21–22. doi: [10.5194/piahs-368-21-2015](https://doi.org/10.5194/piahs-368-21-2015).
- Collier E, Nicholson LI, Brock BW, Maussion F, Essery R and Bush ABG** (2014) Representing moisture fluxes and phase changes in glacier debris cover using a reservoir approach. *The Cryosphere* **8**, 1429–1444. doi: [10.5194/tc-8-1429-2014](https://doi.org/10.5194/tc-8-1429-2014).
- Conway H and Rasmussen LA** (2000) Summer temperature profiles within supraglacial debris on Khumbu Glacier, Nepal. *Debris-Covered Glaciers (Proceedings of workshop held at Seattle, Washington, USA, September 2000)*. IAHS Publication **264**, 89–97. doi: <https://iahs.info/uploads/dms/11685.89-97-264-Conway.pdf>.
- Dobhal DP, Mehta M and Srivastava D** (2013) Influence of debris cover on terminus retreat and mass changes of Chorabari Glacier, Garhwal region, central Himalaya, India. *Journal of Glaciology* **59**(217), 961–971. doi: [10.3189/2013JG12J180](https://doi.org/10.3189/2013JG12J180).
- Duan N** (1983) Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association* **78**(383), 605–610. doi: [10.1080/01621459.1983.10478017](https://doi.org/10.1080/01621459.1983.10478017).
- Essery R, Morin S, Lejeune Y and Ménard CB** (2013) A comparison of 1701 snow models using observations from an alpine site. *Advances in Water Resources* **55**, 131–148. doi: [10.1016/j.advwatres.2012.07.013](https://doi.org/10.1016/j.advwatres.2012.07.013).
- Fick SE and Hijmans RJ** (2017) Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37**, 4302–4315. doi: [10.1002/joc.5086](https://doi.org/10.1002/joc.5086).
- Fyfe CL and 6 others** (2014) A distributed energy-balance melt model of an alpine debris-covered glacier. *Journal of Glaciology* **60**(221), 587–602. doi: [10.3189/2014JG13J148](https://doi.org/10.3189/2014JG13J148).
- Gabbi J, Carenzo M, Pellicciotti F, Bauder A and Funk M** (2014) A comparison of empirical and physically based glacier surface melt models for long-term simulations of glacier response. *Journal of Glaciology* **60**(224), 1140–1154. doi: [10.3189/2014JG14J011](https://doi.org/10.3189/2014JG14J011).
- Galecki A and Burzykowski T** (2013) *Linear mixed-effects models using R; a step-by-step approach*. New York: Springer Science + Business Media. doi: [10.1007/978-1-4614-3900-4_10](https://doi.org/10.1007/978-1-4614-3900-4_10).
- Galwey NW** (2014) *Introduction to mixed modelling; beyond regression and analysis of variance* (2nd ed.). West Sussex: John Wiley and Sons, Ltd. doi: [10.1002/9781118861769](https://doi.org/10.1002/9781118861769).
- Gandrud C** (2015) *Reproducible research with R and R Studio* (2nd ed.). Boca Raton: Chapman and Hall/CRC. doi: [10.1201/9781315382548](https://doi.org/10.1201/9781315382548).
- Gelaro R and 30 others** (2017) The modern-era retrospective analysis for research and applications, Version 2 (MERRA-2). *Journal of Climate* **30**(14), 5419–5454. doi: [10.1175/JCLI-D-16-0758.1](https://doi.org/10.1175/JCLI-D-16-0758.1).
- Global Modelling and Assimilation Office (GMAO)** (2015a) MERRA-2 *const_2d_asm_Nx*: 2d, constants, V5.12.4. *Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC)* Greenbelt, MD, USA. Accessed 2018, doi: [10.5067/ME5QX6Q5IGGU](https://doi.org/10.5067/ME5QX6Q5IGGU).
- Global Modelling and Assimilation Office (GMAO)** (2015b) MERRA-2 *inst1_2d_asm_Nx*: 2-dimensional, 1-hourly, instantaneous, single-level, assimilation, single-level diagnostics, V5.12.4. *Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC)* Greenbelt, MD, USA. Accessed 2018, doi: [10.5067/3Z173KIE2TPD](https://doi.org/10.5067/3Z173KIE2TPD).
- Global Modelling and Assimilation Office (GMAO)** (2015c) MERRA-2 *tavg1_2d_rad_Nx*: 2-dimensional, 1-hourly, time-averaged, single-level, assimilation, radiation diagnostics, V5.12.4. *Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC)* Greenbelt, MD, USA. Accessed 2018, doi: [10.5067/Q9QMY5PBNV1T](https://doi.org/10.5067/Q9QMY5PBNV1T).
- Global Modelling and Assimilation Office (GMAO)** (2015d) MERRA-2 *tavg3_3d_asm_Nv*: 3-dimensional, 3-hourly, time-averaged, model-level, assimilation, assimilated meteorological fields, V5.12.4. *Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC)* Greenbelt, MD, USA. Accessed 2018, doi: [10.5067/WWQSQX8IVFW8](https://doi.org/10.5067/WWQSQX8IVFW8).
- Hagg W and 11 others** (2018) Future climate change and its impact on runoff generation from the debris-covered Inylchek glaciers, central Tian Shan, Kyrgyzstan. *Water (Switzerland)* **10**(1531), 1–26. doi: [10.3390/w10111513](https://doi.org/10.3390/w10111513).
- Hagg W, Brook M, Mayer C and Winkler S** (2014) A short-term field experiment on sub-debris melt at the highly maritime Franz Josef Glacier, Southern Alps, New Zealand. *Journal of Hydrology (New Zealand)* **53**(2), 157–165.
- Hagg W, Mayer C, Lambrecht A and Helm A** (2008) Sub-debris melt rates on Southern Inylchek Glacier, central Tian Shan. *Geografiska Annaler: Series A, Physical Geography* **90**(1), 55–63. doi: [10.1111/j.1468-0459.2008.00333.x](https://doi.org/10.1111/j.1468-0459.2008.00333.x).
- Haidong H, Yongjing D and Shiyin L** (2006) A simple model to estimate ice ablation under a thick debris layer. *Journal of Glaciology* **52**(179), 528–536. doi: [10.3189/172756506781828395](https://doi.org/10.3189/172756506781828395).
- Hao Z and 12 others** (2018) *kableExtra*: construct complex table with 'kable' and pipe syntax. R package version 1.1.0. Available at <https://CRAN.R-project.org/package=kableExtra>.
- Hijmans RJ and 26 others** (2017) raster: geographic data analysis and modelling. R package version 3.1–5. Available at <https://CRAN.R-project.org/package=raster>.
- Hock R** (1999) A distributed temperature-index ice- and snowmelt model including potential direct solar radiation. *Journal of Glaciology* **45**(149), 101–111. doi: [10.3189/S0022143000003087](https://doi.org/10.3189/S0022143000003087).
- Hock R** (2003) Temperature index melt modelling in mountain areas. *Journal of Hydrology* **282**, 104–115. doi: [10.1016/S0022-1694\(03\)00257-9](https://doi.org/10.1016/S0022-1694(03)00257-9).
- Hock R** (2005) Glacier melt: a review of processes and their modelling. *Progress in Physical Geography* **29**(3), 362–391. doi: [10.1191/0309133305pp453ra](https://doi.org/10.1191/0309133305pp453ra).
- Hock R, Radić V and De Woul M** (2007) Climate sensitivity of Storglaciären, Sweden: an intercomparison of mass-balance models using ERA-40 re-analysis and regional climate model data. *Annals of Glaciology* **46**, 342–348. doi: [10.3189/172756407782871503](https://doi.org/10.3189/172756407782871503).
- Immerzeel WW, van Beek LPH, Konz M, Shrestha AB and Bierkens MFP** (2012) Hydrological response to climate change in a glacierized catchment in the Himalayas. *Climatic Change* **110**, 721–736. doi: [10.1007/s10584-011-0143-4](https://doi.org/10.1007/s10584-011-0143-4).

- Immerzeel WW, Wanders N, Lutz AF, Shea JM and Bierkens MFP (2015) Reconciling high-altitude precipitation in the upper Indus basin with glacier mass balances and runoff. *Hydrology and Earth System Sciences* **19**, 4673–4687. doi: [10.5194/hess-19-4673-2015](https://doi.org/10.5194/hess-19-4673-2015).
- Intergovernmental Panel on Climate Change (IPCC) (2014) Key economic sectors and services. In *Climate Change 2014 - Impacts, Adaptation and Vulnerability: Part A: Global and Sectoral Aspects: Working Group II Contribution to the IPCC Fifth Assessment Report*. Cambridge: Cambridge University Press, 659–708, doi: [10.1017/CBO9781107415379](https://doi.org/10.1017/CBO9781107415379).
- Irvine-Fynn TDL, Hanna E, Barrand NE, Porter PR, Kohler J and Hodson AJ (2014) Examination of a physically based, high-resolution, distributed Arctic temperature-index melt model, on Midtre Lovénbreen, Svalbard. *Hydrological Processes* **28**, 134–149. doi: [10.1002/hyp.9526](https://doi.org/10.1002/hyp.9526).
- 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**, 377–386. doi: [10.5194/tc-8-377-2014](https://doi.org/10.5194/tc-8-377-2014).
- Juen M, Mayer C, Lambrecht A, Wirbel A and Kueppers U (2013) Thermal properties of a supraglacial debris layer with respect to lithology and grain size. *Geografiska Annaler, Series A: Physical Geography* **95**(3), 197–209. doi: [10.1111/geoa.12011](https://doi.org/10.1111/geoa.12011).
- Kassambara A (2018) ggpubr: 'ggplot2' based publication ready plots. R package version 0.2. <https://CRAN.R-project.org/package=ggpubr>.
- Kasurak A, Kelly R and Brenning A (2011) Linear mixed modelling of snow distribution in the central Yukon. *Hydrological Processes* **25**, 3332–3346. doi: [10.1002/hyp.8168](https://doi.org/10.1002/hyp.8168).
- Kaur S and Purohit M (2013) *Rainfall statistics of India - 2012*. New Delhi: India Meteorological Department (Ministry of Earth Sciences). ESSE/IMD/HS/R.F.REP/02(2013)/16. Available at [http://hydro.imd.gov.in/hydrometweb/\(S\(ca1icn55kriad55mtvwgw45\)\)/PRODUCTS/Publications/Rainfall Statistics of India - 2012/Rainfall Statistics of India - 2012.pdf](http://hydro.imd.gov.in/hydrometweb/(S(ca1icn55kriad55mtvwgw45))/PRODUCTS/Publications/Rainfall%20Statistics%20of%20India%20-%202012.pdf).
- Kaur S and Purohit M (2014) *Rainfall statistics of India - 2013*. New Delhi: India Meteorological Department (Ministry of Earth Sciences). ESSO/IMD/HS/R.F.REP/02(2014)/18. Available at [http://hydro.imd.gov.in/hydrometweb/\(S\(ca1icn55kriad55mtvwgw45\)\)/PRODUCTS/Publications/Rainfall Statistics of India - 2013/Rainfall Statistics of India - 2013.pdf](http://hydro.imd.gov.in/hydrometweb/(S(ca1icn55kriad55mtvwgw45))/PRODUCTS/Publications/Rainfall%20Statistics%20of%20India%20-%202013.pdf).
- Kayastha RB, Takeuchi Y, Nakawo M and Ageta Y (2000) Practical prediction of ice melting beneath various thickness of debris cover on Khumbu Glacier, Nepal, using a positive degree-day factor. *Debris-Covered Glaciers (Proceedings of workshop held at Seattle, Washington, USA, September 2000)*, IAHS. Publication **264**, 71–81. https://iahs.info/uploads/dms/11683_71-81-264-Kayastha.pdf.
- Khan MI (1989) Ablation on Barpu Glacier, Karakoram Himalaya, Pakistan: a study of melt processes on a faceted, debris-covered ice surface. Thesis submitted in partial completion of the requirements of a Master of Arts, Wilfrid Laurier University, Canada.
- Knowles JE, Frederick C and Whitworth A (2018) merTools: tools for analyzing mixed effect regression models. R package version 0.4.1. <https://CRAN.R-project.org/package=merTools>.
- Konz M, Uhlenbrook S, Braun L, Shrestha A and Demuth S (2007) Implementation of a process-based catchment model in a poorly gauged, highly glacierized Himalayan headwater. *Hydrology and Earth System Sciences* **11**, 1323–1339. doi: [10.5194/hess-11-1323-2007](https://doi.org/10.5194/hess-11-1323-2007).
- Koppes M, Rupper S and Asay M, Winter-Billington A (2015) Sensitivity of glacier runoff projections to baseline climate data in the Indus River basin. *Frontiers in Earth Science* **3**(59). doi: [10.3389/feart.2015.00059](https://doi.org/10.3389/feart.2015.00059).
- Kumar R and 7 others (2016) Development of a glacio-hydrological model for discharge and mass balance reconstruction. *Water Resources Management* **30**, 3475–3492. doi: [10.1007/s11269-016-1364-0](https://doi.org/10.1007/s11269-016-1364-0).
- Kustas WP, Rango A and Uijlenhoet R (1994) A simple energy budget algorithm for the snowmelt runoff model. *Water Resources Research* **30**(5), 1515–1527. doi: [10.1029/94WR00152](https://doi.org/10.1029/94WR00152).
- Kuznetsova A, Brockhoff PB and Christensen RHB (2017) lmerTest Package: tests in linear mixed effects models. *Journal of Statistical Software* **82**(13), 1–26. doi: [10.18637/jss.v082.i13](https://doi.org/10.18637/jss.v082.i13).
- Lambrecht A and 6 others (2011) A comparison of glacier melt on debris-covered glaciers in the northern and southern Caucasus. *The Cryosphere* **5**, 525–538. doi: [10.5194/tc-5-525-2011](https://doi.org/10.5194/tc-5-525-2011).
- Lang H and Braun L (1990) On the information content of air temperature in the context of snow melt estimation. *Hydrology of Mountainous Areas (Proceedings of the Štrbské Pleso, Czechoslovakia, June 1988)*, IAHS Publication **190**, 347–354. http://hydrologie.org/redbooks/a190/iahs_190_0347.pdf.
- Lejeune Y, Bertrand JM, Wagnon P and Morin S (2013) A physically based model of the year-round surface energy and mass balance of debris-covered glaciers. *Journal of Glaciology* **59**(214), 327–344. doi: [10.3189/2013JogG12J149](https://doi.org/10.3189/2013JogG12J149).
- Litt M and 6 others (2019) Glacier ablation and temperature indexed melt models in the Nepalese Himalaya. *Nature Scientific Reports* **9**(5264), 1–13. doi: [10.1038/s41598-019-41657-5](https://doi.org/10.1038/s41598-019-41657-5).
- Loveland TR and 6 others (2000) Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing* **21**(6–7), 1303–1330. doi: [10.1080/014311600210191](https://doi.org/10.1080/014311600210191).
- Lundstrom SC, McCafferty AE and Coe JA (1993) Photogrammetric analysis of 1984–89 surface altitude change of the partially debris-covered Eliot Glacier, Mount Hood, Oregon, USA. *Annals of Glaciology* **17**, 167–170. doi: [10.3189/S0260305500012787](https://doi.org/10.3189/S0260305500012787).
- Mair P (2016) Thou shalt be reproducible! a technology perspective. *Frontiers in Psychology* **7**(1079). doi: [10.3389/fpsyg.2016.01079](https://doi.org/10.3389/fpsyg.2016.01079).
- Masiokas MH and 11 others (2016) Reconstructing the annual mass balance of the Echaurren Norte glacier (Central Andes, 33.5°S) using local and regional hydroclimatic data. *The Cryosphere* **10**, 927–940. doi: [10.5194/tc-10-927-2016](https://doi.org/10.5194/tc-10-927-2016).
- Mattson LE, Gardner JS and Young GJ (1993) Ablation on debris covered glaciers: an example from the Rakhiot Glacier, Punjab, Himalaya. *Snow and Glacier Hydrology (Proceedings of the Kathmandu Symposium, November 1992)*, IAHS Publication **218**, 289–296. <https://iahs.info/uploads/dms/9559.289-296-218-Mattson.pdf>.
- Mayer C, Lambrecht A, Hagg W and Narozhny Y (2011) Glacial debris cover and melt water production for glaciers in the Altay, Russia. *The Cryosphere Discussions* **5**(1), 401–430. doi: [10.5194/tcd-5-401-2011](https://doi.org/10.5194/tcd-5-401-2011).
- Mernild SH, Liston GE and Hiemstra CA (2014) Northern Hemisphere glacier and ice cap surface mass balance and contribution to sea level rise. *Journal of Climate (Boston)* **27**(15), 6051–6073. doi: [10.1175/JCLI-D-13-00669.1](https://doi.org/10.1175/JCLI-D-13-00669.1).
- Mihalcea C and 7 others (2008) 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](https://doi.org/10.3189/172756408784700680).
- Mihalcea C, Mayer C, Diolaiuti G, Lambrecht A, Smiraglia C and Tartari G (2006) Ice ablation and meteorological conditions on the debris-covered area of Baltoro glacier, Karakoram, Pakistan. *Annals of Glaciology* **43**, 292–300. doi: [10.3189/172756406781812104](https://doi.org/10.3189/172756406781812104).
- Möller R, Möller M, Kukla PA and Schneider C (2016) Impact of supraglacial deposits of tephra from Grímsvötn volcano, Iceland, on glacier ablation. *Journal of Glaciology* **62**(235), 933–943. doi: [10.1017/jog.2016.82](https://doi.org/10.1017/jog.2016.82).
- Müller K (2017) here: a simpler way to find your files. R package version 0.1. Available at <https://CRAN.R-project.org/package=here>.
- Nakawo M and Young GJ (1981) Field experiments to determine the effect of a debris layer on ablation of glacier ice. *Annals of Glaciology* **2**, 85–91. doi: [10.3189/172756481794352432](https://doi.org/10.3189/172756481794352432).
- Nakawo M and Young GJ (1982) Estimate of glacier ablation under a debris layer from surface temperature and meteorological variables. *Journal of Glaciology* **28**(98), 29–34. doi: [10.3189/S002214300001176X](https://doi.org/10.3189/S002214300001176X).
- Neuwirth E (2014) RColorBrewer: ColorBrewer palettes. R package version 1.1–2. Available at <https://CRAN.R-project.org/package=RColorBrewer>.
- 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](https://doi.org/10.3189/172756506781828584).
- Nicholson L and Benn DI (2013) Properties of natural supraglacial debris in relation to modelling sub-debris ice ablation. *Earth Surface Processes and Landforms* **38**, 490–501. doi: [10.1002/esp.3299](https://doi.org/10.1002/esp.3299).
- Ohmura A (2001) Physical basis for the temperature-based melt-index method. *Journal of Applied Meteorology* **40**, 753–761. doi: [10.1175/1520-0450\(2001\)040<0753:PBFTTB>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<0753:PBFTTB>2.0.CO;2).
- 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.
- Parajuli A, Chand MB and Kayastha RB (2015) Modified temperature index model for estimating the melt water discharge from debris-covered Lirung Glacier, Nepal. *Remote Sensing and GIS for Hydrology and Water Resources (Proceedings RSHS14 and ICGRHWE14, Guangzhou, China, August 2014)*, IAHS Publication **368**, 409–414. doi: [10.5194/piabs-368-409-2015](https://doi.org/10.5194/piabs-368-409-2015).
- Patel LK, Sharma P, Thamban M, Singh A and Ravindra R (2016) Debris control on glacier thinning - a case study of the Batal glacier, Chandra

- basin, Western Himalaya. *Arab Journal of Geosciences* 9(309) doi: [10.1007/s12517-016-2362-5](https://doi.org/10.1007/s12517-016-2362-5).
- Pellicciotti F, Brock B, Strasser U, Burlando P, Funk M and Corripio J** (2005) An enhanced temperature-index glacier melt model including the shortwave radiation balance: development and testing for Haut Glacier d'Arolla, Switzerland. *Journal of Glaciology* 51(175), 573–587. doi: [10.3189/172756505781829124](https://doi.org/10.3189/172756505781829124).
- Pellicciotti F, Buergi C, Immerzeel WW, Konz M and Shrestha AB** (2012) Challenges and uncertainties in hydrological modelling of remote Hindu Kush–Karakoram–Himalayan (HKH) basins: suggestions for calibration strategies. *Mountain Research and Development* 32(1), 39–50. doi: [10.1659/MRD-JOURNAL-D-11-00092.1](https://doi.org/10.1659/MRD-JOURNAL-D-11-00092.1).
- Pierce D** (2017) *ncdf4: interface to unidata netCDF (version 4 or earlier) format data files*. R package version 1.16. Available at <https://CRAN.R-project.org/package=ncdf4>.
- Pradhananga NS and 6 others** (2014) Estimation of discharge from Langtang River basin, Rasuwa, Nepal, using a glacio-hydrological model. *Annals of Glaciology* 55(66), 223–230. doi: [10.3189/2014AoG66A123](https://doi.org/10.3189/2014AoG66A123).
- Pratap B, Dobhal DP, Mehta M and Bhabri R** (2015) Influence of debris cover and altitude on glacier surface melting: a case study on Dokriani Glacier, central Himalaya, India. *Annals of Glaciology* 56(70), 9–16. doi: [10.3189/2015AoG70A971](https://doi.org/10.3189/2015AoG70A971).
- Radić V and Hock R** (2014) Glaciers in the Earth's hydrological cycle: assessments of glacier mass and runoff changes on global and regional scales. *Surveys in Geophysics* 35, 813–837. doi: [10.1007/s10712-013-9262-y](https://doi.org/10.1007/s10712-013-9262-y).
- Ragetti S and 9 others** (2015) Unraveling the hydrology of a Himalayan catchment through integration of high resolution in situ data and remote sensing with an advanced simulation model. *Advances in Water Resources* 78, 94–111. doi: [10.1016/j.advwatres.2015.01.013](https://doi.org/10.1016/j.advwatres.2015.01.013).
- Ragetti S, Pellicciotti F, Bordoy R and Immerzeel WW** (2013) Sources of uncertainty in modelling the glaciological response of a Karakoram watershed to climate change. *Water Resources Research* 49, 6048–6066. doi: [10.1002/wrcr.20450](https://doi.org/10.1002/wrcr.20450).
- Rana B, Nakawo M, Ageta Y and Seko K** (1998) Glacier ablation under debris cover: field observations on Lirung Glacier, Nepal Himalayas. *Ecology of High Mountain Areas (Proceedings of the International Conference, Kathmandu, Nepal, March 1996)*. 393–403.
- R Core Team** (2017) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at <https://www.R-project.org/>.
- Reid TD and Brock BW** (2010) An energy-balance model for debris-covered glaciers including heat conduction through the debris layer. *Journal of Glaciology* 56(199), 903–916. doi: [10.3189/002214310794457218](https://doi.org/10.3189/002214310794457218).
- Reid TD, Carenzo M, Pellicciotti F and Brock BW** (2012) Including debris cover effects in a distributed model of glacier ablation. *Journal of Geophysical Research* 117(D18105), 1–15. doi: [10.1029/2012JD017795](https://doi.org/10.1029/2012JD017795).
- Réveillet M and 8 others** (2018) Relative performance of empirical and physical models in assessing the seasonal and annual glacier surface mass balance of Saint-Sorlin Glacier (French Alps). *The Cryosphere* 12, 1367–1386. doi: [10.5194/tc-12-1367-2018](https://doi.org/10.5194/tc-12-1367-2018).
- Reznichenko N, Davies T, Shulmeister J, McSaveney M** (2010) Effects of debris on ice-surface melting rates: an experimental study. *Journal of Glaciology* 56(197), 384–394. doi: [10.3189/002214310792447725](https://doi.org/10.3189/002214310792447725).
- RGI Consortium** (2017) *Randolph Glacier Inventory (RGI) – a dataset of global glacier outlines: version 6.0, Technical Report*. Colorado, USA: Global Land Ice Measurements from Space, Digital Media. doi: [10.7265/N5-RGI-60](https://doi.org/10.7265/N5-RGI-60).
- Richardson JM and Brook MS** (2010) Ablation of debris-covered ice: some effects of the 25 September 2007 Mt Ruapehu eruption. *Journal of the Royal Society of New Zealand* 40(2), 45–55. doi: [10.1080/03036758.2010.494714](https://doi.org/10.1080/03036758.2010.494714).
- Rienecker MM and 28 others** (2011) MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate* 24(14), 3624–3648. doi: [10.1175/JCLI-D-11-00015.1](https://doi.org/10.1175/JCLI-D-11-00015.1).
- Sasaki O, Noguchi O, Zhang Y, Hirabayashi Y and Kanae S** (2016) A global high-resolution map of debris on glaciers derived from multi-temporal ASTER images. *The Cryosphere Discussions*, doi: [10.5194/tc-2016-222](https://doi.org/10.5194/tc-2016-222).
- Scherler D, Wulf H and Gorelick N** (2018) Global assessment of supraglacial debris-cover extents. *Geophysical Research Letters* 45, 11,798–11,805. doi: [10.1029/2018GL080158](https://doi.org/10.1029/2018GL080158).
- Schmitt S** (2016) *Cite: an RStudio addin to insert BibTex citation in Rmarkdown documents*. R package version 0.1.0. Available at <https://CRAN.R-project.org/package=Cite>.
- Shadish WR, Cook TD and Campbell DT** (2002) *Experimental and quasi-experimental designs for general causal inference*. USA: Wadsworth: Belmont.
- Shea JM, Immerzeel WW, Wagnon P, Vincent C and Bajracharya S** (2015) Modelling glacier change in the Everest region, Nepal Himalaya. *The Cryosphere* 9, 1105–1128. doi: [10.5194/tc-9-1105-2015](https://doi.org/10.5194/tc-9-1105-2015).
- Shea JM and Moore RD** (2010) Prediction of spatially distributed regional-scale fields of air temperature and vapor pressure over mountain glaciers. *Journal of Geophysical Research* 115(D23107), 1–15. doi: [10.1029/2010JD014351](https://doi.org/10.1029/2010JD014351).
- Shea JM, Moore RD and Stahl K** (2009) Derivation of melt factors from glacier mass-balance records in western Canada. *Journal of Glaciology* 55(189), 123–130. doi: [10.3189/002214309788608886](https://doi.org/10.3189/002214309788608886).
- Singh 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](https://doi.org/10.1017/jog.2019.48).
- South A** (2017a) *rnaturalearth: world map data from Natural Earth*. R package version 0.1.0. Available at <https://cran.r-project.org/web/packages/rnaturalearth/rnaturalearth.pdf>.
- South A** (2017b) *rnaturalearthdata: world vector map data from Natural Earth used in 'rnaturalearth'*. R package version 0.1.0. Available at <https://CRAN.R-project.org/package=rnaturalearthdata>.
- Stahl K, Moore RD, Floyer JA, Asplin MG and McKendry IG** (2006) Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. *Agricultural and Forest Meteorology* 139, 224–236. doi: [10.1016/j.agrfor.2006.07.004](https://doi.org/10.1016/j.agrfor.2006.07.004).
- T Poisot** (2011) The digitize package: extracting numerical data from scatterplots. *The R Journal* 3(1), 25–26. doi: [10.32614/RJ-2011-004](https://doi.org/10.32614/RJ-2011-004).
- Trubilowicz JW, Shea JM, Jost G and Moore RD** (2016) Suitability of North American Regional Reanalysis (NARR) output for hydrologic modelling and analysis in mountainous terrain. *Hydrological Processes* 30, 2332–2347. doi: [10.1002/hyp.10795](https://doi.org/10.1002/hyp.10795).
- Urbanek S** (2013) *tiff: read and write TIFF images*. R package version 0.1–5. Available at <https://CRAN.R-project.org/package=tiff>.
- Wang P, Li Z, Li H, Wang W, Zhou P and Wang L** (2017) Characteristics of a partially debris-covered glacier and its response to atmospheric warming in Mt. Tomor, Tien Shan, China. *Global and Planetary Change* 159, 11–24. doi: [10.1016/j.gloplacha.2017.10.006](https://doi.org/10.1016/j.gloplacha.2017.10.006).
- Wheler BA** (2009) *Glacier melt modelling in the Donjek Range, St. Elias Mountains, Yukon Territory*. MSc thesis, Simon Fraser University.
- Wheler BA, MacDougall AH, Flowers GE, Petersen EI, Whitfield LH and Kohfeld KE** (2014) Effects of temperature forcing provenance and extrapolation on the performance of an empirical glacier-melt model. *Arctic, Antarctic, and Alpine Research* 46(2), 379–393. doi: [10.1657/1938-4246-46.2.379](https://doi.org/10.1657/1938-4246-46.2.379).
- Wickham H** (2010) *ggplot2: elegant graphics for data analysis*. New York: Springer-Verlag. Available from <https://ebookcentral.proquest.com/lib/ubc/detail.action?docID=511468>.
- Wickham H, François R, Henry L and Müller K** (2018) *dplyr: a grammar of data manipulation*. R package version 0.7.8. Available at <https://CRAN.R-project.org/package=dplyr>.
- Xie Y** (2018a) *Bookdown: authoring Books and Technical Documents with RMarkdown*. New York: Chapman and Hall/CRC, doi: [10.1201/9781315204963](https://doi.org/10.1201/9781315204963).
- Xie Y** (2018b) *knitr: a general-purpose package for dynamic report generation in R*. R package version 1.20, <https://yihui.org/knitr/>.
- Yang W, Yao T, Xu B and Zhou H** (2010) Influence of supraglacial debris on summer ablation and mass balance in the 24k glacier, southeast Tibetan Plateau. *Geografiska Annaler: Series A, Physical Geography* 92(3), 353–360. doi: [10.1111/j.1468-0459.2010.00400.x](https://doi.org/10.1111/j.1468-0459.2010.00400.x).
- Yu G** (2019) *ggplotify: convert plot to 'grob' or 'ggplot' object*. R package version 0.0.4. Available at <https://CRAN.R-project.org/package=ggplotify>.
- Zhang Y, Fujita K, Liu S, Liu Q and Nuimura T** (2011) Distribution of debris thickness and its effect on ice melt at Hailuoguo glacier, southeastern Tibetan Plateau, using in situ surveys and ASTER imagery. *Journal of Glaciology* 57(206), 1147–1157. doi: [10.3189/002214311798843331](https://doi.org/10.3189/002214311798843331).
- Zhang Y, Liu S and Ding Y** (2007) Glacier meltwater and runoff modelling, Keqicar Baqi glacier, southwestern Tien Shan, China. *Journal of Glaciology* 53(180), 91–98. doi: [10.3189/172756507781833956](https://doi.org/10.3189/172756507781833956).

Appendix A - References for the observational data

Table 6. References and summary statistics for the reported melt factors

Glacier	<i>n</i>	Maximum	Minimum	Geometric mean	Reference
Ice-cored Fox Glacier Moraine	4	5.8	1.1	2.9	Brook and Paine (2012)
Franz Josef Glacier	6	3.0	1.0	2.1	Brook and others (2013)
Southern Inylchek Glacier	3	4.1	0.9	2.4	Hagg and others (2008)
Koxkar Glacier	16	2.6	0.3	1.1	Juen and others (2014)
Khumbu Glacier	5	18.4	5.3	8.8	Kayastha and others (2000)
Djankuat Glacier	7	4.8	0.5	1.7	Lambrecht and others (2011)
Zopkhito Glacier	4	3.6	2.9	3.2	Lambrecht and others (2011)
Maliy Aktru	6	3.1	0.9	1.6	Mayer and others (2011)
Baltoro glacier	22	5.9	1.2	2.9	Mihalcea and others (2006)
Summit Crater Glacier	9	12.9	2.6	5.3	Richardson and Brook (2010)
24k glacier	8	10.6	1.6	3.7	Yang and others (2010)

Appendix B - Comparison of MERRA-2 and on-site air temperature data

Air temperature was reported with the melt observations in 17 of the source publications. These observations were gathered (*n* = 45) and standardized to daily means to inform the choice of a MERRA-2 air temperature variable to use for calculating new melt factors.

MERRA-2 data products for air temperature at 2 m and 10 m above ground were tested, extracted from the 2-dimensional assimilation instantaneous hourly collection (*inst1_2d_asm_Nx*) (*T2* and *T10*). The variable *T2* might be expected to agree more closely with observations made in, near or lapsed to the first few meters of the glacier surface. However, MERRA-2 2-dimensional assimilation data products represent interactions between the surface and atmosphere by a classification of surface type, where the fractional areas of land, lake, ocean and/or ice compose the surface area of a grid cell. The fractional area of each surface type is based on the aggregation of a 1 km² resolution satellite-based global surface characterization (DISCover,

Table 7. References for the reported melt rates and the number of observations at each glacier

Glacier	<i>n</i>	Reference
Venerocolo Glacier	8	Bocchiola and others (2015)
Pichillancahue-Turbio Glacier	3	Brock and others (2007)
Miage Glacier	24	Brock and others (2010)
Lirung Glacier	10	Chand and others (2015)
Chorabari Glacier	17	Dobhal and others (2013)
Miage Glacier	21	Fyffe and others (2014)
Southern Inylchek Glacier	3	Hagg and others (2008)
Franz Josef Glacier	1	Hagg and others (2014)
Koxkar glacier	3	Haidong and others (2006)
Vernagtferner	6	Juen and others (2013)
Barpu Glacier	71	Khan (1989)
Eliot Glacier	12	Lundstrom and others (1993)
Rakhiot Glacier	6	Mattson and others (1993)
Hinarche Glacier	5	Mayer and others (2011)
Baltoro glacier	8	Mihalcea and others (2008)
Ghiacciaio del Belvedere	3	Nicholson and Benn (2006)
Larsbreen	5	Nicholson and Benn (2006)
Batal glacier	9	Patel and others (2016)
Dokriani Glacier	321	Pratap and others (2015)
Lirung Glacier	15	Rana and others (1998)
Qingbingtan Glacier No. 72	3	Wang and others (2017)
Hailuoguo Glacier	9	Zhang and others (2011)

Table 8. The RMSD of observed air temperatures and MERRA-2 air temperatures adjusted for elevation with lapse rates

	RMSD
<i>T2</i> _{ΓC}	2.56
<i>T10</i> _{ΓC}	2.52
<i>T2</i> _{ΓM}	2.77
<i>T10</i> _{ΓM}	2.74

Γ_C = constant lapse rate, Γ_M = mixed lapse rates, from observations and calculated from MERRA-2 air temperature profile.

Loveland and others (2000)). If debris-covered glaciers were misclassified as land or if glacierization were atypical of a grid cell area, *T2* could be systematically warm biased. In that case *T10* could be a better approximation of point scale, ice-cooled air temperature, assuming a lapse temperature profile in the modeled air temperature fields.

Two approaches were tested to arrive at lapse rates. First, an approximate global average of 6.5 °C km⁻¹ (Γ_C) was assumed for all sites. Second, site-specific lapse rates were acquired as stated in the observation source publications (*n* = 3) and otherwise calculated from mid-layer air temperature of model-levels 72 and 71 from the MERRA-2 *tavg3_3d_asm_Nv* collection (GMAO, 2015d), as follows

$$\Gamma M_j = (T_j^{72} - T_j^{71}) / (E_j^{72} - E_j^{71}) \tag{B1}$$

where *E*⁷² and *E*⁷¹ are the model level mid-layer elevations and *j* is the grid cell. Lapse rates were applied following Eqn (5). Four air temperature variables resulted from this procedure, *T2*_{ΓC}, *T10*_{ΓC}, *T2*_{ΓM} and *T10*_{ΓM}, where C stands for constant and M stands for mixed. The four variables are collectively denoted *T_M*.

The root mean squared deviation, RMSD, was used to evaluate agreement between *T_o* and *T_M*, calculated for every observation, *i*, as follows:

$$\text{RMSD} = \sqrt{\frac{1}{n} \sum_{i=1}^n (T_{o,i} - T_{M,i})^2} \tag{B2}$$

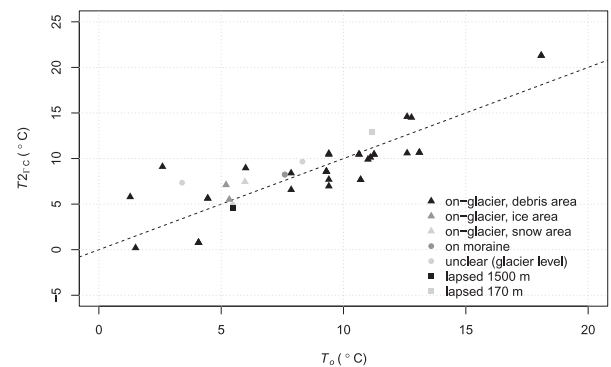


Fig. 6. Variables *T2*_{ΓC} and *T_o*, with symbols distinguishing *T_o* data provenance.

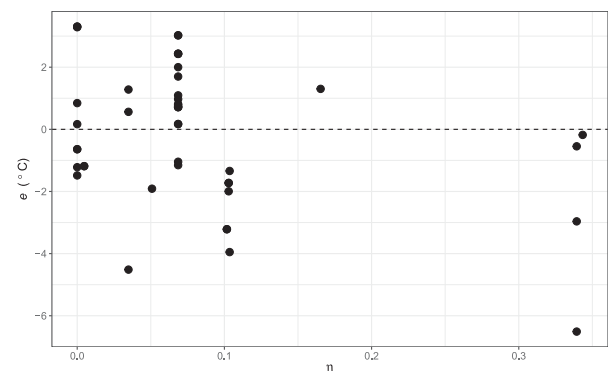


Fig. 7. The difference between *T2*_{ΓC} and *T_o*, *e*, plotted against MERRA-2 grid cell fraction of land ice, *η*.

As seen in Table 8 T_{10} variables were smaller than they were for T_2 but the difference was on the order of just 10^{-4} °C. The variables adjusted with a constant lapse rate were in closer agreement with observations, with RMSD around 0.2 °C smaller for both T_2 and T_{10} . Biases were found to be < 0.01 °C in all cases. The variable $T_{2_{\Gamma C}}$ was chosen for calculating new melt factors in the meta-analysis to be consistent with the common practice of using air temperature data from the glacier near-surface and because the difference between $T_{2_{\Gamma C}}$ and $T_{10_{\Gamma C}}$ was slight.

Air temperature data provenance has been found to affect the consistency of melt factors in some studies. The data for T_o were recorded at automatic weather stations (AWS) that were located on the glacier being studied, on moraine near the melt observation sites, in a base camp in the same glacier basin, or in a local government-operated meteorological station, in which case the data were lapsed to the elevation of the glacier being studied. Figure 6 shows that AWS location did not have a clear affect on the deviation of T_o from perfect agreement with $T_{2_{\Gamma C}}$.

The MERRA-2 grid cell fraction of land ice (η) was nil in six instances, where the glacierized land area is small in reality (Summit Crater Glacier, Pichillancahue-Turbio Glacier, Lirung Glacier, Eliot Glacier, Miage Glacier

and Ghiacciaio del Belvedere). A weak positive relationship between η and the magnitude of the deviation between T_2 and T_o was found, but that relationship did not remain once lapse rates had been applied. The conditions required for glacierization tend to occur at high elevation outside polar regions, and if high elevations are uncommon regionally, the difference between glacier snout elevation and mean regional elevation will be large and the glacierized area will tend to be small. A statistically significant negative linear relationship was found between η and the deviation between $T_{2_{\Gamma C}}$ and T_o ($p = 0.95$), as seen in Figure 7, but there is no physical reason for $T_{2_{\Gamma C}}$ to be warm biased where η is large and cold biased where it is small. Speculatively, the pattern might reflect reduced concentration of observational air temperature data ingested to MERRA-2 with increasing ice area.

At sites near the coast, MERRA-2 air temperature is more strongly influenced by simulated coastal processes, which might overwhelm any signal of local interactions between the air and ice, particularly if the ice area is small. The deviation between T_2 and T_o was found to be slightly higher for glaciers near the coast than glaciers further from the coast but again, that was corrected with the application of elevation lapse rates, because the glaciers near the coast in these data also tend to be at relatively high elevation.