Predictability of ~~mean~~ summertime diurnal winds over ungauged mountain glaciers

Krishnanand J[1], Argha Banerjee[1], Shankar R[2], Himanshu Kaushik[3], Mohd. Farooq Azam[3], and Chandan Sarangi[4]

[1]Earth and Climate Science, Indian Institute of Science Education and Research Pune, Maharashtra, India
[2]The Institute of Mathematical Sciences, Chennai, Tamil Nadu, India
[3]Department of Civil Engineering, Indian Institute of Technology Indore, Madhya Pradesh, India
[4]Department of Civil Engineering, Indian Institute of Technology Madras, Tamil Nadu, India

Correspondence: Argha Banerjee (argha@iiserpune.ac.in)

Abstract. Glacier and valley winds are ~~typical characteristics~~ characteristic features of the microclimate of glacierised valleys. The speed of such winds ~~determines~~ influences the turbulent heat flux, which contributes to ice melt significantly. Sparse in-situ meteorological measurements and the inability of large-scale climate data products to capture such local winds introduce uncertainty into glacier- to global-scale mass-balance calculations. Here, we propose an empirical model having three parameters, namely, the mean wind speed, the sensitivity of the diurnal winds to temperature, and a ~~response time~~time-scale parameter, to predict the mean summertime diurnal wind speed on valley glaciers, based only on reanalysis temperature. Utilising data from 28 weather stations on 18 valley glaciers across the globe, we show that the model, driven solely by reanalysis temperature, reproduces the observed mean summertime diurnal wind speed ~~reasonably well. Furthermore, we show that~~ with a mean RMSE of 0.24 ms^{-1} across stations, which is 8.8% of the mean observed wind speed. Interestingly, the three model parameters can be estimated ~~at any glacier using a few topographic variables, allowing prediction of wind speed on ungauged glaciers~~even on an ungauged glacier based on the mean temperature and regional slope. A leave-one-out analysis of the stations suggests a root-mean-squared error of ~~0.76~~ 0.85 ms^{-1} on average, which is a $\sim$~~300~~200% improvement over a standard reanalysis product. The performance of the model is largely independent of the number of stations available for calibration, as long as it is ~~20~~ 21 or more. ~~More work is needed to explain the physical mechanisms underlying the predictability of the mean diurnal wind speed on ungauged glaciers based solely on reanalysis temperature and a few topographic variables. The presented model~~ The presented empirical parametrisation can improve wind speed estimates on ungauged glaciers, leading to better glacier mass-balance calculations~~at various spatial scales~~.

1 Introduction

The microclimate of a glacierised valley has a characteristic system of local winds consisting of a glacier wind and a valley wind ~~(e.g., ?)~~(e.g., ???). Glacier wind is a shallow~~(of the order of $\sim 10^1$ m)~~, persistent, katabatic wind flowing down-valley~~,~~ ~~consisting of air cooled~~. It's caused by the cooling of air by the ice surface ~~acted upon by a~~ and the associated down-slope buoyancy ~~forcing~~ force (?). Away from the glacier, as the valley floor and the hillsides are heated up during the day by solar insolation, the ~~buoyancy forcing~~ buoyant forces acts up-slope, leading to a daytime ~~up-valley valley windthat is relatively deeper (of the order of~~ anabatic valley wind, which flows up-valley. It is a relatively deeper wind with a depth of $\sim 10^2$ m ~~)~~ ~~(????)~~(**??**). Near the glacier termini, glacier and valley winds compete with each other at the surface level during the day (e.g., **?**). The buoyancy forcing in the valley, and consequently the valley wind, changes sign at night (e.g., **?**). This leads to a diurnal cycle of valley winds with up-valley (down-valley) flow during the day (night). ~~Both of these winds~~

Glacier winds and valley winds fall under the general category of thermally driven slope winds (**?**), which can be described approximately using the steady-state Prandtl model (**?**). This model describes the vertical profile of slope winds via a dynamic equilibrium between buoyant force and friction, and reproduces the characteristic wind speed maxima or jet at a height of a few meters above the surface (e.g., **?**). ~~With a thermodynamic approach to the problem, a simple energy exchange model was also successfully used to describe the effect of katabatic cooling due to glacier wind on the variation of 2m-air temperature along the flow-line of glaciers (??).~~ During the ablation season, a linear relation between instantaneous glacier wind speed and the ambient air temperature has been empirically observed (e.g., **?**). Similar trends have also been reported for valley winds (e.g., **?**). This may imply that the buoyancy forces scale with the horizontal temperature gradient between the valley and the surrounding terrain (**?**). A simple energy exchange model was successfully used to describe the effect of katabatic cooling due to glacier wind on the variation of 2m-air temperature along the flow-line of glaciers (**??**).

Accurate knowledge of instantaneous wind speed on glaciers is important for energy- and mass-balance studies, as surface winds mediate energy and mass exchanges between the atmosphere and glacier surface (**??**). ~~For example, the~~ The latent heat fluxes ~~due to these winds determine sublimation ratesof snow~~resulting from these winds influence snow sublimation rates, which can ~~cause up to $\sim$ 10–90% of the total glacier melt (e.g., ???), and~~ account for $\sim$ 10–20% melt contribution for typical glaciers and up to 90% contribution for glaciers in cold, arid climates (e.g., **??**). Latent heat fluxes are the dominant source of ablation for high-altitude glaciers in arid regions (**?**). The net turbulent heat fluxes, which is the sum of latent and sensible heat fluxes due to these winds, can cause up to $\sim$ 10–35% of the total melt on ~~some~~ glaciers in the Himalaya (e.g., **?**), the Alps (e.g., **?**), and Alaska (e.g., **?**). The turbulent fluxes are also important in the context of the climate response of glaciers. This is because the response of glaciers to the temperature change is largely determined by the corresponding changes in the long-wave radiation balance and those in the turbulent heat fluxes (**?**). ~~The~~ Glacier winds can affect the spatial gradients of near-surface ~~lapse rate on a glacier surface varies depending on the presence of glacier or valley winds (?), which has important consequences for the accuracy of the glacier melt as estimated using simple~~ meteorological variables such as temperature and humidity on glaciers (e.g. **?**). Simple temperature-index ~~models (??). The~~ based methods (**?**) for estimating glacier melt involve extrapolating the near-surface air temperature for the entire glacier surface using lapse rates estimated from nearby weather stations. However, this approach can lead to overestimations of glacier melt because the air near glaciers is cooler than its immediate surroundings. The gradients of temperature, humidity, and wind speed drive heat, moisture, and momentum exchanges between valley winds and glacier winds. The presence of extensive supra-glacial debris cover favours valley winds over the ablation zone (**?**), which ~~also~~ influences the glacier energy balance and the moisture transport along the valley (**?**).

The above discussion highlights the ~~importance~~ need of accurate wind speed estimates for glacier energy- and mass-balance studies. However, the number of glaciers where in-situ meteorological measurements are ~~(publicly)~~ publicly available is probably of the order of $\sim 10^2$ or less. This means ~~a~~ an overwhelming majority of $\sim 2.75 \times 10^5$ glaciers on earth are ungauged. On these glaciers, coarse-scale reanalysis ~~wind-speed~~ wind speed products need to be used to estimate the turbulent heat fluxes (e.g., **??**). Even on the small minority of the glaciers, where weather stations are present, the data are often available for a relatively short period, or have gaps. On such glaciers, coarse-scale reanalysis data (e.g., **??**) have to be used in conjunction with the weather station data. The coarse $\gtrsim$10-km resolution reanalysis products do not capture local-scale winds, and heavily underestimate their speed, and therefore, the associated turbulent fluxes (**?**). The reanalysis temperature estimates on the other hand, owing to a longer spatial correlation length, are relatively more accurate (**?**). Various down-scaling methods are typically applied to these coarse-scale ~~wind-speed~~ wind speed products to estimate the local-scale winds. Statistical down-scaling models (**??**), which essentially fit coarse-resolution climate products to point-scale weather-station data, do not work well on data-scarce mountain glaciers. Physically based methods (e.g., **?**) that rely on simple scaling relations between local and large-scale climate variables cannot be applied on glacierised valleys as they do not incorporate the microclimatic effects of glaciers. Dynamical down-scaling ~~(**???**)~~ (**??**) requires high-resolution regional climate model simulations, which are computationally intensive and are impractical for regional- to global-scale studies. Thus, there is a clear need for a simple~~physically-based~~, parsimonious, and preferably physically motivated model of wind speed on glaciers that can be applied to ungauged mountain glaciers globally.

This study proposes a ~~simple~~ parsimonious model to estimate the mean diurnal cycle of wind speed during summer on any glacier, solely based on easy-to-access reanalysis temperature and topographic data. In-situ data from a set of 28 on- or near-glacier weather stations worldwide are used to calibrate the model. We evaluate the predictive power of this model on an ungauged glacier using a leave-one-out analysis, and also assess its performance as a function of the number of stations used for calibration. We discuss the limits of applicability of the model and its potential utility for ungauged glaciers.

2 Data

2.1 Automatic weather stations

The hourly time series of wind speed and wind direction observed at 28 weather stations (Fig. 1~~a~~) were used in this study. The data were primarily sourced from open-access online databases ~~Zenodo and Pangaea~~like Zenodo (https://zenodo.org) and Pangaea (https://www.pangaea.de), along with a few other publicly available datasets from the literature (Supplementary Table S1). Since we were interested in on- or near-glacier winds, a spatial filter was applied to select ~~the~~ stations within a 1 km buffer around any glacier ~~(Fig.~~ in the same valley. This was done to ensure the stations were located within the same valley, which was visually verified through Supplementary Fig. ~~1b). Only stations with more than 8 weeks of data during the summer months were used. For glaciers located in the northern (southern) hemisphere, the summer was assumed to be 1st June to 30th September (1st December to 31st March).~~

S1. This procedure yielded 28 weather stations located on or near 18 glaciers with an approximately global coverage (Fig. 1a), though the southern hemisphere was under-represented. There were 14 stations from the Himalaya, ~~6~~ six from the Canadian Rockies, ~~5~~ five from the European Alps, and one each from Alaska, the Caucasus, and the Southern Andes. Out of the 28 stations, 27 were within latitudes of 28°–80°N, and only one station was in the southern hemisphere. The elevation of the stations ranged from 196 m to 8323 m above sea level. The available stations sampled a wide variety of mountain glaciers, with their lengths being in the range 2–45 km and the widths in the range 0.3–14.2 km. Six of the glaciers were debris-covered ones, and the remaining 12 were clean glaciers (Supplementary Table S1). The number of days for which data were available ranged between 56–586, with a median of 198. For the 28 weather stations, the hourly time-series of 2-m wind speed for the summer months were selected for all the available years (see Fig. 2a for an example). Only stations with more than eight weeks of data during the summer months were used. For glaciers located in the northern (southern) hemisphere, the summer was assumed to be 1st June to 30th September (1st December to 31st March). The mean summertime diurnal wind speed at a station was calculated by averaging wind speeds for each hour of the day for the entire duration of the dataset (Fig. 2b).

2.2 Climate and topography

~~We used ∼9-km~~ We used $0.1° \times 0.1°$ resolution ERA5-Land (ERA5L) hourly reanalysis data (**?**) from European Centre for Medium-Range Weather Forecasts ~~(**?**)~~ to obtain the 2-m air temperature, matching the observation period and location of each of the 28 weather stations. The temperature data were corrected for the ~~station elevation~~ elevation difference between the station and the corresponding ERA5L pixel using an hourly lapse rate. This lapse rate was derived by regressing ~~the~~ temperature and elevation for all ~~the~~ ERA5 pixels within a 10~~0~~×~~100 km^2 buffer centered~~ 10 grid centred around the station. The 10-m wind speed from ERA5L was used for comparison with the 2-m windspeed predicted using the model developed here. We did not attempt to correct the ERA5L wind speed to bring it to 2-m level, ~~which was not straightforward as the vertical structure of the boundary layer above glaciers are complex (e.g., **?**)~~using the logarithmic law as is typically done in these cases (**?**). Wind profiles of slope wind systems are known to have a maxima, and hence cannot be described with a logarithmic profile. Glacier geometry data from Randolf Glacier Inventory 7 (**?**) and topographic static variables derived from the 30-m resolution digital elevation model (DEM) from ALOS World 3D (**?**) was used for predicting the three model parameters (see Sect. 3.1.1).

3 Methods

3.1 Model

3.1.1 Model for wind speed

Empirically, glacier wind is known to have a characteristic diurnal variation, with the maximum wind speed occurring in the afternoon (e.g., **???**). It has also been noted that the timing of the wind speed maxima lags behind that of the temperature by a few hours (**?**). The time series of the observed wind speed and ERA5-derived temperature for Drang Drung Glacier, the western Himalaya illustrate this behaviour (Fig. 2). It may be seen that the hourly ~~wind-speed~~ wind speed signal is noisy (Fig. 2a). However, the mean diurnal wind speed has a relatively smooth variation, which appears to be a delayed response to the mean diurnal 2-m air temperature on this glacier (Fig. 2b). A similar behaviour is known to be present in ~~land-sea breeze~~

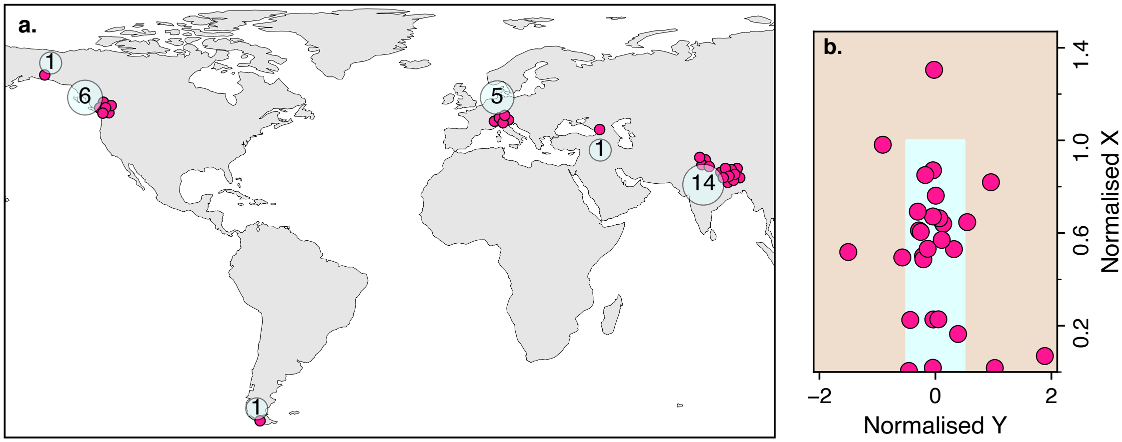

Figure 1. (a) Locations of 28 on- or near-glacier weather stations (red solid circles) used in this study. The circles have been artificially spread out up to about 2° to avoid overlap. The number of stations in different regions are shown within light blue circles. (b) The relative, scaled locations of the weather stations used in this study (red circles) with respect to an idealised glacier (light blue rectangle) having a unit length and width. Here X and Y denotes the longitudinal and transverse distance, respectively, with respect to a coordinate system whose origin is at the headwall of the idealised glacier, and X increases downglacier.

~~and~~ valley-wind systems, and consequently, ~~both of the system~~ it could be described well using simple linear-response models ~~(e.g., ??).~~ (e.g., **?**). A similar linear-response model exists for land-sea breeze (**?**).

Motivated by this, here we attempted to predict the mean diurnal wind speed for both glacier winds and valley winds on or near a glacier, using the corresponding mean diurnal temperature. We restricted our analysis to the summer season as most of the melt happens during this time. ~~Let u denote the~~ We used a coordinate system in which z-axis points vertically up and x-axis is along the glacier flow direction. We assume that winds are generally either in the upslope or downslope direction. We decomposed the mean summertime diurnal wind speed ~~, which was decomposed~~ u, into its mean value $\bar{u}$ and the diurnal anomaly u_d, i.e., $u \doteq \bar{u} + u_d$. Similarly, the mean summertime diurnal temperature was decomposed into $\bar{T}$ and T_d. Then, we modelled u_d as ~~a~~ an empirical linear combination T_d and its time derivative $\frac{d}{dt}T_d$, to obtain the following equation:

$$u = \bar{u} + sT_d - s\tau \frac{d}{dt} T_d, \tag{1}$$

where, s is the sensitivity of the mean diurnal ~~wind-speed~~ wind speed anomaly to that of temperature, and τ is the corresponding ~~response time~~time-scale parameter. Note that this is not a ~~traditional linear response~~ linear-response model (e.g., **??**), where the forcing variable (e.g., T_d) is taken to be ~~a~~ an empirical linear combination of the response variable (e.g., u_d) and its time derivative~~$\frac{d}{dt}u_d$~~. We chose the present empirical version over a traditional linear-response model because the derivative of wind speed was found to be relatively more noisy as compared to that of temperature (Fig. 2b), which resulted in poor fits. In our problem, the dominant frequency ω was diurnal (~ 0.26 hr^{-1}) and the typical ~~response time~~ time-scale parameter τ was

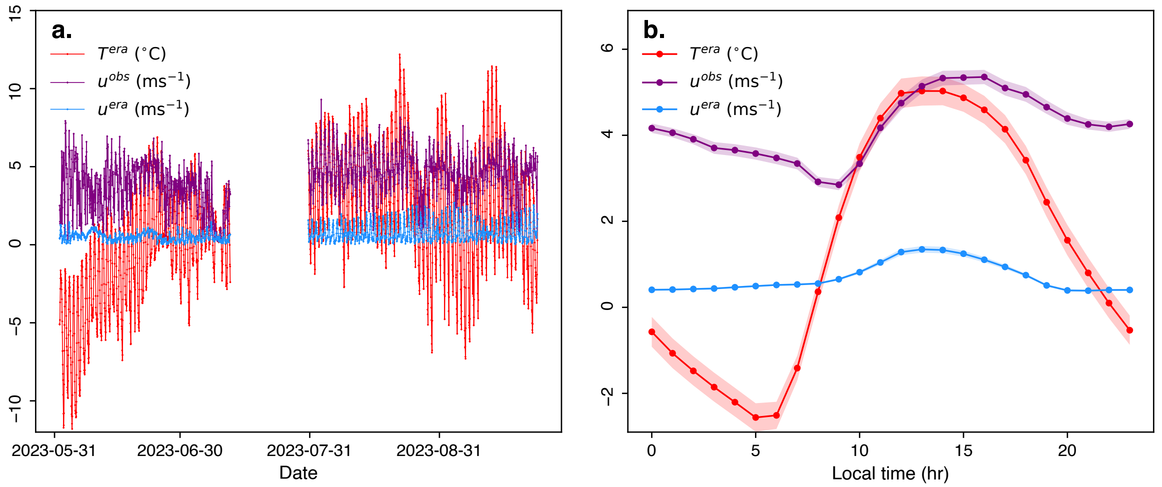

Figure 2. a) Time-series of hourly wind speed (purple line) observed on Drang Drung Glacier in the western Himalaya, along with corresponding ERA5-Land temperature (red line) and wind speed (light blue line) estimates during the summer season of 2023. b) The corresponding mean diurnal cycles, with the standard error of mean shown as shaded bands.

~~a few hours~~ $\sim$ 1 hour, so that ~~$\omega^2\tau^2 \to 0$~~$\omega^2\tau^2 \ll 1$. In this limit, both models were~~in fact~~, in fact, equivalent to each other (Supplementary Text S1).

3.1.2 ~~Model~~ Extracting model parameters for the station

~~To apply the above model to an ungauged glacier, one requires the knowledge of T_d and three model parameters $\bar{u}$, s, and τ for that location. It is possible to obtain T_d from reanalysis datasets. However, the model parameters are not known a priori. Wind speeds on complex terrain have been linked to topography before (??). Motivated by this, we test if the model parameters $\bar{u}$, s, and τ can be estimated for any given glacier as simple linear combinations of static variables associated with glacier geometry and local topography etc., i.e., s~~

$$\bar{u} = \beta_0^u + \beta_1^u \pi_1^u + \beta_2^u \pi_2^u + \cdots,$$

$$s = \beta_0^s + \beta_1^s \pi_1^s + \beta_2^s \pi_2^s + \cdots,$$

$$\tau = \beta_0^\tau + \beta_1^\tau \pi_1^\tau + \beta_2^\tau \pi_2^\tau + \cdots,$$

~~where π_i^x denotes the i^{th} static variable, and β_i^x denotes the corresponding regression coefficient for the model parameter x.~~

~~We searched through a list of 37 static variables (Supplementary Table S2) to come up with the best set of predictors π_i^x. The static variables included 22 valley/glacier-geometry variables (like width and length of the glacier/valley), 12 topographic~~

~~variables (like relief and mean elevation), and three climatic variables (e.g., continentality). For deriving the topographic variables for a station location from the DEM, a combination of QGIS (?) and Google Earth Engine (GEE) (?) was used (Supplementary Table S2).~~

~~The three-step workflow for modelling the mean summertime diurnal wind on an ungauged glacier. (Step 1) Wind speed and ERA5 temperature in each station were fitted to Eq 1 separately, obtaining 28 triplets of best-fit $\bar{u}$, s, and τ. (Step 2) This set of 28 triplets—excluding s and τ from Langenferner Glacier, as discussed in section 4.1.2—were used to obtain the best-fit multi-linear regression coefficients defined in Eqs. 2–4. (Step 3) For any ungauged glacier location, the best-fit Eqs. 2–4 were used to predict the three model parameters, which along with the ERA5L temperature, were fed to Eq. 1 to obtain the wind speed predictions.~~

3.2 ~~Calibrating model~~

3.1.1 ~~Fitting Eq. 1 to observations~~

~~The~~ The model parameter $\bar{u}$ for each of the 28 stations was obtained by averaging the observed mean summertime diurnal wind speed u^{obs}. The remaining two model parameters s and τ were obtained for each station by performing a bilinear least-square regression of T_d and its derivative $\frac{d}{dt}T_d$, against u_d^{obs} (Eq. 1). This obtained 28 triplets of best-fit parameters $\{\bar{u},\, s,\, \tau\}$ – one each for each station. Here, the diurnal variation T_d was obtained by decomposing the ERA5L-derived mean diurnal summertime temperature ~~T^{era}~~T, into its mean and the anomalies. The time derivative of T_d was obtained by calculating the first differences. ~~The root-mean-squared errors (RMSE) of the fits were used to assess the model performance.~~

3.1.1 ~~Fitting Eqs. 2–4 to model parameters~~

~~The best-fit regression coefficients in Eq. 2 were obtained by regressing the 28 values of~~

3.1.1 Estimating model parameters for ungauged glaciers

To apply the above model to an ungauged glacier, one requires the knowledge of T_d and three model parameters $\bar{u}$~~against the static variables . While the RMSE of a multilinear regression generally decreases with the increasing number of static variables, we limited the number of static variables used in Eq.~~, s, and τ for that valley. It is possible to obtain T_d from reanalysis datasets. However, the model parameters are not known a priori. We derived predictors for each model parameter based on some simple dimensional arguments, expressing them in terms of easy-to-access variables like temperature and slope.

Let us denote the air density with ρ, acceleration due to gravity with g, and the difference between the air temperature near the glacier surface, T_g, and the ambient temperature, T_a to be θ. Note that θ is positive for anabatic winds and negative for katabatic winds. For along-slope flow, the vertical component of wind velocity w is related to the horizontal component u, as $w \sim u\varsigma$, where ς is a ratio between a vertical (H) and horizontal length scale (L), e.g., glacier slope. For slope flows, it is assumed that the buoyancy force is balanced by turbulent friction (**?**). From Eq. ~~2 by the following method, to avoid over-fitting. For an n-variable regression model, there were $\binom{37}{n}$ possible choices for the static variables. For~~ 3.3.10 of **?**, the buoyancy term in the momentum equation for slope flows is given by $\frac{g\theta\sin\eta}{T_a}$, where η is the glacier slope in radians. Assuming $\eta \ll 1$ for typical glaciers, we have $\sin\eta \sim \tan\eta \sim \varsigma$. Hence the buoyancy force term scales as $\sim \frac{g\theta\varsigma}{T_a}$. The turbulent friction term is estimated as $\partial_z \overline{u'w'}$, where u' and w' are high frequency components of horizontal and vertical wind speeds (e.g., **?**) . We assume u' and w' to have scales similar to u and w respectively. Assuming that the vertical length scale in turbulent friction term is related to the vertical elevation range H of the ~~given n, starting from $n = 1$, all such possible combinations were systematically tested to find the set of static variables having the lowest RMSE between the predicted and observed $\bar{u}$ for the 28 stations~~ glacier and $w \sim u\varsigma$, we can write $\partial_z \overline{u'w'} \sim \frac{uw\varsigma}{H} \sim \frac{u^2\varsigma}{H}$. Balancing the buoyancy and turbulent friction terms, we have:

$$u^2 \sim \frac{gH\theta}{T_a}. \tag{2}$$

Since the ambient temperature across stations typically varies by at most about 10% of its absolute magnitude ($\sim$ 300 K), we assume T_a to be effectively constant. Applying these assumptions and dropping all the constants, we obtain the following relation:

$$u^2 \sim (H\varsigma)\frac{\theta}{\varsigma}. \tag{3}$$

Steeper glaciers are typically smaller and tend to have smaller elevation range. Therefore, we hypothesise that rather than the factor of $H\varsigma$, the variation in θ/ς controls the variability of u. Finally, we propose the following approximate relation for our model parameter $\bar{u}$.

$$\bar{u} \sim \sqrt{\frac{\theta}{\varsigma}}. \tag{4}$$

The sensitivity of wind speed variation to temperature variation s is likely related to the ratio of the diurnal wind-speed range to the temperature range. Assuming that near the glacier surface, the diurnal wind speed range scales as the horizontal wind speed u, and that the diurnal temperature range scales as θ, we have:

$$s \sim \frac{u}{\theta} \sim \frac{1}{\sqrt{\theta\varsigma}}. \tag{5}$$

The model parameter τ, which is the response-time scale, is assumed to scale with the ratio of some characteristic horizontal length scale (L) to the wind-speed scale (u). This leads to the following approximate relation:

$$\tau \sim \frac{L}{u} \sim \sqrt{\frac{\varsigma}{\theta}}. \tag{6}$$

Assuming the temperature of a melting glacier surface T_g is close to 0°C, θ in Eq.~~This procedure was repeated for n varying from 1, 2, 3, ... The iteration was stopped when the reduction in RMSE upon adding a new variable ($n \to n+1$) was smaller than the reduction achieved in the previous step (Supplementary Fig. S2). The same procedure was then repeated to find the~~ 4–6 is equal to the magnitude of the ambient air temperature, which may be taken as the absolute value of the mean summertime ERA5-Land temperature, denoted by $|\bar{T}|$. Apart from three choices of topographic slope, defined in Supplementary Table S2, we also considered the ratio of any vertical scale to any horizontal scale as a slope. With seven choices for vertical scale, e.g. glacier elevation range and 10-km relief, three choices for horizontal length scale, e.g. glacier length and glacier width, we tested 24 possible candidates for the RHS of Eqs. 4–5, and 72 candidates for the RHS of Eq. 6. Given the limited sample size and the noise level, we relied primarily on visual comparison to assess the performance of these candidates. The 28 model parameter triplets were plotted against their corresponding predictor values and visually assessed. Assuming a linear relation, we performed linear regressions between each model parameter with its best predictor. The resulting parameterisations are represented as:

$$\bar{u} = \beta_1 \sqrt{\frac{|\bar{T}|}{\varsigma}}, \tag{7}$$

$$s = \beta_2 \frac{1}{\sqrt{|\bar{T}|\varsigma}}, \tag{8}$$

$$\tau = \beta_3 \sqrt{\frac{\varsigma}{|\bar{T}|}}, \tag{9}$$

where $\{\beta_1, \beta_2, \beta_3\}$ are the triplet of best-fit ~~regression coefficients for Eqs. 3, 4, which were used to predict s and τ, respectively, for any glacier.~~

linear regression coefficients. If no clear relationship was seen between a model parameter and any of its predictors, the median value of that model parameter across the 28 stations was used as the predictor.

3.2 Model performance

3.2.1 Performance of the calibrated model

The performance of the calibrated model (Eqs.~~1–4), which was calibrated using the procedure described above, was tested~~ 1,7–9), was assessed on each of the 28 weather stations. For each station, the mean summertime diurnal wind speed predicted by the calibrated model at hourly time steps were compared with the corresponding observed values. ~~Note that these predicted values were obtained from the calibrated equations, only using the corresponding ERA5L temperature and the static variables as inputs.~~ The mean, median, and interquartile range of the set of 28 ~~RMSEs~~ root-mean-squared errors (RMSE) between the model predictions and observations were compared with the set of RMSEs between the ERA5L products and observations. Note that these predicted u values were obtained from the calibrated equations, only using the corresponding ERA5L temperature and the static variables as inputs.

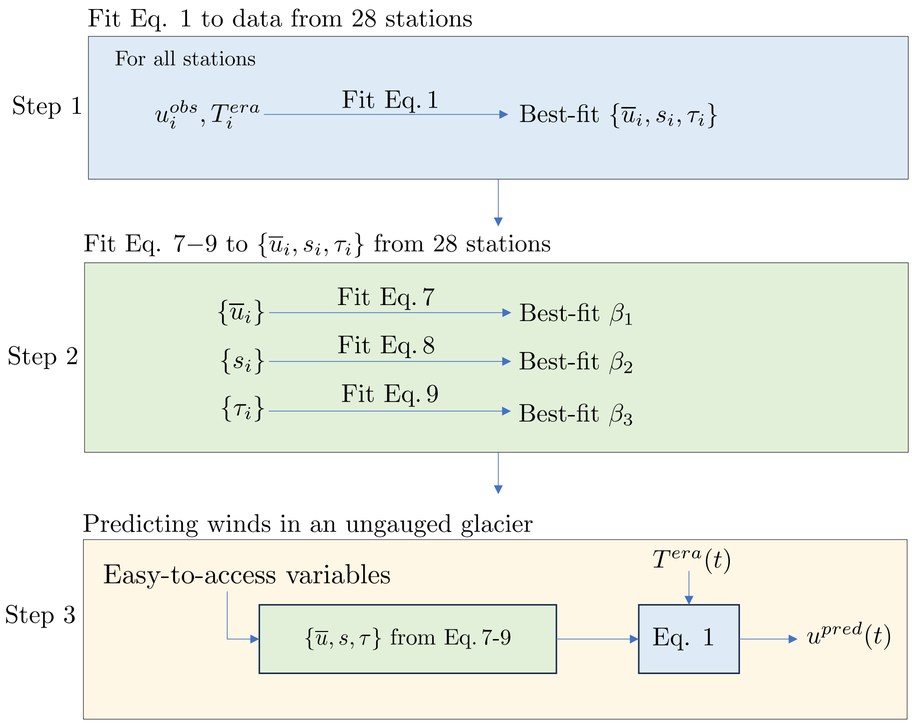

Figure 3. The workflow for modelling the mean summertime diurnal wind on an ungauged glacier employed in this paper. Step 1: Observed wind speed and ERA5 temperature at each station were fitted to Eq.1 separately, obtaining 28 triplets of best-fit $\{\bar{u}, s,$ and $\tau\}$. Step 2: This set of 28 triplets were used to obtain the best-fit regression coefficients defined in Eqs. 7–9. (Step 3) For any ungauged glacier location, the best-fit Eqs. 7–9 were used to predict the three model parameters, which along with the corresponding ERA5L temperature, were fed to Eq. 1 to obtain the wind speed predictions.

3.2.2 Leave-One-Out-Cross-Validation

To evaluate the performance of the model for an ungauged location, we performed a Leave-One-Out-Cross-Validation (LOOCV) (**?**). For this, we calibrated the model following the procedure described in Sect. ~~??~~ 3.1.1 and 3.1.1 with 27 out of 28 weather stations. Note that the LOOCV results may vary depending on whether we search for the best set of predictors at each iteration. However, the present estimates are likely to be conservative, as we may not be working with the best model in some of the iterations, which is likely to increase the RMSEs. The calibrated model was then used to predict the wind speed at the left-out 28th weather station, using only the corresponding ERA5L temperature and ~~the static variables~~slope, and without using any in-situ data from the ~~28th~~ left-out station. We repeated this process 28 times, leaving out a different station each time, and computing the RMSE between the predicted and observed wind speeds for the left-out station. The mean, median, and interquartile range of the set of 28 RMSEs thus obtained were compared with the set of RMSEs between the ERA5L products and observations.

3.2.3 Minimum number of stations required

We tested how sensitive the predictions from the calibrated model were to the number of stations available for calibration. To do this, the model was calibrated only using a randomly selected subset of N stations. ~~The step was repeated for $N = 13, 14, ..., 28$.~~ For each subset, the model was calibrated exactly as discussed in Sect. ~~??~~3.1.1–3.1.1, and then evaluated across all 28 stations by computing the RMSEs between the observed and predicted mean summertime diurnal wind speed. The median RMSE was recorded. The step was repeated for $N = 13, 14, ..., 28$. The exercise was repeated 100 times to obtain a set of 100 mean RMSEs for each N. The resulting distributions of RMSEs were plotted as a function of N, to see if the RMSEs converge to a fixed value with increasing N. The minimum value of N, for which the median RMSE differed by less than 1% from that obtained for $N = 28$, was defined as the minimum number of stations required for model calibration. Note that this exercise was also helpful in ascertaining the model performance on ungauged glaciers, as only a subset of glaciers were used here to calibrate the model, which was then used to predict the wind speed on all 28 stations.

3.2.4 ~~Time scale dependence of prediction~~

~~We tested if the calibrated model, which was derived for mean summertime diurnal winds, could be used to predict wind speeds across different time scales. We first filtered the ERA5L hourly temperature data over different time scales by taking 1-hr to 10-days moving average. For each station and each averaging window, the filtered temperature data were fed into the calibrated model (Sect.**??**) to obtain wind-speed predictions for that station and that time scale. The RMSE between the predicted and observed wind speeds was computed for each station, yielding a set of 28 RMSEs. For this RMSE calculation, the observed hourly wind speed data were also filtered using the same moving-average window.~~

3.3 Estimation of uncertainty

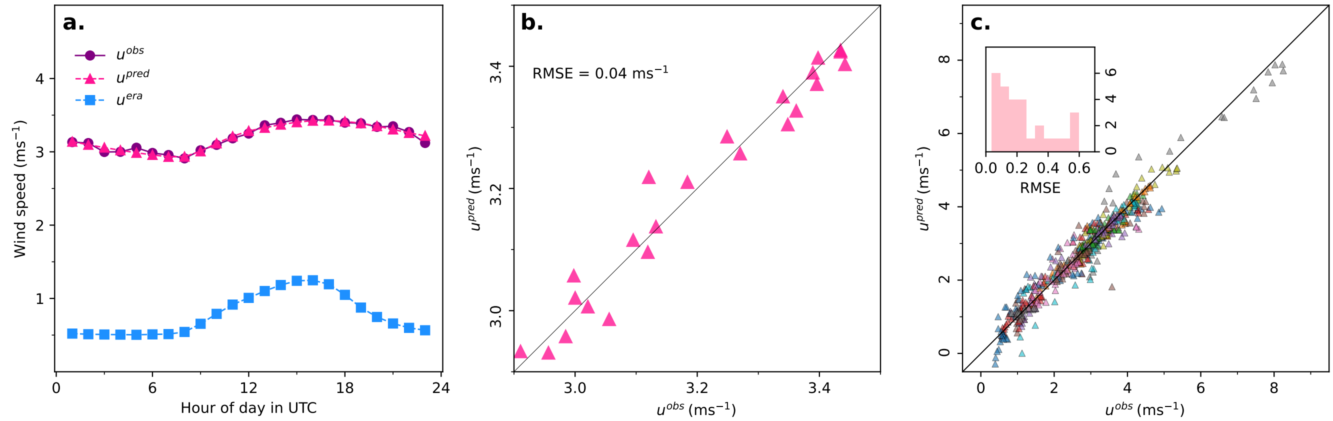

Figure 4. (a) Predicted mean summertime diurnal wind speed (red solid triangle) obtained by fitting Eq. 1 to ~~observations~~ the observed wind speed (purple solid circle) and ERA5L temperature for a selected weather station (Place Glacier in British Columbia, see Supplementary table S1). (b) Comparison of the predicted and observed mean wind speeds for this station. (c) Comparison of the predicted and observed mean wind speed for all 28 station, with different symbol colours denoting different stations. The inset shows the distribution of RMSEs obtained in the fits.

The uncertainty in the predicted windspeed was estimated via a Monte-Carlo simulation with 1000 iterations. In each iteration, the triplet of model parameters were drawn from a corresponding Gaussian distribution. The mean of the Gaussian represented the best model listed later in Eq 10–12. The standard deviations of the Gaussian distributions were chosen either from the fit errors of the linear regressions (Eq. 7–9) or from the IQR of the empirical distribution of the corresponding parameter, as applicable (see Sect. 3.1.1 and 4.1.2). A Gaussian noise was also added to the input values of T^{era}, which had zero mean and the standard deviation was equal to the standard error of mean (Fig. 2b).

4 Results

4.1 Model calibration

4.1.1 ~~Fitting Eq. 1 to observations~~Extracting model parameters for the stations

~~Equation~~ The linear-response equation (Eq. 1~~accurately~~) reasonably captured the mean summertime diurnal wind speed across all the weather stations. The results of fitting Eq. 1 to the observed wind speed and ERA5L temperature on a selected weather station shown in Fig. 4~~a, b~~. Across the 28 weather stations, the quality of fits varied (Fig. 4c)~~, and the~~. The mean (median) RMSE was 0.24 (0.17) ms^{-1}~~with~~, and an interquartile range of 0.12–0.36 ms^{-1} (Supplementary Table S3). In contrast, a direct comparison between the ERA5L wind-speed product with the observed mean summertime diurnal wind yielded a mean (median) RMSE of ~~1.98 (2.09~~1.99 (2.11) ms^{-1}, with an interquartile range of ~~1.29~~1.31–2.57 ms^{-1}.

The distribution of best-fit model ~~parameters $\bar{u}, s$, and τ~~ parameter triplet $\{\bar{u}, s$, and $\tau\}$ obtained for the 28 stations ~~varied significantly~~ showed a significant variability between stations (Supplementary Fig. S2). The best-fit mean wind speed $\bar{u}$ ranged from 0.96–5.12 $\mathrm{ms^{-1}}$ with a mean (median) value of 2.71 (2.96) $\mathrm{ms^{-1}}$. The best-fit sensitivity s ranged from 0.0–1.04 $\mathrm{ms^{-1}{}^{\circ}C^{-1}}$ with a mean (median) value of ~~0.26~~ 0.27 (0.18) $\mathrm{ms^{-1}{}^{\circ}C^{-1}}$. The best-fit ~~response time~~ time-scale parameter τ ranged between ~~0–6.3 hr~~0–3.4 hrs, with a mean (median) value of ~~1.23~~ 1.07 (0.68) hrs.

4.1.2 ~~Fitting Eqs 2–4 to best-fit~~ Estimating model parameters for ungauged glaciers

~~Following the procedure described in Sect. ??, the three model parameters $\bar{u}, s$, and τ yielded the following~~ The predictor $\sqrt{|\bar{T}|/\varsigma_{10}}$ was found to be the most suitable to estimate model parameter $\bar{u}$ (Fig. **??**a), where $|\bar{T}|$ is the absolute value of the mean summertime temperature $\bar{T}$ from ERA5L (Section 2.2), and ς_{10} is the topographic slope over a 10×10-km^2 rectangular grid around the station. A regression between this predictor and the observed best-fit ~~multilinear relationships:~~

$$\cancel{\bar{u} = (2.5 \pm 0.5) + (0.12 \pm 0.02)AR + (4.5 \pm 0.8) \times 10^{-3} R_1 - (1.5 \pm 0.2) \times 10^{-3} R_5,}$$

$$\cancel{s = (0.13 \pm 0.09) + (2.2 \pm 0.6) \times 10^{-4} \bar{Z}_{10} - (1.7 \pm 0.5) \times 10^{-3} Z_s,}$$

$$\cancel{\tau = (0.73 \pm 0.50) + (1.9 \pm 1.4) S_{0.1}.}$$

values of $\bar{u}$ yielded an R^2 value of 0.9 ($p < 0.001$), confirming a strong relationship. The next best model was $\sqrt{|\bar{T}|/\varsigma_g}$, where ς_g is the mean glacier slope, which had an R^2 value of 0.83 but was not significant within $p < 0.01$ level (Supplementary Fig. S3). For the case of model parameter s, no reliable predictor could be found (Supplementary Text. S2). The predictor $1/\sqrt{|\bar{T}|/\varsigma_{10}}$ had the highest R^2 of 0.57 ($p < 0.001$). However, visual inspection indicated that this apparent performance was likely influenced by an outlier (Fig. **??**b). After excluding the outlier, the R^2 value decreased to 0.32, which was not significant at the $p < 0.01$ level. Therefore, we used the median value of s across the 28 stations as its predictor. For the model parameter τ, none of the candidate predictors performed well on visual inspection (Fig. **??**c). The best-performing predictor, $L_g\sqrt{\varsigma_{10}/|\bar{T}|}$, where L_g is glacier length, was not significant, confirming the absence of a strong relationship. We therefore obtained the following parameterisations:

$$\bar{u} = (0.268 \pm 0.017)\sqrt{\frac{|\bar{T}|}{\varsigma_{10}}}, \tag{10}$$

$$s = 0.18 \pm 0.25, \tag{11}$$

$$\tau = 0.7 \pm 1.4. \tag{12}$$

In the ~~above equations, AR is the ratio of the characteristic transverse and vertical scales of the valley, R_x denotes the x-km relief, $\bar{Z}_x$ denotes the mean elevation within an x-km buffer around the station, Z_s is the elevation of the weather station, and $S_{0.1}$ represents the topographic slope within a 0.1-km buffer around the station (Supplementary Table S2). In calculating the local aspect ratio AR, the vertical scale was defined as the standard deviation of elevation within a 1-km buffer around the~~

~~station, and the horizontal scale was the width of the valley measured at the station location (see Supplementary Table S2).~~ above three equations the uncertainty is the standard error of regression in Eqs. 10. Whereas, for Eqs. 11–12, it is the standard deviation of the Gaussian distribution with the same inter-quartile range as the 'observed' distributions of corresponding set of 28 best-fit model parameters as described in Sect. 3.1.1.

The mean wind speed $\bar{u}^{pred}$ predicted using Eq. ~~7~~ 10 compared favourably with the corresponding observed values, with a mean RMSE of ~~0.44~~ 0.57 ms^{-1} (Fig. **??**~~b~~a). In comparison, the corresponding mean RMSE for the ERA5L product was ~~1.88~~ 1.87 ms^{-1}, which was ~~4.2~~ 3.3 times higher. ~~In fact, ERA5L products showed a relatively large negative bias, with a mean bias of -1.7 ms^{-1}. Note that while for 27 out of the 28 stations AR <40, Kennikot Glacier is an outlier with AR of 210. This led to an unphysically large mean wind speed (> 24 ms^{-1}) on this glacier while using Eq. 7. This suggests that the linear relationship between AR and $\bar{u}$ breaks down at such large AR. To avoid this issue, we apply an ad-hoc cut-off on the contribution of the second term of Eq. 7 for Kennikot Glacier, corresponding to a value $AR = 40$. Furthermore, winds at Langenferner Glacier was found to be almost out of phase with temperature (Supplementary Fig. S1), gives rise to unrealistic response times. Therefore, s and τ from this station was left out when fitting Eqs. 3–4 to the 28 triplets of $\{\bar{u}, s, \tau\}$ in section 3.2.~~

~~The hourly anomaly of the mean summertime~~ The diurnal wind speed anomaly u_d^{pred} predicted using Eqs. ~~1, 8–9~~1, 11–12, matched the observations u_d^{obs} reasonably well with an ~~average RMSE of 0.45~~ mean RMSE of 0.49 ms^{-1}~~, although there was some scatter (Fig. 6c).~~. In comparison, ERA5L predicted the diurnal anomaly with an ~~average RMSE of 0.47~~ mean RMSE of 0.46 ms^{-1}~~. While these average RMSEs were comparable, the present model prediction captured the diurnal range of wind speed better, which were systematically underestimated by ERA5L (Fig. 6c).~~, which was comparable to our model.

~~(a-b)The best-fit parameters sensitivity s and response time τ are plotted against their most important static predictors, namely, mean elevation within a 10-km buffer ($\bar{Z}_{10}$), and surface slope within a 0.1-km buffer ($S_{0.1}$). (c) A comparison of the hourly diurnal wind speed anomaly u_d predicted using the model (red solid triangles) and that from the observations. The corresponding comparison for the ERA5L products are also shown (blue solid circles).~~

4.2 Model performance

4.2.1 Performance of calibrated model

The calibrated model (Eqs. 1, ~~7–9~~7–9) was able to capture the mean summertime diurnal wind speed better than the corresponding ERA5L product for all the weather stations (Fig. ~~7~~6). For example, the RMSE of the predicted wind speed had a mean (median) value of ~~0.69 (0.6~~0.87 (0.64) ms^{-1}, and an interquartile range of ~~0.47~~0.42–0~~.87~~.88 ms^{-1}~~(Fig. 7)~~. In comparison, the ERA5L wind speed product had ~~about three~~ two times higher mean ~~(median) of 1.99 (~~of 1.98 ms^{-1}, three times higher median of 2.09 ~~)~~ ms^{-1}, and a two times larger interquartile range of 1.29–2.57 ms^{-1} (Supplementary Table S3).

4.2.2 Leave-One-Out-Cross-Validation

In general, the result of the LOOCV was satisfactory, and the model fitted to data from 27 stations was able to predict the mean diurnal wind speed at the left-out station reasonably well (Supplementary Fig. S4 and Supplementary Table. S3). In comparison

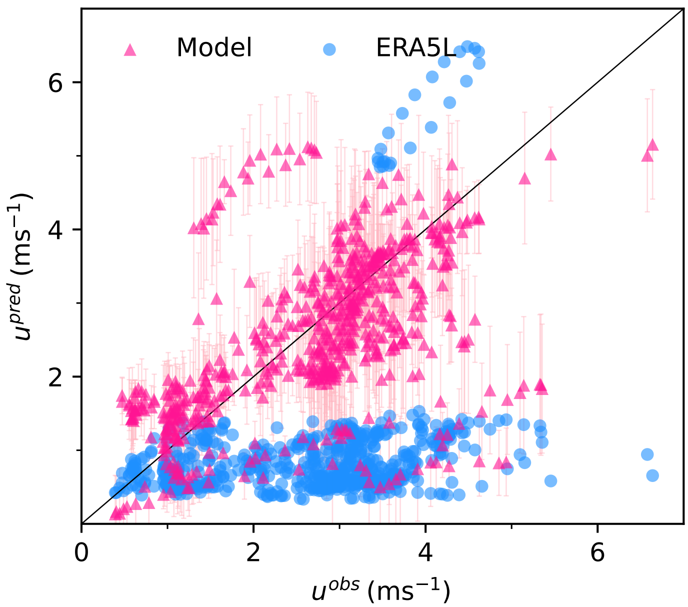

Figure 6. ~~(a)~~ The mean summertime diurnal wind speed predicted using Eqs. 1, 7–9 (red solid triangle) were compared with corresponding observation. The corresponding comparison for ERA5L wind speed are also shown (blue solid circles). The inset shows the distribution of RMSEs for 28 stations for both our model (shown as light pink histogram) and ERA5 (shown as light blue histogram).

to ERA5L, the set of 28 RMSEs of the ~~wind-speed~~ wind speed predicted for the left-out station by the corresponding calibrated model had ~~about 3~~ two times lower mean (~~0.76~~ 0.87 ms^{-1}), ~~median (0.65~~ three times lower median (0.64 ms^{-1}), and three times lower interquartile range (~~0.52–1.01~~ 0.42–0.88 ms^{-1}).

4.2.3 Minimum number of stations

~~As long as the number of weather stations available for calibration N was 20 or more, the predictions of the calibrated model remained relatively unaffected (Fig.~~ 8~~). For example, the~~ A general decline in model performance with decreasing N was observed. The mean RMSE for the predictions over all ~~the~~ 28 stations was ~~0.77~~ 0.853 ms^{-1} for ~~$N = 20$~~ $N = 21$, which was within 1~~3~~% of the mean RMSE value of ~~0.69~~ 0.848 ms^{-1} obtained when all ~~$N = 28$~~ N=28 stations were used for calibration. This indicates that as long as the number of weather stations available for calibration N was 21 or more, the calibrated model was reliable (Fig. 7a).

4.2.4 ~~Time scales of prediction~~

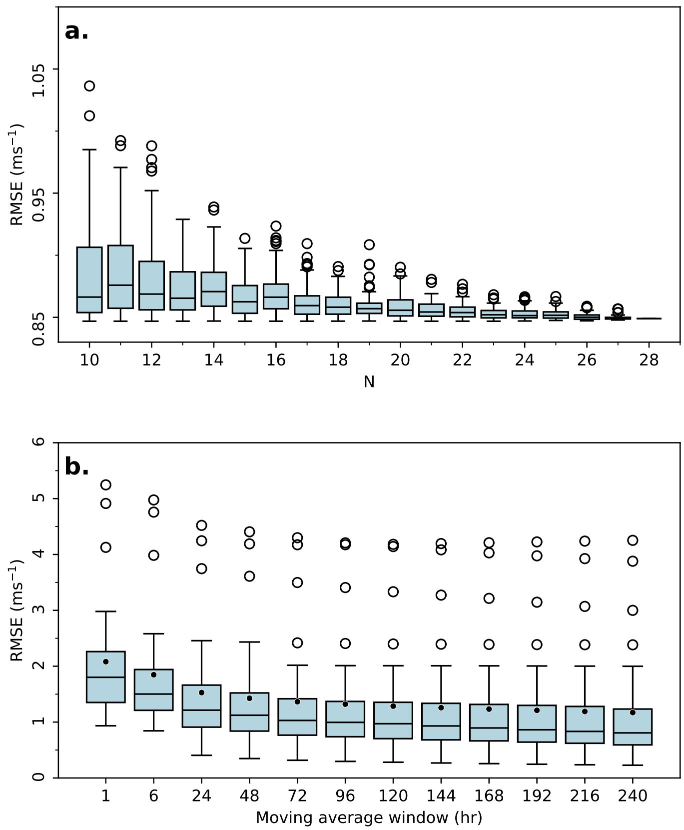

Figure 7. The distribution of RMSEs between predicted and observed wind speed for a) different number of stations N used to calibrate the model, and b) different time-averaging windows. The mean (solid black circle), median (horizontal black line), and outliers (open black circles) of the distributions are shown.

~~While the calibrated model performed better than ERA5L wind speed product across all time-scales, its performance reduced at finer time-resolutions. For moving-average windows of length greater than 96 hr (3 days), the distribution of RMSEs between observed and predicted winds across the 28 stations had a mean value below 1 ms^{-1}. For shorter time-averaging lengths, the mean RMSE increased systemically, with RMSE of $\sim$ 1.7 ms^{-1} for hourly predictions. In comparison, the ERA5L wind speed product had amean RMSE of more than 2 ms^{-1} across all the time-averaging lengths.~~

5 Discussion

5.1 Predicting wind speed on ungauged glacier

The proposed model for the mean summertime diurnal wind speed (Eq. 1), which computed the local wind speed on/around glaciers with ERA5L temperature as input, worked reasonably well for all the 28 weather stations. This may not be unexpected given the existing studies, where thermally driven wind systems such as land-sea breezes and valley-wind circulation have been successfully described as a linear response to diurnal pressure-gradient forcing (e.g., **???**).

The novel result in this study is that the three associated ~~model parameters $\bar{u}, s,$ and τ~~ triplet of model parameters $\{\bar{u}, s,$ and $\tau\}$ could be predicted for any given glacier using the ~~calibrated multilinear~~ relations Eqs. 7–9, without the need for any in-situ data. ~~Compared to the wind-speed predicted directly from the best-fit Eq 1 (Sect. 4.1.1), the wind-speed predicted using the model parameters obtained from the static variable (Eqs. 7–9) were found to be reasonably accurate (Sect. 4.2.1).~~ It is apparent from the LOOCV analysis (Sect. 4.2.2) and the performance of the models calibrated with a smaller subset of stations (Sect. 4.2.3), that the ~~improved performance~~ performance of the present model on ungauged glaciers, ~~as compared to the corresponding ERA5L reanalysis wind-speed products~~ is quite robust. The present model, thus, appears to be a reliable and ~~numerically-efficient~~ parsimonious alternative to the presently available methods and products, ~~which can be used~~ for estimating local wind speed on ungauged glaciers.

Interestingly, the model appears to perform quite well also for the 13 stations that are located near a glacier, and not on it ~~.~~ (Supplementary Fig. S5). In addition, it works well for ~~11~~ five stations located on ~~/near~~ debris-covered glaciers (Supplementary Fig. S1). In all these 16 cases, the station locations are ~~likely to be~~ dominated by valley winds, and not by glacier winds~~(Supplementary Fig. S1).~~ . This suggests that both valley and glacier winds are captured by the same model to a good approximation. ~~Also, it appears that model, which is calibrated for the mean diurnal wind speed over the entire summer season, is still useful at other times scales as well (Sect. 4.2.4). This is because the variability of the mean diurnal wind speed averaged over several days is less than that of the mean wind speed between glaciers (Fig. ??b).~~

5.2 Predictability of model parameters

~~As discussed above that the~~ The strength of the present model, which predicts the mean summertime diurnal wind speed on any glacier with only ERA5L temperature as input, is that the model ~~parameters are known function (Eqs. 7–9) of a set of easy-to-access static topographic variables. For example, the mean wind speed could be predicted entirely based on the local~~

~~relief and the aspect ratio of the valley cross-section. This may imply that in general, the local wind speed on glacierised valleys is strongly controlled by the local topography. In particular, the mean wind speed $\bar{u}$ had strong correlation with local relief (Fig~~parameter $\bar{u}$ can be calculated from mean temperature and regional slope (Eqs. ~~??a). In comparison~~10). However, the predictability of the other two model parameters s and τ was relatively low (Fig. ~~6a, b~~??b,c). In fact, most of the improvements in ~~wind-speed~~ wind speed predicted by the present model, over that obtained from ERA5L, was due to the ability of the present model to capture the variability of the mean wind speed $\bar{u}$ between stations (Sect 4.1.2).

~~While the best multi-linear regressions to obtain the model parameterswere found to be Eqs. 7–9 following the methods discussed in Sect 3.2, there were other multilinear regressions that showed similar performance~~ Perhaps with more data, a larger set of stations, having a higher time resolution, can help improving the predictability of the remaining parameters. ~~See Supplementary Sect. S2 for the two next-best models, which had RMSE values within 1% of that obtained in Eqs 5–7. Although these two multilinear regressions involved different static variables, they were similar in nature in that they were alternative definitions of various aspect ratios, relief metrics, and slope.~~

5.2.1 Time-scales of prediction

We tested if the calibrated model, which was derived for mean summertime diurnal winds, could be used to predict wind speeds across different time-scales. We first filtered the ERA5L hourly temperature data over different time-scales by taking 1-hr to 10-days moving average. For each station and each averaging window, the filtered temperature data were fed into the calibrated model (Eq. 1,10–12) to obtain wind speed predictions for that station and that time-scale. While the calibrated model performed better than ERA5L wind speed product across all time-scales, its performance reduced at finer time-resolutions (Fig. 7b). For moving-average windows of length greater than 96 hours, the distribution of RMSEs between observed and predicted winds across the 28 stations had a median value below 1 ms^{-1}. For shorter time-averaging lengths, the median RMSE increased systemically, with RMSE of $\sim$ 1.8 ms^{-1} for hourly predictions. In comparison, the ERA5L wind speed product had a median RMSE of more than 2 ms^{-1} across all the time-averaging lengths. It appears that model, which is calibrated for the mean diurnal wind speed over the entire summer season, is still useful at shorter times scales as well. This maybe because the variability of the mean diurnal wind speed averaged over several days is less than that of the mean wind speed between glaciers (Fig. **??**b), which the model is able to capture. An analysis of variance of observed and predicted wind speed data across different time-scales (Supplementary Fig. S6) revealed that this improvement in prediction is not simply due to a reduction in variance due to higher smoothing.

5.3 Implication for melt calculations

For a rough estimate of the effects of improved estimates of surface wind speed on ice melt calculations, we used a simple bulk-aerodynamic formulation (e.g., **?**) to compute the hourly sensible (Q_{shf}) and latent (Q_{lhf}) heat fluxes ~~at~~ for the entire data range of all the 28 stations as follows:

$$Q_{shf} = 0.0129\, C^{*} P\, u\, T, \tag{13}$$

$$Q_{lhf} = 22.2\, C^{*}\, u\, (e - e_s). \tag{14}$$

Here, $C^{*} = 0.002$ is a transfer coefficient (**?**), P is the air pressure (in Pa), T is the air temperature (in °C), e is the vapour pressure (in Pa) above the melting surface, and e_s is the vapour pressure (in Pa) at the melting ice surface. The variables P and T were obtained from ERA5. The saturation vapour pressure e_s and actual vapour pressure e were derived from the dew-point temperature obtained from ERA5L ~~(**??**)~~(**?**). The sum of Q_{shf} and Q_{lhf} gave the total turbulent heat flux Q_{thf}. At each station, the turbulent heat flux calculations were repeated thrice for the different estimates of the mean wind speed $\bar{u}$, namely, a) the observed ones, b) those predicted using Eq. ~~7~~10, and c) those obtained directly from ERA5L.

The mean RMSE between the heat fluxes calculated using the observed wind speed ($\bar{u}^{obs}$) and that predicted using Eq. ~~7~~ 10 ($\bar{u}^{pred}$) ~~was 2.2~~ for all stations was 3.5 Wm^{-2} (Fig. ~~**??**~~8), which was about ~~4~~ three times smaller than the mean RMSE of ~~9.8~~ 10.15 Wm^{-2} obtained for the ERA5L derived mean wind speed ($\bar{u}^{era}$). For 17 out of the total 28 stations, where the net Q_{thf} was positive, the heat fluxes increased by a factor of ~~3.4 on average~~ 3.7 on average (Supplementary Table S4). For the remaining 11 stations with net negative turbulent heat fluxes, the magnitude of the fluxes increased by a factor of ~~3.2~~2.3. This suggested that the turbulent heat fluxes on glaciers might be systematically underestimated when ERA5L wind-speed products are used. This is due to ~~a~~ the general underestimation of surface-wind speed by ERA5L at the stations (Fig. ~~**??**b~~6). While Eq. ~~7~~ 10 captures the variability of wind speed between stations, that does not necessarily ensure the corresponding estimates Q_{thf}^{pred} reproduce well the variability of the total turbulent heat fluxes between the stations at a global scale. This is because the temperature and humidity dependent factors in Eq. ~~10,11~~ 13 & 14 can potentially vary strongly between regions. However, for a given region, the spatial variability of turbulent heat flux is likely to be controlled mainly by that of the wind speed. Therefore, the present model can help in improving the turbulent heat flux calculations for a region~~significantly~~, compared to that ~~done using ERA5L-derived~~ derived from ERA5L wind speed. This is also ~~evident from the 4~~ seen from the three times lower mean RMSE between heat fluxes calculated using the model and that using ERA5L-derived winds. ~~It should also be noted~~ We emphasise that the values of turbulent heat fluxes estimated using the bulk aerodynamic method are meant to be illustrative in nature. Although such methods have been used widely and shown to outperform profile-based methods when a wind speed maximum is present (**?**), some recent reports suggest they may perform poorly on glaciers where their underlying physical assumptions may be unrealistic (**???**).

5.4 Applicability of our model

It is important to assess whether the glaciers used in this study are representative of the global glacier population. If not, applying the model to ungauged glaciers may introduce extrapolation errors. The distributions of static glacier variables in our dataset were therefore compared with the global distributions from RGI7 (Fig. 9). Based on this comparison, the model is likely to work better for glaciers with:

a) Mean summertime temperature between −17° to 10°C.

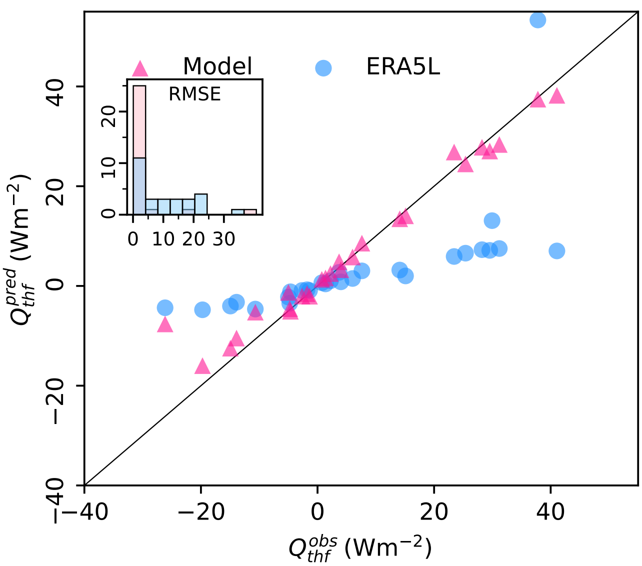

Figure 8. Total turbulent heat flux Q_{thf}^{pred} derived using the mean wind speed predicted in this study (red triangles), and those from ERA5L (blue circles) are compared with the turbulent heat flux Q_{thf}^{obs} computed using observed wind speed. Inset shows the corresponding distributions of RMSEs for the 28 stations.

b) Mean summertime diurnal temperature range between 3° to 10°C.

c) Mean glacier slope $< 30°$.

d) Glacier area > 1 km^2,

e) Glacier length > 2 km,

f) Mean elevation > 2000 m,

Our station distribution is biased toward larger and more gently sloping glaciers, which are generally easier to access. More than 80% of the world's glaciers are smaller than 1 km^2, yet our dataset does not contain a single weather station from glaciers in this size range. Arguably, because such glaciers are typically located near mountain ridges, synoptic wind systems may dominate there rather than thermally driven flows. Also, the small, steep glaciers may not have well developed glacier wind (**?**) . In contrast, high-altitude glaciers in High Mountain Asia, despite being among the most challenging environments to access, are well represented in our dataset.

5.5 Limitations

~~We~~

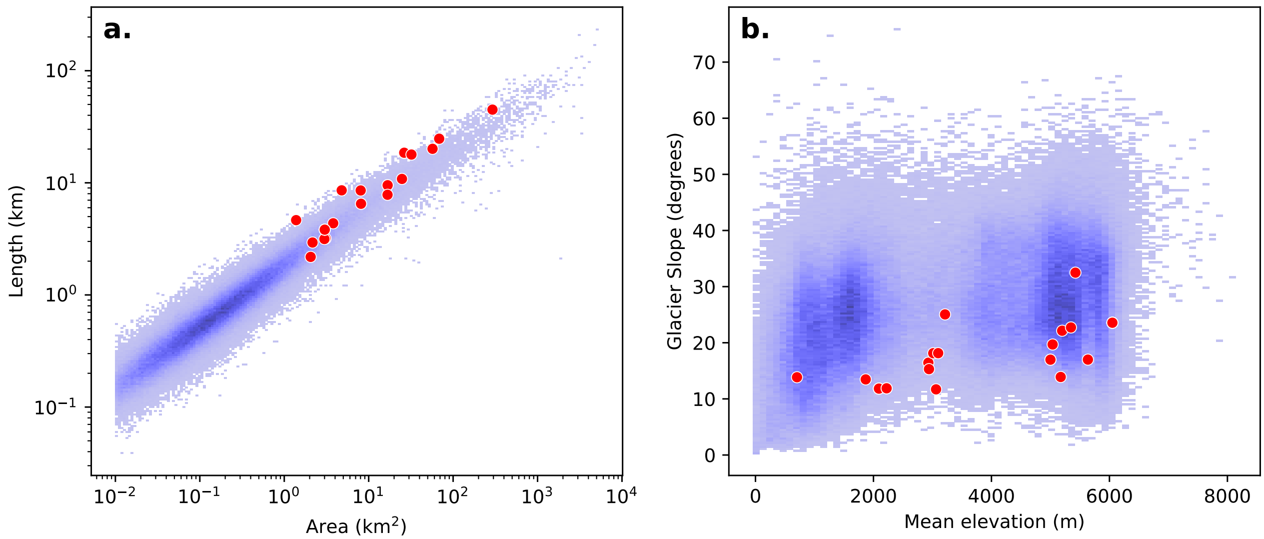

Figure 9. (a) The distribution of glacier length and glacier area for glaciers in this study (red dots) compared with that of the global population of glaciers (blue dots). Similar comparison for mean glacier elevation and glacier slope is shown in (b). Our stations are biased towards larger and gently sloping glaciers, suggesting the possibility of extrapolation errors for smaller steep glaciers.

Ideally, a first-principle derivation for our Eq. 1 is required. We were unable to derive our model systematically from existing theories like the Prandtl model (**?**). However, in general, it is difficult to derive linear-response relations for complex problems. We acknowledge that the present model provides only a first-order solution to the problem of estimating the surface winds on or near a glacier. ~~First of all, the~~ The present results pertain to the mean summertime diurnal wind, while the wind speed is expected to vary over different timescales – hourly, daily,and weekly, (e.g., Fig. 2a) to seasonal and annual. However, our analysis over different time windows (Sect.~~4.2.4~~ 5.2.1) suggests that the magnitude of such temporal variability over ~~time scales~~ time-scales of weeks or longer is often not very large compared to the variability of the mean winds between stations.

An inter-annual variation of the mean summertime wind speed was apparent at the stations with more than one year of data (Supplementary Fig. ~~S5~~S7). Such a temporal variability is, of course, not captured by the present model, which produces a time-independent prediction for mean winds at each station based on Eq. ~~6.~~ 10. For example, the mean value of the fractional absolute inter-annual deviation was less than 10% in all but two of the studied locations (Supplementary Fig. ~~S5~~S7), suggesting that the mean wind speed predicted using Eq. ~~7~~ 10 are still useful. More work is needed, possibly utilising a larger set of stations with longer records, and higher time resolution to see if the observed temporal variability could be incorporated into the present model with the help of additional time-dependent meteorological variables.

~~Another related issue with the present model is that the improved predictions of the mean wind speed are entirely based on a few static topographic variables. This implies that it does not capture the climate response of the wind speed. Given that the glacier wind speed is expected to have some temperature response at different time scales (**?**, ; also see Fig. 1a), more work is needed to incorporate the effects of climate change in Eq. 7.~~

A significant spatial variability of the surface wind within any given glacierised valley is expected (e.g., **?**), which ~~may not be captured fully by the present model~~we do not try to capture with this present model. Only few studies report on within-glacier variation of wind speed and they differ in their conclusion (e.g., **???**). No clear spatial pattern for winds within the same valley was present in our dataset (Supplementary Fig. S8). It should be noted, however, that only four out of 18 glaciers in our study had more than one station in the same valley. Hence, we do not attempt to explain within-glacier variation with our model and rather focus on between-glacier variability. Also, in glacierised valleys, there is a competition between the valley and the glacier winds at the surface level, which flow in opposite directions during the daytime. The present formulation is likely to fail over the region near the glacier terminus where these two winds collide. In fact, the present model gives only wind speed and not the wind direction.

There were some peculiarities in the mean summertime diurnal wind speed at a few stations, which could not be captured by the present model. For example, some stations there were additional early-morning maxima of wind speed (Supplementary Fig. S1), apart from the prominent late-afternoon one. This behaviour could not be reproduced with the present temperature-based linear response model, as the mean diurnal temperature decreased monotonically between sunset and sunrise. At Langenferner Glacier, the temperature and wind speed was out of phase (Supplementary Fig. S1), which was unusual. Here Eq. 1 could not be fitted to the data for a finite τ, and the ~~wind-speed~~ wind speed predictions using Eqs. 1, ~~7–9~~ 10–12 produced a large RMSE (Supplementary Table ~~S4~~S3).

Most of the improvement in mean summertime diurnal wind speed predicted using the present model, vis-a-vis the estimates obtained directly from ERA5L, was due to a better estimation of the mean wind speed $\bar{u}^{pred}$ by the former. As discussed above, the mean RMSEs for $\bar{u}^{pred}$ was ~~4.2~~ three times smaller in the present model, compared to that of the corresponding ERA5L product. In addition, the ERA5L had a significant negative bias, which was 1.9 ms^{-1} on average, for $\bar{u}$. Here we acknowledge that the ERA5L-derived 10-m wind speed is not directly comparable to the 2-m values. However, as discussed before, the interpolation of the value to 2-m level is complex. In contrast, both the present model and ERA5L had similar mean RMSEs for the diurnal anomaly u_d, indicating that there is scope for improving the estimates of the model parameters s and τ. As can be seen from the Supplementary Fig. S2, 64% of the stations have a τ less than the time resolution of our data (1 hour). This might explain why it is difficult to parameterise the variability of τ.

6 Conclusion

In this study, we propose a model to estimate the mean summertime diurnal wind speed on any ungauged glacier. Given the values of three model parameters, the model obtains the wind speed as a linear combination of the mean summertime diurnal temperature and its time derivative. The required input temperature data are derived from a reanalysis product (ERA5-Land).

Among the three model parameters, the parameter $\bar{u}$, which is the mean summertime wind speed, can be estimated from mean reanalysis temperature and regional slope. For the sensitivity parameter s and time-scale parameter τ, we prescribe to use the median values obtained from the corresponding 'observed' values from 28 weath er stations. We show that this model captures well the observed mean summertimediurnal wind speed on 28 on- or near-glacier weather stations across the globe. Interestingly, an analysis of ~~the~~ 28 ~~triplets of best-fit model parameters—excluding s and τ from Langenferner Glacier, as discussed in section 4.1.2—established that the parameters can be approximated as multilinear functions of a few static topographic variables~~weather stations established that mean wind speed parameter $\bar{u}$ can be estimated using simple linear functions of temperature and slope. This implies that the present model can be applied to ungauged glaciers, which is encouraging. A leave-one-out-cross-validation indicated that the method indeed works reasonably on ungauged glaciers, and may improve the RMSE between the observed and estimated mean summertime diurnal wind speeds by about ~~300~~200% on average, compared to the corresponding reanalysis estimates. Most of this improvement is, in fact, due to the ability of the present model to capture the variability of the mean summertime wind between glaciers better than the reanalysis product. The number of stations used for calibrating the model is shown to be adequate for robust predictions. We discuss the limitations of the model in capturing the temporal variability of wind speed over different ~~time scales.~~ time-scales. Although winds are expected to have a within-glacier spatial variability, due to the limited sample size and noise in the data, we could not robustly assess this in our model. Using this model to parameterise the mean summertime diurnal wind speed may improve the turbulent heat flux estimates on glacierised valleys, which tend to be underestimated with reanalysis wind speed products. ~~Understanding the physical basis of the predictability of the mean summertime wind speed on valley glaciers based on a set of topographic~~ Improving the estimates of the other two model parameters, and ~~improving the estimates of the other two model parameters emerge as two of the~~ developing a similar model for smaller timescales, and finding the physical mechanism that explains the success of the empirical model emerge the three most interesting open problems.

. The github repository https://github.com/krrishj2000/Mountain_glacier_winds_parameterisation

contains all codes and processed data used in this work.

. KJ and AB designed the study, conceived the model, and wrote the paper. KJ compiled the meteorological data, did the data analysis, and prepared the figures. RS contributed to the theoretical development. HK and MFA provided field data from Drang Drung Glacier. All authors participated in the discussions, and edited the manuscript.

. Authors declare no competing interests.

. We thank Joseph Shea for sharing weather station data from Place and Weart Glaciers. We acknowledge the contribution of H C Nainwal and his team to the field measurements on Satopanth Glacier. We thank the Ev-K2-CNR Committee for making the data from the weather stations around Khumbu Glacier available. We also thank the two anonymous reviewers of the previous version of the manuscript for their constructive comments. KJ is supported by INSPIRE scholarship from the Department of Science and Technology, Government of India.