

Dear Editor,

We sincerely thank the Editor for the careful assessment of our revised manuscript and for the constructive comments. We greatly appreciate the positive evaluation of the scientific content and the opportunity to further improve the clarity and interpretation of the manuscript.

We have carefully considered all five remaining comments and revised the manuscript accordingly. In particular, we have clarified the distinction between observation-constrained evaluation and independent validation, further clarified the model-derived nature of the observation-constrained reference cumulative mass balance, revised the interpretation of turbulent fluxes during the humid heatwave, corrected the description of the climatic conditions represented by our observational period, and moderated the conclusions regarding possible glacier-type differences in aerodynamic roughness length.

A detailed point-by-point response is provided below. The Editor's comments are reproduced first, followed by our answers and the corresponding revisions made in the manuscript.

Specific comments:

1. Calibration and evaluation using the same EC observations

Editor's comment:

Calibration and evaluation using the same EC observations

In Sect. 3.2 (approximately lines 267–282), the empirical parameters of the Ckat and Clog methods are recalibrated against the quality-controlled 30-min EC observations, while the other methods are also observation-constrained through the EC-derived representative roughness length. These observation-constrained configurations are subsequently evaluated against the EC observations in Sect. 4.3. While I think this is a reasonable approach for assessing and constraining the parameterizations for the Dundee Glacier, the distinction between calibration/observation constraint and independent validation should be made clearer.

Please clarify explicitly that the performance statistics presented for these configurations are not based on an independent validation dataset. Correspondingly, please check the manuscript for statements implying independent validation or overly general conclusions about model performance (e.g. statements that a particular method “performs better” or that the schemes show “good performance”) and qualify these where necessary. I do not think that an additional split-sample analysis is required at this stage; rather, the limitations of evaluating the schemes using the same observational record used to constrain them should be transparently acknowledged.

Answer:

Thank you for this important comment. We agree that the previous version did not sufficiently distinguish between observation-constrained evaluation and independent validation.

We have therefore added an explicit clarification in Sect. 3.2 that the same EC observational record was used both to constrain the turbulent-flux parameterizations and to calculate the comparison statistics. Specifically, the C_{kat} and C_{log} empirical parameters were recalibrated against the quality-controlled EC observations, whereas the C_{Rib1} , C_{Rib2} , and $C_{\text{M-O}}$ methods were constrained using the representative momentum roughness length derived from the same EC record. We now explicitly state that the subsequent statistics therefore do not constitute independent validation and should instead be interpreted as an observation-constrained assessment of model–observation agreement for the Dundee Glacier.

We have also revised Sect. 4.3 to remove the previous description of the EC comparison as “validation” and have replaced several statements concerning model “accuracy,” “good performance,” or one method “outperforming” another with more direct descriptions of model–observation agreement or deviations. The Abstract was similarly revised so that the comparison is explicitly framed as an observation-constrained assessment.

For details, please refer to Sect. 3.2, L285–290 of the revised manuscript:

Because the same EC observational record was used both to constrain the parameterizations—through empirical-parameter recalibration for C_{kat} and C_{log} and through the EC-derived representative z_{0m} for C_{Rib1} , C_{Rib2} , and $C_{\text{M-O}}$ —and to calculate the comparison statistics, the subsequent assessment does not constitute an independent validation. The statistics reported in Sect. 4.3 should therefore be interpreted as an observation-constrained assessment of model–observation agreement for the Dundee Glacier.

The opening of Sect. 4.3 (L486–491) was also revised as follows:

Unless otherwise stated, all results presented in this section are based on these recalibrated or observation-constrained model configurations. Model–observation agreement was quantified using the quality-controlled 30-min EC turbulent-flux observations. Because the same EC observational record was also used to recalibrate C_{kat} and C_{log} and to derive the representative momentum roughness length used by the other methods, the statistics presented here do not represent independent validation and

should be interpreted as an observation-constrained evaluation for the Dundee Glacier.

We also revised the corresponding statement in the Abstract(L23-L27):

Within this observation-constrained assessment, the evaluated schemes showed varying degrees of agreement with the EC observations. Considering multiple timescales, the Richardson-number-based approaches showed comparatively balanced agreement with the observations for both H and LE and more closely reproduced the observed diurnal evolution of turbulent fluxes.

2. Interpretation of the observation-constrained cumulative mass balance

Editor's comment:

I appreciate the substantial clarification made regarding the cumulative mass-balance analysis in Sect. 4.4. The use of the term “observation-constrained reference CMB” is appropriate. However, this reference is still a model-derived quantity: EC fluxes are prescribed where available, while gaps are filled using the recalibrated C_{Rib2} configuration. The independent stake measurement provides useful support, but it covers a somewhat different period from the model simulation.

Please make sure that this distinction remains clear throughout the Results, Discussion, figure captions, and Conclusions, and that the reference CMB is not inadvertently described as an observed mass balance. I would also suggest briefly acknowledging the partial dependence between this reference simulation and the recalibrated C_{Rib2} parameterization when interpreting the comparative CMB results. Again, I do not request additional analysis here; this is primarily an issue of transparency in the interpretation.

Answer:

Thank you for this clarification. We agree that the observation-constrained reference CMB remains a model-derived quantity and that its partial dependence on the recalibrated C_{Rib2} configuration should be stated explicitly.

We have therefore revised Sect. 4.4 to clarify that the reference CMB is generated by EBFM using measured meteorological variables and quality-controlled EC turbulent fluxes where available, while gaps in the EC flux record are filled using the recalibrated C_{Rib2} configuration.

In addition, the caption of Fig. 7 and the caption of Table 4 have been revised to distinguish clearly between directly observed quantities and the model-derived observation-constrained reference CMB.

For details, please refer to Sect. 4.4, L627–639 of the revised manuscript:

Under this framework, the coupled model simulated glacier surface temperature, surface height (SH), and cumulative mass balance (CMB). Because continuous stake-based mass-balance observations were unavailable for the exact modelling period, an observation-constrained reference simulation was constructed to provide a continuous reference CMB time series. In this simulation, the measured meteorological variables and quality-controlled EC turbulent fluxes were prescribed as model forcing. Gaps in the EC turbulent-flux record were filled using the recalibrated C_{Rib2} configuration, whose parameter values minimized the discrepancies between the modeled and observed turbulent fluxes. The CMB simulated under this observation-constrained setup was therefore used as the reference CMB for evaluating the paired simulations. Because the recalibrated C_{Rib2} configuration was used to fill gaps in the EC record, this reference CMB is partially dependent on C_{Rib2} and is therefore not fully independent of the recalibrated C_{Rib2} simulation.

The observation-constrained reference simulation was further compared with independent in situ measurements as a consistency check.

The caption of Fig. 7 (L649-656) was revised to read:

Figure 7: Time series of daily mean turbulent heat fluxes, cumulative mass balance (CMB), surface temperature, and surface height over the Dundee Glacier during the study period, based on the original parameter values (Before), recalibrated parameter values (After), and observations or reference data. “Obs” denotes the EC turbulent-flux observations and the in situ surface-temperature and surface-height observations, whereas “Reference” denotes the model-derived observation-constrained CMB generated using EBFM, with EC turbulent fluxes prescribed where available and gaps filled using the recalibrated C_{Rib2} configuration.. The open star in the CMB panels indicates the independently measured stake-based point surface mass balance of -303 mm w.e.

The caption of Table 4 (L658-661) was revised to read:

Table 4: Comparative statistics for EBFM simulations of surface temperature (T_s) and surface height (SH) against the corresponding in situ observations, and cumulative mass balance (CMB) relative to the model-derived observation-constrained reference CMB, using original (before) and recalibrated (after) turbulent-flux parameters.

3. Statement concerning turbulent fluxes during the humid heatwave

Editor’s comment:

The Abstract currently states that the inability of the schemes to reproduce LE during the humid heatwave indicates “an underestimation of the contribution of turbulent fluxes to glacier melt energy” (approximately lines 36–38). I do not think that this conclusion follows directly from the presented results. During the event, most schemes substantially overestimate the magnitude of negative LE relative to the observations and fail to reproduce the occasional reversal of LE, whereas the representation of H is considerably better.

The results therefore clearly demonstrate difficulties in representing latent heat exchange under extreme warm and humid conditions, but they do not necessarily demonstrate a general underestimation of the total turbulent contribution to melt energy. Please revise this statement in the Abstract, and any equivalent statements elsewhere, so that the conclusion more directly reflects the results. For example, the emphasis could be placed on the inability of the existing parameterizations to accurately represent LE, particularly its weakening and occasional sign reversal, during the humid heatwave.

Answer:

We agree with the Editor. The previous statement concerning an underestimation of the contribution of turbulent fluxes to glacier melt energy may have been stronger than could be directly supported by the presented results. We have therefore revised the Abstract and Conclusions so that the interpretation focuses explicitly on the difficulty of the evaluated parameterizations in representing latent heat exchange during the humid heatwave. This revised interpretation is also consistent with the discussion in Sect. 5.3.

For details, please refer to the revised Abstract, L37–40:

However, none of the evaluated schemes fully reproduced the observed weakening and occasional sign reversal of LE during the humid heatwave, highlighting limitations in representing latent heat exchange under extreme warm and humid conditions.

The corresponding statement in the Conclusions (L888-891) was revised to:

Nevertheless, the estimation errors for LE were non-negligible, and the models failed to reproduce the occasional reversal in LE direction, indicating that current parameterizations have difficulty representing latent heat exchange under extreme meteorological conditions.

4. “Mean climate state” and “multi-year averages” in Sect. 5.3

Editor’s comment:

At the beginning of Sect. 5.3 (approximately lines 744–748), the manuscript states that optimization improves simulations under the “mean climate state (e.g., multi-year averages of T_a and precipitation).” This appears inconsistent with the analysis presented here, which is based on observations collected during May–October 2023. Please revise this wording. I think the intended contrast is between the performance during the broader observation period / non-extreme conditions and performance during the specific humid heatwave event, rather than between a multi-year mean climate and an extreme event.

Answer:

Thank you for identifying this omission. We have therefore revised the beginning of Sect. 5.3 so that the comparison is now between the broader range of conditions represented during the observation period and the specific humid heatwave event. No multi-year climatic interpretation is now implied.

For details, please refer to Sect. 5.3, L760–763 of the revised manuscript:

Our analysis above demonstrates that the optimization of turbulent flux parameterizations markedly enhances the accuracy of glacier surface energy and mass balance simulations under the broader range of conditions represented by the observation period. However, the ability of SEB models to simulate turbulent fluxes under extreme events remains uncertain and warrants further investigation

5. Strength of the conclusion regarding glacier-type dependence of aerodynamic roughness length

Editor’s comment:

The Abstract and Conclusions state that the analysis reveals a “clear glacier type dependence” or that glacier type “exerts a pronounced influence” on aerodynamic roughness length. The comparison is interesting and worth retaining, but these formulations seem somewhat stronger than can be supported by the available observations, particularly given the limited number of long-term EC datasets for continental glaciers—a limitation that the manuscript itself appropriately acknowledges. Please consider moderating this conclusion throughout the manuscript. For example, the results could be described as showing that the available observations indicate a broader and generally lower range of aerodynamic roughness lengths for continental glaciers than for maritime glaciers, or as providing evidence for possible systematic differences between glacier types. This would preserve the interesting result without implying that glacier type itself has been established as the controlling cause.

Answer:

Thank you for this important suggestion. We have therefore moderated this interpretation throughout the manuscript. The previous expressions “clear glacier type

dependence” and “glacier type exerts a pronounced influence” have been replaced by wording that emphasizes possible systematic differences among glacier types.

For details, please refer to the revised Abstract, L31–34:

Our findings suggest that, based on observational data available from previous studies, aerodynamic roughness length may exhibit systematic differences among glacier types, with continental glaciers showing a broader range and generally lower values than maritime glaciers.

Sect. 5.1 was revised as follows:

Available observations suggest possible systematic differences in the characteristic magnitude and range of z_{0m} among glacier types.

The corresponding statement in the Conclusions (L866-872) was revised to:

Our comparison of available aerodynamic roughness-length estimates suggests possible systematic differences among glacier types. Continental glaciers appear to exhibit broader reported ranges and generally lower characteristic roughness lengths than maritime glaciers. This pattern further suggests substantial temporal variability in surface roughness conditions over continental glaciers. However, given the still limited availability of long-term eddy covariance observations in such environments, additional multi-season measurements are required to more precisely constrain the magnitude and range of z_{0m} across different glacier types.

We sincerely thank the Editor again for the careful and constructive comments. We believe that these revisions have improved the transparency of the observation-constrained evaluation, clarified the interpretation of the reference cumulative mass balance, and strengthened the consistency between the presented results and the conclusions drawn from them. We hope that the revised manuscript is now suitable for publication.

Sincerely,

Yichan Xu, on behalf of all authors