

**Dear Dr. Thomas Mölg,**

Please find attached our responses to all reviewers' comments and the revised manuscript. We thank both reviewers for their detailed review that helped us to significantly improve the manuscript. Please note that due to substantial reorganization of the text in the revised manuscript (mainly streamlining of methods and results, following the reviewers' comments) we do not attach a draft with the tracked changes. Instead, we summarize the major changes in the revised manuscript that incorporate comments from both referees:

- 1) Restructured (streamlined) the methods and results sections. In particular, all evaluated bulk methods in the study are clustered into two main types: C-methods and methods based on katabatic models. C-methods consist of four subgroups: C\_log, C\_Rib, C\_M-O and C\_SR, whereas the second cluster consists of the C\_kat method and K\_Int method. All the bulk methods depend on mean meteorological variables only (i.e. there is no application of OPEC data in the bulk methods). All other variants of the bulk methods (i.e. those with OPEC-derived variables) are now presented as part of the sensitivity analysis;
- 2) Removed the section on newly derived stability functions (following the comments from Referee #1);
- 3) Improved the discussion on possible spurious self-correlations and removing the results where the self-correlation was present;
- 4) Introduced a proxy variable for the background temperature lapse rate by using the near-surface air temperature observations from the two nearby meteorological stations, at different altitudes and in the glacier vicinity, in addition to the temperature measurements from the glacier station. The usage of these data led us to identify a dependency of parameters in the K\_Int method on the proxy variable (i.e. difference in off-glacier and on-glacier near-surface air temperature). This empirical relationship allowed us to assess the parameters purely from the mean meteorological variables, instead of using the OPEC-derived stability parameter ( $z/L$ ) as was the case in the initial manuscript;
- 5) Quantified the errors in the modelled sensible heat fluxes resulting from the radiative overheating of the temperature sensor;
- 6) All results (sensible and latent heat fluxes) are now evaluated for the cases with 30-min wind speed exceeding 1 m/s (instead of wind speed exceeding 3 m/s as was the case in the initial manuscript); and
- 7) Revised several figures and added a few that explicitly show dependencies among variables:  $z/L$ , Bulk Richardson number, wind speed and air-surface temperature difference.

Please find below our responses to each referee. The responses are given in bold font, while any quotes from the revised manuscript are copied in italics.

We thank you sincerely for handling our manuscript and look forward to your decision.

Best regards,  
Valentina Radic

## Responses to Referee #1's comments:

• Many useful analyses are presented, and the paper should be of interest to researchers modelling the surface energy and mass balance of mountain glaciers. However, several areas need to be addressed if the most important results for this community are to be highlighted.

In particular, more focus needs to be made on methods that can be implemented solely with mean wind, temperature and humidity, otherwise the results would be more suited to a meteorological journal addressing the underlying theoretical framework for the schemes. Several of results discussed contain ambiguities and potential spurious self-correlation (such as the relationship between  $u^*$  and the newly derived stability functions). These need a much fuller discussion elsewhere, and distract from the main thrust of the paper. That being said, I do think the analysis of turbulence data produces some key results (e.g. that  $u^*$  does not relate well to mean wind speed and that  $z/L$  is poorly predicted by bulk models), which could be presented more explicitly. This could take the form of a dedicated comparison of  $z/L$  with that predicted by M-O, correlation with bulk Richardson stability parameter etc. These results would better set the context for the performance of each parameterisation. One other key result of the paper needs to be further highlighted - the good performance of the simple katabatic model compared to the bulk model.

**We appreciate the constructive criticism aimed at tailoring our presentation of results to the glaciological community readership. We thus revised the manuscript by highlighting more explicitly the key findings related to the bulk methods that can be implemented solely with mean wind, temperature and humidity. Also, we now explicitly present the relation among measured  $u^*$ ,  $U$ ,  $z/L$ ,  $H$  and the bulk Richardson stability parameter (see the figures below). As indicated by the referee, we hope that, by explicitly presenting these relations, we improve the interpretation of the bulk methods' performance.**

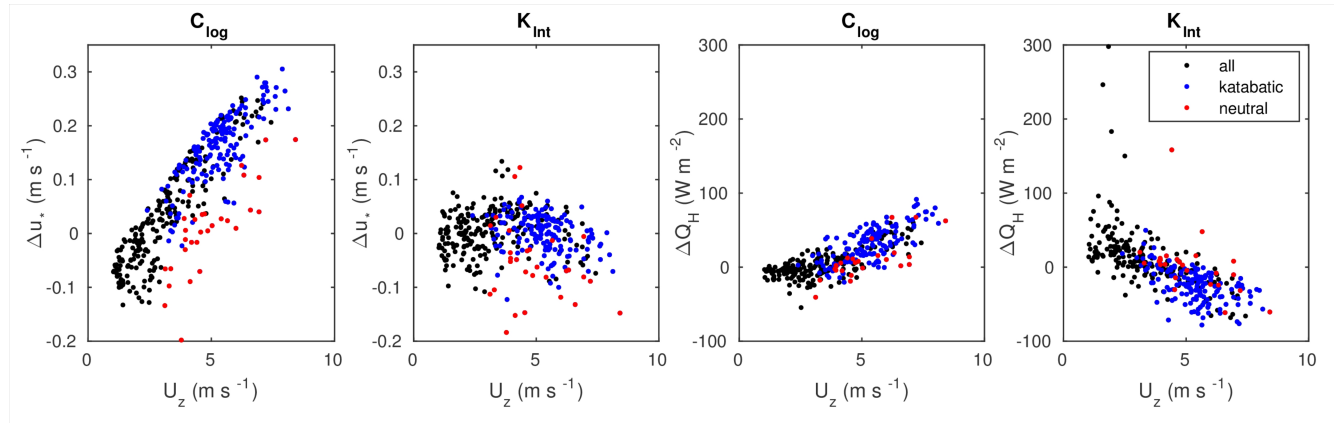


Figure 5: Bias (modeled minus observed values) in  $u^*$  and  $Q_H$  versus the wind speed. Red points show values during the neutral conditions ( $|\frac{z_{v,t}}{L}| < 0.1$ ) only, and blue points indicate values during the prevailing katabatic flow.

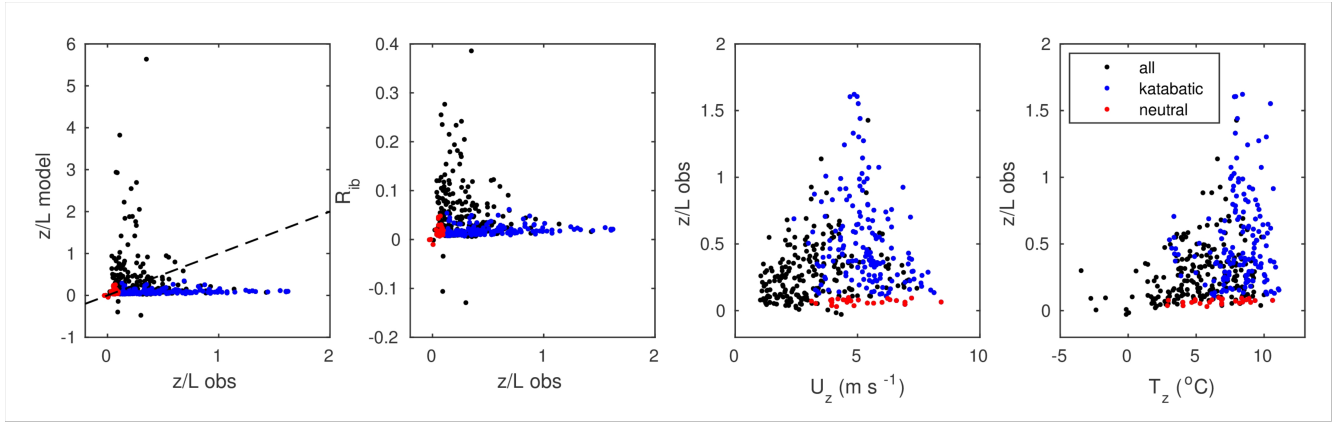


Figure 7: Modeled  $z/L$  with the fixed-point iterative scheme in the  $\mathcal{C}_{M-O}$  method and the Bulk Richardson number ( $R_{ib}$ ) against the OPEC-derived stability parameter ( $z/L$  obs). Dashed black line shows 1:1 line. Also shown is a dependency of the 30-min OPEC-derived stability parameter ( $z/L$  obs) on the wind speed ( $U_z$ ) and the near-surface air temperature ( $T_z$ ).

- The large and somewhat overwhelming quantity of analysis makes the manuscript hard to follow at times. Many of the panels in the later figures (Fig. 9, 10 and 12) deserve to be their own figure as they address a distinct point from other panels. As discussed above, I would prefer to see a dedicated examination of the relationships between  $U$ ,  $u^*$ ,  $H$ ,  $z/L$ , rather than the dispersed results at present. This would help the reader to evaluate different assumptions within the theoretical framework, and point more convincingly toward new theories. At present, there is a danger that readers are confused by the various ways in which the eddy covariance data are used.

**We agree on the overwhelming appearance of the analysis and, following the referee's suggestions, we streamlined the results (e.g. put more focus on explicitly showing the relationships among  $U$ ,  $u^*$ ,  $H$ , and  $z/L$ ) and reorganized the paper to focus on what is most important. We also clustered the bulk methods into two main categories (where only mean variables are used in the method), addressing the remaining bulk methods (with OPEC-derived variables) mainly as a part of a sensitivity analysis.**

- Most of the analyses are clear, appropriate and well supported by theory. The analyses of the new stability function and Kint, however, introduce another layer of theoretical framework that deserves further analysis and discussion. Some of this is beyond the measurements available to the authors (i.e. profile measurements) and could be beyond the scope of the journal. Several aspects of these are quite uncertain and need more discussion. This includes the derivation of  $K_{max}$  and  $H_k$  – does the variation of one, when the other is held constant, reveal the inadequacy of the method? The dependence shown in Fig. 12 (d) and (e) is between  $z/L$  and  $u^*/U$  – and as  $z/L$  contains  $u^*$ , thus there is potential for spurious self-correlation.

**We agree that a more detailed analysis regarding the Kint method in assessing the turbulent fluxes is beyond the scope of this study, mainly because of the limited amount of data on this site. We make a more clear distinction in the revised manuscript between the findings fully supported by our data and the findings which are partly supported but insightful enough to trigger further research when/if the data become available. We now removed the analysis with the self-correlation. Instead, in the revised manuscript, we identified a better way to relate the parameters in  $K_{Int}$  method with the use of near-surface air temperature measurements from**

**the two stations in the glacier's vicinity.**

- In regard to the calculation of new stability function – to properly address this requires profile measurements, otherwise it is simply a circular way to use the measured fluxes to correct the bulk method. It appears that as stability increases, momentum flux decreases while  $H$  does not – this points to the influence of a wind speed maximum, where  $u^*$  goes to zero at close the height of the maximum, which will be lower for lower wind speed (Denby and Greuell, 2000). The authors need to reflect on the sensor heights in regard to flux-profile relationships in presence of katabatic, and how these may be affecting the observed relationships between  $u^*$ ,  $z/L$  and  $H$ .

**We agree that, if one is to derive new stability functions to be used for glacier studies, more measurements (in particular, profile measurements) are needed. Our initial goal was not to derive the new stability functions, but to compare the performance of the optimized stability function (one that assures the best match between modelled and observed fluxes) with the performance of the stability functions commonly used in the glacier studies. However, considering the length of the manuscript and the fact that more observations are needed to adequately analyze the effect of stability, we now removed this section. Instead, we extended the discussion, as suggested by the referee, on the potential influence of a wind speed maximum on the bulk method performance.**

***Page 22/23 in the revised manuscript (copied here in the Latex form):***

*While the original C-methods, however, overestimate  $Q_H$  during the katabatic conditions, the C-methods with stability corrections and measured  $u^*$  underestimate the fluxes (Table \ref{tab:evaluation results} and Fig. \ref{fig: scatter plot with sensitivity tests}). The overestimation of  $Q_H$  during katabatic flows has also been shown in \cite{Denby\_Greuell2000} and explained by a failure of M-O theory in the presence of shallow katabatic wind speed maximum. At the wind speed maximum, measured  $u^*$  approaches zero, while the C-method assumes constant momentum flux in the surface layer and therefore overestimates  $u^*$ . The overestimation is less pronounced for  $Q_H$  than for  $u^*$  because the reduced turbulence at the wind speed maximum leads to an increase in the air-surface temperature difference, and subsequently an increase in the measured  $Q_H$ . However, when measured  $u^*$  is used in the  $\mathrm{C}_{M-O}$  or  $\mathrm{C}_{Rib}$  method, assuming that the eddy diffusivity is as effective as eddy viscosity ( $Pr=1$ ), the C-method underestimates  $Q_H$  since the air-surface temperature difference alone can not compensate for the effect of reduced momentum flux. To correct for this bias in  $Q_H$ ,  $Pr$  would need to decrease, i.e. the C-method would need to account for more effective eddy diffusivity than eddy viscosity at the given height. In the absence of wind profile measurements, we can only assume that these effects take place at our site, but we have no observational evidence for the presence of the wind speed maximum.*

- The authors need to be careful that the key results outlined in the abstract and conclusions are explicitly analysed in the paper. At present, there is some support for the alternate parameterisation schemes presented here, but they depend too much on in-situ turbulence measurements to be used widely. These results are still worth presenting, they just need to be more thoroughly analysed (perhaps elsewhere) before definitive statements can be made. It is good to see movements toward developing new turbulent heat flux parameterisations for mountain glacier environments, which is an essential step for the community.

**A valid point raised by the referee. We streamlined the conclusions to reflect more about our own data and what they show.**

- Specific Comments (page-line):

1-8: “the bulk method” – please clarify what is meant by this term. In general, the terminology used needs clarification. The term K-approach is not likely to be familiar for most readers of The Cryosphere, and is easily confused with the KInt approach introduced later. Consider using a different term here to distinguish the bulk methods in which K scales with  $u^*$ , perhaps “M-O bulk schemes”.

**We now clarify what is meant by the bulk methods in the abstract. The term 'K-approach' is removed, and instead we refer to a more commonly known terminology, i.e. bulk methods based on the gradient transport theory or K-theory (Stull, 1988). We also label these methods in the text as C-methods.**

- 1-12: “The OPEC-derived 30-min momentum flux is linearly related to the measured wind speed, contrary to the proposed quadratic relation by the commonly used bulk methods.” – This result is not shown but rather hinted at (26-4). Needs to be clearly analysed within the paper for this statement to be supported in the abstract.

**Actually, the results have been shown (Figure 10c in the original manuscript) where  $u^*$  is proportional to  $\sqrt{U_z}$ , with a constant of proportionality being a function of the M-O stability parameter. This is equivalent to momentum flux being proportional to  $U_z$ , since the  $u^*$  is defined as the square root of the momentum flux. Nevertheless, since this finding is not the key finding in the study, we removed it from the abstract.**

- 1-15: “In agreement with the katabatic flow model, we show that in a more stable atmosphere the bulk exchange coefficient for momentum is smaller.” Again, the relationship is not analysed specifically, so it is hard to see this a key result. Please revise.

**We added more analysis on this issue in the text (dependency of the parameters in the K\_Int method on the proxy variable for the environmental lapse rate). In the abstract, however, this sentence is now removed since it is not one of the key findings.**

- 1-16: “The sensible heat flux can be more successfully modeled if the bulk exchange coefficients for momentum and heat are allowed to follow different parametrization schemes, rather than assuming equal schemes as is the case in the common bulk methods.” But the data don’t seem to show a large improvement for the more complicated schemes when only mean wind speed and temperature are used. These schemes often rely on measurements of  $z/L$  so aren’t easily transposed in space and time. Please revise.

**The conclusions here follow our findings on the Kint method (section 3.3.4 in the original manuscript): “The Kint -approach, however, does not perform as well for the sensible and latent heat fluxes as it does for the friction velocity. This led us to conclude that while the linear-Gaussian parametrization of  $K(z)$  works well for the momentum flux, it works poorly for the heat flux, indicating a shortcoming in the commonly used assumption that KM and KH can share the same parametrization.” We now improved the analysis on the similarity between eddy viscosity and eddy diffusivity, and provided more discussion on the varying eddy Prandtl number.**

**Page 25 in the revised manuscript (copied here in the Latex form):**

*As already illustrated in Fig.\ref{fig: scatter plots with basic bulk methods}, the  $\mathrm{K}_{\mathrm{Int}}$  method gives the best estimate of the friction velocity and the poorest estimate of  $Q_H$  among all the bulk methods. This performance pattern indicates that  $K_M$  and  $K_H$  do not necessarily share the*

same parametrization. Setting  $Pr=0.75$  for katabatic conditions, the  $\mathrm{K}_{Int}$  method performance is simulating  $Q_H$  improves (Table 3).

- 2-15: References needed here.

**Done. We added Hock (2005) as an appropriate reference here.**

- 2-21: Please consider adding Guo, X., Yang, K., Zhao, L., Yang, W., Li, S., Zhu, M., Yao, T., and Chen, Y.: Critical Evaluation of Scalar Roughness Length Parametrizations Over a Melting Valley Glacier, *Boundary-Layer Meteorology*, 139, 307-332, 2011.

**Included. Thank you for the reference.**

- 3-10: “valley glaciers” – do you mean mountain glaciers? Also, it could be worth consistently referring to mountain or alpine glaciers if the two are to be treated similarly (see 4-26 & 4-27).

**Thank you for pointing out this inconsistency. Yes, we mean mountain glaciers, and have corrected this term accordingly throughout the manuscript.**

- 3-12:  $z_0v$  is a mathematical variable that relates the flux and the gradient, and, as such, is not always related to the turbulence generated by roughness elements at the surface. Outer-layer turbulence can, for instance, increase the momentum flux in the surface layer, thereby changing  $z_0v$ , while the surface roughness elements remain constant. Please revise.

**We agree, but in our case we refer to a turbulence generated by a friction drag. Although the roughness length for momentum is defined as a mathematical concept, it relates to the physical quantity (roughness of surface elements; Stull, 1988) when the turbulence is generated by a friction drag. We revised the text to avoid any ambiguity.**

***Page 3 in the revised manuscript (copied here in the Latex form):***

*While this ‘effective’ roughness length (Braithwaite, 1995) works well as a tuning parameter when the modeled turbulent fluxes are optimized to match the observed ones, it differs from its ‘actual’  $z_0v$  which is dependent only on the geometry and distribution of the roughness elements, assuming turbulence being generated by a friction drag.*

- 5-10: Please provide model numbers of the instruments.

**Added.**

- 6-13: Was the sonic temperature corrected for the effects of water vapour?

**Yes. The sonic temperature was corrected for humidity in the EddyPro software, using a correction based on Schotanus et al. (1983), revised in van Dijk et al. (2004) – this point appears in the revised manuscript.**

- 7-4: WPL corrects for fluctuations in the water vapour density induced by high-frequency changes air temperature, so is associated with turbulent eddies. Please revise.

**Revised. Thank you for this correction.**

- 7-28: The assumption of a melting surface, even on temperate glaciers, does not always hold during the melt season. It would be better to use the SEB, or an air temperature threshold to screen periods in which the surface is likely to be melting to validate the use of this assumption. One period in September 2012 shows air temperature well below freezing, so would almost definitely have lower surface temperature. Also, periods in late August and early September 2012 have air temperature close to 0°C, so the surface temperature is likely to be less than 0°C during these periods. Please discuss further and consider filtering bulk-method results based on periods in which high confidence can be had in the melting surface assumption.

**We agree that it would be better to measure surface temperature directly rather than using this assumption. Unfortunately, we do not have these measurements, while using the SEB closure to derive surface temperature turned out to be unreliable (mainly because the measured radiative fluxes have large errors). In particular, using the NR-Lite net radiometer sensor to estimate the outgoing longwave radiation from the measured net radiation and measured shortwave incoming and reflected radiation, turned out to be erroneous. In the evaluation of bulk methods, we did exclude the 30-min segments for which air temperature was below 1°C (now we clarify this in the text). Note that we also provide an error analysis in the calculated fluxes from the bulk methods assuming random errors in surface temperature.**

**Page 7 in the revised manuscript (copied here in the LaTeX form):**

*In the absence of direct measurements, the surface temperature ( $T_0$ ) was assumed to be at melting point ( $0^\circ\text{C}$ ) and the surface vapor pressure at saturation (6.13 hPa). The assumption of consistent melting is corroborated with the sonic ranger measurements showing persistent surface lowering throughout the observational period. To assure that the assumption holds we use only the data for which  $T_{z_t} > 1^\circ\text{C}$ . In general, assuming that  $T_0 = 0^\circ\text{C}$  works well on temperate glaciers during a melting season, and is more accurate than estimating the surface temperature from the longwave radiation measurements \citep{Fairall\_etal1998} or from a SEB closure \citep{Hock2005}. Nevertheless, when the surface is not consistently melting, SEB closure can give much better results than the assumption of the melting surface (e.g. Conway and Cullen, 2013). Estimating  $T_0$  from our radiation data, proved to be unreliable because of the poor accuracy of NR-Lite net radiometer. As part of our uncertainty analysis, we will quantify errors in our results due to the assumed rather than measured surface conditions.*

- 8-1: Assuming a melting surface in preference to outgoing longwave radiation measurements or SEB closure is predicated on there always being ample energy available for melt. Where this is not the case, SEB closure can give much better results than the assumption of the melting surface (e.g. Conway and Cullen, 2013). I agree that when the surface is most definitely melting, then making the assumption of 0°C is a good way to remove uncertainty in the calculation of surface temperature. Please revise this statement.

**The sentence is now revised to include the point that SEB closure can give much better results than the assumption of the melting surface (e.g. Conway and Cullen, 2013).**

- 8-25: Please refer to the later choice to extend the  $z/L$  range to  $z/L = 2$  (11-13).

**Revised accordingly.**

- 10-5: It would be useful to introduce the filters used to select the 30-min periods used to compare fluxes.

**Revised accordingly.**

- 10-7: The term K-approach is introduced with no background. Is there a less ambiguous term to use for this family of bulk approaches, given that none of the acronyms include K, and you introduce KInt later as a separate method?

**We now label the common bulk method (Clog, C\_M-O, etc) as C-methods. The term 'K-approach' has been removed.**

- 12-7: Why were errors associated with the air temperature not included in the error analysis?

**Our statement was initially incorrect and is now revised to read that the errors in the air-surface temperature differences are used in the error analysis (not just the errors in surface temperature).**

- 15-8: Why were low wind speed periods not included in the analysis? It would be more appropriate to only include the stationarity and wind direction filters, as the other filters are specifically designed for retrieving roughness lengths, rather than removing bad flux estimates. I would expect this to change the comparison significantly, especially the inclusion of low wind speed periods.

**Initially, we included this filter for the consistency to evaluate the bulk methods over the same filtered data used for the derivation of roughness length. In particular, we aimed to include only the 30-min segments with well-developed turbulence ( $u^* > 0.1$  filter) in the presence of relatively strong shearing ( $U_z > 3$  m/s). In the revised manuscript, we now include the 30-min segments with  $U_z > 1$  m/s (but keeping the  $u^* > 0.1$  filter). Note that the additional set of points did not alter our conclusions.**

***Page 9 in the revised manuscript (copied here in the Latex form):***

*To assure that the bulk method evaluation is performed on the high-quality measurements, all of the filters above are applied to the OPEC measured  $u_*$ , sensible heat ( $Q_H$ ) and latent heat ( $Q_E$ ), except the 'neutrality' and 'wind speed' filters. The latter two filters are modified so that that are modified: all runs with  $|\frac{z_{v,t}}{L}| < 2$  are included in the calculation of fluxes, as well as all runs with  $U_z > 1$  m s<sup>-1</sup>. The threshold of  $\frac{z_{v,t}}{L} = 2$  is chosen because the universal stability functions for stable stratification are commonly defined up to  $\frac{z}{L} = 2$  \cite{Foken2008}, which represents strongly stratified stable regime.*

- 23-16: The static stability referred to in the OG model is that of the background airmass. As we have no information on the background stability we don't know if this is necessarily reflected in the stability of the surface layer ( $z/L$ ). Please revise.

**A good point and statement now corrected. We further elaborate on the issue in the results section on the K\_Int method. Note that we altered our analysis of K\_Int dependence on static stability by incorporating the near-surface air temperature measurements from the two stations in the glacier vicinity (AWS\_up and AWS\_low). This new approach now made Kint method independent of the OPEC-derived  $z/L$ .**

- 26-14: Is there a way to evaluate Kint/Kmax without measured  $z/L$ ? For consistency, it would be useful to discuss if this is possible.



Please see the response above. We have modified the way we evaluate  $K_{int}/K_{max}$  in the revised version. Instead of using measured  $z/L$ , we evaluate  $K_{max}$  by its dependency on a proxy variable for background temperature lapse rate, which in our case is a difference between on-glacier and off-glacier near-surface air temperature. This proxy variable is empirically related to the strength of the katabatic flow, therefore, serving as a robust constraint for the  $K_{max}$  parameter.

**Page 14 in the revised manuscript (copied here in the Latex form):**

To adequately determine  $K_{max}$  and  $H_K$ , detailed observations of the wind speed and temperature profiles with height are required. In the absence of these observations, and similarly to  $SC_{kat}$  method, we derive the parameter values through the optimization method that incorporates dependency of the two parameters on the proxy variable for  $\gamma$  (i.e. the difference between on-glacier and off-glacier temperature). Since these two parameters ( $K_{max}$  and  $H_K$ ) are inter-related \citep{Parmhed\_etal2004}, it is sufficient to optimize one parameter while keeping the other constant. The assumed constant value for  $\overline{K_{max}} = 0.8 \text{ m}^2 \text{ s}^{-1}$  (or  $\overline{H_K} = 20 \text{ m}$ ) is taken from its frequently observed values from a field study on an Icelandic glacier \citep{Parmhed\_etal2004} as this was the only study we found that directly measured the two parameters. Keeping one parameter constant, the other is estimated from Eq. \eqref{eq:u\_star\_related\_to\_K\_Int} using 30-min OPEC-derived  $u_*$  and log mean value of OPEC-derived  $z_{0v}$  (from neutral conditions only). In this way we derive optimized  $K_{max}$  (or  $H_K$ ) value for each 30-min segment. The final step is to regress the 30-min optimized parameter values against the proxy variable. The regression is performed on the 2012 data with identified katabatic flow conditions only, while the 2010 data is used to validate the regression model.

- 27-9: Why are those with measured  $z/L$  but not those with measured  $u_*$  included? I think it would be better to only present schemes that do not use any time-varying information from the OPEC system, as these are the parameterisations that are of use to those wishing to use the bulk method.

The methodology has now been modified, so that it addresses the bulk methods that rely on the parametrizations with the mean variables only (i.e. from AWS measurements, and not from OPEC sensor). The OPEC-derived  $z/L$  and  $u_*$  are only applied in the sensitivity analysis in order to identify potential causes of discrepancies in the bulk method performance.

**Page 9 in the revised manuscript (copied here in the Latex form):**

In the gradient-flux relation, the eddy viscosity is parameterized as a function of  $z$ ,  $u_*$  and  $M-O$  stability parameter ( $\frac{z}{L}$ ). Because  $u_*$  and  $\frac{z}{L}$ , in the  $C$ -methods, are modeled rather than directly measured, any error in these modeled values can propagate into the flux estimates. Our goal in this section is to investigate the influence of the two variables,  $u_*$  and  $\frac{z_{v,t}}{L}$ , on the bulk method performance. To do so, we estimate the turbulent fluxes from each of the four bulk schemes using the OPEC-derived  $u_*$  and Obukhov length ( $L$ ).

- 29-15: It is unclear which filtered periods were used here. Please clarify.

We now clearly state which filters are applied on the fluxes used in the inter-comparison. This explanation is given, as suggested by the earlier referee's comment, already in the methods sections so there is no need to reiterate it again in the discussion.

**Page 9 in the revised manuscript (copied here in the Latex form):**

To assure that the bulk method evaluation is performed on the high-quality measurements, all of the filters above are applied to the OPEC measured  $u_*$ , sensible heat ( $Q_H$ ) and latent heat

$(Q_E)$ , except the 'neutrality' and 'wind speed' filters. The latter two filters are modified so that all runs with  $|\frac{z_{v,t}}{L}| < 2$  are included in the calculation of fluxes, as well as all runs with  $U_z > 1 \text{ m s}^{-1}$ . The threshold of  $\frac{z_{v,t}}{L} = 2$  is chosen because the universal stability functions for stable stratification are commonly defined up to  $\frac{z}{L} = 2$  (Foken 2008), which represents strongly stratified stable regime.

- 29-27: There is a need to discuss potential systematic biases in both surface and air temperature and how these could propagate into the calculation of H. This could include additional screen using the SEB to identify melting periods only to compare to OPEC fluxes, and/or a comparison of sonic-temperature with the unventilated air temperature measurements or application of corrections for low wind speed (Huwald et al. 2009) .

**As mentioned earlier, we exclude the 30-min segments that can potentially have surface temperature below 0°C (using the filter for near-surface air temperature to exceed 1°C). In addition to the assessment of errors in the modelled fluxes originating in the random errors in the surface temperature, we now also assess the errors due to the radiative overheating of the unventilated temperature sensor. To do so, we compare sonic-temperature measurements with the measurements from the naturally ventilated sensor, and derive a bias-correction function for temperature measurements.**

**Page 14 in the revised manuscript (Section 2.5 Uncertainty analysis):**

*We quantify uncertainties in the modeled 30-min fluxes due to (1) the choice of roughness lengths, (2) the assumption that the surface temperature is at a melting point, and (3) the systematic error in air temperature due to radiative heating of the temperature sensor.*

- 31-7: “fails to successfully simulate QH.” I would be careful making this statement, as you could argue that it does simulate H fairly well, not just as well as some other, more site-specific schemes. The main failing of the bulk schemes presented here, is the failure to model  $u^*$ . Please revise.

**Revised accordingly.**

- 31-8: “Note that the new stability correction acts in the opposite direction than those commonly used for glacier studies: in our case, the modeled QH needs to increase, rather than be suppressed, as the stability increases.” Yes, but only because  $u^*$  is overestimated. H still decreases relative to stable conditions. Thus, the result is not so much about the effect of atmospheric stability, but the trouble with specifying turbulence in the presence of katabatic flow. Please revise.

**Revised accordingly.**

- 33-4: This paper is not available to the reader at the present time, so it is hard to assess this statement.

**This paper is now published and we cite it accordingly.**

- 33-13 to 16: This is a key result and should feature more highly in the manuscript.

**We agree that this is one of the key findings and now highlight it more clearly.**

- 33-22: “Applying the KInt approach to assess  $u^*$ , which is then used in the K-approach with the newM-O stability function to assess QH gives the best performance across all the bulk methods we tested”. Yes, but the fit between  $u^*$ ,  $U_z$  and Kint is informed by measurements of  $z/L$  (which contains  $H$ ) as is the fit between  $u^*$ ,  $z/L$  and  $H$  in the stability function, so it is not surprising that this function works the best. Please discuss the self-correlation and revise.

**This part of the methodology is now modified (see our responses about the K\_Int method and its relation to the proxy variables for the background temperature lapse rate). In the revised version, the parameters in K\_int method are calibrated independently of measured  $z/L$  (Page 14 in the revised manuscript).**

- Figures 7 and 11: Consider including the numbers assigned to each parameterisation above the columns of each figure panel to aid the reader.

**Revised accordingly.**

- Figure 12: The order of parameterisations in panel (a) needs to be consistent with Table 3 – i.e. the first parameterisations introduced at the top. As with Fig 7, needs to have the numbers assigned to the schemes next to the y-axis labels.

**Revised accordingly.**

- Table 1: Please include the units for the roughness lengths in here.

**Done as suggested.**

- Table 3: Consider removing parameterisations 17, 19, 21 and 23 as they are essentially duplicates of 16, 18, 20 and 22. Also consider adding lines between the sub-sets of parameterisations.

**We revised the table according to the new clusters of bulk methods and have excluded the duplicates.**

We sincerely thank Referee #1 for all the editorial comments. All of them have been incorporated in the revised manuscript.

## **Responses to Referee #2's comments:**

- Generally the paper is well written, but I did find it cumbersome to read, because it includes many different parameterisations that are not easily distinguishable in the text. So I suggest that the authors consider a restructuring of the methods to clarify the difference between the model runs they performed and maybe "cluster" the methods that are similar.

**A great suggestion made by the reviewer. In the revised manuscript, the methods have now been clustered into two main types: C-methods and methods that rely on a katabatic model. We significantly revised the structure of the manuscript to make it more streamlined and less 'cumbersome' to read. We now feel that that paper is easier to read following this clustering and general streamlining.**

- My main concern with the paper is that it neglects the very stable conditions by only looking at conditions where wind speed is  $>3\text{m/s}$  or (the moisture/temperature gradients are large enough). I appreciate that the measurement of turbulent fluxes under very stable conditions are harder to obtain because the mean flow is non-stationary and characterised by brief episodes of intermittent turbulence Mahrt [1989]; Beljaars and Holtslag [1991]; Mahrt [1998]; Cheng et al. [2005]. Considering the significant amount of the periods where low wind speeds occur (Figure 4 in the submitted manuscript), those periods should not be neglected when trying to improve the turbulent fluxes parameterisations over glaciers. A number of studies have been dedicated to finding valid flux-profile relationships for very stable conditions, such as are often found over snow and ice surfaces [e.g. Webb, 1970; Kondo et al., 1978; Lettau, 1979; Brutsaert, 1982; Holtslag and deBruin, 1988; Beljaars and Holtslag, 1991; Cheng and Brutsaert, 2005; Grachev et al., 2007] and those studies have also been applied to snow and ice surfaces [Pomeroy et al., 1998; Jordan et al., 1999; Sharan, 2009; Dadic et al., 2011].

**We only use the near-neutral stability criterion in calculating the roughness lengths, while for the comparison of measured versus modeled turbulent fluxes we include all stability conditions that satisfy  $-2 < z/L < 2$ . In the revised manuscript we now also include the conditions for which wind speed exceeds  $1\text{ m/s}$  (instead of the original threshold of  $3\text{ m/s}$ ). We note, however, that the inclusion of these data did not change the results of the bulk method evaluation. One needs to be careful when assuming that the very stable conditions are present during low wind speeds only → sloped glacier surfaces can have very stable conditions ( $z/L > 1$ ) present during high wind speeds (e.g.  $U_z > 5\text{ m/s}$ ), as we now show in Figure 7 of the revised manuscript (see below).**

**We thank the referee for the list of references. We already have quite extensive list of references but we have added now a selection from this recommended list.**

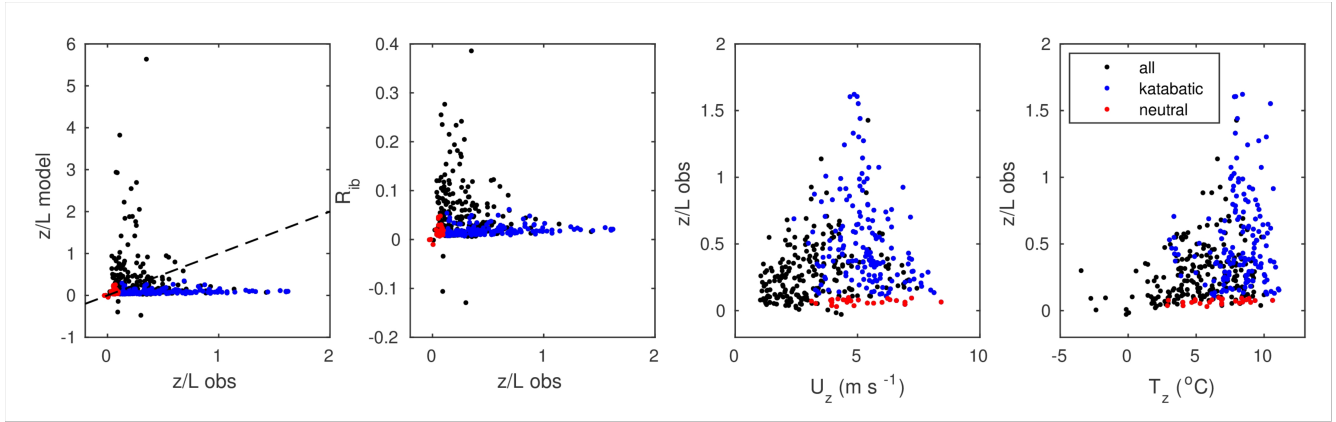


Figure 7: Modeled  $z/L$  with the fixed-point iterative scheme in the  $\mathcal{C}_{M-O}$  method and the Bulk Richardson number ( $R_{ib}$ ) against the OPEC-derived stability parameter ( $z/L$  obs). Dashed black line shows 1:1 line. Also shown is a dependency of the 30-min OPEC-derived stability parameter ( $z/L$  obs) on the wind speed ( $U_z$ ) and the near-surface air temperature ( $T_z$ ).

- All bulk methods assume a logarithmic profile, and they only differ in what stability correction they use. This should be clarified in the manuscript.

This is correct if the stability correction is negligible. However, introducing a stability function into the parametrizations for  $K$  (eddy viscosity) as  $K = k u^* / \phi(z/L)$  does change the logarithmic profile in the bulk method to log-linear profile (under stable conditions). This is why we differentiate between  $C_{log}$  (with logarithmic profile) and  $C_{M-O}$  (with log-linear profile). We prefer to keep this differentiation, and clarify this issue in the revised manuscript.

- Figures 1, 2, and some of Figure 4 (radiation, precipitation, wind direction) are not needed in this paper and can be removed.

We prefer to keep the figures showing the study area and the location of AWS, as well as the photos of the station setup on the glacier. We also prefer to keep radiation and wind direction since we used these variables to help us identify conditions with katabatic flow. Alternatively, we could place these figures in the Supplementary material and will consult ourselves with the editor about whether this is recommended.

- Figure 5: It is of no surprise that pretty much all 4 methods in this Figure have the same results, considering they all use the bulk method at almost neutral conditions. By neglecting the stable conditions, they don't have much reason not to vary. I am therefore not sure what the point of his comparison is.

As noted above, we use all the stability conditions ( $-2 < z/L < 2$ ), not just the near-neutral ones ( $-0.1 < z/L < 0.1$ ). It seems that this was not clearly communicated in the original paper, and therefore we clarified this issue better in the revised manuscript.

**Page 9 in the revised manuscript (copied here in the LaTeX form):**

To assure that the bulk method evaluation is performed on the high-quality measurements, all of the filters above are applied to the OPEC measured  $u_*$ , sensible heat ( $Q_H$ ) and latent heat ( $Q_E$ ), except the 'neutrality' and 'wind speed' filters. The latter two filters are modified so that all runs with  $|\frac{z_{v,t}}{L}| < 2$  are included in the calculation of fluxes, as well as all runs with  $U_z > 1 \text{ m s}^{-1}$ . The threshold of  $\frac{z_{v,t}}{L} = 2$  is chosen because the universal stability

functions for stable stratification are commonly defined up to  $\frac{z}{L} = 2$  *\citep{Foken2008}*, which represents strongly stratified stable regime.

- P18–19: It is not surprising that the "parameterisations" which use measured  $u^*$  as input lead to an increase in fit with the data.  $u^*$  goes into the Q E equation by the power of 4, it's proportional to Q H . It changes L with the power of 3, so will disproportionally decrease  $z/L$ . Some of this discussion (why  $u^*$  has more influence on the turbulent fluxes calculation than  $z/L$ ) might be easier to understand by just looking at the equations and the relevance of the different parameters.

**A good point. We now show the equations that relate the fluxes to  $u^*$  in the bulk method (Equation 19 and 20 in the revised manuscript). Our goal here was to test how well this relation (between the fluxes and mean variables), expected by the theory, is supported by the data. We have now moved this analysis with measured  $u^*$  and  $z/L$  into the sensitivity analysis section and have improved the discussion.**

***Page 9 in the revised manuscript (copied here in the Latex form):***

*In the gradient-flux relation, the eddy viscosity is parameterized as a function of  $z$ ,  $u^*$  and M-O stability parameter ( $\frac{z}{L}$ ). Because  $u^*$  and  $\frac{z}{L}$ , in the C-methods, are modeled rather than directly measured, any error in these modeled values can propagate into the flux estimates. Our goal in this section is to investigate the influence of the two variables,  $u^*$  and  $\frac{z_{v,t}}{L}$ , on the bulk method performance. To do so, we estimate the turbulent fluxes from each of the four bulk schemes using the OPEC-derived  $u^*$  and Obukhov length ( $L$ ).*

- Furthermore, I the observation on page 18 (L1–3) that the C log and C S R methods are not justified in table 3, where the difference between the  $u^*$  models in the correlation coefficient  $r$  is between 0.94 and 0.95 for Q E and between 0.82 and 0.85 for Q H , which is not exactly significant. I am not sure how to address this problem, but I'm sure the authors can come up with more robust conclusions than that.

**Our discussion about the bulk method performance now reflects the results from all evaluation metrics (RMSE, MBE and correlation coefficient), not just the correlation coefficient. This section has also now been modified and moved to the sensitivity analysis section (Section 3.3.1) and the discussion has been improved, i.e. we first intercompare only the bulk methods with mean meteorological variables, and later introduce the sensitivity tests when OPEC-derived  $u^*$  and  $z/L$  are used.**

- p 29, L1-2: Considering that the authors have most SEB components to actually calculate the surface temperature, and that the surface temperature is an important feedback for the TF, the authors should consider calculating the surface temperature and including it in their calculations using the different parametrizations. It would be interesting what effect the different parameterizations have on surface temperature. I do not expect the authors to change all their results now, but maybe it's worth a discussion in the paper.

**Using the SEB closure to derive surface temperature turned out to be unreliable, mainly because the measured radiative fluxes have large errors, in particular, the use of NR-Lite net radiometer sensor to estimate the outgoing longwave radiation from the measured net radiation and measured shortwave incoming and reflected radiation. We now explain this more clearly in the text. Also, we provide an error analysis in the calculated fluxes from the bulk methods assuming random errors in surface temperature.**

***Page 7 in the revised manuscript (copied here in the Latex form):***

In the absence of direct measurements, the surface temperature ( $T_0$ ) was assumed to be at melting point ( $0^\circ\text{C}$ ) and the surface vapor pressure at saturation (6.13 hPa). The assumption of consistent melting is corroborated with the sonic ranger measurements showing persistent surface lowering throughout the observational period. To assure that the assumption holds we use only the data for which  $T_{z_t} > 1^\circ\text{C}$ . In general, assuming that  $T_0 = 0^\circ\text{C}$  works well on temperate glaciers during a melting season, and is more accurate than estimating the surface temperature from the longwave radiation measurements (Fairall et al. 1998) or from a SEB closure (Hock 2005). Nevertheless, when the surface is not consistently melting, SEB closure can give much better results than the assumption of the melting surface (e.g. Conway and Cullen, 2013). Estimating  $T_0$  from our radiation data, proved to be unreliable because of the poor accuracy of NR-Lite net radiometer. As part of our uncertainty analysis, we will quantify errors in our results due to the assumed rather than measured surface conditions.

- p 30, L26-30: Considering that only near-neutral conditions are used for this study, I am not surprised that the stability corrections show very little difference when modelling the fluxes.

**As already mentioned above, we use the stability conditions  $-2 < z/L < 2$ , not just the near-neutral conditions ( $-0.1 < z/L < 0.1$ ). The main reason why the stability corrections did not significantly alter the fluxes is because the modelled  $z/L$ , calculated via the fixed-point iterative scheme, underestimates the OPEC-derived  $z/L$ .**

**Page 20 in the revised manuscript (copied here in the Latex form):**

Intercomparision only across the C-methods (Fig. \ref{fig: scatter plots with basic bulk methods}) reveals that the performance of  $\mathrm{C}_{\text{M-O}}$  method does not significantly differ from  $\mathrm{C}_{\text{log}}$  method. This is because the M-O stability parameter ( $\frac{z_{v,t}}{L}$ ), calculated with the fixed-point iterative scheme of (Munro 1989), is uncorrelated with the OPEC-derived  $\frac{z_{v,t}}{L}$ , underestimates the stability during katabatic conditions and overestimates it during conditions with low speeds (Fig. \ref{fig: modeled versus observed stability}). The stability corrections that depend on calculated  $\frac{z_{v,t}}{L}$ , therefore have a small effect in modifying the fluxes during the katabatic conditions, while during the non-katabatic conditions with the low wind speeds the fluxes are unnecessarily suppressed. Furthermore, we found no correlation between the 30-min OPEC-derived  $\frac{z_{v,t}}{L}$  and any of the mean meteorological variables (e.g. temperature, wind speed; Fig. \ref{fig: modeled versus observed stability}), which explains the failure of the fixed-point iterative scheme that relies on these dependencies. The poor performance of the stability corrections in  $\mathrm{C}_{\text{Rib}}$  method also follows from the lack of correlation between  $R_{\text{ib}}$  and the OPEC-derived  $\frac{z_{v,t}}{L}$  (Fig. \ref{fig: modeled versus observed stability}).

- P31, L11-13: As far as I remember, the reason why the turbulent fluxes are suppressed in Conway and Cullen (2013) is that they assumed the log-linear relationship to be valid under very stable conditions. The log-linear relations, however, do not allow for significant fluxes to occur at very strong stability [Monin and Yaglom, 1971; Mahrt, 1998; Pleim, 2006] and underestimate the turbulent fluxes over these conditions [e.g. Deardorff, 1968; Webb, 1970; Kondo et al., 1978; Louis, 1979; Hogstrom, 1988; Launiainen, 1995; Mahrt, 1998; Jordan et al., 1999; Stossel et al., 2010].

**Yes, this log-linear relationship explains part of the story in the findings from Conway and Cullen (2013). The other part is related to the presence of low wind maximum height. We have now incorporated this explanation more clearly in our discussion. We thank the referee for the provided references, a selection of which we included in the revised manuscript.**

**Page 22/23 in the revised manuscript (copied here in the Latex form):**

While the original C-methods, however, overestimate  $Q_H$  during the katabatic conditions, the C-methods with stability corrections and measured  $u_*$  underestimate the fluxes (Table \ref{tab:evaluation results} and Fig. \ref{fig: scatter plot with sensitivity tests}). The overestimation of  $Q_H$  during katabatic flows has also been shown in \cite{Denby\_Greuell2000} and explained by a failure of M-O theory in the presence of shallow katabatic wind speed maximum. At the wind speed maximum, measured  $u_*$  approaches zero, while the C-method assumes constant momentum flux in the surface layer and therefore overestimates  $u_*$ . The overestimation is less pronounced for  $Q_H$  than for  $u_*$  because the reduced turbulence at the wind speed maximum leads to an increase in the air-surface temperature difference, and subsequently an increase in the measured  $Q_H$ . However, when measured  $u_*$  is used in the  $\mathrm{C}_{\mathrm{M-O}}$  or  $\mathrm{C}_{\mathrm{Rib}}$  method, assuming that the eddy diffusivity is as effective as eddy viscosity ( $Pr=1$ ), the C-method underestimates  $Q_H$  since the air-surface temperature difference alone can not compensate for the effect of reduced momentum flux. To correct for this bias in  $Q_H$ ,  $Pr$  would need to decrease, i.e. the C-method would need to account for more effective eddy diffusivity than eddy viscosity at the given height. In the absence of wind profile measurements, we can only assume that these effects take place at our site, but we have no observational evidence for the presence of the wind speed maximum.



# EVALUATION OF DIFFERENT METHODS TO MODEL NEAR-SURFACE TURBULENT FLUXES FOR A MOUNTAIN GLACIER IN THE CARIBOO MOUNTAINS, BC, CANADA

Valentina Radić<sup>1</sup>, Brian Menounos<sup>2</sup>, Joseph Shea<sup>3,4</sup>, Noel Fitzpatrick<sup>1</sup>, Mekdes A. Tessema<sup>1</sup>, and Stephen J. Déry<sup>5</sup>

<sup>1</sup>Earth Ocean and Atmospheric Sciences Department (EOAS), The University of British Columbia, Vancouver, Canada

<sup>2</sup>Natural Resources and Environmental Studies Institute and Geography Program, University of Northern British Columbia, Prince George, Canada

<sup>3</sup>International Centre for Integrated Mountain Development (ICIMOD), Kathmandu, Nepal

<sup>4</sup>Centre for Hydrology, University of Saskatchewan, Saskatoon, Canada

<sup>5</sup>Natural Resources and Environmental Studies Institute and Environmental Science and Engineering Program, University of Northern British Columbia, Prince George, Canada

*Correspondence to:* Valentina Radić (vradic@eoas.ubc.ca)

**Abstract.** As part of surface energy balance models used to simulate glacier melting, choosing parameterizations to adequately estimate turbulent heat fluxes is extremely challenging. This study aims to evaluate a set of four aerodynamic bulk methods

(labeled as C-methods), commonly used to estimate turbulent heat fluxes for a sloped glacier surface, and two less commonly used bulk methods developed from katabatic flow models. The C-methods differ in their parameterizations of the bulk exchange

5 coefficient that relates the fluxes to the near-surface measurements of mean wind speed, air temperature and humidity. The methods' performance in simulating 30-min sensible and latent heat fluxes is evaluated against the measured fluxes from an open path eddy-covariance (OPEC) method. The evaluation is performed at a point scale of a mountain glacier, using one-level meteorological and OPEC observations from multi-day periods in the 2010 and 2012 summer seasons. The analysis of the two independent seasons yielded the same key findings, which include the following: first, the bulk method, with or without  
10 the commonly used Monin-Obukhov (M-O) stability functions, overestimates the turbulent heat fluxes over the observational period, mainly due to a substantial overestimation of the friction velocity. This overestimation is most pronounced during

the katabatic flow conditions, corroborating the previous findings that the M-O theory works poorly in the presence of a low wind speed maximum. Second, the method based on a katabatic flow model (labeled as the  $K_{Int}$  method) outperforms any C-method in simulating the friction velocity, however, the C-methods outperform the  $K_{Int}$  method in simulating the sensible

15 heat fluxes. Third, the best overall performance is given by a hybrid method, which combines the  $K_{Int}$ -approach with the C-method, i.e. parameterizes eddy viscosity differently than eddy diffusivity. An error analysis reveals that the uncertainties in the measured meteorological variables and the roughness lengths produce smaller errors in the modeled fluxes than are the differences between the modeled and observed fluxes. This implies that further advances will require improvement to model theory rather than better measurements of input variables. Further data from different glaciers are needed to investigate any

20 universality of these findings.

## 1 Introduction

Mountain glaciers of British Columbia are experiencing significant mass loss in response to ongoing climate change and are projected to lose most of their current volume by the end of the century (Radić et al., 2014; Clarke et al., 2015). These projections, however, are highly sensitive to the representation of glacier mass balance in the models, in particular to the model parameterizations used to simulate surface melting. All regional to global scale projections of glacier mass changes rely on a relatively simple representation of glacier melting via so-called 'degree-day' or temperature-index' models (Radić and Hock, 2014). Despite relying on empiricism, these models have relatively successfully dealt with the absence of fine-resolution meteorological input data that are, on the other hand, a prerequisite for a good performance of more physics-based surface energy balance (SEB) models. With rapid development of computational resources and availability of global climate models, SEB models will likely be used in future studies of glacier evolution on regional and global scales. To ensure that future models are firmly rooted in theory, new methods are required to test and parameterize turbulent heat fluxes at the glacier surface.

The turbulent sensible and latent heat fluxes are recognized as important components of the SEB over mid-latitude glaciers worldwide (Hock, 2005). When averaged over periods of weeks to months these fluxes are generally less than net radiation. Over daily to hourly time scales, however, sensible and latent energy transfers can exceed net radiation (Hock, 2005). Highest melt rates at these short temporal scales often coincide with times of greatest turbulent heat flux (Hay and Fitzharris, 1988). Direct measurements of turbulent fluxes, by eddy-covariance method, are relatively rare because they require sophisticated instrumentation requiring continuous maintenance making them unsuitable for long-term operational purposes. Few studies exist that have directly measured turbulent fluxes over mountain glaciers and all of these studies are restricted to shorter time intervals, e.g. weeks to couple of months (e.g., Munro, 1989; Forrer and Rotach, 1997; van der Avoird and Duynkerke, 1999; Cullen et al., 2007; Guo et al., 2011; Conway and Cullen, 2013). A more common approach in the studies of SEB on glaciers is to derive the turbulent fluxes through parameterization schemes that utilize the observations of mean meteorological variables such as near-surface air temperature, wind, and relative humidity (e.g. Hock and Holmgren, 2005; Sicart et al., 2005; Mölg et al., 2008; Reijmer and Hock, 2008; Mölg et al., 2009; MacDougall and Flowers, 2011; Sicart et al., 2011; Huintjes et al., 2015; Prinz et al., 2016). The simplest and most widely applied method is the bulk aerodynamic method (bulk method) (e.g., Braithwaite et al., 1998; Oerlemans, 2000). The bulk method is an integrated form of a gradient-flux relation, based on the theoretical work of Prandtl (1934) and Lettau (1934). The bulk method assumes constant vertical fluxes of momentum, heat and humidity in the surface boundary layer, and horizontal homogeneous conditions of the surface.

Despite the common usage of bulk methods in glacier SEB studies, its application is often problematic. Main limitations of the bulk method applied to glacier SEB studies include:

1. **Surface conditions.** The bulk method for sensible heat fluxes relies on the assumption that the difference between near surface air temperature, often measured at 2 m above the surface, and surface temperature adequately represents the near-surface gradient of air temperature. Furthermore, the surface temperature is rarely measured but assumed or modeled, which can introduce a significant error in the estimates of the turbulent heat fluxes (Conway and Cullen, 2013). Similarly, for the latent heat fluxes, the surface vapor pressure is assumed rather than measured.

2. **Roughness lengths of momentum ( $z_{0v}$ ), temperature ( $z_{0t}$ ) and humidity ( $z_{0q}$ ).** To obtain accurate roughness lengths, detailed measurements at a glacier site are needed, either by an eddy-covariance method or by vertical profile measurements. Without them, roughness lengths from other studies are used assuming that the values are universally applicable on similar glacier surfaces (e.g., Greuell and Oerlemans, 1989; Konzelmann and Braithwaite, 1995). Values of  $z_{0v}$  derived at glacier ice surfaces widely vary in space and time (Bintanja and van den Broeke, 1995; van den Broeke, 1996). There are few directly measured values for the scalar roughness lengths ( $z_{0t}$  and  $z_{0q}$ ) for mountain glaciers, which led to approaches that assume their values approximate  $z_{0v}$ . While this 'effective' roughness length (Braithwaite, 1995) works well as a tuning parameter when the modeled turbulent fluxes are optimized to match the observed ones, it differs from its 'actual'  $z_{0v}$ , which only depends on the geometry and distribution of the roughness elements, **assuming turbulence is generated by frictional drag**. An alternative approach to derive scalar roughness lengths is to relate them to measured  $z_{0v}$  via the surface renewal theory of Andreas (1987). That approach found that  $z_{0t}$  and  $z_{0q}$  over a glacier surface are approximately two orders of magnitude smaller than  $z_{0v}$  in an aerodynamically rough flow (Smeets et al., 1998; Conway and Cullen, 2013), but more studies are needed to investigate whether this finding is invariant in space and time. When the eddy-covariance method is used, accurate roughness lengths can only be confidently derived in near-neutral stability conditions that are rarely present at sloped glacier surfaces, especially during the melt season where stable atmospheric stratification often accompanied by a downslope drainage (katabatic) flow dominates. Better constrained values of roughness lengths are needed, especially considering that order of magnitude changes in  $z_{0v}$  and  $z_{0t}$  alters the modeled turbulent heat fluxes by a factor of two (Munro, 1989; Hock and Holmgren, 1996).
3. **Stability corrections.** Many SEB models applied to snow and ice assume that logarithmic vertical profiles of wind speed, temperature and humidity are valid under prevailing stable conditions and no modification due to the stability effects are necessary (e.g., Munro, 1991; Oerlemans, 2000; Machguth et al., 2006). Experimental and theoretical evidence supports the assertion that Monin-Obukhov (M-O) theory, which requires modification of the fluxes due to the stability effects, is not applicable over sloping glacier surfaces due to violation of the assumptions such as homogeneous, infinite, flat terrain and constant fluxes with height (Denby and Greuell, 2000). When the atmospheric stability is considered in the bulk method, two approaches are commonly used: (a) one that accounts for stability through the bulk Richardson number ( $Ri_b$ ) (e.g., Wagnon et al., 2003; Mölg et al., 2008; Anderson et al., 2010; Gillett and Cullen, 2011) and; (b) another that uses a M-O stability parameter ( $\frac{z}{L}$ ) (e.g., Braithwaite, 1995; Klok et al., 2005; van den Broeke et al., 2005; Hulth et al., 2010). In both approaches, the stability corrections are empirical functions often derived from field studies on a flat and vegetated terrain. Applying the corrections through the two metrics ( $Ri_b$  and  $\frac{z}{L}$ ) can significantly alter the modeled heat fluxes, especially in the presence of large positive near-surface temperature gradients during conditions of low wind speed ( $< 3 \text{ m s}^{-1}$ ) (Conway and Cullen, 2013).
4. **Stable boundary layers accompanied by katabatic flow.** Over a sloping glacier surface, a low-level katabatic jet is a common feature, which means that buoyancy enters the horizontal momentum equation and the turbulence arises independently of the surface roughness (e.g., van der Avoird and Duynkerke, 1999). The Prandtl model for katabatic flow that

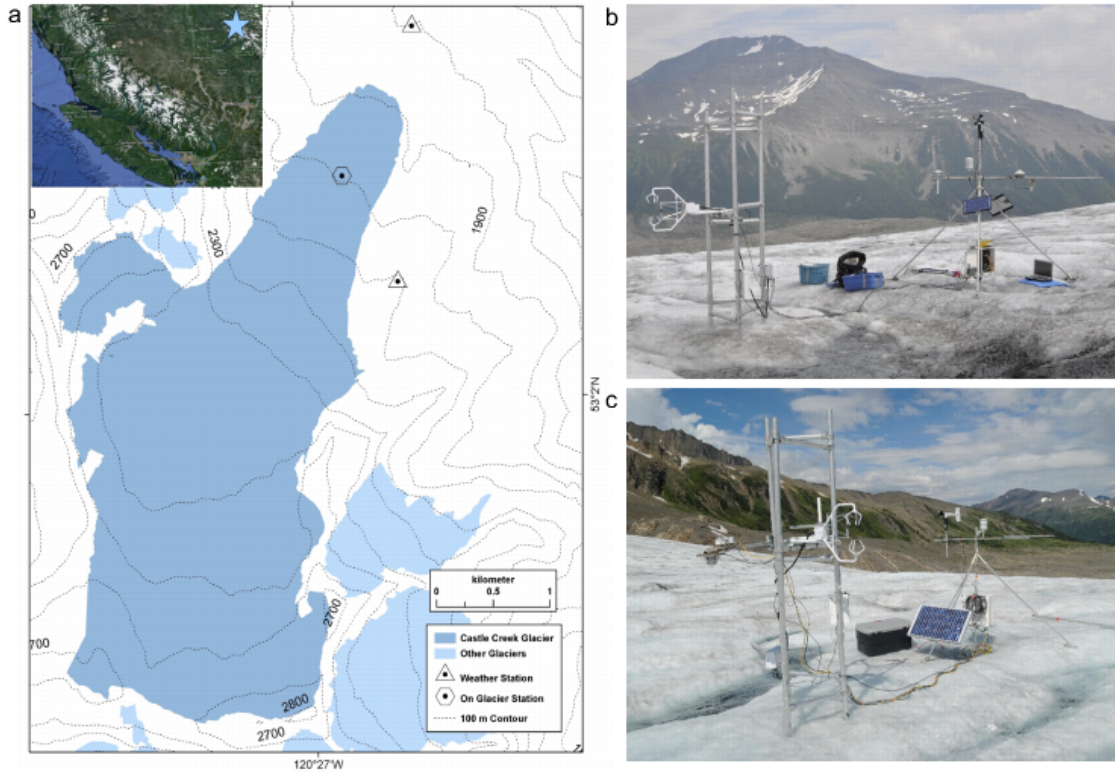
treats eddy viscosity as a constant value (Prandtl, 1942) is unable to correctly describe the sharp near-surface gradients in wind speed and air temperature that are often observed (Munro, 1989; Oerlemans, 1998). Grisogono and Oerlemans (2001) showed that the Prandtl model can be improved if a varying assigned eddy viscosity profile is used instead of a constant value. Using this 'enhanced' Prandtl model, Parmhed et al. (2004) found that modeled turbulent fluxes at a glacier surface compare well with the measured fluxes during katabatic flows. An alternative approach to incorporate the model for katabatic flow into the bulk method for turbulent heat fluxes was developed following the classical Prandtl model for slope flows (Oerlemans and Grisogono, 2002). The simple model only agrees in the overall pattern that shows an increase in sensible heat flux in response to an increase in air-surface temperature difference. Despite its simplicity, application of the model requires calibration with detailed observations of temperature and wind profiles in the surface boundary layer, in addition to the near-surface meteorological observations. Few such measurements exist on glaciers and therefore these models have not been rigorously evaluated so far.

In light of the above, further research is needed to evaluate and develop methods that are suitable for determining near-surface turbulent fluxes over a sloping glacier surfaces subject to katabatic flows. In particular, our objective is to evaluate the existing framework of bulk approaches commonly used for parameterizing the turbulent momentum and heat fluxes at a glacier surface. We achieve this by comparing the modeled fluxes, i.e. output of the bulk approaches, to measured fluxes, i.e. the fluxes derived from an open-path eddy-covariance (OPEC) method. OPEC measurements are obtained from a mountain glacier in British Columbia (BC), over a short-term window (weeks) for two summer seasons. We first provide a brief overview of the field site, data collection, eddy-covariance data treatment and overall methodology. This is followed by a detailed description of the bulk approaches and their evaluation results, and finalized with the discussion and conclusions.

## 2 Data and Methods

### 2.1 Study site and station setup

Castle Creek Glacier lies in the Cariboo Mountains, BC (Fig. 1). The 9.5 km<sup>2</sup> glacier flows north for 5.9 km, has an elevation range of 2827-1810 m a.s.l. and contributes meltwater to Castle Creek, a tributary of the Fraser River. The glacier's annual mass balance has been monitored since 2009 (Beedle et al., 2014). We utilized the data from an automatic weather station on the glacier (AWS<sub>glac</sub>): in summer 2010, the AWS<sub>glac</sub> was installed at 53° 3' 2.99" N and 120° 26' 39.57" W and an altitude of 1967 m a.s.l.; in summer 2012 we installed the AWS<sub>glac</sub> to within ± 10 m of its 2010 position. Two year-round off-glacier automatic weather stations have been in operation since 2007/2008 (Déry et al., 2010). The stations are situated near the glacier's transient snowline and in the glacier forefield, and are herein referred to as AWS<sub>up</sub> (53° 2' 36" N, 120° 26' 18" W, 2105 m a.s.l.) and AWS<sub>low</sub> (53° 3' 45" N, 120° 26' 4" W, 1803 m a.s.l.), respectively. At AWS<sub>up</sub>, the mean annual air temperature over 2007-2010 was -2.6°C, while summer (Jun-Aug) mean air temperature was 6.6°C. Over the same period, precipitation during summer (Jun-Aug) averages 94 mm at AWS<sub>up</sub> although rainfall is likely underestimated due to gauge



**Figure 1.** (a) Map of Castle Creek Glacier with locations of automatic weather station on the glacier (AWS<sub>glac</sub>) and two AWS in the glacier vicinity. AWS<sub>glac</sub> with eddy covariance tower and floating meteorological tripod during installation, Castle Creek Glacier, (b) 2010 (photo by T. Mlynowski), (c) 2012 (photo by V. Radić).

undercatch at the exposed ridge site where mean monthly wind speeds often exceed  $5 \text{ m s}^{-1}$  (Déry et al., 2010). The lower part of the glacier ( $< 2100 \text{ m a.s.l.}$ ), where the AWS<sub>glac</sub> is located, is gently sloping with an approximate mean gradient of  $7^\circ$ .

The AWS<sub>glac</sub> recorded data over a 12-day period (1-12 Aug) in 2010, and a 29-day period (21 Aug - 18 Sep) in 2012. A Campbell Scientific CR1000 datalogger scanned meteorological data with a sampling frequency of 10 s and averaged these measurements every 10 min. A Rotronic HC-S3 T/RH sensor measured temperature and relative humidity at a height  $z_t=1.7 \text{ m}$  above the ice surface, and a RM Young 01503 wind anemometer monitored wind speed and direction at  $z_v=2.0 \text{ m}$  (2010 season) and  $z_v=1.9 \text{ m}$  (2012 season). Kipp & Zonen sensors (CMP6, CM3, CGR3, and NR-Lite) sensed radiation fluxes (incoming and reflected shortwave-, incoming longwave-, and net radiation). We also respectively recorded surface air pressure and liquid precipitation at the AWS<sub>up</sub> and AWS<sub>low</sub>, and we linearly interpolated the measured surface air pressure to obtain surface air pressure at AWS<sub>glac</sub>. The meteorological sensors at AWS<sub>glac</sub>, mounted on a floating tripod (Fig. 1), maintained a constant measurement height above the surface over the observational period. We also installed a Campbell Scientific open-path eddy covariance (OPEC) system and sonic ranger (SR50) on a separate tower drilled into the ice about 3 m upglacier of the AWS<sub>glac</sub> (Fig. 1) (Jarosch et al., 2011). The OPEC system included a sonic anemometer (CSAT-3) and a krypton hygrometer (KH20)

mounted with an initial height of 1.86 m (2010 season) and 2.0 m (2012 season) above the surface. The datalogger recorded three-dimensional (3-D) wind speed, sonic temperature (corrected for water vapor according to Scotanus et al. (1983)) and vapor density fluctuations at 20 Hz. Vapor density fluctuations were later converted to specific humidity fluctuations.

## 2.2 Eddy-covariance data processing

- 5 We applied a series of data processing steps Burba (2013) that are part of the Eddy-Pro Software used to process the OPEC data and calculate the turbulent fluxes. The Eddy-Pro processing output yields turbulence statistics and calculated fluxes for each 30-min segment of raw data. The height of OPEC sensors relative to the ice surface changed due to the surface melting throughout the observational period: the net change in 2010 was 0.7 m, whereas this change equaled 0.75 m in 2012. We applied the following corrections to the raw EC data to correct for differences in the height of the sensors above the ice surface, but
- 10 none of them yielded substantial ( $>1\%$ ) differences in 30-min averaged turbulent fluxes. At the beginning of each observational season we leveled the instruments to horizontal and observed negligible tilt at the end of each observational period. For both seasons, we set the CSAT-3 so that the sensor's arm formed a  $90^\circ$  angle to the prevailing down-glacier wind direction to minimize flow distortion due to air flow through the station structure. Before calculating the turbulence statistics, we applied a coordinate rotation to the velocity data to align the streamlines into the mean flow using a planar fit method (Wilczak et al.,
- 15 2001). This method sets the vertical axis of the CSAT-3 perpendicular to a hypothetical plane produced from averaged wind measurements during the observation period. Because the glacier surface changes with time, we applied sensitivity tests to assess whether the alignment of the plane needed adjustment and whether a different coordinate rotation method (e.g. double or triple rotation scheme) produced different output. We noted no substantial changes, i.e. no difference  $>1\%$ , in the 30-min turbulent fluxes during these tests, however. The latent heat flux data were corrected for oxygen absorption by the  $\text{KH}_2\text{O}$
- 20 (Tanner et al., 1993) and for fluctuations in the water vapour density induced by high-frequency changes in air temperature (Webb-Pearman-Leuning correction; Webb et al., 1980). Finally, we applied a series of standardized processing steps consisting of corrections for potentially high and low frequency loss, and sensor separation (Ibrom et al., 2007; Moncrieff et al., 2004; Horst and Lenschow, 2009).

## 2.3 The algorithm and quality control for roughness lengths

- 25 The covariances of the 3-D wind velocities, temperature and specific humidity, derived by the OPEC system for each 30-min data segment, relate to the friction velocity ( $u_*$ ), temperature scale ( $\theta_*$ ) and specific humidity scale ( $q_*$ ) using:

$$u_* = \left( \overline{u'w'^2} + \overline{v'w'^2} \right)^{\frac{1}{4}}, \quad (1)$$

$$\theta_* = -\frac{\overline{w'T'}}{u_*}, \quad (2)$$

$$q_* = -\frac{\overline{w'q'}}{u_*}, \quad (3)$$

- 30 where  $u'$  and  $v'$  are fluctuations of the horizontal wind component around their 30-min mean values,  $w'$  is fluctuation of the vertical wind component, and  $T'$  and  $q'$  represent temperature and specific humidity fluctuations. We also used the OPEC

system to indirectly measure the Obukhov length ( $L$ ):

$$L = -\frac{T_v u_*^3}{g \kappa \overline{w' T_v'}}, \quad (4)$$

where  $T_v$  is the 30-min averaged virtual temperature (K),  $g$  is the acceleration due to gravity ( $9.81 \text{ m s}^{-2}$ ), and  $\kappa$  (0.40) is the von Kármán constant. Following Cullen et al. (2007), the turbulent scales and  $L$  were then used to derive roughness lengths

5 for each 30-min data segment:

$$z_{0v} = \exp \left[ -\kappa \frac{U_{z_v}}{u_*} - \Psi_v \left( \frac{z_v}{L} \right) \right] z_v, \quad (5)$$

$$z_{0t} = \exp \left[ -\kappa \frac{T_{z_t} - T_0}{\theta_*} - \Psi_t \left( \frac{z_t}{L} \right) \right] z_t, \quad (6)$$

$$z_{0q} = \exp \left[ -\kappa \frac{q_{z_t} - q_0}{q_*} - \Psi_q \left( \frac{z_t}{L} \right) \right] z_t, \quad (7)$$

where  $U_{z_v}$ ,  $T_{z_t}$  and  $q_{z_t}$  are 30-min averages for the wind speed ( $\text{m s}^{-1}$ ), air temperature ( $^{\circ}\text{C}$ ), and specific humidity ( $\text{kg kg}^{-1}$ ) respectively, determined from AWS<sub>glac</sub> measurements at the sensors' heights ( $z_{v,t}$ ). The surface temperature ( $T_0$ ) was assumed to be at melting point ( $0^{\circ}\text{C}$ ) and its vapor pressure at saturation (6.13 hPa) since we did not measure these variables. To reduce errors in our study we only use data for which  $T_{z_t} > 1^{\circ}\text{C}$ . Our  $T_0 = 0^{\circ}\text{C}$  assumption has been found to be more accurate than estimating the surface temperature from the longwave radiation measurements (Fairall et al., 1998) or from a SEB closure (Hock, 2005). Nevertheless, when the surface is not consistently melting, SEB closure can give much better results than the assumption of the melting surface (e.g. Conway and Cullen, 2013). Estimating  $T_0$  from our radiation data, proved to be unreliable because of the poor accuracy of NR-Lite net radiometer. As part of our uncertainty analysis, we quantify errors in our results due to the  $T_0 = 0^{\circ}\text{C}$  assumption. We converted vapor pressure ( $e$ ) to specific humidity using  $q = 0.622 \frac{e}{p}$ , where  $p$  is observed air pressure (hPa).  $\Psi_v(\frac{z_v}{L})$ ,  $\Psi_t(\frac{z_t}{L})$  and  $\Psi_q(\frac{z_t}{L})$  (where  $\Psi_t(\frac{z_t}{L}) = \Psi_q(\frac{z_t}{L})$ ) are integrated forms of universal functions based on M-O theory. Many such functions have been developed, and the most widely used is the Dyer-Businger flux-profile relationship (Businger et al., 1971; Dyer, 1974). Following Conway and Cullen (2013), for stable stratification ( $\frac{z_{v,t}}{L} > 0$ ) we use the expressions of Holtslag and de Bruin (1988) in the following forms:

$$-\Psi_v \left( \frac{z_v}{L} \right) = \frac{z_v}{L} + b \left( \frac{z_v}{L} - \frac{c}{d} \right) \exp \left( -d \frac{z_v}{L} \right) + \frac{bc}{d}, \quad (8)$$

$$-\Psi_{t,q} \left( \frac{z_t}{L} \right) = \left( a + b \frac{z_t}{L} \right)^{\frac{1}{b}} + b \left( \frac{z_t}{L} - \frac{c}{d} \right) \exp \left( -d \frac{z_t}{L} \right) + \frac{bc}{d} - a, \quad (9)$$

where  $a = 1$ ,  $b = \frac{2}{3}$ ,  $c = 5$ , and  $d = 0.35$ . For unstable stratification ( $\frac{z_{v,t}}{L} < 0$ ), which was rarely present during our observational periods, we use the expressions of Dyer (1974).

Turbulence flux data from OPEC systems often contains spurious measurements. The roughness lengths derived from Eq. (5) - (7) widely vary because they rely on several mean and turbulence variables. One way to reduce this scatter is to limit analysis on high-quality data (Andreas et al., 2010). In this study we applied a series of filters to the 30-min data segments (points) obtained during the study period following the approach of Conway and Cullen (2013) and Li et al. (2016). Our filtering algorithm proceeds as follows:



- 5 (a) 'Basic' filter for  $z_{0t}$  and  $z_{0v}$ : Remove the points for which  $\frac{T_{z_t}-T_0}{\theta_*} < 0$  and  $\frac{q_{z_t}-q_0}{q_*} < 0$ , which return unrealistically large roughness lengths since the exponents in Eq. (5) - (7) become much larger than unity. One of the main reasons negative values arise is that the bulk approach (finite differences) is a first-order approximation of the gradient-flux approach, and the surface values for  $T_0$  and  $q_0$  are assumed and kept constant in time rather than measured directly for each 30-min segment.
- (b) 'Stationarity' filter: Only use steady-state runs following the method of Foken (2008), which examines the fluxes for different averaging times. Steady-state conditions can be assumed if 5-min and 30-min averaged fluxes do not differ by more than 30 %.
- 10 (c) 'Neutrality' filter: Select near-neutral conditions ( $|\frac{z_{v,t}}{L}| < 0.1$ ) so that the choice of correction functions ( $\Psi_v, \Psi_t$ ) is not important since  $\Psi_{v,t} \rightarrow 0$  for  $\frac{z_{v,t}}{L} \rightarrow 0$ .
- (d) 'Wind direction' filter: Restrict wind direction to a  $\pm 45^\circ$  sector around the glacier center line to minimize sensor arm interference and ensure the longest on-glacier fetch.
- (e) 'Wind speed' and ' $u_*$ ' filters: Use runs for  $U_{z_v} > 3 \text{ m s}^{-1}$  and  $u_* > 0.1 \text{ m s}^{-1}$  as errors in deriving the roughness lengths become comparatively large for low wind speeds and small friction velocities. In addition, accounting for a relatively strong air flow ( $U_{z_v} > 3 \text{ m s}^{-1}$ ) helps reduce potential errors in  $T_{z_t}$  due to radiative heating of temperature sensor during periods of strong solar radiation (Huwald et al., 2009).
- 15 (f) 'Temperature' gradient' filter for  $z_{0t}$ : To guarantee that a sufficiently large temperature gradient is detected between the surface and measurement height, only use data for which  $T_{z_t} > 1^\circ\text{C}$ . As mentioned earlier, this filter also assures that the assumption of melting surface holds.
- 20 (g) 'Moisture gradient' filter for  $z_{0q}$ : To ensure for sufficiently large moisture gradients between the surface and measurement height, select data that satisfy  $|e_{z_t} - e_0| > 0.66 \text{ hPa}$ , corresponding to a difference between  $T_{z_t} = 0^\circ\text{C}$  and  $T_{z_t} = 1^\circ\text{C}$ .
- (h) 'Small values' filter applied to  $z_{0t}$  and  $z_{0q}$ : Following Andreas et al. (2010) we assumed that the surface exchange of heat and moisture cannot occur at scales smaller than  $10^{-7} \text{ m}$ . This scale is an approximate mean free path of air molecules at sea level. Considering that this is the minimal scale at which molecular diffusion takes place at the surface,  $z_{0t}$  and  $z_{0q}$  smaller than this are treated as unrealistic (i.e. eddy diffusivity can take place only at larger scales).
- 25 (i) 'Large values' filter: No study of turbulent fluxes over glaciers determined roughness lengths larger than 1 m. We thus consider any values of roughness length that exceed 1 m as spurious.

We note that it is common to exclude the data during precipitation events as these greatly increase measurement errors associated with the OPEC system. Nevertheless, this filter was redundant in our study because the above filters already removed the points overlapping with precipitation events recorded at AWS<sub>low</sub>, and because the OPEC system (in particular KH2O) failed

30



to take measurements during most of these events. To assure that the bulk method evaluation is performed on the high-quality measurements, all of the filters above are applied to the OPEC measured  $u_*$ , sensible heat ( $Q_H$ ) and latent heat ( $Q_E$ ) fluxes, except the 'neutrality' and 'wind speed' filters. The latter two filters are modified so that: all runs with  $|\frac{z_{v,t}}{L}| < 2$  are included in the calculation of fluxes, as well as all runs with  $U_z > 1 \text{ m s}^{-1}$ . The upper threshold of  $\frac{z_{v,t}}{L} = 2$  is chosen because the universal stability functions for stable stratification are commonly defined up to  $\frac{z}{L} = 2$  (Foken, 2008), which represents strongly stratified stable regime. The 'measured' 30-min  $Q_H$  and  $Q_E$  are derived from OPEC data as:

$$Q_H = \rho_a c_p \overline{w'T'}, \quad (10)$$

$$Q_E = \rho_a L_v \overline{w'q'}, \quad (11)$$

where  $\rho_a$  is the air density (here OPEC derived),  $c_p$  is the specific heat of air at constant pressure ( $1005 \text{ J kg}^{-1} \text{ K}^{-1}$ ), and  $L_v$  is the latent heat of vaporization ( $2.514 \text{ MJ kg}^{-1}$ ).

## 2.4 Bulk methods

All bulk methods used in this study are rooted in the gradient transport theory or K-theory (Stull, 1988), in which the turbulent fluxes of momentum ( $\tau$ ), sensible heat ( $Q_H$ ) and latent heat ( $Q_E$ ) in the surface boundary layer are proportional to the time averaged gradients of wind speed ( $\overline{U}$ ), potential temperature ( $\overline{\theta}$ ) and specific humidity ( $\overline{q}$ ), expressed by:

$$\tau \equiv \rho_a u_*^2 = \rho_a K_M \frac{\partial \overline{U}}{\partial z}, \quad (12)$$

$$Q_H = \rho_a c_p \frac{1}{Pr} K_M \frac{\partial \overline{\theta}}{\partial z}, \quad (13)$$

$$Q_E = \rho_a L_v \frac{1}{Pr} K_M \frac{\partial \overline{q}}{\partial z}, \quad (14)$$

where  $z$  is the height above the surface,  $K_M$  is the eddy viscosity, and  $Pr$  is the eddy Prandtl number relating the eddy diffusivities for heat ( $K_H$ ) and vapor exchange ( $K_E$ ) to eddy viscosity as  $K_M = Pr K_H$  and  $K_M = Pr K_E$ , respectively. Eq. (12) also defines the friction velocity  $u_*$ , which we use as a surrogate for surface shear stress. We note that by convention in SEB on glaciers, the heat fluxes are positive (negative) if they transport heat towards (away) from the surface.

We provide an overview of the methodology used to compare and analyze the performance of bulk methods in simulating the turbulent fluxes of momentum, heat and humidity. By definition, the bulk method parameterizes the turbulent fluxes with the use of mean meteorological variables (e.g. wind speed, near-surface temperature and relative humidity) and the bulk exchange coefficient. In total we evaluate two groups of bulk methods (i.e. the integrated form of Eq. (12)-(14)): (1) methods that use the bulk exchange coefficient ( $C$ ) as a function of roughness lengths and atmospheric stability, henceforth labeled as C-methods and; (2) methods that incorporate simple katabatic wind models in their parameterizations. As part of the C-methods we analyze four bulk methods that are most commonly used on glacier surfaces: (1)  $C_{\log}$  method - which relates  $C$  only to the roughness lengths assuming logarithmic vertical profiles of wind, temperature and specific humidity, (2)  $C_{\text{Rib}}$  method - which adds stability corrections in  $C$  using the bulk Richardson number ( $Ri_b$ ), (3)  $C_{\text{M-O}}$  - which adds stability corrections in  $C$  using the universal stability functions of  $\frac{z_{v,t}}{L}$ , and (4)  $C_{\text{SR}}$  method - which is the same as  $C_{\text{M-O}}$  in its expression for  $C$ , but uses the

surface renewal model of Andreas (1987) to estimate the scalar roughness lengths ( $z_{0t}$  and  $z_{0q}$ ). If not otherwise stated, the roughness lengths are assumed to be constant in time and equal to the mean log value of the OPEC-derived 30-min roughness lengths observed only during the near-neutral conditions and when all the filters are applied to the OPEC data. The group of bulk methods derived from simple katabatic flow models consists of: (5)  $C_{\text{kat}}$  method - following Oerlemans and Grisogono (2002), and (6)  $K_{\text{Int}}$  method - using an integrated varying eddy viscosity profile (Grisogono and Oerlemans, 2001; Parmhed et al., 2004). Details of each method (1)-(6) are provided in the sections that follow. The modeled 30-min turbulent fluxes ( $u_*$ ,  $Q_H$ ,  $Q_E$ ) are compared to the equivalent OPEC-derived fluxes with the use of standard evaluation metrics: root-mean-square-error (RMSE), mean-bias-error (MBE) and Pearson correlation coefficient (r). The performance of all six methods will be subjected to a set of sensitivity tests to identify key sources of errors. To do so, we will substitute modeled parameter values, such as friction velocity ( $u_*$ ) and stability parameter ( $\frac{z_{v,t}}{L}$ ), with their equivalent OPEC-derived values. We will present results together with the description of sensitivity tests to facilitate readability.

#### 2.4.1 $C_{\text{log}}$ method



According to the mixing-length theory (Prandtl, 1934) for a neutral surface layer, eddy viscosity can be parametrized as (Stull, 1988):

$$K_M = \kappa z u_*. \quad (15)$$

Inserting these parameterizations in the gradient flux relations (Eq. (12) - (14)), treating  $u_*$ ,  $Q_H$ ,  $Q_E$  as constants and integrating the equations with respect to height yield the bulk aerodynamic expressions:

$$u_* = C_v U_{z_v}, \quad (16)$$

$$Q_H = \rho_a c_p \frac{1}{Pr} C_v C_t U_{z_v} (T_{z_t} - T_0), \quad (17)$$

$$Q_E = \frac{0.622}{p} \rho_a L_v \frac{1}{Pr} C_v C_q U_{z_v} (e_{z_t} - e_0). \quad (18)$$

where  $C_v^2 \equiv C_D$  is a dimensionless exchange coefficient ('drag coefficient') for momentum flux and  $C_v C_t \equiv C_H$  and  $C_v C_q \equiv C_E$  are dimensionless exchange coefficients for the sensible and latent heat flux, respectively. Alternatively, one can rewrite Eq. (17) and (18) using  $u_*$  as:

$$Q_H = \rho_a c_p \frac{1}{Pr} C_t u_* (T_{z_t} - T_0), \quad (19)$$

$$Q_E = \frac{0.622}{p} \rho_a L_v \frac{1}{Pr} C_q u_* (e_{z_t} - e_0). \quad (20)$$

Wind speed ( $U_{z_v}$ ), air temperature ( $T_{z_t}$ ) and water vapor pressure ( $e_{z_t}$ ) represent the time averaged values at their measurement heights  $z_{v,t}$  with zero subscript in these variables denoting their surface values (note that  $z_t = z_q$ ). Air density at AWS<sub>glac</sub> ( $\rho_a$ ) is derived as the air density at standard sea-level pressure ( $\rho_0 = 1.29 \text{ kg m}^{-3}$  at  $0^\circ\text{C}$ ) multiplied by the ratio between the air pressure estimated at AWS<sub>glac</sub> ( $p$ ) and the standard sea-level pressure ( $p_0$ ; 1013 hPa). The dimensionless exchange coefficients

for the neutral atmosphere assume logarithmic profile of wind speed, temperature and humidity with height, and take the following form:

$$C_{v,log} = \frac{\kappa}{\ln\left(\frac{z_v}{z_{0v}}\right)}, \quad (21)$$

$$C_{t,log} = \frac{\kappa}{\ln\left(\frac{z_t}{z_{0t}}\right)}, \quad (22)$$

$$5 \quad C_{q,log} = \frac{\kappa}{\ln\left(\frac{z_t}{z_{0q}}\right)}. \quad (23)$$

#### 2.4.2 C<sub>Rib</sub> method

This method assumes the same parameterization for  $K_{M,H,E}$  (Eq. (15)) as the  $C_{log}$  method, but it allows a flux reduction in a stable stratification via the bulk Richardson number ( $Ri_b$ ) defined as:

$$Ri_b = \frac{g (T_{z_t} - T_0)(z_t - z_0)}{T_{z_t} U_{z_v}^2}, \quad (24)$$

- 10 where temperature is given in Kelvin. For stable conditions ( $0 < Ri_b < 0.2$ ), which prevail over a melting glacier surface, the coefficient  $C$  for the property  $y$  (either  $v$ ,  $t$  or  $q$ ) is (Webb, 1970):

$$C_{y,Ri_b} = \frac{\kappa}{\ln\left(\frac{z_y}{z_{0y}}\right)} (1 - 5Ri_b). \quad (25)$$

#### 2.4.3 C<sub>M-O</sub> method

Starting from the parameterization for neutral conditions (Eq. (15)) but allowing  $K_{M,H,E}$  to vary in response to atmospheric

- 15 stability (generally,  $K$  for statically unstable surface layer  $> K$  for neutral  $> K$  for statically stable),  $K_M$  can be expressed as:

$$K_M = \frac{\kappa z u_*}{\psi_v\left(\frac{z}{L}\right)}, \quad (26)$$

where  $\psi_v$  is a dimensionless shear and  $L$  is Obukhov length. Using this approach for  $K$  and integrating the flux-gradient equations, the expression for  $C$  (for the property  $y$ ) becomes:

$$C_{y,M-0} = \frac{\kappa}{\ln\left(\frac{z_y}{z_{0y}}\right) - \Psi_y\left(\frac{z_y}{L}\right)}. \quad (27)$$

- 20 We use the integrated form of M-O stability functions from Holtslag and de Bruin (1988) and Dyer (1974) as expressed earlier in Section 2.3. The calculation of  $\Psi$  requires an estimate of  $L$ , which requires an estimate of  $Q_H$  and  $u_*$ . This method thus becomes a root-finding problem for the three coupled equations, which is either solved by Newton's method or a fixed-point iteration method. We chose the latter since the method is shown to successfully work for this set of equations (e.g., Berkowicz and Prahm, 1982; Lee, 1986) and has been widely used in glacier studies (e.g., Munro, 1989; Conway and Cullen, 2013). In
- 25 this iterative method each 30-min  $Q_H$  and  $u_*$  are initially derived assuming the neutral case ( $\frac{z_{v,t}}{L} = 0$ ). This approach allows an estimate of  $L$  and  $\Psi$ , which in the next iteration yields a new estimate for  $Q_H$  and  $u_*$ . These steps are repeated until no significant change in  $Q_H$  occurs, i.e. the difference becomes  $< \pm 1 \text{ W m}^{-2}$ , a condition which is satisfied within the first five iterations (Munro, 1989).

#### 2.4.4 $C_{SR}$ method



In all the parameterizations for  $C$  we employ a constant value for roughness lengths derived as the mean log value from the 30-min values (Eq. (5) - (7)) calculated for the neutral conditions only. An alternative approach in deriving the scalar roughness lengths ( $z_{0s} \equiv z_{0t,0q}$ ) is through the surface renewal model by Andreas (1987). The model uses simple similarity arguments considering the structure of the viscous sublayer to derive  $z_{0s}$  from known  $z_{0v}$  as a function of Reynolds roughness number ( $Re_* = \frac{u_* z_{0v}}{\nu}$ , where  $\nu$  is the kinematic viscosity of air, equal to  $1.46 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$ ):

$$\ln\left(\frac{z_{0s}}{z_{0v}}\right) = b_0 + b_1 \ln(Re_*) + b_2 [\ln(Re_*)]^2, \quad (28)$$

where  $b_{0,1,2}$  are the polynomial empirical coefficients. This model has been successfully tested for relatively smooth snow and ice surfaces, while for a rough, hummocky glacier surface Smeets and van den Broeke (2008) found that somewhat different values for the polynomial coefficients give better performance.

#### 2.4.5 $C_{kat}$ method



Rather than integrating the flux-gradient relations with a chosen parametrization for  $K$ , Oerlemans and Grisogono (2002) derived a bulk approach for surface heat fluxes through a simplified scaling of the governing equations for heat and momentum balance in a 1-D katabatic flow model (Prandtl, 1942). Their basic assumption was that the katabatic flow is characterized by a well defined wind maximum, setting the exchange coefficient for heat proportional to the maximum wind speed and to the height of the wind maximum. With this scaling approach they determined a 'katabatic bulk exchange parameter' for heat flux ( $C_{kat}$ , in  $\text{m s}^{-1}$ ) as:

$$C_{kat} = -k k_2^2 \Theta_s \left( \frac{g}{T_0 \gamma Pr} \right)^{1/2}, \quad (29)$$

where  $k$  and  $k_2$  are dimensionless empirical constants,  $\Theta_s$  is the temperature deficit at the glacier surface (negative value;  $\theta(0) = \Theta_s$ ,  $\theta(z \rightarrow \infty) = 0$ ), and  $\gamma$  is the background potential temperature lapse rate. Defining the exchange coefficient in this way makes the sensible heat flux increase quadratically with the surface temperature deficit:

$$Q_H = \rho_a c_p C_{kat} \Theta_s. \quad (30)$$

$\Theta_s$  can be replaced by 2 m air-surface temperature difference ( $\Delta T$ ), but it only yields a crude estimate of the actual temperature deficit. The latent heat flux is similarly expressed as:

$$Q_E = \frac{0.622}{p} \rho_a L_v C_{kat} (e_{zt} - e_0). \quad (31)$$

Neither  $Q_H$  nor  $Q_E$  explicitly depend on  $u_*$  because the momentum flux and wind speed are obtained from the katabatic flow model in which the eddy viscosity is parametrized through the background variables ( $\Theta_s$  and  $\gamma$ ).  $C_{kat}$  becomes smaller as the atmospheric boundary layer becomes more stable (larger  $\gamma$ ). As noted in Oerlemans and Grisogono (2002), this is a dynamic effect, which should not be confused with the effect of stratification on the exchange coefficients as predicted by the M-O theory.

Our data are not suitable to adequately test this bulk method since we do not have the vertical profile observations of potential temperature needed to estimate  $\Theta_s$  and  $\gamma$ . Nevertheless, the data are sufficient to investigate the validity of the quadratic relation between  $Q_H$  and  $\Delta T$ . Assuming that the quadratic fit holds, we optimize  $C_{kat}$  for each season independently so that the modeled  $Q_H$  gives the best fit to the observed  $Q_H$ . Although this optimization can be performed in multiple different ways, we chose to set all the parameters constant except the one used for model tuning. The values of constant parameters are taken from the field data from a 10 km long mountain glacier in the Austrian Alps (PASTEX-94; Greuell and Struijk, 1994):  $k_2=1$ ,  $\gamma=0.005 \text{ K m}^{-1}$ ,  $Pr=2$  and  $T_0=273 \text{ K}$ . The chosen optimization parameter  $k$  is then evaluated by minimizing the RMSE between modeled and observed 30-min  $Q_H$  for each season separately, using the air-surface temperature difference ( $T_{z_t} - T_0$ ) as a substitute for  $\Theta_s$ . In addition to this optimization, we investigate whether  $C_{kat}$  can be expressed as an empirical function of a proxy variable for the background potential temperature lapse rate ( $\gamma$ ). We make use of near-surface air temperature measurements from the two stations in the glacier vicinity ( $AWS_{up}$  and  $AWS_{low}$ ) and assume that the proxy variable is the difference between on-glacier ( $AWS_{glac}$ ) and off-glacier temperature. We then regress 30-min  $C_{kat}$  values, obtained from OPEC-derived  $Q_H$  and measured  $\Theta_s$  (Eq. 30), against the proxy variable.

#### 2.4.6 $K_{Int}$ method

Similarly to the  $C_{kat}$  method, this approach assumes a dominant katabatic flow throughout the observational period. Following Grisogono and Oerlemans (2001) and Parmhed et al. (2004) we use the assumed linear-Gaussian profile for eddy viscosity  $K(z)$  as

$$K(z) = \frac{K_{max} \exp(0.5)}{H_K} z \exp \left[ -0.5 \left( \frac{z}{H_K} \right)^2 \right], \quad (32)$$

where  $K_{max}$  is the maximum value of  $K(z)$  (in  $\text{m}^2 \text{s}^{-1}$ ) reached at the height  $H_K$  (in m) above the surface. Near the surface, this parameterization for  $K(z)$  is similar to O'Brien's cubic polynomial approximation applicable to neutral and stable boundary layer (O'Brien, 1970; Stull, 1988). In the original model of Grisogono and Oerlemans (2001),  $K_{max}$  and  $H_K$  are related to the background variables ( $\gamma$ ,  $\Theta_s$ ), slope of the surface, and  $Pr$ . Inserting the parameterizations for  $K(z)$  into the bulk-gradient relations (Eq. (12) - (13)) and integrating them with respect to  $z$  (treating  $K_{max}$  and  $H_K$  independent of  $z$ ), we obtain the following bulk expressions for the turbulent fluxes:

$$u_* = \left( \frac{U_{z_v}}{K_{Int}} \right)^{1/2}, \quad (33)$$

$$Q_H = \rho_a c_p \frac{1}{Pr} \frac{T_{z_t} - T_0}{K_{Int}}, \quad (34)$$

$$Q_E = \frac{0.622}{p} \rho_a L_v \frac{1}{Pr} \frac{e_{z_t} - e_0}{K_{Int}}, \quad (35)$$

where

$$K_{Int} = \int_{z_{0v}}^{z_v} \frac{dz}{K(z)} \approx \frac{H_K}{2 K_{max} \exp(0.5)} \left[ 2 \ln \left( \frac{z_v}{z_{0v}} \right) + \frac{0.5}{H_K^2} (z_v^2 - z_{0v}^2) \right], \quad (36)$$

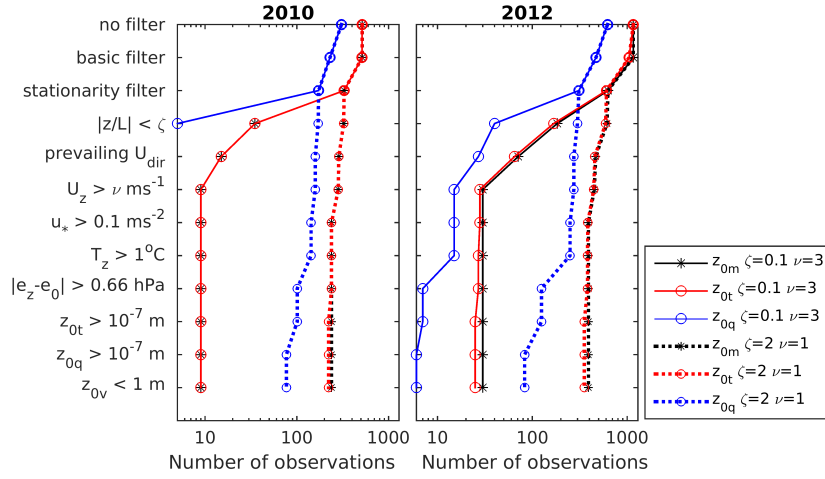
which is valid for  $z_v < H_K$ .

To adequately determine  $K_{max}$  and  $H_K$ , detailed observations of the wind speed and temperature profiles with height are required. In the absence of these observations, and similarly to  $C_{kat}$  method, we derive the parameter values through the optimization method that incorporates dependency of the two parameters on the proxy variable for  $\gamma$  (i.e. the difference between on-glacier and off-glacier temperature). Since these two parameters ( $K_{max}$  and  $H_K$ ) are inter-related (Parmhed et al., 2004), it is sufficient to optimize one parameter while keeping the other constant. We focus therefore only on optimizing  $K_{max}$  while keeping  $H_K$  constant. We set  $\overline{H_K}=20$  m as a mean observed value from a field study on an Icelandic glacier (Parmhed et al., 2004) as this was the only study we found that directly measured the two parameters. Optimized  $K_{max}$  value for each 30-min segment is estimated from Eq. (33) using 30-min OPEC-derived  $u_*$  and mean log value of OPEC-derived  $z_{0v}$  (from neutral conditions only). The final step is to regress the optimized parameter against the proxy variable. The regression is performed on the 2012 data with identified katabatic flow conditions only, while the 2010 data is used to validate the regression model.

## 2.5 Uncertainty analysis

For each 30-min roughness length estimate (Eq. (5) - (7)) we calculate its error through the error propagation method for a multi-variable function (e.g., Bevington, 1969). The measurement error for wind speed ( $\delta U_{z_v}$ ) and assumed errors for air-surface temperature difference ( $\delta \Delta T$ ) and air-surface difference in water vapor pressure ( $\delta \Delta e$ ) are propagated (retaining the linear terms in Taylor expansion) into the final error for the roughness length. The following error values are used:  $\delta U_{z_v}=0.11$  m s<sup>-1</sup>,  $\delta \Delta T=0.5^\circ\text{C}$ , and  $\delta \Delta e=0.23$  hPa.

We quantify uncertainties in the modeled 30-min fluxes due to: (1) the uncertainty in the roughness lengths; (2) the assumption that the surface temperature is at a melting point; and (3) the systematic error in air temperature due to radiative heating of the temperature sensor. To estimate errors in (1), for each modeled 30-min flux, we use a Monte Carlo approach so that the bulk method is run 1000 times with randomly perturbed roughness length values. The roughness length values are randomly picked from a derived normal distribution of their log values, using the OPEC-derived mean log value and standard deviation (applying all the filters from Section 2.3) as the distribution parameters. The error in each 30-min modeled flux is represented by a standard deviation of the 1000 ensemble runs. The same Monte Carlo approach is used to quantify the uncertainty in (2). For each Monte Carlo run, we randomly prescribe  $T_0$ , assuming that  $T_0$  has a normal distribution with a mean of  $0^\circ\text{C}$  and a standard deviation of  $0.5^\circ\text{C}$ . Total errors in the modeled fluxes due to (1) and (2) are calculated as root-mean-sum of all individual 30-min errors, and are labeled as  $\delta Q_{z_0}$  and  $\delta Q_{T_0}$ , respectively. Finally, to assess errors in (3), we make use of the air temperature from the sonic anemometer ( $T_{sa}$ ), where susceptibility to radiative overheating is minimal. The values of  $T_{zt}$  are subtracted from  $T_{sa}$ , giving  $T_{bias}$ , which during nighttime periods (no solar radiation) is calculated to be  $2^\circ\text{C}$ . The nighttime bias is removed from  $T_{bias}$ , which is then regressed against the incoming shortwave radiation ( $K_{down}$ ) within a set of given ranges of measured wind speed (see below). The rationale here is that, for a given  $K_{down}$ , the higher the wind speed, the more efficient the natural ventilation of the sensor, leading to smaller  $T_{bias}$  (e.g., Harrison, 2010). The linear regression, applied on



**Figure 2.** Initial number of observations for estimating the roughness lengths (from 30-min averages of OPEC data) and remaining number of observations after each filter is applied (ordered from top down) for 2010 and 2012 observational periods. Solid line represents the filters used to determine the roughness lengths, and dotted line represents the filters used to determine the fluxes of momentum, temperature and humidity.

the 2012 data, yields the following relationship between  $T_{bias}$  and  $K_{down}$ :

$$T_{bias} = 0.0026 K_{down} + 0.035, \text{ for } 1 < U_{zv} < 2,$$

$$T_{bias} = 0.0018 K_{down} + 0.049, \text{ for } 2 < U_{zv} < 3,$$

$$T_{bias} = 0.0010 K_{down} + 0.063, \text{ for } 3 < U_{zv},$$

(37)

- 5 where  $K_{down}$  is in  $\text{W m}^{-2}$  and  $U_{zv}$  is in  $\text{m s}^{-1}$ . Corrected air temperatures are then calculated by subtracting  $T_{bias}$  from  $T_{zt}$ . Differences between the modeled fluxes, derived by the bulk methods with and without corrected temperature, represent the 30-min errors, while their root-mean-sum gives the total error over the observational period ( $\delta Q_{T_{bias}}$ ).

### 3 Results

#### 3.1 Estimates of the roughness lengths

- 10 We apply the filters discussed in Section 2.3 to the OPEC data to obtain the roughness lengths using the Eq. 5 - 7. For the 12-day observational period in 2010 there are 515 unfiltered measurements for  $z_{0v,0t}$  and 307 for  $z_{0q}$ . The fewer initial data points for  $z_{0q}$  arises from high sensitivity of the KH2O sensor to the build-up of droplets on the sensor window during precipitation events. During these events, the KH2O sensor produced spurious values or failed to record the data, while CSAT-3 recorded observations, albeit with questionable quality. After filtering, usable data are reduced by 98 %, giving nine high-
- 15 quality estimates for  $z_{0v,0t}$ , but none for  $z_{0q}$  (Fig. 2). Similarly, the initial number of data points in the 29-day observational

**Table 1.** Mean and  $\pm$  standard deviation ( $\sigma$ ) of log roughness lengths of momentum ( $z_{0v}$ ), temperature ( $z_{0t}$ ) and humidity ( $z_{0q}$ ) derived for near-neutral stability conditions ( $|\frac{z}{L}| < 0.1$ ) during 2010 and 2012 observational period (units are in m). Also shown is weighted mean  $\pm$  standard deviation ( $\sigma_w$ ; see text) and number of points ( $N$ ) used to assess these values, i.e. number of points remaining after all the filters have been applied to the original data.

| Parameter | mean $\pm \sigma$  | weighted mean $\pm \sigma_w$ | $N$  | mean $\pm \sigma$  | weighted mean $\pm \sigma_w$ | $N$  |
|-----------|--------------------|------------------------------|------|--------------------|------------------------------|------|
|           | 2010               | 2010                         | 2010 | 2012               | 2012                         | 2012 |
|           | $\log_{10}$ values | $\log_{10}$ values           |      | $\log_{10}$ values | $\log_{10}$ values           |      |
| $z_{0v}$  | $-2.67 \pm 0.66$   | $-2.63 \pm 0.67$             | 9    | $-2.56 \pm 0.75$   | $-2.54 \pm 0.76$             | 30   |
| $z_{0t}$  | $-5.08 \pm 1.11$   | $-4.98 \pm 0.98$             | 9    | $-4.58 \pm 0.87$   | $-4.52 \pm 0.79$             | 25   |
| $z_{0q}$  |                    |                              | 0    | $-4.11 \pm 1.15$   | $-3.89 \pm 1.09$             | 6    |

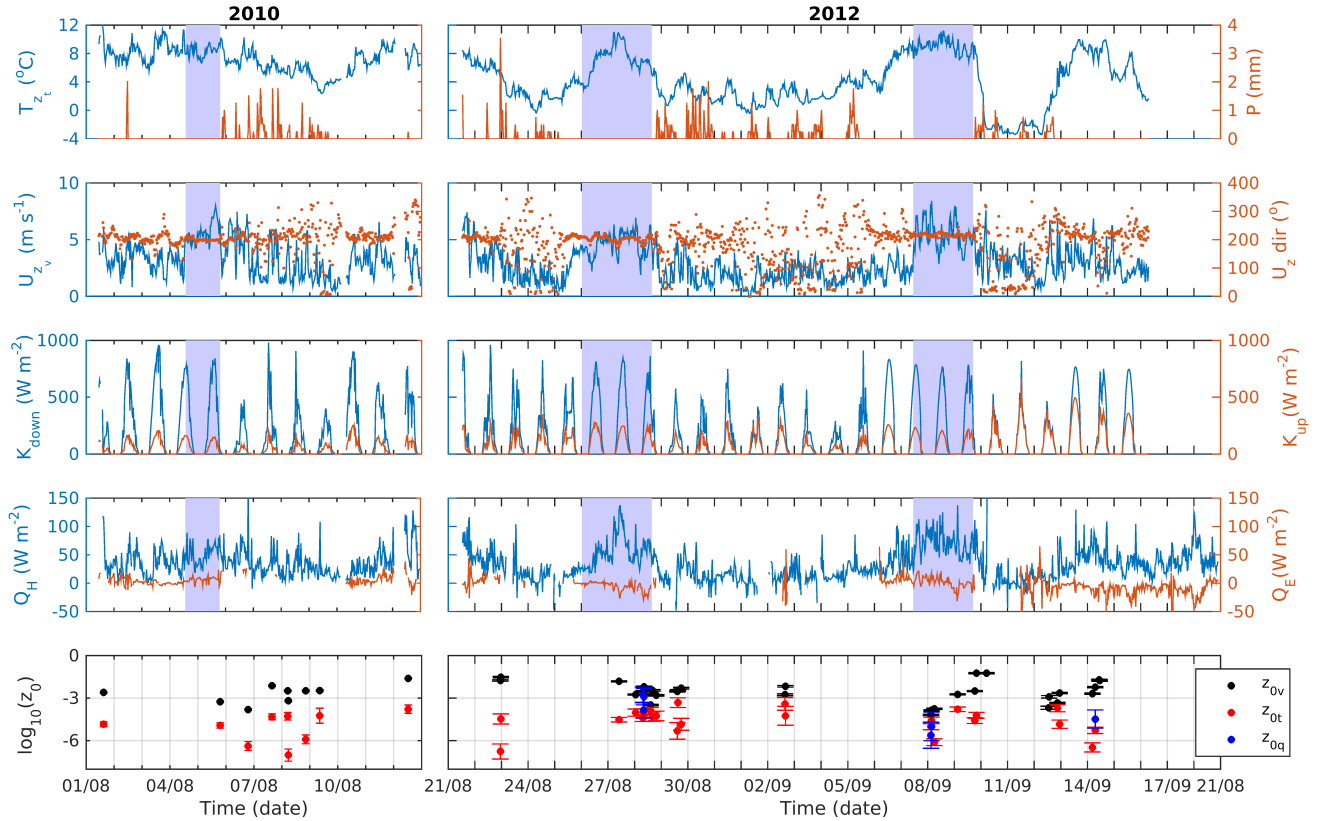
period in 2012 is 1161 for  $z_{0v,0t}$  and 621 for  $z_{0q}$ , while the post-filtering number of data points are 30 for  $z_{0v}$ , 25 for  $z_{0t}$  and six for  $z_{0q}$ .

Among all the filters, the 'neutrality filter' ( $|\frac{z_{v,t}}{L}| < 0.1$ ) eliminates most of the data points, followed-by the 'stationarity' filter. Despite the significant loss of viable data, these filters are essential to achieve accurate determination of roughness lengths with the bulk method. OPEC-derived fluxes are assumed valid only during the steady-state or stationary conditions (i.e. in the absence of intermittent turbulence and/or gravity waves). For the near-neutral conditions, the choice of the stability correction functions has negligible effect on the calculated fluxes (the correction function approaches zero as  $|\frac{z_{v,t}}{L}|$  approaches zero). Furthermore, the near-neutral stability conditions are more commonly observed during overcast and windy conditions rather than during clear-sky conditions with prevailing katabatic flows. This fact minimizes the chance of selecting 30-min intervals with katabatic flows for which the bulk method, used to determine the roughness lengths, might be deficient.

We determine the time windows of prevailing katabatic flows (shaded areas in Fig. 3) by identifying longer periods (several hours to days) with almost stationary down-glacier wind direction ( $200^\circ$  for our site) and relatively high mean wind speeds ( $> 3 \text{ m s}^{-1}$ ), preferably reaching maximum 30-min mean values  $> 5 \text{ m s}^{-1}$ . With this rough identification criteria we may miss some shorter time intervals with katabatic flows, but a more restrictive approach reduces the uncertainty that originates from the selection of katabatic flow periods in our further analysis. The identified prevailing katabatic periods (Fig. 3) overlap with clear sky conditions and with relatively high near-surface air temperatures ( $T_{zt}$ ), as is commonly expected for katabatic flows at glacier surfaces (e.g., Oerlemans et al., 1999).

Each roughness length estimate has its respective error (error bars in Fig. 3 derived from the error propagation (Section 2.5). To provide the mean value for roughness lengths during neutral conditions we used the average of their logged values to account for the measurements' log-normal distribution. To incorporate the individual error of each estimate into the mean value, we derive a weighted mean roughness length, where the weights are inversely proportional to the individual errors. The results (Table 1) show that the weighted mean values do not substantially differ from the mean values and we thus use the non-weighted mean values in the remaining analysis. For both seasons, our results show that mean  $\log z_{0v}$  is of the order of



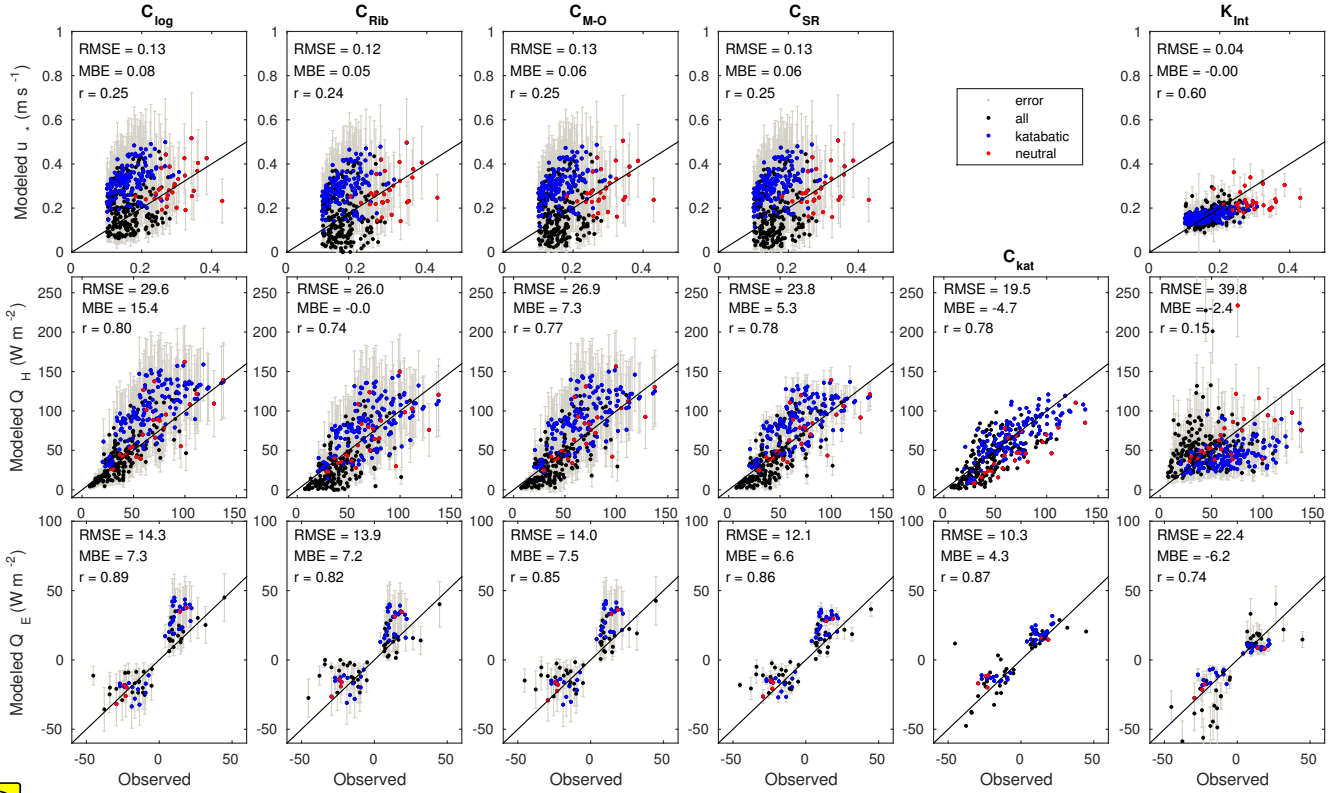


**Figure 3.** Thirty minute averages of meteorological variables and SEB fluxes measured at the  $\text{AWS}_{\text{glac}}$  in 2010 (left panel) and 2012 (right panel): near-surface air temperature ( $T_z$ ), accumulated precipitation ( $P$ ; measured at  $\text{AWS}_{\text{low}}$ ), wind speed ( $U_z$ ), wind direction ( $U_{\text{dir}}$ ), incoming ( $K_{\text{down}}$ ) and reflected ( $K_{\text{up}}$ ) shortwave radiation, and OPEC-derived sensible ( $Q_H$ ) and latent ( $Q_E$ ) heat flux. Bottom panel shows the OPEC-derived roughness lengths for momentum ( $z_{0v}$ ), temperature ( $z_{0t}$ ) and humidity ( $z_{0q}$ ), after filtering the OPEC data, and their estimated errors. Shaded in blue are the identified time intervals of prevailing katabatic flow (see text).

$10^{-3}$  m and two orders of magnitude larger than mean  $\log z_{0t}$ . The overall error of these mean values, expressed as one sigma ( $\pm$  a standard deviation), shows uncertainty of one order of magnitude around the mean value (Table 1). To determine  $z_{0q}$  for 2010 season, which had no high-quality measurements for  $z_{0q}$ ; we thus assume  $z_{0q} \approx z_{0t}$  since the equality holds for the mean log values in the 2012 season.

### 5 3.2 Inter-comparison of bulk methods

We compare 30-min turbulent fluxes of momentum, sensible and latent heat, derived from the OPEC data (observed fluxes) with the modeled fluxes from each bulk scheme. To ensure that the evaluation is performed on high quality data, only the 30-min fluxes that pass the filters in Section 2.3 are used, however, without the 'neutrality' filter and allowing for  $U_z > 1 \text{ m s}^{-1}$  (Fig.

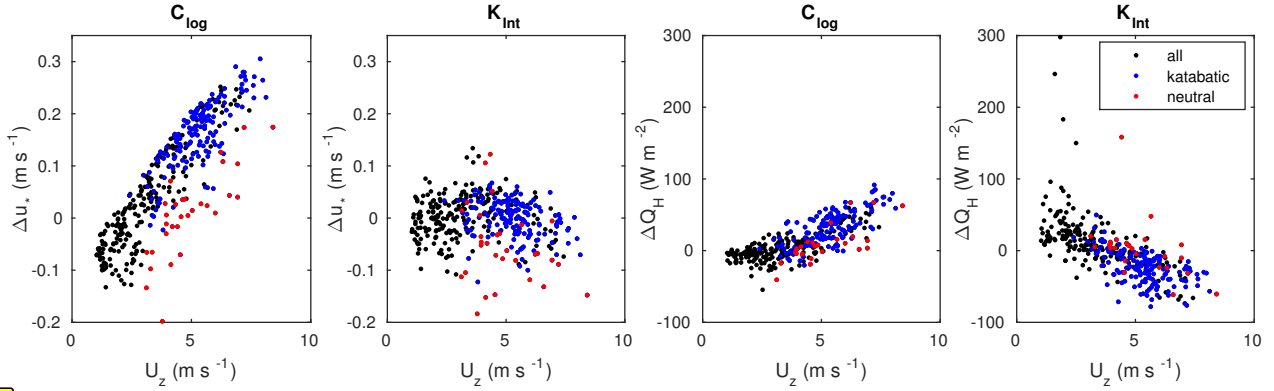


**Figure 4.** Comparison of 30-min observed (OPEC-derived) versus modeled friction velocity ( $u_*$ ), sensible ( $Q_H$ ) and latent ( $Q_E$ ) heat flux, according to the six bulk methods:  $C_{log}$ ,  $C_{Rib}$ ,  $C_{M-O}$ ,  $C_{SR}$ ,  $C_{kat}$ , and  $K_{Int}$ . Dots and error bars respectively show the mean and standard deviation of Monte Carlo ensembles (see text). No error bars are given for  $C_{kat}$  method since it is independent of roughness lengths. Red points denote values during the near-neutral conditions ( $|\frac{z_{v,i}}{L}| < 0.1$ ) only, and blue points indicate values during katabatic flow. Root-mean-square-error (RMSE;  $W m^{-2}$ ), mean bias error (MBE;  $W m^{-2}$ ), and Pearson correlation coefficient ( $r$ ) are provided for each scatter plot.

2). This filtering yielded 239, 222, and 77 values for  $u_*$ ,  $Q_H$ , and  $Q_E$ , respectively in 2010, while for 2012 season we obtained 393, 353, and 83 values, respectively. In all methods, if not otherwise mentioned, we assume  $Pr=1$ , i.e.  $K_M = K_H = K_E$ .

Modeled fluxes for the 2012 season (y-axis in Fig. 4) are the mean values from the ensemble of Monte Carlo runs with the randomized roughness lengths, whereas the error bars are represented with  $\pm$  one standard deviation ( $\sigma$ ) of the ensemble. All

- 5 C-methods poorly simulate  $u_*$  ( $r=0.25$ ), with substantial overestimation during katabatic flow, and underestimate  $u_*$  during non-katabatic conditions and relatively low wind speeds ( $U_z < 3 m s^{-1}$ ). This type of dependency of the bias (modeled minus observed  $u_*$ ) on the wind speed is present in each C-method (Fig. 5). The  $K_{Int}$  method provides the best simulation of  $u_*$  ( $r=0.60$ ), while the bias dependence on the wind speed is less pronounced than in the C-methods (Fig. 5). Good performance of  $K_{Int}$  method is expected since the parameter  $K_{max}$  is optimized to maximize the match between modeled and observed  $u_*$ .
- 10 All C-methods simulate  $Q_H$  and  $Q_E$  equally well ( $r \approx 0.77$ ) despite their equally poor performance in simulating  $u_*$ , while



**Figure 5.** Bias (modeled minus observed values) in  $u_*$  and  $Q_H$  versus the wind speed. Red points denote values during the near-neutral conditions ( $|\frac{z_{v,t}}{L}| < 0.1$ ) only, and blue points indicate values during katabatic flow.

the  $K_{Int}$  method performs poorly ( $r=0.15$ ) for  $Q_H$ .  $C_{kat}$  method performs best in simulating  $Q_H$  since its parameter has been optimized with the observed  $Q_H$ . The scatter plots for the 2010 season (not shown) are similar to these for the 2012 season, while the correlation values are smaller in 2010 due to smaller sampling period (Table 2).

The performance of the  $C_{SR}$  and  $C_{M-O}$  methods accord because the OPEC-derived values for scalar roughness lengths (near-neutral conditions only) largely agree with the estimated values from the surface renewal model (Fig. 6). Large roughness Reynolds numbers ( $Re_*$ ) estimated for our site indicate a rough flow regime for which the surface renewal model (Andreas, 1987) predicts a scalar roughness length two orders of magnitude smaller than  $z_{0v}$ . This prediction is similar to our observed values, especially when the mean log values of roughness lengths rather than their individual 30-min estimates are used. Despite the large scatter of 30-min OPEC-derived  $\frac{z_{0,t}}{z_{0v}}$  values around the predicted values from the surface renewal model, more than 50 % of the observed values fall between the predictions of Andreas (1987) and those of Smeets and van den Broeke (2008) expected for hummocky ice with  $z_{0v} > 1 \times 10^{-3}$  m (Fig. 6).

Intercomparison of the C-methods (Fig. 4) reveals that the performance of  $C_{M-O}$  method is similar to the  $C_{log}$  method. This similarity arises in part because the M-O stability parameter ( $\frac{z_{v,t}}{L}$ ), calculated with the fixed-point iterative scheme of Munro (1989) underestimates the stability during katabatic conditions and overestimates it during conditions with low wind speeds (Fig. 7). The stability corrections that depend on  $\frac{z_{v,t}}{L}$  thus have a small effect in modifying the fluxes during the katabatic conditions, while the fluxes are suppressed during the non-katabatic conditions. We also found no correlation between the 30-min OPEC-derived  $\frac{z_{v,t}}{L}$  and any of the mean meteorological variables (e.g. temperature, wind speed; Fig. 7), which likely explains the failure of the fixed-point iterative scheme that relies on these dependencies. The poor performance of the stability corrections in the  $C_{Rib}$  method also follows from the lack of correlation between  $Ri_b$  and the OPEC-derived  $\frac{z_{v,t}}{L}$  (Fig. 7). The lack of a relation between  $Ri_b$  and  $\frac{z_{v,t}}{L}$  disagrees with the proposed stratification-dependent conversion of  $\frac{z_{v,t}}{L}$  into  $Ri_b$  as presented in Arya (2001). Defining  $Ri_b$  through  $U_z$  and  $\Delta T$  (Eq. 24) assumes that the mean variables are a good proxy for the local stability: during the near neutral conditions, ideally characterized by overcast weather with strong winds and relatively



**Table 2.** Comparison between modeled and OPEC-derived sensible ( $Q_H$ ) and latent ( $Q_E$ ) heat fluxes, expressed as root-mean-square-error (RMSE), mean bias error (modeled minus observed; MBE), and Pearson correlation coefficient (r) for 2012 and 2010 (values in parenthesis) observational period, given for a set of six bulk methods and their variants (see text).

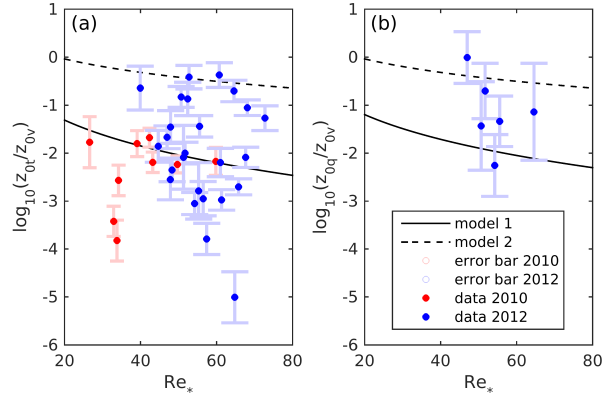
| Method                           | $Q_H$             | $Q_H$             | $Q_H$         | $Q_E$             | $Q_E$             | $Q_E$         |
|----------------------------------|-------------------|-------------------|---------------|-------------------|-------------------|---------------|
|                                  | RMSE              | MBE               | r             | RMSE              | MBE               | r             |
|                                  | $\text{W m}^{-2}$ | $\text{W m}^{-2}$ |               | $\text{W m}^{-2}$ | $\text{W m}^{-2}$ |               |
| (1) $C_{\log}$                   | 29.6 ( 28.7 )     | 15.4 ( 20.6 )     | 0.80 ( 0.57 ) | 14.3 ( 11.5 )     | 7.3 ( 9.7 )       | 0.89 ( 0.65 ) |
| $C_{\log} u_*$                   | 15.2 ( 11.4 )     | -2.3 ( -1.4 )     | 0.87 ( 0.77 ) | 6.7 ( 2.9 )       | 1.5 ( 0.6 )       | 0.94 ( 0.86 ) |
| $C_{\log} K_{\text{Int}}$        | 16.4 ( 13.0 )     | -4.4 ( -1.7 )     | 0.82 ( 0.66 ) | 7.4 ( 4.4 )       | 1.0 ( 1.3 )       | 0.92 ( 0.73 ) |
| (2) $C_{\text{Rib}}$             | 26.0 ( 23.7 )     | 0.0 ( 3.1 )       | 0.74 ( 0.46 ) | 13.9 ( 8.7 )      | 7.2 ( 4.5 )       | 0.82 ( 0.43 ) |
| $C_{\text{Rib}} u_*$             | 26.6 ( 21.5 )     | -21.9 ( -17.0 )   | 0.84 ( 0.65 ) | 9.8 ( 5.0 )       | 2.8 ( -2.7 )      | 0.88 ( 0.65 ) |
| (3) $C_{\text{M-O}}$             | 26.9 ( 25.3 )     | 7.3 ( 11.5 )      | 0.77 ( 0.51 ) | 14.0 ( 10.0 )     | 7.5 ( 7.1 )       | 0.85 ( 0.54 ) |
| $C_{\text{M-O}} u_*$             | 20.3 ( 13.7 )     | -13.6 ( -8.7 )    | 0.84 ( 0.79 ) | 6.5 ( 3.0 )       | 1.2 ( -0.6 )      | 0.94 ( 0.85 ) |
| $C_{\text{M-O}} \frac{z}{L}$     | 20.1 ( 18.3 )     | -5.3 ( 1.4 )      | 0.78 ( 0.53 ) | 9.9 ( 6.4 )       | 4.3 ( 3.0 )       | 0.88 ( 0.57 ) |
| $C_{\text{M-O}} \frac{z}{L} u_*$ | 22.3 ( 15.1 )     | -15.0 ( -9.7 )    | 0.80 ( 0.74 ) | 7.0 ( 3.3 )       | 0.9 ( -1.0 )      | 0.93 ( 0.82 ) |
| $C_{\text{M-O}} u_* \text{ Pr}$  | 16.5 ( 12.5 )     | -8.0 ( -6.2 )     | 0.86 ( 0.78 ) | 6.9 ( 3.3 )       | 1.7 ( 0.4 )       | 0.94 ( 0.81 ) |
| (4) $C_{\text{SR}}$              | 23.8 ( 28.3 )     | 5.3 ( 16.1 )      | 0.78 ( 0.52 ) | 12.1 ( 11.7 )     | 6.6 ( 9.0 )       | 0.86 ( 0.54 ) |
| $C_{\text{SR}} u_*$              | 18.7 ( 10.6 )     | -11.7 ( -2.3 )    | 0.85 ( 0.80 ) | 6.4 ( 3.3 )       | 1.3 ( 1.6 )       | 0.94 ( 0.86 ) |
| $C_{\text{SR}} \frac{z}{L}$      | 18.7 ( 19.0 )     | -6.3 ( 6.2 )      | 0.80 ( 0.54 ) | 8.9 ( 7.4 )       | 3.6 ( 4.7 )       | 0.90 ( 0.61 ) |
| $C_{\text{SR}} \frac{z}{L} u_*$  | 20.8 ( 12.2 )     | -13.1 ( -2.8 )    | 0.81 ( 0.73 ) | 6.9 ( 3.5 )       | 1.0 ( 1.2 )       | 0.93 ( 0.82 ) |
| (5) $C_{\text{kat}}$             | 19.5 ( 17.9 )     | -4.7 ( 1.4 )      | 0.78 ( 0.44 ) | 10.3 ( 5.2 )      | 4.3 ( 3.2 )       | 0.87 ( 0.76 ) |
| (6) $K_{\text{Int}}$             | 39.8 ( 31.5 )     | -2.4 ( 6.8 )      | 0.15 ( 0.20 ) | 22.4 ( 13.6 )     | -6.2 ( 4.7 )      | 0.74 ( 0.46 ) |
| $K_{\text{Int}} \text{ Pr}$      | 37.7 ( 30.6 )     | 2.9 ( 8.7 )       | 0.26 ( 0.23 ) | 22.2 ( 13.6 )     | -6.2 ( 5.5 )      | 0.77 ( 0.45 ) |

low  $\Delta T$ ,  $Ri_b$  should be close to zero. On a sloped glacier surface, however, strong winds commonly occur during clear sky conditions with a katabatic flow (Fig. 7).  $Ri_b$  close to zero is thus more indicative of strong katabatic flow than neutral stability.

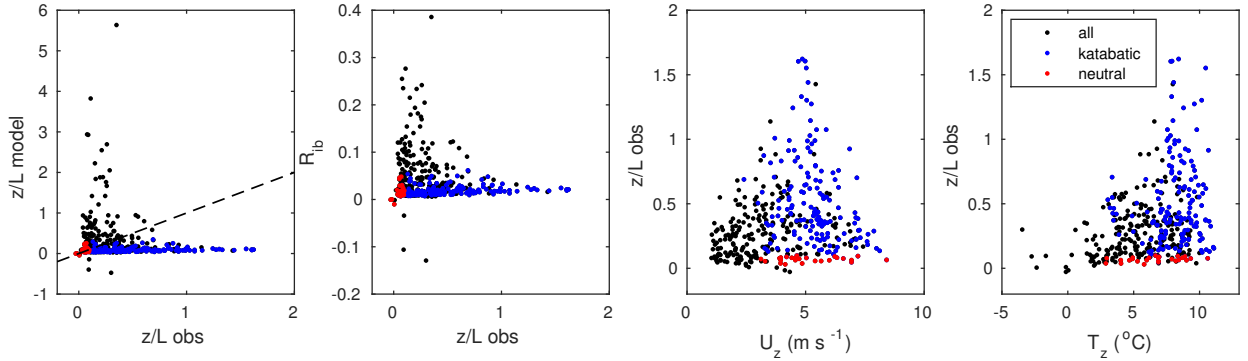
### 3.3 Analysis of bulk method performance

#### 3.3.1 C-methods

- 5 In the gradient-flux relation, the eddy viscosity is parameterized as a function of  $z$ ,  $u_*$  and M-O stability parameter ( $\frac{z}{L}$ ). Because  $u_*$  and  $\frac{z}{L}$ , in the C-methods, are estimated rather than directly measured, any error in these modeled values can propagate into the flux estimates. Our goal in this section is to investigate the influence of the two variables,  $u_*$  and  $\frac{z_{u,t}}{L}$ , on the bulk method performance. To do so, we estimate the turbulent fluxes from each of the four bulk schemes using the OPEC-derived  $u_*$  and Obukhov length ( $L$ ). Henceforth the notation  $u_*$  and/or  $\frac{z}{L}$  at the end of each method's name indicates the use of OPEC-derived



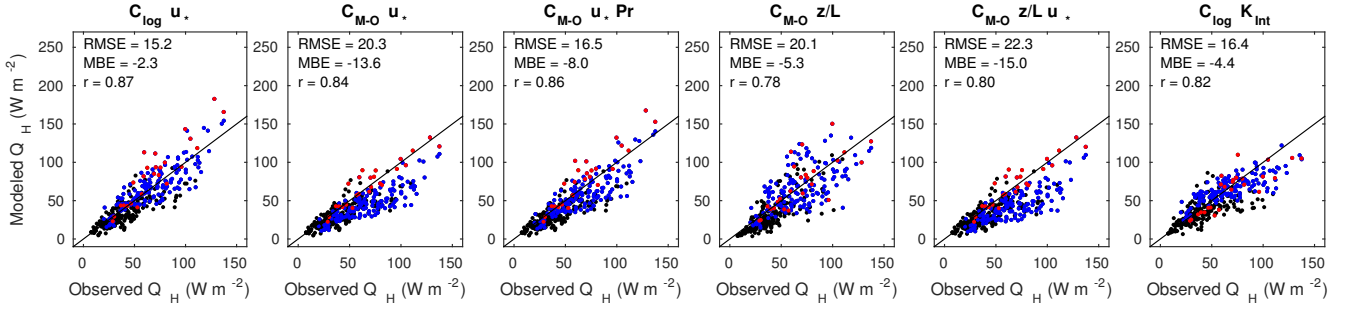
**Figure 6.** The ratio of roughness lengths for temperature ( $z_{0t}$ ) and momentum ( $z_{0v}$ ) (panel a) and the ratio of roughness lengths for humidity ( $z_{0q}$ ) and  $z_{0v}$  (panel b) versus roughness Reynolds number ( $Re_*$ ). Model 1 (solid line) is the theoretical prediction of Andreas (1987), and model 2 (dashed line) is the prediction of Smeets and van den Broeke (2008). Mean log value of  $z_{0v}$  specific to each season (Table 1) is used to calculate  $Re_*$ .



**Figure 7.** Modeled  $z/L$  with the fixed-point iterative scheme in the  $C_{M-O}$  method and the Bulk Richardson number ( $Ri_b$ ) against the OPEC-derived stability parameter ( $z/L$  obs). Dashed black line shows 1:1 line. Also shown is a dependency of the 30-min OPEC-derived stability parameter ( $z/L$  obs) on the measured wind speed ( $U_z$ ) and the near-surface air temperature ( $T_z$ ). Colors are the same as in Fig. 4 and 5.

variables (e.g.  $C_{\log} u_*$  method,  $C_{M-O} u_* \frac{z}{L}$  method,  $C_{SR} \frac{z}{L}$  method). In each scheme we apply mean log roughness lengths (Table 1).

The following results are found for both the 2010 and 2012 seasons:  $C_{\log} u_*$  and  $C_{SR} u_* \frac{z}{L}$  performed best (lowest RMSE) in simulating the heat fluxes among the C-methods with measured  $u_*$  and/or  $\frac{z_{v,t}}{L}$  (Table 2). The variance error (scatter of residuals around 1:1 line) in the  $C_{\log} u_*$  method is also smaller than in the  $C_{M-O} \frac{z}{L}$  method (Fig. 8). Both results suggest that in order to better simulate  $Q_H$ , correct estimates of  $u_*$  are more important than accurately estimating  $\frac{z_v}{L}$ . The  $C_{\log} u_*$  method



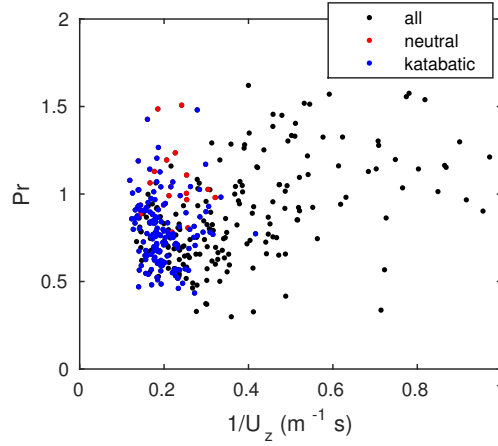
**Figure 8.** Comparison of 30-min values of observed (OPEC-derived) and modeled sensible heat flux ( $Q_H$ ) using six different bulk methods:  $C_{\log u_*}$ ,  $C_{M-O} u_*$ ,  $C_{M-O} u_* Pr$ ,  $C_{M-O} \frac{z}{L}$ ,  $C_{M-O} \frac{z}{L} u_*$ , and  $C_{\log K_{Int}}$ . Red (blue) points show values during near-neutral (katabatic) flow conditions.

outperforms any  $C_{M-O}$  method, even with OPEC-derived  $u_*$  and  $\frac{z}{L}$ . The results for  $Q_E$  agree with the results for  $Q_H$ . Despite  $Q_E$  being less affected by the M-O stability corrections than  $Q_H$ , the modeled latent heat fluxes are closer to those observed if the OPEC-derived  $\frac{z_v}{L}$  are used. While the original C-methods overestimate  $Q_H$  during the katabatic conditions, the C-methods with stability corrections and measured  $u_*$  underestimate the fluxes (Table 2 and Fig. 8).

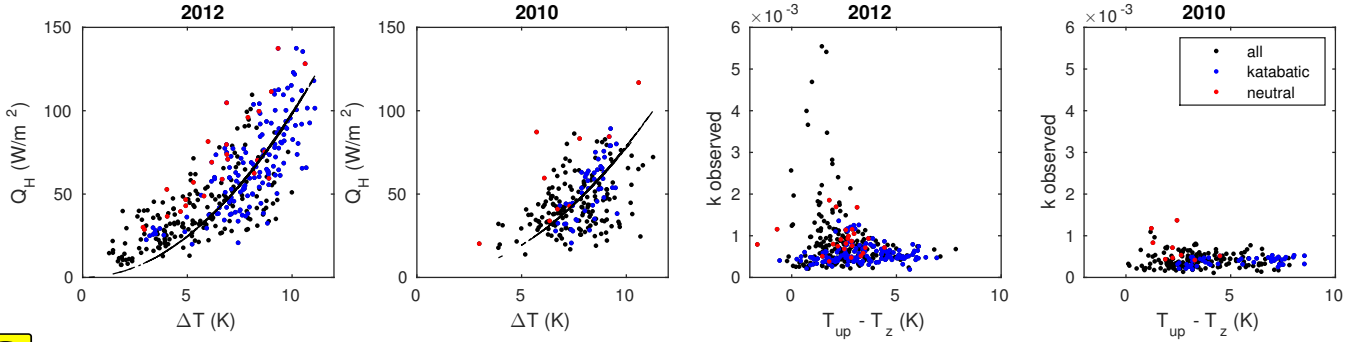
- 5 Overestimation of  $Q_H$  during katabatic conditions has also been shown by Denby and Greuell (2000) and explained by a failure of M-O theory in the presence of a shallow katabatic wind speed maximum. At the wind speed maximum, measured  $u_*$  approaches zero, while the C-method assumes constant momentum flux in the surface layer and therefore overestimates  $u_*$ . The overestimation is less pronounced for  $Q_H$  than for  $u_*$  because the reduced turbulence at the wind speed maximum leads to an increase in the air-surface temperature difference, and subsequently an increase in the measured  $Q_H$ . When measured  $u_*$
- 10 is used in the  $C_{M-O}$  or  $C_{Rib}$  method, however, the C-method underestimates  $Q_H$  since the air-surface temperature difference alone cannot compensate for the effect of reduced momentum flux. To correct for this bias in  $Q_H$ ,  $Pr$  would need to decrease (i.e. the C-method would need to account for more effective eddy diffusivity than eddy viscosity at the given height). In the absence of wind profile measurements, we can only assume that these effects take place at our site, but we have no observational evidence for the presence of the wind speed maximum. By plotting  $Pr$ , derived from Eq. (19) using OPEC-derived  $Q_H$  and  $u_*$ ,
- 15 against  $1/U_z$ , we show that  $Pr < 1$  during katabatic conditions with high wind speeds (Fig. 9), yielding an average  $Pr=0.79$  during the prevailing katabatic flow. Due to a self-correlation between  $Pr$  and M-O stability parameter, we do not assess  $Pr$  dependence on OPEC-derived  $\frac{z_v}{L}$ , however, several studies showed that as the local stability increases Prandtl number decreases, leading to  $Pr < 1$  for  $\frac{z}{L} > 1$  (Grachev et al., 2007). Setting  $Pr=0.79$  in the  $C_{M-O}$  method with measured  $u_*$  (labeled as  $C_{M-O} u_* Pr$ ) during the katabatic conditions only, improves the performance of the C-method in simulating  $Q_H$  (Fig. 8).

## 20 3.4 $C_{kat}$ method

The optimization of the  $k$  parameter yields  $k = 4.00 \times 10^{-4}$  and  $k = 5.04 \times 10^{-4}$  for the 2010 and 2012 seasons, respectively. When this optimized  $C_{kat}$  is used to derive sensible and latent heat fluxes (Eq. (30) and (31)) the modeled fluxes simulate the



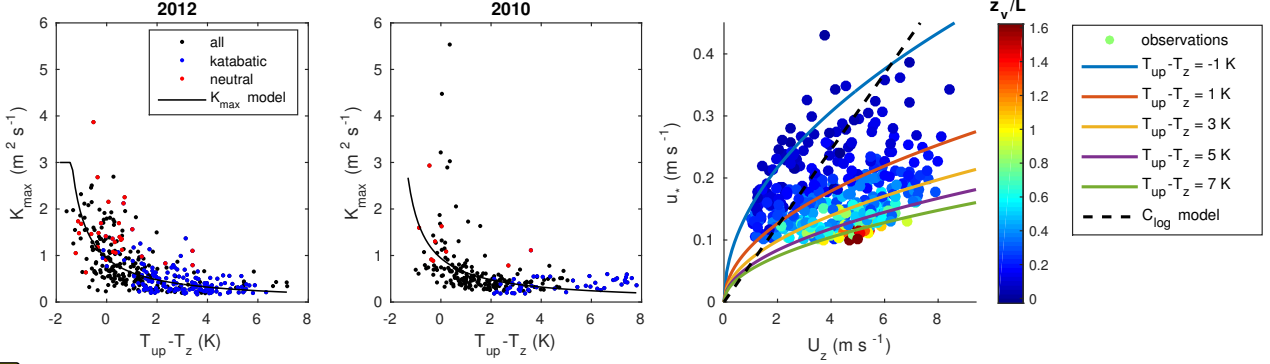
**Figure 9.** Measured eddy Prandtl number versus the inverse of the wind speed.



**Figure 10.** Observed 30-min sensible heat flux ( $Q_H$ ) versus observed air-surface temperature difference ( $\Delta T$ ), and modeled 30-min  $Q_H$  fluxes (black line) according to the  $C_{kat}$  method, for the 2012 and 2010 seasons. Dependence of the 'observed'  $k$  parameter, calculated from OPEC-derived  $Q_H$  and measured  $\Delta T$ , on the 2-m air temperature difference between  $AWS_{up}$  and  $AWS_{glac}$  ( $T_{up} - T_z$ , for the 2012 and 2010 seasons).

observed ones better (smaller RMSE and MBE) than any of the C-methods (1)-(4) (Table 2). There is a quadratic fit between OPEC-derived  $Q_H$  and  $\Delta T$  (Fig. 10), which explains 61 % (19 %) of the variance in the 2012 (2010) observations. Our results do not guarantee the validity of Eq. (30), but demonstrate that the simple katabatic model of Oerlemans and Grisogono (2002) reproduces the basic characteristics of katabatic flow at our site. This model, when calibrated, can simulate the turbulent heat exchange more successfully than a common bulk method (C-method).

In the original model (Oerlemans and Grisogono, 2002),  $C_{kat}$  depends on the environmental lapse rate, such that  $C_{kat}$  decreases as the static stability increases. Our proxy for stability is the difference between on-glacier ( $AWS_{glac}$ ) and off-glacier temperature ( $AWS_{up}$  or  $AWS_{low}$ ). We assess the correlation between this proxy variable and the 'observed' parameter  $k$ ,



**Figure 11.**  $K_{max}$  versus air temperature difference between  $AWS_{up}$  and  $AWS_{glac}$ , for the 2012 and 2010 seasons. Black (blue) circles represent all (katabatic) observations, and black line is the empirical fit model for  $K_{max}$ . OPEC-derived 30-min friction velocity,  $u_*$ , versus observed wind speed  $U_{zv}$ , for the 2012 season, where each observation point is colored according to its 30-min value of OPEC-derived  $\frac{z_v}{L}$  (colorbar). Solid colored lines are modeled  $u_*$ , using the  $K_{Int}$  method with  $\overline{H_k}=20$  m and optimized  $K_{max}$ , for five different stability conditions ( $T_{up}-T_z=-1, 1, 3, 5$  and  $7$  K). Dashed line is modeled  $u_*$  using the  $C_{log}$  method.

whose values are derived from OPEC-derived  $Q_H$  and measured  $\Delta T$  (Eq. 30) using the same parameter values in Eq. 29 as in the prior optimization method. For the 2012 dataset, the correlation in the log-log regression between the 'observed'  $k$  and  $T_{up} - T_z$  ( $r=-0.35$ ,  $p$ -value  $< 0.01$ ) is higher than in the case with  $T_{low} - T_z$  ( $r=-0.14$ ,  $p < 0.01$ ). The power-law relationship between the 'observed'  $k$  and  $T_{up} - T_z$  reveals that as the temperature difference increases, indicative of higher static stability, the  $k$  parameter decreases (Fig. 10). This finding agrees with the 2010 dataset, however, the correlations are not statistically significant at 95 % confidence level (Fig. 10).

### 3.5 $K_{Int}$ method

The relation between the optimized 30-min  $K_{max}$  and the proxy variable is strongly non linear (Fig. 11). The best fitting model ( $r=0.61$ ,  $p$ -value  $< 0.01$ ) that relates  $K_{max}$  to the difference in air temperature at  $AWS_{up}$  and  $AWS_{glac}$  yields the following power-law relation:

$$K_{max} = 1.86 (T_{up} - T_{zt} + 2)^{-0.98}, \quad (38)$$

where the units for  $K_{max}$  are in  $m^2 s^{-1}$ . Note that this empirical fit is assumed valid within the range of commonly observed values for  $K_{max}$  taken from Parmhed et al. (2004):  $0.03 m^2 s^{-1} < K_{max} < 3 m^2 s^{-1}$ . The power-law model is calibrated on the 2012 data only (Fig. 11), while the 2010 data are used for the model validation. The validation yields statistically significant correlation coefficients between the 30-min modeled and observed  $K_{max}$  values ( $r=0.42$ ,  $p$ -value  $< 0.01$ ). This result indicates that, for our study site, the same empirical model is well applicable to both field seasons (Fig. 11).

The modeled  $u_*$  derived from  $K_{Int} \frac{z}{L}$  method (Eq. (33)) successfully captures two features observed in the data (Fig. 11): (i) a non-linear relation between friction velocity and wind speed; and (ii) a functional dependence of this non-linear relation



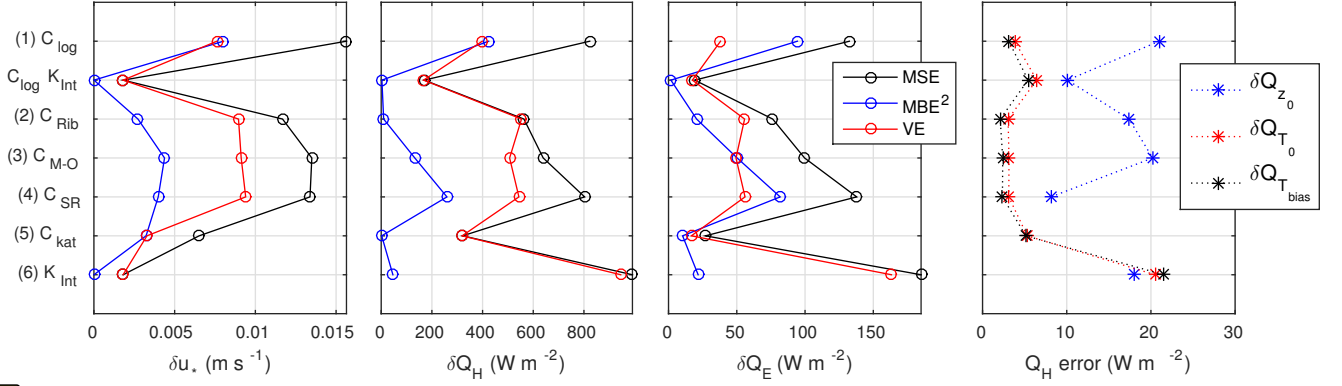
on the proxy variable for  $\gamma$ , i.e. environmental static stability. For comparison, we also show modeled  $u_*$  versus  $U_{z_v}$  where the  $C_{\log}$  method is used to determine the friction velocity (Eq. (16)). For the near-neutral conditions ( $|z_v/L| < 0.1$ ), the  $C_{\log}$  method performs well because of the near-linear dependence between the OPEC-derived  $u_*$  and wind speed. As stability increases, however, the linearity is replaced by a relation that more closely resembles  $u_* \propto \sqrt{U_{z_v}}$ , where the strength of the proportionality depends on the environmental static stability, approximated here with the off-glacier and on-glacier temperature difference ( $T_{up} - T_z$ ; Fig. 11).

As already shown (Fig. 4), the  $K_{Int}$  method gives the best estimate of the friction velocity and the poorest estimate of  $Q_H$  among the bulk methods. This performance pattern indicates that  $K_M$  and  $K_H$  do not necessarily share the same parametrization. Setting  $Pr=0.79$  for katabatic conditions, the  $K_{Int}$  method performance in simulating  $Q_H$  improves (Table 2). As our final test, we combine the best performing C-method with the  $K_{Int}$  method to investigate whether the top performance methods in simulating  $Q_H$  and  $u_*$ , respectively, when combined, give the best performance overall. This variant of the bulk method calculates the friction velocity from the  $K_{Int}$  method (Eq. (36)), while the heat fluxes are calculated from the  $C_{\log}$  method following Eq. (22) and (23). The results show that this  $C_{\log}K_{Int}$  'hybrid' method indeed outperforms all the methods in the study (Fig. 8).

### 3.6 Error Analysis



We summarize uncertainties between modeled and observed fluxes for the six bulk methods used in the study (four types of C-methods and two types of methods based on the katabatic models) including the 'hybrid'  $C_{\log}K_{Int}$  method (Fig. 12). Error estimates are shown for the 2012 season only, and agree well with the estimates for the 2010 season. The differences in model versus measured fluxes are presented as a Mean Square Error (MSE, which is equal to  $RMSE^2$  from Table 2) decomposed into a squared MBE and a variance error (VE), satisfying  $MSE = MBE^2 + VE$ . We also show the total errors due to: (1) the uncertainty in the roughness lengths ( $\delta Q_{z_0}$ ); (2) the assumption that the surface temperature is at a melting point ( $\delta Q_{T_0}$ ); and (3) the systematic error in air temperature due to radiative heating of the temperature sensor ( $\delta Q_{T_{bias}}$ ). The  $C_{\log} K_{Int}$  hybrid method has the lowest MSE among all bulk methods (Fig. 12). Across the methods, the total error (MSE) is more dominated by the contribution from the variance error (VE) than by the mean bias error. Modeled  $Q_H$  with the  $C_{\log}$  and  $C_{M-O}$  method has the largest sensitivity to perturbations in roughness lengths (largest  $\delta Q_{z_0}$ ), likely because the uncertainties in both  $z_{0v}$  and  $z_{0t}$  propagate into the modeled  $\frac{z_{v,t}}{L}$  and then, via the fixed-point iterative method, into estimated  $Q_H$ . Errors due to perturbations in  $T_0$  ( $\delta Q_{T_0}$ ) and bias in  $T_z$  ( $\delta Q_{T_{bias}}$ ) are small among all of the methods, except the  $K_{Int}$  method. Most importantly, our analysis reveals that the three sources of errors in modeled  $Q_H$  (C-methods) are all smaller than the RMSE (Table 2) between modeled and observed values. In other words, the 'internal' model errors are smaller than the bias in the model performance when compared to the measured fluxes. In short, further advances will require improvement to model theory rather than better measurements of input variables.



**Figure 12.** First three panels: Mean Square Error (MSE), squared Mean Bias Error (MBE<sup>2</sup>) and Variance Error (VE) for the six main bulk methods and the 'hybrid' C<sub>log</sub> K<sub>Int</sub> bulk method, for 2012 dataset. Note that MSE = MBE<sup>2</sup> + VE. Last panel shows the error in  $Q_H$  for each bulk method due to: (1) the uncertainty in the roughness lengths ( $\delta Q_{z_0}$ ); (2) the assumption that the surface temperature is at a melting point ( $\delta Q_{T_0}$ ) and; (3) the systematic error in air temperature due to radiative heating of the temperature sensor ( $\delta Q_{T_{bias}}$ ).

#### 4 Discussion

We discuss our findings in the order that we introduced uncertainties in the bulk methods, particularly in simulating  $Q_H$  on sub-daily scales. Uncertainties originate from (1) assumed rather than measured surface temperature, (2) estimation of roughness lengths, (3) representation of stability corrections in the bulk methods, and (4) an absence of a successful model applicable during very stable conditions and prevailing katabatic flow on sloped glacier surfaces.

1. In the absence of reliable measurements of surface temperature,  $T_0$ , we assumed the glacier surface to be at the melting point throughout the observational period. We quantified the uncertainty of this assumption using Monte Carlo sampling. Although significant for the 30-min  $Q_H$  estimates, model error ( $\delta Q_{T_0}$ ) is negligible over the whole observational period (Fig. 12). This error estimate does not account for any systematic biases in the surface temperature that might occur throughout the observational period (e.g. overnight cooling or refreezing of the surface; changes in surface temperature due to a debris cover or a formation of water channel at the study site). As long as surface temperature fluctuations are small, random, and centered at 0°C, assuming the glacier surface is at the melting point does not cause any substantial error in the simulated fluxes over the entire observational period.

2. To estimate roughness lengths ( $z_{0v}$ ,  $z_{0t}$ ,  $z_{0q}$ ) we used detailed eddy-covariance measurements subjected to a series of corrections and filters. Since the C<sub>M-O</sub> method is used to derive roughness lengths, it is crucial to assess the roughness lengths during the conditions when the best performance of this method is expected (i.e. during the near-neutral stability conditions when  $|\frac{z_{v,t}}{L}| < 0.1$ ). These conditions rarely occurred in our study, resulting in significantly fewer data. We also assumed the mean log values of  $z_{0v,0t,0q}$  are representative of the entire observational period. For both seasons, we

obtained  $z_{0v}$  that scales to  $10^{-3}$  m, which is characteristic for hummocky ice, while our  $z_{0t}$  and  $z_{0q}$  estimates turned out to be two orders of magnitude smaller than  $z_{0v}$ . These findings are in agreement with the previous studies on glaciers (e.g., Smeets et al., 1998; Conway and Cullen, 2013), while the latter finding is corroborated with the surface renewal theory of Andreas (1987). Similar estimates for  $z_{0v,0t}$  from the two different summers at the same site, gives some confidence in our method used to determine roughness lengths.

Taking into account small sample size, the absence of any model for predicting a time-varying roughness lengths, and the lack of any significant changes in the surface properties during the observational period, we used constant values for  $z_{0v,0t,0q}$  throughout each season. The uncertainty in the modeled turbulent fluxes due to potentially time-varying roughness lengths was quantified via Monte Carlo sampling ( $\delta Q_{z_0}$ ) and was found to be smaller than the RMSE between modeled and observed fluxes. The modeled fluxes are thus poorly simulated due to the poorly defined parameterizations, rather than the poorly constrained roughness lengths. In our Monte Carlo assessment of  $\delta Q_{z_0}$ , we assumed the errors in 30-min  $z_{0v,0t,0q}$  to be random and small, rather than having a systematic error due to any potential changes in the actual surface roughness (e.g. snowfall; development of a surface drainage system; inhomogeneous surface ablation). The assumed random errors are attributed to instrumental (measurement) error and to the choice of the threshold stability value for the near-neutral conditions (i.e.  $|\frac{z_{v,t}}{L}| < 0.1$ ). We posit that the large variance error (the scatter in the graphs showing 30-min modeled versus observed turbulent fluxes) arises from the variability in the meteorological mean variables (temperature, wind speed, local stability) rather than the variability in the roughness lengths in time. If, instead of the constant mean log values for  $z_{0v,0t,0q}$ , we use their observed values specific to each 30-min segment with the near-neutral conditions, the resulting overfit would not improve the performance over all the points because: (i) the overfitted points represent  $< 5$  % of points in the observation period and; (ii) the model performs poorly outside near-neutral conditions.

3. When testing the stability corrections in the bulk methods, we assumed that the main predictor of the time-varying bulk exchange coefficient is the time-varying local stability, not the changes in the surface roughness. Using the two common parameterizations to account for the stability corrections ( $C_{Rib}$  and  $C_{M-O}$  method) we found that neither scheme significantly improves estimates of turbulent fluxes relative to the scheme without the stability corrections ( $C_{log}$  method). Our analysis suggests two reasons for this: (i) the empirical stability functions ( $\Psi_{v,t,q}$ ) taken from the literature perform poorly on our data and; (ii) if not directly measured via the eddy-covariance system,  $\frac{z_{v,t}}{L}$  metric is impossible to accurately obtain because no functional relationship is found between the metric and the mean meteorological variables (e.g. temperature and wind speed). The latter reason may also explain why measured  $Ri_b$  could not be expressed as a function of OPEC-derived  $\frac{z_{v,t}}{L}$ . Our findings corroborate previous experimental and theoretical evidence that the M-O theory is poorly applicable over sloping glacier surfaces (Denby and Greuell, 2000). During katabatic flow, the bulk methods that rely on the M-O stability theory strongly overestimate the momentum flux and less strongly overestimate the heat flux. These biases in modeled fluxes are expected in the presence of low-level wind speed maxima that are characteristic during katabatic flow (Denby and Greuell, 2000). We cannot test this expectation since we lack observations of vertical wind profiles for our study, however. The  $K_{int}$  method, based on a simple katabatic model, best simulates the friction

velocity among all the methods in the study. This method under-performs any other method in simulating  $Q_H$ , however, revealing a shortcoming in the assumption that the eddy diffusivity ( $K_H$ ) and eddy viscosity ( $K_M$ ) share the same parametrization. One way to improve for this shortcoming, as indicated in our results, would be to introduce a vertically varying eddy Prandtl number,  $Pr(z)$ .

- 5 4. Under stable stratification and in the presence of a low level jet (katabatic flow), intermittent turbulence and gravity waves are present, and steady-state conditions do not exist (e.g., Foken, 2008; Axelsen and van Dop, 2009). We removed the data in our study that failed to satisfy the criteria of stationarity ('stationarity' filter). Assuming that our filtered data sample reflects the steady-state conditions, we found that the  $K_{Int}$  method can simulate the 30-min  $u_*$  better than the C-methods. Improved performance is due to the parametrization of  $K(z)$  which, being independent of  $u_*$ , yields the bulk method with  $u_* \propto \sqrt{U_z}$ . This function  $u_*(U_z)$  yields a better fit to the observations than the function  $u_* \propto U_z$  characteristic for the C-methods. Despite the absence of vertical profile measurements to adequately test the model for  $K(z)$ , our derivation of  $K_{max}$  as a function of a proxy variable for environmental static stability (i.e. the difference between off-glacier and on-glacier near-surface air temperature), agrees well with the expected dependencies of this parameter on the background potential temperature lapse rate,  $\gamma$  (Parmhed et al., 2004). While our results are promising in terms of the applicability of the  $K_{Int}$ -approach to determine the friction velocity, it remains unclear how our static stability variable relates to  $\gamma$ . A possible extension of our work would be to investigate whether a more exact parameterization for  $K$  (as a function of  $z$ ,  $\frac{\partial U}{\partial z}$ , and  $\frac{\partial T}{\partial z}$ ) in the surface layer can be determined as a solution to the governing equations (heat and momentum balance) for katabatic flow.

## 5 Conclusions

- 20 The main objective of the study was to evaluate commonly used bulk methods (C-methods) for simulating turbulent heat fluxes at a sloped glacier surface and compare these with approaches based on a simple katabatic flow model. We investigated the stability components of the bulk methods, and attempted to improve upon these, assuming site-specific roughness length values are available. In addition to the four C-methods ( $C_{log}$ ,  $C_{Rib}$ ,  $C_{M-O}$  and  $C_{SR}$  method) with different parameterizations for the bulk exchange coefficient ( $C$ ), we evaluated two types of bulk methods developed from katabatic flow models ( $C_{kat}$  and  $K_{Int}$  method; Grisogono and Oerlemans, 2001; Oerlemans and Grisogono, 2002). These six bulk schemes, for simulating 30-min turbulent fluxes of momentum, sensible and latent heat, were evaluated against the equivalent fluxes obtained by an open path eddy-covariance (OPEC) method. The evaluation was performed at a point scale of a mountain glacier in BC, using one-level meteorological and OPEC observations over multiple days for two melt seasons. The OPEC data were subjected to a set of quality control corrections and filters prior to analysis. Few 30-min data segments met the criteria for the near-neutral static stability and the steady-state turbulence conditions, revealing a challenge to perform this type of analysis at sites dominated by the stable conditions and drainage (katabatic) flows. Despite the small OPEC data sample available to determine surface roughness lengths ( $z_{0v,0t,0q}$ ) and the high-quality 30-min turbulent fluxes, the evaluation results derived from the two independent field seasons accord. A summary of our main findings include:

Bulk exchange coefficients can be derived from the OPEC-derived roughness lengths for neutral stratification, where the mean log values for  $z_{0v,0t,0q}$  over the entire period are assumed representative of the actual surface roughness and constant in time. The widely used bulk methods facilitated by the standard meteorological observations on glaciers, provide a good approximation of the OPEC-derived 30-min turbulent fluxes during the neutral stability conditions;

5 The Monin-Obukhov stability functions widely used in glacier studies perform relatively poorly on our site. The C-methods, with or without the stability corrections, overestimate the sensible heat fluxes ( $Q_H$ ) over the observational period. This overestimation arises from a positive mean bias in the modeled friction velocity ( $u_*$ ), and the bias increases with the katabatic wind speed. Without an adequate model for  $u_*$ , the use of the stability corrections for  $Q_H$  will likely fail. Similar findings derived from a different glacier in this region (Fitzpatrick et al., 2017) suggests the transferability  
10 of this finding to other mountain glaciers subject to similar climate;

If OPEC data are not available,  $\frac{z_{v,t}}{L}$  cannot be successfully predicted from the mean meteorological observations (e.g. wind speed, air-surface temperature difference). The proposed fixed-point iterative method for predicting  $\frac{z_{v,t}}{L}$  suggested by others (Berkowicz and Prahm, 1982; Lee, 1986; Munro, 1989) fails to provide a good approximation of OPEC-derived  $\frac{z_{v,t}}{L}$  at our site;

15 **Contrary to the K-theory predictions** (Stull, 1988), we found that the relation between the OPEC-derived 30-min  $u_*$  and observed wind speed is non linear. Our data reveal  $u_* \propto \sqrt{U_z}$ , where the constant of proportionality is a nonlinear function of the environmental static stability;

The  $K_{Int}$  method outperforms any other bulk scheme in simulating  $u_*$  partly because  $u_* \propto \sqrt{U_z}$ . The  $K_{Int}$  method performs poorly in simulating  $Q_H$ , however, revealing a shortcoming in the commonly used assumption that eddy viscosity  
20 and diffusivity can use the same parametrization. Applying the  $K_{Int}$  method to assess  $u_*$ , which is then used to assess  $Q_H$  in the  $C_{log}$  method, yields the best performance across all the bulk methods. The 'hybrid' method, combining a katabatic model to derive the friction velocity and a simple C-method to derive  $Q_H$ , works the best in simulating the sensible heat fluxes; and

25 **The bulk exchange coefficient in the  $K_{Int}$  method is related to a difference between off-glacier and on-glacier near-surface air temperature, which is a proxy variable for the background potential temperature lapse rate ( $\gamma$ ) in the original model of Grisogono and Oerlemans (2001).** As expected by the original model and confirmed by our results, the bulk exchange coefficient decreases as the atmospheric stability increases (larger  $\gamma$  in the model or larger temperature difference). Multi-level observations of wind speed and temperature within the boundary layer above a glacier surface are needed to adequately evaluate this model, however.

30 *Author contributions.* VR participated in the 2012 field campaign, developed and performed the analysis, and wrote the initial version of the manuscript. BM provided financial and field support and helped prepare and revise the manuscript. JS led the 2010 field campaign and

contributed to the manuscript refinement. NF processed the OPEC data and contributed to the manuscript refinement. MAT contributed to the method development. SJD provided logistical support, meteorological data and contributed to the manuscript refinement.

*Acknowledgements.* Funding supporting this study was provided through the Natural Sciences and Engineering Research Council (NSERC) of Canada (Discovery grants to V. Radić, B. Menounos, S.J. Déry), the Canada Research Chairs Program and Science Horizons. Our eddy  
5 covariance system and meteorological equipment were supported by NSERC Research Tools and Instruments grant and Canada Foundation for Innovation grant (S.J. Déry). T. Mlynowski is thanked for the field setup in 2010, and Andrew Duncan for the field support in 2012. Special thanks to Zoran Nesić for all the technical support and training. Branko Grisogono, Christian Schoof, Elisa Mantelli and the two reviewers, Jonathan Conway and Ruzica Dadic, are thanked for discussions and constructive criticism on the initial version of the manuscript.

## References

- Anderson, B., Mackintosh, A., Stumm, D., George, L., Kerr, T., Winter-Billington, A., and Fitzsimons, S.: Climate sensitivity of a high-precipitation glacier in New Zealand, *Journal of Glaciology*, 56, 114–128, 2010.
- Andreas, E.: A theory for the scalar roughness and the scalar transfer coefficient over snow and sea ice, *Boundary-Layer Meteorology*, 38, 159–184, doi:10.1007/BF00121562, 1987.
- Andreas, E. L., Persson, P. O. G., Jordan, R. E., Horst, T. W., Guest, P. S., Grachev, A. A., and Fairall, C. W.: Parameterizing Turbulent Exchange over Sea Ice in Winter, *Journal of Hydrometeorology*, 11, 87–104, doi:10.1175/2009JHM1102.1, 2010.
- Arya, S. P.: *Introduction to Micrometeorology*, Academic Press, San Diego, 415 pp, 2001.
- Axelsen, S. L. and van Dop, H.: Large-eddy simulation of katabatic winds. Part 1: Comparison with observations, *Acta Geophysica*, 57, 803–836, doi:10.2478/s11600-009-0041-6, 2009.
- Beedle, M. J., Menounos, B., and Wheate, R.: An evaluation of mass-balance methods applied to Castle Creek Glacier, British Columbia, Canada, *Journal of Glaciology*, 60, 262–276, doi:10.3189/2014JoG13J091, 2014.
- Berkowicz, R. and Prahm, L. P.: Evaluation of the profile method for estimation of surface fluxes of momentum and heat, *Atmos. Environ.*, 16, 2809–2819, 1982.
- Bevington, P. R.: *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York, 1969.
- Bintanja, R. and van den Broeke, M.: Momentum and scalar transfer coefficients over aerodynamically smooth Antarctic surfaces, *Boundary-Layer Meteorology*, pp. 89–111, 1995.
- Braithwaite, R.: Aerodynamic stability and turbulent sensible-heat flux over a melting ice surface, the Greenland ice sheet, *Journal of Glaciology*, 41, 562–571, 1995.
- Braithwaite, R., Konzelmann, T., Marty, C., and Olesen, O.: Reconnaissance study of glacier energy balance in North Greenland, 1993–94, *Journal of Glaciology*, 44, 239–247, 1998.
- Burba, G.: *Eddy Covariance Method for Scientific, Industrial, Agricultural, and Regulatory Applications: A Field Book on Measuring Ecosystem Gas Exchange and Areal Emission Rates*, LI-COR Biosciences, Lincoln, NE, USA, 331 pp, 2013.
- Businger, J., Wyngaard, J., Izumi, Y., and Bradley, E.: Flux-profile relationships in the atmospheric surface layer, *Journal of Atmospheric Sciences*, 28, 181–&, doi:10.1175/1520-0469(1971)028<0181:FPRITA>2.0.CO;2, 1971.
- Clarke, G. K. C., Jarosch, A. H., Anslow, F. S., Radić, V., and Menounos, B.: Projected deglaciation of western Canada in the twenty-first century, *Nature Geoscience*, 8, 372–377, doi:10.1038/NGEO2407, 2015.
- Conway, J. and Cullen, N.: Constraining turbulent heat flux parameterization over a temperate maritime glacier in New Zealand, *Ann. Glaciol.*, 54, 41–51, doi:10.3189/2013AoG63A604, 2013.
- Cullen, N. J., Mölg, T., Kaser, G., Steffen, K., and Hardy, D. R.: Energy-balance model validation on the top of Kilimanjaro, Tanzania, using eddy covariance data, *Annals of Glaciology*, 46, 227–233, doi:10.3189/172756407782871224, 2007.
- Denby, B. and Greuell, W.: The use of bulk and profile methods for determining surface heat fluxes in the presence of glacier winds, *Journal of Glaciology*, 46, 445–452, doi:10.3189/172756500781833124, 2000.
- Déry, S. J., Clifton, A., MacLeod, S., and Beedle, M. J.: Blowing Snow Fluxes in the Cariboo Mountains of British Columbia, Canada, *Arctic Antarctic and Alpine Research*, 42, 188–197, doi:10.1657/1938-4246-42.2.188, 2010.
- Dyer, A.: A review of flux-profile-relationships, *Boundary-Layer Meteorology*, 7, 363–372, 1974.

- Fairall, C., Persson, P., Bradley, E., Payne, R., and Anderson, S.: A new look at calibration and use of Eppley precision infrared radiometers. Part I: Theory and application, *Journal of Atmospheric and Oceanic Technology*, 15, 1229–1242, doi:10.1175/1520-0426(1998)015<1229:ANLACA>2.0.CO;2, 1998.
- Fitzpatrick, N., Radić, V., and Menounos, B.: Surface Energy Balance Closure and Turbulent Flux Parameterization on a Mid-Latitude Mountain Glacier, Purcell Mountains, Canada, *Frontiers in Earth Science*, 5, doi:10.3389/feart.2017.00067, 2017.
- Foken, T.: *Micrometeorology*, Springer, Berlin, Heidelberg, 308 pp, 2008.
- Forrer, J. and Rotach, M.: On the turbulence structure in the stable boundary layer over the Greenland ice sheet, *Boundary-Layer Meteorology*, 85, 111–136, doi:10.1023/A:1000466827210, 1997.
- Gillett, S. and Cullen, N. J.: Atmospheric controls on summer ablation over Brewster Glacier, New Zealand, *International Journal of Climatology*, 31, 2033–2048, doi:10.1002/joc.2216, 2011.
- Grachev, A. A., Andreas, E. L., Fairall, C. W., Guest, P. S., and Persson, P. O. G.: On the turbulent Prandtl number in the stable atmospheric boundary layer, *Boundary-Layer Meteorology*, 125, 329–341, doi:10.1007/s10546-007-9192-7, 2007.
- Greuell, J. W., V. d. B. M. R. K. W. R. C. S. P. and Struijk, I.: PASTEX: A Glacio-Meteorological Experiment on the Pasterze (Austria), in: *Field Report*. Institute for Marine and Atmospheric Research, Utrecht University, p. 34, 1994.
- Greuell, W. and Oerlemans, J.: Energy balance calculations on and near Hintereisferner (Austria) and an estimate of the effect of greenhouse warming on ablation., in: *Glacier fluctuations and climatic change*, edited by Oerlemans, J, vol. 6 of *Glaciology and Quaternary Geology*, pp. 305–323, 1989.
- Grisogono, B. and Oerlemans, J.: Katabatic flow: Analytic solution for gradually varying eddy diffusivities, *Journal of the Atmospheric Sciences*, 58, 3349–3354, doi:10.1175/1520-0469(2001)058<3349:KFASFG>2.0.CO;2, 2001.
- Guo, X., Yang, K., Zhao, L., Yang, W., Li, S., Zhu, M., Yao, T., and Chen, Y.: Critical Evaluation of Scalar Roughness Length Parametrizations Over a Melting Valley Glacier, *Boundary-Layer Meteorology*, 139, 307–332, doi:10.1007/s10546-010-9586-9, 2011.
- Harrison, R.: Natural ventilation effects on temperatures within Stevenson screens, *Q. J. R. Meteorol. Soc.*, 136, 253–259, doi:10.1002/qj.537, 2010.
- Hay, J. and Fitzharris, B.: A comparison of the energy-balance and bulk-aerodynamic approaches for estimating glacier melt, *Journal of Glaciology*, 34, 145–153, 1988.
- Hock, R.: Glacier melt: a review of processes and their modelling, *Progress in physical geography*, 29, 362–391, doi:10.1191/0309133305pp453ra, 2005.
- Hock, R. and Holmgren, B.: Some aspects of energy balance and ablation of Storglaciaren, northern Sweden, *Geografiska Annaler*, 78A, 121–131, doi:10.2307/520974, 1996.
- Hock, R. and Holmgren, B.: A distributed surface energy-balance model for complex topography and its application to Storglaciaren, Sweden, *Journal of Glaciology*, 51, 25–36, doi:10.3189/172756505781829566, 2005.
- Holtslag, A. and de Bruin, H.: Applied modeling of the nighttime surface energy balance over land, *Journal of Applied Meteorology*, 27, 689–704, doi:10.1175/1520-0450(1988)027<0689:AMOTNS>2.0.CO;2, 1988.
- Horst, T. W. and Lenschow, D. H.: Attenuation of Scalar Fluxes Measured with Spatially-displaced Sensors, *Boundary-Layer Meteorology*, 130, 275–300, doi:10.1007/s10546-008-9348-0, 2009.
- Huintjes, E., Neckel, N., Hochschild, V., and Schneider, C.: Surface energy and mass balance at Purogangri ice cap, central Tibetan Plateau, 2001–2011, *Journal of Glaciology*, 61, 1048–1060, doi:10.3189/2015JoG15J056, 2015.



- Hulth, J., Rolstad, C., Trondsen, K., and Rodby, R. W.: Surface mass and energy balance of Sorbreen, Jan Mayen, 2008, *Annals of Glaciology*, 51, 110–119, 2010.
- Huwald, H., Higgins, C. W., Boldi, M.-O., Bou-Zeid, E., Lehning, M., and Parlange, M. B.: Albedo effect on radiative errors in air temperature measurements, *Water Resources Research*, 45, doi:10.1029/2008WR007600, 2009.
- 5 Ibrom, A., Dellwik, E., Flyvbjerg, H., Jensen, N. O., and Pilegaard, K.: Strong low-pass filtering effects on water vapour flux measurements with closed-path eddy correlation systems, *Agricultural and Forest Meteorology*, 147, 140–156, doi:10.1016/j.agrformet.2007.07.007, 2007.
- Jarosch, A. H., Anslow, F., and Shea, J.: A versatile tower platform for glacier instrumentation: GPS and Eddy Covariance Measurements, pp. 52–55, extended abstracts and recommendations from the IASC Workshop on the Use of Automatic Measuring Systems on Glaciers, 10 23–26 March 2011, Pontresina, Switzerland, 2011.
- Klok, E. J., Nolan, M., and Van den Broeke, M. R.: Analysis of meteorological data and the surface energy balance of McCall Glacier, Alaska, USA, *Journal of Glaciology*, 51, 451–461, doi:10.3189/172756505781829241, 2005.
- Konzelmann, T. and Braithwaite, R.: Variations of ablation, albedo and energy balance at the margin of the Greenland ice sheet, Kronprins Christian Land, eastern north Greenland, *Journal of Glaciology*, 41, 174–182, 1995.
- 15 Lee, C. B.: Simple model and climatological aspects of the structure of the convective boundary layer, *Atmos. Environ.*, 20, 705–714, 1986.
- Lettau, H.: *Atmosphärische Turbulenz*, Leipzig: Akademische Verlagsgesellschaft., 1934.
- Li, Z., Lyu, S., Zhao, L., Wen, L., Ao, Y., and Wang, S.: Turbulent transfer coefficient and roughness length in a high-altitude lake, Tibetan Plateau, *Theoretical and Applied Climatology*, 124, 723–735, doi:10.1007/s00704-015-1440-z, 2016.
- MacDougall, A. H. and Flowers, G. E.: Spatial and Temporal Transferability of a Distributed Energy-Balance Glacier Melt Model, *Journal of Climate*, 24, 1480–1498, doi:10.1175/2010JCLI3821.1, 2011.
- 20 Machguth, H., Paul, F., Hoelzle, M., and Haeberli, W.: Distributed glacier mass-balance modelling as an important component of modern multi-level glacier monitoring, *Annals of Glaciology*, 43, 335–343, doi:10.3189/172756406781812285, 2006.
- Mölg, T., Cullen, N. J., Hardy, D. R., Kaser, G., and Klok, L.: Mass balance of a slope glacier on Kilimanjaro and its sensitivity to climate, *International Journal of Climatology*, 28, 881–892, doi:10.1002/joc.1589, 2008.
- 25 Mölg, T., Cullen, N. J., Hardy, D. R., Winkler, M., and Kaser, G.: Quantifying Climate Change in the Tropical Midtroposphere over East Africa from Glacier Shrinkage on Kilimanjaro, *Journal of Climate*, 22, 4162–4181, doi:10.1175/2009JCLI2954.1, 2009.
- Moncrieff, J., Clement, R., Finnigan, J., and Meyers, T.: Averaging, detrending, and filtering of eddy covariance time series, In: *Handbook of Micrometeorology: A guide for Surface Flux Measurement and Analysis* (eds X. Lee, W. Massman, B. Law, X. Lee, W. Massman, and B. Law). Kluwer Academic Publishers, the Netherlands, pp. 7–31, 2004.
- 30 Munro, D.: Surface roughness and bulk heat transfer on a glacier: comparison with eddy correlation, *Journal of Glaciology*, 35, 343–348, 1989.
- Munro, D.: A surface energy exchange model of glacier melt and net mass balance, *International Journal of Climatology*, 11, 689–700, 1991.
- O’Brien, J. J.: A note on the vertical structure of the eddy exchange coefficient in the planetary boundary layer, *J. Atmos. Sci.*, 27, 1213–1215, 1970.
- 35 Oerlemans, J.: The atmospheric boundary layer over melting glaciers, In: *Clear and Cloudy Boundary Layers*, Royal Netherlands Academy of Arts and Sciences, 129–153, 1998.
- Oerlemans, J.: Analysis of a 3 year meteorological record from the ablation zone of Morteratschgletscher, Switzerland: energy and mass balance, *Journal of Glaciology*, 46, 571–579, doi:10.3189/172756500781832657, 2000.

- Oerlemans, J. and Grisogono, B.: Glacier winds and parameterisation of the related surface heat fluxes, *Tellus Series A - Dynamic Meteorology and Oceanography*, 54, 440–452, doi:10.1034/j.1600-0870.2002.201398.x, 2002.
- Oerlemans, J., Bjornsson, H., Kuhn, M., Obleitner, F., Pálsson, F., Smeets, C., Vugts, H., and De Wolde, J.: Glacio-meteorological investigations on Vatnajökull, Iceland, Summer 1996: An overview, *Boundary-Layer Meteorology*, 92, 3–26, doi:10.1023/A:1001856114941, 5 1999.
- Parmhed, O., Oerlemans, J., and Grisogono, B.: Describing surface fluxes in katabatic flow on Breidamerkurjökull, Iceland, *Quarterly Journal of the Royal Meteorological Society*, 130, 1137–1151, doi:10.1256/qj.03.52, 2004.
- Prandtl, L.: *The mechanics of viscous fluids (Aerodynamic theory)*, Berlin: Durand (ed.), vol. III, division G., 1934.
- Prandtl, L.: *Führer durch die Strömungslehre*, Vieweg und Sohn, 648 pp., 1942.
- 10 Prinz, R., Nicholson, L. I., Mölg, T., Gurgiser, W., and Kaser, G.: Climatic controls and climate proxy potential of Lewis Glacier, Mt. Kenya, *The Cryosphere*, 10, 133–148, doi:10.5194/tc-10-133-2016, 2016.
- Radić, V. and Hock, R.: 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, 2014.
- Radić, V., Bliss, A., Beedlow, A. C., Hock, R., Miles, E., and Cogley, J. G.: Regional and global projections of twenty-first century glacier 15 mass changes in response to climate scenarios from global climate models, *Climate Dynamics*, 42, 37–58, doi:10.1007/s00382-013-1719-7, 2014.
- Reijmer, C. H. and Hock, R.: Internal accumulation on Storglaciären, Sweden, in a multi-layer snow model coupled to a distributed energy- and mass-balance model, *Journal of Glaciology*, 54, 61–72, doi:10.3189/002214308784409161, 2008.
- Scotanus, P., Nieuwstadt, F., and Debruin, H.: Temperature measurement with a sonic anemometer and its application to heat and moisture 20 fluxes, *Boundary-Layer Meteorology*, 26, 81–93, doi:10.1007/BF00164332, 1983.
- Sicart, J., Wagnon, P., and Ribstein, P.: Atmospheric controls of the heat balance of Zongo Glacier (16 degrees S, Bolivia), *Journal of Geophysical Research-Atmospheres*, 110, doi:10.1029/2004JD005732, 2005.
- Sicart, J. E., Hock, R., Ribstein, P., Litt, M., and Ramirez, E.: Analysis of seasonal variations in mass balance and meltwater discharge of the tropical Zongo Glacier by application of a distributed energy balance model, *Journal of Geophysical Research-Atmospheres*, 116, 25 doi:10.1029/2010JD015105, 2011.
- Smeets, C., Duynkerke, P., and Vugts, H.: Turbulence characteristics of the stable boundary layer over a mid-latitude glacier. Part I: A combination of katabatic and large-scale forcing, *Boundary-Layer Meteorology*, 87, 117–145, doi:10.1023/A:1000860406093, 1998.
- Smeets, C. J. P. P. and van den Broeke, M. R.: The parameterisation of scalar transfer over rough ice, *Boundary-Layer Meteorology*, 128, 339–355, doi:10.1007/s10546-008-9292-z, 2008.
- 30 Stull, R.: *An Introduction to Boundary Layer Meteorology*, Kluwer Acad. Publ., Dordrecht, Boston, London, 666 pp., 1988.
- Tanner, B. D., Swiatek, E., and Greene, J. P.: Density fluctuations and use of the krypton hygrometer in surface flux measurements, in: *Management of irrigation and drainage systems: integrated perspectives*, edited by: Allen, R. G., American Society of Civil Engineers, New York, NY, 1993.
- van den Broeke, M.: Characteristics of the lower ablation zone of the west Greenland ice sheet for energy-balance modelling, *Annals of 35 Glaciology*, 23, 160–166, 1996.
- van den Broeke, M., Van As, D., Reijmer, C., and Van De Wal, R.: Sensible heat exchange at the antarctic snow surface: A study with automatic weather stations, *International Journal of Climatology*, 25, 1081–1101, doi:10.1002/joc.1152, 2005.

- van der Avoird, E. and Duynkerke, P.: Turbulence in a katabatic flow - Does it resemble turbulence in stable boundary layers over flat surfaces?, *Boundary-Layer Meteorology*, 92, 39–66, 1999.
- Wagnon, P., Sicart, J., Berthier, E., and Chazarin, J.: Wintertime high-altitude surface energy balance of a Bolivian glacier, Illimani, 6340 m above sea level, *Journal of Geophysical Research-Atmospheres*, 108, doi:10.1029/2002JD002088, 2003.
- 5 Webb, E., Pearman, G., and Leuning, R.: Correction of the flux measurements for density effects due to heat and water vapour transfer, *Quarterly Journal of the Royal Meteorological Society*, 106, 85–100, doi:10.1002/qj.49710644707, 1980.
- Webb, E. K.: Profile relationships: the log-linear range, and extension to strong stability, *Quarterly Journal of the Royal Meteorological Society*, 96, 67–&, doi:10.1002/qj.49709640708, 1970.
- Wilczak, J., Oncley, S., and Stage, S.: Sonic anemometer tilt correction algorithms, *Boundary-Layer Meteorology*, 99, 127–150, doi:10.1023/A:1018966204465, 2001.
- 10