Referee review for manuscript tc-2017-80, under review for The Cryosphere.

"EVALUATION OF DIFFERENT METHODS TO MODEL NEAR-SURFACE TURBULENT FLUXES FOR AN ALPINE GLACIER IN THE CARIBOO MOUNTAINS, BC, CANADA" by Valentina Radic, Brian Menounos, Joseph Shea, Noel Fitzpatrick, Mekdes A. Tessema, and Stephen J. Déry.

Reviewer: J Conway

General Comments

This manuscript addresses the turbulent fluxes of sensible (*H*) and latent heat (*LE*) over small mountain glaciers using in-situ measurements and theoretical models. The topic is not only extremely relevant to efforts to resolve cryosphere-climate relationships, but has many unresolved questions. The manuscript is novel in that it presents new measurements that add to the relatively small body of literature on turbulent fluxes in these environments. The roughness length parameter used in the most common Monin-Obukhov (M-O) bulk approaches to modelling these fluxes is successfully derived from measured fluxes in ideal neutral-stability conditions. The derived values agree well with other estimates from the literature, with the roughness length for momentum being 2 order of magnitude larger than that for temperature. The authors also derive the roughness length for humidity – a seldom performed activity – and show it to be of the same order as temperature. The authors then delve into modelling *H* with the M-O bulk method, showing a general overestimation of *H* and the friction velocity (u^*). The authors test various methods to represent *H* and u^* using a combination of measured and modelled turbulence variables within a variety of theoretical frameworks. Some of these give promising results, but overall, no theoretical framework is markedly better at modelling *H* using only mean wind speed and temperature at one height above the glacier surface. The best method presented is the simplest, relating *H* to a quadratic of the 2-meter air temperature (Oerlemans and Grisogono, 2002). The measured values of u^* and the M-O stability parameter (z/L) are shown to be poorly predicted by bulk models. If u^* is known, then *H* can be predicted well by M-O bulk schemes.

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.

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 convincing 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.

Most of the analyses are clear, appropriate and well supported by theory. The analyses of the new stability function and K_{int}, 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.

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.

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.

In summary, the manuscript should make a good contribution to the literature on this subject if a number of issues are addressed.

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 K_{Int} 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".

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.

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 analyses specifically, so it is hard to see this a key result. Please revise.

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.

2-15: References needed 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.

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).

3-12: *z0v* 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 *z0v*, while the surface roughness elements remain constant. Please revise.

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

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

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.

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.

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.

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

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

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 K_{Int} later as a separate method?

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

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.

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.

26-14: Is there a way to evaluate K_{int}/K_{max} without measured z/L? For consistency, it would be useful to discuss if this is possible.

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.

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

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) .

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.

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.

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

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

33-22: "Applying the K_{Int} 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 K_{int} 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.

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

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.

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

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.

Editorial Comments

2-9: "recourses" -> "resources"

4-28: "is monitored" -> "has been monitored"

4-31: -> "In the glacier vicinity, two year-round automatic weather stations have been in operation since 2007/2008 (Déry et al., 2010). The stations are situated on the lateral and terminal moraines, and are referred to as AWSup …. And AWSlow…. respectively."

5-12: Do you mean AWSlow?

7-5: "potentially high and low frequency loss," -> "potential loss of high and low frequency signal"

17-3: -> "Obukhov length (*L*)"

23-7: -> "for 2010 and 2012" or "for 2012 and 2010".

26-9: -> "best estimate of the friction velocity or momentum flux among all the bulk schemes we tested so far (compare Fig. 11 to Fig. 5)." Need to help the reader to navigate between the results.

26-23: It would be consistent with the presentation of other schemes to include (22) and (23) before (18) and (19), respectively. i.e. first using iterative z/L, then using measured. The same for (13) and (14).

26-23: "empirical function for K_{max}" and "empirical function for H_k" - please refer to equations 35 and 36 here.

27-1: -> "hybrid methods drops (Table 3)."

30-34: -> "resolve issue (i)"

33-3: -> "modeled friction velocity"