Comment on gmd-2021-272
Anonymous Referee #2

Referee comment on "Validation of turbulent heat transfer models against eddy covariance flux measurements over a seasonally ice covered lake" by Joonatan Ala-Könni et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-272-RC2, 2021

The authors consider turbulence measurements over a lake in Finland. They compare different bulk paranetrizations of the fluxes with Eddy Covariance measurements, which had been obtained during four winter episodes. They find small differences between all parametrization models and conclude that the best agreement with the EC measurements is found for a method using a constant transfer coefficient without any stability correction.

The paper is well written and clearly organized. However, I think some improvement is necessary before its publication.

Major Revisions

1) Equation (14) is not correct. A logarithm is missing in the first term of the r.h.s. Please check your codes if they are correct (with logarithm).

2) The main conclusion (see abstract of the manuscript) is that the assumption of a constant transfer coefficient being independent on stratification is the best one. This needs much better explanation. Modellers might get the idea to ignore the stability correction in their runs in general. But this is against all previous experience over decades from observations, theory, and Large Eddy Simulation. So, if this is really the result, then the reader must be better convinced that it is not an artefact. Possible reasons might be conditions violating Monin Obukhov similarity or problems with the accuracy of measurements (e.g. due to influences of the boxes near the small 'tower') and many others.

To better convince the reader I find it necessary to show results (fluxes) obtained by the EC method as a function of z/L or of the Richardson number (as in Grachev et al., 2007) or in many other papers (e.g. most recently Srivastava, Gryanik et al.). It would be helpful to show results of the phi- function (eqs. (9,10) and that behind the SHEBA equation (14) (see Grachev et al., 2007) as a function of z/L.

3) It is difficult to interprete the differences between all three schemes based only on the scatter plots (Figure 5).

4) When the final result remains unchanged, it needs to be explained more careful. It
should be written that further research is necessary to test the robustness of this result. E.g. measurements over other lakes are necessary before a general suggestion to modellers can be given. Such results could depend on the lake size, where the flow over small lakes might be more inhomogeneous than the flow over large lakes and inhomogeneity might hide the stability dependence.

5) The underestimation of fluxes might be due to errors in the roughness lengths (especially the ratio between roughness length for momentum and for scalars is uncertain).

**Minor Revisions**

1) The introduction is interesting but the description of main goals and their explanation are rather short.

2) Line 66 with 'the' bulk...

3) line 78: I would not say that the EC measurements really overcome heterogeneity problems.

4) line 82: present 'a' unique

5) Considering figure 1 and description, I was not sure if data were selected with wind directed only along the fjord to minimize inhomogeneity impact.

6) equation (9): insert space between zeta and 'for'

7) Sometimes, the article is missing, e.g. line 247 (The surface energy ...)

8) Line 247: better give equation for the energy balance

9) Figure caption 3. Sign of flux vice versa as compared with line 280 ?

10) Line 13: is the underestimation a problem due to the phi-function or due to the roughness assumption?

11) Line 295 and similar .... better: results over the melting surface

12) Line 310: patterns

13) paragraph starting at line 388: This comes very late although it is the main result (see abstract and above).

**References**
