This author response to comments from Anonymous Referee #3 is structured as follows:

Referee comments

Author response

Changes in manuscript or references to "tc-2019-8_MSchanges.pdf" where these changes can be found

This manuscript presents a comparison of surface energy balance for a water track and for two reference locations in the Taylor Valley during 26 days in summer of 2012-2013. The main aim of the study was to evaluate the hypothesis that water tracks alter the surface energy budget, mainly due to the higher moisture content of the soil. The energy balance estimation relied on eddy-covariance (EC) method, net radiation measurements and soil temperature profiles. I am not an expert on these cold ecosystems, yet I do have some experience on the EC method and will concentrate on EC related issues in my review. It seems that the other two reviewers are experts on the Antarctic ecosystems and hence our fields of expertise complement each other. In general the paper is well-written and it has a clear structure which is easy to follow. The available data set is not large, only 26 days, and hence far reaching conclusions cannot be drawn based on it. Nevertheless, the authors do their best and it should be also acknowledged that this is the first EC study in this remote location. However, there is one major flaw in this study which is related to forcing the energy balance to be closed. This should be changed before the manuscript can be accepted for publication. Please see more details below. In my view the manuscript is otherwise good and hence I recommend accepting it for publication after minor revisions suggested below, in addition to the modifications suggested by the other two reviewers.

We thank Referee #3 for appreciating the presentational quality of our results. It was very helpful to get expert feedback on our methodology and we could improve methodological description and discussion thanks to your comments and suggestions. Your comments on the limitations of the interpretations of our findings because of seasonal differences and shortness of the recording period helped us improve our discussions, although we captured the major part of the extremely short Antarctic summer and are confident that we can draw conclusions from our measurements to some degree. We appreciate your critical assessment of our assumption of closed energy balance. However, we claim that we can keep our energy balance concept because the results match expected physical behavior. Forcing a closed energy balance is therefore rather a limitation to our interpretations than a per se incorrect assumption.

Specific Comments

1) The authors estimated QIT (the energy input to the permafrost) as a residual of the surface energy balance, meaning that they force the energy balance to be closed. However, there is a plethora of publications out there that show that at EC sites the energy balance is almost

never closed, meaning that incoming energy exceeds the sum of outgoing energy and energy stored in the system (see e.g. Reed et al., 2018; Stoy et al., 2013; Hendriks Franssen et al., 2010; Leuning et al., 2012). Usually 10-30 % of the energy is missing. It is still unclear why the energy balance is not typically closed at EC sites, yet it is often hypothesized to be related to sampling mismatch between instruments, instrumental problems, under sampled large eddies transporting heat, terrain heterogeneity or energy storage in soils, air and biomass below the measurement height (Wilson et al., 2002; Leuning et al., 2012). The reasons might also be different for different sites. Be that as it may, this issue should be acknowledged also in this manuscript in question. Hence, I argue that the energy balance residual cannot be assumed to be equal to QIT and the authors should analyze it only as energy balance residual, not some specific real energy flux.

Thank you for the comment. We are, of course, aware of the residual in the surface energy balance commonly observed for eddy-covariance based evaluations of it and have ourselves published on it. This residual certainly has implications for estimating the energy which may go toward deepening the thawed layer by moving the ice-table depth further into the ground toward the permafrost layer depth as the summer season progresses. As you correctly mentioned, almost all experimental studies have reported a residual with few exceptions only (see *Mauder and Foken* (2006)). However, we do not consider our approach a flaw, but rather a limitation to the interpretation of the residual as Q_{IT}. There is a plethora of publications suggesting various approaches how to eliminate the residual by splitting it across the sensible and latent heat fluxes according to the e.g. Bowen ratio, surface characteristics, ratio of actual to potential evapotranspiration etc. We intentionally did not consider applying any of these approaches, since the ecosystem of the MDV is different from the majority of other non-permafrost ecosystems investigated using eddy-covariance based approaches in that the energy flux directed toward deepening the melt-front may be the dominant term or of similar magnitude compared to the sensible and latent heat fluxes (e.g. our Fig. 5). In our case, Q_{IT} estimated as the residual corresponds to about 30 % of the net radiation and therefore corresponds to the maximum values reported in literature. However, permafrost-related fluxes are simply non-existent at these non-permafrost sites and thus their residuals may well be linked to the mechanisms proposed in the recent literature (see list in *Foken* (2008)). In other words, residuals may be observed at different sites for different well-justified reasons. Therefore the interpretation of the residual needs to be site-specific. The temporal changes of Q_{IT} as well as differences across water track and off-water track observations strongly suggest that our assumption of equating the residual to the melt energy flux at ice-table depth may be justified to a large extent. We do not believe that it makes sense to assume a generic value between 10 and 30 % of the net radiation for the residual reported in the literature, and then proceed to estimating the soil heat flux at ice table depth. However, we acknowledge that some portion of our residual may be unrelated to the melt-energy flux, questioning the validity of our assumption. To date, there is no experimental technique using direct flux measurements which can eliminate the residual. The residual was found to be greatly reduced when using large-aperture scintillometry based upon similarity theory or large eddy simulation to estimate SEB components. For comparison, any surface energy balance model would by definition assume the SEB to be closed, and models are routinely constrained by observations suffering from the observational residual. If those are flawed, then either the constraints or the nudging procedures are invalid, and the model results are also flawed rendering land-surface models useless. In summary, we are aware of our far-reaching assumption of interpreting the residual in the SEB as the melt-energy flux and have incorporated explanations and a discussion about our concept and the related issues in the methods (p.9, l. 7–33) and in the discussion (p.17, l.2–p.18, l.11) of the revised manuscript.

Furthermore, if QIT is related to energy input to the permafrost shouldn't it be directed downwards (i.e. it should be negative) since melting ice requires energy? Now the QIT estimated from the residual is mostly positive which seems physically implausible. Also Anonymous referee #1 pointed out some unphysical behavior of the estimated QIT when compared with dSTL which might stem from the fact that the energy balance residual is not equal to the real QIT.

You are perfectly correct in that an energy flux used for deepening the thawed layer should be directed into the ground. Please note, however, that we use the sign-convention commonly applied in micrometeorology to all energy fluxes including turbulent air fluxes and the molecular ground heat fluxes. Since the latter are directed into the ground and thus away from the surface, they are given a positive sign. The observations of unphysical behavior by referee #1 were clarified in our response to referee #1's comment on p.9, l.6 of the discussion paper.

2) The measurement campaign lasted only 26 days and hence far-reaching conclusions cannot be based on it. This is something that should be discussed in the text. I mean, there are likely seasonal changes in the energy fluxes in both areas, reference and water tracks, and these seasonal changes are not necessarily the same. The authors should discuss what part of the summer these measurements likely represent well and how the energy fluxes are likely changing during the summer.

The summer in this extraordinary high-latitude ecosystem is unusually short. The observations over 26 days were taken during the peak summer season between the end of December and January. If defining 'summer' as the period during which the soil at some depth over the course of the day is at temperatures above freezing, then the length of the summer equates to 95 (0 cm depth), 82 (5 cm depth) or 67 (10 cm depth) days based upon long-term climate measurements for 1993-2011. The summer period is centered around December 21 due to maximum insolation. There is little variation across sites within the MDV. The 26 day period of observations therefore represents a substantial portion of the MDV summer season during peak warming. We adjusted our manuscript to clarify that we observed energy exchange in the summer season. In particular, we justified the seasonal representativeness (p.12, l.11–p.13, l.2) and added a few thoughts on seasonal variation (p.13, l.16–20).

Minor Comments

page 1 line 4 To me "during the Antarctic summer of 2012-2013" sounds like that the measurements lasted for several months. Please explicitly mention the length of the measurement campaign (26 days) here.

Added the measurement period (p.1 l.8).

p. 4 Section 3 More information about the EC setup is needed. What was the sonic anemometer measurement height? Was the gas analyser next to the sonic or below it? What was the horizontal and vertical separation between the two sensors? Were the EC instruments clearly above the surface roughness elements (e.g. rocks)? How rough was the surface in the three locations? Give estimates e.g. for the roughness lengths

We added the requested information in tabular form (Tab. 1).

p. 5 Section 4 Did you filter out the low turbulence periods from the EC data? During low turbulence EC is not measuring accurately the surface gas exchange and hence these periods should be filtered out e.g. by removing periods with low friction velocity. This applies also to water vapor fluxes. This is somewhat linked to the non-closure of energy balance at EC sites (Wilson et al., 2002)

We did not observe a single averaging period with dynamic stable stability. Most periods were neutral or unstable lending to the unusual surface characteristics of low albedo and the sun not setting. We added a short statement in the Methods section (p.10, l.33–p.11, l.3)

p. 5 l. 11 Buoyancy correction by Schotanus et al. (1983) is slightly wrong and one should follow van Dijk et al. (2004) instead. Schotanus et al. (1983) is missing a term 0.51¡q¿¡Ts'w'¿ from the right-hand side of their equation (8). Here brackets denote temporal averaging, q air specific humidity (kg kg-1), Ts temperature (K) measured with the sonic anemometer and w vertical wind speed (m s-1).

Thanks for bringing this to our attention. In fact, we used the equation from *Liu et al.* (2001) which has the correct equations (p.8, l.20).

p. 5 l. 13 Foken et al. (2004) quality flagging scheme gives quality flags between 1 (best quality) and 9 (worst quality). Here you mention that periods with quality flags equal or smaller than 1 are filtered out. Please clarify

This is correct. We used the simplified quality flag scheme (0=1,2,3; 1=4,5,6; 2=7,8,9) and discarded all data with flags exceeding 1. The wording was adjusted (p.8, l.24).

p. 6 l. 5-7 It is a bit unclear what was the overall data coverage after filtering. Please clarify and mention it explicitly for each site.

We added information on data coverage after the filtering and gap-filling process for each station to this section (p.10 l.6–9). Additionally we show time series of the filtered and gap-filled energy fluxes in supplementary Figure 2, with marking of the data used for comparison of individual energy fluxes between the stations and with marking of data used for SEB calculations.

p. 6 l. 11 This 79 % is quite peculiar value. Why not to use nice round number like e.g. 80 %?

We decided to use the median of the contribution of the Water Track surface to the energy fluxes to select for representative values, which is 79 %. We agree that a rounded number looks nicer and therefore rounded off to 80 % (p.10, l.27–31).

p. 6 l. 8-21 and elsewhere Please use consistent naming for the different surface types. For example, here "stream channel" equals "River" in Fig. 3, right?

We consistently applied the term "stream channel" now and checked for other inconsistent naming of surface types.

p. 7 l.12 p.8 l. 2 It is unclear what is done here. What is the difference between the ratios reported here e.g. for QH? Please rewrite and clarify.

We referenced Figure 4 at the wrong position on p.8 l.1. With this being corrected now (p.14, l.17–19), the difference between values estimated from linear models vs. from daily totals are clearer.

p. 8. l. 3 How much later QH peaked at the water track? Two hours? Please add this information to make it more concrete

The lags were about 1 hour. We added them in the text (p.14, l.20–21).

Technical Corrections

p.4 l. 10 You are referring to Fig. 3 before Fig. 2. You need to go in order.

Since we cannot refer to Figure 2 before that, we removed the reference to Figure 3 now.

p. 4 l. 27 You can remove "and carbon dioxide" from the sentence since those measurements were not used in this study.

Removed.

p. 5 l. 10 Please replace "spectrally correction" with "corrected for low- and high-pass filtering"

Replaced.

p. 5 l. 12 You can remove "and carbon dioxide fluxes" from the sentence since those measurements were not used in this study.

Removed.

p. 5 l. 9-10 To be exact the method by Wilczak et al. (2001) does not eliminate the mean vertical wind. It tries to eliminate the mean vertical wind caused by tilted sonic anemometer.

We rephrased the sentence (p.8, l.16–18).

p. 6 l. 21 “Each class was assigned an adequate momentum surface roughness length from literature and used in footprint calculations”, right?

Replaced (p.10, l.25–26).

p. 7 Fig.3 “Glacial Till” and “Ice” cannot be separated from each other in grayscale figure. Please change the colors.

Used a dark gray for “ice” instead of gold.

p. 7 l. 5 remove “was” after water track.

Removed.

References

Foken, T., The energy balance closure problem: an overview, *Ecol. Appl.*, *18*(6), 1351–1367, doi:10.1890/06-0922.1, 2008.

Liu, H., G. Peters, and T. Foken, New equations for sonic temperature variance and buoyancy heat flux with an omnidirectional sonic anemometer, *Bound.-Lay. Meteorol.*, *100*(3), 459–468, doi:10.1023/A:1019207031397, 2001.

Mauder, M., and T. Foken, Impact of post-field data processing on eddy covariance flux estimates and energy balance closure, *Meteorol. Z.*, *15*(6), 597–609, doi:10.1127/0941-2948/2006/0167, 2006.