Answer to Anonymous Referee #2

The comments of the reviewer are in black, our reply is coloured blue.

This manuscript presents an analysis of methane (CH4) eddy covariance (EC) data measured above a wetland in NE Siberia. The manuscript focuses on CH4 fluxes during night time in non-turbulent and low-mixing conditions when the EC measurement level is decoupled from the surface. Wavelet methods developed in a companion paper are used to estimate fluxes with 1 min time resolution over one summer and this high frequency flux time series is used to identify and classify high CH4 flux events during the analyzed period. These events are then speculated to be linked with atmospheric mesoscale circulation taking place in these nocturnal low-mixing conditions. However, large part of the abstract, introduction and some other sections are discussing ebullition and other non-related topics, whereas results and conclusions are all about nocturnal low-mixing conditions. The authors should modify the beginning of the manuscript so that it matches with the end, so that the text forms one coherent entity. There are also other shortcomings in the text and description of data processing. Please see below.

As it stands the manuscript is interesting and shows promise but requires major revisions (see below) before publication. Once revised, it should be of interest also for the wider community working with micrometeorological flux measurements and hence the study is within the scope of ACP. Besides the shortcomings mentioned above, the presentation quality is good, although some figures need adjustment. I recommend the publication of this manuscript after major revision based on the comments below.

We thank Anonymous Referee #2 for his constructive comments. According to his remarks we revised our manuscript as described in the following reply.

GENERAL COMMENTS

1) Please modify the abstract and introduction so that they match with the results. In my opinion these sections should be largely rewritten since now they are quite disconnected from the rest of the manuscript. The results are about gas fluxes under nocturnal low-mixing conditions and the abstract and introduction should be written about this topic, not about arctic wetland CH4 emission dynamics. As you know, these problems related to low-mixing conditions are universal, not only related to arctic wetlands.

We agree that there is some kind of disconnection between the specific process of ebullition, which is presented in abstract and introduction, and the results of our manuscript. Nonetheless we think that it is important to consider that the scientific discussion on methane emissions in Arctic permafrost wetlands mentions ebullition as an important pathway. Thus the main reason of our data analysis was to find signs of ebullition using the wavelet approach – in our case studies, we detected other reasons for all found events, but no signs of ebullition. It seems that ebullition, occurring as heterogeneous single events on the spatial scale of the EC footprint of our towers, is not detectable. We think, that this finding might be also important for the scientific community.
We will rewrite parts of abstract and introduction as requested, so that it will be clear that ebullition is not the main topic of the manuscript, but we decided not to remove our remarks on ebullition completely due to its importance in Arctic permafrost wetlands.

2) The wavelet method is presented in the manuscript as more accurate than EC and

reliable reference for the EC fluxes. This is a strong statement, which should be supported by convincing evidence. In order to make that kind of statement you should show that when the data is processed using standard EC methodologies spurious data points are left in the flux time series, yet with the wavelet methods these problematic periods are handled better. I suspect that most of the low-mixing conditions would have been filtered out by quality screening and friction velocity filtering the data. Hence CH_4 budgets derived using standard EC processing are likely not affected by these spurious fluxes during low-mixing. Please show a comparison of fluxes (e.g. monthly CH_4 budgets) derived with standard EC processing (including quality screening and friction velocity filtering) and fluxes calculated with your wavelet methods to support your statement. Alternatively, you should phrase the text differently so that it is clear for the reader that it is not possible to say that the wavelet method is more accurate or reliable than standard EC.

Your question is quite similar to specific comment 8 by Norbert Pirk. Thus, we answer both as follows: using a mother wavelet with an exact resolution in time domain like the Mexican hat wavelet yields the exact flux by integrating over a short time interval, e.g., 1 min. In this case also all parts in the frequency domain that contribute to the flux are included and considered (Percival and Walden, 2000). Summing up these 1 min fluxes to the typical EC averaging interval of 30 min, the results of both EC and wavelet method must be equal as long as the lowest contributing periods < 30 min. That is the case during steady-state conditions (Foken and Wichura, 1996; Foken et al., 2004, 2012) and Schaller et al. (2017) could prove that comparing both wavelet analysis and EC.

In the case of non steady-state conditions with contributing periods > 30 min, the EC quality control tests should flag those cases to be excluded (Foken et al., 2012). Additionally, in those cases also the ogive test (Desjardins et al., 1989; Foken et al., 1995; Oncley et al., 1990) yields contributions to the flux for periods > 30 min. Besides that, the Mexican hat wavelet will yield nonetheless in every case correct and trustworthy fluxes, also for periods > 30 min, if the integration interval in the period domain is chosen big enough (Percival and Walden, 2000; Torrence and Compo, 1998).

We will add this information to the revised manuscript.

As you correctly mention, the event cases in our study would be excluded and/or replaced most probably during the gap filling analysis due to poor quality flags or the friction velocity u* being too low. Usually, gap filling algorithms determine the regression by binning the data into classes and calculating the median. This procedure leads to an exclusion of “outliers” containing very big or very small fluxes – and especially these usually rejected “outliers” which might contain distinctive, significant fluxes are targeted in our manuscript. There are a few phenomena that could explain such processes, like ebullition, free or wet convection, low level jets, breaking gravity waves or advection during calm wind, which have partly also been addressed in previous single studies.

Our manuscript explicitly investigates single events that would usually be removed by gap-filling algorithms, but does not target long-term results. A long-term study using EC on our dataset was already published by Kittler et al. (2017). An extension of our manuscript by long-term studies on the wavelet flux including error analysis would be beyond the scope of this article. Nonetheless, we definitely agree that a comparison of wavelet fluxes and EC fluxes as well as EC gap filling approaches over a longer period should be performed to quantify its effect-term balances. This would be an interesting additional future paper, which is already in preparation in our working group. **We will add this information to the revised manuscript.**

3) In many occasions the observed high flux events are linked with mesoscale motions (gravity waves, low-level jets etc.), yet the connections are quite speculative. This is understandable since these mesoscale motions are difficult to quantify with just one flux tower and the authors also clearly state this in the discussion section of the manuscript. However in contrast to the discussion, the abstract and the conclusions are written in such way that the connections are obvious based on the data. Please rephrase the text so that it is clear for the reader that the role of mesoscale motions is quite speculative and additional instrumentation would be needed for a proper identification of these flow patterns.

Yes, we generally agree that it is difficult to link the found high flux events to the identified mesoscale processes. Your question is similar to specific comment 7 by reviewer Norbert Pirk, which we both answer as follows:

Based on our meteorological measurements, we conclude that the identified mesoscale processes triggered the observed high-flux events. We agree that it is somehow “a stretch” or “speculative”, to identify these mesoscale phenomena without measurements of spatial (vertical / horizontal) profiles. Usually, periods consisting of phenomena like the found mesoscale processes, are replaced by gap filling algorithms during the standard eddy covariance processing. Long-term measurements of the atmospheric boundary layer, including devices like SODAR-RASS or LIDAR as well as arrays of other vertical / horizontal gradient measurements, could fill this gap. Long-term measurements are necessary here, because these phenomena are not detected very often and any statistical analysis is nearly impossible then.

Nonetheless we think, that it will be worth to publish observations of such relatively rare phenomena if it is possible to observe them. The detection, identification and differentiation of such mesoscale phenomena require considerable experience. Such experience is fortunately available in our working group, considering, e.g., the publications by Foken et al., (2012) or Serafimovich et al. (2018), where surface flux measurements and boundary-layer measurements were available. Of course, we carefully discussed the identified processes within our working group, to be as sure as possible in our statements regarding the available data. For the identification of these mesoscale processes we used the data from both towers (distance: 600 m). Here it is important to know that the identified phenomena are always visible at both tower 1 and 2, which supports our findings. **This was mentioned in the manuscript only in the first sentence of section 3.3, so we will add this finding also to the conclusion to highlight it.**

To sum up, the classification of the different mesoscale processes is not completely satisfying, but also not fully speculative. **We will carefully revise the manuscript, so that it is clear for the reader why the identification is not just speculative.**

4) More details about data processing are needed. Did you do coordinate rotation and how did you do it? The regular 2D-coordinate rotation (align u with mean wind and nullify mean w for each 30 min period) does not necessarily work well during low-turbulence since mean w is not necessarily zero when mesoscale motions are at play. Hence I hope that you used planar fitting (Wilczak et al., 2001) and defined the plane used in the coordinate rotation using high quality data. On the other hand if you did not do any coordinate rotation (like in Schaller et al., 2017) then the fluxes might be seriously compromised, since sonic anemometers are always slightly tilted no matter how carefully they are aligned with the surface below. Also did you correct for the time lag between gas analyzer data and sonic data? Lag time determination is always difficult for periods with low and intermittent turbulence. Therefore, please

provide additional details about EC processing. Related to the wavelet analysis, how did you take into account the time series edges and their effect on the results? Did you zero-pad the data and then estimate the cone of influence (Torrence and Compo, 1998)? This is important especially for the low frequencies. Please add details, since they are missing also from the companion paper (Schaller et al., 2017). As it is, it is difficult to judge whether the data were processed in a proper way.

Thanks for addressing these important points, we agree that in the current state of the manuscript it is difficult to be sure that the data (pre-) processing was done in a completely correct way, so we will add these information as follows.

Coordinate rotation – Norbert Pirk also addressed this topic in his specific comment #25. So we answer both your and his question as follows: Due to a sloped terrain, a non-exact alignment of the sonic anemometer or flow distortion effects, the streamlines of the wind might be tilted. In this case, usually a coordinate rotation is conducted to move the coordinate system into the streamlines and to solve this problem. For our study, we carefully inspected the measured wind vectors during times with well-developed turbulence and good flow conditions, i.e., sufficiently high wind velocities. We could not detect any disturbance of the streamlines due to the terrain or other influences, i.e., we are sure that the assumption of a negligible mean vertical wind component ($\langle w \rangle = 0$) is valid. Due to that finding, the complete study is based on w-data without tilt correction.
Especially due to the fact that the found phenomena were connected to a distinct vertical wind component, double rotation might lead to irregularly high rotation angles (this was investigated in an master thesis by Matthias Mauder in 2002 for the EBEX-2000 experiment, flat terrain). Instead, planar fitting (Wilczak et al., 2001) is a proper choice in such situations, since here the rotation angles are based on a long term averaging period, so that short time periods, where $\langle w \rangle \neq 0$, does not affect them. In our study we still did not apply planar fitting, because 1) it is not a long term study to make a careful analysis of the optimal interval for planar-fit rotation (Siebicke et al., 2012) and 2) at our towers fortunately the assumption of $\langle w \rangle = 0$ was proven as valid for well-developed turbulence. **We will add this information to the revised manuscript.**

Time lag correction – the time lag correction between the sonic anemometer and the gas analyser was conducted by maximisation of the covariances by cross correlation for every 30-minutes-interval. Because of the conditions with low turbulence the time lag may be different for each time series. No constant time lag was applied. **We will add this information to the revised manuscript.**

Edge effects in wavelet analysis – the filtering of erroneous results of the wavelet calculations at the beginning and end of the cross scalogram is of crucial importance to obtain results of high quality without border effects (Torrence and Compo, 1998). Our wavelet analysis was conducted on the whole available dataset from 1st June to 15th September 2014 using a windowed approach. For the Mexican hat wavelet, the lag between two subsequent window calculation start points was set to 6 hours while the length of each calculated window was 12 hours, i.e., it overlaps the window start/end points. The data was zero-padded and the cone of influence estimated. At the end, the calculated windows were cut at the borders and merged, so that border effects in the final result only occur at the beginning and end of the whole time series. These border times were excluded from further analysis. **We will add this information to the revised manuscript.**

SPECIFIC COMMENTS

page 7, line 24-27 This part is unclear. Do you mean extreme outliers in the 20 Hz data or in the 1-min flux data? It is difficult to understand why the outliers could explain the difference in the amount of events observed with the two towers. Please clarify and rephrase.

In this section of our manuscript we refer to the statistics of detected events, i.e., the number of detected event-minutes. The statistics is based on the result of wavelet analysis, i.e., the 1-min Mexican hat wavelet flux.
At tower 1 7.7 % of all data were out of the range from Q_1-1.5(Q_3-Q_1) to Q_3+1.5(Q_3-Q_1), where Q_1 denotes the 25%- quantile and Q_3 denotes the 75%- quantile. At tower 2 5.4 % of the data are out of this range, i.e., 2.3 % less “outliers” than at tower 1.
As the MAD test is a robust estimator, it is quite resilient against such outliers. In consequence, due to the statistical properties of the data, the number of outliers (event minutes) detected by the MAD test is greater at tower 1. **We will modify the text in lines 24 – 27 to clarify that.**

Sect. 3.3 I would like to see an analysis using Richardson number (Ri), since Ri is typically used to indicate dynamic stability of the flow. Moreover, if Ri exceeds so called critical Richardson number (Ric) then the turbulence is strongly dampened or even almost completely wiped out (e.g. Grachev et al., 2013). Ric is typically said to be around 0.25, although this is debated (Galperin et al., 2007). It would be interesting to see how Ri is affected by these events you identified and if Ri>Ric always before the events. The analysis you did on stability parameter (z/L) is somewhat similar, however I would prefer Ri since you cannot determine a turbulence cutoff with z/L the same way as with Ri. I suggest you use gradient Richardson number for the analysis, however in case you are missing the needed vertical gradients, then use flux Richardson number.

The flux Richardson number (Rf) can be converted exactly into z/L (e.g., Arya, 2001; Foken, 2017); the gradient Richardson number should not be used under stable conditions because of possible decoupling between the both measurement heights. It can be assumed that the critical Richardson number is approximately equal to z/L=–1. The calculation for Rf and z/L is extremely affected by errors because for stable conditions u or u* are near zero, and small errors can generate large effects on the stability parameter. For the accuracy under such conditions see also Högström (1996). Additionally, in our studies we use z/L only as a general classifier for the stability, i.e., the exact numerical values should not have an important impact on our findings.

p. 8, l. 20-25 Why the analysis with relative humidity? I would guess that it is not relevant for the topics at hand.

Especially ecologists use the relative humidity (combined with air temperature) quite often for studies on atmosphere-biosphere exchange. We think that the information on humidity provides meaningful information that should remain in the manuscript.

p. 8, l. 29-32 Referring to my comment before, did you do coordinate rotation? It should be always done, regardless of how flat the terrain is since anemometers are always at least slightly tilted. If coordinate rotation is not done, then w data is compromised by horizontal wind speed fluctuations.

Please see our answer to your general comment 4).

p. 9, l. 6-8 It is difficult to understand why there would be unstable stratification during night. Could it be because during these events EC is not working properly and hence you have erroneous heat fluxes and therefore also erroneous values for z/L? Did you have also negative vertical gradients in air temperature (decrease with height) during these periods?

We agree that usually during times with short-wave incoming radiation ≈ 0 unstable stratification is quite unlikely to happen. In the high-latitude Arctic zone (68.78° N), depending on the date, there are just about a few hours between sunset and sunrise. In our study, we defined "night" by the fixed time span 21:00 to 9:00. **We will add this information to the revised manuscript.**

Additionally, we revisited all events that occurred during our fixed nighttime span 21:00 to 9:00. Unstable stratification was observed only in times where also short-wave incoming radiation > 20 Wm^{-2} was measured. In such cases it is not unlikely to observe at least slightly unstable situations. **We will add this information to the revised manuscript.**

p. 9 l. 16 This title should be modified. Based on the evidence shown it is not possible to say that there was advection of CH4 to the study domain. In order to make such a statement you should have measured also CH4 concentration horizontal gradients.

Please see in general our answer to your general comment 3, where we show why our statements on the mesoscale processes are still not completely satisfying, but definitely not fully speculative. Due to the fact that there are only rare studies on such kind of mesoscale events in the Arctic, we think that it is worth to publish them.
Specifically to your question on section 3.4 labeled "Nighttime advection": We agree, that it is much more difficult to detect advection without direct gradient profile measurements. On the other hand, Figure 6 in the manuscript supports this finding and gives evidence, because it directly shows the approaching fog bank. The occurrence and arrival of this fog bank at the towers exactly in time with the observed event in the data clearly indicates advection. Also in literature such kinds of fog events are classified directly as advection fog (e.g., Stull, 1988).

Sect. 3.4 The event that is analyzed in this section was already analyzed in the companion paper (Schaller et al., 2017). For instance Fig. 4 here is partly the same as Fig. 5 in Schaller et al. (2017) and also the text is quite similar. It would be better to concentrate on some other event in this study, now this analysis is a bit redundant.

We agree that there is a bit of redundancy between this manuscript and Schaller et al. (2017), due to the fact that the event discussed intensively in our manuscript was already shown partly there. But, the focus between the two publications is completely different: Schaller et al. (2017) show the methodology behind the wavelet flux calculation and validate the method against EC for times of well-developed turbulence and stationary conditions. In section 3.2.2 of that publication we gave a very short insight to the capability of wavelet analysis to resolve also fluxes in times, where the steady-state condition is not fulfilled. In that manuscript, we "promised" the reader that there will be a follow-up paper, which investigates such kind of events as well as this specific event in detail. We think that we should keep this "promise" we gave in our companion paper, so we decided not to change our focus to another event.
In the revised paper we will clarify that this example was analyzed in Schaller et al. (2017) **with respect to the method, while this article investigates the underlying process.**

p. 10 l. 12-15 You analyse here a period that lasts for two hours, right? Can you then extent the maximum wavelet period above 120 min? If you can, then how accurate the results are at these very low frequencies, given that your time series does not cover even one whole wavelet when the wavelet period is above 120 min? On a related note, shouldn't you also take into account the cone of influence (those regions of the wavelet spectrum that are significantly affected by the edges; see Torrence and Compo, 1998) in your cross-scalograms and in the corresponding analysis? If you used two hour long time series in this analysis, then wavelet periods above 120 min are definitely within the cone of influence and hence unreliable.

Please see our remarks on the edge effects on your general comment 4). If there are only two hours of data, a wavelet analysis on this dataset will not yield to trustworthy data of that length due to parts of the results being within the cone of influence. None of the data we used or showed in our studies are influenced by edge / border effects, as already explained. **We will add this information to the revised manuscript.**

p. 11 l. 4-5 This sentence should be modified. One cannot claim that EC fluxes were systematically overestimated since you do not have an absolutely correct reference. For instance damping of the signal within the cone of influence (Torrence and Compo, 1998) might decrease the wavelet based fluxes. This could partly explain the observed difference.

Please see our answer on your general comment 2). Due to the mathematical properties of the wavelet analysis in general the steady-state assumption has not to be fulfilled, while eddy covariance results might be significantly biased on the same time. Additionally, the chosen mother wavelet (Mexican hat) allows a very good resolution in the time domain and thus, as there is no influence by edge effects in this period, there is no damping of the signal caused by the cone of influence. **It is not possible to set up an "absolutely correct reference" under field conditions. To consider that, we remove the word "strongly" in the revised manuscript.**

p. 11 l. 31 Please replace "by an eddy-covariance system" with "with these wavelet algorithms".

Because the footprint does not depend on the data analysis we believe that "EC system" is already a more general name than "EC method" or "wavelet method". **To be sure that the reader will not misunderstand that, we changed the sentence to "by this EC setup with a sensor height ≥ 4.9 m above ground".**

Sect. 3.52 & 3.5.3 & 3.5.4 Difference between these three categories is difficult to see, especially the description of 3.5.3 and 3.5.4 looks similar. Try to emphasize more the differences in meteorological forcings between these event categories. As it reads now, combining the events with different mesoscale flow patterns seems rather subjective.

Please see our answer to your general comment 3), which also addresses the differentiation between the meteorological forcings. Concerning the mentioned sections, we carefully revisited them: the examples 3.5.1, 3.5.3, 3.5.4 look similar, but the reason is different. 3.5.1 is due to the constant wind direction and the observed fog clearly an advection situation. 3.5.3 and 3.5.4 show both an increase of turbulence, but a change of the wind direction is typical for a low level jet and not for braking gravity waves. Example 3.5.2 has all criteria for a passage of a frontal system, see e.g. pressure. After carefully revisiting the events, **we decided to keep the differentiation, also due to our answer to your general comment 3).**

p. 12 l. 12-13 As you probably know, ebullition is often hypothesized to be connected with falling (Tokida et al., 2007), but sometimes also increasing atmospheric pressure (Chen and Slater, 2015). Could this be somehow connected to this daytime event?

We thoroughly inspected the observations for that daytime event again. The course of the air pressure, which was decreasing before the event and increasing by 2 hPa/hour with and after the event clearly supports the assumption of a front passing by, which is also observable in all fluxes, not only the methane flux.

p. 13 l. 20-21 Onset of turbulent mixing in the morning has been shown to cause CH4 flux peaks also in other studies (e.g. Peltola et al., 2015). Did these events that you identified to be connected with the onset of turbulent flow take place in the morning?

All events that were determined to be caused by the onset of turbulent flow were observed during night time, but none of them in the morning. It should be noted here that due to high geographical latitudes there was never a complete sunset down to darkness and also the sunrise does not occur that rapidly as, e.g., in the middle latitudes like in Central Europe.

p. 14 l. 8 How did you define which events were influenced by advection? These periods discussed here are most likely non-stationary and would have been filtered out from standard EC data.

Please see our answer to your general comment 3). In the usual EC processing these periods would be filtered out, and subsequently treated by gap filling algorithms as they are in fact non-stationary and would fail the steady-state test.

p. 14 l. 11-18 This is a good point and it would have been nice to see this idea used in the prior analysis.

Such an analysis needs an extended experimental setup. We think that this analysis should be done using data from a long-term study including more towers, additional boundary layer measurements as well as arrays of other vertical / horizontal gradient measurements.

p. 14 l. 24 How did you determine this 15 min limit for identifying events that are affected by advection?

In fact, this limit is somehow a helpful “practical rule of thumb” for the events observed at our measuring site considering the available. On the other hand, we also see that is difficult or even impossible to specify an exact minimum event duration, where you can be sure that advection is the driver. A long-term study would be necessary to prove that, **therefore we will remove this sentence in our manuscript.**

Sect. 4.1.1 I would add here text about CH_4 concentration profiles since large part of this manuscript discusses flushing of previously stored CH_4 below the EC level. With detailed concentration profile you could measure this.

Yes, profile measurements would be definitely a good measure to support findings on meso-scale processes like in our study. In addition to our answer to your general comment 3), **we will also add this information to our manuscript.**

p. 15 l. 1-2 Why the analysis on cluster events was not possible?

We assume, you mean p. 16 l. 1-2. We decided to postpone this analysis to a follow-up study, because the complete, detailed investigation would be beyond the scope of this manuscript. It is quite difficult and due to the lack in additional boundary layer measurements maybe even impossible to get reliable evidence on the exact meso-scale processes triggering these events. Another major reason for our decision to exclude these events is the scope of the manuscript, which focuses on events that occur at short timescales, i.e., last only for minutes or some tens of minutes. **We agree, that this reasons were not stated clearly yet, so we will modify section 4.1.2, so that is clear, why the analysis might be possible but beyond the scope.**

Figure 2 Mean w is around 0.15 m/s, which is quite high value. Did you do coordinate rotation? You should definitely do it. Another thing: you could add here the Richardson number, like I suggest above.

Please see our answer on your general remark 4) and on your specific remark on section 3.3.

Figures 2 & 3 These two figures are complicated and should be explained better. For instance how did you define “Unstable”, “Stable” and “Neutral”? Where the stability is shown? How can you have EC data in the bottom plot with different quality classes at the same time?

We will improve the caption text of both figures, so that everything is well explained. For the definition of the stability we followed the recommendations on error analysis by Foken and Wichura (1996) and Foken et al. (2004, 2012) to be consistent. These recommendations derive from the fact that here the universal functions are nearly equal to 1 (Foken and Skeib, 1983). Stability was labeled “unstable” for $zL^{-1} < -0.0625$, “stable” for $zL^{-1} > 0.0625$, “neutral” for $-0.0625 \leq zL^{-1} \leq 0.0625$. In the plot the stability is shown for every 30-minute-interval directly underneath the cross-wavelet scalogram of Mexican hat.

The quality class of the EC data is labeled by small rectangles in the bottom panel of Fig. 4f for each 30 minute interval, i.e., the timestep 0:00 – 0:30 in Fig. 3 is labeled as quality class 4 – 6. There are no different quality classes at the same time, but maybe here the Figure is misleading the reader. **Therefore, we will revise the bottom planel, so that is easier to determine the EC quality class.**

Figures 2, 3 & 4 You most likely have change in flux sign at some certain color in the cross-scalograms (e.g. negative fluxes at blue colors and positive at red colors). Please highlight the zero flux lines in the cross-scalograms with e.g. white contour lines. Also, is the color scale the same in both subplots? If not, then please try to use one color scale per figure. Add also the cone of influence (Torrence and Compo, 1998) to all subplots.

Yes, the values in cross-scalogram are negative or positive valued, depending on the direction of flux. **We will change the colors in the revised manuscript, so that intensity and algebraic sign are easy to read off. Also the color scale for the subplots in Figures 2, 3 & 4 will be changed. There is no cone of influence (see your general comment 4), but we will note this in figure captions.**

TECHNICAL CORRECTIONS

p. 4, l. 12 You defined the abbreviation EC here, but you defined it already on page 2 line 19. Use the abbreviation everywhere in the text after you define it. Also, you use both “eddy covariance” and “eddy-covariance”, replace both with EC.

We will change that.

p. 7, l. 14 and other places Please give dates in a consistent manner and try to follow the journal recommendations.

We will change that.

References

Arya, S.P., 2001. Introduction to Micrometeorology, 2nd ed. Academic Press, San Diego.

Desjardins, R.L., Macpherson, J.I., Schuepp, P.H., Karanja, F., 1989. An Evaluation of Aircraft Flux Measurements of Co2, Water-Vapor and Sensible Heat. Boundary-Layer Meteorology 47, 55–69.

Foken, T., 2017. Micrometeorology, 2nd ed. Springer, Berlin.

Foken, T., Dlugi, R., Kramm, G., 1995. On the determination of dry deposition and emission of gaseous compounds at the biosphere-atmosphere interface. Meteorol. Z. 91–118.

Foken, T., Göckede, M., Mauder, M., Mahrt, L., Amiro, B.D., Munger, J.W., 2004. Post-field data quality control, in: Lee, X., Massman, W., Law, B. (Eds.), Handbook of Micrometeorology: A Guide for Surface Flux Measurement and Analysis. Kluwer Academic Publishers, Dordrecht, pp. 181–208.

Foken, T., Leuning, R., Oncley, S.R., Mauder, M., Aubinet, M., 2012. Corrections and Data Quality Control, in: Eddy Covariance: A Practical Guide to Measurement and Data Analysis, Springer Atmospheric Sciences. Springer, Dordrecht, pp. 85–131.

Foken, T., Skeib, G., 1983. Profile measurements in the atmospheric near-surface layer and the use of suitable universal functions for the determination of the turbulent energy exchange. Boundary-Layer Meteorology 25, 55–62. https://doi.org/10.1007/BF00122097

Foken, T., Wichura, B., 1996. Tools for quality assessment of surface-based flux measurements. Agricultural and Forest Meteorology 78, 83–105. https://doi.org/10.1016/S1352-2310(96)00056-8

Högström, U., 1996. Review of some basic characteristics of the atmospheric surface layer. Boundary-Layer Meteorology 78, 215–246. https://doi.org/10.1007/BF00120937

Kittler, F., Heimann, M., Kolle, O., Zimov, N., Zimov, S., Göckede, M., 2017. Long-Term Drainage Reduces CO_2 Uptake and CH_4 Emissions in a Siberian Permafrost Ecosystem: Drainage

impact on Arctic carbon cycle. Global Biogeochemical Cycles 31, 1704–1717. https://doi.org/10.1002/2017GB005774
Oncley, S.P., Businger, J.A., Itsweire, E.C., Friehe, C.A., Larue, J.C., Chang, S.S., 1990. Surface layer profiles and turbulence measurements over uniform land under near-neutral conditions, in: 9th Symp on Boundary Layer and Turbulence. Amer. Meteor. Soc., Roskilde, Denmark, pp. 237–240.
Percival, D.B., Walden, A.T., 2000. Wavelet methods for time series analysis. Cambridge University Press, Cambridge.
Schaller, C., Göckede, M., Foken, T., 2017. Flux calculation of short turbulent events -- comparison of three methods. Atmospheric Measurement Techniques 10, 869–880. https://doi.org/10.5194/amt-10-869-2017
Serafimovich, A., Metzger, S., Hartmann, J., Kohnert, K., Zona, D., Sachs, T., 2018. Upscaling surface energy fluxes over the North Slope of Alaska using airborne eddy-covariance measurements and environmental response functions. Atmospheric Chemistry and Physics 18, 10007–10023. https://doi.org/10.5194/acp-18-10007-2018
Siebicke, L., Hunner, M., Foken, T., 2012. Aspects of CO2 advection measurements. Theoretical and Applied Climatology 109, 109–131. https://doi.org/10.1007/s00704-011-0552-3
Stull, R.B., 1988. An Introduction to Boundary Layer Meteorology. Kluwer Academic Publishers, Dordrecht, Boston, London.
Torrence, C., Compo, G.P., 1998. A Practical Guide to Wavelet Analysis. Bulletin of the American Meteorological Society 79, 61–78. https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2
Wilczak, J.M., Oncley, S.P., Stage, S.A., 2001. Sonic anemometer tilt correction algorithms. Boundary-Layer Meteorology 99, 127–150. https://doi.org/10.1023/A:1018966204465