We like to thank the reviewer for the time spent and the valuable comments. They will help us to improve the manuscript. Please, see below our answers to the single comments.

Reviewer 3.

General:

I think that the introduction and discussion are a bit narrow when it comes to referencing related work. The manuscript mostly refers to studies previously performed at Stordalen mire, however, there is not much discussion about related work around the Arctic.

Thank you for bringing this up. We will add other studies, not directly related to Abisko-Stordalen to the discussion.

Specific comments:

l. 62: Do you refer here to surface-near air temperature or soil temperature or permafrost temperature?

We referred to the near-surface air temperature. We will specify this in the text of revised version.

l. 104-109: Pleae refer here in the introduction to previous studies that have conducted similar comparisons, e.g., Hommeltenberg et al. (2014), Rößger et al. (2019), Kim et al. (2020). Rößger et al. (2020) investigated methane fluxes from a heterogenous tundra ecosystem; thus this article would be quite appropriate for comparison also in other regards

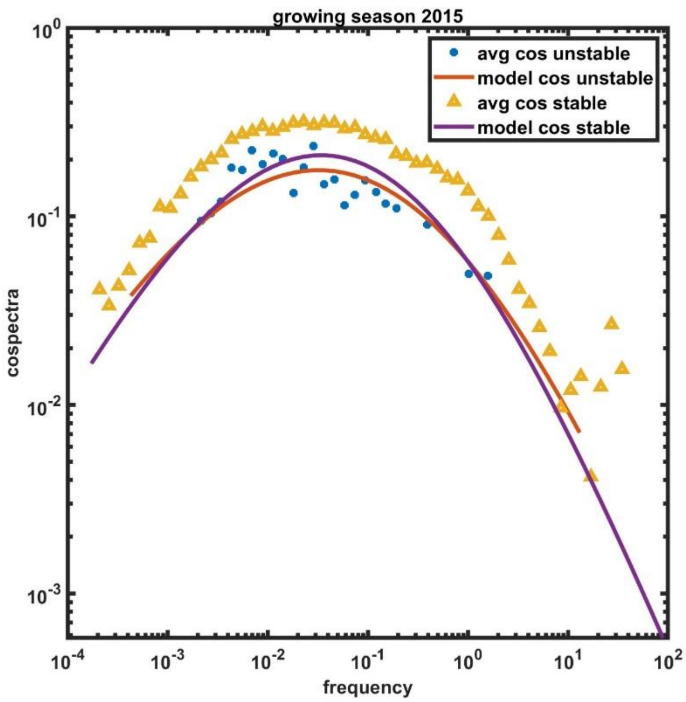
Thank you for the suggestions. We will add those papers to the introduction

l. 123: Specify if surface-near air temperature is meant.

We will clarify that air temperature was meant. We will replace the sentence "The mean annual temperature in this region has been increasing… " with "The mean annual air temperature in this region has been increasing… "

l. 153: The tube length is very long. Can you assure that flow was turbulent throughout the tube? What ist he high-frequency attenuation of the fluxes due to the tube transport effects?

We analyzed this with the cospectra of the CH_4 and w. This does not show a dampening effect at the high frequency (see figure below), thus the high frequency attenuation does not seem to be very large. Furthermore, the postprocessing software we used to calculate fluxes includes correction for high-frequency losses We will add a statement on this in the text.



l. 168-181: Please describe better the locations of the ancillary soil measurements. In a heterogenous mire landscape peat temperature can have large spatial variability. Particularly of interest is what site you choose as being representative for the heterogeneous eastern area composed of drier palsas and thawed wetter sites.

Soil pit temperature and moisture probe for western sector is located on a palsa plateau. However, the WTL probe is located in a pond approximately 10 m away from the soil temperature and soil moisture sensors, as there is no water table above the permafrost of palsas. Soil pit temperature and moisture probe for the eastern sector is located in the wet area. WTL probe is located in the wetter area located approximately 10 m away.

l. 202-203: I do not understand this approach of removing flux values when two consecutive data points originated from different wind direction sectors? Which flux values where then removed? Why was this done?

We will change the text from “Also, flux values when two consecutive data points originated from different wind direction sectors were removed.” to

“Also, consecutive data points originating from the two pre-defined wind direction sectors were removed to avoid influences from unstationary conditions”.

l. 213-215: Have you tried to model also 30 min fluxes? Why not modelling the 30 min flux data (as Rößger et al. (2019))?

We have not tried 30 minutes’ gap-filling by ANN, but did so with Jena-tool. We portioned data in the different way that Rößger et al. (2019). In the comparison of the model, results from the model based on the 30 minutes (“Jena”) does not show any better performance than models designed on the daily averages. As most of the sub-daily variation of CH_4 fluxes seems to be due to random turbulent variability, modelling this variability may not be very fruitful, at least for gapfilling purposes.

l. 222-223: Please describe in more detail how the 30 min data were „aggregated" to annual footprint climatologies.

We followed the standard approach to derive footprint climatologies (e.g. Kljun et al. 2015), where footprint function values are aggregated for each grid cell (50 cm x 50 cm). We will add more detailed description to the revised manuscript.

l. 235-237: How is the weighting calculation exactly done? Which quantity of the „climatology" was used for weighting the contribution of a mosaic pixel?

Again, we refer to Kljun et al. (2015) where the derivation of a footprint climatology and its application to remote sensing data is described. The flux contribution per pixel is weighted with the footprint function value.

l. 245-246: The statement that "...methane emissions … do not show diel cycle" is too bold. Figure S4 shows that there is systematic diurnal variability – even for whole-year data. Indeed, it would be good to analyse diurnal variability month by month.

Figure S4 shows indeed slightly lower emission during mid-day hours. However, the difference is very small compared to the short term variation in the fluxes as indicated by the interquartile range. Thus, for the purpose of gapfilling this effect could be negligible in calculating daily averages. However, it is interesting to observe this type of diel cycle, with minima at daytime, and as reviewer suggested, it could have its origins on temperature cycle of the top peat layer. This could affect the methanotrophy, while the methanogenesis occurring at slightly deeper layers would be less affected. This would lead to higher methanotrophy at daytime and thus lower emission. We will include this discussion to the revised version of the manuscript

l. 274: Rößger et al. (2019) applied ANN for a heterogenous tundra; the paper might be interesting for comparison

Thank you for the reference. We will incorporate it in the manuscript.

*l. 323-333: I think that the equations (1) and (2) are only valid under rather strong assumptions that should be clearly stated. In my view equation (1) and equation (2) can be considered valid for the 30-min periods for which the footprint contributions of the two contrasting landcover types fp and ft are estimated. However, using the same form of the equation for annual averages is only valid if the time series of the footprint contributions fp and ft, respectively, are uncorrelated with the temporal development of the emission factors Ep and Et, respectively. If they would show some correlation, the average of the product f*E would equal the product of the averages of f and E, respectively, plus the covariance of f and E. Therefore, one maybe important assumption is that f and E of the* *respective landcover types are uncorrelated. Other important assumptions are that the average methane fluxes of the palsa sites in the eastern area and in the western area are equal and that the average methane fluxes of the thawed sites in the eastern area and in the western area are equal. It would be good if this assumption could be backed by more comprehensive description of microtopography, hydrology and vegetation of palsa and thawed sites of the western and eastern areas, respectively.*

Yes, we will clearly state the assumptions in the revised version of the manuscript. Of course this is estimate has its caveats but we feel it is useful as an attempt to separate the effect of different surfaces to aggregate flux. Figure below shows monthly averages of the contribution of surface classes to the footprint. According to the figure relative contribution of the different surface classes is stable during each year and also between years.

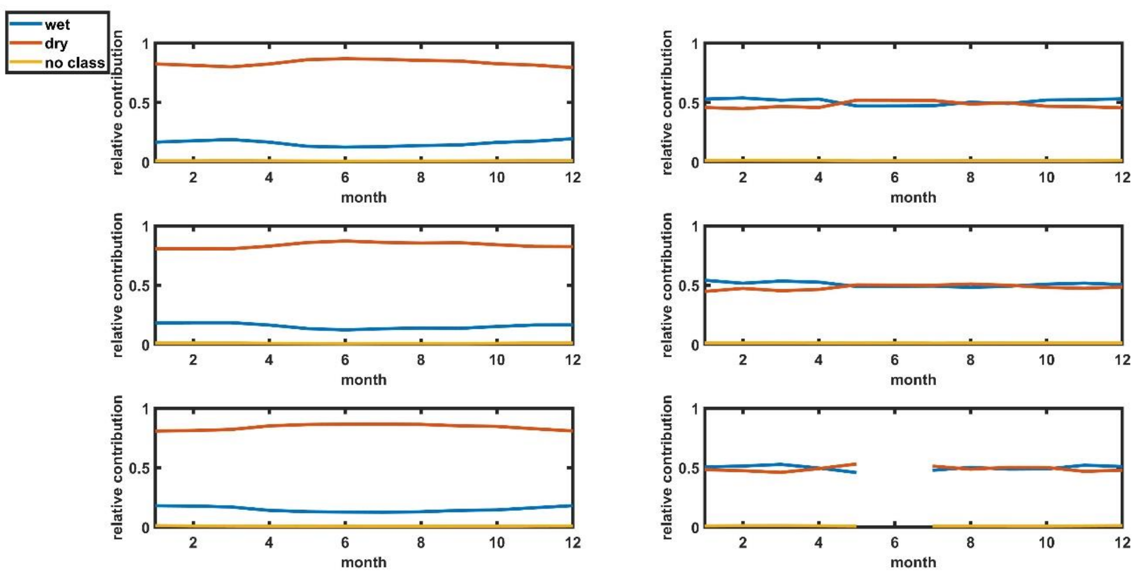


Figure1. Relative contribution of the different surface classes to footprint

l. 360-361: Sentence too vague: Where and when the western sector is colder? Which depth? Time scale?

We will change the text from "The western sector, was colder than the eastern sector, causing the existence of permafrost. " to

"During our investigation period (2014-2016),the peat temperature from 30 to 50 cm below ground was colder in the western sector, than those of the eastern sector, corresponding to the existence of the permafrost."

l. 369, Figure 2: The lowest panel does not show water depth but water table height. However, it would be more suitable to show water table depth or height referenced to a reference point at the ground surface in the investigated mire. The jump in water table *height between the summer of 2014 and the summer of 2015 appears unrealistic. Please check the water level times series for biases and inhomogeneities in the time series.*

The sensors which measure WTL are taken out every autumn to prevent damage through freezing. Unfortunately, the exact installation depth is not completely documented for 2014, only showing that it is not comparable to the following years. The jump seen in WTL is thus artificial. The reference point at that year is different than in the other years, and unfortunately there is no record of it. However, there is another two-year data set (2014-2015) we can use to introduce a correction factor to WTL data from 2014. We will make this change to our data and re-run the analysis with WTL. After correction the WTL data is in line with soil moisture data, indicating the 2014 to be the driest year (Figure below).

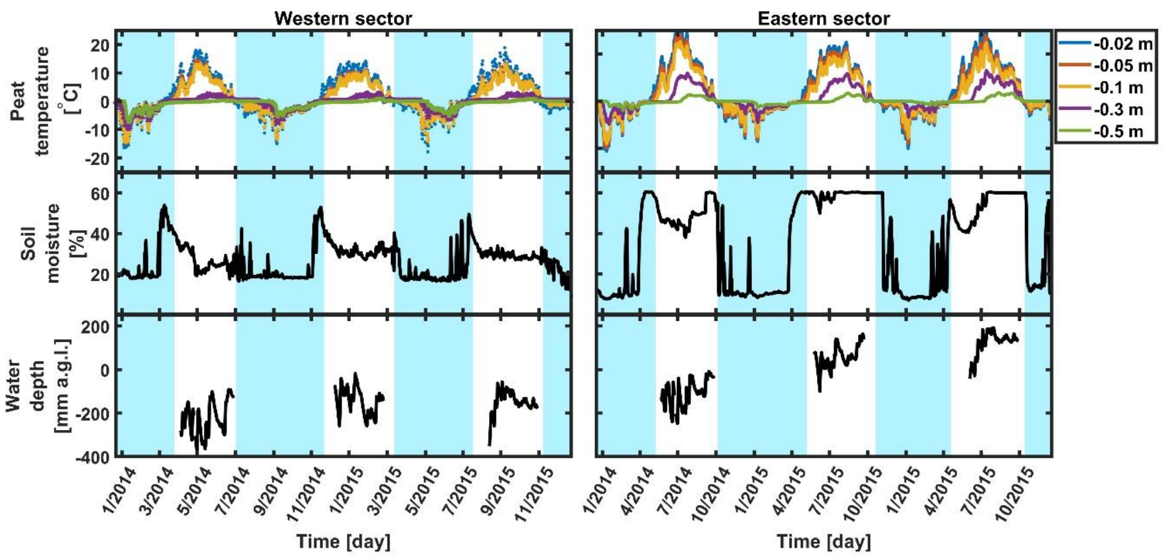


Figure1: Figure 2 of the manuscript, with corrected WTL data.

l. 387-390: Please write more specific, e.g. "...on average over all three years more than 90% to the fluxes measured at the eddy covariance tower."

You need to write at least "We replace…", or something, to begin the answer.

"Footprint and flux contribution of drier and wetter areas are presented in Figure 3. The dry areas (yellow) contribute to more than 90 % fluxes from the western sector. In the eastern sector, the wetter (blue) and drier areas contribute almost equally to the fluxes. The contributions of the wet and dry areas to the fluxes in both sectors were stable across the three study years" to.

"Footprint and flux contribution of drier and wetter areas are presented in Figure 3. The dry areas (yellow) contribute *on average over all three years more than 90% to the fluxes measured at the eddy covariance tower* from the western sector. In the eastern sector, the wetter (blue) and drier areas contribute almost equally to the fluxes. The contributions of the wet and dry areas to the fluxes in both sectors were stable across the three study years" *l.392, Caption figure 3: Please describe more precise what is shown in the bottom panel. How are these average contributions of contrasting landcover types calculated. Generally, I would prefer another diagram type that allows evaluation of the variability of footprint contributions of the two landcover types.*

We prefer to retain the figure as it is now. Variability of the different land covers is quite stable during the whole measurement period as presented on the figure above. We will add the figure on the variability to supplementary material, and refer to it in the revised manuscript.

l. 419-420: How did you deal with the autocorrelation in the time series? Serial dependence of data points could lead to biased results of the Wilcoxon test.

Thank you for this comment. Autocorrelation existed up to 8 days. With this information I divided winter data to the subsets where every 9th day was taken. We tested the difference of those subsets without autocorrelation to zero with Wilcoxon rank sum test. Tests made in this way also rejected hypothesis that winter fluxes are equal to zero.

l. 455: Unclear what "breakout week" means

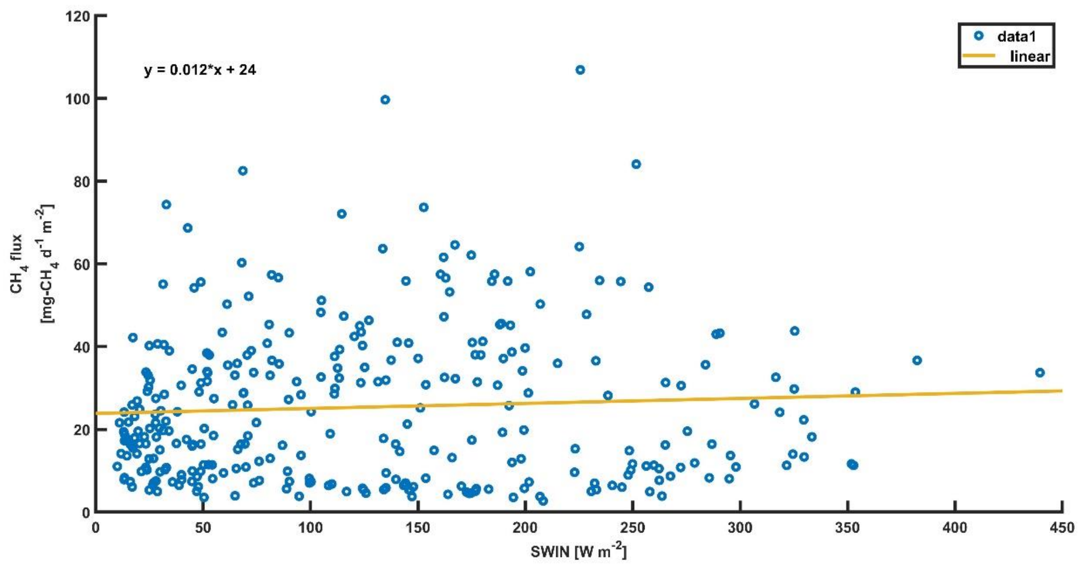
We will change the text from

"Figure 5. Weekly averages of CH4 fluxes vs surface peat temperature (top panels), vs the best correlated layer (middle panels), and vs the deeper layer (bottom panel). Data were divided into the beginning of the growing season (blue dots) and end of the growing season (orange triangles), where breakout week was the week with the highest emission." to:

"Figure 5. Weekly averages of CH4 fluxes vs surface peat temperature (top panels), vs the best correlated layer (middle panels), and vs the deeper layer (bottom panel). Data were divided into the beginning of the growing season (blue dots) and end of the growing season (orange triangles). Weeks with emission before reaching the maximum weekly averaged emission was defined as the beginning of the growing season. Weeks with emission after reaching the maximum weekly averaged emission was defined as the beginning of the growing season. "

l. 565-572: I find the discussion of the explanatory strength of incoming shortwave radiation confusing. The GLM parameters in Table S2 for the explanatory variable incoming shortwave radiation are all negative, indicating that methane emissions were lowered under high incoming shortwave radiation. Thus, the GLM results do not suggest strong relations between shortwave radiation, photosynthesis, substrate supply and CH4 production. Or was the sign convention for incoming radiation different than I assumed?

Thank you for the comment. We will rewrite this part of the discussion. The negative contribution of shortwave radiation can be due to the slight diel cycle of CH4 emission, with lowest values at daytime. Mechanistically we can think that the solar irradiance will heat the top of the peat layer, thus leading to increased methanotrophy at daytime (see discussion above on diel cycle). We can lead to situation where the methanotrophy is higher in sunny days with warm surface and lower in cloudy days. The role of photosynthesis for the substrate supply of methanogenesis is likely to act in the seasonal time scale, where its effect can be masked by the strong correlation between peat temperature and CH4 emission.



l. 724-725: The hysteresis-like behaviour can be also explained by the phase lags between different temperatures in air, ground surface and different soil depths

This is actually our point. In the papers by Chang et al., (2020; 2021) hysteresis with air and shallow peat temperature are shown, and as explanation, precursor availability is used. We want to show that the existence and direction of hysteresis-like behavior can depend on which depth the temperature is measured. As reviewer writes, the phase lag between the temperatures at different depths is likely explanation, at least partially.

Figure S1: Specify in the caption if soil or air temperature is shown. At which height in the atmosphere?

We will specify this in the text.

Technical:

We will apply all technical comments.