Comment on bg-2020-488
Joshua Ratcliffe (Referee)

General comments:

In this study the authors present growing season (May-September) CO2 and CH4 flux from a boreal mire in the south of Finland. They conclude that the effect of footprint heterogeneity on fluxes is negligible and that May-September temperatures seem to be the most important factor in determining the seasonal variance of the fluxes, with warmer temperatures leading to greater net CO2 uptake and greater CH4 emissions.

The study is quite novel as a multi-year CH4 and CO2 flux record from a boreal peatland, of which only a handful exist. I like the consideration of footprint heterogeneity in the analysis, which is novel, and the degree of detail the authors have provided in the plots and the results, which make it easier than usual to assess both the quality and variability of fluxes. While the dataset contains some very large gaps, especially in 2011 and 2013, the authors have discussed this in detail and have partially considered this when drawing their conclusions from the dataset, including presenting a reasonable estimate of gapfilling error.

While the study is interesting and I would ultimately like to see it published, I have a few critical points, including one major critique about how the flux driver data has been interpreted. I am particularly concerned that the relationship between temperature and fluxes may be an artefact of the measurement gaps in the timeseries. I also think that if the temperature/flux relationship is real, then the authors should explore this in more detail, and determine whether this is related to growing season length or to more fundamental biological processes.

Specific comments:

My main concern is that the seasonal trend in data gaps may invalidate the analysis done in Figure 12 and thus the conclusions about the flux drivers. For example, 2013 appears 6-7 degrees warmer than 2012 in Figure 12, this must be mainly due to the differences in data coverage. Perhaps more concerning fluxes are lowest in years where the authors had the best data coverage at either end of the growing season, periods which will also have lower fluxes. As such, the same time period is not being compared in each year and naturally the fluxes are highest in the years where data is missing from the early and late season. The authors could account for this by only selecting a period where there is data.
in all years (July August?). Alternatively, the authors could account for the seasonal influence by looking at the anomaly for the period in question, for example presenting the value of NEE/Ta etc. for one year, minus the mean value for all other years for the months which data is available.

If the temperature relationship is real I would like to see some more exploration of this. Is this effect due to growing season, in which case the authors could look at a metric such as degree days above zero, or PAR above zero, or is due to fundamental biological processes processes? Perhaps the authors could look at light response curves of NEE during different temperature conditions inorder to show this.

I find it puzzling that the authors choose to talk about the u star threshold and energy balance closure in the first paragraph of the discussion. Neither of these are the focus of the study and lack of energy balance closure in peatlands is often seen and might even be expected when soil heat flux is omitted. The ustar analysis is standard. I think this can be omitted or moved to methodology section.

The large gaps in 2011 and 2013 (appears to be around 50%) make these numbers questionable as seasonal estimates, for table 5 the authors might want to include an additional row with the averages excluding these years Technical comments:

L41: “large area” is a bit subjective, suggest being more specific or removing

L41: More recent estimates show that drained peatlands have now tipped this balance into a net warming effect. Suggest a qualifier such as “undrained peatlands” or “natural peatlands”

L51-53: I agree that chambers are unsuitable for this, but it would be good to back this up a little better. Can you cite some studies that show a divergence between EC and chamber estimates?

65-66: I agree with what is written here, i.e. “fairly wide spread” for flux totals. But in my view this contradicts several later statements L417 and L560 where the results are described as "similar to other bogs "or typical of other boreal bogs. Maybe these later statements should be amended to, “within the range seen in other boreal bogs” or “typical of some other boreal bogs such as x,y,z”

L69-71: These terms are all very subjective, warm temperature, ample sunshine etc. can they be more tightly described here?

L75: suggest “WTD is an important driver as it controls the thickness of the oxic zone”

L98. This seems unfinished. It’s been analysed in detail and what did they find?

L132: “Standard schemes and quality control” is rather vague, and the cited references offer several different options in this regard, such as Moncrieff or Fratini spectral corrections. Our own work on boreal peatlands has also shown the form of timelag compensation used (optimization vs. maximisation) can have a large impact on the processed fluxes, especially when fluxes are low, and it is not clear from Sabatini et al., 2018 which of these the authors used.

L179-180: There was presumably some impact from trampling in 2011? Can the authors state here if this was the case?

L217-220: I am not sure respiration or photosynthesis can be well modelled in peatlands using Q10 or Micaelis-Menthen, however depending on the site and combined with the
sliding window approach it is probably acceptable. Given the gaps in the dataset I am also not convinced alternative models or techniques would perform any better. I would encourage the authors to think about using alternative techniques such as ANN or random forest in future work.

L263: In one of the earlier figures the water table is shown lower than this, -25 cm

L307-311: I suggest this data is presented in a table (possibly as SI) and maybe the authors can replace this with a summarised version, stating what a plausible range for the winter fluxes may be (even if this is as simple as extrapolating median, upper and lower quartile daily fluxes)

Figure 6: I really like this figure, but can the authors include monthly tick marks? I really struggled determining exactly when the gaps occurred

Figure 7, 8: In really like these, but I would suggest having a consistent unit of time, probably months

L371-372: it’s not clear to me how summertime differs from growing season here, can you clarify this?

379-390: It’s great to read about the gapfilling uncertainty and how high it is, this seems entirely reasonable given the gaps in the data. Gapfilling uncertainty is only one source of error, choice of u star threshold, filtering thresholds and measurement error are all also significant. I suggest the authors justify why only gapfilling uncertainty has been calculated and state how they think a more comprehensive assessment may differ.

L393: please define “very low”

L395: the negative impacts of what on what?

L295: It would be good here to talk about whether low WTD is affecting GPP or ER and by how much for how long, what is written seems really vague.

L416-417: please add the calculated uncertainties to this

L415_430: This reads rather like a long list of sites and numbers with little discussion. Can the authors comment if there are any clear trends or distinctions across these sites, For instance, why is there a higher emission at Plotnikovo, or do we not know?

L425: I am not sure the Vompersky et al., 2000 reference is appropriate, the title appears to be referencing CO2 not methane and the study also pre-dates modern Eddy Covariance measurements of CH4, Fribourg and Roulet are also rather old studies now, from the early days of CH4 Eddy Covariance. Some more up-to-date comparisons would be good.

L433: how much is a “small decline”?

L455: I would say your figures show this clearly, this sounds rather uncertain.

L449: This sentence seems unfinished

L514-517: This reads like a list of different findings, can it be synthesised a little more?

L557: This section seems to be missing a concluding sentence that ties it all together.

L551: How is it that boardwalks are overestimates, compared to other features? This is
not clear to me

L552: but presumably, people were walking over the locations where the boardwalks were... I am not sure you can dismiss their impact for this reason

L559-560: Again, if you earlier sate how variable fluxes are then say Siikaniva is typical it seems like a contradiction. Maybe re-write.

L577: Perhaps the authors can also comment on how these limitations can be overcome?

Technical corrections

L295: possible typo “from dome”

L455: should be “a net annual emission”

L376: should be weekly to seasonally