Response to the reviews of the preprint bg-2020-488, Alekseychik et al. 2021 "Carbon balance of a Finnish bog: temporal variability and limiting factors"

The responses are in blue.

Reviewer 1 (Joshua Ratcliffe)

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 C02 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.

Dear Joshua, thank You very much for the positive evaluation of this work and the many useful comments and criticisms. We made efforts to fully address each point you have mentioned.

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.
This is absolutely correct. These issues resulting from the use of measured fluxes should be articulated better in the discussion of Fig 12. We chose to assess the effect of the drivers on the cumulative measured fluxes, instead of on cumulative gap-filled fluxes due to the uncertainties in the model during the long gaps (this mainly concerns 2011 & 2013).
We did experiment with the gap-filled fluxes and tried using them to demonstrate the effects of environmental drivers. We abandoned method as it was difficult to tell how much role the model component played in the results, which is quite high in some of the years as you have noted. However, the versions of Fig. 12 using the original flux (currently in the draft) and using the gap-filled flux

(previous) looked similar (see below). This may also be taken as an indirect implication that our modeling approach is viable.

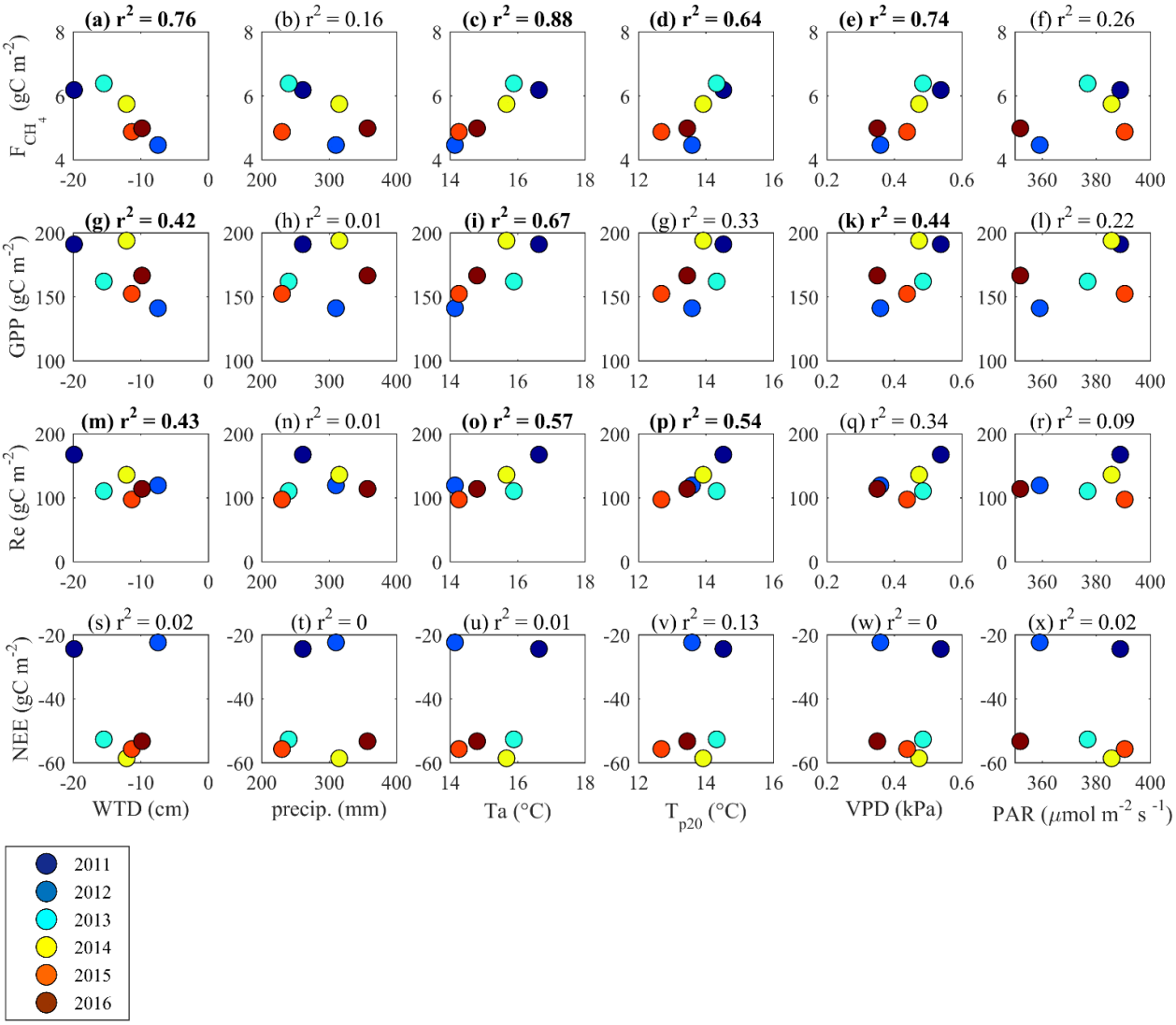

Fig. 1. Old version of Fig. 12, using gap-filled June-August fluxes.

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.

The suggestion to use monthly data to show the anomalies in fluxes and drivers is very good, and we will add a corresponding figure. The record is best broken down into months because of the uneven distribution of gaps.

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 in order to show this.

This is shown by the temporal trends in the NEE model parameters k and Pmax (Fig. 7) and the response of Pmax to the variation in WTD (Fig. 8).

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.

We would like to stress that the result shown in Fig. 2 is novel and the use of EC CH_4 flux gives promise of better performance in the determination of the u_* threshold than EC NEE. Soil heat flux is omitted as the focus is on the turbulent heat fluxes that change dynamically with u*. However, as this section appears to be out of place, we will move it to the Appendix.

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.

This sounds reasonable indeed, will be added.

Technical comments:

L41: "large area" is a bit subjective, suggest being more specific or removing – here I refer to their large proportion in the Boreal landscape land area, will be rephrased

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" – will be specified, thank You for the suggestion. Siikaneva-2 bog is a fully natural mire, and the entire discussion is certainly of relevance mainly for the natural, undrained mires.

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? We have to admit that this statement looks somewhat EC-centric and biased. Chamber studies did produce such estimates – although the possible interpretations, of course, differ from the case of EC. The attempts to reconcile the EC with the chamber fluxes in this very site can be found in several papers by Korrensalo et al. and Männisto et al. These will be mentioned.

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"- I totally agree, "within the range seen in other boreal bogs" sounds better, given the large spread. The basic problem here is that the "bogs" are such a diverse group of ecosystems that any averaging across them is, strictly speaking, wrong. Most studies do it anyway, due to the lack of data.

L69-71: These terms are all very subjective, warm temperature, ample sunshine etc. can they be more tightly described here? – will be rephrased. However, these are simply relative qualifiers: "favourable conditions" are such that lead to the highest net C sink possible for the given ecosystem, by definition.

L75: suggest "WTD is an important driver as it controls the thickness of the oxic zone" – will be rephrased as suggested

L98. This seems unfinished. It's been analysed in detail and what did they find? – I will make this part about the study of Tuovinen et al. (2019) more specific.
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. – The fluxes were processed using EddyUH software which is summarized in Mammarella et al. (2016) and combines the methods from other literature cited in this paragraph. Concerning the time lag approach: maximization of cross-covariance is used.
L179-180: There was presumably some impact from trampling in 2011? Can the authors state here if this was the case? I was intending to comment on this later on in the manuscript but forgot. It is an interesting question. We are only able to test for the possible effects of trampling post-factum by looking at the data. The lack in the directional different in flux model parameters between 2011 and the later years suggests a minor or absent effect of trampling; had the opposite been true, one would have observed an increase in the reference fluxes of Re and F_{CH4}, and maybe a drop in Pmax in the western sector starting from 2012 when the boardwalks were build. Besides, there are no significant traces of trampling around the present western boardwalks.
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. – We assure you that those fluxes conform very well to the Q10- and Michaelis-Menthen-type functions and this holds throughout the season. However, we will definitely try to provide alternatives using other techniques in future studies.
L263: In one of the earlier figures the water table is shown lower than this, -25 cm. Indeed, Fig. 5e shows the lowest WTD of 2011 equal to about -25 cm. Will be corrected
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) – Thank You for the idea, we will organize these results as a table.
Figure 6: I really like this figure, but can the authors include monthly tick marks? I really struggled determining exactly when the gaps occurred – Sure, apologies for the difficulties with interpreting this figure, it will be improved. The ticks will be added. Note that the grey background marks the May-Sep period.
Figure 7, 8: In really like these, but I would suggest having a consistent unit of time, probably months – Fig. 8 will be remade with months in x-axis.
L371-372: it's not clear to me how summertime differs from growing season here, can you clarify this? – Summertime is June-August, growing season May-Sep. These are typically used by us but may be not as obvious to the others. To be specified.
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. Gapfiiling 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. – Gapfilling error is maybe the easiest to estimate as it "only" requires a set of several gapfilling trials using different approaches. Moffat et al. (2007) estimate it at 25 g C / m2 year – similar to my assessment for Siikaneva-2. The other error sources are notoriously difficult to approach as they require collocated EC sensors. As shown in this prepring (Fig.1) the u* threshold is well defined so it cannot introduce

a large error. Systematic errors related to measurement and EC data postprocessing were probably smaller than that from that induced by gapfilling.

L393: please define "very low" – based on Fig. 11 I would say the fully "tolerable" WTD limit is at about -15…-20 cm.
L395: the negative impacts of what on what? Of WTD on the reference fluxes.
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. Seasonal and sub-seasonal WTD variation is not in phase with the variation in fluxes, so we can only speak of seasonal averages.
L416-417: please add the calculated uncertainties to this – will be added
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? – I will rework this section. However, the vast differences in bog subtype and vegetation cover complicate such comparisons.
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. – I will try to find more recent relevant studies, but this is basically all that is available for boreal bogs at the moment.
L433: how much is a "small decline"? – 22% of the cumulative May-Sep NEE, in relative expression (47 g C m-2 in 2011 vs. 60 g C m-2 on average). Using only Jun-Aug results in a NEE decline of 33% (understandably as the drought occurred in these months). I called this a small decline as NEE suffers from dry weather much more in some other mires.
L455: I would say your figures show this clearly, this sounds rather uncertain. – I will refer to the figures, thanks for the suggestion. I agree that this result is apparent.
L449: This sentence seems unfinished – "during drought" is missing.
L514-517: This reads like a list of different findings, can it be synthesised a little more? – I will work to make this paragraph flow better.
L557: This section seems to be missing a concluding sentence that ties it all together. – will be added.
L551: How is it that boardwalks are overestimates, compared to other features? This is not clear to me – boardwalks are thin linear features stretching over the surface of the bog. They have a width of ca. 30 cm, and so are essentially sub-grid features as the map resolution is 1 m. However, the resolution-coarsening algorithm picked them up and assigned the "boardwalks" surface type to all 1-m pixels through which they passed. This results in ca. 5-fold artificial increase in their area, which directly applies to the boardwalk contribution shown in Fig. 9. This is different from the natural surface cover, which typically consists of patches of >1m^2, and are also averaged over a much greater area, so the systematic error in their contribution is much smaller.

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 - That is correct, we cannot completely rule out the possible effects of the boardwalks' installation. This will be mentioned.
L559-560: Again, if you earlier sate how variable fluxes are then say Siikaniva is typical it seems like a contradiction. Maybe re-write. – Will be rewritten.
L577: Perhaps the authors can also comment on how these limitations can be overcome? – Our attempts will be briefly summarized here, although we cannot offer a definitive solution at this point.

Technical corrections
L295: possible typo "from dome" – done (see the response to Reviewer 2)
L455: should be "a net annual emission" - done
L376: should be weekly to seasonally - done