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 2 (anonymous)

I was excited to read this paper given the huge amounts of data and the relatively long measurement period (6 growing seasons! Both CO2 and CH4 measurements), as well as having really all the important environmental variables and fluxes measured concurrently. I think that the conclusions are more or less supported by the results in a logical way and given what I know about the site from reading other papers (Korrensalo, et al….). However, as a peatland expert who knows a lot about C fluxes and modeling but doesn't use EC techniques, the results section manuscript was incredibly challenging to read and to follow. Fortunately, the discussion section mostly redeemed it; the authors did a nice job of integrating the results of this study with earlier studies at this site and across northern peatlands.
The challenge was that I was not convinced on the appropriateness of the modeling and the subsequent analysis of the model parameters. This is partly a result of the framing; I thought that these parameters were simply used for gap-filling (e.g. Table 5) but instead these made up the bulk of the results. An analysis of the modelling parameters used in the flux calculations comprised the whole of Sections 3.2, 3.3, 3.4, 3.5. This was problematic because the explanation of the modelling was insufficient and unclear and the justification for the approach, both theoretical and practically, was quite weak. First, the parameters are not even named (defined) or explained (Section 2.4). Then, it is not clear how these parameter values were determined. Or rather, it was clear until I read the results section and looked at Figure 7 (which shows something different than Table 4), 10, and 11 which shows that these parameters are dynamic over time, then I was completely lost. What, how and why the modeling was done in this way must be clear. This analysis comprises the bulk of the results section so the explanation needs to be clear, include the relationship with time, and can take some space.

Furthermore, I was not convinced that the model structure is completely appropriate until I read the explanation in the discussion and got a refresher about other results from Siikaneva; there is no justification for the use of these models for this site other than an earlier study used similar methods. It isn't apparent that the authors tested alternative model structures for either Re or CH4 that might include other known controls on CH4 flux (like water table). From the discussion, a bit more insight emerges as to why the authors chose these particular models for the C flux parameters but this needs to be justified in the model description section with references to the earlier studies from this site.

Thank You for making these important remarks! We regret that the model/gapfilling method description came though as overly complicated. Perhaps a clearer explanation will improve this part and provide the necessary justification.

The somewhat unconventional modeling/gapfilling method pursues two aims:

1) good performance given a dataset characterized by a mixture of long and short gaps

2) provision of temporally varying parameters which can be used to analyze the possible drivers (namely, the model parameters)

Regarding (1), the present approach is a result of a lengthy work with the Siikaneva bog dataset. We did attempt to model GPP, Re and CH_4 flux with a range of "single-fit" models involving WTD, θ, Ta, RH, surface (~stomatal) conductance and so forth, some of which are mentioned on L292-295. However, the performance of those was questionable. There are two problems with the "single-fit" approach: a) moisture availability (WTD, θ, RH), conductance (gs), and other quantities of potential importance have a comples relationship with the fluxes – this was verified by residual analysis (not shown); b) the parameters of the "single-fit" models are most influenced by the seasonal cycle in the drivers, as this time scale provides the highest range in the driver values – but the drivers are highly correlated with each other on this time scale, meaning it's impossible to properly separate their effects on the fluxes (see e.g. Fig. 8).

This last leads directly to the point (2). The lack of clarity in the exact flux-driver functional relationship motivated the extensive use of the Re, GPP and CH_4 model parameters. We note that the reference fluxes of the exponential models of Re and CH4 emission and the maximum photosynthesis are qualitatively the same as normalized fluxes: their variation shows not the course of the actual flux, but the flux normalized by the model driver, i.e. T_{soil} for CH_4 flux and Re, and PAR for GPP. Thus, the model parameters lack a seasonality induced by these drivers and are more informative than actual fluxes in the case of a gappy record such as ours. Rinne et al. 2007, Rinne et al. 2018 and Rinne et al. 2020 have successfully used normalized fluxes to elucidate the controls – we aimed to do something similar without copying their methodology.

Despite the apparent complexity of the modelling method explanation, it is based around a simple idea of combining two different models:

- Model 1, using the "single-fit" approach to fill the long gaps (the parameters of those are found in Table 4). As the variation of the model parameters in long gaps cannot be established, it would be the safest to assign constant values. Certainly, this leads to potential error which You have pointed out and which is acknowledged in the manuscript.
- Model 2, based on recalculation of the model parameters in a moving time window with a daily step (Eqs. 1-3). This model provides the temporally varying series of the model parameters (e.g. Fig. 7). Given the time window width of 5-15 days (differs among GPP, Re and F_{CH4}), the resulting parameter time series are representative of variations on a similar time scale.
- The fluxes are gapfilled using the combination of the two models.

As said above, the model parameter time series provide information similar to the mode widely used normalized fluxes. We aimed to investigate variation of the model parameters on weekly to monthly to seasonal scales and look for matching variation in the environmental parameters that caused it. It did prove to be challenging, but the insights that seemed most reasonable are discussed in the text.

We are grateful to You for suggesting several references to support the Materials & Methods and Discussion. In fact, I independently found a similar Gaussian behavior in model parameter seasonalities as shown in Rößger et al. (2019) (Scaling and balancing…). I even made a gapfilling experiment, fitting Gaussian curves to the model parameters which seemed to provide good estimates of the parameters outside the measured period or during long gaps. However, I later realized that the wide interannual variation (Fig. 7) still cannot be captured, so this Gaussian modeling doesn't have an advantage over assigning constant parameter values for the whole gap (i.e. using a single fit as detailed above).

Justification and clarification based around the arguments given above will be incorporated in the text.

Finally,
Table 1 shows that many years had only a small amount of data meeting the QC criteria. I'm a bit concerned about the circularity of using modelled fluxes (including the gap-filled data) that have been modelled given prescribed controls (with insufficient justification for the use of the models) to look at the controls of the fluxes, given that these fluxes were modelled using temperature. A more rigorous analysis is justified.

This is a reality of Siikaneva-2 and many other sites – poor data coverage caused by technical issues. The current study, to a large extent, is an exercise in interpreting a gappy dataset and trying to tease out meaningful information. Even with such a challenging dataset, one is still tempted to inquire into which drivers caused the interannual variations. An approach we took to minimize the effect of model domination was using only measured fluxes in Figs. 12-13 and the related discussion. Table 5 reports gapfilled fluxes, but it doesn't have the aspect of interannual variability.
We will add a plot of monthly averages vs. drivers (similar to Figs. 12-13), using only the month covered well enough by data.

Specific comments:
Introduction: I really disliked that the background for this manuscript relied only on information and background from other EC measurements in peatlands, particularly from bogs. Linking fluxes to sub-surfaces controls was originally done using chamber measurements; these have really laid the foundation for understanding environmental controls on fluxes using EC, including at this site, and some did this more than 10 years ago.

We thank you for this substantial criticism, and will improve the Introduction by adding a paragraph on the results of chamber studies and in-situ sensor campaigns.

75: Why are all these referred to as "potential" drivers? These are known drivers, at least at other sites. – Will be rephrased as appropriate.
162: Wasn't 2011 hot and dry? Is this appropriate? Using the mean LAI course is the best solution we could think of. Interannual LAI did not show any clear relationship with the environmental drivers – neither peak LAI nor mean LAI. Surprisingly, this was also true for GPP. However, the seasonality of Pmax is undoubtedly related to that of LAI.

254: This is really only the result section given that there is a later discussion section - Corrected
295: Dome? – supposed to be "from the typical dome-shape"
Section 3.3: could use figure references. How and where was this non-growing season flux determine? – Figure references will be added. The non-growing season fluxes were estimated based on the small amount of data available outside the May-September periods (can be seen in e.g. Fig. 6)
Figure 7: this is confusing (see main points above) given also Table 4. – with the explanations I have included in the beginning of my response, this should become clearer. Fig. 7 essentially shows the outcome of the second model, the one using a moving time window approach.
350-353: Confusing – these fractions refer to the range between maxima and minima of the curves plotted in Fig. 10. Rephrased.
Table 5: Why not add indicate the error here? Especially because of disclaimer on line

359? – Error estimates added
385: 30-40% of what? – this is relative uncertainty, so fraction of the seasonal cumulative NEE value.
4.1: where is this shown? – this refers to Fig. 1. As per the request of the other reviewer, we decided to make this Appendix 1. The result is novel but doesn't fit in with what follows.
433: interannual difference controlled by temperature, but how is this related to and dependent on the model used here? – as discussed above, the model does make an impact - especially in the gappy 2011 and 2013. We will additionally stress this limitation.
441: Not shown in Fig. 13 – apologies, should be Fig. 11.
481: other studies have shown differently (e.g. King et al., 1997). – This reference will be added. Admittedly, direct conflicts between the previous studies do not make things any easier!
507-8: Unclear what this paragraph is referring to? Maybe include some references to figures and tables. – I think You may have specified wrong line numbers, but if this is about the non-growing fluxes, then the relevant results were mentioned in L304-311 and Fig.6.
555: include a figure reference here. – Added, should be Fig. 10.
560: Could this be related to the gap-filling or modeling methods? – I made three attempts to do gap-filling, with different approaches, before the one described in the current preprint. All produced very similar 6-year mean May-Sep cumulative fluxes. The 30-40% relative uncertainty on NEE that was mentioned refers to the cumulative values of the individual years; the 6-year mean NEE must have a lower uncertainty which I would estimate as ±5 g C or about 10%.
Data availability: Given that this is 2021 and there are many opportunities for data publishing and indeed this is generally required, contacting the author for data is really not an acceptable route for data availability. – Of course, we understand and will publish the original data used in this MS on an official repository with a DOI.

Additional references to consider:

Chadburn, S. E., Aalto, T., Aurela, M., Baldocchi, D., Biasi, C., Boike, J., ... & Westermann, S. (2020). Modeled microbial dynamics explain the apparent temperature sensitivity of wetland methane emissions. Global Biogeochemical Cycles, 34(11), e2020GB006678.
Helbig, M., L. Chasmer, N. Kljun, W. Quinton, C. Treat, O. Sonnetag (2016). The positive net radiative greenhouse gas forcing of increasing methane emissions for a rapidly thawing boreal forest-wetland landscape, *Global Change Biology 23*: 2413-2427, doi: 10.1111/gcb.13520.
King, J. Y., & Reeburgh, W. S. (2002). A pulse-labeling experiment to determine the contribution of recent plant photosynthates to net methane emission in arctic wet sedge tundra. Soil Biology and Biochemistry, 34(2), 173-180.
Rößger, N., Wille, C., Veh, G., Boike, J., & Kutzbach, L. (2019). Scaling and balancing methane fluxes in a heterogeneous tundra ecosystem of the Lena River Delta. Agricultural and Forest Meteorology, 266, 243-255.
Rößger, N., Wille, C., Holl, D., Göckede, M., & Kutzbach, L. (2019). Scaling and balancing carbon dioxide fluxes in a heterogeneous tundra ecosystem of the Lena River Delta. *Biogeosciences*, *16*(13), 2591-2615.
Treat, C.C., J. Bubier, R.K. Varner, P. Crill (2007). Time-scale dependence of environmental and plant-mediated controls on CH4 flux from a temperate fen. *Journal of Geophysical Research- Biogeochemisty* 112: G01014.