Point-by-point response to reviewer comments

Vainio et al., Biogeosciences

This point-by-point response to the reviews is based on the earlier responses given during the open discussion.

Main changes made in the manuscript:

- The modelling approach for the CH4 fluxes was changed: we model the CH4 fluxes directly with the RF technique and using soil moisture as a driver in the model.
- The Discussion section was completely re-organized and subtitles were added, also some new text was added in the Discussion based on the referee comments.

Referee #1

1. The modeling framework

RC: Line 225: How many observations did you have for May-July and August-October? Did you have many measurements from one point in your model (e.g., early May measurement, late July measurement)? Can you be certain that the soil moisture measurements conducted within one study period (e.g. in early May and late July) can be directly compared and used in the same model, even though the soils tend to get drier during the summer? I think it's important that the measurements that you use for your response variable (i.e., soil moisture) are fully comparable with each other.

AC: There were around 6 measurements (median) from each sample point in both of the seasons. Our data is not temporally very comprehensive, and thus we opted to do two static periods. Still, we wanted to take into account that the soil does, in fact, get drier towards autumn. We assume that this two-seasons-strategy is good compromise with the data we have. We actually started the measurements in late May (line 132), so we missed the wettest period in the spring, and the most active measurement period was in June–August (line 135). However, we still ended up having approximately same amount of data for both seasons.

RC: 225: Could you also describe why you decided to use soil moisture as a predictor of fluxes instead of using the different topographic indices directly? Also, did you consider creating a continuous vegetation type raster based on your vegetation classes and the gridded layers for the study domain? This could have been a useful predictor for CH4 fluxes as well.

AC: We re-evaluated our modelling approach and changed the modelling: we model the CH4 fluxes directly with the RF technique and using soil moisture as a driver in the model (see Fig.1 in this document). This change was motivated by the fact that bulk of the criticism from the reviewers was directed towards the CH4 flux modelling approach. The manuscript was modified accordingly.

We considered but did not create a continuous vegetation type raster, because there are many drivers affecting the vegetation, and we did not have data of all such drivers. There is no direct connection between e.g. vegetation and soil moisture (Fig. 4), nor with topography. A continuous vegetation type raster would have probably required more thorough mapping of ground vegetation of the entire area.

RC: 256: Why didn't you use the similar framework that you used for soil moisture to predict CH4 fluxes with soil moisture? You could have created a RF (or a GLM/GAM or some other) model with the measured soil moisture as a predictor, and then used that model to predict fluxes across the landscape using the predicted soil moisture. And this could have been repeated over the different bootstrapped soil moisture maps to get CH4 flux uncertainty map as well.

AC: See the previous response, we changed our modelling approach to follow this referee suggestion.

RC: 240: Could you add the response graphs (partial dependence plots) describing the relationship of these indices and soil moisture to the Appendix?

AC: These were added to the Appendix.

2. Description of the model performance

RC: Line 20: Somewhere here I would add a sentence about how the statistical models performed, and how reliable your results are.

AC: We added something more about the model to the abstract.

RC: 377: I would be interested to see a scatterplot of the observed and predicted (CV) fluxes to see how well the model predicts high and low soil moisture values. Same applies to CH4 flux (line 406).

AC: We added the scatterplot showing cross-validation results to the Appendix.

RC: 542: Somewhere in the Discussion you should also discuss how well your upscaling performed compared to previous studies. What are the main uncertainties, and how can these uncertainties be reduced? What predictors are you missing? How about other RS-derived indices, such as NDVI?

AC: According to our current impression, there are not many published studies trying to upscale CH4 fluxes from forest stand like ours. The previous study of (Sundqvist et al. 2015) used simple soil wetness and temperature relationship to upscale CH4 fluxes, whereas Kaiser et al., (2018) and Warner et al., (2019) studied more southern ecosystems. Furthermore, comparisons of e.g. cross-validation results are hampered by different cross-validation techniques used in different studies. In this manuscript, we utilized distance-blocked cross-validation since it is argued to produce more realistic estimates of cross-validation metrics than other techniques (Roberts et al., 2017). For instance, with traditional leave-one-out cross-validation the predictive performance of our RF model would seemingly improve (e.g. r^2 increase from 0.51 and 0.26 to 0.67 and 0.56 for May–July and August–October, respectively). Therefore, we argue that direct comparison of cross-validation metrics between studies using different cross-validation strategies is not feasible. We added a section to the discussion on how these predictors are missing in upscaling exercises. The scale where our measurements are carried out cannot be directly applied to similar spatial scale than remote sensing methods that is used to estimate NDVI, despite the rapid development of satellite products.

RC: Title: I would consider adding the word “statistical modelling” somewhere in the title

AC: We changed the title to “Topography-based statistical modelling reveals …”

RC: Line 19: I would say “using digital elevation model-derived topographic indices” instead of “topography”

AC: We modified this as suggested.

RC: 28: I was a bit surprised to see this methodological suggestion as a final sentence concluding your study. I would consider changing this to something more broader, e.g. to the sentence on line 565-572.

AC: We modified this as suggested.

RC: 71: “Large amount of measurement points” is a rather subjective statement as for some people this might mean hundreds or thousands of observations. Maybe define the rough amount of measurement points instead, and mention that this is more than has previously been used.

AC: We clarified this sentence in the following way: “In this study, we used relatively high number of measurement points (60 points on an area of ca. 10 ha) in order to fully cover the small-scale spatial variability in the CH4 flux and its driving forces. Similar type of studies using chamber measurements are rarely based on more than 20 measurement points.”

RC: 75: With one driving parameter (i.e. soil moisture), right? You didn’t have many driving parameters to make the upscaled CH4 flux map?

AC: Yes, originally we used soil moisture only. However, due to the Referee #1 comment we re-evaluated this approach and based on this re-evaluation opted to follow reviewer suggestion to use RF model also for CH4 fluxes.

RC: 76: But what about Kaiser et al., 2018?

AC: We modified the text to: “Only a few studies (Kaiser et al., 2018; Warner et al., 2019) have applied similar approach, of which Kaiser et al. (2018) at a boreal coniferous forest, emphasizing the novelty of this study.”

RC: 105: You could add an index map to this figure showing e.g. the location of Hyytiälä, too.

AC: In the end, we opted not to add a new map.

RC: 210: I would describe these gridded layers in their own paragraph, similar to the other environmental measurements, and dedicated this one to the models only.

AC: Good suggestion, we separated these under a new subheading.

RC: 220: Could you provided a little bit more information about what parameters were chosen for the different indices? For example, TWI can change quite a bit depending on what parameters you use in the calculation.

AC: We added more information on these to the methods.

RC: 370: This figure could be moved to the supplementary – it’s not so useful for the reader because there are so many different points.

AC: We think that this figure gives a nice overview of the measured CH4 flux and its variation at the measurement points, and it is rather easy to see the spatial variation from this figure. (E.g. Fig. 4 is not giving this information, and a table would be more difficult to read.)

RC: 400: This is just an idea, but you could also replace these two maps by a map that describes the mean summer soil moisture and a map that describes its change over the growing season. It might be easier for the reader to spot the areas that are drying this way.

AC: This is an interesting thought and good to consider. However, we think that it might be more difficult for readers to understand what was done in this paper based on such a figure.

RC: 431: I would use the same color scheme that you used in Fig. 8 for the Fig. 9 as well, to make sure that you are using different color schemes for soil moisture and CH4 fluxes.

AC: Diverging colormaps (as the one used in Fig. 8) are suitable for data sets containing negative and positive values, because this way it is possible to emphasize the difference from zero. However, based on

Referee #2 comment, we modified Figs. 7 & 9 to show relative uncertainties, and as a result Fig. 9 colormap was changed.

RC: 440: The discussion is rather long and without subtitles it is a little bit hard to follow. Could you consider adding a few subtitles and structuring it according to your main aims of the study (spatial variation, drivers and upscaling, hot spots)?

AC: We divided the discussion under subtitles by re-organizing the text.

RC: 560: Again, I was a bit surprised to see this discussion here as it was not motivated in your introduction or it wasn't one of your main aims of the paper. Maybe include it to the introduction or remove it completely?

AC: This is a good remark. We added shortly to the introduction that previously usually fewer measurement points have been used in soil chamber CH_4 measurements, with the assumption that they are representative for a larger area.

RC: 567: If you want to discuss the sampling strategy, I would provide some more details here. E.g., how should the sample points be selected (e.g. systematic grid, gradient, random sampling, researcher-defined)? What is the number of temporal replicates required to understand spatiotemporal variability in this system? Further, in the abstract you mention that capturing the environmental variability requires 15-20 sample points. But do you think using statistical methods (e.g. random forest) with 15-20 points is reliable?

AC: Thank you for this comment. We added shortly in this paragraph about the selection of sample points: in our opinion, e.g. the elevation maps would be useful when selecting the sample points. With this study we were not able to reveal more high-frequency temporal variability, but we can add a sentence here that more frequent measurements would be needed for that – ideally it would require automatic chambers measuring e.g. once per day, or at minimum manual measurements every week.

The idea of the analysis shown in Fig. 10 in the manuscript was to evaluate how many chamber measurement locations were needed to get an accurate estimate of landscape-level flux by only averaging over the measured chamber data without any upscaling with RF. This was clarified in the manuscript, as this was also commented by Referee #2.

References:

Kaiser, K. E., McGlynn, B. L. and Dore, J. E.: Landscape analysis of soil methane flux across complex terrain, Biogeosciences, 15(10), 3143–3167, doi:10.5194/bg-15-3143-2018, 2018.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F. and Dormann, C. F.: Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure, Ecography (Cop.)., 40(8), 913–929, doi:10.1111/ecog.02881, 2017.

Sundqvist, E., Persson, A., Kljun, N., Vestin, P., Chasmer, L., Hopkinson, C. and Lindroth, A.: Upscaling of methane exchange in a boreal forest using soil chamber measurements and high-resolution LiDAR elevation data, Agric. For. Meteorol., 214–215, 393–401, doi:10.1016/j.agrformet.2015.09.003, 2015.

Warner, D. L., Guevara, M., Inamdar, S. and Vargas, R.: Upscaling soil-atmosphere CO2 and CH4 fluxes across a topographically complex forested landscape, Agric. For. Meteorol., 264, 80–91, doi:10.1016/j.agrformet.2018.09.020, 2019.

Referee #2

RC: There are two larger general comments: 1) Concerning the upscaling method:

AC: We re-evaluated our modelling approach based on these Referee comments, and changed the modelling to follow the suggestion by Referee #1: we model the CH4 fluxes directly with the RF technique and using soil moisture as a driver in the model.

RC: On top of this, there is never a discussion of why the approach of modeling soil moisture and then CH4 flux is advantageous. Furthermore, the authors dedicate a large portion of this manuscript to the upscaling exercise, but barely, if at all, discuss whole plot scale fluxes. It would be interesting to hear how much the estimated net CH4 sources offset the plot level sink between the two time periods, and how uncertain their plot level fluxes are. After all, the primary purpose of upscaling is not to accurately predict CH4 flux at every individual point, it is to enhance our predictive capability of large-scale CH4 exchange in a way that reflects soil heterogeneity.

AC: See the previous response, we changed the modelling approach.

We agree with the referee that the primary purpose of upscaling is to get an accurate estimate of landscape fluxes and not at individual points. We revised the text so that this message gets across to the reader clearly, and added some discussion on the whole area flux and the modelling. This was actually one of the main reasons why the analysis on chamber location bias was done (Fig. 10). Typically, mean of CH4 fluxes observed at a handful of chamber locations is reported and considered as representative of ecosystem CH4 exchange, however, this neglects any bias stemming from non-representative sampling locations. By using mean upscaled CH4 flux as a reference, we were able to show that 15–20 randomly selected chamber measurement locations (out of 60 locations available) were able to produce as accurate estimate of the landscape CH4 flux as averaging over all the chamber data. This information should prove useful when designing future chamber measurement campaigns in similar locations aiming to achieve accurate landscape-level flux estimates without upscaling with e.g. machine learning techniques.

RC: It would be more useful to randomly subsample the flux observations and build soil moisture-CH4 relationships from the random subsets. Then see how upscaled fluxes based on these relationships compare to the predictions made using the whole dataset.

AC: We acknowledge that the wording on lines 422-425 in the original manuscript needed clarification, and this has maybe caused misunderstanding. The idea of the analysis shown in Fig. 10 in the manuscript was to evaluate how many chamber measurement locations were needed to get an accurate estimate of landscape-level flux by only averaging over the measured chamber data without any upscaling with RF. Here the mean upscaled CH4 flux was used as a reference since it accounts for the soil heterogeneity (see above). It was shown that average over 15-20 locations resulted in a similar bias as average over all the chamber measurement locations. This should be useful information for future chamber measurements in similar locations. We revised the text so that this analysis is clearer and describe the methodology briefly also in Materials and Methods.

RC: The authors should also report their modeled CH4 predicted fluxes for pixels corresponding to the sample sites, which would help explain whether differences in upscaled fluxes are caused by a model bias or because of the heterogeneity of the predictor variable domain.

AC: Thank you for this comment, we added the modelled flux of the sample point locations (pixels) to Table 2.

Specific Comments

RC: Abstract: If possible, add some descriptive statistics (i.e. mean, min, max) for what CH4 fluxes were observed in each season.

AC: These were added.

RC: 21: The wording "as well as on the related ground vegetation" is confusing to me.

AC: Corrected "as well as from the related ground vegetation".

RC: 33-35: I do not believe that this is the current paradigm. Observed methane fluxes are the net sum of both opposing processes occurring in the soil.

AC: While it is the prevailing paradigm that the availability of oxygen mainly controls these processes in nature, and thus in general upland soils are no favourable place for CH4 production, there are possibilities for CH4 production taking place in these soils too, as we state in the next paragraph. The recent study of Treat et al. (2018) reported that during non-growing season upland soils can be a significant source of methane although they are generally considered as sinks during the growing season.

RC: 40: This is not an instant effect, however, which has implications on the influence of both total soil moisture and its temporal variability. It would be useful to note this here.

AC: We agree that the soils do not turn immediately sources of CH4 after inundation, but there are time lags between these processes. We added the following sentence: "However, there are likely notable time lags between the start of inundation and methanogenesis, complicating the analyses of dependencies between these processes.".

RC: 53-56: I think this is a great point, but I am confused why it is in this paragraph.'

AC: Thank you for this constructive comment but our impression is that this is logically introduced as it is.

RC: 57-62: This paragraph is important but it could be written more clearly. Are the authors trying to say that we often consider ecosystem fluxes in large-scale models but have not adequately accounted for heterogeneous sources/sinks within the ecosystem (which likely respond to environmental changes differently)?

AC: Yes, the sources and sinks within the ecosystems are not adequately known or accounted. We added a sentence in this paragraph, clarifying this.

Methods:

RC: 152: Is there an explanation for why one measurement would be so large?

AC: The sample point was located on small water pond, where the water table level was most of the time above the peat surface. The CH4 emissions from this sample point were at the same level with typical peatland emissions. One possible explanation for the highest emission would be ebullition.

RC: 177: It would be good to provide a +/- range for what types of temperature variation were observed here as justification.

AC: We added the range of average temperature at the sample points in late June 2013, when the temperature measurements at all the sample points were started.

RC: 220: Was the TWI then resampled/interpolated or left at coarse resolution? Additionally, I would hesitate to say that the TWI is "not accurate" at fine scales since it is simply a statistical metric and not a measurement of anything. It does, however, have limited application for estimating soil moisture on very high resolution DEMs because the metric itself is very sensitive to surface microtopography and noise.

AC: Yes, TWI was then interpolated with bilinear interpolation to the finer grid (5x5). We clarified the text.

RC: 225-242: I believe this paragraph could be rewritten to be better related to this study. A lot of the information on the inner workings of the RF algorithm can be condensed with an appropriate citation and a brief note on the advantages/disadvantages of RF over simpler techniques like multiple regressions. More information on why the model parameters (ntree = 300, mtry = 2) and predictor variables were selected would be helpful.

AC: We believe that some description is needed since the method is not necessarily very familiar to the chamber flux measurement community. Minimum number of observations per tree leaf was set to 2, due to limited amount of data. Small number for this parameter helps when trying to capture dependencies at both ends of the training data distribution, especially when using limited dataset. Amount of trees in the RF model (ntree) was initially set to 300 based on prior experience with RF and value for this hyperparameter did not significantly influence results. However, the number of variables randomly sampled as candidates at each split (mtry) was not changed from its default value (one third of the total number of variables). The predictor variables were selected based on Spearman's rank correlation coefficient.

Results:

RC: 304-306: Having S.D. values or some indicator of variability in soil moisture besides these means would be helpful here and in other parts of the results section.

AC: We added the SD values in text with the reported mean values.

RC: 326-327: What was the data of the highest emission outlier? It could be nice to see where it and other CH_4 measurements fall on the time series graphs above.

AC: The CH_4 highest emission (1080 µmol m^{-2} h^{-1}) was measured on 5 June 2013. We added the date of the measurement in the text (line 326). It is unfortunately difficult to present an approachable timeseries plot of the measured CH_4 fluxes due to high variation (large emissions vs. other data) – furthermore, we think there is no need to add any more figures to this paper.

RC: 333-334 and Fig 3: This is unclear to me. Is this a temporally static correlation between the mean of all CH_4 fluxes at each point and the mean of soil moisture at each point?

AC: Yes.

RC: 339: Is "September" supposed to be "October" here?

AC: Thank you for pointing this out. These were actually calculated for May–Sep, and was corrected for May–October as follows: "The mean measured CH_4 flux in May–October at the site was −4.88 µmol m^{-2} h^{-1} (median −6.43 µmol m^{-2} h^{-1}) in 2013 (n=339), and −6.46 µmol m^{-2} h^{-1} (median −5.90 µmol m^{-2} h^{-1}) in 2014 (n=373), however, the difference was not statistically significant."

Furthermore, we noticed that in line 290 we had reported mean air temperature in May–September for different years, and these were also changed to cover May–October, as well. Mean air temperature for May–Oct in 2010–2017 has been 10.0–13.2 C-degrees, in 2013 12.4 and in 2014 12.7 C-degrees.

RC: Fig 4: It would be nice to break these plots up by May-July and August-October observations.

AC: Thank you for this suggestion, we replaced Fig. 4 with a new one with the two seasons separately, and modified the text accordingly. We also added information about statistically significant differences between the seasons among each vegetation group (plus signs). Moreover, we noticed that there was a mistake in the figure caption of Fig. 4: the triangles are actually medians, whiskers are 25th and 75th percentiles, and asterisks are means – these were corrected in the caption.

RC: 361-367: Again, reporting the only the mean is limiting, also report SD (or some other metric of variability) within these sample groups.

AC: We added SD values.

RC: Fig 5: Was this variability maintained between the early to late summer transition? It would be good to show both groups on this plot, but might make things too cluttered.

AC: We explored the possibility to plot the seasons separately in two different plots (see below). However, it does not provide much new information compared to the modelled flux maps – mainly the wettest sample points shift from emission to uptake of CH4. Also, the figure splitted into two seasons is not very easy to read, as the Referee expected, and moreover the purpose of this figure is to show the spatial variation at sample points. Hence, we think it is best to keep this figure as it is.

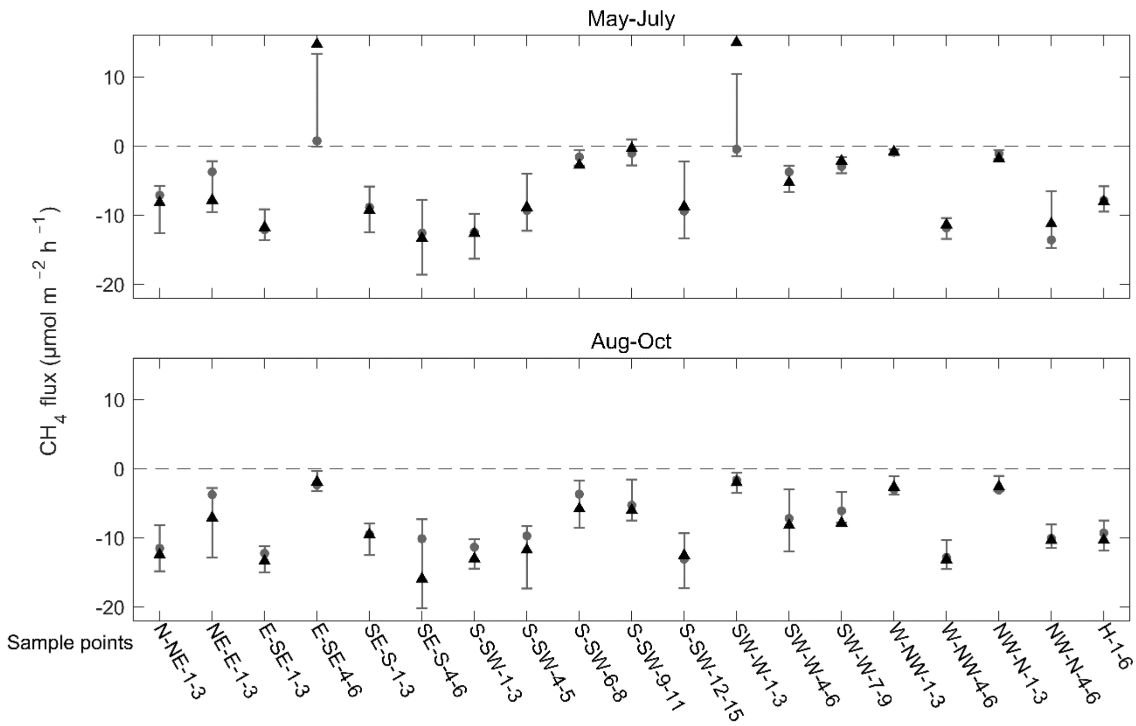

Fig. 5 of the manuscript splitted into two seasons.

RC: Table 2. It would be very helpful here to report the modeled statistics for both the whole area and at the sample points. Currently it is unclear whether the modeled soil moisture is systematically lower and therefore causing systematic overestimations of CH4 uptake, or if the domain of the entire study area happens to be drier on average leading to a larger estimated CH4 uptake.

AC: We added the modelled values at the sample point locations to Table 2.

RC: Fig 7: Normalizing the uncertainty at each pixel by its predicted value would help communicate the spatial patterns in the consistency of the RF ensemble output. I would also suggest that the authors add a note on interpreting this uncertainty, which is more of a measurement of the agreement of predictions among multiple RF iterations than the error between predictions and observations like RMSE. I am a major supporter of reporting ensemble uncertainty along with model metrics like RMSE, but the wording can get very confusing!

AC: Thanks for this comment. We plotted the relative uncertainties instead of absolute uncertainties as the Referee suggests, however it is good to note that this will then alter the message conveyed with the figure. We also modified the text so that it is clear that this figure shows the uncertainty related to the upscaling procedure alone, and does not include uncertainties related e.g. to possibly biased training data due to biased sampling locations. For the overall uncertainty, cross-validation metrics such as RMSE are better, like the Referee points out.

RC: 418-421: This is another place where normalizing the uncertainty of the ensemble predictions is useful.

AC: Figure 9 was modified as the referee suggests, as well as Fig. 7, however note that for locations with close to zero fluxes the relative uncertainties will inflate. This is one of the reasons why we opted to plot absolute uncertainties in the first place.

RC: 422-424: I may have missed it, but I do not remember seeing this approach described in the methods and it is kind of unclear here. I am also confused by what this is supposed to demonstrate.

AC: We acknowledge that the wording on lines 422–425 in the original manuscript needed clarification, and this has maybe caused misunderstanding. The idea of the analysis shown in Fig. 10 in the manuscript was to evaluate how many chamber measurement locations were needed to get an accurate estimate of landscape-level flux by only averaging over the measured chamber data without any upscaling with RF. Here the mean upscaled CH4 flux was used as a reference since it accounts for the soil heterogeneity (see above). It was shown that average over 15–20 locations resulted in a similar bias as average over all the chamber measurement locations. This should be useful information for future chamber measurements in similar locations. We revised the text so that this analysis is clearer and describe the methodology briefly also in Materials and Methods.

Discussion:

RC: 444: This is unclear. The RF model was just used to estimate spatial distributions of soil moisture, which were then used to predict CH4 flux based on a linear model, correct?

AC: Due to the Referee #1 comment we re-evaluated the modelling approach, and opted to follow reviewer suggestion to use RF model also for CH4 fluxes.

RC: 451-454: I am not sure what these lines are doing in this paragraph. They seem disconnected from the point.

AC: We assume that the discussion on the relationship between vegetation and soil moisture suits here, although, we have now completely rearranged the Discussion section.

RC: 457: How are these two species different in terms of phenology, growth form, and root structure? If they are similar, I would hesitate to infer that the vegetation is affecting CH4 flux rather than soil properties other than moisture.

AC: This sentence is not based on just two species, but the vegetation classes. This proposition is very moderate suggestion that there might be other effects than soil moisture, something related to vegetation may affect the CH4 flux.

RC: 469-470: Yes, but variability within point clusters was not communicated to the readers. It would be very useful to include.

AC: It is true that there is variation in the CH4 fluxes within sample point groups (Fig. 5). We add a sentence about that in the Discussion (lines 492–494).

RC: 470-472: I do not agree with this. The points created the domain of the training data, so we would expect the model output to be constrained by that domain. Additionally, the mean of the data only tells part of the story. It would be much more useful to compare the distributions of prediction values vs. observations.

AC: We partly agree with this comment. RF model cannot predict outside the range of values in the training data and the same applies to our CH4 upscaling procedure. However, we argue that upscaling can fix biases caused by skewed sampling locations (as long as the training data contains at least some data points at both extremes of the pdf) and hence agreement between the mean values is not trivial. We agree also that mean of the data tells only part of the story, however analysis of mean flux is important if accurate landscape-scale CH4 budgets are strived for. Hence, we would like to keep this part of the manuscript as it is, but included a sentence about the fact that mean tells only part of the story, but comparison of means is important if the target is accurate landscape-scale CH4 budget (line 572).

RC: 494-496: This is interesting, what differences in the ecosystems/soil types may account for this?

AC: Unfortunately, we cannot really find a good explanation to this based on the collected data, but think it is relevant to mention.

RC: 503-509: It would be useful to communicate whole plot scale CH4 flux estimates, but net sums and total source and sink strength.

AC: If we understand correctly, the Referee suggests to add discussion about the whole-site mean values presented in Table 2. We tried to emphasize the main message in the discussion.

RC: 513-514: This is could also be due to reduced activity of methanogens in deeper soil layers/microsites.

AC: Yes, due to lower soil moisture. At least for the wet areas. We added this to the same sentence.

RC: 568-571: Yes. But why did this study focus on modeling soil moisture and not directly modeling CH4 flux based on landscape features?

AC: See our answers to the main questions raised.

RC: Conclusions: This section could be filled out more completely. Differences in CH4 flux based on vegetation type was an interesting finding, for example.

AC: Thank you for this comment. We wanted to keep the Conclusion paragraph short.

References:

Treat, C. C., Bloom, A. A. and Marushchak, M. E.: Nongrowing season methane emissions–a significant component of annual emissions across northern ecosystems, Glob. Chang. Biol., 24(8), 3331–3343, doi:10.1111/gcb.14137, 2018.