Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2018-155-AC1, 2018

Interactive comment on "Interpreting eddy covariance data from heterogeneous Siberian tundra: land cover-specific methane fluxes and spatial representativeness" *by* Juha-Pekka Tuovinen et al.

Juha-Pekka Tuovinen et al.

juha-pekka.tuovinen@fmi.fi

Received and published: 20 July 2018

We thank the referee for the extensive criticism. We respond to each comment in the following.

First, we should mention that our original plan was to submit another, more data-oriented manuscript at the same time. This paper would present longer-term data and its processing in more detail and constitute a more conventional study of flux variations and their environmental controls. Unfortunately, the preparation of that manuscript was

delayed and thus a concurrent submission was not possible. As we did not find it appropriate to refer to a "manuscript in preparation", we removed such citations at a final stage. Thus it is possible that some essential information is missing and we need to complement the manuscript with material that supports our arguments and presentation.

MAJOR COMMENTS

Comment:

1.) constant flux rates: I was extremely surprised when read- ing the results section to find that temporal variability in flux rates is ignored, instead a constant flux rate is assigned to each of the chosen LCC groups. This constitutes an over-simplification of the tundra ecosystem in the first place, but also limits the presented approach to extremely basic overall findings. Considering that fluxes over a period of 8 weeks from the late growing season of an Arctic ecosystem are used, besides short and mid-term climate variability the fluxes will be influenced by slowly varying conditions such as e.g. thaw depth or soil temperatures. Accordingly, it cannot be assumed that even mean flux levels for moving windows remain constant over such a long period. Even more important in this context, I would strongly assume that the different LCC groups will show a different trend in phenology in this part of the sea- son, i.e. some may be subject to earlier senescence, and also some of them may react more strongly to environmental stress such as water limitation, or first nights with freez- ing temperatures. This implies that the differences/ratios between flux rates from LCC groups will not be constant over time. So what is the value of providing just a single mean flux rate per LCC group? Would the differences still be significant if short-term temporal variability, and shifts in flux rate ratios, would be taken into account? Regard- ing the applicability of the results, under the given circumstances the output of the flux decomposition is of no value for any other purpose than investigating a potential sensor location bias (and even here the impact is limited). To reach a broader audience, tem- porally vary- ing fluxes per biome with functional links to environmental controls would have to be provided. Summarising this item, it is obvious that the chosen approach is based on a strong simplification of the actual flux patterns. This should be discussed in detail in a revised version of the manuscript. This discussion needs to be supple- mented by a demonstration how variable measured net flux rates are over time, broken up into the chosen LCC groups. If it cannot be proven that the ratios between flux rates from different groups remain largely constant over time, the approach cannot be ap- plied as is. In general, I strongly urge the authors to consider extending their approach, so temporal variability in flux rates can be considered.

Reply:

While it is plausible to assume that there should be significant temporal flux variability driven by environmental controls, such as soil temperature, there is obviously no guarantee that these would constitute the main factor controlling the variations in the observed fluxes. At Tiksi, with a heterogeneous landscape, this variation during the study period is overshadowed by variations in the flux footprint, i.e. the varying contribution of different land cover types. The following arguments support this conclusion: (1) There is no significant correlation between the measured CH4 flux and either soil temperature or air temperature. There is no such correlation either if the data are limited to land cover class (LCC) group-specific cases in which a LCC group dominates. (2) Figure 5 shows that the observed CH4 flux depends strongly on wind direction. This dependency reflects the variability of flux footprint rather than environmental factors. The variance of measured 30-min fluxes is 0.049 microg2/m4/s2 . Calculated from the binned mean fluxes (50 classes in Fig. 5), the variance related to the directional pattern is 0.040 microg2/m4/s2, i.e. 82% of the total variance. (3) The regression model presented shows that LCC proportions explain 80% of the flux variation (p.15, l.32). We conclude that these results indicate that the short-term variability in the observed fluxes is predominantly due to footprint variability and any environmental control can have a secondary effect only.

This does not rule out the role of phenology and longer-term trends in soil tempera-

ture, for example, in affecting the longer-term variations of CH4 fluxes. In the revised manuscript, we will analyse the data also on a weekly basis. This analysis will indicate that the differences between the LCC group-specific fluxes persist systematically but also that the fluxes indeed vary even within the rather limited study period; in the revised manuscript we will show how this relates to the trends in soil temperature and LAI. The weekly mean fluxes will be used for upscaling, too, and we can show that our conclusion about the spatial representativeness of the EC measurements is robust. The temporally resolved analysis also provides material for an enhanced discussion of the results, for example indicating that no statistically significant estimates can be obtained when the data set becomes too limited.

Comment:

2.) uncertainty estimates: The manuscript misses to even discuss some essential sources of uncertainty that influence the given approach, and those few aspects that are treated (e.g. uncertainties in maps) are only covered qualitatively. Even very easy components, such as e.g. assigning an uncertainty to the input flux rates from the EC system, which is then projected onto the modelled LCC flux rates, is missing. Most importantly, there is no uncertainty estimate for the footprint approach. It is obvious that any source weight function can only be an approximation of the actual field of view of the sensor, as many footprint validation studies have shown in the past. In this study, however, footprint simulations are treated as a given fact. There are uncertainties in all the input parameters used to feed the footprint model, there are uncertainties associated with parameterizations/assumptions inherent to the footprint model, and there are uncertainties related to the methodology (e.g. horizontal homogeneity, stationary flow, and so on). For a modified version of this study, the authors need to provide a convincing concept to constrain the uncertainties in computed source weight functions, and how these influence the results obtained by the flux decomposition approach. In addition, the uncertainty concept should, as mentioned above, also involve the flux data uncertainty, and also the uncertainties inherent to the maps used in this study should be involved, and quantified.

Reply:

It is incorrect to state that we do not address or quantify uncertainties. The LCC group-specific flux estimates are presented with the 95% confidence intervals (Table 5), which are based on a heteroskedasticity and autocorrelation consistent estimator (Sect. 2.3.1). These error estimates are used for upscaling the fluxes, the results of which are also shown with quantitative uncertainty estimates (Table 6 and Figure 6). We also report the LCC classification accuracies (Sect. 2.1.3) and discuss the related uncertainty (Sect. 3.2, end). As explained in Sect. 2.3.1, we assume that the confidence intervals represent the integrated effect of different error sources, including the measurement data, LCC classification and footprint modelling. We believe that this is a more appropriate method than a bottom-up approach, in which individual error estimates should be first allocated to each error source assumed and then propagated to estimate the total uncertainty. Whereas it is obvious that footprint modelling has uncertainties, both structural and input related, it is not obvious at all how this could be quantified for a meaningful error estimate. However, we agree that the uncertainties related to footprints should be made more explicit. We will add discussion that explains that the footprint dimensions and representativeness metrics presented in the manuscript are affected by model uncertainties, including citations to footprint validation studies.

Comment:

3.) Validation of flux rates: In Section 3.2, the authors include a good paragraph (p.16 ll.20ff) that supports the negative flux rates found for bare soil. As part of the line of argumentation, chamber measurements from a previous study are cited. I think it's fair to assume that this study did not only measure those 32 data points cited here for bare soil, but also other components within the Tiksi landscape. Why are those not used? Having flux chamber results for the different LCC groups would be the best way to

validate that the flux composition actually produced realistic results. Also, the Tiksi flux tower has been running for several years now - why restrict this study to just 8 weeks? Why not use more data, so the database is more representative, and can also resolve temporal variability? Why not split the dataset into training and validation sets, so any finding can actually be evaluated?

Reply:

The reviewer is right that the chamber measurements cited in the manuscript were not limited to bare soil. We also agree that such data would be useful for validating the present results. The reason for not using the chamber data more extensively is that the experimental design was incomplete: the number of chamber plots was modest, and the reach from the EC mast was limited due to the use of an online gas analyser. Moreover, the chamber plots did not fully correspond to the land cover classification that was subsequently developed and used in the present study. Thus it is not possible to use these data for a proper validation of the flux decomposition. The bare soil data were introduced to provide support for the surprisingly high uptakes rates observed. However, we can report here that the overall pattern the chamber data depicts is consistent: wet fens appear as strong CH4 emitters (two plots, 32 observations at each plot, means of 0.56 and 3.8 microg/m2/s) and dry fens as moderate emitters (four plots, 31/32 observations, mean 0.25 microg/m2/s).

It is true that the Tiksi tower has been running for several years now. As indicated above, a paper analysing longer-term data is in progress, while the aim of the present manuscript is different and achievable with data from a more limited period representing a well-defined season.

As for training/validation, we did validate the results by splitting the dataset into training and validation sets. The approach is described in Sect 2.3.1 (end). The validation statistics show a good performance against independent data (p.16, l.6-8).

Comment:

4.) scope of this study: With the limitations of the chosen concept (constant fluxes) as mentioned above, the authors should clearly restrict the scope of the study to an estimate to constrain sensor location bias. I do not see any other application of their method besides this (I would be glad to get convinced otherwise, e.g. by a thorough discussion ..). I do not think they can claim to provide land cover specific CH4 flux rates, since they present one set of mean flux rates for a single period of time, nothing more. They also do not interpret EC data, since obviously there's no temporal variability, no links to environmental controls, no interpretation why certain LCC groups show different fluxes than others. What is being provided here is an extremely simplified approach to estimate flux rates per LCC group, and check if, given these flux rates, the net fluxes represent the emissions from a larger area (aka sensor location bias). Since there is also no discussion which aspects influence the performance of this approach (e.g. length scale of variability in terrain features, differences in flux rates between LCC groups, footprint variability, etc), there is no way of telling if this approach could be applied at other sites as well.

Reply:

Unfortunately, we fail to see the logic of the comment that we cannot "claim to provide land cover specific CH4 flux rates, since [we] present one set of mean flux rates for a single period of time, nothing more." It is true that we presented average fluxes, but any averaging is irrelevant to the question whether we provided (statistically significant estimates of) land cover specific fluxes. We obviously did. However, in the revised version we will address temporal variation by analysing the data in shorter periods.

Similarly, the comment "[t]hey also do not interpret EC data, since obviously there's no temporal variability, no links to environmental controls, no interpretation why certain LCC groups show different fluxes than others" seems odd. The meaning of "interpretation" in the context of this manuscript should be clear already from the latter part of the title and the objectives listed in the introduction, the latter of which the reviewer commends later in the review: "the 3 different sets of objectives are clearly formu-

lated". These objectives do not include temporal variability or environmental controls; in contrast, they explicitly refer to mean fluxes (p.4, l.25). Please note also that in conclusions (p.19, l.9-10) we state that the environmental controls and long-term data will be studied in a follow-up paper. We realise this is possible and even necessary for the present manuscript, and the temporally resolved analysis that will be included will address these aspects.

As for the last alleged omission, i.e. no discussion about the different fluxes among the LCC groups, this is formulated, on the basis of a literature survey, as a statistical hypothesis. This is explained in Sect. 2.3.2. The success of the regression model, in terms of both statistical significance and the logic of results, confirms this hypothesis and consequently provides credence to the assumptions behind the flux differences. This is discussed in Sect. 3.2, in which we report previously measured fluxes for different tundra surfaces. We will complement this discussion by outlining mechanisms that are known to explain the differences.

Comment:

5.) a thorough discussion is simply missing! What is the implication of the findings? How could the presented approach be used? Where are the weaknesses, which factors limit the interpretation of the findings?

Reply:

While some implications of the study are presented in the conclusions ("An important implication emerging from our results…", p.18,l.26), we agree that the discussion related to the applicability of the approach presented is too limited. We will add discussion about the feasibility of the statistical model, drawing upon the new analysis of temporal variations to be included.

GENERAL COMMENTS

Comment:

1.) The introduction is well structured overall, and the 3 differ- ent sets of objectives are clearly formulated. The paragraph on methane (starting p.2 l.25) is rather confusing, though, and should be revised.

Reply:

We will revise this paragraph.

Comment:

2.) Section 2.2 needs a complete overhaul. Many sections, e.g. most parts of 2.2.1, are textbook knowledge, and do not need to be shown in detail herein. Section 2.2.2 is much too detailed for what actually needs to be described. You project a source weight function on gridded maps, and accumulate the weights of individual cells, sorted by categories, nothing more. Overall, this whole section is much too long. I suggest to revise it to the following structure: - 2.2.3 should be moved to the front - 2.2.2 should be shortened, and simplified, coming next - 2.2.1 should be discarded entirely - 2.2.4 should be moved as part of the results section

Reply:

Our idea was to show that the source area can be defined in different ways, and our results show that it is important to present an exact definition when reporting footprint dimensions or referring to the study area. There are numerous papers in which the footprint concept is used very loosely, even in a misleading way, but we refrained from specifying them in the text. However, to improve the focus of the manuscript, we will remove Sect. 2.2.1 and include a shortened version as an appendix.

We do not understand why Sect. 2.2.2, which only covers less than 1.5 manuscript pages, would be "much too detailed". It provides the mathematical definition of the variables we use in our analysis – "nothing more", to cite the reviewer, and we do not claim otherwise in the manuscript. We feel that an exact description of the methods should be encouraged rather than discouraged. Mathematical formulation facilitates

such exactness (and someone so inclined may even find it useful to see the EC-related averaging process formalised as in Eq. 8, for example).

Comment:

3.) In Section 2.3, the ordering of the information should be revised. Many pieces of information given in 2.3.2 were needed to interpret the text in 2.3.1, for example

Reply:

We will check the ordering of information in this section and revise the text accordingly.

Comment:

4.) Results Section 3.1: The first part on general footprint characteristics should be removed (P1, P2). After all, what you basically state here is the obvious fact that footprint areas grow with stable stratification. The authors may move Table 3 to the appendix, and refer to it in the main text in case any reader wants to see the details, but this is clearly not part of the main story. The center part, highlighting the heterogeneity of surface characteristics within the footprint, reads well (P3-P5). The last part (P6+) should be revised - it is informative to describe a sensor location bias using the different surface characteristics, but the current format is confusing, using too many versions of a reference area (also Table 4 should be reduced).

Reply:

While we agree that it is well-known that the footprint area increases with atmospheric stability, this is not presented as a conclusion of the present study; we state that the results show "expected qualitative features" (p.13, l.11). Rather, we present quantitative dimensions of the flux footprint for this site, which we believe constitute information that is essential for further studies that use the EC data from this site. Furthermore, we compare different source area definitions, showing substantial differences, and conclude that it is important to report an exact definition (see above). However, we agree that this may appear a side track and will move Table 3 to an appendix and modify the first paragraphs of Sect. 3.1. We will also reduce the number of reference areas and simplify Table 4 and the related text in Sect 3.1. accordingly.

Comment:

4.) p.16 l.10ff: I don't think it makes much sense to compare the Tiksi flux rates against values from other sites without also comparing environmental conditions, and the measurement approaches

Reply:

We think it is a common procedure to compare new data to previous results, even if there may be differences in site characteristics and environmental conditions. We compare the LCC group-specific fluxes with an extensive set of chamber-based measurements (cf. Major comment 3) during the growing season (as detailed on p.16, l.10-13), indicating that our LCC group-specific flux estimates are reasonable. We also present corresponding EC results from comparable sites and observe that the mean CH4 flux at Tiksi is within the variation in the mean summer flux among these sites.

Comment:

5.) results section 3.3: It is confusing, and actually not understandable, why so many different reference areas have been used to compare the footprint LCC composition to. This actually leaves the impression that the authors were searching for a nice configuration that can demonstrate that the EC measurements are actually well representative (e.g. p.17 l.27 ' the sensor location bias could be minimized by reducing the radius to 800– 1000 m'). What is the value in such an exercise? People who are interested in using EC data want to know how well they represent a LARGE area.

Reply:

In practice, an EC study site is defined as a more or less arbitrary area surrounding the EC mast. In our case, we selected this area according to a seemingly objective criterion based on a footprint dimension (95% coverage in neutral conditions), but obviously we

could have chosen a different target area. In any case, the spatial representativeness of the EC data collected depends on the corresponding footprint climatology. Our analysis shows that our CH4 flux measurements are representative of the original reference area, given the uncertainty range of the upscaled flux. However, the mean sensor bias would be smaller for a smaller target area, which would be as logical a choice as the original one. It is important to know how well the EC data represent a large area, which we also assess, but first it is important to understand what is the surface configuration that you actually are measuring. We will reduce the number of different reference areas and try to clarify this discussion in the revised version.

TECHNICAL COMMENTS

Comment:

p.3 l. 16: I don't see a connection between spatial hetero- geneity and the need for long-term measurements ...??

Reply:

The idea was to imply that for representative sampling more data are needed for heterogeneous than homogeneous surfaces. We will remove 'long-term'.

Comment:

Section 2.1.1: A bit too brief. Soil types could be mentioned, and it should be mentioned that vegetation is given in a different subsection. Section 2.1.2: The outline of the QC is too short. What exactly was done regarding instationarity, for example? How were unphysical outliers defined? And how were the gaps treated in the end?

Reply:

We agree that these sections are too brief. This relates to the delayed paper mentioned above, which would have provided more details. We will add more information about the soil types and data processing.

Comment:

Section 2.1.3: I suggest moving the definition of PCTs into a table. It should be mentioned that the dominant vegetation, and other characteristics, are described later in 2.1.4

Reply:

This may introduce some repetition, but we will prepare a table summarising the LCC properties.

Comment:

p.6 l.16: The authors should decide if they want to use LCC or PCT as a term for this classification. Using both is very confusing!

Reply:

Plant functional types (PFTs) and land cover classes (LCCs) do not refer to the same thing. The vegetation was surveyed as PFTs, and each of the LCCs may contain several PFTs. The confusion probably stems from the inexact formulation on p.5. (l.26-), which aims to list the LCCs rather than the PFTs. We will rephrase this sentence.

Comment:

Fig.4: the lower 3 panels are not necessary , since they show the same patterns as above, only normalized against the black dashed line

Reply:

The reviewer is right: the patterns in the lower panels are the same as in the upper ones. We will remove the lower panels and add new right axes for bias (%) to the upper ones.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2018-155, 2018.