

The authors present two years of CH<sub>4</sub> flux data from a drained peatland forest site which have been collected with an automated chamber system consisting of six chambers connected to a high resolution gas analyser. The aim of the study is twofold. First, the flux data series is used to test whether CH<sub>4</sub> fluxes from these chamber measurements are better analysed with linear or non-linear regression. As conclusion, the authors recommend to first calculate all fluxes by linear regression and then to recalculate high fluxes with an exponential regression. High CH<sub>4</sub> fluxes were defined by a site-specific threshold. Second, the study analyses the variation in CH<sub>4</sub> flux rates from the forest floor at various scales (diurnal, seasonal, inter-variation) and annual balances are presented as well.

The manuscript is very well written and this study provides an important flux dataset. High resolution gas analysers for CH<sub>4</sub> measurements are a recent development and the number of studies combining automated chambers and these analysers for long-term measurements are still scarce. This study has the potential to result in an excellent paper providing new insights into CH<sub>4</sub> flux dynamics and the methodological challenges associated with gathering these data. However in my opinion, the manuscript has two major flaws. On the one hand, the authors do not fully explore the potential of the dataset from a methodological standpoint and should expand this part of the manuscript more. On the other hand, they provide a lengthy description of the flux differences between the single chambers, but the experimental design does not really allow a proper discussion of the fluxes from an ecological standpoint. Thus, I recommend major revision and will detail my concerns below.

#### General comments

- I find the threshold of 3.5 µg CH<sub>4</sub> m<sup>-2</sup> h<sup>-1</sup> quite arbitrary. Based on the data presented here, I am not convinced to use such a threshold as decision for which regression to use. Why should this method be more appropriate than using a statistical criterion like e.g. AIC and to decide based on that criterion for each flux measurement separately which regression method to use?
- I am missing more details (including figures) about the effect of different closure times on the flux calculation results.
- You have only one replicate per vegetation type. Based on this setup, it is not really possible trying to understand the differences between the chambers from an ecological standpoint. Furthermore, as additional data besides vegetation composition, you only seem to have soil temperature for the single plots. Why wasn't the water table measured at each chamber? Do you have any knowledge about the soil profiles at the different locations? How has the porosity of the soil changed due to the drainage? Are your six locations really representative for the chosen vegetation types and the soil conditions at the site?
- Despite the number of replicates, the dataset is very suitable to study diurnal variations in CH<sub>4</sub> fluxes. This should be a separate section in the discussion and be more focused on the underlying processes causing these variations. Right now, this part has good references, but is mainly descriptive. In general, the manuscript has a good reference list, but often you only write which correlations were found in other studies. You need to go a step further and discuss more the processes involved. Process descriptions often stay too vague and general.

#### Specific comments

- Page 1, line 17: CH<sub>4</sub>
- Page 2, line 14: „...thus turning in particular well-drained peatlands...”
- Page 2, line 22: Methane oxidation rates are also strongly controlled by the methane concentration in the soil, not only the oxygen concentration.
- Page 3, line 19: Add that exponential regression is especially sensitive to disturbances at the beginning of the measurement.

- Page 3, lines 19-20: It is not generally true that you need more than five data points to fit an exponential regression. It depends on the flux strength. See for example the paper by Pedersen et al. (2010) and Forbrich et al. (2010) which you cite in your manuscript. Thus, it is not uncommon to perform non-linear regression on datasets derived from syringe sampling. A high resolution gas analyser is not typically required. The great advantage of the high resolution gas analyser is that it reduces the uncertainty of the estimated slope of the flux curve, it does not necessarily change the mean estimate.
- Page 3, line 23: Specify ‘high temporal resolution’. You probably mean both the sampling rate during one chamber measurement, and the total number of chamber measurements you can perform per day.
- Page 3, line 31: Do you know how much fertilizer was applied?
- Page 4, line 7: For the first species in the brackets write the full name. The way you have written it now, “S.” stands for Sphagna and not Sphagnum.
- Page 4, line 20: Specify the type of fan used. What was the volume turnover inside the chamber?
- Page 4, line 31: What does “w = 1 cm” mean?
- Page 5, line 1: Is “Linak, 2009” a reference? If yes, it is not in your reference list. In general, be more consistent when mentioning product names. Include the company name as well as the associated city and country.
- Page 5, line: Wasn’t the flow rate quite low? What was the actual tubing length and tubing diameter for the chambers?
- Page 5, line 7: Specify the type of sensor you used for soil temperature measurements.
- Sections 2.3 & 2.4: Which software did you use for the flux calculations and data analysis?
- Page 5, line 13: “bihourly” = twice per hour or every two hours?
- Page 5, line 14: If I understand correctly, you did not discard any data points from the measurement start. Why was it in your case justified to not apply a deadband to the flux data? How can you be sure that you had proper headspace mixing immediately after chamber closure?
- Page 5, line 17 – page 6, line 1: I don’t quite understand this part. What exactly is the purpose of equation 3? Are these parameter estimates inserted into equation 4? Also, is the Kutzbach model applicable to CH<sub>4</sub> since it was developed for CO<sub>2</sub>?
- Page 7, line 1: insert “CO<sub>2</sub>” in front of “concentration”
- Page 7, line 10: Were these hardware problems of the gas analyser? If yes, it might be interesting information for other users.
- Page 8, line 2: insert “it” before “usually”.
- Page 8, lines 4-12: You base a lot of the following sections on these results. Provide example figures of single flux concentration curves so that the reader can judge for himself/herself.
- Page 8, line 9: How do you know that it is an underestimation?
- Page 8, line 16: Shouldn’t it be “<”?
- Page 8, lines 18-19: This sentence is mainly a repetition of the previous sentences. And is the data removal really the only reason for the observed differences?
- Page 8, lines 21-23: I am not convinced of this based on the presented data.
- Page 10, line 30: Friction velocity is an important parameter, but it has not been mentioned at all before this section. I assume, u\* is based on the eddy tower measurements?
- Page 12, line 5: I don’t see a reason for mentioning the CO<sub>2</sub> data here.
- Page 12, lines 10-11: This is an obvious observation when using relative differences.
- Page 13, line 1: Also discuss the importance of the water table depth. The low temperature does reduce metabolic activity, but methanogens are favoured by the increasing soil moisture content.

- Page 13, lines 14-23: This section is a very good example of the weakness in your study. You are lacking data on (potentially) important environmental variables and are just speculating here.
- Page 15, line 5: Could the lack of correlation be due to a lack of grass species (e.g. root exudates as food source for microorganisms) in comparison to the other study?
- Table 1: Do you have a reference for  $VGA_{max}$ ? Include more details about the sampling method and the sampling time.
- Table 3: Did you also perform correlation tests on the entire dataset without dividing it into seasons?
- Figure 1: At what height was air temperature measured? What were the standard errors of the average water table depths? How far were the chambers away from the WTL measurement points? Maybe provide a map of the experimental site setup as a supplement.
- Figure 2: Did you also check the relationship for each year separately?
- Figure 3: It looks a bit like  $CH_4$  uptake sometimes was even higher than  $-40 \mu g CH_4 m^2 h^{-1}$  and the fluxes just went off scale. Also, what was the uncertainty of the single fluxes on average?
- Figure 5a: This bar chart is quite meaningless without some indication of the uncertainty for each bar. Do you also have a cumulative error estimate for Figure 5b?
- Figure 6: Is the daily flux just upscaled from the average hourly flux or does it represent the cumulative hourly fluxes per day? Also, I find it really difficult to distinguish between the black and blue points. It would be nice to have these plots for the other five chambers as supplement.