Discussion of “Forest-floor greenhouse gas fluxes in a subalpine spruce forest: Continuous multi-year measurements, drivers, and budgets”

Author Response to Referee 1 comments

Krebs et al.

November 30, 2023

In the following, reviewer comments are given italics, author comments are given in normal font.

1. General comments

Luana Krebs and co-authors present an impressive soil GHG flux dataset from a subalpine forest. Multi-year datasets from such ecosystems are rather scarce and therefore definitely deserve publication. CO2, CH4 (and N2O) were measured in high temporal resolution year-round during 4 years and annual CO2eq budgets were calculated showing higher net emission during the warmest year 2022. I have some suggestions to improve the manuscript.

Thanks for the comment and your suggestions.

It should be considered removing N2O from the manuscript. Obviously N2O data is available only for 2 years and the extremely low fluxes are only shown in the supplements and not included in the budgets or any other calculations. It appears that the 180 sec chamber closure likely were not long enough for serious N2O flux calculations (extremely low 10th percentile R2 in Table.2) – as mentioned in the manuscript, using such a poor fit to calculate fluxes should be avoided. I don’t know the actual cause. It is unlikely that the subalpine soil does not show any, or not measurable N2O emissions throughout the whole year. It seems more likely that there appeared problems with the laser. I don’t operate one by myself but heard from colleagues that the real measurement accuracy in the field is not really the 0.06 ppb for N2O that is suggested. There might be drift or whatever else. How and how often was the laser calibrated? If you do not have 100% trust in the N2O data, I would take them completely out of the manuscript. Currently you state that there was no soil N2O efflux at this site – are you 100% sure?

The suggestion to remove the N2O fluxes from the manuscript seems to be based on five reasons:
1) data available only for 2 years,
2) very low N2O fluxes,
3) 180 sec closure for flux measurements and low R² values,
4) unlikely that alpine soils show such small fluxes, and
5) problems with the laser (accuracy, drift).

In the following we want to address those five points.
1) Long-term N2O fluxes in forests are very rare, due to the difficulties measuring year-round. Thus, two years of N2O flux data are actually extraordinary, and there is no reason to delete such rare fluxes.
2) We agree that the N₂O fluxes were low, and we did not use them for further driver analysis, one of the main objectives of the manuscript. This decision was made mainly because of the low magnitude of the fluxes and their irrelevance for the forest-floor GHG budget (using the mean N₂O flux measured with the automatic chambers over the two years, 0.63 nmol m⁻² h⁻¹, we arrive at an annual budget of 0.066 g CO₂-eq m⁻² yr⁻¹ which represents 0.003% of the annual forest-floor GHG budget). However, we think that the result of the irrelevance of N₂O fluxes for the annual forest-floor GHG budget of the forest does not diminish the importance of these fluxes.

3) In order to check for the validity of our chamber measurements, we have performed measurements using static chambers with the dimensions of d = 30 cm and h = 30 cm (Hutchinson and Mosier, 1981). We used eight static chambers, i.e., four chambers next to the automatic chambers, and four chambers placed randomly within the research area. Soil collars were installed two weeks prior to the first measurement campaign. Four rounds of sampling were done on two measurement days in October 2023 (n=32), when soil temperatures were between 5.5-10 °C, well above the long-term mean, and soil moisture values above 8%, favoring microbial activities. Three collars were irrigated between the first and second sampling round on the two days to simulate a heavy rainfall event, favoring denitrification. We left the chambers closed for 1 h and sampled the air in the headspace every 20 min. The fluxes that we measured with the eight static chambers were low (mean±SD = 2.9±31.1 nmol m⁻² h⁻¹). Furthermore, they agreed very well with the fluxes which we have measured using the automatic chambers over two years (mean±SD = 0.63±58.6 nmol m⁻² h⁻¹), and in October (mean±SD = 10.2±14.7 nmol m⁻² h⁻¹). We are thus confident that the low fluxes measured using the automatic chambers are real and that an insufficient closure time is not the reason for the low fluxes.

The low R² values can be due to different reasons, high variability during closure time or small fluxes (see point 4). We would like to point out that the static chamber fluxes show rather low R² values too. However, we found a clear positive relationship between the R² value and the magnitude of the flux for both measurement techniques. Therefore, we suggest looking at a different criterion to assess the N₂O flux quality, such as the mean square error (RMSE) which is not dependent on the flux magnitude itself.

4) At the site, N supply to plants and microorganisms is limited, so it is to be expected that N₂O fluxes are low as well. Foliage N concentrations indicate N limitation for spruce (foliar N concentration are about 1% in 0- and 1-yr-old needles as opposed to the optimum range of N content in needles between 1.5 and 2.3 %; Thimonier et al., 2010; Ingestad, 1959), N concentrations in the soil are low (1.4% in the organic layer and 0.4% in 10-20 cm depth, Jörg, 2008), N deposition is low (about 10 kg N ha⁻¹ yr⁻¹; Thimonier et al., 2019; Gharun et al., 2021). Thus, our site is rather low in N which could be used for microbial transformations, competing with plant uptake (Schulze et al. 2019).

Moreover, we measured N₂O fluxes using eddy covariance above the forest in 2016 and 2017 where we found that the forest emitted 0.047 g N₂O m⁻² yr⁻¹ which corresponds to 0.122 µmol N₂O m⁻² h⁻¹. These low fluxes further back up the notion that the soil N₂O fluxes are very low at the site. Furthermore, a study conducted a boreal spruce forest with low nitrogen deposition rates (about 5 kg N ha⁻¹ yr⁻¹), reported very low mean N₂O fluxes of around 0.02 µmol N₂O m⁻² h⁻¹ agreeing well with our results (Rütting et al., 2021).

5) The flux detection limit has been assessed based on the precision of the laser and the closure time. Our flux detection limit is a lower bound, i.e., under ideal conditions, we could detect fluxes as low as 29 nmol N₂O m⁻² h⁻¹. The fluxes within the range of ±29 nmol N₂O m⁻² h⁻¹ are therefore not significantly different from zero. It is correct though, that the real measurement precision is likely to be higher than 0.03 ppb. This basically just means that our flux detection limit would be higher. A study by Nemitz et al. (2018) reports the precision of the instrument as 0.09 ppb. This would give a flux detection limit of 87.2 nmol N₂O m⁻² h⁻¹, which is still a very low flux. The flux detection limit could be reduced with longer chamber closure time. According to this logic, we can say that the higher the measured flux, the
more confidence we have in the measurements. Overall, we can say that we are confident that our \( \text{N}_2\text{O} \) measurements are very low and not relevant for the GHG budget of the forest floor.

Based on the explanations given above, we would like to keep and discuss the \( \text{N}_2\text{O} \) fluxes in the manuscript. Long-term and high-resolution \( \text{N}_2\text{O} \) chamber measurements in forest ecosystems are very rare which is why we think showing this dataset is very important for the scientific discourse. We would like to redo the quality assessment of the \( \text{N}_2\text{O} \) fluxes using the RMSE instead of the \( R^2 \) which is dependent on the slope of the linear regression. Generally, we suggest moving the \( \text{N}_2\text{O} \) figure (Fig. A.1) from the appendix to the main text, add a panel (d) showing the boxplot from the static chamber measurements (Fig. 1), and discuss these measurements more in detail later on.

![Forest-floor \( \text{N}_2\text{O} \) fluxes (nmol m\(^{-2}\) h\(^{-1}\)) for the years 2017 (a) and 2020 (b). Black lines show means over four chambers, grey bands show standard deviations among four chambers. Boxplot showing distribution of means over four automatic chambers (c) and \( \text{N}_2\text{O} \) fluxes from static chamber measurements (d). The dotted lines depict the minimum flux which could be detected by the Dual Quantum Cascade Laser spectrometer.](image)

Some information should be added to the method section about how the snow in the chamber was treated in flux calculations. The volume of the water(ice) of the snow cover must be subtracted from the chamber volume during snow cover. How was that done? Was snow porosity measured or assumed somehow? If the snow volume was not subtracted, it is no wonder that CO2 emissions became the lower the more snow was in the chamber. Or was only the volume above the snow surface used for calculation? (In this case it would be flux from the snow surface, not the forest floor). Just of interest, what had happened in spring? Typically opaque chambers warm up faster and snow melts much earlier in and around them.

Thanks for the comment. Indeed, we did not provide enough details on how we treated the presence of snow in the chamber during flux calculations. As the reviewer rightly guessed, we subtracted the snow volume from the chamber volume during snow covered periods. We also accounted for the varying chamber volume due to the chamber frames we had installed during winter. Thus, the original flux calculations were correct. We will add the following formula to the manuscript, which describes the volume calculation for the flux calculations:

\[
V = 0.75 \times 0.75 \times (0.5 + \text{frame number} \times 0.5 - \text{snow depth})
\]
Where the chamber area was 0.75 m * 0.75 m. The height of the chamber was 0.5 m. However, the height of the volume depended on the number of additional frames added to allow flux measurements during snow covered periods as well as the snow depth inside the chamber. The additional frames were made of the same white PVC as the chambers themselves with a height of 0.5 m. Up to two additional frames were added.

With our chamber design we are confident to have avoided potential side effects on the environmental conditions as much as possible. The chambers were white (high albedo), very large (reducing edge effects), and in the open position, they moved far away from the soil collar (avoiding shading). We do not have data on the timing of snow melt inside of the chambers. However, we observed that snow melted rather later inside the chamber compared to outside. To assess potential chamber effects on environmental conditions, we measured soil temperature inside and outside of each of the four chambers. We did not find any significant differences inside vs. outside the chambers (Fig. 2a). In winter (Dec-Mar), the temperatures inside the chambers (orange) were only 0.1-0.5 °C lower than outside the chambers (blue, depending on the chamber). However, the differences were only significantly different from zero in December and February-March (Fig. 2b). In these months, soil temperatures are around 0 °C. When looking at the temperature response curve of forest-floor respiration (Fig. 2c), we can say that at such low soil temperatures, a difference of 0.5 °C in soil temperature has a minimal effect on the magnitude of forest-floor respiration. Therefore, we think that the effect of our chambers on the forest-floor respiration is negligible. We will add this info in the appendix of the revised manuscript.

![Fig. 2](image)

**Fig. 2:** a) Soil temperatures (Tsoil) at 5 cm inside (orange) and outside (lightblue) the chambers over the course of a year. b) Difference in Tsoil at 5 cm between inside and outside the chambers over the course of a year. Lines show means, bands show standard deviations over all chambers and three years (2017, 2020 and 2021). c) Q10 model showing the relationship of daily means of Tsoil at 5 cm and forest-floor respiration of all chambers and years.

It is often stated that high temporal resolution measurements have the advantage to capture hot-moments in GHG fluxes. Well – have such hot moments been observed? Currently it does not really seem so. Did freeze-thaw periods occur? What happened during these periods with CH4 fluxes? It would be important to zoom out some such hot moments from the long-term datasets and show them separately. Even if there was freeze-thaw and no peak in CH4 flux occurred – this could be shown. Were there any CH4 emissions at all, at least short-term? We for instance once observed a small CH4 and huge N2O peak during freeze-thaw in a deciduous forest (Schindlbacher Biogeochemistry 2022)
but very similar CH4 pattern with just lower uptake during winter as in your study (Heinzle AgrForMet 2023) in a spruce mountain forest. Another possibility for hot moments is after rain post drought periods. Did you observe any flux peaks of any GHG during such periods? Seems there were some very short term peaks and some longer-term positive CH4 fluxes in the “observed fluxes” in Fig. A.3.

Thanks for the suggestion to look deeper into hot moments. We actually have already described bursts of CH4 emissions coinciding with snow fall and snow melt in section 3.1 of the original manuscript. Furthermore, we showed daily flux data in Fig. 1 of the originally submitted manuscript where such emission periods can be observed. Since the manuscript is already quite dense, we had planned to write a separate paper concerning short-term variations in CH4 as well as CO2 fluxes, using the full resolution of the data (3-hourly) and showing the data per chamber separately. In the next manuscript, we would like to not only investigate the short-term effects of snowfall, snow melt and freeze-thaw events, but also effects of drought and precipitation (e.g., rewetting events) on the fluxes. Here, we focused on daily flux data as well as long-term means over all chambers to represent the entire forest. We will rephrase the sections in the manuscript where we talk about hot moments.

There appeared very high CO2 fluxes during a period in summer 2022 – any idea why? If possible the advantages of the high resolution GHG dataset should be worked out – but probably the advantage is not as great as can be seen from the fact that the simple Q_{10} driven model produced more or less the same CO2 budgets as the random forest model with a lot more input parameters than soil temperature.

Thank you for the comment about the very high CO2 fluxes in 2022. As shown in the manuscript, summer 2022 temperatures were very high at the site; summer 2022 was the warmest summer since we started our measurements at the site in 1997. It is well known that temperature is a major driver for any respiratory process (Davidson et al., 2006; Amthor, 2000). Also our Random Forest (RF) driver analysis revealed that soil temperature was the main driver for forest floor CO2 fluxes. Furthermore, a study by Anjileli et al. (2021) has shown that heat extremes can increase the soil respiration by 25%. Therefore, it is not surprising that we measured very high forest floor CO2 fluxes during summer 2022.

Since we found that the CO2 fluxes at our site are mainly driven by soil temperature, it is not surprising that the Q_{10} model worked reasonably well (R^2 = 0.86) and generally captured the temporal dynamics of respiration at our site. However, we would like to point out that this is not self-evident. Many studies have shown that Q_{10} models do not reproduce measured fluxes well when additional drivers impact the fluxes, for instance when soil moisture, frost, or carbohydrate limitations come into play (e.g., Rühr and Eugster, 2009; Reichstein et al., 2013; Mitra et al., 2019).

In our study, the very high fluxes in summer 2022 were not accurately reproduced by the Q_{10} model, while the RF model estimated them well. Therefore, we argue that Q_{10} models are not able to capture extreme fluxes which might be caused by more drivers than temperature alone. In contrast, our high-resolution dataset coupled with machine learning offered a more comprehensive model, which included multiple environmental variables and at the same time was able to consider chamber specific characteristics, and thus was able to capture the extreme fluxes. Thus, we think that the reliability of the RF budget is higher compared to the Q_{10} budget. In order to make this point clearer, we will also discuss this in more detail in the revised manuscript.
Moreover, unlike CO$_2$ fluxes, CH$_4$ fluxes cannot be effectively modelled using a simple and reliable approach like the Q$_{10}$ model. Therefore, we favor high-resolution data and machine learning approaches to gain insights into the underlying drivers and obtain reliable GHG budget estimates.

2. Line Comments

1.1 Title

“Continuous” is a bit confusing since there were no measurements done in 2018 and 2019

We rephrased the title to: “Forest-floor respiration, N$_2$O and CH$_4$ fluxes in a subalpine spruce forest: Drivers and annual budgets”.

1.2 Intro: P2 50-60:


We agree that soil warming experiments are not relevant to the current study. We will change this part of the introduction and incorporate the mentioned references in the introduction.

1.3 Methods:

Calibration of laser?

The instrument has not been calibrated during the measurement campaign. However, to ensure fit quality, the laser temperatures and tuning rates were adjusted on a regular basis. After the measurement campaign in 2017, the instrument was sent to Aerodyne laser instrument for repair and maintenance; the laser measuring N$_2$O was replaced in 2019. We did not calibrate the instrument since we were not interested in absolute concentrations. Furthermore, comparable instruments used for N$_2$O measurements with eddy covariance are commonly not calibrated on a regular basis either. In a study by Rannik et al. (2015) comparing the available equipment for N$_2$O flux measurements employing the EC technique and evaluating their performance, ability to detect small fluxes, and assessing long-term stability in determining the N$_2$O exchange with a comparable instrument (CW-TILDAS-CS, Aerodyne), no calibration was performed during the campaign.

1.4 Methods: Tab1:
For snow depth usually rather the maximal depth is provided than the mean depth

Thank you for the comment, we will add the maximum depth to the table.

1.5 Methods: Line 140

It is mentioned that 2% of the data were discarded after step 1 but not how many data were discarded after step 2 and 3. Please add.

Step 2 removed 0.2% and 0.7% of CO$_2$ and CH$_4$ fluxes, respectively. Step 3 excluded 6% and 9% of CO$_2$ and CH$_4$ fluxes, respectively. We will add this information to the manuscript.

1.6 Fig.1

Fig 1 and similar cases: It took me a while until I figured out that the record is not consecutive and that 1.1.2020 follows the 31.12.2017. Please indicate somehow that there is a 2 year gap in the dataset (e.g. by a gap between the panels)

Thank you for the comment. We will add a gap between the panels as suggested and update the figure caption as well. Following the updated figure:
Fig. 3: Daily mean a) air temperature and soil temperature at 5 cm depth, b) water-filled pore space at 5 cm depth (left axis) and daily sum of precipitation (right axis), c) snow depth, and daily mean forest-floor d) respiration fluxes (not gap-filled), and e) CH₄ fluxes (not gap-filled), for the years 2017, 2020, 2021, and 2022. Please note the gap in measurements between 2017 and 2020. Black lines show means over four chambers, grey bands show standard deviations among four chambers. All data shown were quality-checked as described in the main text.

1.7 Line 230

I'd rather write "after snowmelt" instead of late winter

We are not sure what the reviewer refers to since we did not use the expression “late winter” in line 230. However, did you mean line 232 where we wrote “at the end of winter”? This we can change to “after snowmelt”.

1.8 Lines 300-305

When discussing the budgets terms such as “higher-than-usual” “considerably higher” etc. should be avoided. If you want to make a solid statement that the 2022 fluxes were higher than average, then it would be necessary to apply statistics to the budget data. Otherwise no scientifically valid conclusion can be drawn.

Thank you for the comment. We will back up our statements with statistics in a revised manuscript. We will add to the text that the 2022 CO₂ budget and the annual mean Tₕ of 2022 fall outside of the 95% confidence intervals (±1.96SD, i.e., for the CO₂ budget: ± 392 g CO₂ m⁻² yr⁻¹). In case of CH₄, the 2022 budget does not fall outside of the 95% confidence interval, therefore, we will delete the word “considerably” in line 302.

1.9 Line 305

delete “exceptionally”

We will change this, thank you.

1.10 Discussion: Line 315

There is a bulk of literature that shows that CO₂ emissions are depressed under really dry conditions – please rephrase here (the WFPS in the manuscript are generally rather low, but this is likely a matter of the very low bulk densities, which might have been underestimated a bit?). Anyway, was the soil in summer 2022 really so dry?

Thanks for the comment. We would like to point out that it is difficult to obtain reliable absolute values of soil moisture at our site due to a very heterogeneous soil, and many roots and rocks in the upper horizons (please see our response on SWC to reviewer 2). Therefore, in our discussion we would rather not focus on absolute SWC or WFPS values, but rather on relative changes in soil moisture. Annual means of WFPS were indeed lowest for 2022 compared to the other years, but also soil temperatures were very high (see above). We do not see any indication of soil moisture limitation of respiration in
our data (see driver analysis). Bulk densities are indeed low at the site which is likely due to high organic matter content in the upper 5 cm. Bulk density calculations were done using soil data collected according to ICOS RI standards which gives us high trust that these values are correct (Arrouays et al., 2018, please see also Saby et al., 2023). We will rewrite the mentioned text section.

### 1.11 Table 4

Table 4 can be moved into the supplements, or also the results = measured fluxes from all the studies are shown in the table for comparison with the current ones.

Since long-term measurements for multiple greenhouse gas emissions are scarce, we would like to add the flux rates to Table 4 and keep it in the main text.

### 3. References


