

Dear Reviewer #2,

We would like to thank you for your time and for the critical and constructive feedback you have given our manuscript. We have attempted to address your comments in the manuscript as best as possible. References to line numbers refer to the preprint.

The article deals with the effect on liming across different forest types in Germany. Liming was generally performed to reduce the acidity and increase forest productivity. Given the contrasting results present in the scientific literature about this topic, the authors try to clarify the effect of this operation using both space-for-time substitution and chronosequence approaches. The topic is worth to be investigated and suitable for this journal.

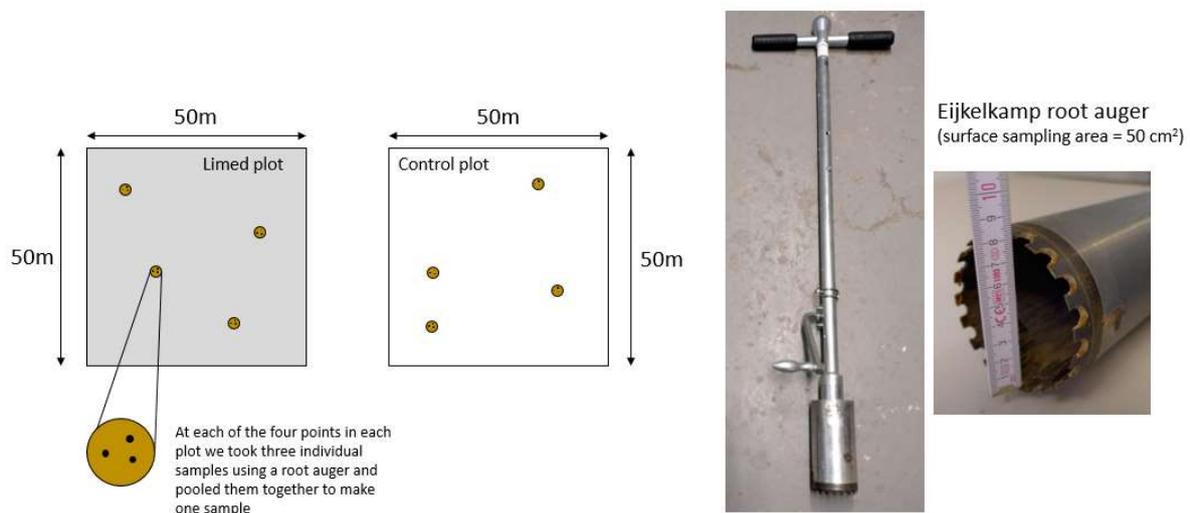
The material and method section is very clear and the methods applied both for the sampling (litter layer and mineral soil) are sounds. My only concern is about the sampling of the organic horizon that was collected on a very small area since the root auger used for sampling has a diameter of 8 cm only.

Author's response:

To clarify, at each of the four points where we sampled in each plot we took three individual samples using a root auger ($\sim 150 \text{ cm}^2 * 4 = 600 \text{ cm}^2$) and pooled them together to make one sample. The picture below demonstrates how and where samples were taken.

To make this clearer we changed the wording in the sentence (lines 115-116) to:

"At each of the four sampling locations per plot, three samples were taken in close proximity to another for each depth and pooled."



In the section where the flux measurements are described, some more info are required. For instance, it is no clear to me why the measurements were performed only in the limed sites (prior and after the liming). Why for the baseline it was not used the same approach as for the soil sampling, namely a control site and a limed site? From line 472 it seems that you measured the soil flux in both control and limed plots, isn't it? So, probably some clarification in the text are needed.

Author's response:

To clarify, we measured soil CO₂ and CH₄ fluxes in both limed and control plots prior and after liming. We have updated the text to make this clearer (line 163).

The Discussion section is well developed and each of the three different subsections clearly address the impact of liming on forest floor and mineral soil carbon stocks. Scientific literature is updated. Figures and tables are easy to understand.

Conclusions are sounds and in line with the hypothesis done at the end of the introduction section.

In general, I found the article very easy to follow and well organized.

Specific comments:

Table 1: Why the soil features in this tables are reported only for the 0-5 cm depth? Are these measures coming from the previous study or are from some historical data? Similarly, why the texture is reported for the 30-60 cm layer?

Author's response:

The soil pH and base saturation values reported in Table 1 were measured from samples taken during the last sampling campaign. We have updated the Table title to reflect this. Next, we reported soil pH and soil base saturation in the upper most soil layer (0-5 cm) because this is where the effect of liming was most pronounced. On the other hand, we reported the soil texture from a deeper soil depth (30-60 cm) because the subsoil texture better reflects the growing medium of both sites. Because of earthworm bioturbation, soil texture differences between limed and control plots were evident in the upper soil layers at some sites.

line 190 – 2044: some more info about the 13 C measurements such as the standards and type of instruments used for measurements would be welcome.

Author's response:

We have elaborated on the section describing the $\delta^{13}\text{C}$. The text now reads:

"The carbon isotope signature ($\delta^{13}\text{C}$) of CO₂ was determined by isotope ratio mass spectrometry after gas chromatographic separation, the $\delta^{13}\text{C}$ of the added lime was analyzed using an isotope ratio mass spectrometer coupled to an elemental analyzer, both at the Centre for Stable Isotope Research and Analysis (KOSI) at the University of Göttingen."

We did not include information on the standards as the precise $\delta^{13}\text{C}$ signature value is not important for the mixing pool model calculation, here only the relative differences between limed and control plots are needed.

Similarly, the samples where these measurements were performed should also be indicated (e.g. sites and number of samples).

Author's response:

We have elaborated the methods section to better explain where and how the samples were taken. The text now reads:

"The proportion of lime-derived CO₂ to the overall CO₂ flux, was determined using δ¹³C stable isotope approaches and a two-pool mixing model at the same three sites where soil GHG fluxes were measured. Unlike the soil GHG measurements, we collected gas samples for δ¹³CO₂ analysis every second measurement campaign. Samples were collected two minutes and 62 minutes after chamber closure"

Line 254: it would be nice to know how long back in time the historical data are referring. 10 yr? 20 yr? more? From figure 2 it seems that the historical data are referring to the previous 20 yr, isn't it? I appreciate table A1 at the end where all the dates of the previous sampling are indicated

Author's response:

We have updated the text in line 254 indicating when the historical data was collected. The text now reads:

"At a subset of experimental sites where historical data were available, most dating back to 1990 (Table S1), ..."

We have however moved Table A1 to the supplementary section (which is a separate document), in line with the reviewer's earlier comments to streamline the manuscript. While it is 'nice-to-have', this table simply provides background information, and we felt it was not important enough to maintain in the main manuscript.

459-464: the possible impact of earthworms could be partly evaluated looking at the C concentrations. Since earthworms move vertically more homogeneous C distribution between the different depths should be present in sites where earthworms' activity is higher compared to control plots where liming was not applied.

Author's response:

We looked into this and found that indeed there were slightly (significantly) higher C contents in the limed plots at 0-5 cm in deciduous forest sites ($8.2\% \pm 7.1\%$; mean \pm 95% confidence interval). However no trend was evident when all plots were included. Below 5cm depth no difference was evident.

We have now mentioned this in the discussion section (underlined text).

"Next, improvements in both the biochemical environment and litter palatability will likely have stimulated earthworm bioturbation (Persson et al., 2021), as is evident from the higher C-contents measured in the top 5 cm of soil in the limed plots in the broadleaf forest plots (data not shown)."