Authors' Response:

We thank the two anonymous reviewers for their thorough and detailed reports on our manuscript. Their feedback has been especially helpful and their advice has enabled us to improve our paper. We have combined our response to both reviews in this document, and address their comments directly below, in bold text. We have made amendments as advised wherever possible; however, where there is a conflict between the reviewers, we make specific reference to this and explain what action we have taken.

Reviewer #1 Report

Summary:

This paper examines greenhouse gas (GHG) fluxes taken between late July and August 2015 in a commercial conifer plantation on peat soils in Northumberland following recent felling. This study addresses current uncertainties surrounding GHG emissions from recently felled woodland on organic-rich soils, an area where the spatial and temporal variability of fluxes are poorly constrained, particularly for CH_4 and N_2O. The authors aim to improve understanding of the importance of this variability and its underlying drivers to improve predictions of GHG emissions within these systems.

The authors employ a suite of GHG flux measurement techniques including Eddy Covariance towers (EC), automated chambers sampling three microtopographic features (with a litter addition experiment), and manual chamber measurements. Fluxes measured by the automatic samplers were upscaled using remote sensing to match the contributions from ridges, hollows and drainage ditches measured under the footprint of the EC tower. Environmental and physio-chemical conditions were also measured at each chamber measurement location.

Overall, the study appears methodologically sound and represents a valuable contribution to the field. The study is novel, the conclusions and interpretations are well supported by the results. I have no major concerns regarding the results, the principal findings or the suitability of this work for publication in Biogeosciences. However, I found the supplementary materials linked were from a related, previously published study cited in the manuscipt. Further data could not be found following the EIDC link provided despite searching, please ensure this data is made available.

We have updated the information re the incorrect link. The data deposit with EIDC is being finalised at time of publication (ref EIDCHELP-107114)

There are a few minor errors in the text (identified in the specific comments), mostly editing oversights, and nothing major. However, there are a few aspects of the methodology that would benefit from improved clarity, and some areas that could be integrated more into the discussion (highlighted in the specific comments but also below).

Comments:

The authors appropriately refer the reader to related publications describing the methodology and quality control procedures for EC tower measurements. However, some of these should be summarised briefly within the manuscript or included in the supplement so it can stand alone, for the benefit of readers who are not familiar with these earlier publications.

Info added to supplementary materials

While the location of the EC tower is shown in figures, there is little information about the location of the Skyline2D automated chambers, or the locations where manual chamber measurements were made. This makes it difficult for the reader to assess the extent to which the differences between each method reflects spatial heterogeneity within the sampling location, although it is clear the EC tower and the Skyline overlap occasionally as shown in Figure 4. Figure 6 shows that the measurement locations differed substantially in terms of their soil moisture and chemistry, which is discussed in terms of their probable effect on GHG emissions seen between measurements, however this spatial context would be useful.

Fig 1 updated to include locations of various chambers etc.

Where used, there is little information given about the statistical signficance tests applied in this study or whether these were carried out on the complete dataset of on the daily mean values. Based on the figures, the latter is likely, but this could stated explicitly for reproducibility.

Section 2.8 added to clarify statistical techniques

It appears that the automated chamber fluxes were upscaled according to a simple area-weighted approach according to the proportional coverage of each micro-topographical type beneath the footprint of the EC tower. However, this is not explicitly stated or described. Such an approach may cause limitations, as EC system will measure a weighted average flux, with greater contributions coming from areas closer to and upwind of the EC tower. This footprint also changes continuously in response to changes in wind speed, direction etc., meaning that relative contributions of each land-use cover types vary both spatially and temporally. It would be worth discussing this and the effect they may have on comparison between methods and uncertainty.

The section describing differences in the diurnal fluxes is interesting but is not explored in detail in the discussion. Considering the implications for studies using only daytime manual chamber measurements, this aspect deserves more consideration.

Agreed- text amended. Please see response below re section 4.2

The automated chambers measured fluxes directly from ditch water, whereas the manual chambers sampled from the margins of the ditch. It is appreciated that the manuscript does not make direct statistical comparisons between these measurements, and that fluxes from ditch-edges are relatively uncommon and therefore valuable. However presenting these cumulative fluxes from contrasting sources alongside one another and not explicitly emphasising their differences could be misleading. The distinctions between the measurements made could be made clearer and discussed in greater detail.

Agreed- text amended. Please see response below re section 4.2

Specific comments:

Figure 2: Text size is difficult to read. Check if EC tower position is correct and remove “[I think!]” from figure caption. Frequency axis would be more useful if re-orientated horizontally.

Checked and caption text amended

Line 120 -121: It is assumed that these are the values used in the upscaling? Do the authors consider the weighted-average flux made by the EC tower depending on distance, wind speed etc., or is it a simple area weighted calculation?

We have clarified this in the text:

ridges, hollows and water bodies, which covered 43%, 46% and 11%, respectively, of the mean EC tower flux footprint for the measurement campaign

Line 185: How far upwind of the EC tower?

Clarified: ca 80 m

Line 187: Some clarification required here, that the SkyLine2D essentially acted as a floating chamber at this time? Has this been done before, and is it possible that the chamber design may have introduced some artefacts e.g. dampened surface turbulence or flow and what the effect might have been on measured fluxes? Was ebullition an issue at all?

See response to reviewer #2: text added to discussion related ebullition

Section 2.4.2: Approximately where were these collars in relation to the EC tower and SkyLine2D? Were the randomly placed collars located across different microtopography types as with the automatic chambers? The rationale (and method) for selecting random locations is not clear.

The locations of chambers have now been added to the revised Fig 1

Line 230: NEE abbreviation not spelled out beforehand.

Abbreviation now explained

Figure 3. A minor complaint, but the graphs are not perfectly aligned, making visual comparison difficult.

<u>Panel sizes have been adjusted</u>

Line 287 and elsewhere: Statistical significance tests used are not described in the text- please provide some details

A section has been added to the methods section:

Unless otherwise stated in the data processing sections above, statistical analyses were conducted using SAS 9.4 (SAS Institute, NC USA). Tests for differences were performed using paired T-test for comparison of the two EC systems and mixed effects models for tests between the effects of topography and litter addition (as fixed effects) on GHG fluxes, with block and timepoint as random factors. Differences between fixed effects were determined by comparison of least square means. For all tests, an alpha of $p < 0.05$ was used to determine significance.

Figure 7: Add units to box plot. Remove “(add units?)” from figure caption.

Units added. Caption amended.

Line 350: remove “Figure11Error! Reference source not found.”

Text amended

Reviewer #2 Report

General comments:

The paper addresses several relevant scientific questions, related to the contributions of different topographic elements to GHG fluxes in drained forests after clear-cutting and how flux estimates in such topographically diverse sites depend on the method used. By presenting data from eddy covariance and an advanced automated chamber system as well as more traditional manual chambers, the authors both provide useful interpretations on the dynamics of GHGs in the study site and comparisons between methods commonly used within the discipline but rarely assessed within a single study. The authors demonstrate that waterbodies and depressions in drained forests contribute disproportionately to landscape-scale GHG fluxes, highlighting the need for adequate stratification when manual chamber sampling is used to represent such sites.

Thus, the study represents a valuable contribution to the scientific field by improving our understanding of GHG dynamics across spatiotemporal heterogeneity and forest management choices based on state-of-the-art flux measurement methodologies.

With this being said, the study is initially framed to assess the GHG balance after tree felling, but the relatively short measurement duration (five weeks) cannot appropriately evaluate the full (annual) GHG balance. It is appreciated that the multimodal dataset presented in this study is highly resource demanding to obtain and may therefore be difficult to operate for a full year. A suggestion would therefore be to frame the study more clearly as a process- and methodology-oriented investigation of landscape heterogeneity and how it is captured by different flux measurement approaches, rather than as an assessment of the overall GHG balance following tree felling.

In extension of the above, it is suggested to make a clearer distinction between observations made during the approximately five-week summer measurement campaign and any extrapolation of these measurements to annual GHG balances. A clearer description of how cumulative fluxes were calculated and a stronger emphasis on the seasonal limitations of the dataset would strengthen the interpretation of the results and help avoid overextension of conclusions beyond the observational period.

Agreed- we have clarified these points (please see responses to specific comments below)

Further, the article would benefit from clarifying the assumptions underlying the flux estimates derived from Skyline2D and manual chamber measurements, especially considering the contribution of ebullition (if present) and diffusive fluxes in the ditches. Further, as I understand it, Skyline2D consisted only of clear (transparent) chambers while the manual chambers were only opaque (dark). I understand that after clear cutting, photosynthesis is expected to severely decrease, but could this methodological difference influence the comparability of results between these two systems? I understand that no negative NEE values were measured with the Skyline2D, but I suggest addressing this explicitly in the methods or discussion.

Text added to discussion (see comment below re section 4.111 & 4.2) to cover both ebullition and different chamber designs

While the interpretations are generally supported by the reported results, the paper would be improved by more emphasis and detailed discussions on implications for spatiotemporal stratification of sampling locations and the comparability between sampling methods which seems to be a strength of the paper. I suggest that section 4.2 ("Upscaling GHG fluxes and appropriate chamber measurement strategies") includes more discussion on how the results can help guide appropriate measurement strategies, i.e. how would the manual chamber measurements have been more optimally designed and is it possible to adequately represent this landscape with manual chambers? I think this would be interesting since many studies do not have resources to set up EC or purchase or operate systems such as Skyline 2D.

Agreed- section added to discussion as recommended:

In presenting cumulative flux estimates from a diverse range of techniques, it is vital to consider the limitations inherent due to some key differences between the methods. These include: ***i.*** *variation in chamber design; the SkyLine2D chamber was transparent, whereas the manual chambers were opaque.* CO_2 *fluxes measured with the manual chambers therefore were ecosystem respiration and the SkyLine2D fluxes were NEE. In this specific case, there was no net uptake of* CO_2 *from the SkyLine2D chambers as the chambers included no living photosynthetic material, but this key difference in design should be noted;* ***ii.*** *Chamber placement vs EC footprint. The EC footprint for each half hourly flux period is dynamic, depending on the wind speed and direction, whereas the chamber footprint is fixed. In the case of comparing SkyLine2D fluxes to EC data, the upscaling was made on the basis of a weighted average of the topography within the mean EC footprint over the study period. However, there will have been periods where the relative contributions of the topography varied in EC flux data and this must be considered as a limitation of the comparison. The manual chambers were placed randomly, without stratification and this will also be a source of error in comparisons with fluxes from both SkyLine2D and EC. Further, no water bodies were included in the manual flux data and this should be viewed with this in mind;* ***iii.*** *Diurnal variation was an important influence in* CO_2 *fluxes. This was captured by the automated techniques, EC and SkyLine2D but not the manual chambers. A manual sampling strategy should account for this and aim to measure during a window which captures the daily mean flux. At this site, the lack of diurnal variation seen in both* N_2O *and* CH_4 *fluxes mean that this was not such an issue for these two GHGs.*

Specific comments:

Abstract

A key question related to GHG balance is formulated in the beginning of the abstract, leading one to think that this manuscript is about annual balances, whereas in the conclusions section it is more about spatiotemporal variability and challenges of capturing this with different methods. Also, the campaign is only five weeks long and this short period naturally does not represent the annual flux as non-growing season fluxes are also important. So, it is a little unclear from the abstract if this paper is a process-oriented study aiming at identifying environmental drivers in this variable and manipulated landscape or if it is about post-harvest GHG balances. I would argue it is first and foremost the former.

Also, cumulative fluxes are mentioned but not how they were obtained, which is important because of the apparent disconnect between key question of annual balance and a study campaign of five weeks.

Therefore, a clearer formulation of the aim of this paper and the methodology behind estimation of the cumulative fluxes should be included in the abstract. This aim is already given in line 78-80 in the introduction and could be reused in the abstract in condensed form.

Abstract has been updated to include this information

Introduction

Line 40-41: Unclear formulation regarding the area of deep peat soil. It would be helpful to also mention the area in km2 of forested land on deep peat soils, so the importance of understanding post-harvest effects on these 2000 km2 are better framed.

Defined 'deep' peat (as > 50 cm).

L84: Use "GHG", since this abbreviation has already been presented.

M&M

Figure 1. I suggest to merge figure 1 and 2 and in the current fig. 2 also add the approximate placement of the SkyLine2D and manual chambers. Wind rose could be put in supplemental materials.

These have now been merged and the chamber locations added.

L125 (Figure 2 caption): Delete the "[I think!]".

Deleted

L131: Delete space before reference.

Deleted

L163: Missing "."

Added

Line 172-173: Nice with a comparison of the two EC systems, but it is unclear what the percentages exactly mean – why not write the linear regression slopes in numbers? The potential offset (intercept) between the two EC systems is not mentioned and maybe because it was not significantly different than zero? However, a significant offset is relevant to mention, so I think this could be included here.

Text added to include this information

L191: What was the applied rate of fresh litter? This may be useful information in the discussion of the effect of litter on GHG emissions.

Information added

Line 207-210: Can you clarify here if also CH4 and N2O fluxes were discarded when R2 for CO2 was less than 0.9? I assume you are using CO2 as an indicator of the quality of the flux measurement. So if CO2 is bad then the other fluxes will be as well? This could perhaps be clarified more in this section.

Clarification has been added.

Line 223-228: It is unclear where these CO_2 data are presented in the manuscript. Are they presented together with the data for the 8 manual chambers (line 212-222)? If so, please provide the justification for this and make it clearer here and in the results that this is the case.

Clarification made in the methods- CO_2 data from the 8 manual chambers are not included in this manuscript- only CH_4 and N_2O (and therefore 'CO_2' has been deleted from the text in line 217); the manual CO_2 fluxes presented are from the 25 survey collars alone.

Line 232-236: Please indicate clearly that it is the cumulative fluxes in the campaign period you estimate here and not annual balances.

Clarification- text added 'for the campaign period'

Results

Figure 3: Please include in the M&M how you measured PAR. It is not currently described in the M&M

New section added on environmental variables and PAR measurements added to EC section to explain how these data were derived.

Line 260-263: In fig. 4C a peculiar pattern of CH4 fluxes from water bodies (SkyLine measurements) is observed. Seems there is a constant baseline of 60 nmol m-2 s-1. This seems odd, and is not mentioned in the text. Can you comment on that here and possibly also in the text to provide a reasonable explanation for this. This constant elevated CH4 emission from the ditches must also be the reason why they are significantly higher than the ridge and furrows (fig. 5a).

The CH_4 fluxes from the water bodies are displayed using the right-hand axis on this figure which is offset to try to make the differences more clear. The caption has been amended to make this more explicit in this regard.

L313: Should it be "litter did **not** increase cumulative CO2 emissions however…"? This would be in line with $p < 0.06$ in figure 7 (if alpha 0.05 is used).

Clarified this in the text:

There was a trend towards increasing cumulative CO_2 emissions with addition of brash though this did not meet the statistical threshold of significance which we use here: the total flux of 5.62 ± 1.70 mol m^{-2} from the collars receiving brash (equivalent to 3.10 ± 0.27 g C m^{-2}) was ca. 33% greater than 4.23 ± 1.97 mol m^{-2} those without (equivalent to 2.31 ± 0.31 g C m^{-2}, $p< 0.06$ ***Error! Reference source not found.****)*

Figures 9 & 10 are really nice and great result to see the prevailing direction of CH4 emissions aligns with the ditches. However, I would suggest somehow to merge these two figures as they convey a lot of the same information. A suggestion could be to duplicate fig. 10 with a DEM overlain with the N2O diagram in fig. 9. Since day and night fluxes did not differ for CH4 and N2O current fig. 9 could be put in supplemental.

We have made the changes to the Fig 10 (now Fig 9) and moved old Fig 9 to supplementary as suggested

Line 373-378: Why do you need to scale the summer GHG balance to annual values? I think this is somewhat misleading as these upscaled numbers will probably be overestimated compared to if you had actually measured the GHG balance for a whole year (e.g. lower fluxes in colder seasons). I would rather scale to monthly values and clearly state this is summer months, e.g. to not indicate that these balances are representative for a whole year.

Agreed- we have re-scaled the cumulative flux estimate to monthly as this is more appropriate

Discussion

L403: Did this reference also measure in hollows? If not, it may have been a spatial bias in addition to the mentioned temporal bias.

It did measure in the hollows

Line 410-427: These fluxes are high and as mentioned a factor 2 higher than Peacock et al. However, it is not clear to me how exactly the chambers were installed in the ditch or if the Skyline2D chamber landed directly on the water (see description in line 186-192). You mention a floating collar (line 424), but this is not described in M&M and the potential biases, incl. ebullition. Nielsen et al. 2026 (https://doi.org/10.5194/essd-18-33-2026) also used the SkyLine2D to measure on ditches and described that the chamber could lead to forced ebullition which was removed from the dataset as this represents a bias. Can you elaborate on how you quality checked your ditch measurements from the SkyLine2D system and importantly how you distinguished between chamber artefact CH4 fluxes and "true" CH4 fluxes? Right now it seems from the text that there were no issues with the ditch measurements using the SkyLine2D, but it is not currently possible for the reader to understand how the ditch data was treated and if you filtered out all ebullition fluxes and only considered what you classify as diffusive fluxes. Provided the magnitude of the ditch fluxes and its importance for cumulative upscaled fluxes, a better description of the quality of the ditch CH4 is warranted as it could be suspected that a large part of the high CH4 fluxes may in fact be ebullitions. And it can be discussed whether a chamber like this is at all suitable for measuring ebullition fluxes, which are important components of the total CH4 flux from these types of water bodies.

We have now updated the methods to address the floating collar:

In each of the water measurement points, a floating circular landing collar was anchored in place using the same metal guides used in terrestrial plots. To minimise disturbance the chamber was programmed to close slowly when landing on the floating collar.

We have included a section in the discussion to address ebullition:

The periods of highest methane fluxes occurred following the previously mentioned rainfall events, though the response time from the waterbodies was about 24- 36 hours longer than the terrestrial increases following rainfall. We noted that these fluxes coincided with increased turbulence caused by faster flow through the drainage ditches, and therefore may lead to degassing similar to the effect reported from waterfalls (Rust et al., 2025). In light of this, we have not differentiated between diffusive and ebullitive fluxes, although we indicate a threshold above which others have calculated that a distinction could be made (Fig 3c): the calculation of Villa et al., (2021) equates to ca. 350 nmol m^{-2} s^{-1}. We suggest that interpreting the magnitude of these fluxes should be done within the context of this threshold.

L413-415: Along the lines of the previous comment regarding L373-378, I am not convinced that upscaling summer measurements to an annual flux and comparing them to annual fluxes based on seasonally explicit data is comparable. It may be more relevant to compare your data to other measurements made under similar weather conditions.

Please see previous responses re upscaling cumulative fluxes

L463-469: Was the rate of the retained residues in this study comparable to the rates in the other articles referred to? Could you comment on the difference between the estimate if litter derived CO2 in this study (33%) and the global estimate referred to (6.9%)?

Additional comments made in discussion:

though it should be noted that the amount of residue is not reported in this study and may vary considerably with tree species and stocking denisities etc.

Line 489-491: This conclusion of remarkably well agreement of EC and Skyline CH4 when omitting ditches is another indication that the ditch CH4 fluxes with SkyLine should be treated with care and at least presented in a way so it is possible to understand if there are many ebullitions or if is dominated by diffusive fluxes.

Please see previous response re ebullition

L509-515: While Vanguelova et al. (2010) found WTH to increase soil C, others, e.g. Clarke et al. (2020) (10.1016/j.foreco.2020.118877) found WTH to reduce SOC across Nordic and UK coniferous forests. It may be useful to nuance this part of the discussion since it is possible to find simultaneous increases CO2 emissions and SOC storage in response to enhanced litter inputs.

Clarification made to reflect this nuance:

Long term, WTH can increase soil carbon (C) and nitrogen (N) stocks (Vanguelova et al., 2010); however the opposite may also occur since decomposition of litter left behind at harvest can have a priming effect increasing decomposition of existing organic material (Clarke et al., 2021)

L518-524: I suggest to change this paragraph somewhat. The extrapolated C losses (12 t C y-1) may not be reliable on an annual basis due to the limited seasonal coverage. I suggest to report the C losses over the study period and compare this number to the losses during site preparation from the earlier study. This would make a similar point without having to extrapolate.

Agreed- text altered to reflect this:

The CO_2 losses from this study period alone were ca. 3 T ha^{-1} which is comparable to the 2- 4 t C per year estimated to be incurred during preparation (e.g. draining and ploughing) of this soil type in the first place (Hargreaves et al., 2003). Such an extrapolation would be oversimplistic, and a longer-term study at this site showed that losses slowed considerably in the following months so that the actual accumulated C loss in the first-year post felling was ca. 7 t C (Xenakis et al., 2021).