Discussion of “Seasonal variation in vegetation-climate interactions shape the CO_2 exchange in a degraded raised bog”

By Behrens et al.

Reviewer’s comments are marked in blue.

Author Response to Reviewer #3

This paper is well-written, methodologically strong, and engaging. Although it presents only three years of eddy covariance data from a drained, shrub-dominated raised bog, the authors maximize the value of this dataset. This type of abandoned, unmanaged peatland is both climatically relevant and significantly underrepresented in the literature. The analytical framework is thorough for a single-site study: the combination of anomaly regression, ecophysiological response curves, and the sequential filtering approach to separate TA and VPD effects demonstrates strong methodological ambition and careful attention to confounding factors. The central message, that the timing of a warm spring matters as much as its intensity for the annual carbon balance, is clearly conveyed and well supported by the data. The figures are effective, the writing is precise, and the authors are transparent about the limitations of a three-year dataset. The comments below primarily clarify and suggest ways to tighten the framing, none of which undermine the core findings.

Response: We thank the reviewer for their encouraging and positive feedback. We would like to take the opportunity to streamline the title to better reflect the main message by changing it to:
“The timing of warming matters as much as its intensity for the annual carbon balance of a degraded raised bog.”

General Comments

G1. Length of the measurement record

The eddy covariance record spans only three years, which limits the ability to draw robust conclusions about long-term drivers or to place the observed interannual variability in a broader climatic context. The authors are aware of this and acknowledge it in the conclusions, which is appreciated; this is a classic, project-related constraint. That said, the three years captured do include substantial climatic contrasts (wet/warm 2024 vs. dry 2025/2023), which partially offset one another. No additional action required, but explicitly framing this as a limitation when discussing generalisability, particularly for the spring warming and VPD findings, would strengthen the paper.

Response: We agree that the three year span is a limitation of the paper. We explicitly mention this limitation in the conclusion in lines 596-599:
“While this study spans three years, the period captured substantial inter-annual variability, providing valuable insight into peatland responses to the climate. Continued observations over longer timescales would further strengthen the assessment of whether the observed patterns—particularly the spring warming response and the seasonal divergence in VPD effects—represent persistent characteristics of this ecosystem.”

G2. Absence of remote sensing or greenness indicators for phenology

Given that vegetation phenology is a central element of the study, it is somewhat surprising that no remote sensing datasets or optical greenness indicators (e.g., NDVI, PhenoCam, or near-surface RGB camera indices) were used to independently corroborate the GPP-derived phenological transition dates. Such datasets are now widely available at high temporal resolution and would provide an empirical cross-check on the SOS/POS/EOS dates derived from the double-sigmoid fit. The authors should, at a minimum, discuss whether such data exist for the site and, if not, why they were not used, or acknowledge this as a gap.

Response: We agree that the inclusion of other flux-independent sources of phenology would strengthen the analysis. There is in fact a Phenocam installed at the site. However, the data has issues in 2023 and a data gap near the expected transition point in February/March 2024. Further, we tested remote sensing indices like daily MODIS NDVI to additionally derive the phenological transition dates but the lower temporal resolution and missing data due to noisy pixels in the satellite data product provided a worse resolution than deriving phenology from our GPP measurements. We will add these points into the Methods section 2.4 where we justify our seasonality derivation.

G3. Scale mismatch in the WTD discussion

The manuscript argues (Section 4.2.1) that the minor role of WTD anomalies on daily NEE contrasts with “its widely accepted role as the primary driver of CO_2 emissions across sites.” However, the cited cross-site studies (Evans et al., 2021; Ma et al., 2022; Tiemeyer et al., 2020) establish WTD–CO_2 relationships at the interannual or cross-site scale, not at the ±3-day anomaly scale used here. These are fundamentally different analytical contexts. A lack of daily covariation in anomalies does not contradict WTD as the dominant structural control on absolute emission levels. The framing should be revised accordingly.

Response: We recognize this point which was also raised by reviewer #2. As stated above we will rephrase this as follows:
“While WT is widely accepted as the primary driver of CO_2 emissions on the annual timescale and across sites (Evans et al., 2021; Ma et al., 2022; Tiemeyer et al., 2020), anomalies in WTD had little influence on changes in daily NEE (Fig. 7).”

Additionally we will - following the suggestion of Reviewer #1 and #2 - change the notation of groundwater level from WTD to WT.

Line Comments

l. 130 - Site heterogeneity

The site is described as moderately heterogeneous, with varying cover of Calluna, Molinia, Erica, graminoids, and individual Betula trees. Nothing wrong with that for an EC study, but it is somewhat surprising that spatial heterogeneity does not appear to be addressed further, for instance, through footprint-weighted vegetation composition or a comparison of flux behavior across dominant footprint sectors. This may be out of scope for this paper, but a brief sentence acknowledging it as a caveat would be appropriate.

Response: While there is some moderate heterogeneity, especially through some patches of dwarf birches, the vegetation is not well separated in sectors. This is visible to some extent in Figure 1. The canopy consists of ericaceous shrubs with interspersed graminoids such as Molinia Caerulea and birch seedlings throughout the footprint. We therefore do not expect different flux behaviour from different sectors of the footprint. We will make sure this point is made clear in the Site Description.

I. 140 Disappearance of SWC and TS as covariates

Soil water content (SWC) and soil temperature (TS) are measured at the site (two profiles at 5 cm depth) and reported here, but they do not appear as predictor variables in the gap-filling or driver analysis frameworks. Given that SWC and TS are direct controls on microbial decomposition and that Reco is included in the analysis, their absence warrants explanation. Was spatial heterogeneity too large for these point measurements to be representative of the footprint? Or were they excluded because of data gaps? Please clarify.

Response: This is a point that warrants clarification. During inspection of the data we found that both TS and SWC sensors had more data gaps in them than our final product of TA and WTD data. Due to the collinearity of TA and TS (R= 0.83) and SWC and WTD (R= 0.56) we decided to only incorporate one of each into the driver analysis. We will make sure this point is clarified in the Methods section to avoid any confusion.

-

I. 191 - No hydrological predictor in gap-filling

Water table depth is not included among the gap-filling predictors (TA, VPD, SWIN, hour, month). For a deeply drained site where WTD fluctuates substantially (Fig. 3e), this is surprising, even if WTD is ultimately found not to be a dominant daily flux driver (Section 3.5), excluding it from the gap-filling model means the model cannot learn any WTD signal that may exist, which could become circular when the driver analysis subsequently finds no WTD effect. At a minimum, the rationale for excluding WTD from the gap-filling predictors should be explicitly stated, even if it is not standard for non-peat-related sites.

Response: In the gapfilling model we stuck to the model structure used in the study by Vekuri et al., 2025. However, we agree that including WTD as a predictor is reasonable and we will incorporate it in the revised version.

I. 192 - Test data split for gap-filling evaluation

“20% of the data was set aside as test data to evaluate the model's performance.”

Please clarify whether this 20% was drawn by random sampling or a temporal block holdout. Random sampling of half-hourly data yields optimistically biased R^2 due to autocorrelation between adjacent time steps. A block holdout would better reflect real gap-filling skill, especially for the long equipment-failure gaps (19 and 29 days). Again, all depends on the gapfilling objective; it is very legit for small gaps.

Response: We acknowledge this difference in testing scopes for ML inference models. In the current version the test samples were randomly drawn. We will add an iterated temporal-split validation with temporal blocks of different sizes being held out during training to give reliable validation metrics for the model for different gap lengths. Further we will additionally report the RMSE and the bias in addition to only the R^2.

l. 195 - Neural network architecture

Two hidden layers with 50 nodes each represent considerable model capacity for a single-site problem with 5–6 input predictors. The ensemble approach and Laplace uncertainty quantification partially address the risk of overfitting, but it would help to be more explicit about what this model is being asked to do in practice. Outside the two major equipment failures, gaps are presumably short (hours to a few days), for which a simpler model would likely perform equally well. Does the reported $R^2 = 0.85 \pm 0.03$ reflect performance on short gaps, long gaps, or both combined? Clarifying the distribution of gap lengths and model performance by gap-length category would help readers assess whether the architecture is justified.

Response: Currently the R^2 reflects performance on randomly drawn test data which is withheld during model training, thus it likely resembles short gaps. As mentioned in the response above, we will incorporate a more robust testing procedure, providing information on how well the model fills gaps of different lengths in the flux data, and add these results to the Supplementary section.

l. 222–224 - Description of L1 and L2 in Eq. (2)

“L1 and L2 are the times of the early summer and later summer plateauing of GPP.”

This is incorrect. In Eq. (2), L1 and L2 are the asymptotic amplitude parameters (same units as GPP, not time). The timing parameters are t0 and t1. Please correct.

Response: Thanks for detecting this error. The description of the function parameters will be corrected.

l. 305 - Figure 2: no long-term SWIN reference

Long-term climatological context is provided for TA and P from the nearby DWD station, but not for SWIN (Fig. 2c). Is this because SWIN is not available from the meteo station over a comparable period, or is it because the data are not comparable? If so, this should be stated - without a long-term reference, it is harder to assess whether some of the radiation conditions in 2023–2025 were anomalous.

Response: Indeed, there is no radiation data from a nearby DWD station available. We will move the mention of long-term mean temperature and precipitation to the main text (Section Site Description).

l. 315 - Figure 3: 2024 WTD

The 2024 WTD time series (Fig. 3e) looks markedly different from 2023 and 2025, with a not-so-deep summer drop. A good reason to investigate this effect for the 2024 carbon balance, which looks very different from the others.

Response: Indeed the WTD in 2024 was markedly higher in July and August compared to the other years due to some high precipitation events. Looking at Figure 6, we found that this is not the period where NEE differs from the other years, therefore it seems that WTD did not rise enough to effectively suppress soil respiration. In contrast, Reco was lowest in 2025 (Fig. 6), when WTD dropped to its lowest levels (Fig. 3). We will make sure this point is addressed in the Discussion.

Fig. 4 (l. 353–360) - Double-sigmoid fit: envelope, bias, and phenological metric

The double-sigmoid fits capture the overall seasonal envelope of GPP well across all three years. However, the fitted curves tend to run slightly above the bulk of the observed daily values (blue dots), suggesting a modest but systematic overestimation of GPP around peak season, visible in all three panels. Since SOS, POS, and EOS are derived from the third derivative of this fit, a consistent upward bias in the envelope need not affect transition timing per se, but it may influence the partitioning of flux budgets between EGS and LGS if the peak date shifts accordingly. The authors should comment on whether this affects the seasonal budgets in Table 1, and ideally report confidence intervals for SOS, POS, and EOS.

Response: We will provide confidence intervals for the derived seasonal transitions by bootstrapping the double sigmoid function and the respective derived third derivatives. We can then present the confidence intervals for the transition dates by adding error bands in Figure 4 and reporting the values in the text. We will test how much the flux budgets vary when using the maximum versus minimum lengths of the derived seasons, and will report these ranges in the Results sections 3.4, where the seasonal budgets are presented.

Additionally, the phenological transitions are expressed in day-of-year. Given that the spring onset in this study is strongly linked to accumulated warmth (the warm spring of 2024 being the key example), it would be worth discussing whether expressing the phenological axis in accumulated degree days - standard in crop phenology - would provide a more mechanistically interpretable picture of the growing season dynamics.

Response: We agree that accumulated degree days can provide a more mechanistic representation of phenological development, particularly in systems where temperature accumulation strongly controls seasonal transitions. However, in this study we chose to express phenological transitions as day-of-year to maintain consistency with the broader peatland and ecosystem flux literature, where DOY-based metrics are more commonly reported and facilitate comparisons across years and sites.

In addition, our objective was to characterize the temporal dynamics of ecosystem fluxes and seasonal boundaries rather than to develop a temperature-driven phenological model. While the unusually warm spring of 2024 likely accelerated seasonal onset through enhanced heat accumulation, translating the seasonal metrics into degree days would introduce an additional modeling layer and assumptions regarding temperature thresholds and accumulation periods that are beyond the scope of the present study.

l. 375–378 - Table 1: season date ranges

Given that SOS/EOS dates differ substantially across years (SOS ranges from DOY 80 to 103), it would help readers to include the actual DOY ranges for NGS, EGS, and LGS for each year directly in the table as additional rows, rather than requiring a cross-reference to Sect. 2.4.

Response: Thank you for this suggestion. We will add the information of the seasonal DOY ranges to Table 1.

l. 420 - Reco–VPD relationship: confounding with SWC/WTD

Caution is warranted when interpreting the relationship between Reco anomalies and VPD, as VPD and SWC/WTD are typically correlated during dry periods, making it difficult to isolate a direct VPD effect on Reco from an indirect moisture-limitation effect. The authors acknowledge the negative SWC–flux correlation in the filtered data (Fig. S4, l. 544–545), but this concern applies equally to the Reco–VPD relationship highlighted here. The reader would benefit from a clearer statement that the VPD effect on Reco cannot be fully decoupled from concurrent soil moisture conditions, which is also the context eventually provided at l. 540.

Response: We acknowledge this issue. We employed a predictor selection process to limit the inclusion of confounding variables. This is outlined in the Methods (Section 2.6 Driver Analysis). Working on the daily scale and with anomalies, further seems to reduce the coupling of VPD and SWC. We checked for the correlation of SWC and VPD and the correlation of their anomalies in all seasons. The correlations between the actual measurements of SWC and VPD are 0.12, -0.4 and 0.06 while for the anomalies they are -0.08, -0.21 and -0.03 in the NGS, EGS and LGS respectively. We will provide this information in the Supplementary.

l. 475 - High GPP and footprint heterogeneity

The attribution of high GPP to the dense shrub canopy is plausible, but a footprint-based analysis - comparing flux signatures under wind directions dominated by the denser shrub patches versus the Betula patches or Molinia areas - could provide more direct empirical support. The authors may well be working on this as a separate study, and if so, a brief note to that effect would be appropriate. If not, the claim should be framed as a suggestion rather than a conclusion.

Response: Perhaps our description of the site was misleading. As mentioned in our previous response, the canopy is not distinctly divided into separate sections with different vegetation types. It is rather a patchwork structure largely dominated by shrubs with intergrowing grasses and individuals of trees, the dense shrub canopy is therefore not a local aspect of the ecosystem but the dominant characteristic of it overall. One stand of birches in the north-west of the tower is the only region that differs from the rest of the footprint, however the main wind directions are SW and NE, thus the impact of it on the fluxes measured is marginal. We will rephrase our site description to make sure this point is clear.

References

Vekuri, H., Tuovinen, J.-P., Kulmala, L., Aurela, M., Thum, T., Liski, J., and Lohila, A.: Improved uncertainty estimates for eddy covariance-based carbon dioxide balances using deep ensembles for gap-filling, Agric. For. Meteorol., 371, 110558, https://doi.org/10.1016/j.agrformet.2025.110558, 2025.