Final author response to manuscript ID egusphere-2025-5778, titled “Ecosystem respiration during snowmelt and soil thaw leads to a rare annual CO_2 net loss in a boreal fen”

Karoliina Särkelä, Timo Vesala, Torben R. Christensen, Juval Cohen, Angelika Kübert, Xuefei Li, Hannu Marttila, Jouni Pulliainen, Eeva-Stiina Tuittila, Efrén López-Blanco

30.4.2026

We thank the editor and the referees for their time and effort in reviewing our manuscript, as well as for their valuable comments and suggestions, which have helped improve the manuscript. We have addressed all the points raised and revised the manuscript accordingly. Please find below our point-by-point responses to the referees' comments, first addressing the comments of Referee 1 and then those of Referee 2. Please note that the line numbers refer to the clean (non-marked) version of the manuscript.

Referee #1's comments bolded:

Särkelä et al. describe the carbon balance of the Siikaneva wetland from 2005 to 2021, focusing on the role of the non-growing season. They find that while this fen is in principle a sink for CO2, that it became a source in 2016 due to an unexpected large release of CO2 at the end of the cold season, and they attribute this to the release of CO2 built-up during the winter that remained trapped due to the frozen soil. I find this study well-written and focused, and the proposed mechanism for the large release of CO2 late in the cold season makes sense to me. I have a few suggestions for improvement, but most of them are minor.

We thank REF#1 for the time and effort invested in reviewing our manuscript and for the helpful comments and suggestions, which are very useful in improving the manuscript.

1. **My main concern is the use of the word growing and non-growing season, because I'm not sure that the term is correctly used here when using NEE as a threshold. For example, the authors mention that the non-growing season already started on July 30th back in 2018. However, this is a famously warm year with an extensive drought which of course stimulated ecosystem respiration. This is also clear from Figure S3. Did the plants really stop growing in the middle of summer, or did ecosystem respiration simply**

outweigh GPP? The authors refer to Körner et al. for their definition, but that paper suggests a different term when using NEE as a threshold: the productive season. This is of course an easy terminological fix, and I suggest using the productive season – or productive growing season if need be – to avoid confusion. The term non-growing season would have to be altered accordingly, because the same issue applies there.

Reply 1.

We thank REF#1 for this useful comment: following their suggestion, we changed the wording to 'productive season' and 'non-productive season' to follow the definition given by Körner et al. (2023) when specifically referring to the budgets and periods that utilize this definition. However, to our understanding, Körner et al. has defined four different ways to define the growing season, and although all the definitions have their unique terms (growing season sensu stricto, phenological season, productive season and meteorological season), they could all be referred to as growing season. Therefore, we chose to stick with the non-growing season/growing season terminology elsewhere, when not specifically referring to the parts of the results that utilize the productive season definition.

We edited the section 2.2 that explains the season definitions as follows:

> Lines 93-103:
>
> 2.2 Non-growing season definitions: non-productive season and thermal non-growing season
>
> Many different definitions exist for the growing season and non-growing season, which can be based on environmental variables such as air temperature (referred to as the thermal growing season when persistently above 5 °C (Kollo et al., 2023; Ruosteenoja et al., 2016)), snow cover period (Arndt et al., 2020; Rafat et al., 2022), or thresholds in gross primary production (Böttcher et al., 2014). In this study, we define growing season and non-growing season based on the ecosystem carbon exchange (NEE), following one of the definitions suggested by Körner et al. (2023), and specifically referred to as productive or non-productive season. The productive season corresponds to the period when the ecosystem acts as a net carbon sink (daily NEE < 0), and the non-productive season when it is a net carbon source (daily NEE > 0). We used a threshold of three days (following the sink period definition by Aurela et al. (2004)): the productive season starts on the first of three consecutive days with negative NEE and ends on the first of three consecutive days with positive NEE.

We additionally compared the duration of the thermal non-growing season with the non-productive season, which is described more in detail in Reply 2. We clarified now how the productive season definition affects the season start and end dates, specifically during the extreme years (i.e., that positive NEE does not indicate lack of photosynthesis but rather that respiration outweighs potential photosynthesis, both during the late summer 2018 and spring 2016).

We added the following explanation of the 2018 to the discussion section:

> Lines 370-376: The two years, 2016 and 2018, with a positive annual CO_2 balance (i.e., when the ecosystem acted as a net source of CO_2 rather than a sink) showed strong deviations between the thermal and productivity-based season definitions. In 2016, the non-productive season persisted unusually late into early June (ending on 3 June), whereas in 2018 it began exceptionally early, already on 30 July. The unusually early onset of the non-productive season in 2018 has previously been linked to extreme summer drought conditions (Rinne et al., 2020). However, it is important to note that this early start of the non-productive season does not necessarily indicate a lack of photosynthesis or actual plant growth, but rather a net carbon loss at the ecosystem level.

We describe how the late start of the productive season in 2016 relates to the thermal growing season in Reply 2, and GPP in detail in Reply 3.

2. **Also, it would be good to see what the meteorological growing season would look like. For this, it would be very helpful if a figure similar to Figure 2 could be made that shows air temperature instead of NEE, which is much clearer than what's shown in Figure S1. Similarly, it would be helpful to have such figures for GPP and Reco as well.**

Reply 2.

A couple of things here. First, we had chosen initially to focus only on the NEE as it is the direct measurement. However, we now examined GPP during the spring 2016, which is described in more detail in Reply 3. Second, great suggestion about a potential figure including a link with thermal growing season: we added the suggested figure (see Fig. 2b) in the revised manuscript. Our findings suggest that 2016 and 2018 are the years with largest deviations of the productive season from the thermal growing season.

We added an explanation of the methodology to the method section:

> Lines 104-108: We defined the thermal growing season for each year as the period between the day of year when the cumulative sum of daily mean air temperature above a 5 °C threshold (air temperature °C − 5 °C) reached its minimum, corresponding to the point before sustained daily temperatures consistently exceeded 5 °C, and the day of year when this cumulative sum

reached its maximum, corresponding to the point after which daily temperatures no longer contributed to net accumulation above the 5 °C threshold (Ruosteenoja et al., 2016).

Additionally, we added the following text to the results section:

> Lines 207-214: On average, the thermal non-growing season accounted for approximately 53% of the year (195 ± 17 days), beginning on October 13th (±10) and ending on April 26th (±9), therefore being 31 days longer than the non-productive season. The mean absolute difference between the non-growing season periods based on the two definitions was 27 days for the onset and 12 days for the end. The largest delay between the onset of the thermal growing season and the start of the productive season occurred in spring 2016 (thermal growing season onset April 27th vs productive season onset June 3rd), whereas the earliest onset of the non-productive season relative to the thermal non-growing season occurred in 2018 (July 30th non-productive season onset vs. October 22nd thermal non-growing season onset).

In the discussion section, we added a following paragraph about the non-productive and non-thermal seasons, and how they compared:

> Lines 367-376: The non-productive season lasted on average from mid-September through the end of April (Fig. 2a) and was a month shorter than the thermal non-growing season (Fig. 2). The 5 °C threshold used to define the thermal growing season was a better predictor of the onset of the growing season than of its end, indicating a closer link between temperature conditions and spring onset than autumn termination. The two years, 2016 and 2018, with a positive annual CO_2 balance (i.e., when the ecosystem acted as a net source of CO_2 rather than a sink) showed strong deviations between the thermal and productivity-based season definitions. In 2016, the non-productive season persisted unusually late into early June (ending on 3 June), whereas in 2018 it began exceptionally early, already on 30 July. The unusually early onset of the non-productive season in 2018 has previously been linked to extreme summer drought conditions (Rinne et al., 2020). However, it is important to note that this early start of the non-productive season does not necessarily indicate a lack of photosynthesis or actual plant growth, but rather a net carbon loss at the ecosystem level.

Additionally, in the discussion section, we also clarified how the 2016 thermal growing season start was not delayed:

> Lines 438-441: Additionally, the thermal growing season started on April 27th (Fig. 2b) one day after the long-term mean, indicating that spring thermal

conditions were within the typical range and therefore unlikely to have constrained the onset of ecosystem activity or reduced early-season carbon uptake.

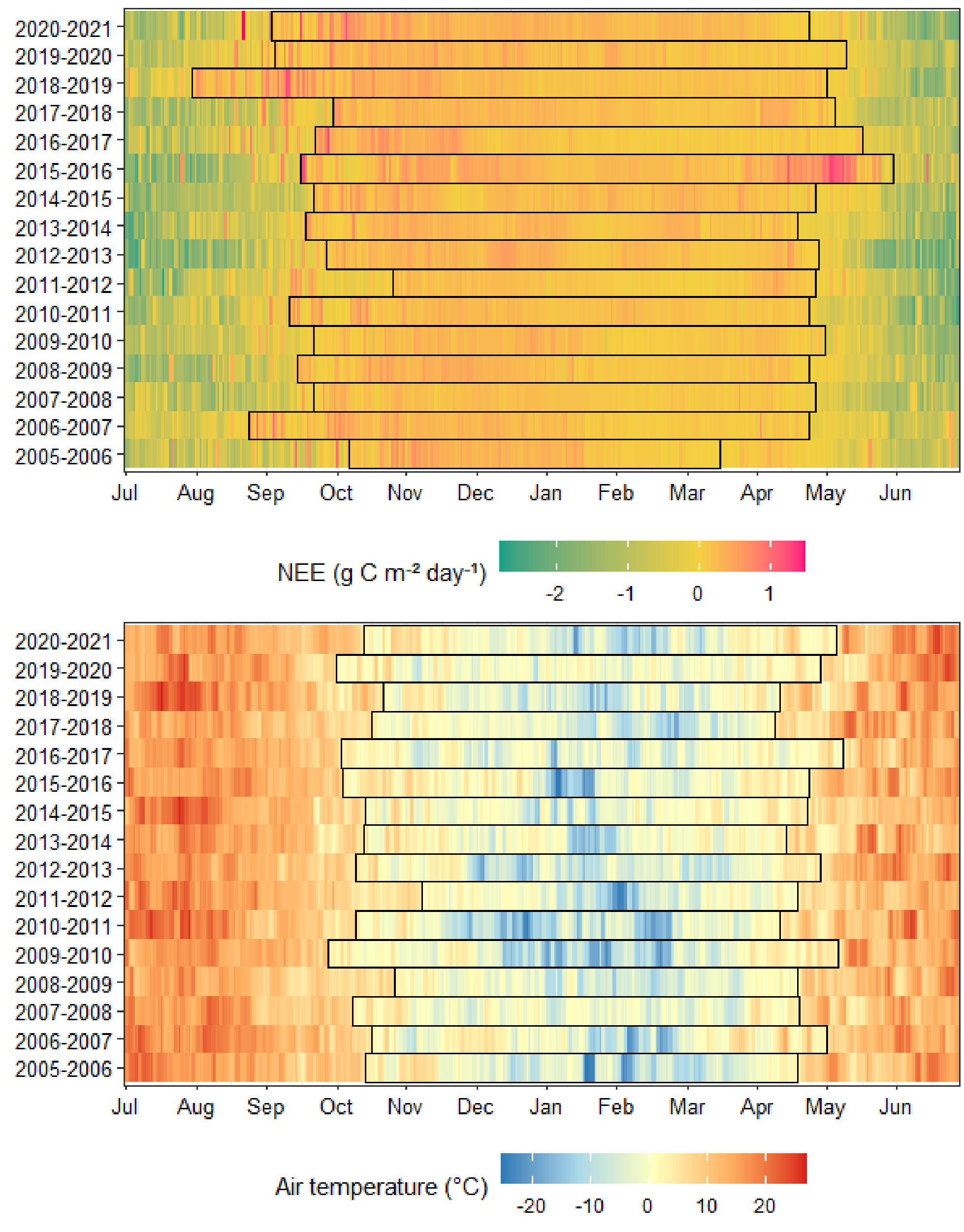

Figure 2. Figure 2. (A) Daily net ecosystem exchange of CO_2 (NEE; g C m^{-2} day^{-1}) in 2005–2021 and the non-productive season (black rectangular outline) periods centered in the winter. Each row represents one year, beginning in July and ending in June of the following year. Colors indicate daily mean NEE values, where the yellow–green gradient reflects negative values (net carbon sink) and the yellow–pink gradient reflects positive values (net carbon release). Black rectangular outlines mark the NGS periods, defined as the period beginning after the first three consecutive days of positive NEE and ending after the first three consecutive days of negative NEE. On average, the NGS started on September 17th and ended on April 30th.

(B) Daily mean air temperature for 2005–2021, with the thermal non-growing season centered on winter. Each row represents one year, beginning in July and ending in June of the following year. Black rectangular outlines mark the thermal non-growing season following the definition of Ruosteenoja et al., 2016. The season boundaries are determined from the annual cumulative sum of daily temperature deviations from a 5 °C threshold: the growing season begins when the cumulative sum reaches its minimum and ends when it reaches its maximum, indicating that temperatures remain persistently above or below 5 °C. On average, the thermal non-growing started on October 13th and ending on April 26th.

3. **Otherwise, the event that occurred in early 2016 is very reminiscent of the conditions of a frost drought. This can lead to extensive shrub and tree damage. While the sedges, rushes and mosses present at Siikaneva are probably quite resilient to such events, it is possible that some vegetation got damaged and that this delayed the uptake of carbon after snowmelt. Was the onset of GPP later than in other years when comparing to, for example, growing degree days or a simple temperature sum?**

Reply 3.
We thank the REF#1 for this helpful comment. As suggested, we examined the dynamics of GPP in 2016 to evaluate whether reduced photosynthetic activity could explain the spring NEE patterns. We compared the GPP during spring 2016 with the long-term behavior across the study period and found that GPP in spring 2016 closely followed the long-term mean and did not exhibit anomalous behavior (see Fig S9 below, now also added to the Supplementary material). To further assess the onset of photosynthetic activity, we determined the first occurrence of five consecutive days with GPP exceeding 1 g m^{-2} day^{-1}. In 2016, this occurred on 6 May, five days earlier than the 17-year average. These observations indicate that photosynthetic recovery was likely not delayed, and therefore reduced GPP is unlikely to explain the anomalous NEE patterns.

We added the following text to the result section describing the conditions in 2016 spring:

> Lines 304-310: To assess whether the high positive values of NEE were related to a delayed onset of photosynthetic activity, we evaluated GPP (See Fig. S9, in the Supplement). Overall, GPP in spring 2016 closely followed the long-term mean and did not exhibit anomalous behaviour. The onset of photosynthetic activity, defined as the first occurrence of five consecutive days with GPP exceeding 1 g C m^{-2} d^{-1}, occurred on May 6th, which is five days earlier than the 17-year average and within the range of interannual variability. This timing coincided with the decline in NEE and the increase in CH_4 emissions (See Fig. S8, in the Supplement), suggesting the onset of biological activity. However, the NEE stayed positive indicating a net release of CO_2 until June 3rd, when it turned permanently negative (See Fig. 2b).

And we further added the following text bit to the discussion section 4.2 reviewing the 2016 event:

> Lines 431-441: Given the severity of the cold spell in January and the formation of a surface ice layer, it is possible that vegetation was affected by these extreme conditions, potentially delaying the onset of photosynthetic activity and contributing to the positive high NEE observed in spring. To evaluate whether elevated NEE was driven by increased respiration or reduced photosynthetic uptake, we examined GPP dynamics (Fig. S9). As GPP followed the long-term average and exceeded 1 g C m^{-2} d^{-1} already on May 6^{th}, four weeks before NEE turned permanently negative, this indicates that the onset of photosynthetic activity was not delayed. Therefore, the persistently high positive NEE values observed in spring were not primarily driven by reduced photosynthetic uptake, but instead reflect elevated ecosystem respiration. Additionally, the thermal growing season started on April 27^{th} (Fig. 2b), one day after the long-term mean, indicating that spring thermal conditions were within the typical range and therefore unlikely to have constrained the onset of ecosystem activity or reduced early-season carbon uptake.

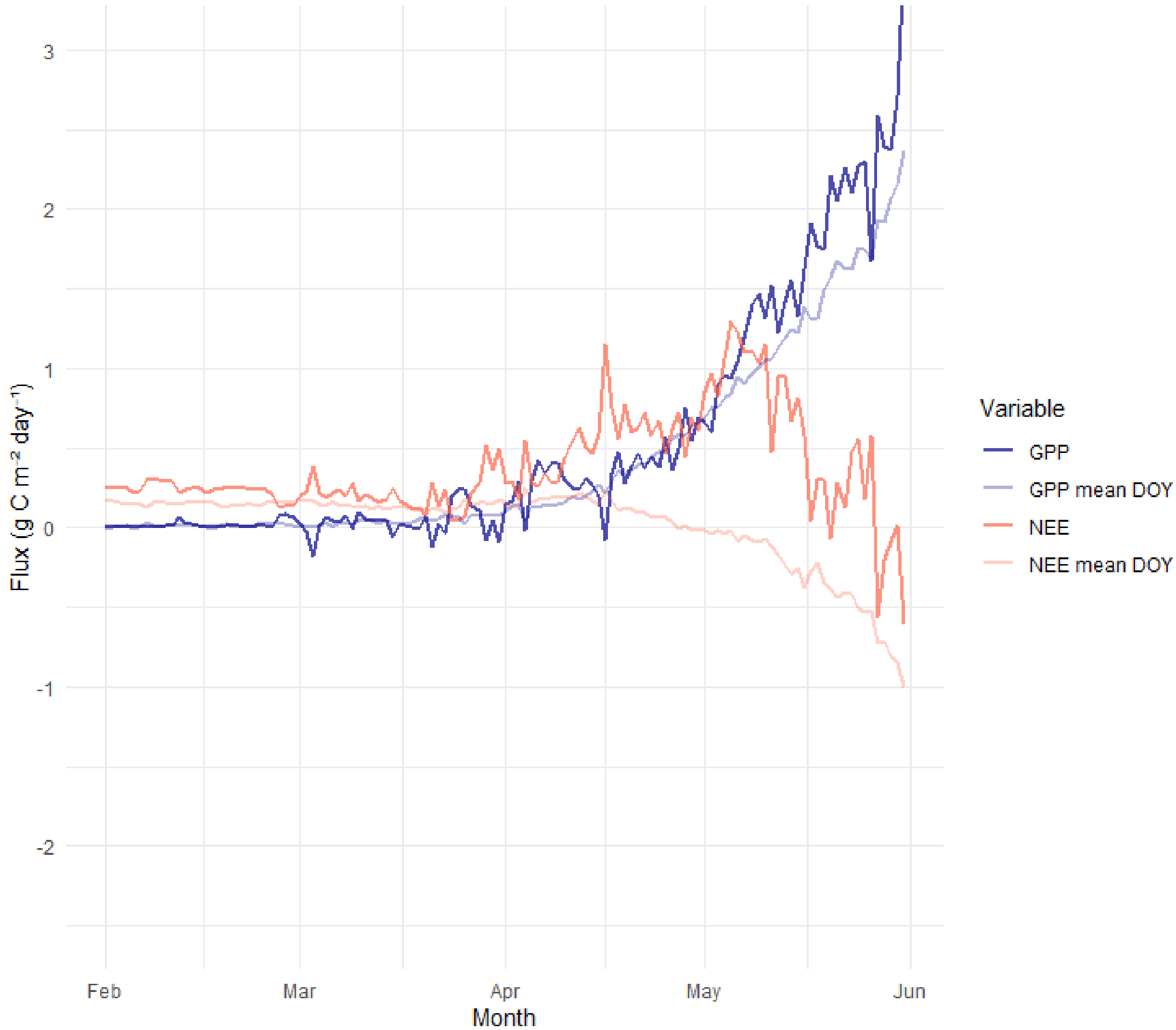

Figure S9. Daily GPP and NEE of CO_2 during spring 2016 (February–June). Solid lines show the observed daily fluxes in 2016, while the semi-transparent lines represent the long-term (2005-2021) mean. Mean fluxes were calculated by averaging fluxes for each calendar day across the full dataset and is shown to illustrate how the 2016 flux dynamics compare with the typical seasonal pattern.

Based on this, we suggest that the strong spring emissions were instead primarily driven by two complementary mechanisms:

1. Release of CO_2 accumulated under ice: Freezing of the surface layer restricted gas diffusion, allowing CO_2 produced at depth to accumulate. Higher-than-average soil temperatures at depth likely enhanced microbial activity, contributing to a larger pool of stored CO_2 released once transport pathways reopened.
2. Increased instantaneous respiration during thaw: Because GPP did not indicate delayed spring onset, we suggest that the increased respiration is a result of either enhanced decomposition of labile carbon released when ice disrupted the

peat matrix and plant tissues, or by oxygen-rich snowmelt water stimulating aerobic respiration (as suggested by Arndt et al., 2020).

These interpretations are supported accordingly by additional data sources:

- Phenocamera observations (placed in the Supplementary: Fig. S10): Images document an ice layer during winter which is not observer other years. The first CO_2 emission peak occurred when the ice layer started melting while soil temperature at 5 cm remained below 0 °C, suggesting that this peak was primarily driven by the release of previously accumulated CO_2. A second emission peak occurred after the zero-curtain period, when soil temperatures rapidly increased above 0 °C, indicating enhanced instantaneous respiration.
- Methane (CH_4) flux dynamics (placed in the Supplementary: new Figure S8): CH_4 emissions also increased during snow- and ice-melt and were higher than in other years, supporting the hypothesis of gas accumulation beneath the ice layer (see new Figure S8). However, CH_4 subsequently declined while CO_2 emissions continued to increase (Figure S8). We suggest this divergence is likely due to either enhanced decomposition of labile carbon or oxygen-stimulated respiration, as both processes increase CO_2 fluxes but would not necessarily increase CH_4 production. Comparing CO_2 and CH_4 fluxes thus help us differentiate between contributions of accumulated gas release (transport) and enhanced microbial respiration (production).

Based on the reasoning above, we edited the discussion as follows:

Lines 408-430: Overall, our observations indicate that the 2016 events combined the release of CO_2 accumulated during winter with an additional increase in respiration during spring thaw, together enhancing total emissions.

During snowmelt, CO_2 and CH_4 emissions rose sharply while surface soil temperatures remained near 0 °C, suggesting release of stored gases from the snowpack and soil, possibly aided by ice cracking (Raz-Yaseef et al., 2017). The conditions during the previous autumn and winter were favourable for gas buildup in soil. Phenocamera observations show the presence of an ice layer during winter (see Fig. S10 in the Supplement), which appeared more extensive than in other years and likely formed after a sudden cold spell in January following a warm autumn. This ice layer would have restricted gas diffusion, allowing CO_2 and CH_4 produced at depth to accumulate. Warm conditions in late autumn 2015 and early winter kept deep soil layers warmer than average throughout the winter months (Fig. 4b), enhancing CO_2 and CH_4 production during the non-productive season and increasing the pool of gases available for release once transport pathways reopened. The random forest analysis confirmed a soil temperature legacy effect, with late-autumn and early-winter conditions influencing fluxes during spring thaw (Fig. 5c).

After snow cover disappeared, CH_4 fluxes declined while CO_2 emissions remained elevated, indicating a shift from transport-dominated fluxes to enhanced respiration. This increase in respiration was likely driven by the decomposition of labile organic matter made available through freeze–thaw damage to the organic matter in the soil. Soil freeze-up can create mechanical disturbance and break down biomass in the soil, which could promote faster decomposition in the following spring (Brooks et al., 2005; Byun et al., 2021). Freezing can also damage the winter microbial populations, which releases easily degradable carbon and nutrients to the surviving community, boosting respiration post thaw (Brooks et al., 2005). In addition to responses related to soil freezing, infiltration of oxygen-rich snowmelt water has been shown to stimulate aerobic microbial activity at depth (Arndt et al., 2020), further amplifying respiration during thaw.

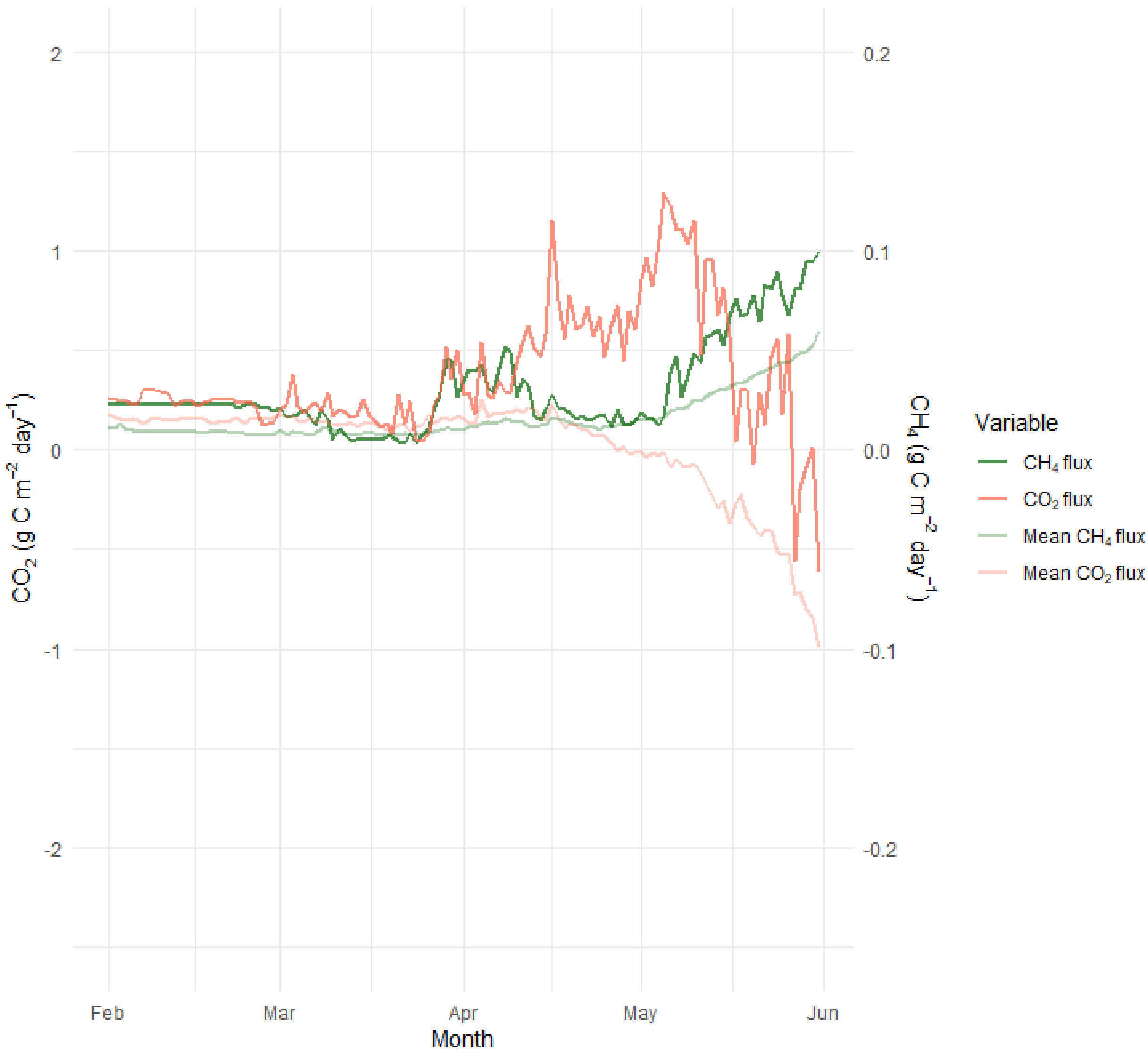

Figure S8. Daily CH_4 and NEE of CO_2 during spring 2016 (February–June). Solid lines show the observed daily fluxes in 2016, while the semi-transparent lines represent the long-term (2005-2021) mean. Mean fluxes were calculated by averaging fluxes for each calendar day across the full dataset and is shown to illustrate how the 2016 flux dynamics compare with the typical seasonal pattern.

Detailed remarks:

Line 68: "no-growing" should be "non-growing" (or non-productive)

Agreed and edited accordingly.

Line 100: This looks like a sentence remnant that should be removed.

Agreed and edited accordingly.

Referee #2's comments bolded:

The manuscript by Särkelä et al analyses a 17-year eddy covariance CO2 flux dataset from a boreal fen to investigate the drivers of non-growing season CO2 fluxes. The authors present an interesting paper, use a unique, long-term dataset, and report novel findings. The research question is highly relevant since a better understanding of the non-growing season fluxes is needed to better understand interannual variability in carbon uptake and its response to climate change. However, in my opinion, the manuscript could be substantially improved if the following issues can be addressed.

We thank the REF#2 for the careful reading of our manuscript and for the constructive comments and appreciate the time and effort invested in evaluating our work. The suggestions and comments provided have been very helpful in improving the clarity and robustness of the manuscript.

1. **One main working hypothesis and finding of the manuscript is that warm fall temperatures and a cold spell during mid-winter cause enhanced CO2 losses turning the fen from a CO2 sink to a source. First, I think the manuscript would benefit from a more detailed discussion of the underlying processes supporting this synthesis. Second, it remains unclear to me how the "release of the stored CO2 in the peat" can lead to an increase in emissions. If the freezing of the topsoil led to accumulation of CO2 in the peat column, then the resulting fluxes would be merely shifted. Mid-winter fluxes would be smaller and spring fluxes higher. For annual fluxes to be increased, respiration during the NGS must be higher, which could be driven by warmer soil temperatures. I think the distinction between CO2 production and transport should be discussed.**

Reply 1.
We thank the reviewer for this important and constructive comment. We agree that merely shifting fluxes from winter to spring would not by itself increase annual emissions, as this would reflect a transport effect rather than enhanced production. Therefore, additional processes must have contributed to the higher annual CO_2 emissions observed in 2016. One of such processes suggested in the current manuscript is the higher soil temperatures observed at depth, but the impact on instantaneous fluxes (production) during spring was not fully discussed.

In response, we revised the Discussion section to more clearly distinguish CO_2 production from transport processes. To further address this point, we examined additional data sources (phenocamera imagery, methane flux dynamics during the spring emission, and GPP which is discussed in more detail at Reply 2) to evaluate whether the observed spring emissions reflect the release of previously accumulated gases or enhanced respiration during thaw.

We now clarified in the revised manuscript that the strong spring emissions likely resulted from two complementary mechanisms:

1. Release of CO_2 accumulated under ice: Freezing of the surface layer likely restricted gas diffusion, allowing CO_2 produced at depth to accumulate. Importantly, soil temperatures at deeper layers during winter were higher than average, likely enhancing microbial activity and CO_2 production, which contributed to a larger pool of stored CO_2 released once transport pathways reopened.
2. Increased instantaneous respiration during thaw: Because GPP did not indicate delayed spring onset (discussed more on the Reply 2), we suggest this was driven by either enhanced decomposition of labile carbon released when ice disrupted the peat matrix and plant tissues, or infiltration of oxygen-rich snowmelt water stimulating aerobic respiration (as suggested by Arndt et al., 2020).

These interpretations are supported accordingly by additional data sources:

- Phenocamera observations (placed in the Supplementary: Fig. S10): Images document an ice layer during winter which is not observer other years. The first CO_2 emission peak occurred when the ice layer started melting while soil temperature at 5 cm remained below 0 °C, suggesting that this peak was primarily driven by the release of previously accumulated CO_2. A second emission peak occurred after the zero-curtain period, when soil temperatures rapidly increased above 0 °C, indicating enhanced instantaneous respiration.

- Methane (CH_4) flux dynamics (placed in the Supplementary: Figure S8): CH_4 emissions also increased during snow- and ice-melt and were higher than in other years, supporting the hypothesis of gas accumulation beneath the ice layer. However, CH_4 subsequently declined while CO_2 emissions continued to increase (see Fig. S8 in the Supplement). We suggest this divergence is likely due to either enhanced decomposition of labile carbon or oxygen-stimulated respiration, as both processes increase CO_2 fluxes but would not necessarily increase CH_4 production. Comparing CO_2 and CH_4 fluxes thus help us differentiate between contributions of accumulated gas release (transport) and enhanced microbial respiration (production).

Based on the reasoning above, we edited the discussion as follows:

Lines 408-430: Overall, our observations indicate that the 2016 events combined the release of CO_2 accumulated during winter with an additional increase in respiration during spring thaw, together enhancing total emissions.

During snowmelt, CO_2 and CH_4 emissions rose sharply while surface soil temperatures remained near 0 °C, suggesting release of stored gases from the snowpack and soil, possibly aided by ice cracking (Raz-Yaseef et al., 2017). The conditions during the previous autumn and winter were favourable for gas buildup in soil. Phenocamera observations show the presence of an ice layer during winter (see Fig. S10 in the Supplement), which appeared more extensive than in other years and likely formed after a sudden cold spell in January following a warm autumn. This ice layer would have restricted gas diffusion, allowing CO_2 and CH_4 produced at depth to accumulate. Warm conditions in late autumn 2015 and early winter kept deep soil layers warmer than average throughout the winter months (Fig. 4b), enhancing CO_2 and CH_4 production during the non-productive season and increasing the pool of gases available for release once transport pathways reopened. The random forest analysis confirmed a soil temperature legacy effect, with late-autumn and early-winter conditions influencing fluxes during spring thaw (Fig. 5c).

After snow cover disappeared, CH_4 fluxes declined while CO_2 emissions remained elevated, indicating a shift from transport-dominated fluxes to enhanced respiration. This increase in respiration was likely driven by the decomposition of labile organic matter made available through freeze–thaw damage to the organic matter in the soil. Soil freeze-up can create mechanical disturbance and break down biomass in the soil, which could promote faster decomposition in the following spring (Brooks et al., 2005; Byun et al., 2021). Freezing can also damage the winter microbial populations, which releases easily degradable carbon and nutrients to the surviving community, boosting respiration post thaw (Brooks et al., 2005). In addition to responses related to soil freezing, infiltration of oxygen-rich snowmelt water has been shown to stimulate aerobic microbial activity at depth (Arndt et al., 2020), further amplifying respiration during thaw.

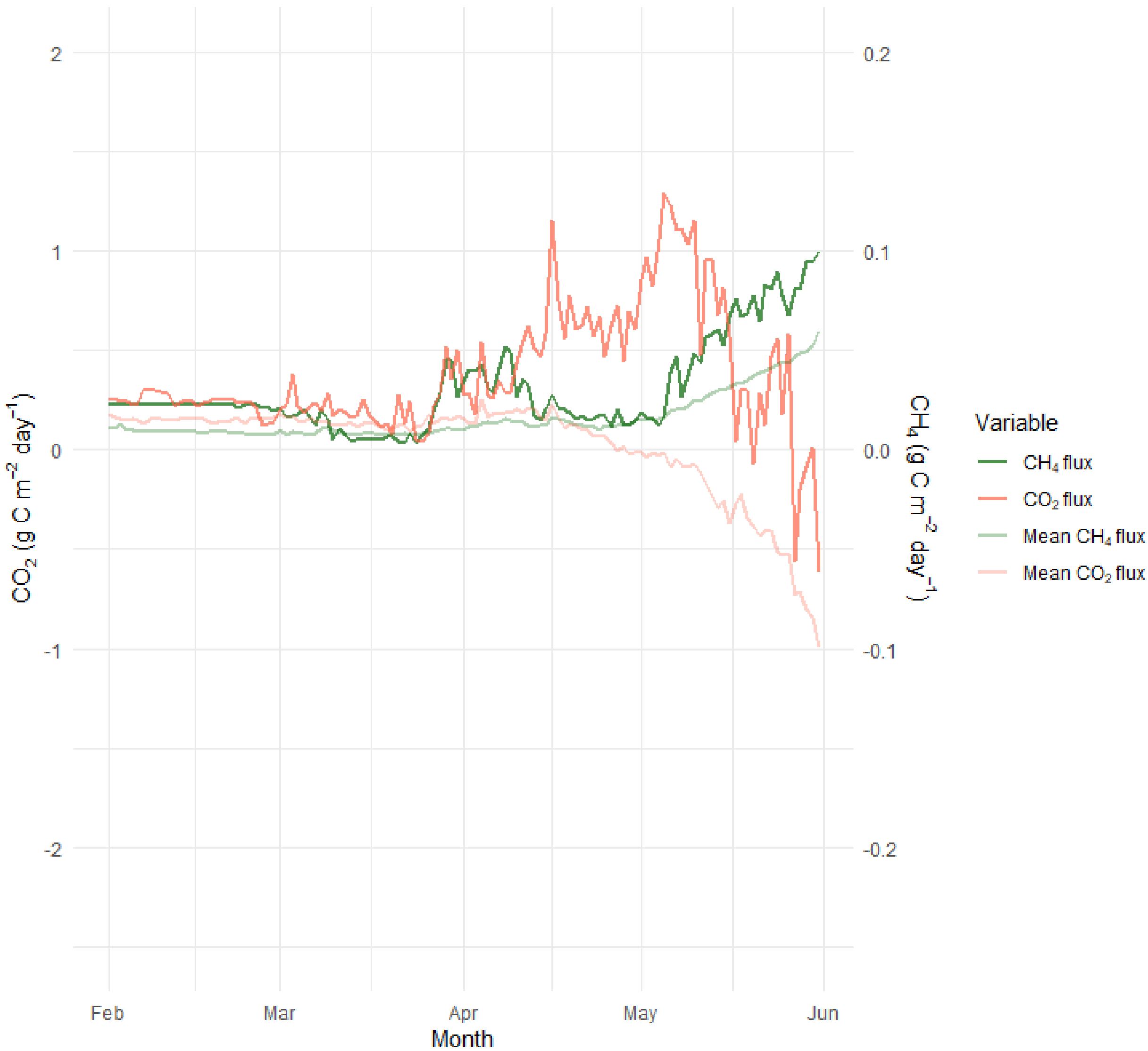

Figure S8. Daily CH_4 and NEE of CO_2 during spring 2016 (February–June). Solid lines show the observed daily fluxes in 2016, while the semi-transparent lines represent the long-term (2005-2021) mean. Mean fluxes were calculated by averaging fluxes for each calendar day across the full dataset and is shown to illustrate how the 2016 flux dynamics compare with the typical seasonal pattern.

2. **The authors mention that they partitioned NEE into GPP and Reco. However, they do not discuss the dynamics of the component fluxes and how they contribute to the 2016 event. While GPP and Reco are not directly measured, they can nevertheless help explain patterns in NEE. Furthermore, the differences in April/May might be, at least partially, be explained by decreased GPP. There is the possibility that the cold period with barely any snow cover caused frost damage leading to lower GPP in the spring. The cold event was very extreme compared to the other years. However, in the current manuscript, it is less discussed than the warm fall.**

Reply 2.
We thank the REF#2 for this helpful comment. We initially chose to focus only on the NEE as it is the direct measurement, and the partitioned fluxes are only used in the gap-filling procedure. We also clarified in the revised version that we used the readily available processed data (which utilized gap-filling algorithms and GPP and Reco partitioning).

Following REF#2 comment, we now further examined the dynamics of GPP in 2016 in more detail to evaluate whether reduced photosynthetic activity could explain the anomalous spring NEE patterns.

We compared the GPP in 2016 with the long-term behaviour across the study period. GPP in spring 2016 closely followed the long-term mean and did not exhibit anomalous behaviour (see new figure Fig. S9 below, which is also placed in the Supplement). To further assess the onset of photosynthetic activity, we determined the first occurrence of five consecutive days with GPP exceeding 1 g m^{-2} day^{-1}. In 2016, this occurred on 6 May, which is five days earlier than the 17-year average, well within the normal range of interannual variability. This timing also coincides with the point when NEE begins to decline and CH_4 emissions start to increase (See Fig. S8 above and in the Supplement). We included this explanation and the figure in the manuscript, see details below.

All in all these observations suggest that photosynthetic recovery in spring 2016 was not delayed relative to other years, and therefore reduced GPP is unlikely to explain the anomalous NEE patterns. Instead, the results support the interpretation presented in Reply 1, where the spring event is attributed to transport of accumulated gases during snowmelt **and** increased instantaneous respiration during thaw, likely enhanced by the availability of labile carbon following frost-related disturbance or oxygen-rich snowmelt water stimulating aerobic respiration.

We added the following text to the result section describing the conditions in 2016 spring:

> Lines 304-310: To assess whether the high positive values of NEE were related to a delayed onset of photosynthetic activity, we evaluated GPP (See Fig. S9, in the Supplement). Overall, GPP in spring 2016 closely followed the long-term mean and did not exhibit anomalous behaviour. The onset of photosynthetic activity, defined as the first occurrence of five consecutive days with GPP exceeding 1 g C m^{-2} d^{-1}, occurred on May 6th, which is five days earlier than the 17-year average and within the range of interannual variability. This timing coincided with the decline in NEE and the increase in CH_4 emissions (See Fig. S8, in the Supplement), suggesting the onset of biological activity. However, the

NEE stayed positive indicating a net release of CO_2 until June 3rd, when it turned permanently negative (See Fig. 2b).

And we further added the following text bit to the discussion section 4.2 reviewing the 2016 event:

Line 431-441: Given the severity of the cold spell in January and the formation of a surface ice layer, it is possible that vegetation was affected by these extreme conditions, potentially delaying the onset of photosynthetic activity and contributing to the positive high NEE observed in spring. To evaluate whether elevated NEE was driven by increased respiration or reduced photosynthetic uptake, we examined GPP dynamics (Fig. S9). As GPP followed the long-term average and exceeded 1 g C m^{-2} d^{-1} already on May 6th, four weeks before NEE turned permanently negative, this indicates that the onset of photosynthetic activity was not delayed. Therefore, the persistently high positive NEE values observed in spring were not primarily driven by reduced photosynthetic uptake but instead reflect elevated ecosystem respiration. Additionally, the thermal growing season started on April 27th, one day after the long-term mean, indicating that spring thermal conditions were within the typical range and therefore unlikely to have constrained the onset of ecosystem activity or reduced early-season carbon uptake.

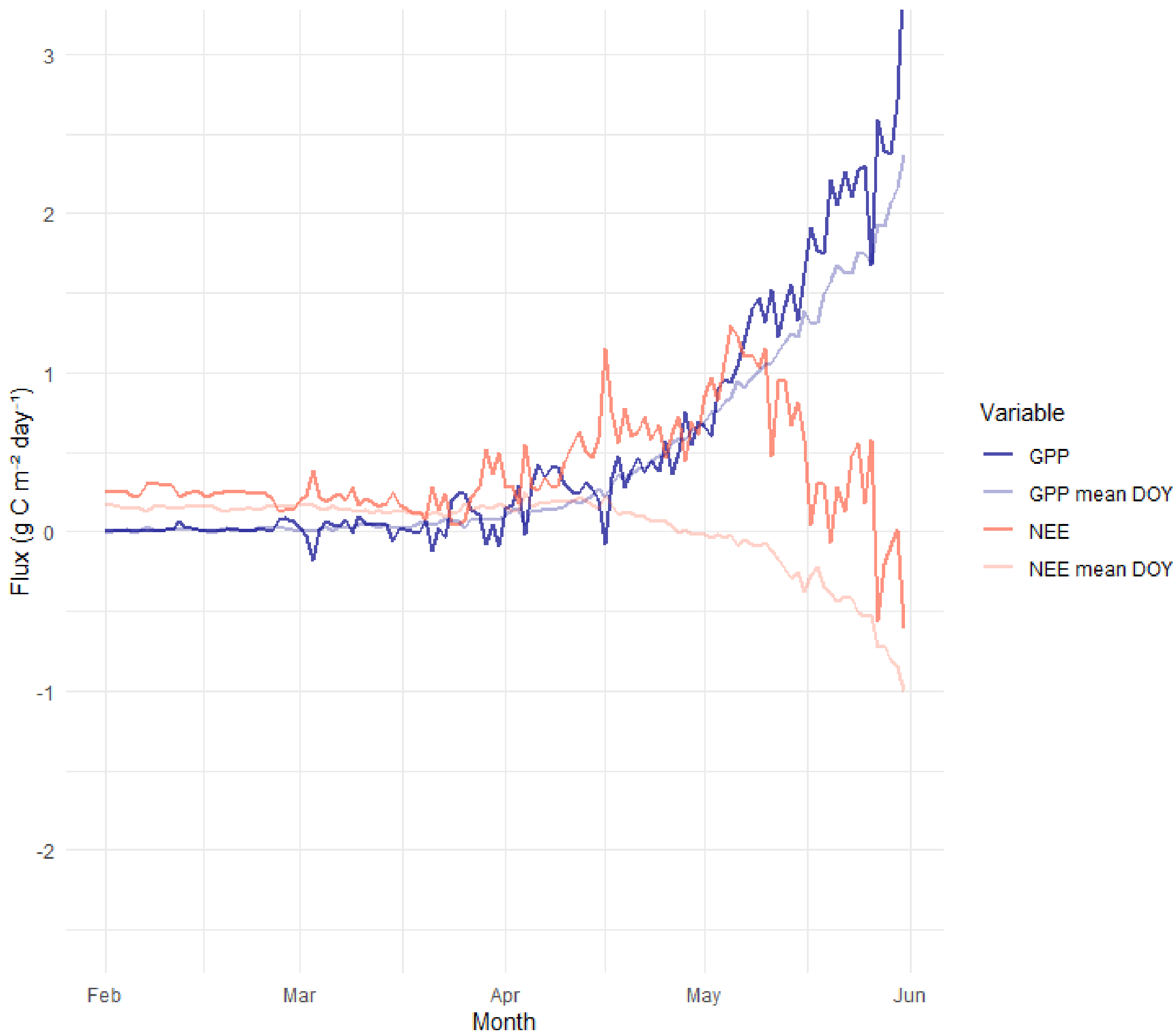

Figure S9. Daily GPP and NEE of CO_2 during spring 2016 (February–June). Solid lines show the observed daily fluxes in 2016, while the semi-transparent lines represent the long-term (2005-2021) mean. The DOY mean was calculated by averaging fluxes for each calendar day across the full dataset and is shown to illustrate how the 2016 flux dynamics compare with the typical seasonal pattern.

3. **The authors provide a clear definition of the growing season and non-growing season. Their definition is identical with carbon uptake period. The authors could consider if the definition should rather be based on GPP, which is directly related to plant productivity, and thus a better indicator of growing conditions. Since NEE includes soil respiration, the authors' definition of GS may include periods of vegetation growth.**

Reply 3.

We thank the REF#2 for this thoughtful suggestion, REF#1 also pointed to the similar issue. To address both comments, we changed the wording to ‘productive season’ and ‘non-productive season’ when referring to the seasonal budgets to follow more closely the definition given by Körner et al. (2023) and to avoid implying that the period is solely based on vegetation growth. We now clarifed how this definition affects the season start and end dates, specifically during the extreme years (i.e., that positive NEE does not indicate lack of photosynthesis but rather that respiration outweighs potential photosynthesis, both during the late summer 2018 and spring 2016).

In order to better describe the growing conditions, we included a new sub-figure (Figure 2b, see below) similar to the original heatmap of NEE (Figure 2. In the manuscript) showing now also air temperature together with thermal growing season when using 5 °C as a threshold. We believe this addition provides a direct meteorological indicator of growing conditions without relying on GPP, which is not a direct measurement. Our findings suggest that 2016 and 2018 are the years with largest deviations of the productive season from the thermal growing season.

We added an explanation of the methodology to the method section:

> Lines 104-108: We defined the thermal growing season for each year as the period between the day of year when the cumulative sum of daily mean air temperature above a 5 °C threshold (air temperature °C − 5 °C) reached its minimum, corresponding to the point before sustained daily temperatures consistently exceeded 5 °C, and the day of year when this cumulative sum reached its maximum, corresponding to the point after which daily temperatures no longer contributed to net accumulation above the 5 °C threshold (Ruosteenoja et al., 2016).

Additionally, we added the following text to the results section:

> Lines 206-213: On average, the thermal non-growing season accounted for approximately 53% of the year (195 ± 17 days), beginning on October 13th (±10) and ending on April 26th (±9), therefore being 31 days longer than the non-productive season. The mean absolute difference between the non-growing season periods based on the two definitions was 27 days for the onset and 12 days for the end. The largest delay between the onset of the thermal growing season and the start of the productive season occurred in spring 2016 (thermal growing season onset April 27th vs productive season onset June 3rd), whereas the earliest onset of the non-productive season relative to the thermal non-growing season occurred in 2018 (July 30th non-productive season onset vs. October 22nd thermal non-growing season onset).

In the discussion section, we added a following paragraph about the non-productive and non-thermal seasons, and how they compared:

> Lines 370-376: The two years, 2016 and 2018, with a positive annual CO_2 balance (i.e., when the ecosystem acted as a net source of CO_2 rather than a sink) showed strong deviations between the thermal and productivity-based season definitions. In 2016, the non-productive season persisted unusually late into early June (ending on 3 June), whereas in 2018 it began exceptionally early, already on 30 July. The unusually early onset of the non-productive season in 2018 has previously been linked to extreme summer drought conditions (Rinne et al., 2020). However, it is important to note that this early start of the non-productive season does not necessarily indicate a lack of photosynthesis or actual plant growth, but rather a net carbon loss at the ecosystem level.

In the discussion section, we clarified how the 2016 thermal growing season start was not delayed:

> Lines 438-441: Additionally, the thermal growing season started on April 27th (Fig. 2b), one day after the long-term mean, indicating that spring thermal conditions were within the typical range and therefore unlikely to have constrained the onset of ecosystem activity or reduced early-season carbon uptake.

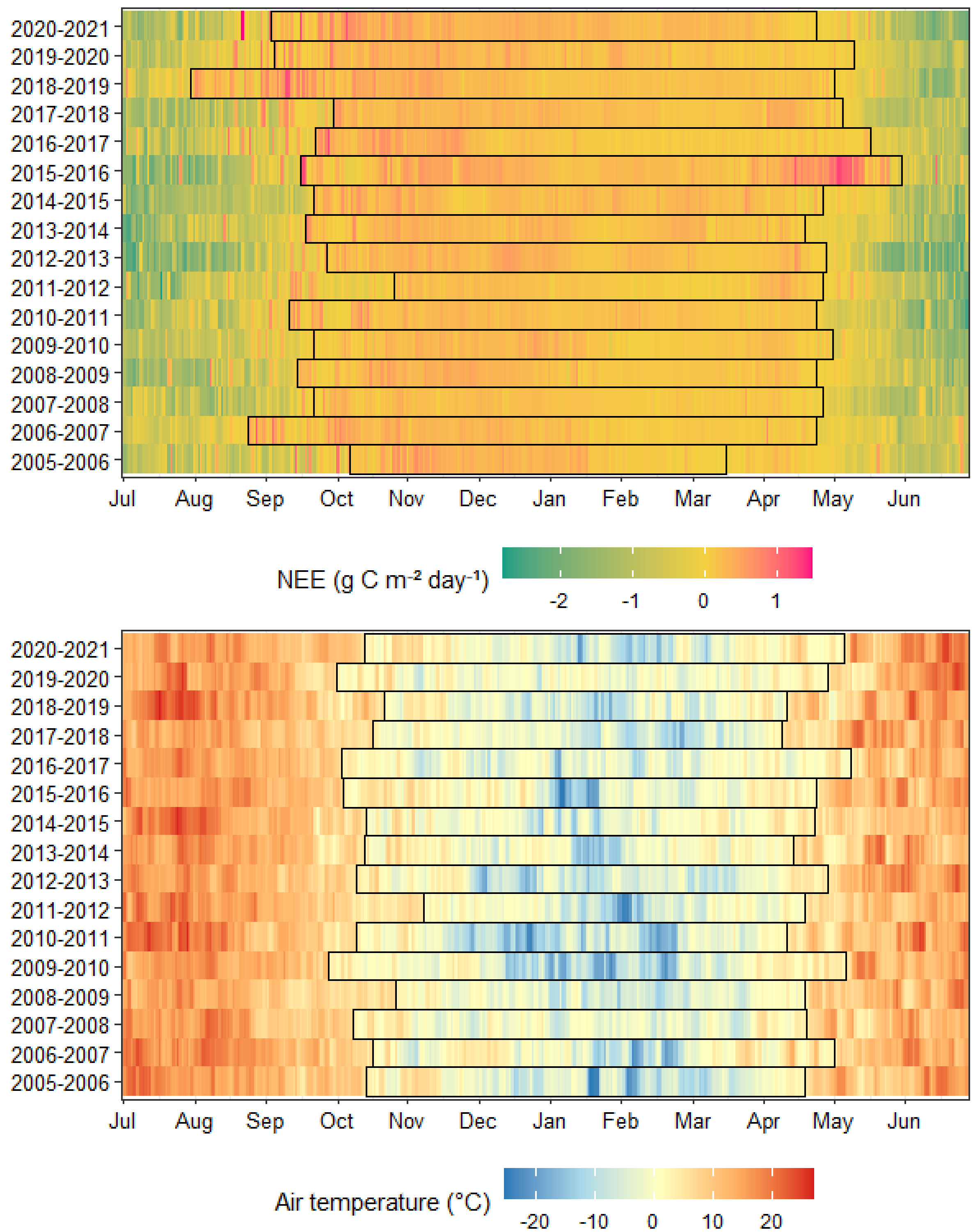

Figure 2. (A) Daily net ecosystem exchange of CO_2 (NEE; g C m^{-2} day^{-1}) in 2005–2021 and the non-productive season (black rectangular outline) periods centered in the winter. Each row represents one year, beginning in July and ending in June of the following year. Colors indicate daily mean NEE values, where the yellow–green gradient reflects negative values (net carbon sink) and the yellow–pink gradient

reflects positive values (net carbon release). Black rectangular outlines mark the NGS periods, defined as the period beginning after the first three consecutive days of positive NEE and ending after the first three consecutive days of negative NEE. On average, the NGS started on September 17th and ended on April 30th.
(B) Daily mean air temperature for 2005–2021, with the thermal non-growing season centered on winter. Each row represents one year, beginning in July and ending in June of the following year. Black rectangular outlines mark the thermal non-growing season following the definition of Ruosteenoja et al., 2016. The season boundaries are determined from the annual cumulative sum of daily temperature deviations from a 5 °C threshold: the growing season begins when the cumulative sum reaches its minimum and ends when it reaches its maximum, indicating that temperatures remain persistently above or below 5 °C. On average, the thermal non-growing started on October 13th and ending on April 26th.

4. **The NEE anomaly in 2016 was unfortunately preceded by a period of missing data. The authors address this and even detect differences in budgets when disregarding the period of missing data. However, if the CO2 emissions were shifted from mid-winter to spring, then the mid-winter period is crucial since emissions would be suppressed. Additionally, the gap-filling algorithm (which seems to be a photosynthesis and respiration model) might underperform due to the extreme conditions during this period, which are therefore not included in the model fitting. Could the authors assess the gap-filling performance shortly after the end of the data gap in 2016?**

Reply 4.

We fully understand REF#2 point. Because of that, we further assessed the gap-filling quality after the period of missing data. and added the following statement to the manuscript:

> Lines 260-263: Additionally, we evaluated the gapfilling performance after the data gap and during the anomalously high flux values. The adjusted R^2 between the gap-filled NEE and the highest quality measured data from spring 2016 (February 26th – May 30th) was 0.78, indicating a generally good agreement (See Fig. S7 in the Supplement).

To further illustrate the quality of the gap-filling, we now provide the following figure (placed in the Supplementary: Fig S7) showing the gap-filled values alongside the highest-quality raw data during spring 2016.

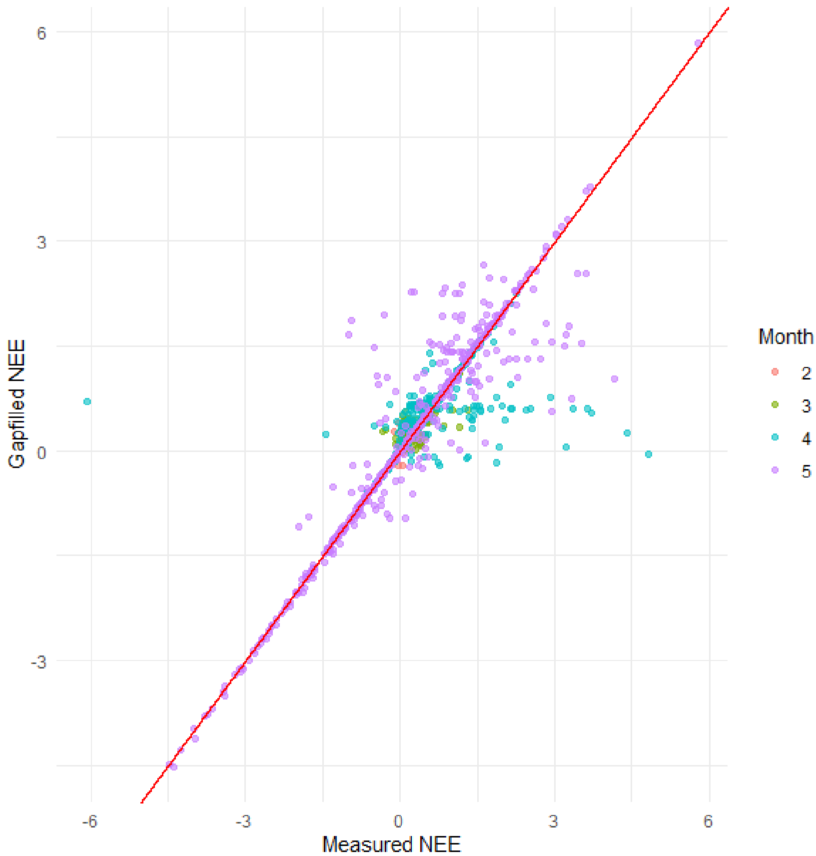

Figure S7. Half-hourly NEE observations plotted against gap-filled NEE values during spring 2016. Only the highest-quality measurements from the period 26 February to 31 May were included. This period corresponds to the time after the data gap until the end of the anomalous emission event. Colours indicate the month of observation.

Additional comments:

line 18: Could the other provide the standard deviation for this estimate of 57%?

We edited the sentence accordingly.

> Lines 18-19: On average, the non-productive season accounted for 60% of the year (226 ± 27 days), ranging from mid-September to late April, and offset 57% (± 33%) of the subsequent productive season's CO_2 uptake.

line 105: Why did the authors choose a 75% flux footprint. More commonly, 80% or 90% is used. Their definition leaves a quarter of the signal to originate from other surfaces.

We thank the REF#2 for pointing this out as it has been mistakenly left here. This sentence and its reference for the 200-meter flux footprint representing 75% was accidently placed under the section describing flux measurements, when it originally was meant to justify the use of 200-meter area for the freeze-thaw status classification in section 2.5. We will move the sentence to Section 2.5:

> Lines 155-157: Median backscatter time series (VV, VH, and VH/VV ratio) were extracted for a 200 m zone around the Eddy Covariance mast and the broader Siikaneva Fen. The 200-meter zone represents ~75% flux footprint area (Riutta et al., 2007).

The 80-90 % footprint area used for flux calculations varies between tens of meters to a maximum of 200 meters based on atmospheric conditions. We added a sentence about this remark to the section 2.3:

> Lines 112-114: 80-90% footprint was used for flux calculations, which varies between tens of meters to a maximum of 200 meters depending on atmospheric conditions.

line 111: Was the WPL term applied for a closed-path system? Why not calculating fluxes using mixing ratios?

We thank the REF#2 for pointing this out. We must admit there was a mistake in this description, and the REF#2 is right: the fluxes have been calculated using mixing ratios. We edited this sentence accordingly:

> Lines 118-120: Raw 10 Hz data were processed in EddyUH (Mammarella et al., 2016), including despiking, double coordinate rotation, sonic temperature correction, frequency response correction, and were calculated using mixing ratios.

line 115: It remains unclear what the length of the "period" is.

Sorry for not being clear. The period is the preceding two-week period- we now edited this sentence to improve clarity:

> Lines 123-124: GPP was modelled as a function of photosynthetically active radiation (PAR) and air temperature, with parameters optimized for the preceding two-week period.

section 2.5: Please add citations to this section. What is the temporal resolution of the satellite data?

The temporal resolution was, on average, two acquisitions per week. Based on REF#2 comment, we added several citations for the revised version and edited the main text as follows:

> Lines 157-161: The temporal resolution was, on average, two acquisitions per week. Due to changes in liquid water content, soil freezing typically reduces the

measured backscatter intensity by a few decibels (Hallikainen et al., 1985; Ulaby et al., 1982; Cohen et al., 2021), and the presence of wet snow further reduces the backscatter by an additional few decibels (Nagler and Rott, 2000; Luojus et al., 2007).

section 2.6: In my opinion, the random forest analysis only contributes little to the explanation of observed dynamics. The authors should consider providing more information on the model performance. Additionally, it seems as if they mostly focus on soil temperature as a driver. Would there be other simpler statistical analysis to investigate the relationship between soil temperature and lagged effects? Or could the three soil temperature aggregations be used in the same statistical model to test, which aggregation has the most predictive power? As it is presented now, it remains unclear how legacy effects and instantaneous affects are clearly separated.

We thank the REF#2 for this very helpful comment. While the random forest analyses were originally used to rank relative variable importance, we have now also assessed their overall performance to provide quantitative metrics and to determine which soil temperature aggregation performs best as a predictor. We also included now these new results and discussion in the manuscript. For each two-month period, three models differing only in the soil temperature aggregation were compared. The model with the highest explanatory power indicates which aggregation is most relevant during that period.

Our results allow both a comparison of soil temperature aggregations and a description of model performance. We note that our primary goal remains the ranking of key ecological drivers rather than maximizing flux prediction.

Model performance shows a moderate seasonal pattern, where different representations of soil temperature led to the best-performing models at different times of the year. During winter, models using current soil temperature perform best (Jan-Feb: $R^2 \approx 0.53$; Nov–Dec: $R^2 \approx 0.38$), indicating a stronger influence of present conditions. In contrast, during spring and early summer, models using soil temperature averaged over earlier periods (approximately 2–4 months prior) perform best (Mar–Apr: $R^2 \approx 0.52$; May–Jun: $R^2 \approx 0.67$), suggesting that conditions during the preceding winter influence fluxes later in the year. During peak summer (Jul–Aug), models using soil temperature averaged over the previous ~2 months show the highest explanatory power ($R^2 \approx 0.71$), highlighting the importance of accumulated thermal conditions over recent weeks. In autumn (Sep–Oct), differences between the representations are minimal ($R^2 \approx 0.62$–0.65 across models).

We added the following summary of the model performances to the results rection:

> Lines 339-344: Overall, model performance was moderate to high (R^2=0.38–0.71) and varied depending on the soil temperature representation. The best-performing model for Nov–Dec and Jan–Feb periods used concurrent soil temperature (R^2=0.38 and R^2=0.53, respectively; model used in Fig. 5a), for Mar–Apr and May–Jun the 2–4 month lagged soil temperature performed best (R^2=0.52 and R^2=0.67, respectively; model used in Fig. 5c), and for Jul–Aug the 60-day running mean showed the highest explanatory power (R^2=0.71; model used in Fig. 5b), while differences between representations were small in Sep–Oct (R^2=0.62–0.65).

Finally, regarding the comment that the analysis appears to focus mostly on soil temperature, we clarify that it was not treated as inherently more important than other predictors. Soil temperature was the only variable for which we explicitly tested legacy effects. Other predictors were included in the models but were not subjected to lagged aggregation analysis – however, to address the comment below, we have now tested WTD with the lagged aggregations as well.

line 180: Could there also be legacy effect from WTD?

We thank the REF#2 for this suggestion. We initially focused on soil temperature because it provides a clear mechanism for legacy effects particularly during periods of snow cover, when soil temperature can drive respiration that is released later.

Following the REF#2's comment, we also examined potential legacy effects of water table depth (WTD) using similar temporal aggregations as for soil temperature. The analysis showed that the instantaneous WTD was the most important, while the importance of lagged WTD aggregations declined and mostly overlapped with other variables. In other words, although a WTD legacy effect is possible, our analysis did not reveal a clear or interpretable signal. We added the figure in the supplementary material (see new Fig. S11).

line 215: Which test at which level was used here?

We edited this sentence. This was not assessed with a statistical test, therefore the wording “significant” was misleading in this context as it has been merely a qualitative statement on the absolute differences. Based on this comment, we edited the sentence as follows:

> Lines 245-246: Mean CO_2 emissions during October–December were 75% higher than those in January–March (8.4 vs. 4.8 g C m^{-2}; Fig. 3b)

Fig. 4: October data is not shown. However, the authors discuss the October/November anomaly. Additionally, the delayed topsoil thawing could result in supressed GPP, which is not discussed.

.

We thank the REF#2 for pointing this out. After carefully reviewing the data, we noticed that in fact the mistake was in the text (that talks about the November-October anomaly), when it was supposed to be November-December. We changed this issue in the text, and kept the figure as it is.

> Lines 283-284: The late autumn of 2015 was notably warm, with November-December reaching the highest mean temperature in the 2005–2021 record at 1.69 °C, 3.22 °C above the long-term average (Fig. 4c).

Moreover, the impact of the topsoil freezing and GPP was discussed in the revised version as described in Reply 1 and Reply 2.