The authors thank the reviewer for the constructive and helpful comments. We have carefully revised our manuscript based on the suggestions provided by the reviewer, as follows (reviewer's comments in bold):

This study addresses an interesting and important topic in the methane community, the seasonality of CH4 flux, and its causes, emphasizing the thawed period. The study makes use of observational results at two high-latitude sites and previously published modeled results for those sites and further analyzed the differences in CH4 flux and its dependencies on temperature and substrate, microbial biomass before and after the highest temperature. My major comments are as below:

1. The thawed period is used for the analysis; however, it is not clearly defined. I assume it is different from growing season, which is determined based on vegetation. The thaw period is defined with temperature, precisely soil temperature. I did see how it is defined. As we know that the soil temperature has a very long fluctuation around zero degrees in the shoulder season, how that is used to define the thawed period. Please clarify.

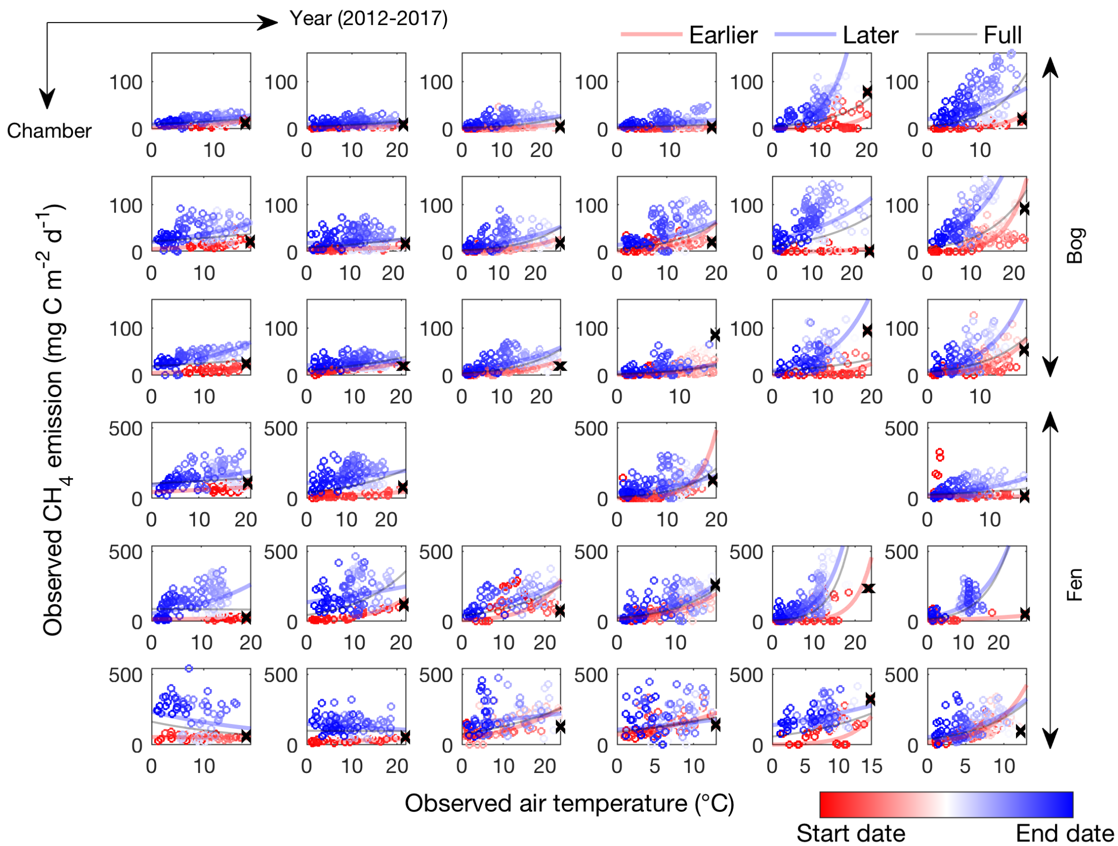
We defined thawed season as the period when the temperature being analyzed is above 1 °C (L 152-155). We agree with the reviewer that the length of thawed season may vary substantially with different temperature thresholds; however, our finding that CH_4 production becomes higher later in the thawed season is not sensitive to the definition of thawed season. For example, consistent CH_4 emission hysteresis is observed when pairing measured CH_4 emissions with soil surface temperature (when soil surface temperature is above 1 °C, Fig. 1) and air temperature (when air temperature is above 1 °C, Supplementary Fig. 1). We have examined the sensitivity of the daily mean 0-20 cm soil temperature used in our thawed season definition, and found consistent hysteretic temperature responses when using 1 °C (Fig. 2) and 0 °C (Supplementary Fig. 3) as the temperature threshold. We have improved the clarity of our thawed season definition to address the reviewer's concern (L 153-159).

We have added the following in the revisions:

L 152-159

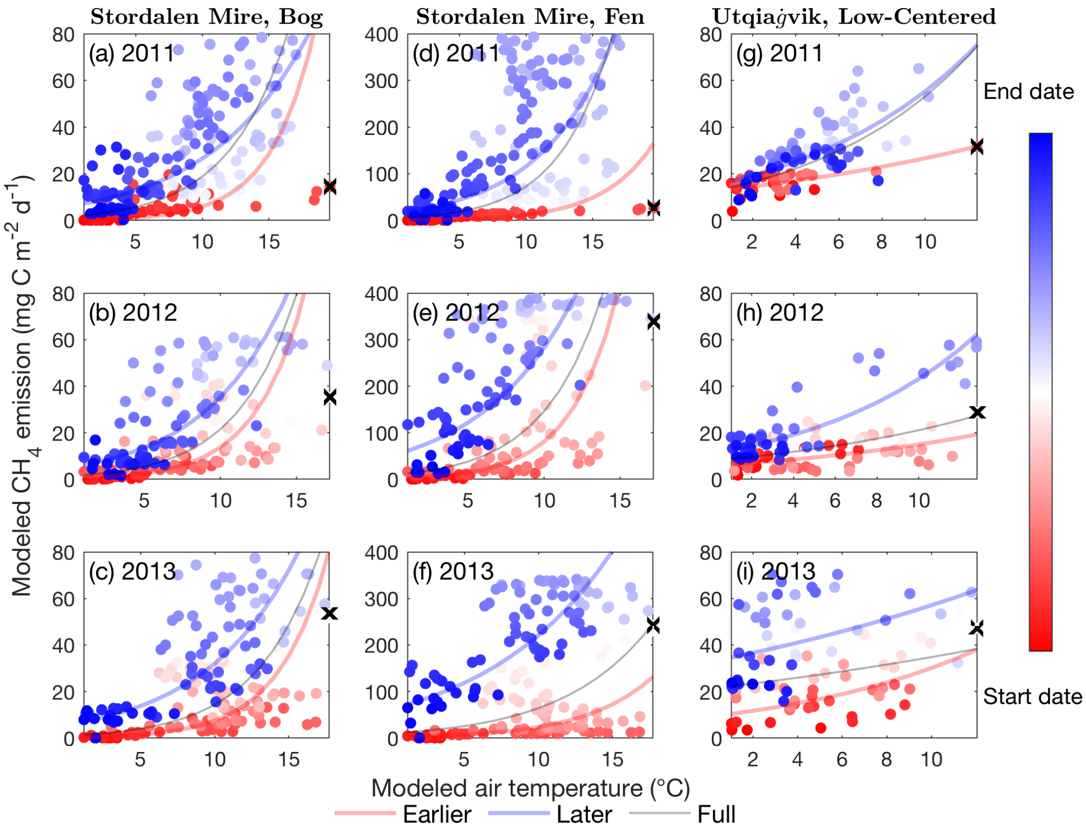
Thawed seasons were defined as the time period when measured or modeled temperatures are at least 1 °C to avoid low CH_4 emissions in the 0 – 1 °C temperature window that can alter the base reaction rate of our Boltzmann-Arrhenius functions. Four types of temperature were used in our analysis: (1) measured soil surface temperature (e.g., Fig. 1), (2) modeled vertical mean 0 – 20 cm soil temperature (e.g., Fig. 2), (3) measured air temperature (e.g., Supplementary Fig. 1), and (4) modeled air temperature (e.g., Supplementary Fig. 2).

Supplementary Fig. 1



Supplementary Figure 1. CH_4 emissions are hysteretic to air temperature measured in individual automated chambers in the Stordalen Mire bog (top three panels) and fen (bottom three panels) sites from 2012 to 2017 thawed seasons (left to right). Open circles and lines represent the daily data points and the fitted apparent CH_4 emission temperature dependence, respectively. The earlier, later, and full-season periods are colored in red, blue, and black, respectively. Earlier and later periods are defined as the time before and after the seasonal maximum air temperature denoted by black cross signs. Start date and end dates represent the beginning and ending of a thawed season defined as the period when measured daily air temperature is above 1 ˚C, respectively.

Supplementary Figure 2



Supplementary Figure 2. CH_4 emissions are hysteretic to air temperature modeled at the Stordalen Mire bog (a to c) and fen (d to f) and the Utqiaġvik low-centered polygon (g to i) from 2011 to 2013 thawed seasons. Dots and lines represent the daily data points and the fitted apparent temperature dependence, respectively. Earlier, later, and full-season periods are colored in red, blue, and black, respectively. Earlier and later periods are defined as the time before and after the seasonal maximum air temperature denoted by black cross signs. Start date and end dates represent the beginning and ending of a thawed season defined as the period when modeled daily air temperature is above 1 ˚C, respectively.

2. The authors used the highest temperature to separate the two periods; this needs to be justified. The strong fluctuation of soil temperature in one year, even the highest degree can be in a few days how to distinguish the temperature difference as < 0.1 degree in two days, particularly when those two similar temperatures are in a few days apart. I think it might be good to use a running average of the soil temperature.

The earlier and later periods of a thawed season is separated by the highest daily temperature observed or modeled in that season; however, it is just a qualitative measure describing the intra-seasonal variability detected in apparent temperature dependence of

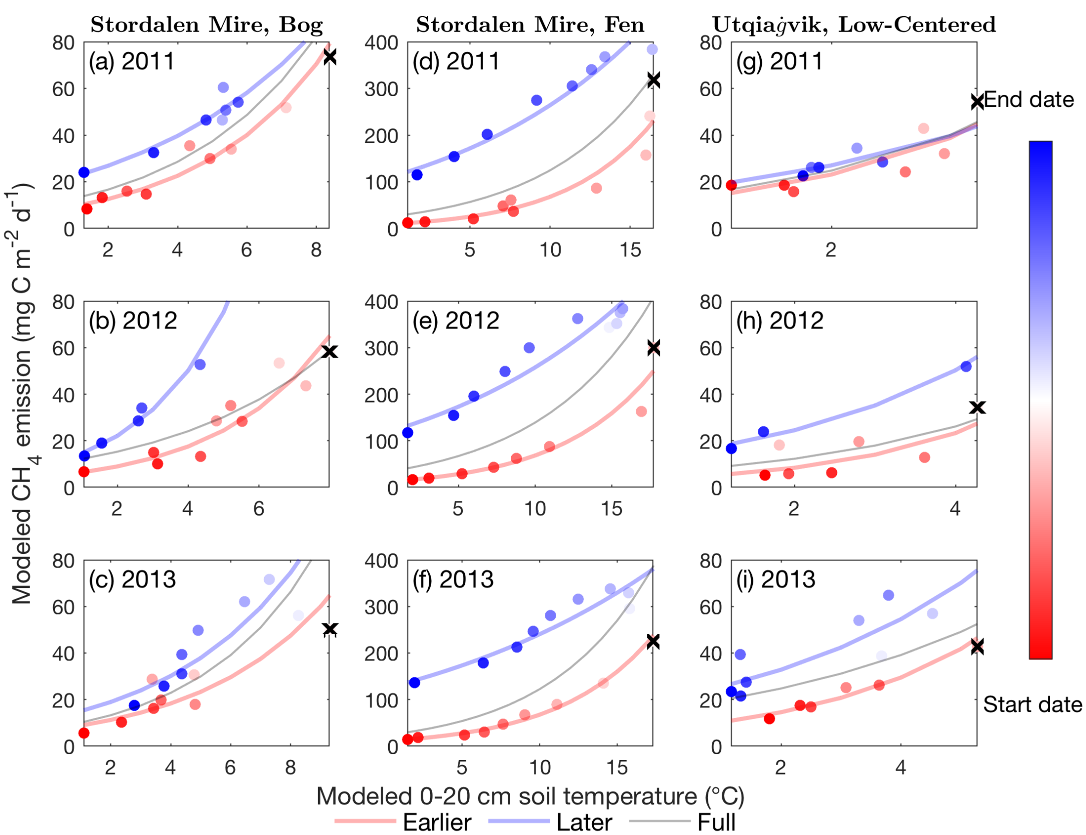
CH_4 emissions (i.e., quantifying the counterclockwise hysteresis loop shown in the scatters in Fig. 1 and 2). To address this reviewer comment, we have included the temporal variations in apparent temperature dependence of CH_4 emissions at weekly timescales (Supplementary Fig. 4) and also found higher CH_4 emissions later in a thawed season at the same temperature. Therefore, the hysteretic apparent temperature dependence of CH_4 emissions found in our study is not sensitive to the selection of earlier and later periods, nor to the temporal resolution used in representing the process. We have included the discussions above in the revised manuscript (L 253-255) to clarify this point.

We have added the following in the revisions:

L 253-255

Consistent hysteresis patterns are found at weekly timescales (Supplementary Fig. 4), suggesting that the apparent CH_4 emission hysteresis is not sensitive to temporal resolution nor the timing of maximum seasonal temperature.

Supplementary Fig. 4



Supplementary Figure 4. Weekly CH_4 emissions are hysteretic to weekly soil temperature modeled in the Stordalen Mire bog (a to c) and fen (d to f) and the Utqiaġvik low-centered polygon (g to i) from 2011 to 2013 thawed seasons. Dots and lines represent the daily data points and the fitted apparent temperature dependence, respectively. Earlier, later, and full-season periods are colored in red, blue, and black, respectively. Earlier and later periods are defined as the time before and after the seasonal maximum 0-20 cm soil temperature denoted by black cross signs. Start date and end dates represent the beginning and ending of a thawed season defined as the period when modeled weekly 0-20 cm soil temperature is above 1 ˚C, respectively.

3. Line 154, both air and soil temperature, are used to define the thawed season. It needs a very clear definition on that. In the figure, authors used ground temperature in some places; please keep consistent of air temperature, soil temperature, and ground temperature, which one is used and what it represents. Is the soil temperature < 5cm? is the ground temperature surface temperature? Did air temperature consistent with soil temperature? If not, how are they correlated? How many days of delay in terms of the highest temperature?

We have improved our descriptions in the type of temperature used in our analysis (L 155-159). Our results showed that CH_4 emissions are hysteretic to both air and soil temperatures at different temporal resolutions (e.g., Fig. 1 and Supplementary Fig. 1), suggesting that the CH_4 emission hysteresis is more sensitive to seasonal cycles in temperature than short-term variations in temperature (e.g., time lags between air and soil temperatures).

4. Although two sites are claimed to be used in the analysis, they are not in equal weight in the analysis. The authors claimed that one site has strong variation, while the other does not. This is not a solid justification.

We have revised the manuscript to specify that we are presenting a detailed analysis in the Stordalen Mire fen site, although similar hysteresis patterns can be found in the Stordalen Mire bog and Utqiaġvik sites (L 30-36; 93-99). We address this reviewer comment, which also was pointed out by Reviewer #1, by indicating that results collected at Utqiaġvik are described as a case study to represent the robustness of the modeled CH_4 emission hysteresis, because similar hysteretic responses to temperature were found at that site also. An important reason that we focused our discussion on the Stordalen Mire is that we previously validated the modeled CH_4 production pathway by the relative abundance of acetoclastic and hydrogenotrophic methanogens inferred from 16S rRNA gene amplicon sequencing data (Chang et al., 2019), which is now mentioned on L 97-99.

We have added the following in the revisions:

L 30-36

Here, we show that apparent CH_4 emission temperature dependencies inferred from year-round chamber measurements exhibit substantial intra-seasonal variability, suggesting that using static temperature relations to predict CH_4 emissions is mechanistically flawed. Our model results indicate that such intra-seasonal variability is driven by substrate-mediated microbial and abiotic interactions: seasonal cycles in substrate availability favors CH_4 production later in the season, leading to hysteretic temperature sensitivity of CH_4 production and emission.

L 93-99

We used a comprehensive biogeochemistry model (*ecosys*) to investigate observed intra-seasonal changes in apparent CH_4 emission temperature dependence at two high-latitude sites: Stordalen Mire (68.2 °N, 19.0 °E) and Utqiaġvik (formerly Barrow, 71.3 °N, 156.5 °W). We focus most of the detailed analysis at Stordalen Mire, where we recently validated the modeled CH_4 production pathways using acetoclastic and hydrogenotrophic methanogen relative abundance inferred from 16S rRNA gene amplicon sequencing data (Chang et al., 2019).

5. This paper highlights the substrate control, but both acetate and H2 were not validated against to the observational data. How to prove the robustness of the study? Please clarify.

The temporal changes in CH_4 production dynamics and the relative abundance of acetoclastic and hydrogenotrophic methanogens modeled in the Stordalen Mire have been validated in (Chang et al., 2019), suggesting that the model can reasonably represent the observed seasonal cycles in CH_4 cycling. Although the substrate mediated CH_4 production hysteresis inferred from our model data is consistent with laboratory incubations (Updegraff et al., 1998), we do not have acetate and hydrogen measurements to support the seasonal cycles in modeled substrate concentrations. We have revised the manuscript to clarify that the aim of this model-based study is to shed light on future CH_4 model development (i.e., substrate dynamics should be properly represented), and further measurements are required to examine the substrate mediated CH_4 production hysteresis proposed here (L334-342).

We have added the following in the revisions:

L334-342

Although the CH_4 emission rates and CH_4 production pathways modeled in the Stordalen Mire fen have been examined (Chang et al., 2019), continuous substrate concentration measurements are lacking for validating the substrate-mediated hysteretic temperature responses proposed here. Wide ranges of acetate and hydrogen concentrations have been reported from incubation experiments studying methanogenesis (e.g., Hines et al., 2008; Tøsdal et al., 2015; Zhang et al., 2020); however, those values may not be used to validate the time and space specific substrate concentrations modeled at our study sites. Therefore, further studies and additional field measurements are needed to test our proposed hypothesis of the causes of observed CH_4 emission hysteresis.

6. As the conceptual diagram shows in figure7, why the figures 1 – 2 were not plotted in the similar format to clearly show the hysteretic response. The current plotting is not straightforward in terms of supporting the figure 7.

The Arrhenius fits were included in Fig. 1 and 2 to quantify the differences in apparent activation energy for CH_4 emissions inferred from different periods, and to make it easier to compare with previously published data (e.g., Yvon-Durocher et al., (2014)). In the revised manuscript, we use lighter colors for the Arrhenius fits and highlight the counterclockwise apparent hysteresis in the scatters to make it more intuitive to compare with the conceptual diagram Fig. 7

We have added the following in the revisions:

Fig 1

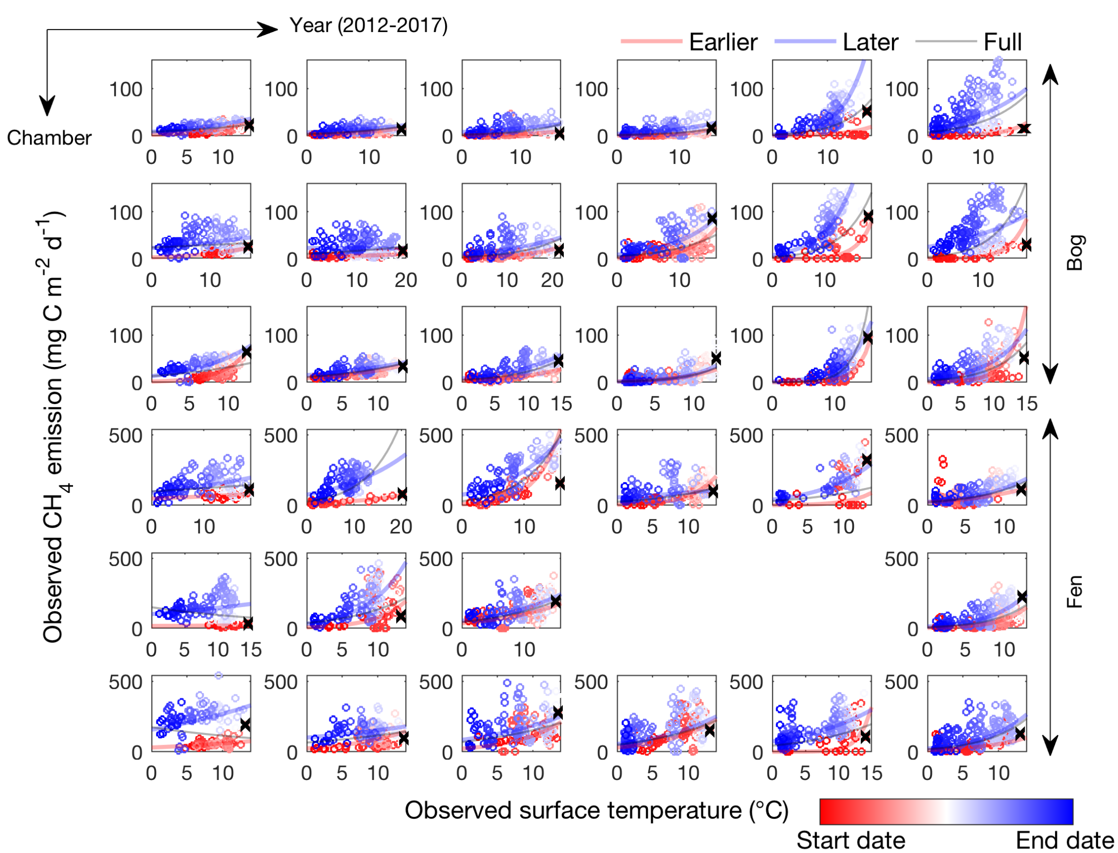


Figure 1. CH_4 emissions are hysteretic to soil surface temperature measured in individual automated chambers at the Stordalen Mire bog (top three panels) and fen (bottom three panels) sites from 2012 to 2017 thawed seasons (left to right). Open circles and lines represent the daily data points and the fitted apparent CH_4 emission temperature dependence, respectively. The earlier, later, and full-season periods are colored in red, blue, and black, respectively. Earlier and later periods are defined as the time before and after the seasonal maximum soil surface temperature denoted by black cross signs. Start date and end dates represent the beginning and ending of a thawed season defined as the period when measured daily soil surface temperature is above 1 ˚C, respectively.

Fig 2

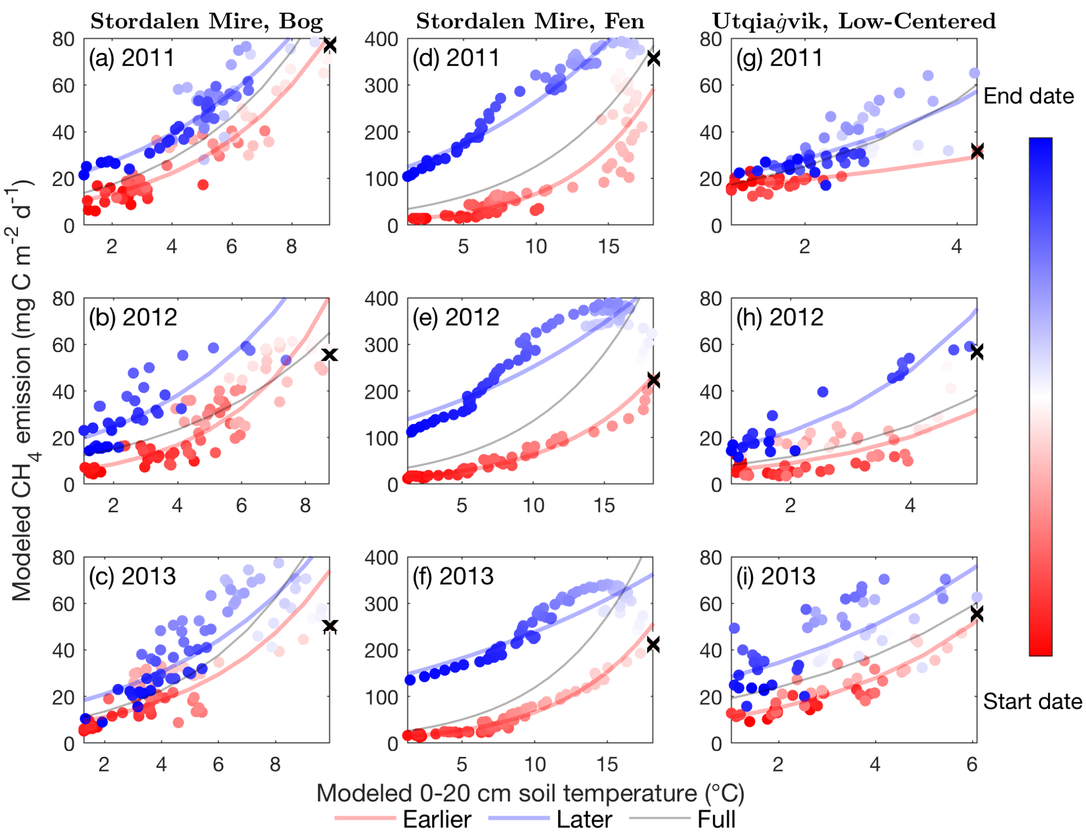


Figure 2. CH_4 emissions are hysteretic to soil temperature modeled in the Stordalen Mire bog (a to c) and fen (d to f) and the Utqiaġvik low-centered polygon (g to i) from 2011 to 2013 thawed seasons. Dots and lines represent the daily data points and the fitted apparent temperature dependence, respectively. Earlier, later, and full-season periods are colored in red, blue, and black, respectively. Earlier and later periods are defined as the time before and after the seasonal maximum 0-20 cm soil temperature denoted by black cross signs. Start date and end dates represent the beginning and ending of a thawed season defined as the period when modeled daily 0-20 cm soil temperature is above 1 ˚C, respectively.

7. Figure 9 might need to be clearly defined, see my previous comments, and put in the first section of the paper. It is the foundation of the whole manuscript.

We have improved the description of Fig. 9 based on the reviewer's comments (L 387-389; 394-397). We agree with the reviewer that it is important to point out that the observed CH_4 emission hysteresis is unlikely caused by delayed CH_4 emissions. Nevertheless, we would like to keep the current structure because it may be more straightforward to people that are not familiar with this research field.

We have added the following in the revisions:

L 387-389

In the sensitivity test, we turned off CH_4 production during the later part of the thawed season so the later-period CH_4 emissions modeled in this run are driven by lagged releases of earlier-period CH_4 production.

L 394-397

Therefore, our results suggest that lagged CH_4 emissions from residual CH_4 produced in the earlier period are not a dominant factor leading to the observed CH_4 emission hysteresis, although lagged CH_4 emissions may amplify the apparent CH_4 emission hysteresis detected in the system.

8. The figure legend of blue color to red color representing the start date to end date, does the highest temperature is in the exact middle of the thawed period? Can you mark the highest temperature on that legend and in the figures?

We have revised the figures so that the highest temperature is in the exact middle of the color bar for the thawed period. The highest temperature has been marked with black crosses in each subplot.

9. The writing is confusing in some sentences; please revise for clarity purposes.

We have reviewed the manuscript and improved the writing to address the reviewer's concern on clarity. Please refer to the highlighted texts in the revised manuscript.

10. There are a few duplicate references in the bibliography.

We have reviewed the bibliography and fixed the duplicate references.

Reference:

Chang, K.-Y., Riley, W. J., Brodie, E. L., McCalley, C. K., Crill, P. M., & Grant, R. F. (2019). Methane Production Pathway Regulated Proximally by Substrate Availability and Distally by Temperature in a High-Latitude Mire Complex. *Journal* *of Geophysical Research: Biogeosciences*, 2019JG005355. https://doi.org/10.1029/2019JG005355

Hines, M. E., Duddleston, K. N., Rooney-Varga, J. N., Fields, D., & Chanton, J. P. (2008). Uncoupling of acetate degradation from methane formation in Alaskan wetlands: Connections to vegetation distribution. *Global Biogeochemical Cycles*, *22*(2). https://doi.org/10.1029/2006GB002903

Tøsdal, A., Urich, T., Frenzel, P., & Marianne, M. (2015). Metabolic and trophic interactions modulate methane production by Arctic peat microbiota in response to warming. *Proceedings of the National Academy of Sciences*, E2507–E2516. https://doi.org/10.1073/pnas.1420797112

Updegraff, K., Bridgham, S. D., Pastor, J., & Weishampel, P. (1998). Hysteresis in the temperature response of carbon dioxide and methane production in peat soils. *Biogeochemistry*, *43*(3), 253–272. https://doi.org/10.1023/A:1006097808262

Yvon-Durocher, G., Allen, A. P., Bastviken, D., Conrad, R., Gudasz, C., St-Pierre, A., et al. (2014). Methane fluxes show consistent temperature dependence across microbial to ecosystem scales. *Nature*, *507*(7493), 488–491. https://doi.org/10.1038/nature13164

Zhang, L., Liu, X., Duddleston, K., & Hines, M. E. (2020). The Effects of pH, Temperature, and Humic-Like Substances on Anaerobic Carbon Degradation and Methanogenesis in Ombrotrophic and Minerotrophic Alaskan Peatlands. *Aquatic Geochemistry*, (0123456789). https://doi.org/10.1007/s10498-020-09372-0