**Response to the Reviewer #1:**

We thank the reviewer for the constructive comments and suggestions that have helped us rethink and improve the manuscript. We have revised the manuscript according to the reviewer's comments (see point-by-point responses below). Note there are some differences between the responses here and that within the authors' comments uploaded earlier. Specifically, we now decided to follow the reviewer's suggestion on replacing our originally modified ELM moisture scalars with new scalars and then conducted more simulations together with sensitivity analysis to parameters related to carbon decomposition and methane model. We have updated our results and discussions accordingly; our major conclusions about the model improvements in simulating soil temperature and zero-curtain period remain the same, and the improvements regarding simulating $CO_2$ and $CH_4$ fluxes still hold, although with different optimized parameterizations.

For reference, our response to comment "n" by the reviewer is labeled "R1C[n]" where R1 represents Reviewer #1. Throughout this document, the reviewer's comments are reproduced in their entirety in black, and our responses are given directly afterward in blue.

This study considers cold-season CO2 and methane emissions from Arctic tundra, which has been recently identified as an important component of the tundra carbon budget. In this work, the authors modify an ESM land surface scheme (ELMv1-ECA) to better represent freeze/thaw processes in the soil, to better represent the impact of soil moisture and temperature on soil carbon decomposition, and to allow transport of methane out of the soil during the cold season. This addresses three factors that they identified as causing the model to produce too little carbon emission during the cold season. Namely, that winter soil temperature (and therefore also liquid water) were underestimated, that the CO2 production was too little even with the correct soil temperature, and that methane transport out of the land surface was not possible during winter. They find that the model is significantly improved with these modifications, although emissions during the ZCP are now slightly overestimated instead of underestimated. They also look at long term trends.

In general the paper is clearly written and logically ordered. It is clear that the model does improve between the initial and final model versions, and now can simulate cold season emissions better than previously. I liked the fact that they tested a large number of different functions for the decomposition response, since this is certainly a major source of uncertainty in modelling.

However, before considering this for publication I would ask for some substantial justification and clarity about the changes that have been made. I would also ask for some additions to the text to consider other (potentially) important factors.

We appreciate the reviewer's constructive comments. We have provided particularly detailed elaboration for the issues pointed out by the reviewer in the revised manuscript as discussed by the following responses.

**General comments:**

=============

1) Carbon/substrate. There is plenty of evidence that availability of carbon substrate is important for controlling methanogenesis (e.g. Strom et al 2012), and soil respiration in general (e.g. Brooks et al 2004). This is not discussed in this paper. It is not clear whether the model simulates the soil carbon dynamically, or whether (dynamic or not) the soil carbon takes appropriate values in the model. For one site (IVO) there is some discussion of this (Section 3.2), but it sounds like this soil

carbon data is only used to set the soil thermal/hydraulic properties. Does it also form the substrate for soil respiration?

If the substrate is not correctly simulated, then you may compensate for this with incorrect choice of decomposition functions. I would ask for some more analysis of the soil carbon - for example compare it to observed values at the site and identify if this can be a source of bias (I would suggest add to a plot or table in the supplementary material and discuss in the main text).

**R1C1:** As we pointed out in the manuscript (lines X - Y), the dependencies of soil thermal and hydraulic properties on soil organic carbon are embedded in the model via linear relationships (Lawrence and Slater, 2008). The relationships between soil properties and organic carbon content are initialized at the beginning of simulations.

We also agree with the reviewer about the importance of carbon substrate on influencing methanogenesis activity and soil respiration, and ELMv1-ECA does account for these impacts. We now have clearly stated that in the revised manuscript (within the Appendices).

> "ELMv1-ECA explicitly simulates carbon cycle dynamics (both plant and soil) and accounts for the limitation of nutrient (i.e., nitrogen and phosphorus) availability for plant growth and the nutrient competition between plants and microbes (Burrows et al., 2020; Zhu et al., 2019; Golaz et al., 2019; Zhu et al., 2020). The ELMv1-ECA uses a Century-like soil carbon decomposition cascade model with vertically resolved soil biogeochemistry (Koven et al., 2013), and explicitly accounts for the influence of substrate and nutrient availability on soil respiration (both root and microbes) (Zhu et al., 2019)." (Lines 703 - 707 in the revised manuscript)

> "ELMv1-ECA considers the availability of carbon substrate as an important driver of methanogenesis activity and methane production (Riley et al., 2011; Xu et al., 2016)." (Lines 741 - 742 in the revised manuscript)

We would also like to conduct more analysis on soil carbon; however, due to the lack of site-level observations, we cannot compare simulated soil carbon with observation for this study. We will explore available *in situ* soil carbon datasets and incorporate these analyses in our future studies.

2) Snow. The main problem with ESM's underestimating winter soil temperature is often related to representation of snow, so I was surprised that this was not discussed in more detail, and only the phase change was considered as leading to underestimated winter soil temperatures (although snow is mentioned once in the results). For example, Burke et al (2020) show the offset between air and soil temperature in CMIP5 and CMIP6 ESM's as a function of snow depth. In models that poorly represent snow, the offset can be up to 10 degree C biased - meaning at 10 degree C cold bias in winter soil temperature. In models that improved their snow scheme between CMIP5 and CMIP6, there is a huge improvement in this. Even in a model that does not represent latent heat *at all*, the winter soil temperature offset against air temperature is substantially smaller than in models with a poor snow insulation scheme. I would therefore strongly recommend that snow is considered in terms of the simulation of winter soil temperatures. I suggest that at the point where snow depth is discussed in the results (see specific comments, below), an assessment of how well the snow is simulated should be presented with supplementary figures.

**R1C2:** We agree with the reviewer on the importance of accurately simulated snow conditions (e.g., snow thermal insulation and snow coverage-related impact). In our following paper (Tao et al., 2021), we have conducted more experiments investigating how biases in simulated snow variables (i.e., snow depth, snow water equivalent (SWE), and snow coverage) are propagated to biases in soil temperature and soil moisture, and then translated into biases in heterotrophic respiration and cold-season $CO_2$ and $CH_4$ emissions over pan-Arctic permafrost regions. Please also see our response to Reviewer #2 (R2C1). For this study, the snow depth measurements at the study sites are problematic, showing about 30 cm snow depth during summer times (see raw data by Oechel and Kalhori (2018)). For instance, at ATQ, a site that shows the most reasonable snow depth observations, there are suspicious snow depth measurements in summer (Figure R1.1b). We also checked Snow Telemetry (SNOTEL) sites but did not find one close enough to our sites for a better comparison. Indeed, continuous quality snow measurements, especially SWE, are extremely challenging to obtain (Pirazzini et al., 2018; McGrath et al., 2019).

Still, as the reviewer suggested, we have provided comparison results of simulated and observed snow depth despite the suspicious measurements (Figure R1.1 as Figure S9 in the supplementary file); we also added discussion on how sensitive the simulated soil temperature and carbon fluxes are to snow depth in the supplementary file (Figure S9). We also added a sentence to the revised manuscript.

> "Sensitivity analysis demonstrates large impacts of snow depth on simulated winter soil temperature, summer soil moisture, heterotrophic respiration, and $CO_2$ fluxes (Figure S9)." (Lines 588 - 589 in the revised manuscript)

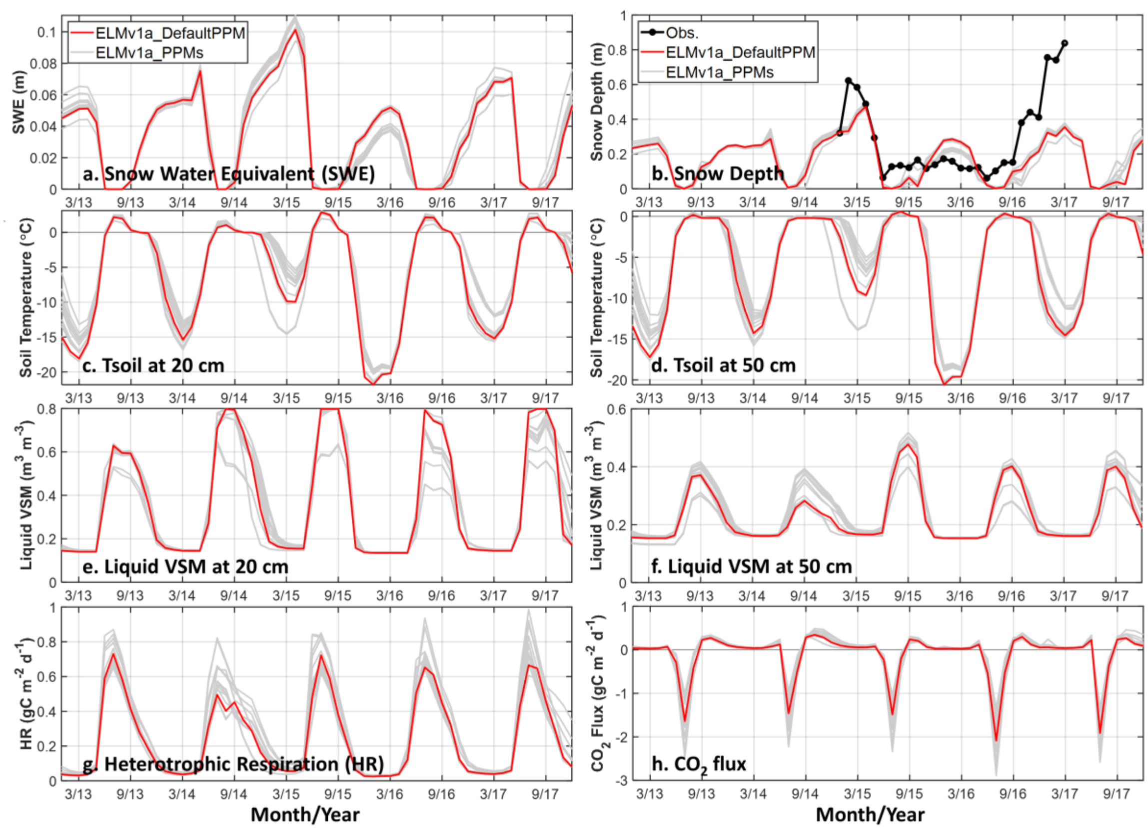


**Figure R1.1:** (Figure S9 in the revised manuscript) Sensitivity of ELM simulated (a) SWE, (b) snow depth, (c and d) soil temperatures, (e and f) liquid volumetric soil moisture (VSM), (g) heterotrophic respiration (HR), and (h) $CO_2$ flux at ATQ to precipitation-phase partition methods (PPMs). Suspicious snow depth measurements appear during summertime (black line; b). Red lines indicate the simulations with the ELM default PPM, and grey lines are simulations with different

PPMs as tested by Jennings and Molotch (2019). Different PPMs result in large discrepancies in the snowfall portion of total precipitation, leading to considerable differences in simulated SWE (a) and snow depth (b). The sensitivity of soil temperatures to snow depth and SWE is affected by 1) saturation of snow thermal insulation capacity (i.e., the levels of snow thermal insulation will not increase with snow depth if it exceeds an effective snow depth) (Slater et al., 2017); 2) rainfall fraction of total precipitation, which will not only influence snow compaction and accumulation process (thus snow depth) and snow mass (thus SWE), but also severely impacts soil water contents; 3) snow coverage, which impacts surface albedo and absorbed solar radiation, outgoing longwave radiation, and thus net radiation; 4) snow thermal conductivity schemes; and 5) active layer thickness and bottom boundary conditions of soil temperature, which resulted from lumped impacts of snow-or-rain partition over the long-term period (1901 to 2017 here). Also, the maximum supercooled liquid water content in frozen soils (e and f) does not dramatically decline with decreases in soil temperature if it drops to a certain level, and thus the substrate availability to microbial respiration does not accordingly decrease either. Thus, the integrated impacts of precipitation-phase partition methods and simulated snow depth on winter $CO_2$ emissions are smaller than that on warm-season $CO_2$ flux.

3) Justification of the phase change modification. While I have no problem that the main modification to the phase change calculation (allowing temperature to fall below zero during the phase change) is physically sound, I am not so convinced by the phase change "efficiency" parameter that was introduced. This is referenced to some papers where such a parameter was included in a model previously, but those papers are extremely brief in the justification of this and there is no reference to some observation or physical theory. Additional justification is therefore required for this "efficiency" parameter (or removal of this parameter if it is not fully justified).

**R1C3:** Currently, we employed this phase-change efficiency to effectively solve ELM's problem of overestimating phase-change rates while maintaining the current modeling structure as much as possible. In such a manner, the same modification can be easily adapted by other studies using ELM or CLM. We agree with the reviewer that the reference papers for the phase-change efficiency are extremely brief in their justification. Below we provide detailed justification for the phase-change efficiency and the virtual soil temperature; we have also reworked our appendices and added the justification to the revised Appendix A.

The heat (conduction) transfer equation (Eq. R1.1) can be rewritten as (Eq. R1.2) with soil-water freezing phase change,

$$c\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\left(\lambda\frac{\partial T}{\partial z}\right), \tag{Eq. R1.1}$$

$$c\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\left(\lambda\frac{\partial T}{\partial z}\right) - L_f\rho_{liq}\frac{\partial\theta_{liq}}{\partial t}, \tag{Eq. R1.2}$$

where $T$ is soil temperature (K), $c$ is the volumetric soil heat capacity (J m$^{-3}$ K$^{-1}$), $\lambda$ is soil thermal conductivity (W m$^{-1}$ K$^{-1}$), $z$ is the soil depth (m) of the model soil layers, $L_f$ is the latent heat of fusion (J kg$^{-1}$), $\theta_{liq}$ is soil liquid water content (m$^3$ m$^{-3}$), and $\rho_{liq}$ is the density of liquid water (kg m$^{-3}$). Instead of $-L_f\rho_{liq}\frac{\partial\theta_{liq}}{\partial t}$, We can also use the increasing rate of ice content ($\theta_{ice}$) with opposite sign ($+ L_f\rho_{ice}\frac{\partial\theta_{ice}}{\partial t}$) as in the manuscript. Here, we keep $\frac{\partial\theta_{liq}}{\partial t}$ for the discussion, and rewrite (Eq. R1.2) as,

$$\left(c + L_f\rho_{liq}\frac{\partial\theta_{liq}}{\partial T}\right)\frac{\partial T}{\partial t} = \frac{\partial}{\partial z}\left(\lambda\frac{\partial T}{\partial z}\right), \tag{Eq. R1.3}$$

By introducing an apparent heat capacity $c_{app} = c + L_f\rho_{liq}\frac{\partial\theta_{liq}}{\partial T}$, (Eq. R1.3) can be rewritten as,

$$c_{app}\frac{\partial T}{\partial t}=\frac{\partial}{\partial z}\left(\lambda\frac{\partial T}{\partial z}\right), \quad \text{(Eq. R1.4)}$$

which is the same as the heat (conduction) transfer equation (Eq. R1.1) with the actual soil heat capacity $c$ replaced with the apparent heat capacity $c_{app}$. That is, the apparent heat capacity incorporates the latent heat released by soil water freezing into the actual soil heat capacity.

To solve (Eq. R1.4), we need to compute the derivative of the soil freezing characteristic curve ($\theta_{liq}(T)$) with respect to temperature ($\frac{\partial\theta_{liq}}{\partial T}$). Here, we approximate the $\theta_{liq}(T)$ curve by combining the freezing point temperature-depression equation (Eq. R1.5) (Fuchs et al., 1978) and the soil water retention curve (Eq. R1.6) (Clapp and Hornberger, 1978). This leads to the supercooled water formulation (Eq. R1.7) (also the Eq. A3 in our manuscript) (Niu and Yang, 2006), shown below,

$$\psi(T)=\frac{10^3L_f(T_f-T)}{gT}, \quad \text{(Eq. R1.5)}$$

$$\psi(\theta_{liq})=\psi_{sat}\left(\frac{\theta_{liq}}{\theta_{sat}}\right)^{-B}, \quad \text{(Eq. R1.6)}$$

$$\theta_{liq}(T)=\theta_{sat}\left[\frac{10^3L_f(T_f-T)}{gT\psi_{sat}}\right]^{-1/B}. \quad \text{(Eq. R1.7)}$$

Computing $\frac{\partial\theta_{liq}}{\partial T}$ requests the soil freezing curves $\theta_{liq}(T)$ to be continuous and differentiable for a range of temperatures during the freezing process (Kurylyk and Watanabe, 2013; Hansson et al., 2004). Here, we follow the existing ELM framework to implement Eq. R1.7 in a much simpler way. Specifically, the baseline ELM first solves the heat transfer equation (Eq. R1.1) without consideration of soil water phase change; then it estimates the energy involved ($H$) for adjusting the initially solved soil temperature to the freezing point ($T_f$), and then estimates the mass change needed ($H_m$) based on $H$. The model then readjusts the soil liquid water, ice content, and soil temperature according to the actual available liquid water for freezing given the maximum supercooled liquid water allowed under current soil temperature ($\theta_{liq}(T)$, see example in Figure R1.2). The numerical representation for readjusted soil temperature is given by (Eq. R1.8) (or Eq. A5 in the manuscript),

$$T_i^{n+1*}=T_f+\frac{\Delta t}{c_i\Delta z_i}H_{i^*}=T_f+\frac{L_f(w_{ice,i}^{n+1}-w_{ice,i}^{n})}{c_i\Delta z_i}. \quad \text{(Eq. R1.8)}$$

This uncoupled two-step implementation significantly overestimates soil water freezing rates. Two reasons are responsible for the overestimation. First, the freezing point ($T_f$ = 0°C) is used to determine the occurrence of soil water phase change under all conditions. To further freeze supercooled soil liquid water, as Figure R1.2 shows, the soil temperature has to be colder than the virtual soil temperature ($Tv$), as we introduced by (Eq. A8) in the manuscript. Second, due to the steep slope (especially close to $T_f$ = 0°C) as shown in Figure R1.2, the estimated ice mass increase (i.e., $w_{ice}^{n+1}-w_{ice}^{n}$ or $w_{liq}^{n}-w_{liq,max}^{n+1}$, see (Eq. A5) in the manuscript) most often exceeds the required mass change, i.e., $H_m=-c_i\frac{\Delta z_i}{L_f}(T_f-T_i^{n+1})$, and thus soil liquid water freezes quickly in a large chunk. Soon, the liquid water available to be frozen becomes too small to release sufficient latent heat to compensate for the required energy deficit ($T_f-T_i^{n+1}$).

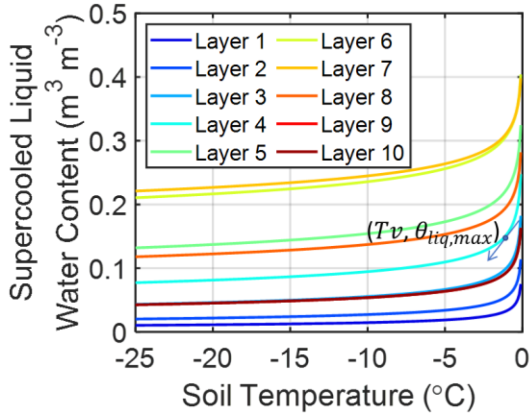


**Figure R1.2:** Maximum supercooled soil liquid water as a function of soil temperature at BEO.

Thus, through multiplying the initially estimated mass change ($H_m$) by the phase change efficiency ($\varepsilon$), we replace the freezing point in (Eq. R1.8) with an efficiency-weighted average of the initially solved soil temperature ($T_i^{n+1}$) and the freezing point,

$$T_i^{n+1*} = T_f + \frac{\Delta t}{c_i \Delta z_i}\left(-c_i \frac{\Delta z_i}{\Delta t} T_{inc}\varepsilon_i + \frac{L_f\left(w_{ice,i}^{n+1} - w_{ice,i}^{n}\right)}{\Delta t}\right)$$

$$= T_f - \left(T_f - T_i^{n+1}\right)\varepsilon_i + \frac{L_f\left(w_{ice,i}^{n+1} - w_{ice,i}^{n}\right)}{c_i \Delta z_i} \qquad \text{(Eq. R1.10)}$$

$$= (1-\varepsilon_i)T_f + \varepsilon_i T_i^{n+1} + \frac{L_f\left(w_{ice,i}^{n+1} - w_{ice,i}^{n}\right)}{c_i \Delta z_i}.$$

The two changes effectively improve the soil water freezing process simulations and prevent soil from becoming irreversibly too cold quickly as simulated by the baseline phase change scheme. We have reworked our Appendix and methodology section and better developed the elaboration for our modification. In the future, we will bypass this two-step implementation by solving (Eq. R1.4) in a coupled mode using the Apparent Heat Capacity Parameterization (AHCP) method (Kitover et al., 2016; Rawlins et al., 2013; Wang et al., 2010; Nicolsky et al., 2007; Hinzman et al., 1998; Endrizzi et al., 2014; Guymon et al., 1980; Mottaghy and Rath, 2006), along with our incorporation of advective heat transfer.

4) Environmental modifiers. The soil moisture function is modified in two ways: 1 to decrease respiration at high water contents, which replaces the oxygen-availability modifier, and 2 to continue to have respiration at zero water contents. Firstly, why would you replace the process-based oxygen availability with an empirical function that would presumably represent the process less well? And secondly, setting non-zero respiration at zero water contents is dubious and the functions chosen (shown in Fig S1) are strange-looking. There is no evidence that respiration occurs at zero water contents, rather that it continues over winter because there is non-zero liquid water in frozen soil. It would make much more sense to change the function so that it reaches zero at zero water contents, i.e. shift the curves to the left. This would also look a lot more like existing literature, e.g. Yan et al 2018. So, I am not convinced by these modifications.

**R1C4:** Thank you. We decided to follow the reviewer's suggestions. We now have removed our originally modified ELM moisture scalars, and proposed another way to improve the moisture scalars (Appendix B and Figure S1). Specifically, we simply changed the $\psi_{min}$ in $f_W$ to prevent zero respiration within the top ~50 cm during the cold season.

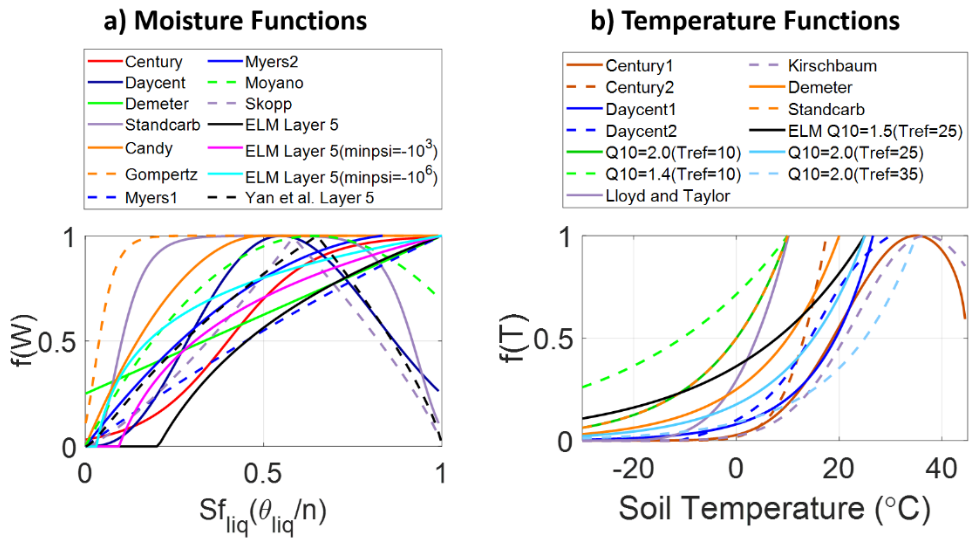


**Figure R1.3:** (New Figure S1 in the supplementary file) The moisture and temperature functions tested in this study are shown in (a) and (b), respectively. ELM's original moisture scalar predicts zero respiration in subfreezing soils (e.g., ELM layer 5, black; a) when soil liquid water content becomes small under frozen conditions. With revised minimum water potentials (magenta and cyan), ELM's f(w) would not drop to zero unless soil temperature becomes cold enough to have a very small supercooled liquid water content and a soil water potential less than the prescribed $\psi_{min}$, and thus preventing zero respiration within top ~50 cm soils during the cold season. Other functions have been tested and documented in Yan et al. (2018) and Sierra et al. (2015).

We then conducted more simulations testing the newly modified moisture scalars and other moisture functions, including the one proposed by Yan et al. (2018). Throughout the revised manuscript, we have updated all the results with the newly tested simulations. Our new Table S2 shows 22 moisture scalar and 37 temperature scalars, which composites 814 (=22x37) environmental modifiers (f(W)×f(T)×f(O)×f(D)) in total. Some moisture scalars already incorporate the oxygen stress to account for the inhibition of decomposition in the wet anoxic conditions, including Standcarb (Harmon and Domingo, 2001), Daycent (Kelly et al., 2000), Skopp (Skopp et al., 1990), and Moyano (Moyano et al., 2013), and (Yan et al., 2018). When using these moisture-oxygen scalars, the total environmental modifier function is f(W)×f(T)×f(D) without f(O). We do notice an instability problem of ELM-estimated oxygen stress, which, however, does not cause a very large problem because of the minimum threshold (0.2) imposed to f(O) (Appendix B; (Oleson et al., 2013)).

5) CH4 cold season transport. Again this is poorly justified and the equations are missing. In the appendix line 700-702 it states that "We integrate the emissions from ice cracks and remnants of aerenchyma tissues with (Eq. C14) by removing temperature limitation and applying a small Taere during winter time" Firstly, there was no mention of temperature limitation, so what does this part refer to? Secondly, what is "a small Taere"? (ie what is the value and why did you choose it?). It would really be useful to give the equation that you use in the model, instead of just this unclear description. In the methods it is justified by reference to a paper that there could be more conduction of methane through snow. However, the change made to the diffusion inside the soil is set as an arbitrary value and a sufficient justification is not given for changing it. Potentially, increasing Taere would increase the emissions enough (if you choose the right value) that this non-justified change to diffusion rate would not be needed.

**R1C5:** Thanks. In the baseline model, methane emissions through aerenchyma were turned off when the soil temperature is below 0°C. We have added detailed elaboration (including relevant equations, parameters, and description) regarding the methane transport mechanism into Appendix C. Also, the reviewer is correct about the changes to diffusion within frozen soils. We now have removed this modification in the newly tested simulations and updated our results.

Also, inspired by the reviewer, we now have conducted sensitivity experiments on methane parameters (including the parameter determining cold-season $T_{aere}$; see Table S3). Then, we now optimized the model simulations through two steps. Specifically, 1) we first evaluated the simulations using the (814) environmental modifiers to the base decomposition rate that assembled commonly used empirical soil temperature- and moisture-dependency functions (**R1C4**). These simulations used the newly modified methane model with the default parameters (Table S3). We selected the common decomposition schemes that provided satisfactory results of CO2 flux for all the sites (i.e., $NSE_{CO2}$> 0.5). 2) Then, we iteratively repeated the simulations with the common carbon decomposition schemes along with the seven methane parameterizations (Table S3). Among all these simulations ("NewPC_NewDecomNewCH4"; Table 2), we identified an optimal simulation for each site that has the smallest distance from ($NSE_{CH4}$, $NSE_{CO2}$) to (1, 1).

The updated Figure 7 (now is Figure 5 in the revised manuscript) is shown below.

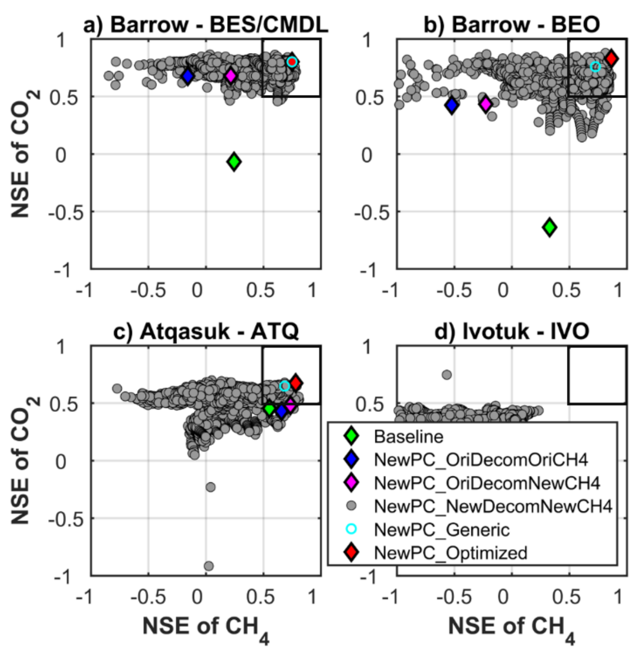


**Figure R1.3:** (New Figure 5 in the manuscript.) Scatter plot between the Nash–Sutcliffe Efficiency (NSE) of simulated monthly $CO_2$ and $CH_4$ emissions. The grey dots represent all the tested (1934) simulations. (see Table 2 for configurations for all the experiments.)

6) Missing out of IVO. It is shown that IVO cannot simulate reasonable methane emissions even with the correct temperature and moisture, however CO2 emissions did not suffer from this problem. Therefore I don't fully understand why CO2 from IVO cannot be included in the analysis. The optimisation would have to be done only on CO2 but it could still be optimised, is that right?

**R1C6:** Thanks. We have added back the $CO_2$ results at IVO (shown by the figures below). We also have changed the line styles and colors as suggested by the reviewers, and we have made all the figures consistent in plotting styles (see figures).

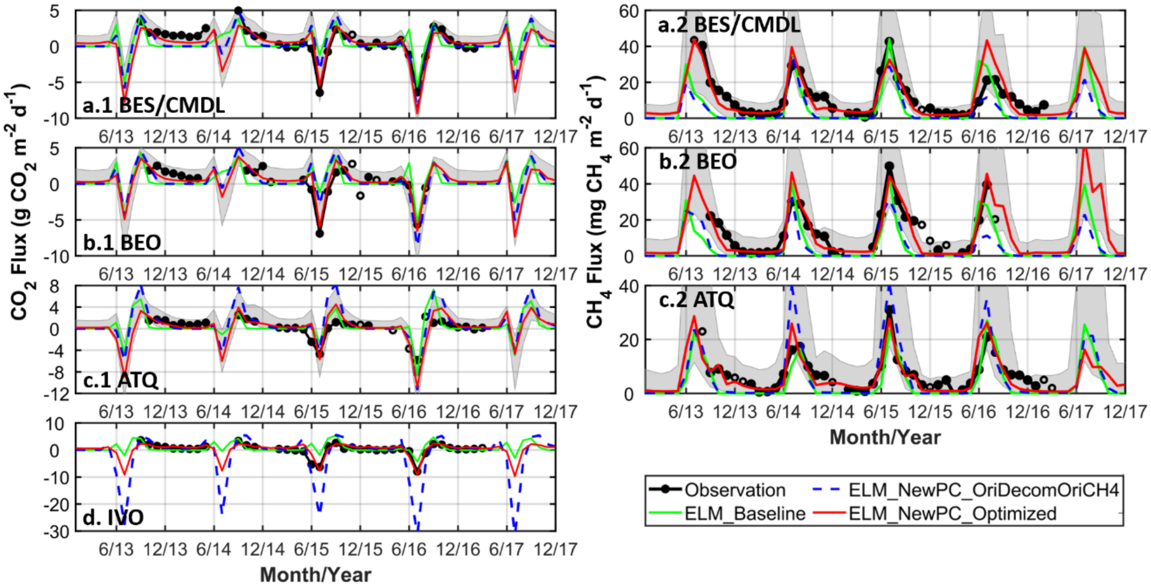


**Figure R1.4:** (New Figure 7 in the manuscript.) Monthly time-series comparison results now also include $CO_2$ results at IVO. For a better illustration, we now have moved all the intermediate simulation results into the supplementary file.

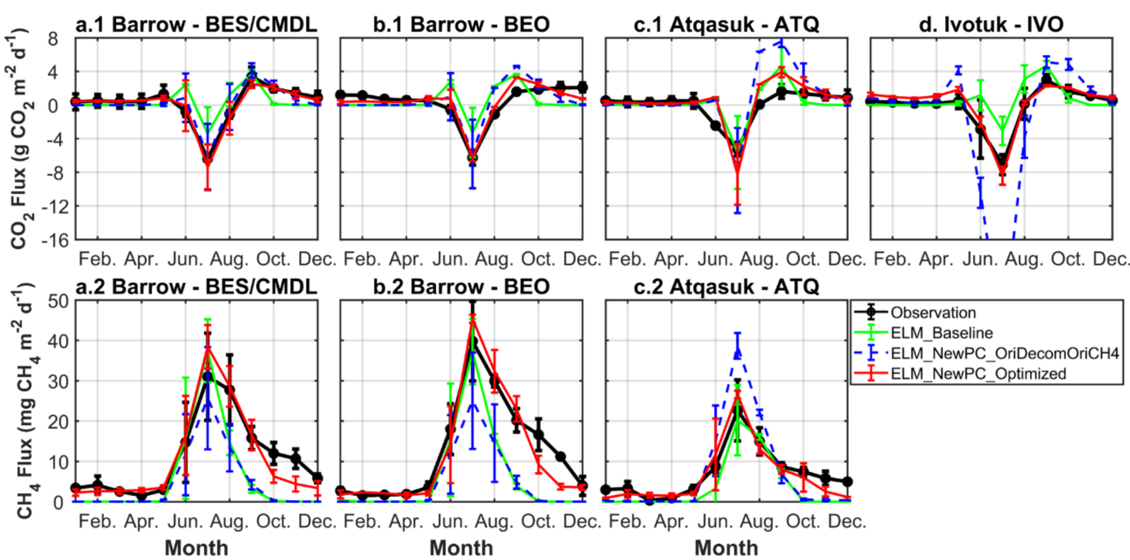


**Figure R1.5:** (New Figure 8 in the manuscript.) Comparison of multi-year (2013-2017) averaged monthly mean also with $CO_2$ results at IVO.

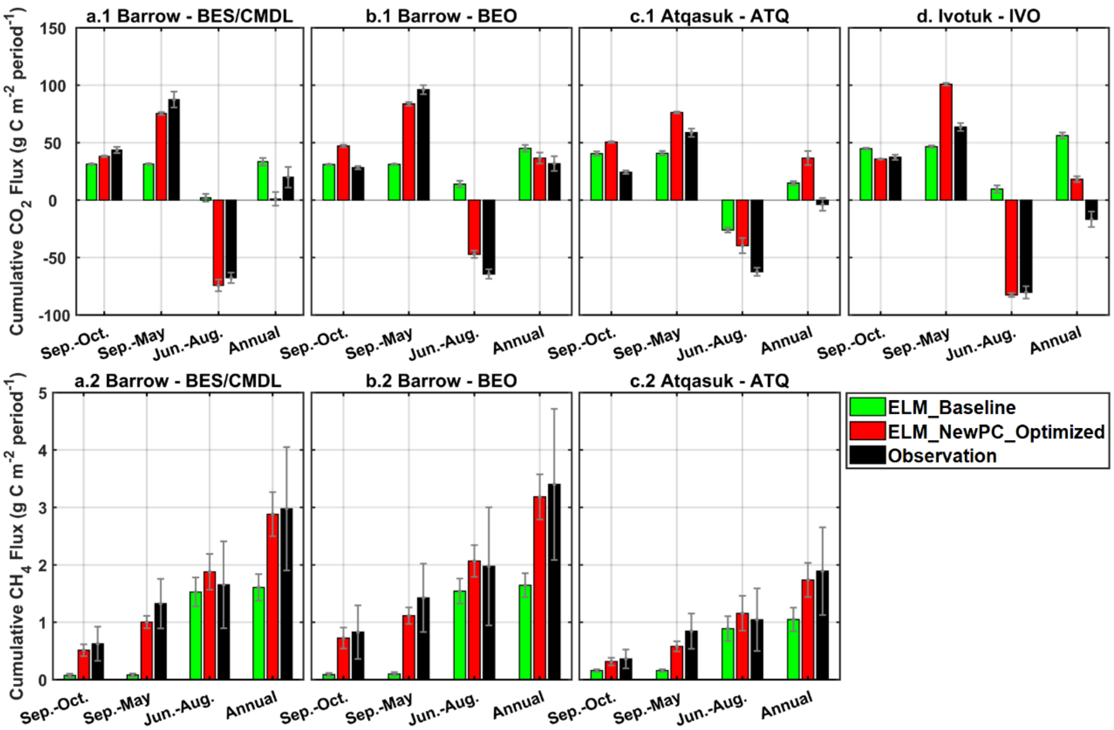


**Figure R1.6:** (New Figure 9 in the manuscript.) Multi-year (2013-2017) averaged cumulative $CH_4$ emissions and $CO_2$ net fluxes during the early cold season (Sep. and Oct.), cold-season period (Sep. to May), warm-season period (Jun. to Aug.), and the annual cycle (Sep. to Aug.) at our study sites.

7) Additional analysis: For further investigation of the temperature/moisture functions in frozen conditions, you could plot the emissions against temperature and moisture, instead of over time. Then you can see if the models and observations are producing similar functions. (This would show up, for example, if there is hysteresis in the observations which would make it difficult for any single function to capture the dynamics, and would be worthwhile to know.)

**R1C7:** As suggested by the reviewer, we had checked the emissions vs. temperature and moisture (included in the authors' comments uploaded earlier), and we found that results could not well capture the observed hysteresis effects of methane temperature response at IVO. In the future, we will apply a Macromolecular Rate Theory (MMRT)-based temperature sensitivity approach, which uses a quadratic relationship to approximate the $CH_4$ - temperature dependencies and thus can address the $CH_4$ hysteresis effect (Chang et al. 2020, 2021), which, however, is out of the scope of this study. We have added this plan into our revised manuscript.

> "In the future, we will apply a Macromolecular Rate Theory (MMRT)-based temperature sensitivity approach (Chang et al., 2020; 2021) to address the hysteresis effect on $CH_4$ emissions …" (Lines 423 - 425 in the revised manuscript)

**Specific comments**

=============

**Abstract**

Line 16-17 "simulated cold-season emissions at three tundra sites were improved by 84% and 81%" - it is not clear what metric the 84% and 81% refer to, is this the mean absolute error? Please specify.

**R1C8:** Thanks for pointing out this confusion. We have revised this sentence as below. (We now have mentioned "MAE" earlier in the abstract; also, the results have been updated, and thus all the relevant calculations have been updated accordingly.)

> "Furthermore, the MAEs of simulated cold-season emissions at three tundra sites were improved by 72% and 70% on average for $CH_4$ and $CO_2$, respectively." (Lines 16 - 17 in the revised manuscript)

Line 17-19 "...zero-curtain period in Arctic tundra, accounted for more than 50% of the total emissions" This statement is slightly misleading. This is the case in the model, but the study showed that this part was overestimated compared to the observations. I would therefore add something like "in the model, compared with around 45% (30-60%) in the observations"

**R1C9:** Thanks for the suggestion. We have revised this sentence as below.

> "Overall, $CH_4$ emitted during the early cold season (Sep. and Oct.), which often includes most of the zero-curtain period in Arctic tundra, accounted for more than 50% of the total emissions throughout the entire cold season (Sep. to May) in the model, compared with around 49.4% (43-58%) in observations." (Lines 17 - 19 in the revised manuscript)

**Introduction**

Lines 60-62 " However, current land models tend to significantly underestimate soil temperature during the cold season over permafrost regions (Dankers et al., 2011; Tao et al., 2017; Nicolsky et al., 2007; Yang et al., 2018b). One possible reason is that many land models fail to appropriately account for the latent heat released during soil water freezing" It is true that many land surface models did underestimate soil temperatures but, more recently, improved snow schemes have removed a lot of this problem. For example, your first reference Dankers et al (2011) has a followup study Burke et al (2013) which includes a multilayered snow scheme and removes the majority of the winter cold bias - although a small cold bias remains. I highly recommend adding some discussion of snow here to make it clear that the latent heat is not the only (or even the biggest) factor. Most recent LSM's (e.g. in CMIP6) do represent latent heat, if not particularly well, I suggest clarifying that to "One possible reason is that while many land models account for latent heat released during soil water freezing, they do not treat and distribute this heat appropriately"

**R1C10:** Thank you very much for the suggestion. We have changed the sentences as the reviewer suggested.

> "One possible reason is that while many land models account for latent heat released during soil water freezing, they do not treat and distribute this heat appropriately or/and do not simulate soil moisture correctly." (Lines 60 - 62 in the revised manuscript)

Lines 69 "many land models cannot accurately capture the ZCP length due to their underestimation of soil temperatures" This is not really an accurate statement. Many land models cannot accurately capture the ZCP length (true), but this is because they don't have enough soil moisture or an adequate representation of latent heat, not "due" to underestimated temperatures. Rather, underestimated temperatures can arise as a *result* of not simulating the ZCP.

**R1C11:** We agree with the reviewer on this point and have modified the sentence as below.

> "Nevertheless, many land models cannot accurately capture the ZCP length due to inaccurately simulated soil moisture and/or inadequate representation of latent heat, thus underestimating soil temperature and cold-season $CO_2$ emissions (Commane et al., 2017) and $CH_4$ (Zona et al., 2016)." (Lines 68 - 70 in the revised manuscript)

**Study sites and data**

Line 93. " CARVE CO2 measurements were not available;" should this be "...were not available from 2015-2017;" ? Currently this part is unclear.

**R1C12:** To accurately clarify this, we have revised the sentence as below.

> "The CARVE $CO_2$ measurements were not available at the data archive we used here; therefore, monthly winter-time $CO_2$ flux data at the same towers assembled by Natali et al. (2019) are included to complement $CO_2$ observations from 2013 to 2014." (Lines 93 - 95 in the revised manuscript)

Line 116-117 "Due to the discontinuity of observed soil moisture over time and along with the vertical depth, evaluating ELMv1-ECA simulated liquid water content at layer node-depth was limited." This sentence does not make sense to me, please clarify.

**R1C13:** We have modified the sentence as below.

> "The observed soil moisture is only available at two or three depths that are quite different from model layer node-depths, and also show discontinuities in time. Thus, evaluating ELMv1-ECA simulated liquid water content was limited." (Lines 116 - 117 in the revised manuscript)

**Methodology**

Line 152-153 "The underlying assumption here is that the liquid water of soil resists freezing as the freezing process proceeds and Sfnliq,i decreases, analogous to how dry soils resist getting drier due to capillary force." This is the explanation given for the efficiency factor (see comment 3, above): However, this capillary force in freezing soils is represented by the non-zero liquid water contents at sub-freezing temperatures, and it is not clear to me that it needs an additional factor. The efficiency factor, I guess (although it is not clear what it actually does - see next comment) corresponds to a 'loss' of some of the energy produced by latent heat. It does not make sense that energy would just disappear. Please explain/justify.

**R1C14:** Thanks. Please also see our response R1C3. We have removed this sentence as we incorporate our description in R1C3 into the manuscript.

Line 153-154 "We applied the phase change efficiency to the initially estimated energy and mass change involved, i.e., $H_i$ and $H_m$ (see (Eq. A4) in the Appendix)". It is not clear what 'applied' means here, did you multiply some part of these equations by the efficiency factor? The easiest thing to do would be to include the equations in the Appendix that you actually used (i.e. rewrite those equations with the efficiency factor included, instead of leaving it to our imagination).

**R1C15:** We multiply the initially estimated energy and mass change involved (see (Eq. A4)) by the phase change efficiency (see (Eq. A9)). We now have modified Appendix A by incorporating all the equations related to the changes. Please also see our response **R1C3**.

Table S2: I am missing where the moisture functions ModifiedELMâˇARˇ S1 ModifiedELMâˇARˇ S2, etc are documented?

**R1C16:** We now have removed the originally modified ELM moisture scalars, and thus this comment does not apply anymore. Please also see our response **R1C4**.

Line 264-266 "We confirmed that ELMv1-ECA's PFT dataset was a good compromise between representing the site-scale ecosystem and other global parameters and surface datasets within ELM. " Firstly, what is ELMv1-ECA's PFT dataset? This is not mentioned. Secondly, how did you assess whether it was 'good'? I recommend adding more information here.

**R1C17:** The ELMv1-ECA's default plant type function (PFT) dataset was derived from satellite-based data by Lawrence et al. (2007), and the model's global surface datasets have been calibrated against satellite observations (Oleson et al., 2013). Results show that the simulated GPP generally agrees with satellite-based products (Figure R1.7). In addition to PFTs, many other surface datasets (including other vegetation-related parameters, soil colors, soil organic matter content, soil sand and clay percent, topography, slope, and inundation-related parameters, etc.) also need to be calibrated for an ideal site-level simulation. In the future, we plan to embed an advanced calibration tool into the ELM framework to enhance its capability and efficiency of multi-objective optimizations. Here, to address the reviewer's comment, we have added sentences below to the revised manuscript.

> "The ELMv1-ECA's default plant type function (PFT) dataset was derived from satellite-based data by Lawrence et al. (2007). We analyzed the vegetation composition from the closet survey plot to the flux tower and examined the rationality of ELMv1-ECA's percentage of PFT for the site-scale simulation through testing different PFT datasets derived from this vegetation survey (Davidson and Zona, 2018). We found that these PFT datasets generally are not superior to the original PFT dataset, which generally reproduced satellite-based GPP (Figure S4). We thus confirmed that ELMv1-ECA's PFT dataset was a good compromise between representing the site-scale ecosystem and other global parameters and surface datasets within the model." (Lines 263 - 269 in the revised manuscript)

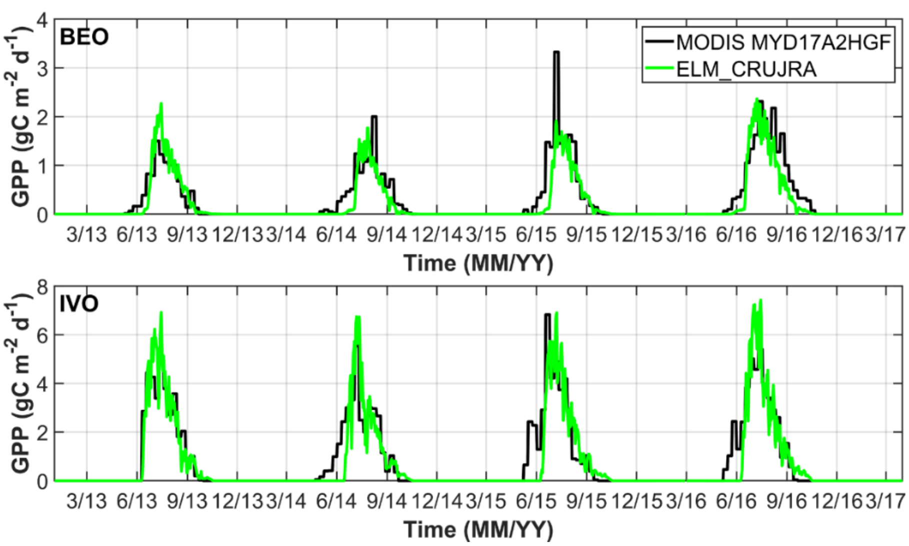


**Figure R1.7:** (Figure S4 in the revised supplementary file) Comparison between ELMv1-ECA simulated daily GPP and MODIS GPP product (500m 8-day MYD17A2HGF) at two of our sites.

Line 266-267 "The simulated saturated and unsaturated CH4 emissions were weighted with the estimated inundation fractions at the footprint of ABoVE eddy-covariance flux towers" Surely the *unsaturated CH4 emissions should be weighted with the *noninundated fraction in the footprint? I guess this is probably what you did, it's just not written very clearly, it currently sounds like both saturated and unsaturated CH4 emissions were multiplied by the inundated fraction.

Readers will not know that the model simulates methane separately from saturated and unsaturated grid cell fractions, therefore I suggest making that point here.

**R1C18:** Thanks for pointing this out. We have modified the sentence as below.

> "The surface $CH_4$ emission is a weighted average of simulated saturated and unsaturated components using predicted inundation and non-inundation fractions. To compare simulated $CH_4$ emissions with ABoVE measurements at the site scale, we use the estimated inundation fractions at the footprint of ABoVE eddy-covariance flux towers (see details in (Xu et al., 2016))." (Lines 269 - 272 in the revised manuscript)

**Results**

Line 326-343 Here you talk about the improvement to the ZCP. Looking at the plots, there is a great improvement in deeper soil layers but not so much in the surface (for 3/4 sites). I suggest that the text should recognise this fact about the surface being less well simulated.

**R1C19:** We have added an additional sentence at the end of this paragraph.

> "In general, the improvements in ZCP are larger in deeper layers than topsoils, with the top layer showing only marginal improvement." (Lines 364 - 365 in the revised manuscript)

Line 346 Reference to Figure 3 should be Figure 4.

**R1C20:** Thanks. We have changed this. (Now, Figure 4 becomes Figure 3 in the revised manuscript due to the removal of Figure 2.)

Line 357-358 "The deeper active layer simulated by NewPC implies more soil water storage capacity, resulting in lower soil moisture in shallow soil layers and higher soil water in deep layers" This pattern is not really seen with most of the sites, either the new simulation seems to have lower soil moisture in general, or in the case of IVO it is greater or similar in almost every layer in the new simulation. There is also the claim of soil moisture being improved - this is true because the timing of thaw and freezeup is better, but actually the level of saturation in general seems to now be lower and in several cases the old scheme was better in that regard. This is just a suggestion, but I am aware of more than one land surface scheme that has found their scheme of dealing with saturation of soil moisture leads to water being forced out of the top of the soil during the freeze-up period. I was just wondering if simulating a longer ZCP might lead to more water being lost in this way, and would therefore explain why the new model is drier. There are several possibilities, of course!

**R1C21:** The "claim of soil moisture being improved" was made based on the RMSE of simulated soil liquid water content (Table S5). The pattern (i.e., lower soil moisture in shallow soil layers and higher soil water in deep layers) is shown in Figure 3 (Figure 2 in the revised manuscript) by the magenta vs. green lines during summertime when the active layers reach the deepest thaw depths.

Our model should be able to simulate the mechanism of "water being forced out of the top of the soil during the freeze-up period", given the differences in porosity and ice volume. But at the tested sites, soils usually either are unsaturated or get drained before the freeze-up period and thus we did not find that mechanism to be the case here.

Line 383-384. As I discussed above, the snow is important and I suggest that this is the place to present some additional analysis rather than simply referring to "underestimated snow depth (not shown)".

**R1C22:** Thanks. We have added the figure included in our response **R1C2** into the supplementary file and add the relevant discussion. Please also see our response to Reviewer #2 regarding this point, i.e., **R2C1.**

Line 416, having checked that using observed soil moisture and temperature does not improve the CH4 simulation, saying that including advective heat transport would likely improve the simulations is surely incorrect, since this would just improve the soil temperature, which you found did not help. I would also be surprised if a better wetland simulation would help if using observed soil moisture did not improve the simulation. Geological seepage is certainly a possibility though.

**R1C23:** Due to discontinuities in soil moisture observations, as we have discussed, imposing soil moisture on the model is quite tricky. Unless through a well-designed data assimilation method, the control experiments still cannot correctly simulate surface standing water and wetland inundation dynamics. We now have removed this sentence from the manuscript.

In addition, including advective heat transport would not only improve simulated soil temperature, it would also impact active layer thickness, the impermeable bottom boundary (frozen) layer for water fluxes, and thus soil moisture redistributions and microbial activities. Especially during the thawing season, advective heat transport of warm spring rain could trigger early thawing onset, speed-up thawing rate, and then stimulate $CH_4$ emissions (Neumann et al., 2019), thus controlling the seasonal cycle of $CH_4$ emissions. However, we do agree with the reviewer that the advective heat transport might play a lesser role here, but we feel this is still an important factor. We now have added the sentences below.

> "The failure of simulated $CH_4$ emissions to capture the methane seasonality at IVO (as indicated by Figures S8) might occur because of the lack of 1) a reasonable wetland module that can adequately account for inundated hydro-ecological dynamics, 2) advective heat transport at the air-ground interface through rainfall infiltration and within subsurface soils through water transfer, and 3) the geological micro-seepage emission of $CH_4$, as reported in previous studies (Anthony et al., 2012; Etiope and Klusman, 2010; Russell et al., 2020). For instance, Lyman et al. (2020) showed large temporal variability of $CH_4$ at natural gas well pad soils, similar to the observations at IVO (Anthony et al., 2012). The advective heat transport not only impacts soil temperature but also affects soil moisture redistribution, substrate availability, and microbial activity; this mechanism and wetland inundation dynamics together would cause hysteretic effects on $CH_4$ emission response to soil temperatures (Chang et al., 2020; 2021). In the future, we will apply a Macromolecular Rate Theory (MMRT)-based temperature sensitivity approach (Chang et al., 2020; 2021) to address the hysteresis effect on $CH_4$ emissions and explore more on the contribution of geological micro-seepage emission."(Lines 415 - 425 in the revised manuscript)

Line 420-423. This is missing the information that the performance of CH4 is degraded at the BES/CMDL and BEO sites. It's somewhat misleading to only mention the improvements.

**R1C24:** We do not think the degraded $CH_4$ performance here matters very much since we only changed the phase change scheme but kept the carbon decomposition and methane modules the same as the baseline which we know are problematic. Improved ZCP durations and soil temperature and moisture do not necessarily guarantee improvements in $CO_2$ and $CH_4$ fluxes, given the problematic carbon decomposition and methane modules. We feel our discussion here is not misleading.

Line 442-444. This part is unclear. When you say "soil properties", do you mean soil thermal/hydraulic properties? And is the improvement of the ELMv1-ECA's moisture scalars due to the function being based on the suction rather than the volumetric soil moisture content? Please clarify this. Also, please give evidence that ELMv1-ECA "reasonably explained the varying influence along the vertical soil profile".

**R1C25:** Soil properties here mean the soil texture-dependent porosity and saturated soil water potential, and yes, the reviewer is correct that ELMv1-ECA uses soil suction rather than soil moisture content for moisture scalars. Over high-latitude peatlands, soil porosity varies with depth due to vertically stratified organic matter contents, which leads to large differences in soil pore spaces and substrate affinity. This is a key strength of ELMv1-ECA's moisture scalars that differs from other models' moisture scalars. The Fig. 1 in Niu and Yang (2006) is good evidence supporting that ELMv1-ECA "reasonably explained the varying influence along the vertical soil profile". Besides ELMv1-ECA's moisture scalars, we also implemented the moisture function of Yan et al. (2018) for each soil layer using the soil properties of each layer. We have revised the sentence as below.

> "For the Sierra et al. (2015) empirical moisture functions, the influence of liquid moisture content on heterotrophic respiration is uniformly applied to all active soil layers, even though the soil properties (e.g., porosity and saturated soil water potential) are quite different vertically. ELMv1-ECA's moisture scalars (including the original scheme) that use soil water potential of each layer, in contrast, reasonably explained the varying influence along with the vertical soil profile (Niu and Yang, 2006). " (Lines 448 - 452 in the revised manuscript)

Line 444 "Thus, the simulations..." I suggest removing 'thus' because overestimation isn't implied from the previous sentence.

**R1C26:** Thanks. We have removed "thus" here.

Line 450 "assigns small thresholds for the moisture scalar" this is unclear. Did you mean "assigns small minimum values for the moisture scalar"?
Line 451: Same problem as 450.

**R1C27:** We now have removed the modified ELM moisture scalars. Please also see our response **R1C4**.

Line 457-459 ", at ATQ, where cold-season temperatures are relatively warmer than at BES/CMDL and BEO, simulations with the original ELMv1-ECA environmental modifier (i.e., "NewPC_OriDecom_NewCH4"; discussed in Section 3.1.2), already released much more CO2 and CH4 throughout the cold season than in the baseline simulations, " Can you add a reference to a Table or Figure that shows this happens more at ATQ than the other sites? It isn't very clear to me on Figure 8. And in fact on Figure 6 (c3 and c4), it looks like the cold season production of CO2 and CH4 still goes to zero at ATQ with NewPC_OriDecom_NewCH4.

**R1C28:** We have added reference to this statement, i.e., “NewPC_OriDecomNewCH4” in Figure S8. The NewPC_OriDecomNewCH4 simulated $CO_2$ and $CH_4$ still go to zero after the ZCP ends because the original decomposition scheme still predicted zero respiration under frozen conditions, but the accumulated cold-season emissions are still larger than the baseline results because with the improved phase-change scheme it has a much longer ZCP compared to the baseline. Note, for a better illustration, we now did not include the intermediate simulations in the figures and moved all the intermediate results into our supplementary file (Figure S8).

Line 471. vary -> varies

**R1C29:** We have changed it as suggested.

Line 471-474 " For cold sites (i.e., BES/CMDL and BEO), the sensitivity of simulated CH4 to Q10 values is larger than the sensitivity of CO2 net flux to Q10 because cold temperature suppresses vegetation growth (i.e., CO2 uptake); while for the warm site (i.e., ATQ), both CH4 and CO2 net flux are very sensitive to the Q10 values." Can you refer here to some numbers/figures that show this? I see that in Table S4 this is apparent if you compare the lines with the same soil moisture function but different Q10’s.

**R1C30:** Due to the removal of the originally modified moisture scalar, the related discussion has also been removed, and thus the comments here do not apply anymore. Please see our discussion regarding the new results in the revised manuscript.

Line 485/Section 4.3. I suggest you start this section with a clarification that throughout this section you are analysing the results with optimal decomposition scheme for each site (and therefore different parameters are used for each site).

**R1C31:** Thanks for the suggestion. We have added a sentence as below.

> “Throughout this section, we only retain and discuss the identified optimal simulation results (i.e., ELM_NewPC_Optimized) for each site.” (Lines 501 - 502 in the revised manuscript)

Line 496 slightly -> slight

**R1C32:** Thanks. We have revised this.

Line 511-512 "We find that the simulated cold-season CO2 emissions were larger than the warm-season CO2 net uptake at all three sites" Please specify during which time period. (Presumably they are in balance during the spinup, but will become out of balance later in the simulation due to changing climate, so it makes sense to note the time period here)

**R1C33:** Thanks. We have modified the sentence as below.

> “We find that the simulated cold-season $CO_2$ emissions were larger than the warm-season $CO_2$ net uptake during the analyzing period (2013-2017) at all four sites (Figure 8).” (Lines 534 - 535 in the revised manuscript)

**Summary**

I suggest that you additionally mention the potential issues of using the heterotrophic respiration to estimate CH4 production. For example, this means that CH4 emissions may drop as the soil becomes more saturated (once soil moisture passes the optimum), whereas in fact the highest CH4 emissions should be in saturated conditions.

**R1C34:** Thanks. We have added the discussion in the summary section as suggested.

> "Moreover, due to lacking representations of wetland hydro-ecological dynamics, the model uses simulated upland heterotrophic respiration to estimate $CH_4$ production (Riley et al., 2011), which might cause underestimations of $CH_4$ emissions especially under wet conditions." (Lines 594 - 596 in the revised manuscript)

Line 529 "by updating upper boundary resistance" Was this the only change? What about the change you made to the diffusion through the soil? I don't think that is related to the upper boundary? Please check this to make sure it's summarizing accurately.

**R1C35:** We apologize for missing information here. We have revised the sentence as below.

> "We further refined the cold-season methane processes by mimicking emission pathways through ice cracks and remnants of aerenchyma tissues, reducing the maximum oxidation rate constant, and reducing upper boundary (snow) resistance that allows CH4 to be emitted from frozen soils through snow to the atmosphere" (Lines 556 - 559 in the revised manuscript)

Line 546 "the identified an" -> "the identified"

**R1C36:** Thanks. We have revised this as suggested.

Line 562 add "due to microbial dynamics" or similar, for clarity

**R1C37:** Thanks. We have revised the sentence as below.

> "In addition, the single static multiplicative function used to parameterize the impact of environmental conditions on respiration might not be appropriate, because the environmental impact also depends on maximum respiration rate, soil texture, soil carbon content and quality, and microbial biomass (Tang and Riley, 2019)." (Lines 591 - 593 in the revised manuscript)

Line 571-573 "The increasing rate of cold-season heterotrophic respiration (releasing CO2) may become larger than the trend of warm-season vegetation CO2 uptake under future climate" In fact in your simulations, the cold season respiration already became larger than the warm season CO2 uptake by 2017, is that right? This point could be made stronger with that information.

**R1C38:** Thanks. Although our simulations show net $CO_2$ sources, the observed net annual $CO_2$ budgets at ATQ and IVO still indicate net uptake (see Figure 8) despite the discontinuity.

Appendix Eq. A3. What does * mean in this equation? Eq. A7. as already discussed in my comments on the Methodology, you need to show where/how these factors are applied in the model - via equations would be easiest.

**R1C39:** The * in Eq. A3 (and A5) is a typo, and we have removed it. The variables with * mean the ultimately adjusted variables (e.g., $T_i^{n+1*}$). We have also reworked our Appendix to make everything consistent and include all the relevant equations. Please also see our response in **R1C3**.

Line 642. This equation isn't entirely consistent with equation A3, the 10ˆ3 is on the bottom and g is missing. Can you check both of these?

**R1C40:** That is because of different units used for soil water potential (mm or MPa), but both equations are correct. We have made the equations consistent in the revised manuscript.

Line 692. It is not entirely clear what A(z) represents, is this the total methane emission to the atmosphere or just the part from aerenchyma?

**R1C41:** It is the $CH_4$ transport via aerenchyma from layer z. We have revised the sentence as below.

> "Vascular plant aerenchyma tissues serve as diffusive pathways to transport $CH_4$ from soil to the atmosphere. The $CH_4$ transport via aerenchyma from soil layer $z$ ($A(z)$, mol $m^{-2}$ $s^{-1}$) is calculated as …" (Lines 749 - 750 in the revised manuscript)

Line 700: Apologies if this is common knowledge but I don't know what "amount of carbon per tiller" means. Is this correct?

**R1C42:** This is described in Wania et al. (2010), and we have added this reference here. For clarification here, we have revised the sentence as below.

> "and the factor 0.22 represents average observed tiller biomass (gC per tiller) (Wania et al., 2010; Schimel, 1995)." …" (Lines 756 - 757 in the revised manuscript)

Line 700-702: This needs more explanation/equations, see comment (5) in general comments, above.

**R1C43:** We have revised this as suggested. Please also see our revised Appendix C.

Line 710: "Table 2" should be Table 1, I think

**R1C44:** Yes, thanks. We have changed it.

Line 710: Please specify which parameter in the equations you are changing. The table refers to it as "scale_factor_gasdiff_snow" and it's not clear where this fits in Eq. C13 (if it all)

**R1C45:** This modification was made to the upper boundary layer resistance, and we have also added the relevant equations in the Appendix C.

Line 711: presents -> is present.

**R1C46:** Thanks. We have revised this as suggested.

**Figures**

General: Firstly, it is common to plot the observations in black and the model versions in colours (or at least use a different style of line), which I would recommend here since it would add clarity to the plots. Secondly, there appears to be a slight difference in the CO2 when the methane transport modifications are introduced, particularly for ATQ. I did not see any way that the methane transport would influence the CO2 simulation - could you explain this difference?

**R1C47:** We have replotted our figures according to reviewers' suggestions, i.e., using black for observations and other colors for simulations with different line styles. An example is shown in Figures R1.4 and R1.5.

Regarding the second point, modifications in $CH_4$ transport will impact $CH_4$ concentration in soils and thus influencing $CH_4$ oxidation rate (see details in Riley et al. (2011)), which will further impact $CO_2$ emissions.

Figure 1 is not super clear which labels are refering to which sites, since there are more labels than red dots. Could you add lines or arrows to indicate for certain which site is in which location.

**R1C48:** We have modified this figure as suggested and have replaced the original Figure 1 with the new one.

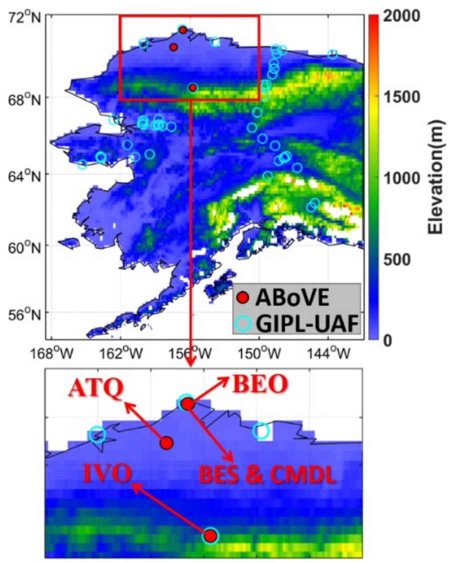


**Figure R1.8:** New Figure 1 with lines and arrows.

Figure 2 is a bit of a mess and it does not seem to be logically organised. For example, the split between green circles appears to be between methane and "every other form of carbon", perhaps it would make sense to separate vegetation and soil (non root) carbon? Most of the arrows are brown and seem to represent "some kind of influence". To me it is important to show the flows of carbon between the different spheres, some of which is shown in black (e.g. CO2 emission), some in brown (e.g. heterotrophic respiration producing CH4), and some not shown at all, such as the flow of carbon between plants and soil. This diagram needs to be revisiting to get a complete and coherent presentation.

**R1C49:** After careful consideration, we decided to remove Figure 2.

Papers mentioned


Burke et al 2013 https://link.springer.com/article/10.1007%252Fs00382-012-1648-x
Burke et al 2020 https://tc.copernicus.org/articles/14/3155/2020/
Brooks et al 2004 https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2486.2004.00877.x
Strom et al 2012 https://www.sciencedirect.com/science/article/abs/pii/S0038071711003385
Yan et al 2018 https://www.nature.com/articles/s41467-018-04971-6

## References


Anthony, K. M. W., Anthony, P., Grosse, G., and Chanton, J.: Geologic methane seeps along boundaries of Arctic permafrost thaw and melting glaciers, Nat Geosci, 5, 419-426, 2012.

Burrows, S. M., Maltrud, M., Yang, X., Zhu, Q., Jeffery, N., Shi, X., Ricciuto, D., Wang, S., Bisht, G., Tang, J., Wolfe, J., Harrop, B. E., Singh, B., Brent, L., Baldwin, S., Zhou, T., Cameron-Smith, P., Keen, N., Collier, N., Xu, M., Hunke, E. C., Elliott, S. M., Turner, A. K., Li, H., Wang, H., Golaz, J. C., Bond-Lamberty, B., Hoffman, F. M., Riley, W. J., Thornton, P. E., Calvin, K., and Leung, L. R.: The DOE E3SM v1.1 Biogeochemistry Configuration: Description and Simulated Ecosystem-Climate Responses to Historical Changes in Forcing, J Adv Model Earth Sy, 12, 2020.

Chang, K. Y., Riley, W. J., Crill, P. M., Grant, R. F., and Saleska, S. R.: Hysteretic temperature sensitivity of wetland CH4 fluxes explained by substrate availability and microbial activity, Biogeosciences, 17, 5849-5860, 2020.

Chang, K. Y., Riley, W. J., Knox, S. H., Jackson, R. B., and al., e.: Substantial hysteresis in emergent temperature sensitivity of global wetland CH4 emissions, Nature Communications, 2021.

Clapp, R. B., and Hornberger, G. M.: Empirical equations for some soil hydraulic properties, Water Resources Research, 14, 601-604, 1978.

Commane, R., Lindaas, J., Benmergui, J., Luus, K. A., Chang, R. Y. W., Daube, B. C., Euskirchen, E. S., Henderson, J. M., Karion, A., Miller, J. B., Miller, S. M., Parazoo, N. C., Randerson, J. T., Sweeney, C., Tans, P., Thoning, K., Veraverbeke, S., Miller, C. E., and Wofsy, S. C.: Carbon dioxide sources from Alaska driven by increasing early winter respiration from Arctic tundra, P Natl Acad Sci USA, 114, 5361-5366, 2017.

Davidson, S. J., and Zona, D.: Arctic Vegetation Plots in Flux Tower Footprints, North Slope, Alaska, 2014, ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1546, 2018.

Endrizzi, S., Gruber, S., Dall'Amico, M., and Rigon, R.: GEOtop 2.0: simulating the combined energy and water balance at and below the land surface accounting for soil freezing, snow cover and terrain effects, Geosci Model Dev, 7, 2831-2857, 2014.

Etiope, G., and Klusman, R. W.: Microseepage in drylands: Flux and implications in the global atmospheric source/sink budget of methane, Global and Planetary Change, 72, 265-274, 2010.

Fuchs, M., Campbell, G., and Papendick, R.: An analysis of sensible and latent heat flow in a partially frozen unsaturated soil, Soil Sci Soc Am J, 42, 379-385, 1978.

Grant, R. F., Mekonnen, Z. A., Riley, W. J., Arora, B., and Torn, M. S.: Mathematical Modelling of Arctic Polygonal Tundra with Ecosys: 2. Microtopography Determines How CO2 and CH4 Exchange Responds to Changes in Temperature and Precipitation, J Geophys Res-Biogeo, 122, 3174-3187, 2017a.

Grant, R. F., Mekonnen, Z. A., Riley, W. J., Wainwright, H. M., Graham, D., and Torn, M. S.: Mathematical Modelling of Arctic Polygonal Tundra with Ecosys: 1. Microtopography Determines How Active Layer Depths Respond to Changes in Temperature and Precipitation, J Geophys Res-Biogeo, 122, 3161-3173, 2017b.

Golaz, J. C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., Abeshu, G., Anantharaj, V., Asay-Davis, X. S., Bader, D. C., Baldwin, S. A., Bisht, G., Bogenschutz, P. A., Branstetter, M., Brunke, M. A., Brus, S. R., Burrows, S. M., Cameron-Smith, P. J., Donahue, A. S., Deakin, M., Easter, R. C., Evans, K. J., Feng, Y., Flanner, M., Foucar, J. G., Fyke, J. G., Griffin, B. M., Hannay, C., Harrop, B. E., Hoffman, M. J., Hunke, E. C., Jacob, R. L., Jacobsen, D. W., Jeffery, N., Jones, P. W., Keen, N. D., Klein, S. A., Larson, V. E., Leung, L. R., Li, H. Y., Lin, W. Y., Lipscomb, W. H., Ma, P. L., Mahajan, S., Maltrud, M. E., Mametjanov, A., McClean, J. L., McCoy, R. B., Neale, R. B., Price, S. F., Qian, Y., Rasch, P. J., Eyre, J. E. J. R., Riley, W. J., Ringler, T. D., Roberts, A. F., Roesler, E. L., Salinger, A. G., Shaheen, Z., Shi, X. Y., Singh, B., Tang, J. Y., Taylor, M. A., Thornton, P. E., Turner, A. K., Veneziani, M., Wan, H., Wang, H. L., Wang, S. L., Williams, D. N., Wolfram, P. J., Worley, P. H., Xie, S. C., Yang, Y., Yoon, J. H., Zelinka, M. D., Zender, C. S., Zeng, X. B., Zhang, C. Z., Zhang, K., Zhang, Y., Zheng, X., Zhou, T., and Zhu, Q.: The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution, J Adv Model Earth Sy, 11, 2089-2129, 2019.

Guymon, G. L., Hromadka, T. V., and Berg, R. L.: A One Dimensional Frost Heave Model Based Upon Simulation of Simultaneous Heat and Water Flux, Cold Reg Sci Technol, 3, 253-262, 1980.

Harmon, M., and Domingo, J.: A users guide to STANDCARB version 2.0: a model to simulate the carbon stores in forest stands, Dep. of For. Sci., Oreg. State Univ., Corvallis, OR, 2001.

Hansson, K., Simunek, J., Mizoguchi, M., Lundin, L. C., and van Genuchten, M. T.: Water flow and heat transport in frozen soil: Numerical solution and freeze-thaw applications, Vadose Zone J, 3, 693-704, 2004.

Hinzman, L. D., Goering, D. J., and Kane, D. L.: A distributed thermal model for calculating soil temperature profiles and depth of thaw in permafrost regions, Journal of Geophysical Research-Atmospheres, 103, 28975-28991, 1998.

Hugelius, G., Tarnocai, C., Broll, G., Canadell, J. G., Kuhry, P., and Swanson, D. K.: The Northern Circumpolar Soil Carbon Database: spatially distributed datasets of soil coverage and soil carbon storage in the northern permafrost regions, Earth Syst Sci Data, 5, 3-13, 10.5194/essd-5-3-2013, 2013.

Jennings, K. S., and Molotch, N. P.: The sensitivity of modeled snow accumulation and melt to precipitation phase methods across a climatic gradient, Hydrol Earth Syst Sc, 23, 3765-3786, 2019.

Kitover, D. C., van Balen, R. T., Vandenberghe, J., Roche, D. M., and Renssen, H.: LGM Permafrost Thickness and Extent in the Northern Hemisphere derived from the Earth System Model iLOVECLIM, Permafrost Periglac, 27, 31-42, 2016.

Koven, C. D., Riley, W. J., Subin, Z. M., Tang, J. Y., Torn, M. S., Collins, W. D., Bonan, G. B., Lawrence, D. M., and Swenson, S. C.: The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4, Biogeosciences, 10, 7109-7131, 2013.

Kurylyk, B. L., and Watanabe, K.: The mathematical representation of freezing and thawing processes in variably-saturated, non-deformable soils, Adv. Water Resour., 60, 160-177, 10.1016/j.advwatres.2013.07.016, 2013.

Kelly, R., Parton, W., Hartman, M., Stretch, L., Ojima, D., and Schimel, D.: Intra-annual and interannual variability of ecosystem processes in shortgrass steppe, Journal of Geophysical Research: Atmospheres, 105, 20093-20100, 2000.

Koven, C. D., Riley, W. J., Subin, Z. M., Tang, J. Y., Torn, M. S., Collins, W. D., Bonan, G. B., Lawrence, D. M., and Swenson, S. C.: The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4, Biogeosciences, 10, 7109-7131, 2013.

Lawrence, D. M., Thornton, P. E., Oleson, K. W., and Bonan, G. B.: The partitioning of evapotranspiration into transpiration, soil evaporation, and canopy evaporation in a GCM: Impacts on land-atmosphere interaction, J. Hydrometeorol., 8, 862-880, 2007.

Lawrence, D. M., and Slater, A. G.: Incorporating organic soil into a global climate model, Clim Dynam, 30, 145-160, 10.1007/s00382-007-0278-1, 2008.

Lyman, S. N., Tran, H. N., Mansfield, M. L., Bowers, R., and Smith, A.: Strong temporal variability in methane fluxes from natural gas well pad soils, Atmospheric Pollution Research, 2020.

McGrath, D., Webb, R., Shean, D., Bonnell, R., Marshall, H. P., Painter, T. H., Molotch, N. P., Elder, K., Hiemstra, C., and Brucker, L.: Spatially Extensive Ground-Penetrating Radar Snow Depth Observations During NASA's 2017 SnowEx Campaign: Comparison With In Situ, Airborne, and Satellite Observations, Water Resources Research, 55, 10026-10036, 10.1029/2019WR024907, 2019.
Mottaghy, D., and Rath, V.: Latent heat effects in subsurface heat transport modelling and their impact on palaeotemperature reconstructions, Geophys J Int, 164, 236-245, 2006.
Moyano, F. E., Manzoni, S., and Chenu, C.: Responses of soil heterotrophic respiration to moisture availability: An exploration of processes and models, Soil Biol Biochem, 59, 72-85, 2013.
Natali, S. M., Watts, J. D., Rogers, B. M., Potter, S., Ludwig, S. M., Selbmann, A.-K., Sullivan, P. F., Abbott, B. W., Arndt, K. A., Birch, L., Björkman, M. P., Bloom, A. A., Celis, G., Christensen, T. R., Christiansen, C. T., Commane, R., Cooper, E. J., Crill, P., Czimczik, C., Davydov, S., Du, J., Egan, J. E., Elberling, B., Euskirchen, E. S., Friborg, T., Genet, H., Göckede, M., Goodrich, J. P., Grogan, P., Helbig, M., Jafarov, E. E., Jastrow, J. D., Kalhori, A. A. M., Kim, Y., Kimball, J. S., Kutzbach, L., Lara, M. J., Larsen, K. S., Lee, B.-Y., Liu, Z., Loranty, M. M., Lund, M., Lupascu, M., Madani, N., Malhotra, A., Matamala, R., McFarland, J., McGuire, A. D., Michelsen, A., Minions, C., Oechel, W. C., Olefeldt, D., Parmentier, F.-J. W., Pirk, N., Poulter, B., Quinton, W., Rezanezhad, F., Risk, D., Sachs, T., Schaefer, K., Schmidt, N. M., Schuur, E. A. G., Semenchuk, P. R., Shaver, G., Sonnentag, O., Starr, G., Treat, C. C., Waldrop, M. P., Wang, Y., Welker, J., Wille, C., Xu, X., Zhang, Z., Zhuang, Q., and Zona, D.: Large loss of CO2 in winter observed across the northern permafrost region, Nat Clim Change, 10.1038/s41558-019-0592-8, 2019.
Nicolsky, D. J., Romanovsky, V. E., Alexeev, V. A., and Lawrence, D. M.: Improved modeling of permafrost dynamics in a GCM land-surface scheme, Geophys Res Lett, 34, 2007.
Niu, G. Y., and Yang, Z. L.: Effects of frozen soil on snowmelt runoff and soil water storage at a continental scale, J. Hydrometeorol., 7, 937-952, 2006.
Neumann, R. B., Moorberg, C. J., Lundquist, J. D., Turner, J. C., Waldrop, M. P., McFarland, J. W., Euskirchen, E. S., Edgar, C. W., and Turetsky, M. R.: Warming effects of spring rainfall increase methane emissions from thawing permafrost, Geophys Res Lett, 2019.
Oechel, W. C., and Kalhori, A.: ABoVE: CO2 and CH4 Fluxes and Meteorology at Flux Tower Sites, Alaska, 2015-2017, https://doi.org/10.3334/ornldaac/1562, 2018.
Oleson, K. W., Lawrence, D., Bonan, G., Drewniak, B., Huang, M., Koven, C., Levis, S., Li, F., Riley, W., and Subin, Z.: Technical Description of version 4.5 of the Community Land Model (CLM)(NCAR Technical Note No. NCAR/TN-503+ STR). Citeseer, National Center for Atmospheric Research, PO Box, 3000, 2013.
Pirazzini, R., Leppanen, L., Picard, G., Lopez-Moreno, J. I., Marty, C., Macelloni, G., Kontu, A., von Lerber, A., Tanis, C. M., Schneebeli, M., de Rosnay, P., and Arslan, A. N.: European In-Situ Snow Measurements: Practices and Purposes, Sensors-Basel, 18, Artn 2016

10.3390/S18072016, 2018.
Rawlins, M. A., Nicolsky, D. J., McDonald, K. C., and Romanovsky, V. E.: Simulating soil freeze/thaw dynamics with an improved pan-Arctic water balance model, J Adv Model Earth Sy, 5, 659-675, 2013.
Riley, W. J., Subin, Z. M., Lawrence, D. M., Swenson, S. C., Torn, M. S., Meng, L., Mahowald, N. M., and Hess, P.: Barriers to predicting changes in global terrestrial methane fluxes: analyses using CLM4Me, a methane biogeochemistry model integrated in CESM, Biogeosciences, 8, 1925-1953, 10.5194/bg-8-1925-2011, 2011.
Russell, S. J., Bohrer, G., Johnson, D. R., Villa, J. A., Heltzel, R., Rey-Sanchez, C., and Matthes, J. H.: Quantifying CH4 concentration spikes above baseline and attributing CH4 sources to hydraulic fracturing activities by continuous monitoring at an off-site tower, Atmos Environ, 117452, 2020.

Schimel, J. P.: Plant-Transport and Methane Production as Controls on Methane Flux from Arctic Wet Meadow Tundra, Biogeochemistry, 28, 183-200, 1995.

Sierra, C. A., Trumbore, S. E., Davidson, E. A., Vicca, S., and Janssens, I.: Sensitivity of decomposition rates of soil organic matter with respect to simultaneous changes in temperature and moisture, J Adv Model Earth Sy, 7, 335-356, 2015.

Slater, A. G., Lawrence, D. M., and Koven, C. D.: Process-level model evaluation: a snow and heat transfer metric, Cryosphere, 11, 989-996, 2017.

Skopp, J., Jawson, M., and Doran, J.: Steady-state aerobic microbial activity as a function of soil water content, Soil Sci Soc Am J, 54, 1619-1625, 1990.

Tang, J. Y., and Riley, W. J.: A Theory of Effective Microbial Substrate Affinity Parameters in Variably Saturated Soils and an Example Application to Aerobic Soil Heterotrophic Respiration, J Geophys Res-Biogeo, 124, 918-940, 2019.

Tao, J., Zhu, Q., Riley, W. J., and Neumann, R. B.: Snow-to-Rain Shifts Regulate Cold-Season Carbon Emissions From pan-Arctic Permafrost, TBD. (In Preparation), 2021.

Wang, L., Koike, T., Yang, K., Jin, R., and Li, H.: Frozen soil parameterization in a distributed biosphere hydrological model, Hydrol Earth Syst Sc, 14, 557-571, 2010.

Wania, R., Ross, I., and Prentice, I. C.: Implementation and evaluation of a new methane model within a dynamic global vegetation model: LPJ-WHyMe v1.3.1, Geosci Model Dev, 3, 565-584, 2010.

Wang, Y. H., Yuan, F. M., Yuan, F. H., Gu, B. H., Hahn, M. S., Torn, M. S., Ricciuto, D. M., Kumar, J., He, L. Y., Zona, D., Lipson, D. A., Wagner, R., Oechel, W. C., Wullschleger, S. D., Thornton, P. E., and Xu, X. F.: Mechanistic Modeling of Microtopographic Impacts on CO2 and CH4 Fluxes in an Alaskan Tundra Ecosystem Using the CLM-Microbe Model, J Adv Model Earth Sy, 11, 4288-4304, 2019.

Xu, X. Y., Riley, W. J., Koven, C. D., Billesbach, D. P., Chang, R. Y. W., Commane, R., Euskirchen, E. S., Hartery, S., Harazono, Y., Iwata, H., McDonald, K. C., Miller, C. E., Oechel, W. C., Poulter, B., Raz-Yaseef, N., Sweeney, C., Torn, M., Wofsy, S. C., Zhang, Z., and Zona, D.: A multi-scale comparison of modeled and observed seasonal methane emissions in northern wetlands, Biogeosciences, 13, 5043-5056, 10.5194/bg-13-5043-2016, 2016.

Yan, Z. F., Bond-Lamberty, B., Todd-Brown, K. E., Bailey, V. L., Li, S. L., Liu, C. Q., and Liu, C. X.: A moisture function of soil heterotrophic respiration that incorporates microscale processes, Nat Commun, 9, 2018.

Zhu, Q., Riley, W. J., Tang, J. Y., Collier, N., Hoffman, F. M., Yang, X. J., and Bisht, G.: Representing Nitrogen, Phosphorus, and Carbon Interactions in the E3SM Land Model: Development and Global Benchmarking, J Adv Model Earth Sy, 11, 2238-2258, 2019.

Zhu, Q., Riley, W. J., Iversen, C. M., and Kattge, J.: Assessing Impacts of Plant Stoichiometric Traits on Terrestrial Ecosystem Carbon Accumulation Using the E3SM Land Model, J Adv Model Earth Sy, 12, 2020.

Zona, D., Gioli, B., Commane, R., Lindaas, J., Wofsy, S. C., Miller, C. E., Dinardo, S. J., Dengel, S., Sweeney, C., and Karion, A.: Cold season emissions dominate the Arctic tundra methane budget, Proceedings of the National Academy of Sciences, 113, 40-45, 2016.

**Response to the Reviewer #2:**

We thank the reviewer for the constructive comments and suggestions that have helped us rethink and improve the manuscript. We have revised the manuscript according to the reviewer's comments (see point-by-point responses below). (Note, we now decided to follow reviewer #1's suggestion on replacing our originally modified ELM moisture scalars with new scalars and then conducted more simulations together with sensitivity analysis to parameters related to carbon decomposition and methane model. We have updated our results and discussions accordingly; our major conclusions about the model improvements in simulating soil temperature and zero-curtain period remain the same, and the improvements regarding simulating $CO_2$ and $CH_4$ fluxes still hold, although with different optimized parameterizations. )

For reference, our response to comment "n" by the reviewer is labeled "R2C[n]" where R2 represents Reviewer #2. Throughout this document, the reviewer's comments are reproduced in their entirety in black, and our responses are given directly afterward in blue.

The manuscript by Tao et al. describes an improved capacity of the ELM land surface model to simulate the zero-curtain period and cold season greenhouse gas emissions. The paper is well-written and the changes made to the model are well-described. I don't see large shortcomings to this paper but, like the other reviewer, it would be nice to have a few more clarifications on why certain approaches were chosen and to place the results in a broader context.

We appreciate the reviewer's constructive comments. We have provided particularly detailed elaboration for the issues pointed out by the reviewer in the revised manuscript as discussed by the following responses.

First of all, the model is only tested on four sites in Alaska. Two are from the same area, while the other two are further inland. I'm not convinced that this climatic gradient is sufficient to capture the dynamics of the cold season across the Arctic, which is the stated goal by the authors for their next paper. Especially since the model does not capture the soil temperature during the cold season at IVO. This may be due to the model setup (e.g. soil conditions or atmospheric forcing), but could also be due to an incorrect simulation of the insulation of the snow as suggested by the other reviewer. In any case, this does not add confidence that the model will perform well in, for example, central Siberia or in the sub-Arctic, where winter conditions are quite different from the north slope of Alaska. This regional bias needs to be considered in the text since it is essential to judge the performance of the model.

**R2C1:** Thank you for the comments. Below we address the reviewer's concerns from three aspects.

1) Site number. We agree with the reviewer that the site number is limited. We would also like to test more sites, however, sites in permafrost regions that have all the necessary measurements needed for this study, including snow depth, soil temperature, soil moisture, year-round $CO_2$ and $CH_4$ fluxes (especially over the cold season), are quite rare. In the future, we will increase our site number by exploring newly published $CO_2$ and $CH_4$ datasets over the pan-Arctic permafrost domain (e.g., BAWLD-CH4, McKenzie et al. (2021), FLUXNET-CH4, Delwiche et al., 2021, etc.).

2) We agree with the reviewer about the importance of soil conditions, atmospheric forcing, and snow conditions. At IVO, the model indeed well reproduced zero-curtain duration and soil temperatures, as shown in Figures 2, 3, and 4, although still underestimating winter soil temperature.

We have provided the comparison results of simulated snow depths against measurements (which show problematic measurements, though) and sensitivity analysis of simulated carbon fluxes to snow conditions in the supplementary file (Figure S9). In our flowing paper (Tao et al., 2021b), we will exclusively focus on the snow impacts on carbon fluxes. Please also see the figure and discussion in our response to the snow issues raised by Reviewer#1 (R1C2).

3) Regional bias. Actually, here we aimed to identify a generic parameterization that can be applied to Alaskan Arctic tundra, not over the whole pan-Arctic domain across a variety of climate and ecosystem gradients, although it is our ultimate goal. In a recent paper focusing on regional simulation over Alaskan Arctic tundra, we used the generic decomposition scheme identified in the original manuscript and found significant improvements in regional $CO_2$ budgets over the Alaska North Slope tundra compared to baseline results (Tao et al., 2021a). The newly identified generic parameterization is superior to the early version. Here, we applied the newly identified generic parameterization to six high-latitude sites and compared simulation results with historical measurements (Figure R2.1; Figure S10 in the supplementary file). ELMv1-ECA simulated results with the identified generic parameterization (red; Figure R2.1) demonstrate improvements in the seasonality of $CH_4$ emissions at these sites compared to baseline results (green; Figure R2.1). These results can support the applicability of the identified generic parameterization as a reasonable initial scheme for simulating $CO_2$ and $CH_4$ emissions over other high-latitude regions when using ELMv1-ECA. In the future, we will optimize regional simulations against recently published spatial datasets of $CO_2$ and $CH_4$ upscaled from in-situ measurements over the pan-Arctic domain.

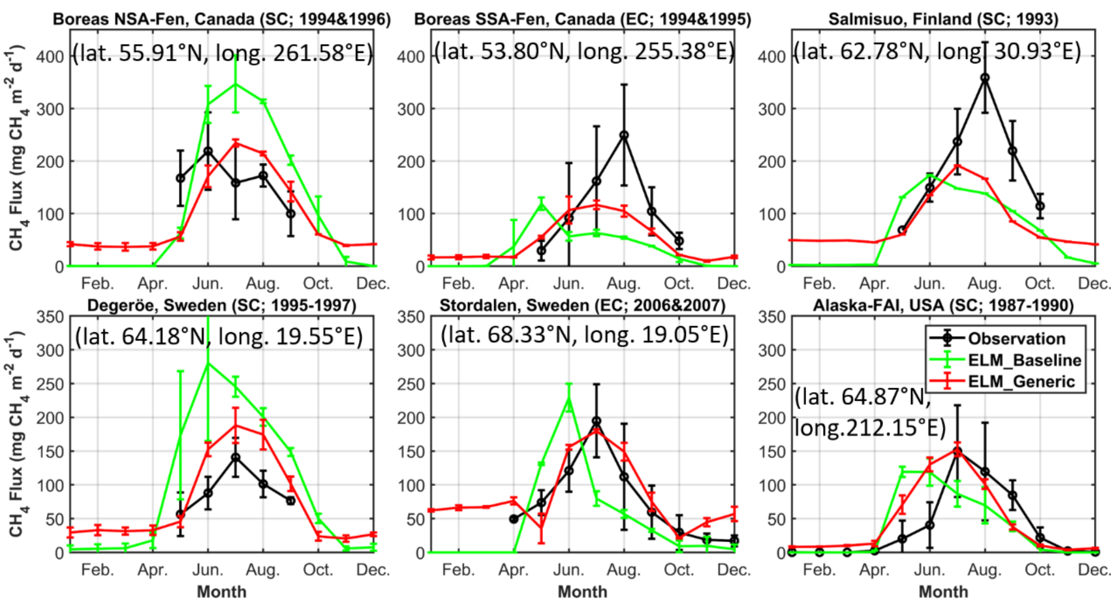


**Figure R2.1:** (Figure S10 in the supplementary file) Comparison of observed and simulated CH4 emissions at six high-latitude sites as reported in Riley et al. (2011) and Xu et al. (2016).

Overall, to address the reviewer's concerns, we have added into the revised manuscript the discussion about the representativeness of the tested tundra sites, how the generic decomposition scheme identified here can be possibly transferred to regional scale, e.g., over the Alaskan North Slope tundra, and our plan for expanding the study over the pan-Arctic domain.

> "Despite the small site number and the limited spatial representativeness of each site, the identified generic decomposition scheme might be applied to the Alaska North Slope tundra. Nevertheless, the generic decomposition scheme might induce uncertainty in simulations and might not be the optimal regional scheme over other ecosystems or given

different climate forcing and soil conditions. Still, we conclude that when using ELMv1-ECA, the generic decomposition scheme can serve as a reasonable initial scheme for estimating $CO_2$ and $CH_4$ emissions over other high-latitude areas (e.g., Figure S10). In the future, we will explore more sites from newly published $CO_2$ and $CH_4$ datasets from pan-Arctic ecosystems, e.g., BAWLD-CH4 (Kuhn et al., 2021) and FLUXNET-CH4 (Delwiche et al., 2021; Knox et al., 2019)." (Lines 476 - 481 in the revised manuscript)

"In the future, we will examine the generic model parameterization at more sites over the pan-Arctic; we will also optimize regional simulations against spatial datasets of $CO_2$ and $CH_4$ upscaled from in-situ measurements over pan-Arctic permafrost domain (Natali et al., 2019; Virkkala et al., 2021; Zeng et al., 2020; Peltola et al., 2019), and discuss the uncertainty of estimated trends of the spatially averaged $CO_2$ and $CH_4$ emissions associated with snow impact and model parameterizations." (Lines 547 - 551 in the revised manuscript)

Second, the simulation of cold season greenhouse gas emissions is much improved but, again, with only a few sites used for validation this may be getting the right numbers for the wrong reasons, when the model has been specifically optimized for these sites. The addition of cracks and plant remnants to act as conduits to the atmosphere makes sense, but this is a rudimentary solution that does not enable the simulation of sudden bursts of CO2 and CH4 which have been observed across the Arctic during the cold season – including at Barrow (Mastepanov et al., 2008; Pirk et al., 2017; Raz-Yaseef et al., 2017). A discussion on why the model is not able to do this, and how this may lead to a systematic bias would be warranted.

**R2C2:** We were aware of the sudden bursts of $CO_2$ and $CH_4$ during the freeze-up period because the gases are pushed out of freezing soils (Mastepanov et al., 2008; Pirk et al., 2017). During the spring period, the $CO_2$ and $CH_4$ emission pulses released from the accumulated gases trapped underneath over the winter period due to thawing of frozen surface and enhanced soil cracking associated with rain-on-snow events (Raz-Yaseef et al., 2017) are much less than the sudden bursts during the freeze-up period at the tested sites (Figures 7 and 8 in the revised manuscript). Indeed, our results perform better in reproducing spring-thaw period emissions than the sudden bursts of carbon emissions during freeze-up periods (Figures 7 and 8 in the revised manuscript). Currently, we mimicked the freeze-up sudden burst mechanism by preventing $CO_2$ and $CH_4$ from dissolving in the soil ice fraction (Riley et al., 2011). As suggested by the reviewer, we have added discussion on how we simulate this mechanism and why it currently cannot well capture the sudden burst, and also our plan to address this issue in the future.

"As for the underestimations of post-ZCP carbon emissions, one critical reason is the lack of sudden bursts of $CO_2$ and $CH_4$ within the model, i.e., the gases are pushed out of freezing soils during the freeze-up period (Mastepanov et al., 2008; Pirk et al., 2017). Currently, the ELMv1-ECA mimics this sudden burst mechanism by preventing $CO_2$ and $CH_4$ from dissolving in the soil ice fraction (Riley et al., 2011), which could capture some burst emissions (e.g., $CH_4$ emissions in Oct. and Sep. of 2013 at ATQ; Figure 6); but it still shows an overall underestimation for sudden-burst emissions especially at colder sites (e.g., BES/CMDL and BEO; Figures 6 and 7). We will improve this mechanism in the future by explicitly simulating ice encroaching soil pores and pushing out gases and liquid water out of the soil matrix." (Section 4.2, Lines 490 - 497 in the revised manuscript)

"The underestimated emissions during post-ZCP months (Oct. to Nov.) are mainly caused by the lack of sudden bursts of $CO_2$ and $CH_4$ during the freeze-up period." (Section 5, Lines 591 - 592 in the revised manuscript)

Finally, the paper is incredibly detailed, which is generally welcome, but in this case there are simply too many figures and tables. The information presented in Figure 5 overlaps with Figure 3 and Table 4, for example. I suggest that some of these figures and tables are moved to the supplemental information, especially when they're only briefly discussed in the text.

**R2C3:** We agree with the reviewer about the repeated information from Figure 5 and Table 4, and thus we have moved Table 4 to the supplementary file as (Table S4) and retained Figure 5 in the manuscript. We use Figure 3 to analyze the simulated freezing process and how the revised soil water phase-change scheme improves soil temperature simulations, explicitly highlighting the better simulated zero-curtain periods, which is critical to our manuscript.

We have also removed Figure 2 and Figure 6, and moved Table 5 to the supplementary file since it repeats information that Figure 9 depicts.

Also, some of the figures are incredibly busy because several parameters are plotted together but this makes it confusing to me what I'm looking at without continuously checking the legend. The colors are hard to distinguish from each other, especially the yellow color when printed. It would also help if the observations are plotted with a clear black line or dashed vs continuous, for example, and that soil moisture and temperature are also plotted with different line types.

**R2C4:** Thank you for the suggestions. We have modified our figures accordingly. Specifically, we have replotted time-series plots (including the original Figures 3, 8, and 9), using black lines for observations, and also changed yellow to other colors. Particularly for Figure 3 (Figure 2 in the revised manuscript), we have changed the soil moisture lines to dashed lines with contrasting colors (shown below).

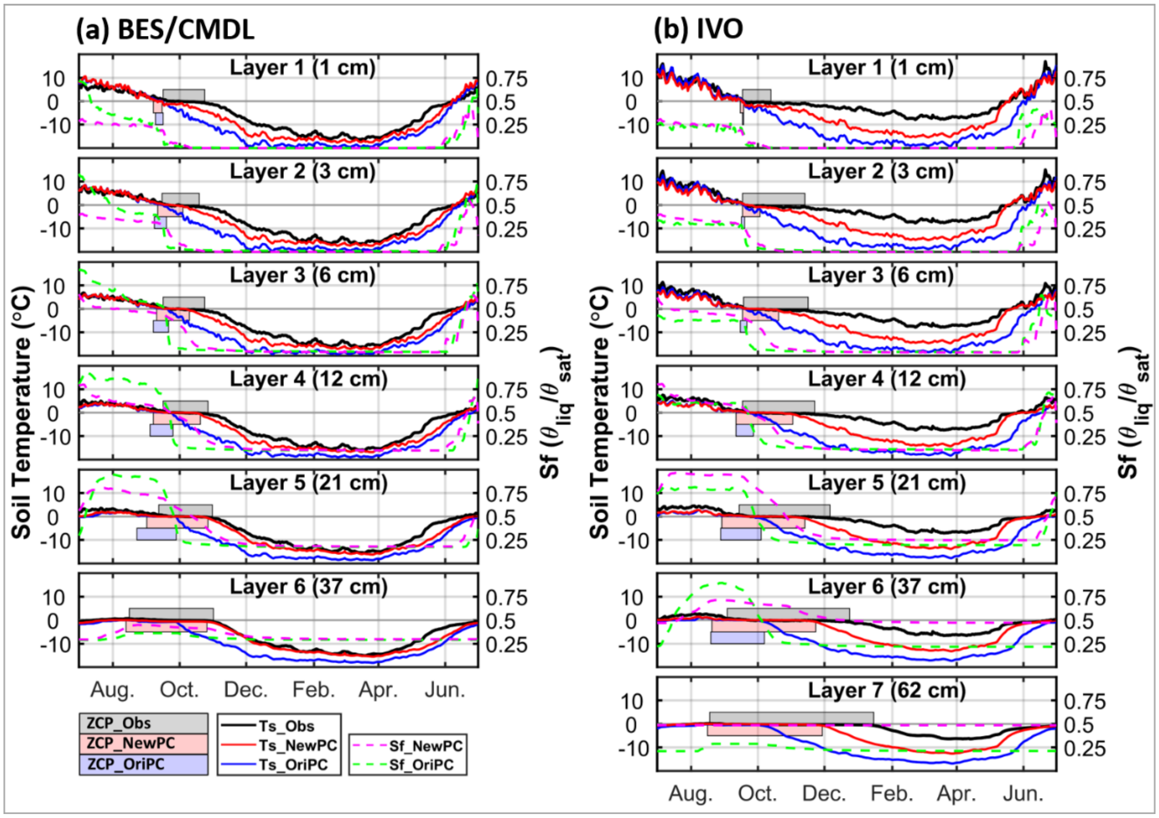


**Figure R2.2:** (New Figure 3 in the revised manuscript) Comparison of multi-year (2013 - 2017) averaged daily soil temperatures observed (Ts_Obs, black) and simulated with the original (Ts_OriPC, blue) and improved (Ts_NewPC, red) phase-change schemes at BES/CMDL (a) and IVO (b). Simulated moisture saturation with the original (Sf_OriPC; green) and improved (Sf_NewPC; magenta) schemes are shown on the right hand axes. The horizontal axes indicates days from July to June, with ticks represent the first day of each month. Hatched areas represent durations of zero-curtain periods observed (ZCP_Obs, gray) and simulated (ZCP_OriPC, blue; ZCP_NewPC, red). No baseline ZCP is shown in the 6$^{th}$ layer for BES/CMDL and the 7$^{th}$ layer for IVO because the maximum annual temperature is below 0°C.

Minor comments:

Page 4, Line 110: were these gaps large? If gaps were only a few days this is fine, but it would be good to know if weeks or months of data needed to be gap-filled.

**R2C5:** The original sentence is copied below:

> "We first filled missing gaps vertically by fitting a polynomial to the soil temperature profile (Kurylyk and Hayashi, 2016) on a daily scale, then screened out outliers by examining the daily time series." (Lines 109 - 110 in the revised manuscript)

The ABoVE/CARVE in situ measurements are available at depth 5 cm, 10 cm, 15 cm, 20 cm, 30 cm, and 40 cm, and UAF GIFL soil temperatures are also available at various depths. Here, by "gaps" we meant the discontinuities in measurements along the vertical soil depths. For instance, sometimes the measurements at 20 cm are missing, then we filled this missing data by fitting a polynomial to the soil temperature profile, i.e., measurements available at other depths. We only perform this gap-filling if we have at least one measurement at depths above the missing measurement depth and at least one measurement at depths below the missing depth. For gaps in time, we did not perform gap-filling in case introducing artifacts.

Page 6, line 176-177: no need to specify that the 'S' stands for supplemental. This is rather standard knowledge.

**R2C6:** We have removed this sentence.

Page 13, line 405: it's unclear to me why there's an ensemble of grey dots for the NewPC_NewDecom_NewCH4 but not for the other dots? This is not well-described in the caption or the text.

**R2C7:** The grey dots represent all the tested (1934) parameterizations of carbon decomposition schemes and methane parameters listed in Tables S2 and S3, as indicated in the annotation of Table 2. We have added clear clarification in the caption of Figure 7 (now is Figure 5 in the revised manuscript). Thanks for the suggestion.

Page 18, line 562: please elaborate on why the single static multiplicative function would not be appropriate.

**R2C8:** We have added more elaboration on this and change the sentence as below.

> "In addition, the single static multiplicative function used to parameterize the impact of environmental conditions on respiration might not be appropriate because the environmental impact also depends on maximum respiration rate, soil texture, soil carbon content and quality, and microbial biomass (Tang and Riley, 2019)." (Lines 591 - 593 in the revised manuscript)

**References**

Mastepanov, M., Sigsgaard, C., Dlugokencky, E. J., Houweling, S., Ström, L., Tamstorf, M. P. and Christensen, T. R.: Large tundra methane burst during onset of freezing, Nature, 456(7222), 628–630, doi:10.1038/nature07464, 2008.

Pirk, N., Mastepanov, M., López-Blanco, E., Christensen, L. H., Christiansen, H. H., Hansen, B. U., Lund, M., Parmentier, F.-J. W., Skov, K. and Christensen, T. R.: Toward a statistical description of methane emissions from arctic wetlands, Ambio, 46(1), 70–80, doi:10.1007/s13280-016-0893-3, 2017.

Raz-Yaseef, N., Torn, M. S., Wu, Y., Billesbach, D. P., Liljedahl, A. K., Kneafsey, T. J., Romanovsky, V. E., Cook, D. R. and Wullschleger, S. D.: Large CO2 and CH4 emissions from polygonal tundra during spring thaw in northern Alaska, Geophysical Research Letters, 44(1), 504–513, doi:10.1002/2016GL071220, 2017.

## References For Author's Response


Delwiche, K. B., Knox, S. H., Malhotra, A., Fluet-Chouinard, E., McNicol, G., Feron, S., Ouyang, Z., Papale, D., Trotta, C., and Canfora, E.: FLUXNET-CH4: A global, multi-ecosystem dataset and analysis of methane seasonality from freshwater wetlands, Earth System Science Data Discussions, 1-111, 2021.

Knox, S. H., Jackson, R. B., Poulter, B., McNicol, G., Fluet-Chouinard, E., Zhang, Z., Hugelius, G., Bousquet, P., Canadell, J. G., Saunois, M., Papale, D., Chu, H., Keenan, T. F., Baldocchi, D., Torn, M. S., Mammarella, I., Trotta, C., Aurela, M., Bohrer, G., Campbell, D. I., Cescatti, A., Chamberlain, S., Chen, J., Chen, W., Dengel, S., Desai, A. R., Euskirchen, E., Friborg, T., Gasbarra, D., Goded, I., Goeckede, M., Heimann, M., Helbig, M., Hirano, T., Hollinger, D. Y., Iwata, H., Kang, M., Klatt, J., Krauss, K. W., Kutzbach, L., Lohila, A., Mitra, B., Morin, T. H., Nilsson, M. B., Niu, S., Noormets, A., Oechel, W. C., Peichl, M., Peltola, O., Reba, M. L., Richardson, A. D., Runkle, B. R. K., Ryu, Y., Sachs, T., Schafer, K. V. R., Schmid, H. P., Shurpali, N., Sonnentag, O., Tang, A. C. I., Ueyama, M., Vargas, R., Vesala, T., Ward, E. J., Windham-Myers, L., Wohlfahrt, G., and Zona, D.: FLUXNET-CH4 Synthesis Activity: Objectives, Observations, and Future Directions, Bulletin of the American Meteorological Society, 100, 2607-2632, 10.1175/Bams-D-18-0268.1, 2019.

Kuhn, M. A., Varner, R. K., Bastviken, D., Crill, P., MacIntyre, S., Turetsky, M., Walter Anthony, K., McGuire, A. D., and Olefeldt, D.: BAWLD-CH 4: A Comprehensive Dataset of Methane Fluxes from Boreal and Arctic Ecosystems, Earth System Science Data Discussions, 1-56, 2021.

Natali, S. M., Watts, J. D., Rogers, B. M., Potter, S., Ludwig, S. M., Selbmann, A.-K., Sullivan, P. F., Abbott, B. W., Arndt, K. A., Birch, L., Björkman, M. P., Bloom, A. A., Celis, G., Christensen, T. R., Christiansen, C. T., Commane, R., Cooper, E. J., Crill, P., Czimczik, C., Davydov, S., Du, J., Egan, J. E., Elberling, B., Euskirchen, E. S., Friborg, T., Genet, H., Göckede, M., Goodrich, J. P., Grogan, P., Helbig, M., Jafarov, E. E., Jastrow, J. D., Kalhori, A. A. M., Kim, Y., Kimball, J. S., Kutzbach, L., Lara, M. J., Larsen, K. S., Lee, B.-Y., Liu, Z., Loranty, M. M., Lund, M., Lupascu, M., Madani, N., Malhotra, A., Matamala, R., McFarland, J., McGuire, A. D., Michelsen, A., Minions, C., Oechel, W. C., Olefeldt, D., Parmentier, F.-J. W., Pirk, N., Poulter, B., Quinton, W., Rezanezhad, F., Risk, D., Sachs, T., Schaefer, K., Schmidt, N. M., Schuur, E. A. G., Semenchuk, P. R., Shaver, G., Sonnentag, O., Starr, G., Treat, C. C., Waldrop, M. P., Wang, Y., Welker, J., Wille, C., Xu, X., Zhang, Z., Zhuang, Q., and Zona, D.: Large loss of CO2 in winter observed across the northern permafrost region, Nat Clim Change, 10.1038/s41558-019-0592-8, 2019.

Peltola, O., Vesala, T., Gao, Y., Raty, O., Alekseychik, P., Aurela, M., Chojnicki, B., Desai, A. R., Dolman, A. J., Euskirchen, E. S., Friborg, T., Gockede, M., Helbig, M., Humphreys, E., Jackson, R. B., Jocher, G., Joos, F., Klatt, J., Knox, S. H., Kowalska, N., Kutzbach, L., Lienert, S., Lohila, A., Mammarella, I., Nadeau, D. F., Nilsson, M. B., Oechel, W. C., Peichl, M., Pypker, T., Quinton, W., Rinne, J., Sachs, T., Samson, M., Schmid, H. P., Sonnentag, O., Wille, C., Zona, D., and Aalto, T.: Monthly gridded data product of northern wetland methane emissions based on upscaling eddy covariance observations, Earth Syst Sci Data, 11, 1263-1289, 10.5194/essd-11-1263-2019, 2019.

Riley, W. J., Subin, Z. M., Lawrence, D. M., Swenson, S. C., Torn, M. S., Meng, L., Mahowald, N. M., and Hess, P.: Barriers to predicting changes in global terrestrial methane fluxes: analyses using CLM4Me, a methane biogeochemistry model integrated in CESM, Biogeosciences, 8, 1925-1953, 10.5194/bg-8-1925-2011, 2011.

Tang, J. Y., and Riley, W. J.: A Theory of Effective Microbial Substrate Affinity Parameters in Variably Saturated Soils and an Example Application to Aerobic Soil Heterotrophic Respiration, J Geophys Res-Biogeo, 124, 918-940, 2019.

Tao, J., Zhu, Q., Riley, W. J., and Neumann, R. B.: Warm-season net CO2 uptake outweighs cold-season emissions over Alaskan North Slope tundra under current and RCP8.5 climate Environmental Research Letters, (Accepted), 2021a.
Tao, J., Zhu, Q., Riley, W. J., and Neumann, R. B.: Snow-to-Rain Shifts Regulate Cold-Season Carbon Emissions From pan-Arctic Permafrost, TBD. (In Preparation), 2021b.
Virkkala, A. M., Aalto, J., Rogers, B. M., Tagesson, T., Treat, C. C., Natali, S. M., Watts, J. D., Potter, S., Lehtonen, A., and Mauritz, M.: Statistical upscaling of ecosystem CO2 fluxes across the terrestrial tundra and boreal domain: regional patterns and uncertainties, Global Change Biol, 2021.
Watts, J. D., Natali, S., Potter, S., and Rogers, B. M.: Gridded Winter Soil CO2 Flux Estimates for pan-Arctic and Boreal Regions, 2003-2100, https://doi.org/10.3334/ORNLDAAC/1683, 2019.

Xu, X. Y., Riley, W. J., Koven, C. D., Billesbach, D. P., Chang, R. Y. W., Commane, R., Euskirchen, E. S., Hartery, S., Harazono, Y., Iwata, H., McDonald, K. C., Miller, C. E., Oechel, W. C., Poulter, B., Raz-Yaseef, N., Sweeney, C., Torn, M., Wofsy, S. C., Zhang, Z., and Zona, D.: A multi-scale comparison of modeled and observed seasonal methane emissions in northern wetlands, Biogeosciences, 13, 5043-5056, 10.5194/bg-13-5043-2016, 2016.
Zeng, J. Y., Matsunaga, T., Tan, Z. H., Saigusa, N., Shirai, T., Tang, Y. H., Peng, S. S., and Fukuda, Y.: Global terrestrial carbon fluxes of 1999-2019 estimated by upscaling eddy covariance data with a random forest, Scientific Data, 7, 2020.