**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 will revise the manuscript according to the reviewer's comments (see point-by-point responses below). 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 will provide 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:** We agree with the reviewer that the site number is limited. We would also like to test more sites, however, sites over 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, are quite rare. We did test the updated ELM at Alaska SNOTEL sites and all the cold-season $CO_2$ flux sites over pan-Arctic permafrost regions as reported by Natali et al. (2019). But, we only found one pair of SNOTEL sites (AK-968) and cold-season $CO_2$ flux sites that are co-located with each other and have reasonable observations during an overlapping period. We will definitely explore more sites in the future.

At IVO, the model well reproduced cold season soil temperatures, as shown in Figures 3, 4, and 5. We agree with the reviewer about the importance of soil conditions, atmospheric forcing, and snow conditions. We tested multiple reanalysis forcing, including CRUNCEP (Climatic Research Unit and NCEP Reanalysis; QianFilled) and GSWP3 (Global Soil Wetness Project Phase 3), and we will provide simulation results at the tested tundra sites with other climate forcing in the supplementary file. We will also provide a comparison 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. Please also see our response to Reviewer#1 (**R1C2**).

In a recent paper focusing on regional simulation over Alaskan Arctic tundra, we used the generic decomposition scheme identified in Figure 7 that provided the best overall performance for all the sites. We found that the updated ELMv1a demonstrates significantly improved performance in the simulated regional mean of cold-season $CO_2$ emissions over the Alaska North Slope tundra, showing a 55% reduction in RMSE compared to the bassline results (0.14 versus 0.31 gC $m^{-2}$ $day^{-1}$) (Tao et al., 2020).

We also tested the updated ELM (again with the generic decomposition scheme) with multiple reanalysis forcing datasets over pan-Arctic permafrost regions. We found that, compared to a machine-learning derived spatial dataset of $CO_2$ emissions based on ground in situ measurements (denoted N2019 here) (Natali et al., 2019; Watts et al., 2019), the simulated $CO_2$ emissions results driven by GSWP3 show generally smaller biases over pan-Arctic permafrost than that driven by CRUJRA (**Figure R2.1**). We will discuss these biases and possible reasons in our following paper (Tao et al., 2021).

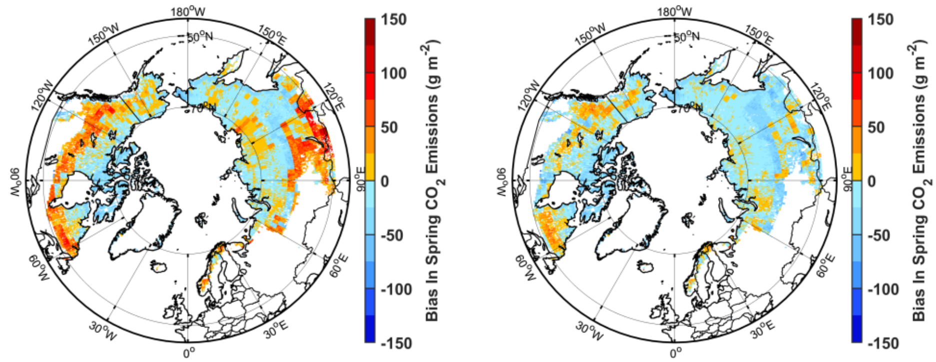


**Figure R2.1:** Bias in multi-year averaged (2003 to 2017) spring $CO_2$ Emissions over pan-Arctic permafrost. Left: Bias = Simulations by ELMv1 driven by CRUJRA – N2019; Right: Bias = Simulations by ELMv1 driven by GSWP3 – N2019. (Preliminary results subject to change.)

We will also add into the revised manuscript some discussion about the representativeness of the tested tundra sites and how the generic decomposition scheme identified here can be transferred to a larger scale across a variety of climate and ecosystem gradients.

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. Currently, we mimicked this 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 will add 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.

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 will move Table 4 to the supplementary file and retain 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 will also remove Figure 2 (see our response to Reviewer #1 R1C49).

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 will modify our figures accordingly. Specifically, we will replot time-series plots (including Figures 3, 8, and 9), using black lines for observations, and also change yellow lines to other colors or in different line styles. Particularly for Figure 3, we will change the soil moisture lines to another line type with contrasting colors.

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."

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 will remove this.

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 (200) new decomposition schemes listed in Table S2 as indicated in the annotation of Table 2. We will add clear clarification in the context and the caption of Figure 7. Thanks for the suggestion.

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

**R2C8:** We will add 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 microbial biomass (Tang and Riley, 2019)."

**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 Comments**

Kurylyk, B. L., and Hayashi, M.: Improved Stefan Equation Correction Factors to Accommodate Sensible Heat Storage during Soil Freezing or Thawing, Permafrost Periglac, 27, 189-203, 2016.

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 Arctic tundra under current and RCP8.5 climate Environmental Research Letters, (Under Review), 2020.

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.

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.