

PRIME PF Reviewer Comments

Reviewer comments

Authors responses

Reviewer 1

General comments

The authors have conducted a study on the possibility of the northern high latitudes becoming a net carbon source under future climate conditions using the novel PRIME framework – an ESM emission-to-impacts emulator, which includes recent advancements in the representation of permafrost physics together with dynamic vegetation not yet available in ESMs. The study shows that the inclusion of permafrost physics and vegetation dynamics is essential in determining the temperature range at which the northern latitudes could transition from a carbon sink to a carbon source.

The reviewer has a good understanding of our study and we thank them for the insightful and helpful review.

Specific comments

Line 89: Is it possible to provide the actual climate sensitivities to which the percentiles correspond?

This version and calibration of FaIR is constrained to reproduce the IPCC AR6 assessed distribution of ECS, so the percentiles used in this study broadly correspond to the percentiles of this distribution. As the PRIME framework uses FaIR in emissions-driven mode and takes percentiles from the temperature distribution at a single timepoint (2100, SSP1-2.6), the selected percentiles do not precisely align with the percentiles from the ECS distribution, but are designed to span this uncertainty range (see Mathison et al. 2025; Figure 3). We have amended the following line in the methodology:

*“This distribution was sub-sampled by selecting the following percentiles of the **global mean temperature in 2100 from SSP1-26: 0th, 1st, 5th, 25th, 50th, 75th, 95th, 99th, 100th** (representing low to high climate sensitivities).”*

Mathison, C., Burke, E. J., Munday, G., Jones, C. D., Smith, C. J., Steinert, N. J., Wiltshire, A. J., Huntingford, C., Kovacs, E., Gohar, L. K., Varney, R. M., and McNeall, D.: A rapid-application emissions-to-impacts tool for scenario assessment: Probabilistic Regional Impacts from Model patterns and Emissions (PRIME), *Geosci. Model Dev.*, 18, 1785–1808, <https://doi.org/10.5194/gmd-18-1785-2025>, 2025.

Lines 172-174: It is not clear to me how the individual lines within each colour in Fig. 2 represent different spatial patterns. Do these correspond to individual grid cells? How is the spatial sensitivity of climate patterns calculated?

PRIME is used to calculate the spatial patterns using the SSP5-8.5 scenario with 18 CMIP6 Earth system models (listed in Table S1). By spatial patterns, we mean the different spatial patterns of warming seen in our ensemble based on the different warming patterns in each CMIP6 ESM at the same global mean temperature change (i.e. different magnitudes of Arctic amplification). See lines 94 to 99 in the Methods. We have made this more clear in the Methods by adding the following text:

“The spatial patterns of change were calculated on a monthly basis using the SSP5-8.5

scenario with 18 CMIP6 ESMs, which were selected from the full CMIP6 ensemble and span the available range of uncertainty (Table S1). **The spatial patterns represent the different patterns of change in the climate variables used to drive JULES seen in each CMIP6 ESM at the same global mean temperature change (i.e. different magnitudes of Arctic amplification).**

JULES was run for each ESM pattern individually with the global mean temperature change and CO₂ concentrations sampled from the FaIR distribution.

Lines 199-201: Based on Fig. S6, this statement is true only for high climate sensitivity scenarios. It would be interesting to elaborate on the reasons why simulations with low climate sensitivity sustain the carbon sink longer in a high emission/high warming scenario than in a low emission/low warming scenario, maybe also in continuation of the analysis provided in the next section (lines 232-238).

The reviewer is correct that we highlight the result of the high climate sensitivity simulations and their impact on the sink-to-source transition more than the low climate sensitivity simulations. This is because our focus is on the sink-to-source transition and when we see this transition. However, it is also interesting that with a low climate sensitivity we see a sustained carbon sink, and that with low climate sensitivities the sink is more resilient to temperature changes under increased emission scenarios.

We edited the following text to account for this statement only being true for high climate sensitivity scenarios:

“This suggests that high warming (from either high emissions or high climate sensitivity) not only accelerates the timing of the sink-to-source transition but also amplifies the magnitude of carbon loss from northern high-latitude ecosystems.”

And we have added the following text to acknowledge this finding on climate sensitivity.

Line 195:

“Generally, the ensemble members suggest that the lower the climate sensitivity, the greater the resilience of the land carbon sink.”

Line 204:

“The results here highlight how the timing of the sink-to-source transition depends individually on the climate sensitivity and climate scenario (Table 3). This means that the sensitivity of the carbon sink is a balance between changes in CO₂ and temperature, and how they change with respect to one another. For example, in low climate sensitivity ensemble members (0th to 5th percentiles), the carbon sink is shown to be more resilient in high emission scenarios (SSP2-4.5, SSP5-8.5) compared to lower emission scenarios (SSP1-2.6, SSP5-3.4-OS). This is likely due to the higher CO₂ and temperature levels providing more optimal conditions for vegetation growth and CO₂ fertilisation. However, if we compare results from low climate sensitivity (0th to 5th percentiles) and high emission scenario (SSP2-4.5, SSP5-8.5) with medium climate sensitivity (25th to 75th percentiles) and lower emission scenario (SSP1-2.6, SSP5-3.4-OS), the timing of the sink-to-source transition is shown to vary by more than 200 years despite similar amounts of global warming seen (i.e. not transitioned by 2300 versus transitioning during 22nd century; Table 3). This implies that under the same global warming level, the CO₂ fertilisation effect negates the carbon loss from the soil under higher emissions, but not in the lower emission scenarios.”

Line 231: Could you explain how you calculated the 62% likelihood? It seems that at 2 degrees 100% of the simulations in SSP2-4.5 are showing a carbon sink.

The sentence on line 231 is comparing the likelihood of a sink-to-source transition occurring at 2 degrees in the SP2-4.5 scenario with the SSP5-3.4-O scenario. The reviewer is correct that

100% of simulations show a carbon sink in SSP2-4.5, and we are saying this means a carbon sink is 62% more likely at 2 degrees in this scenario compared to SSP5-3.4-OS.

Lines 232-238: Consider including here an analysis on the impact of climate sensitivity on why in high emission scenarios the carbon sink is more resilient towards temperature changes (see previous comment). It seems that in low emissions/low warming scenarios the temperature changes are still low enough not to enable forest expansion, while in high emissions/high warming scenarios the low to moderate temperature increase (Fig. S3, lower percentiles) can already trigger the northward expansion. So, the trigger point (in terms of temperature) is not clearly defined and maybe varies with other factors (e.g., CO₂ fertilization).

This comment was addressed in the first section of the results 'Carbon sink-to-source transition' as this is where we explicitly consider the simulations with different climate sensitivities (**addressed above see line 204**). The reviewer made a good point with this and we have also added the following to the conclusions to highlight the differing effects of warming based on climate sensitivity as well as emission scenario.

"It is found that this balance is dependent both on the climate sensitivity of the system, as well as the climate scenario followed."

Lines 299-300: Do the other configurations of JULES implicitly account for the fire emissions from soil? Would it be possible to give an estimate of these emissions (or at least the order of magnitude compared to burning vegetation)?

The JULES-INFERNO configuration is the only configuration which explicitly accounts for fire emissions, however only burns from the litter pools so does not implicitly account for fire emissions from the soil. To use the JULES-INFERNO configuration, the model is tuned against the vegetation distribution and thus implicitly includes the vegetation burning. The other JULES configurations implicitly account for fire emissions for the land surface, but not explicitly from the soil versus vegetation.

Line 323: Is the temperature "exceeding" or rather "going under" the optimum for growth in overshoot scenarios?

Yes we agree with the reviewer that this is misleading wording here. This sentence has been changed to the following:

"This source behaviour arises because vegetation productivity declines rapidly once temperatures drop back below ~~exceed~~ the optimum for growth, while delayed soil respiration continues to rise."

Technical corrections

Line 8: Consider changing to "increase the risk of northern high latitudes becoming a net carbon source by more than 50%"

Changed.

Line 20: Add "and" after "CO₂ concentrations".

Changed.

Line 21: "climate change mitigation".

Changed.

Line 48: "differently"?

We agree this sentence was a bit awkward to read, changed to:

"Earth system models (ESMs) are routinely used to produce large-scale climate

projections under a range of emissions scenarios representing different potential future climate policy pathways.

Line 49: Coupled Model Intercomparison Project Phase 6 (CMIP6)

Changed.

Line 64: Probabilistic Regional Impact from Model patterns and Emissions (PRIME) framework by Mathison et al. (2025)

Changed.

Lines 66-67: Consider changing to: "One advantage of PRIME is the possibility to use JULES configurations..."

Changed.

Line 82: abbreviation inserted above

Changed to define acronym on first use in introduction rather than in methods.

Line 103: consider giving the full name of the JULES abbreviation the first time it is mentioned in the text

JULES is now defined in the Introduction.

Line 181: Do you mean "from higher climate sensitivity percentiles"?

Yes, changed.

Line 190: "than they take up"; Table 3: please, note that when printed the light grey is not visible.

Changed. Thank you for pointing out the issue with the light grey, as it is visible in a PDF format we will leave as it is for now.

Line 259: Consider changing to: "it is argued that... (Alfaro-Sanchez et al., 2024)"

Changed.

Line 329: "create"

Changed.

Lines 329-330: creates a balance between what and what?

We agree this sentence is a bit hard to read, it has been changed to the following:

"Overall, rising atmospheric CO₂ concentrations and global mean temperatures create a balance between enhanced vegetation productivity, which strengthens the terrestrial carbon sink, and increasing permafrost thaw, which releases previously frozen carbon and weakens the sink; where the magnitudes of these individual responses remain uncertain across modelling studies."

Line 333: "This highlights..."?

Changed.

Reviewer 2

This study probabilistically quantifies the patterns of the sink-to-source carbon transition in the northern high latitudes under various emission scenarios (including the overshoot scenario) by means of the PRIME-JULES framework that explicitly couples permafrost physics, dynamic vegetation and fire processes. This study could fill the gap that key carbon cycle processes have not been jointly coupled in the current CMIP6 models.

The reviewer correctly summarises the paper and identifies the gap we are trying to fill, we thank the reviewer.

As an ESM emulator, the PRIME framework can only simulate effects of climates on terrestrial carbon cycle, and fails to simulate the feedback of carbon cycle to climate system. That is, it fails to capture the two-way carbon-climate coupling effects. Furthermore, it does not couple the biophysical feedbacks of terrestrial ecosystems to climate (e.g., the impacts of vegetation albedo and evapotranspiration changes on regional temperatures), which may lead to biases in the simulation of long-term carbon cycle feedbacks.

We agree this is a limitation of using land surface models offline in this way, opposed to the computationally expensive coupling into an Earth system model. However, using the PRIME framework has allowed us to investigate the carbon cycle dynamics including processes yet to be fully coupled in ESMs, i.e., permafrost and dynamic vegetation; processes which we have shown to be especially important in the northern high latitudes. Additionally, PRIME allowed us to produce a large ensemble size compared to what would be possible with an ESM, enabling the sink-to-source transition to be investigated on a much more granular level. This limitation of our framework is highlighted and addressed in the Conclusions section, where we state: *“Additionally, simulations with coupled models will allow for an evaluation of the ecosystem carbon balance including the additional biophysical feedbacks, a current limitation with the PRIME framework.”*

The northward expansion of dynamic vegetation in the model is only driven by climate and CO₂, with no incorporation of factors such as soil nutrient limitations (excluding nitrogen), physical constraints of permafrost instability on plant growth, and human activities (e.g., forest management). In the real world, these factors can significantly inhibit the northward expansion of boreal forests, which may result in an overestimation of the enhancement effect of dynamic vegetation on the carbon sink.

The reviewer correctly identifies processes that are not currently incorporated in land surface models, which will also influence the northward expansion of vegetation. This is discussed in the manuscript in the following text:

“In reality there are additional sources of uncertainty which determine the net carbon balance, such as the amount of boreal forest expansion that will be seen (Rotbarth et al. 2023). For example, there are uncertainties in the amount of tree growth potential within permafrost ecosystems, where it is argued that the allocation of nutrients within a tree will prioritise remaining upright on unstable grounds rather than growth in these regions (Alfaro et al. 2024). The spatial patterns will also be influenced by complex interactions with water and nutrients which could limit growth potential (See et al. 2024; Zhao et al. 2022), and additionally, the impact of fire creates large amounts of uncertainty to the spatial carbon loss (Virkkala et al. 2025).”

Simulations in this study are solely based on the single land surface model JULES. The model results have not been validated with in-situ observational data (e.g., long-term carbon flux observation stations in tundra and boreal forest regions), leaving the generalizability of the results to be verified.

The results from this study are based on a single land surface model and the importance of additional land surface models including high northern latitude processes to enable inter-model comparison is highlighted in the conclusions:

“Investigating the balance of carbon fluxes in this region in additional land surface models will allow for inter-model comparison and help to further understand the sensitivity of such fluxes to climate change (Padron et al. 2022).”

However, it is noted that the novel methodology presented here is the large ensemble which provides a probabilistic analysis of the net land carbon cycle flux, which has not been previously possible with a model such as JULES. This work demonstrates the potential to expand to other land surface models used in ESMs.

Additionally, the manuscript already included a section in the methods which evaluates the JULES-pf configuration (see section 2.3). In this section, JULES-pf is evaluated against CARDAMOM v14 (an observationally-informed model-data fusion analysis), and we additionally compared contemporary NEP estimates with both top-down and bottom-up approaches presented in Hugelius et al. 2024 (see lines 147 to 154). We also highlight the large uncertainty reflected in the data with this value.

“Hugelius et al. 2024 used three broad methodological approaches to quantify the NEP: land surface models, atmospheric inversion and in-situ upscaling (including JULES-ES and CARDAMOM V13), each of which has a large uncertainty range. The total plausible range of NEP spans from a net source of -0.66 to a net sink of 1.37 Gt yr⁻¹; the large range reflects data uncertainties, including whether fire is explicit or implicit, and methodological uncertainties.”

The 95% percentile range of CARDAMOM appears to be overly broad, and thus it is understandable that the results of JULES-pf fall within this range. However, this cannot prove that the results of JULES-pf are reliable or superior to those of other models. Please clarify what the 95% percentile range of CARDAMOM indicates.

The evaluation of JULES-pf is not to imply superiority over other models, it is to use CARDAMOM's observationally informed uncertainty to qualify JULES-pf as reliable. The evaluation assesses output variables against observationally-informed values to evaluate its performance with respect to the information carried in current state-of-the-art geospatial datasets. We do not evaluate against field-based estimates due to the mismatch in spatial scales (observation <1 km², JULES simulation ≈200-300 km²). Whereas satellite-based Earth Observation and upscaled geospatial datasets give the opportunity to evaluate at an appropriate spatial resolution. However, these datasets have been shown to be inadequate to evaluate land surface models as different products that estimate the same ecological variable (e.g. GPP, AGB) do not agree with one another often several times their reported uncertainties (e.g., Seiler et al., 2022).

Our solution is to evaluate using an observationally-informed systemic analysis, enforcing consistency between the observational datasets. On this basis the evaluation shows that JULES-pf's contemporary performance is plausible. For detailed description see Section 1.1 in Supplementary Information. In brief, CARDAMOM is a Bayesian calibration framework applied at pixel-scale, thus a local analysis, to retrieve ensembles of parameters consistent with observations and their uncertainties for an intermediate complexity terrestrial ecosystem model (DALEC) providing a data-informed diagnostic analysis. The 95 % confidence interval is directly estimated from the pixel-level ensemble of parameters and the resulting DALEC simulations. Thus representing a realistic uncertainty range as informed by a variety of datasets (inc. biomass, leaf area, absorbed photosynthetic radiation, soil C, geospatial priors on plant traits).

Seiler, C., Melton, J. R., Arora, V. K., Sitch, S., Friedlingstein, P., and Anthoni, P.: Are terrestrial biosphere models fit for simulating the global land carbon sink? *Journal of Advances in Modeling Earth Systems*, 14, e2021MS002 946, 2022

Line 64: Mathison et al. (2025) should be cited in the correct reference format.
Corrected.