Reply letter to reviewer 2

We thank the reviewer for the thorough review and valuable suggestions for improvement. We have addressed all points in the following. All responses are in italics and changes made in the manuscript are highlighted in bold. Updated figures are shown, but may not be exactly as shown later in the revised version due to ongoing graphics editing.

General Comments

The article describes a new permafrost model, labelled CryoGridLitle, focused on minimizing the computational and forcing resources required to be executed while achieving a realistic representation of permafrost areas. Furthermore, the analysis presents a comprehensive estimate of the uncertainties associated to the model, with the only exception of the uncertainties due to the external forcing and the model itself. The paper is clearly written and structured but I have concerns about the suitability of the model evaluation process followed by the authors, as it is limited in scope. I think that the topic is adequate for the journal, and that the development of a permafrost model able to perform multi-century simulations is relevant for the climate and cryosphere communities. I recommend publication after major revisions.

We thank the reviewer for his thorough review of our manuscript and note the concerns about the limited evaluation of the model. In response, we have made substantial improvements to the manuscript and have included additional material focusing on both the assessment of the model physics and a more detailed assessment at key sites representing different permafrost conditions and histories, as suggested by the reviewer.

Major points

M1- The authors cite in the Introduction a series of works in which permafrost models are used in paleoclimate studies. Nevertheless, the performance of the CryoGridLitle is not compared to that from other models applied at paleoclimate scales. Is the CryoGridLitle framework performing better? What are the advantages of this new model in comparison with previous models used in paleoclimate studies? What are the disadvantages? The answer to these questions should appear in an article introducing a new modelling framework when other models are already available.

We agree with the reviewer that the advantages, limitations, and goals of our modeling approach should be stated more clearly in the introduction. In the revised version of the manuscript, we have rewritten and expanded the relevant paragraph as follows:

Here, we present and evaluate a computationally efficient numerical permafrost model (CryoGridLite) designed to provide insights into the evolution of the thermal state of permafrost and active layer thickness over many centuries for the Arctic permafrost region. With the CryoGridLite model, we aim to bridge the gap between very sophisticated permafrost process models and reduced schemes used for paleoclimatic simulations. Compared to comprehensive process models like CryoGrid3 \citep{westermann2016} or the CryoGrid Community Model \citep{westermann2023cryogrid}, CryoGridLite is approximately three orders of magnitude faster, enabling the execution of long-term simulations spanning hundreds to thousands of years. The enhanced efficiency is achieved through two key components: (i) the utilization of an implicit solution scheme and (ii) a streamlined representation of the underlying processes.
Our approach aligns with the objective of providing plausible ranges of permafrost states rather than focusing on highly precise results for specific locations. This is accomplished through the application of multi-parameter ensemble simulations, which have become feasible due to the substantial enhancements achieved in computational efficiency. By employing this methodology, we can capture a broader range of potential permafrost conditions, encompassing the inherent uncertainties and variability associated with real-world scenarios. In the simulations we account for uncertainties in soil parameters such as water and ice content, and uncertainties in snowpack properties.

M2- The ERA Interim reanalysis and the MK3Lv.1.2 model are now outdated. ERA5 and ERA5-Land are state-of-the-art reanalyses with better spatial resolution. And the same can be said about paleosimulations included in the PMIP4/CMIP6 projects. Why were these old products selected as forcings of the simulations? As the authors indicate in the manuscript, the spatial resolution of the forcings is one of the aspects affecting the performance of the CryoGridLite model, and those forcings need to be interpolated before running the model. Would not make more sense to use new, high resolution renalyses and paleosimulations for this study?

We appreciate the reviewer's comment regarding the use of older forcing data in our study. While we acknowledge that more recent reanalysis and paleo simulations exist, our study was based on ERA Interim and the MK3Lv.1.2 forcing data due to their availability within our data framework and their widespread acceptance when we started developing CryoGridLite. We understand that it is desirable to use the most recent forcing data, but we would like to stress that new forcing data do not necessarily invalidate the applicability of older data. For instance, potential biases have been reported for ERA5-Land data in the Arctic (Cao et al., 2020), while applicability of ERA Interim has been validated in numerous studies.
Our study aims to demonstrate the capabilities of our modeling approach to perform long-term permafrost simulations that also consider highly uncertain and variable soil and snow parameters using large ensembles. We agree that future studies focusing on investigating the response of permafrost within specific climate periods or related to specific climatic events in the past should use state-of-the-art forcing data. However, for this study, which focuses on model applicability and validity, we think that the use of older forcing data is acceptable. Preprocessing and repeating all ensemble simulations would be a very excessive task and using a different forcing will not change the broad scale thermal history. Nonetheless, we appreciate the reviewer's critique, and we will explicitly mention in the revised manuscript that more recent forcing data should be used in follow-up studies focusing on model application rather than general model validation.

Cao, B., Gruber, S., Zheng, D., and Li, X.: The ERA5-Land soil temperature bias in permafrost regions, The Cryosphere, 14, 2581–2595, https://doi.org/10.5194/tc-14-2581-2020, 2020.

M3- After reading the two first sections, it is clear that the main advantage of the CryoGridLitle is the speed of execution in comparison with more comprehensive permafrost medelling schemes. To this end, many important processes are not included in the model, such as water advection or runoff. Nevertheless, there is no estimate about the performance gain of CryoGridLitle in comparison with other schemes. Comparing the time required to simulate the 1400 years of this study in the CryiGridLitle and in another scheme (maybe CryoGrid3?) would allow the reader to assess if the gains in execution time really compensate for the absent processes.

We agree with the reviewer. We performed a runtime comparison and added it to the relevant paragraph (see response above). In addition, we have performed a thorough performance test of the implicit solution scheme and added the following description to the method section, including an additional figure:

The numerical performance and stability of the applied implicit method is evaluated in comparison with commonly used numerical integration methods: Crank-Nicolson and TRBDF2 (2nd order diagonally implicit Runge-Kutta methods), Radau IIA5 (5th order fully implicit tableau method), Heun (canonical explicit trapezoidal method), and SSPRK43 (4th order stability-preserving explicit Runge-Kutta method). The test simulations under periodic freeze-thaw conditions result in a mean absolute error of about \qty{0.014}{K} for time steps of 24 hours. With this acceptable level of accuracy, the CryoGridLite scheme is about an order of magnitude faster than the other numerical integration schemes.

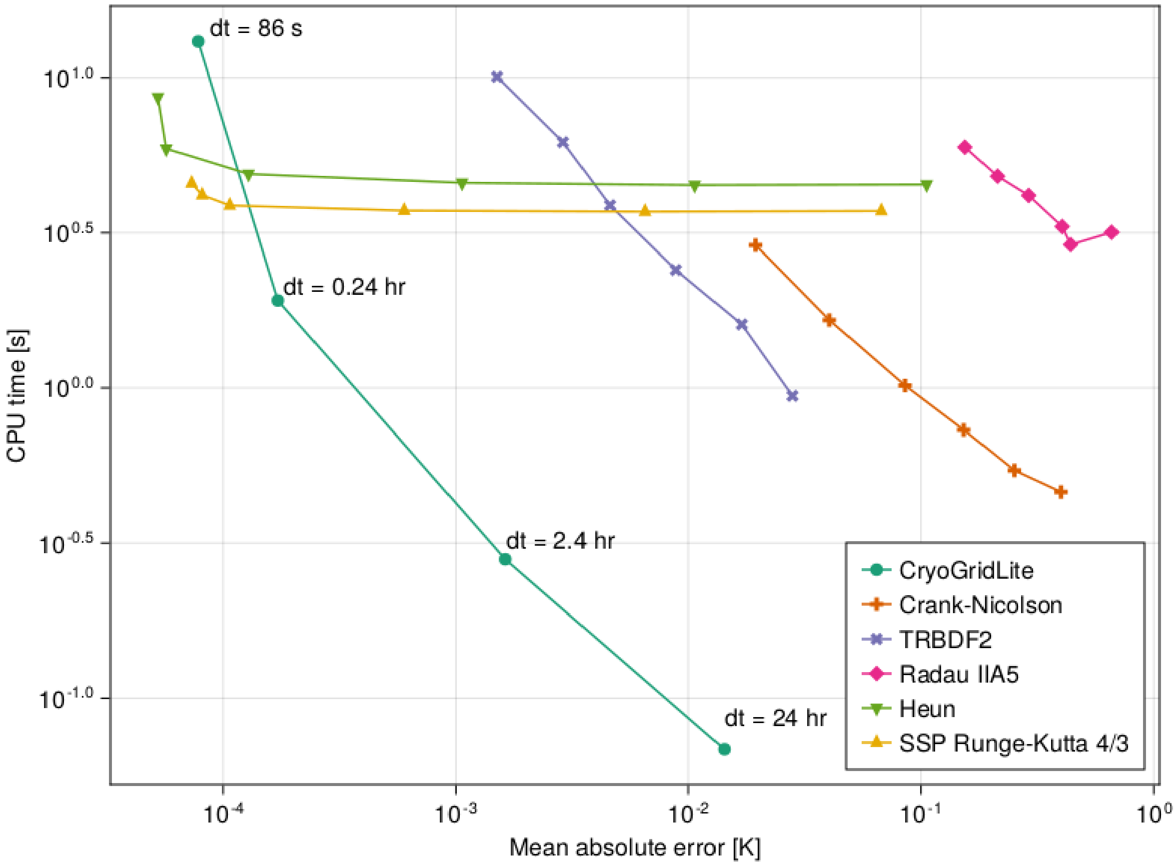

Fig. xyz: Work-precision plot comparing different numerical integration methods with CryoGridLite on heat conduction with phase change with periodic upper boundary. The y-axis shows the computational demand in CPU time of each numerical solve while the x-axis shows the global temperature error of solution with respect to a high-resolution reference simulation using a 10-stage, 4th order strong stability preserving (SSP) Runge-Kutta with a 60 second timestep. All of the algorithms (except CryoGridLite) use adaptive time stepping schemes with tolerances changed from 10^-6 to 10^-1. Since CryoGridLite is a fixed-time step method, we instead use a timestep of 24 hours scaled from 10^-3 to 10^0 for comparison. With the largest timestep of 24 hours CryoGridLite achieves an acceptable global error of approximately 0.014 K while being an order of magnitude more efficient than all other methods.

M4- The authors should consider the possibility of adding some theoretical tests to evaluate the ability of the model to simulate physical processes, such as propagation of heat with different thawing conditions. I am aware that the CryoGrid3 model is able to successfully reproduce such processes, but since the formulation for heat propagation in the CryoGridLitle model has been modified, there should be some proof that the model is able to correctly reproduce phase change and basic heat diffusion though the ground.

We thank the reviewer for this recommendation, which we have included in the revised version of the manuscript. In the appendix we will present a theoretical test against an analytical solution for heat transfer without phase change that shows the ability of the model to correctly reproduce heat diffusion. We will also show a test against the Stefan solution that demonstrates the ability of the model to correctly simulate the progression of the thawing front.

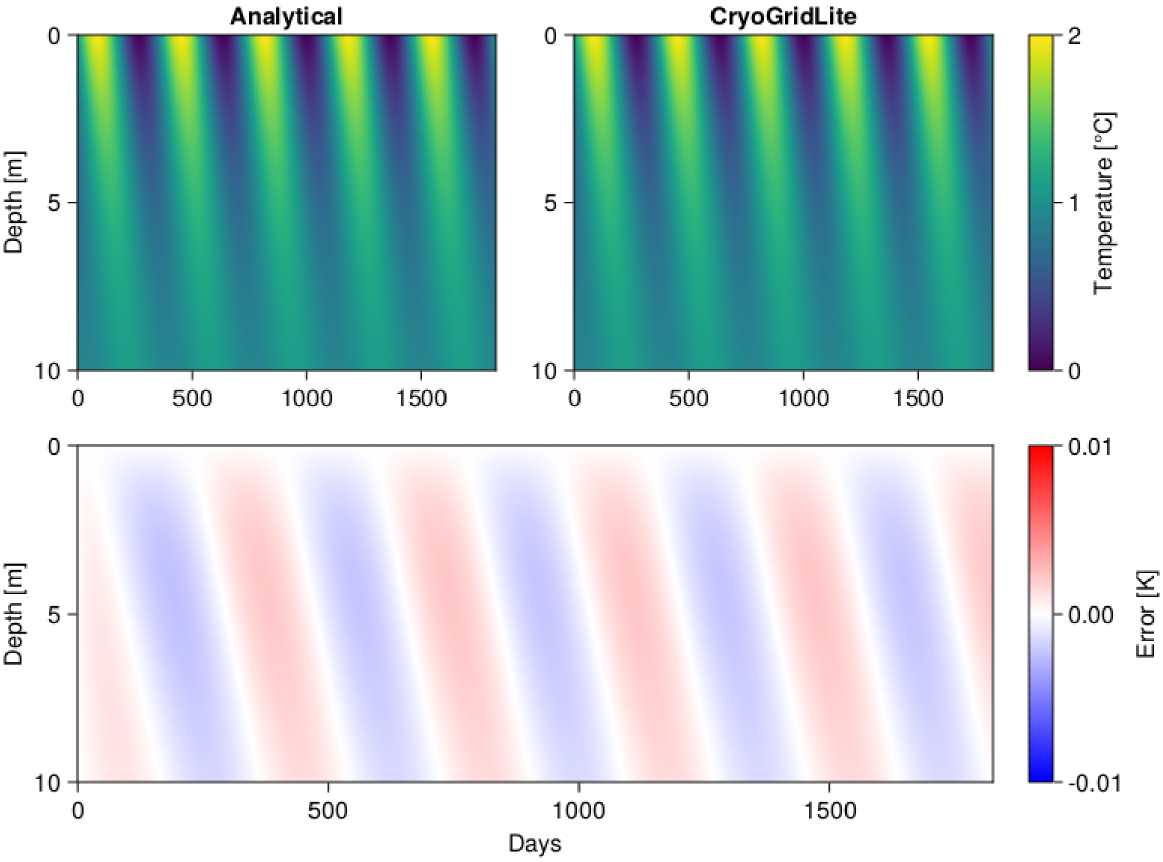

Fig. xyz: Evaluation of CryoGridLite in simulating heat diffusion with sinusoidal upper and zero flux lower boundary conditions without phase change. The analytical solution (a) for a semi-infinite half-space is compared to the solution obtained using CryoGridLite (b). Analysis of the differences (c) reveals a relatively small periodic bias of less than ±0.01K at deeper depths due to the spatial discretization. These results demonstrate that the numerical scheme effectively captures the heat diffusion process, yielding results that closely align with the analytical solution.

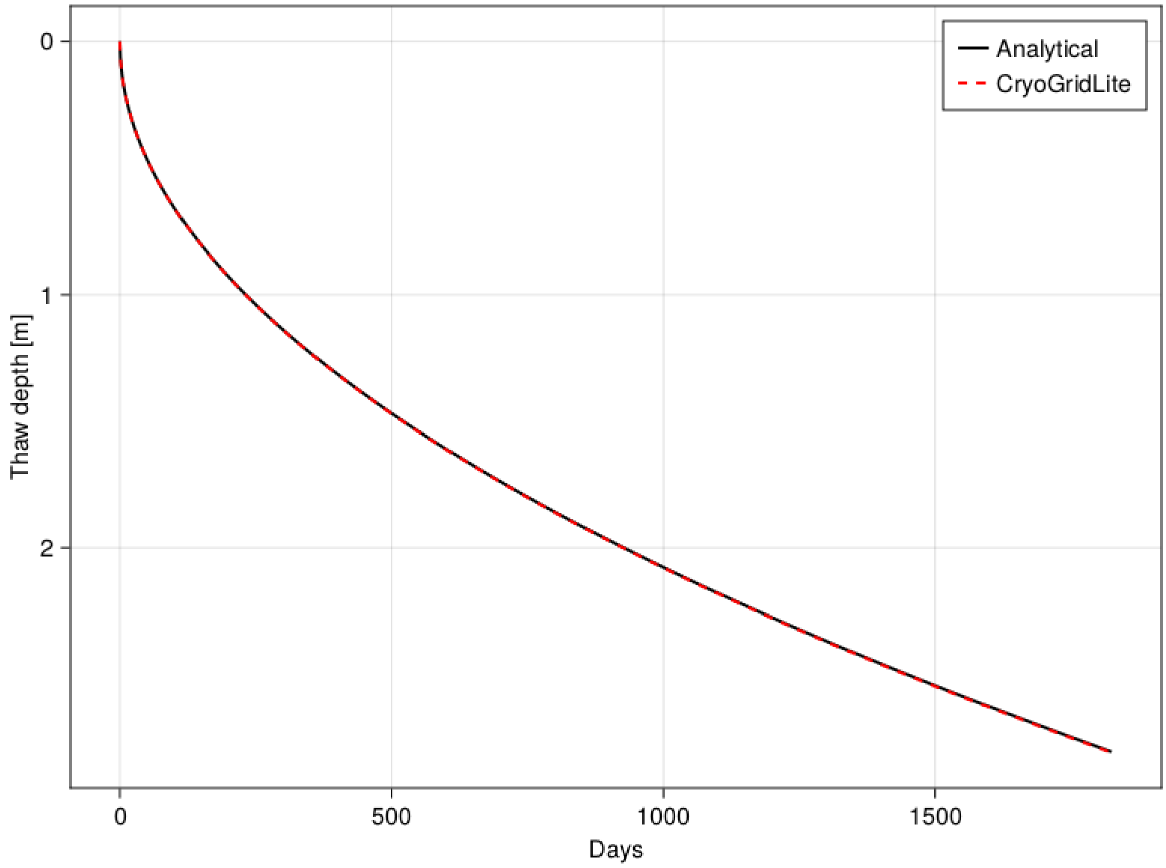

Fig. xyz: Comparison of the analytical solution to the two-phase Steefan problem and CryoGridLite for the one-sided thawing of a frozen soil column. The soil temperature is initialized at -0.02°C, and a constant upper boundary temperature of 1°C is applied for a period of 5 years. The spatial domain is discretized into a nonuniform rectangular grid cell with minimum spacing of 0.01 m down to a depth of 0.5 m. The good agreement between the analytical solution and the CryoGridLite simulation demonstrates that the model is able to accurately represent phase change of water.

M5- Another concern is the evaluation of the model against measurements of soil temperature and active layer thickness (ALT) at CALM sites. I wonder if evaluating soil temperatures and ALT from CALM is enough to "validate" the model. The CALM network is indeed a very valuable resource for model assessment, but it is also limited geographically and in time. Otherwise, there are several estimates of near-surface permafrost area in the present, some of the relevant articles have been cited in the manuscript, and several indices have been developed in order to use meteorological data as a benchmark for evaluating permafrost models (e.g., 1-2). Then, why is the evaluation of CryoGridLite restricted to comparing with CALM stations? At the very least, a comparison with estimates of permafrost area during recent times should appear in the article, as this is a relevant factor to evaluate the quality of the simulations.

We appreciate the reviewer's comment and would like to clarify that while we are validating the model for Active Layer Thickness (ALT), we are also validating it for thermal state representation using borehole temperature measurements. Unlike other modeling studies, we are not limiting our validation to temperature alone but also testing the reproducibility of trends in the soil thermal regime. However, we agree that further validation can be done to improve the model's robustness. As suggested by both reviewers we have collected additional data, especially from long-term borehole temperature measurements, to improve our model's validation. Furthermore, we will also validate the model's ability to reproduce recent permafrost extent as suggested by the reviewer.

A more detailed model evaluation is being conducted using borehole measurements, where long-term data spanning over 10 years are available for selected sites. Air and ground temperature observations from four sites in Canada (Salluit, Quaqtaq, and Ellesmere Island) \citep{allard2020borehole} and Russia (Urengoy) \citep{smith2022} are being obtained and compared to the corresponding model data from the nearest grid cell within the model domain. For the sites in the Quebec region (Salluit and Quaqtaq), the measured air

temperatures systematically show a difference of about \qtyrange{1}{2}{K} compared to the ERA interim forcing. This can be well explained by small scale variability e.g. due to orographic effects or proximity of the measurement sites to the coast. For both sites, the observed mean annual ground temperatures (MAGT) at \qty{10}{m} depth during the 2000s and 2010s are very well within the modeled ensemble range and even agree well with the model mean, especially for Quaqtaq, suggesting that the model ensemble captures site-level conditions despite a bias in the forcing data. For a depth at \qty{20}{m}, additional borehole observations are available for the late 1980s and early 1990s \citep{allard1995recent}. The time series of measurements, however, is not long enough to confirm the pronounced negative temperature anomaly simulated before 1980 for Quaqtaq. For a high Arctic site on Ellesmere Island (Fig. Appendix c), the long-term observations and simulations also appear to agree reasonably well, although the natural variability is not fully captured by the model. The long-term measurements from Urengoy (Fig. Appendix d) in Russia show very cold borehole temperatures, and only the coldest simulations in the ensemble show such low temperatures. Another borehole is located in the immediate vicinity which shows much warmer temperatures but has a shorter time series.

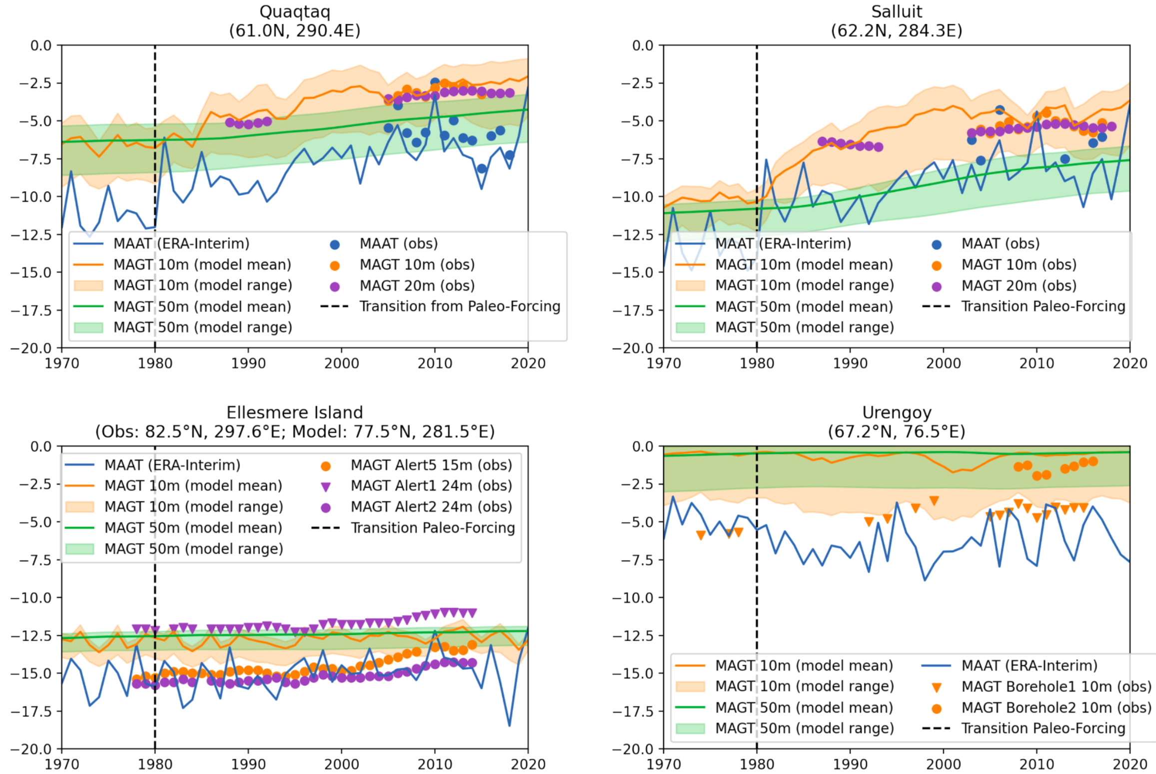

Figure Axyz: Model evaluation based on long-term borehole temperature and near-surface air temperature measurements in different regions. (a) Quaqtaq, (b) Salluit in the Quebec region, (c) a high Arctic region on Ellesmere Island, and (d) Urengoy in northern West Siberia, Russia. If multiple borehole measurements are available, they are indicated separately.

M6- Lines 309-319: There are some works providing ground surface temperature histories from inversions of deep subsurface temperature profiles in the North American part of the Arctic (e.g., 3-6). Nevertheless, the authors choose proxy-based paleoreconstructions of surface air temperature to compare with permafrost temperature changes in Alaska. Inversions of subsurface temperature profiles are more adequate for this comparison, since these are estimates of changes in ground surface temperature.

We thank the reviewer for this comment and have gratefully incorporated the paleoreconstructions of surface temperatures into the discussion of the revised version:

For the Québec region observations and model results agree well for the L20C period (Fig. Appendix a,b). However, we note that in the model forcing data before 1980 are significantly colder than those after 1980. This temperature shift is attributed to the applied paleo-temperature forcing. In particular, the cooling trend reconstructed by \citet{chouinard2007recent} for the decades around 1950 is not clearly visible in our forcing

data. Nevertheless, the modeled temperature anomalies (Fig. \ref{fig:maps-MAGT}) indicate that the L20C period is much warmer than the L19C period. This finding is consistent with the observations and supported by regional climate reconstruction \citet{chouinard2007recent} revealing a longer-term warming trend in northern Québec starting in the 18th century.

Minor points

m1- The title should be changed to reflect the model evaluation part of the study. Something like "The evolution of Arctic permafrost over the last three centuries using a new, fast permafrost model".

We appreciate the suggestion and modified the title which now reads: "**The evolution of Arctic permafrost over the last three centuries from ensemble simulations with the CryoGridLite permafrost model**"

m2- Lines 103-106: It is not clear how the model can produce deviations larger than the threshold and still conserve energy. Could you please give more detail about this point?

The model employs the energy form to solve the heat equation, from which the temperature profile is derived. The heat flux within the soil profile is determined using a tridiagonal solver that calculates the redistribution of heat by providing the gradients based on which the energy is distributed. This iterative procedure continues until the solved temperature profile reaches equilibrium with the temperature profile diagnosed from the energy profile. However, due to the discontinuity caused by the freezing curve of water, there may be discrepancies between the temperature profile and the energy profile. In most integration steps, this temperature discrepancy is below the selected threshold (<0.01 K), but in some cases, satisfactory conversion cannot be achieved within an acceptable number of iterations (500). This only affects three grid cells around which the actual freezing front is located and occurs only if a cell is almost completely thawed/frozen and the thawing/freezing front is transferred to the neighboring cell. This discrepancy only indicates the energy is not exactly distributed as expected from the solution of the temperature profile. The total energy content resulting from the fluxes across the boundaries and the previous heat content remains conserved for each time step.

m3- The definition of 1850-1900 as preindustrial period is not consistent with the consensus in paleoclimate publications. Indeed, the period 1750-1800 is a better option for preindustrial times (3). Please, consider a change of labelling, perhaps with 1850-1900 as later 19th century (L19C).

We thank the reviewer for this remark and have changed the labeling accordingly throughout the manuscript.

m4- I am unable to see the mentioned short-term changes in permafrost area due to volcanic eruptions in Figure 6. Is there any other way of presenting the results of Figure 6 that shows more clearly the effect of the eruptions?

We agree with the reviewer. The effects of volcanic eruptions are not well visible in the figure. Also, in response to a comment from reviewer 1 and a suggestion from the initial editorial review, we have created a new diagnostic to evaluate the presence of near-surface permafrost. This now clearly shows the effects of volcanic eruptions on permafrost occurrence probability.

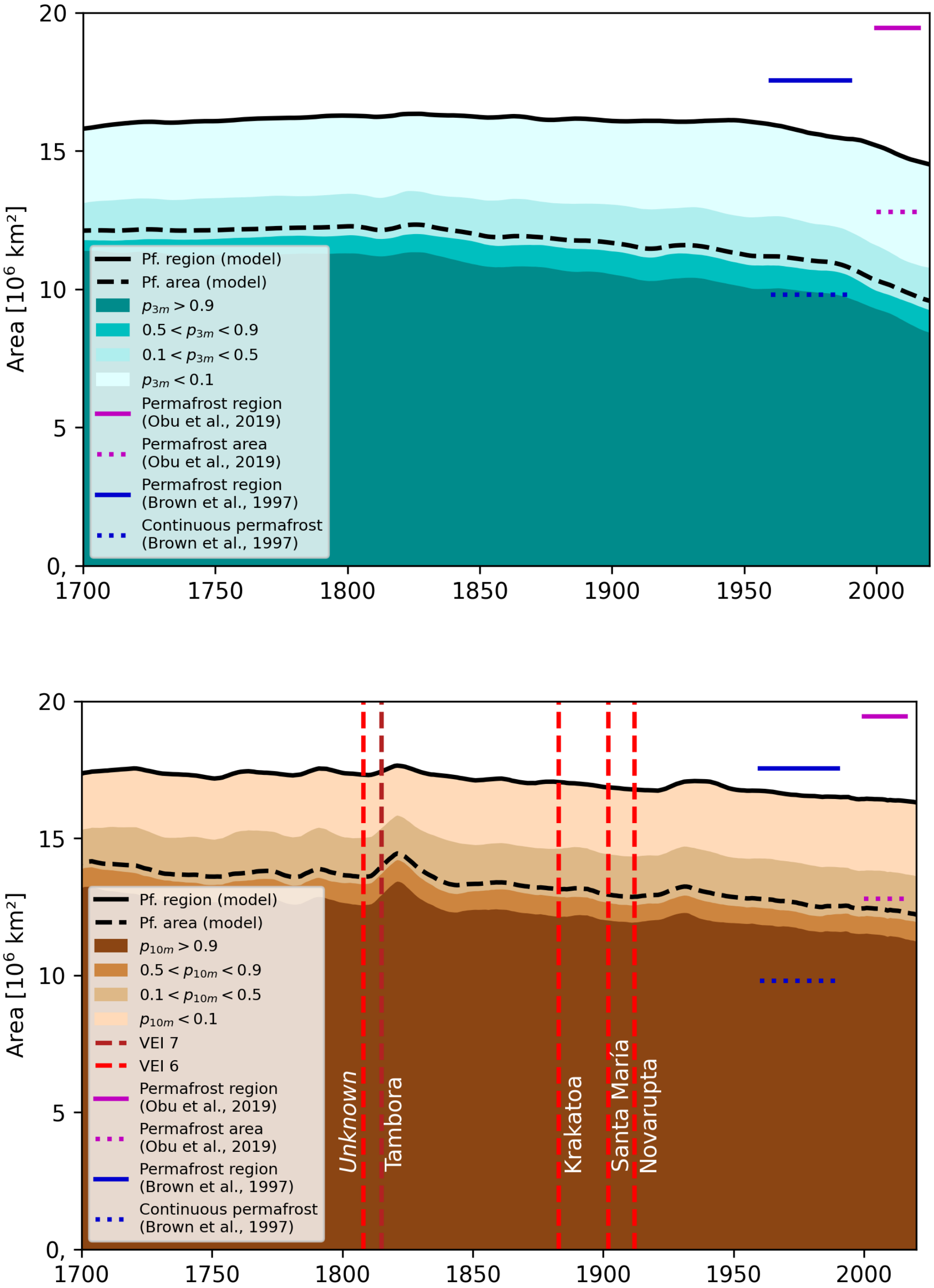

Figure 6. Total areas of probability of near-surface permafrost occurrence according to two different diagnostics focusing on (a) 3 m and (b) 10 m ground depth. The zones of occurrence probability are derived from the parameter ensemble simulations. Horizontal lines mark areas of permafrost regions as delineated by Brown et al. (1997) and Obu et al. (2019) for the early 21st century. Dashed red vertical lines mark strong volcanic eruptions events (Volcanic Eruption Index (VEI) ≤ 6) which are represented in applied climate forcing data.

References

1- Koven, C. D., Riley, W. J., and Stern, A.: Analysis of Permafrost Thermal Dynamics and Response to Climate Change in the CMIP5 Earth System Models, Journal of Climate, 26, 1877–1900, https://doi.org/10.1175/JCLI-D-12-00228.1, 2012/10/01.

2- Burke, E. J., Zhang, Y., and Krinner, G.: Evaluating permafrost physics in the Coupled Model Intercomparison Project 6 (CMIP6) models and their sensitivity to climate change, The Cryosphere, 14, 3155–3174, https://doi.org/10.5194/tc-14-3155-2020, 2020.

3- Majorowicz, J. A., Skinner, W. R., and Šafanda, J.: Large ground warming in the Canadian Arctic inferred from inversions of temperature logs, Earth and Planetary Science Letters, 221, 15–25, https://doi.org/https://doi.org/10.1016/S0012-821X(04)00106-2, 2004.

4- Taylor, A. E., Wang, K., Smith, S. L., Burgess, M. M., and Judge, A. S.: Canadian Arctic Permafrost Observatories: Detecting contemporary climate change through inversion of subsurface temperature time series, Journal of Geophysical Research: Solid Earth, 111, https://doi.org/https://doi.org/10.1029/2004JB003208, 2006.

5- Taylor, A. E. and Wang, K.: Geothermal inversion of Canadian Arctic ground temperatures and effect of permafrost aggradation at emergent shorelines, Geochemistry, Geophysics, Geosystems, 9, https://doi.org/https://doi.org/10.1029/2008GC002064, 2008.

6- Jaume-Santero, F., Pickler, C., Beltrami, H., and Mareschal, J.-C.: North American regional climate reconstruction from ground surface temperature histories, Climate of the Past, 12, 2181–2194, https://doi.org/10.5194/cp-12-2181-2016, 2016.

7- Hawkins, E., Ortega, P., Suckling, E., Schurer, A., Hegerl, G., Jones, P., Joshi, M., Osborn, T. J., Masson-Delmotte, V., Mignot, J., Thorne, P., and van Oldenborgh, G. J.: Estimating Changes in Global Temperature since the Preindustrial Period, Bulletin of the American Meteorological Society, 98, 1841–1856, https://doi.org/10.1175/BAMS-D-16-0007.1, 2017.

Citation: https://doi.org/10.5194/egusphere-2022-473-RC2