In their manuscript, García-Pereira et al. perform historical and future simulations with the MPI-ESM, which was modified to account for a dry and wet configuration of Arctic soils, as well include various model soil depths. With the help of these simulations, the authors investigate the differences of these setups for soil surface temperature, active layer depths and permafrost extent for past and future scenarios. The work is an important effort and relevant for the journal Cryosphere. I believe the work is close to publication, but I have a few major comments that, in my opinion, should be addressed:

- I believe the manuscript can be somewhat stream-lined to stress novel results from this specific study. Some of the extensively discussed points are already addressed in other papers (e.g. aspects of the importance of freeze-thaw processes and soil hydrology for the Arctic heat and water balance). There is also some very technical information on model versions and reasoning for these changes since MPI-ESM of CMIP6 simulations (which is not applied here), which I believe is maybe not that interesting to readers not directly involved with the model (and would be anyways better fitted for a model development paper). I would probably just focus on what is present in the model versions presented here.

- Inclusion of observational constraints: The authors compare modelled permafrost extent to an observational-derived product (with moderate/poor agreement for both dry and wet configurations), but do not compare soil temperatures and active layer depths shown here with similar available products. I recommend consistently including ground temperature and active layer thickness observational products as well (https://climate.esa.int/en/projects/permafrost/), to better grasp the model performance in simulating soil temperature and active layer depths. Since these datasets are also based on thermal models and limitations should probably also be briefly mentioned.

- The simulation setups presented are idealized to potentially represent dry and wet soils, and it is unclear which set up is actually more realistic to represent the Arctic as a whole, both from an observational standpoint as well as in terms of process representation. Since there are both relatively dry and wet areas in the Arctic, would it be possible to combine the results from the wet and dry setup offline, using wetland coverage maps to weight every grid cell to give an estimated combined of permafrost thaw in the future?

- A bit of an open question: the authors stress the differences in permafrost extents by the different approaches/simulation setups, which I think is fair to grasp a magnitude of impact that can be communicated as a single number. However, the differences in permafrost extent likely are differences of areas of very deep permafrost (I assume mostly >10m depth?), which would be probably disconnected from surface fluxes and at which depths I would assume that mostly bedrock or mineral soil is found. Is the thawing in these depths really relevant for a discussion on climate impacts of permafrost on surface hydrological feedbacks for the climate and carbon degradation? In my opinion the importance deepening of close-to-surface carbon-rich layers (such as the Yedoma domain) should be stressed, even if this is not reflected in a full permafrost retreat of the soil column.

- It is not completely clear to me from the abstract and conclusions, what the final message was regarding the effects of the difference model soil depth configurations on the trends of averaged active layer thickness and PE. Looking at the effects of various model soil depths in Figure 4, 6 and Figure 10, this effect seems small (and probably not worth it in terms of additional computational expenses for coupled models?).

- In terms of significance, would it be worth to include compute back-of-the-envelope calculations implications for carbon release from the permafrost (e.g. from typical spatial carbon contents?).

Minor Comments

L24 -> nearly four times ?

L48 I think PE is not defined anywhere.

L70 yedoma -> Yedoma

L74 Also the QUINCY model

Lacroix, F., Zaehle, S., Caldararu, S., Schaller, J., Stimmler, P., Holl, D., Kutzbach, L., & Göckede, M. (2022). Mismatch of N release from the permafrost and vegetative uptake opens pathways of increasing nitrous oxide emissions in the high Arctic. Global Change Biology, 28, 5973–5990. https://doi.org/10.1111/gcb.16345

L118 "An enhanced vertical resolution accounts for a better representation of hydro-thermodynamic processes near the surface (Chadburn et al., 2015)." -> Reference to your Table 1?

L188 "In case this depth does not specifically match a certain mid-layer depth value, JSBACH yields ALT using linear interpolation."

When would this be the case? I would imagine this would apply if a layer is partly frozen, but the authors explicitely define the >273.15 °C definition.

L293 Define winter offset in the text and why it is important.

L301 "Both results reflect weaker insulation in the standard snow model."
Is this because of different parameterizations of the snow model (heat constants?), or added processes?

Wouldn't it be possible to back this up with data from the simulation?

Figure 1. I know this is a conceptual figure, but the bedrock seems very deep? The bed rock in the Arctic is usually found around 1-2 m (unless in valleys).