Review of „Representation of soil hydrology in permafrost regions may explain large part of inter-model spread in simulated Arctic and subarctic Climate“ by Philipp de Vrese et al. submitted to *The Cryosphere:*

The article by Philipp de Vrese et al. presents a set of coupled simulations using the MPI Earth system model in which parameterizations of the soil hydrology within the permafrost region have been varied. The effects of these variations on land-atmosphere fluxes in the permafrost region are investigated and also effects on the global climate are discussed. The authors compare the variation among the presented simulations to variations among the projections of CMIP6 models and interpret that a the inter-model spread in key climate variables of the high latitudes can be attributed to differences in the models' parameterizations of soil hydrology in areas affected by permafrost.

The question of whether the high latitudes become wetter or drier under future climate change is unanswered and at the same time highly relevant for Arctic climate feedbacks such as the permafrost carbon feedback. While the present article does not answer this question, it does provide new insights which help to identify sources of uncertainty of climate model predictions and points to assumptions and parameterizations in current model which require more attention and careful consideration by climate modellers. Thus, the article is very relevant within the scope of TC and strongly builds on works which have previously been published in this journal.

The methodology of the study is explained well and adequate to address the research objectives. The results are described clearly and mostly explained in an understandable way. The overall presentation quality is good, but some figures should be revised (see comments below). Less clear to me was the distinction between Results, Discussion, and Conclusions which appear to be mixed to certain degree. Here, the paper would benefit from restructuring and rebalancing of the presented contents (see comments below).

Overall, the paper addresses a relevant topic, the results are substantial and presented well. In principle I support its publication in TC. Before the paper can be accepted, however, the following comments need to be taken into consideration by the authors:

General comments

- A main point of the study is, that different parameterizations of the soil hydrology in permafrost regions could explain inter-model spread in climate variables. The authors explain in detail how the soil hydrology is parameterized in MPI-ESM and how they modified it to come up with two contrasting but plausible setups (WET and DRY). However, the authors do not establish a connection to the respective parameterizations in other state-of-the-art land surface schemes. In order to not only quantitatively (e.g. Fig. 6), but also qualitatively substantiate the explanation of inter-model spread due to soil hydrology parameterizations, it would be very helpful if the authors would (i) introduce briefly typical assumptions and parameterizations of the soil hydrology in the current generation of LSMs (in the Introduction; an overview Table could be helpful for this), and (ii) discuss the findings using the two setups of MPI-ESM in relation to typical setups of other land surface schemes in CMIP6 models (in the Discussion section).

- The description of the modifications to the model (section 2.1) and its setups (section 2.2) would benefit a lot from a schematic illustration of the soil hydrology processes and their parameterizations. For example, such a figure could illustrate the differences between runoff, infiltration, drainage, and how these fluxes are affected by the state of the surface and subsurface. Such a schematic figure might also be used to illustrate the different assumptions made for the WET and DRY setups.

- I am wondering how useful the approach of applying the modified parameterizations only to a pre-defined sub-region of the land model domain (the present-day permafrost extent based on Hugelius et al. 2014) is. The biogeophysics and biogeochemistry follows universal laws such that it should be possible to come up with general parameterizations which can be applied over the entire land domain and not only in an pre-defined region whose borders are set "arbitrarily". While I see that the approach is sufficient to address the research questions of this study, it brings a lot of shortcomings with it. For example, it introduces additional inter-model uncertainties due to the definition of the region which makes inter-model comparisons hard. In addition, the definition of the region is based on a present-day snapshot of permafrost occurrence which in turn is dynamic and would be very different under past or future climate conditions. I would appreciate a discussion of this issue and if the authors could line out a way towards a more sustainable modelling approach for modified permafrost representations.

- The authors conduct and present a set of simulations for which the climate reaches stabilization at an GMST increase of 1.5° the results of which I find very insightful (Fig. 10). In fact, these are two coupled simulations for the WET and DRY setup, respectively, which are using overshoot scenarios in addition to the two simulations under SSP5-8.5. Therefore, I find it rather surprising that the results of these simulations are presented within the "Conclusions" section (which reads more like a Discussion section). It would help the overall structure and accessibility of the article, if (i) an appropriate research objective would be provided in the Introduction which motivates these simulations, (ii) the simulation results would be presented in a dedicated Results section which addresses that research question.

- I appreciate the effort of the authors to also conduct numerical experiments with ICON-ESM to substantiate their findings. However, I am not convinced that the way this is currently presented in the article is helpful for the readers. The ICON-ESM simulations are described in a dedicated Methods section 2.4, but not mentioned thereafter, neither in the results, nor in the discussion or conclusions. If it is a specific objective of the authors to substantiate their findings also with the simulations of ICON-ESM, then these should also be mentioned in the Results and Discussion sections. Otherwise, the authors might consider to omit presenting the ICON-ESM simulations at all in the present study.

- The final section which is labelled "Conclusions" provides mostly a discussion of the results, and in addition introduces further model results obtained using overshoot scenarios. However, the section does not present any concrete conclusions of the present work in a comprehensive and condensed way. I strongly suggest to carefully revise the structure and contents presented in this section. Connecting to my comment regarding the 1.5° simulations above, I suggest that these results should be presented in a dedicated subsection of the "Results" section. All results (land–atmosphere interactions, comparison to the CMIP6 spread, (optionally the ICON-ESM comparison), and the comparison of the overshoot runs) could then be discussed in a dedicated Discussion section building on the current "Conclusions" section. An additional section which provides the study's conclusions in a condensed way would further improve the article's accessibility and impact.

- The Discussion of the results does not sufficiently consider the parameterization of soil hydrology in other land surface schemes (see first comment above). An additional paragraph in which the present work with MPI-ESM is related to other model configurations could help to substantiate the findings regarding the inter-model spread also qualitatively.

Specific comments

- l.109ff: The authors state that the model representation of soil physics is largely based on the modifications by Ekici et al. and explain the differences. However, it was not clear to me how the model version used relates to that described in de Vrese et al. (2021) where several modifications to the soil physics and biogeochemistry have been described. Maybe I overlooked this, but are these modifications also included in the model version used for the present study?

- l.124f: "but includes their influence on the hydrological soil properties." Does this refer to the modifications described in sections 2.1.2.-2.1.4 or are there further ways in which the organic matter influences the hydrological soil properties (e.g. hydraulic conductivities). If the latter is the case, all equations should be provided in the main text or the appendix. Otherwise it should be made clear, that this statement refers to the modifications described later on.

- Section 2.1.1. mentions a lot of issues and problems which faced by soil hydrology parameterizations, in particular in JSBACH. However, it is not always clear which of these issues are being addressed in the present study and which are not.

- l.158: It is not clear how infiltration, especially into frozen ground, is controlled in the current version. The authors only state that it is "exclusively controlled by the saturation of the near-surface soil layers ...", but they do not provide any equations or references to other work where this is described.

- l.225: The authors state that the field capacity constitutes the upper limit of the soil water content in JSBACH, which I found confusing. I would have expected that it is possible to have soil water contents exceeding the field capacity, for example under saturated conditions. For example, in l.248 it is written that "only the water that exceeds the layer's field capacity is added to the drainge flux" which in my view contradicts the formulation in l.225. Am I missing something here? A clarification would be appreciated.

- l.292: It is not clear to which version „JSBACH's original formulation“ refers. It is stated that the original formulation would prohibit infiltration at sub-zero temperatures, but I understood that this had been only introduced into the model by Ekici et al. Also in light of my comment below on using the term "reference model", it might be helpful to clearly designate the different JSBAHC model versions/configurations referred to in this paper.

- l.314ff: It is not clear how the subsurface parameters have been set, since the deep model configuration deviates from the standard one. I suspect that the same datasets as in González-Rouco et al., 2021 and Steinert et al. 2021a have been used? This should be clarified in order to allow reproducibility.

- l.414: Why not stating the proportion of simulations contained in two standard deviations (~95%) if these are used?

- l.422/Fig. 5: Here and later the authors often refer to "trends" which are shown in Fig. 5. How have these trends been estimated? (linear regression?, which time period?, ...). Please specify this, ideally already in the Methods section.

- l.433ff: I do not understand the argumentation regarding the effect of excess ice. I would expect that, in a first-order approximation, the inclusion of excess ice and associated ground subsidence would result in wetter conditions if thaw depths increase, since the ground subsidence counteracts the possibility of deeper infiltration (see Fig. 8 in Andresen et al. (2020)).

- The final paragraph of the results section (l.538ff) introduces an additional simulation setup (W2D). Admittedly, I had a hard time understanding this paragraph. This part would benefit from a bit more detailed explanations of the motivation for and the interpretation of the additional simulations. It might also help to replace some of the rather technical language which uses symbols as parts of the sentences by verbal descriptions.

- l.589: It might be interesting to explicitly state the high latitudes become warmer under the DRY relative to the WET setup while the tropics and mid latitudes are colder under the WET setup relative to the dry setup.

- l.603f: Please provide references for the statement regarding the effect of sea ice.

- l.615: It should be stated in the text that the subsea permafrost was diagnosed using a different model configuration.

- l.624: It is not clear what "negatively affected" means in this context.

- l.662: The work of Aas et al. (2019) might also be relevant in this context.

- l.664f: It would be interesting to discuss whether ICON-L/JSBACH4 provide this functionality.

- To my opinion, Figure 3 is not crucial for understanding the methodology and should go into the appendix or supplement.

- The caption of Fig. 4 refers to a "reference model" abbreviated REF, which is, however, not introduced at any point in the main text. It is not clear whether this is the standard configuration of MPI-ESM/JSBACH3.2 or whether it includes the modifications by Ekici et al. Please clarify.

- Fig. 5/6: While Fig. 5 shows the trends in total water content, Fig. 6i shows the aggregated changes in liquid water content. However, it would also be interesting to see how the absolute values of total water content evolve during the 21st century and how the two setups compare to each other to get the full picture.

- Fig. 6: The differences in permafrost volume (panel h) are quite huge among the two setups. It would be interesting to also show the simulated permafrost areal coverage and mention how it compares to estimates for the present day coverage. Are other estimates contained within the spread between the WET and DRY scenarios?

- Fig. 7: I do not understand why this figures uses a mix of "area" and "line" plot features. It would make more sense if the IQR and 2sigma would be indicated as areas, and the Delta as lines. But showing all quantities as a line plot should also work.

- Fig. 7/8: To be consistent and to provide the full picture, you might consider to include the plots and values for the region north of 50°N also for precipitation and evapotranspiration.

- Fig. 9: The scenario (SSP5-8.5) should be mentioned in the caption.

- Tab. 2 should go into the supplement.

Technical corrections

- l.29: Should be "as contained in the Earth's atmosphere."

- l.137: "be" missing in "could derived from observations."

- l.150: "see below" --> Please be more specific and refer to specific sections.

- l.157: typo: "soil" instead of "seoil"

- Fig. 5: I don't see why you write "EVAP.", "PR." and "H20" instead of "Evaporation", "Precipitation" and "Total soil water" in the heading.

References

Aas, K. S., Martin, L., Nitzbon, J., Langer, M., Boike, J., Lee, H., Berntsen, T. K., & Westermann, S. (2019). Thaw processes in ice-rich permafrost landscapes represented with laterally coupled tiles in a land surface model. The Cryosphere, 13(2), 591–609. https://doi.org/10.5194/tc-13-591-2019

All other references are contained in the manuscript.