We thank the reviewer for evaluating our manuscript, and for the numerous comments which help to improve it. We provide answers to the comments below.

Reviewer comments are in *blue italics*.

Author replies are in normal font.

Extracts from the manuscript are in **bold**, and modifications in the revised manuscript are **highlighted yellow**.

Anonymous Referee #2

Received and published: 24 August 2020

Authors implement micro- and meso-scale spatial grid resolution in the 1D CryoGrid model to illustrate the spatial effect of microtopographic feature on the rate of permafrost thaw. Authors found that implementing higher spatial resolution in the model leads to "more realistic possibilities (L13)". Now sure what type of possibilities they have in mind? Improving spatial representation of the polygonal tundra in the ESM type models is important. However, the current version of the manuscript lacks clarity. I found it hard to follow and the central Figure 2 looks like an electrical circuit diagram. If the take-home message is that very ESM needs to have micro-, meso-scale permafrost tundra representation, then it needs to be clearly stated. Maybe including recommendations on how authors think that can be done easily, in their opinion, using the current approach.

Overall, this is a timely and important work that needs to be published. However, the description, terminology, and flow require more work. I have a hard time reading and understanding the concept laid in the paper. I understand that much of the tiling concept was introduced in previous work (Aas et al.). However, the recap could be extremely helpful in setting up the stage in this study. Also, talking about uncertainties between different tiling approaches might be useful too. For example, if we average the overall effect from individual polygons, it could have the same carbon footprint as representing the polygonal tundra heterogeneity in one tile. When could that be or not be true? The comments below illustrate the lack of my knowledge of the presented scaling method. I hope the authors would not be discouraged by my comments and try to help me better understand their work in the revised version of the manuscript.

We appreciate that the reviewer acknowledges the importance and timeliness of our work. From the comments we understand that the main criticism of the reviewer is, that the description of our methodological approach lacks clarity and that the implications of our work with respect to large-scale models are not sufficiently explained. We think that these issues could be related to the fact that we presupposed that readers would be familiar with preceding works in which the concept of laterally coupled tiles was applied in a permafrost context (Langer et al., 2016, Aas et al.,2019, Nitzbon et al., 2019, and Nitzbon et al., 2020). In these papers, the concept is introduced and applied to investigate different subgrid-scale processes in permafrost ecosystems. In particular, Nitzbon et al. (2019) introduced a three-tile setup for ice-wedge polygonal tundra and evaluated it using field observations from Samoylov Island in the Lena River delta. In the present study, we use that setup to represent micro-scale heterogeneity of ice-rich lowlands, and extended it by a representation of meso-scale heterogeneity. We show that the combined representation of micro- and meso-scale heterogeneities gives rise to pathways and feedbacks of landscape evolution which qualitatively agree with observations, but which have not been simulated with a numerical model before.

In order to underline the original and novel contributions of the present study, we carefully revised the manuscript and took care that it stands for its own. In particular, we revised the objectives, extended the description of the tiling approach, reworked the figures, conducted additional simulations, restructured and extended the results sections, and extended the discussion with respect to large-scale modelling. Hereafter, we address the specific comments of the reviewer in a point-by-point style.

Abstract Can be shortened and cleaned. There are too many we found. . ., also found. . ., for example. . ., our results suggest. . . It was really hard to wrap my head around what exactly was found and how that helps science, stakeholders, economy, etc.

We agree that the abstract contained a lot of detailed information. We revised and shortened it for the revised manuscript.

L50. What is tile-based modeling approach? Need to define.

We added a paragraph in the Introduction which introduces the „tiling approach“ and summarizes previous work:

Many of the above-referenced modelling studies employed a so-called “tiling approach” to account for subgrid-scale heterogeneities of permafrost terrain. Instead of discretizing extensive landscape domains on a high-resolution mesh, the landscape is partitioned into a low number of characteristic landscape units, each of which is associated with a representative “tile” in the model. Thereby, geometrical characteristics (e.g., areas, distances, perimeters) are used to parameterize lateral fluxes among the landscape units. For example, Langer et al. (2016) used a two-tile model setup to investigate the effect of lateral heat fluxes in a lake-rich permafrost landscape, and Nitzbon et al. (2019) suggested a three-tile model setup to represent the micro-scale heterogeneity associated with ice-wedge polygon tundra. Schneider von Deimling et al. (2020) applied a five-tile setup to represent the interaction of linear infrastructure such as roads with underlying and surrounding permafrost. To date, the tiling method has not been applied to simultaneously represent permafrost landscape heterogeneities and their interactions across multiple spatial scales.

Moreover, we extended the description of the tiling scheme in the Methods section so that the methodology is now understandable without being familiar with the tiling concept in advance:

We used the concept of laterally coupled tiles to represent subgrid-scale spatial heterogeneities of permafrost terrain (Langer et al., 2016; Aas et al., 2019; Nitzbon et al., 2019). In general, the tiling concept involves the partitioning of real-world landscapes into a certain number of characteristic units, which are associated with the major surface and subsurface heterogeneities found in the landscape. Each of these units is then represented by a single ``tile'' in a permafrost model, and multiple tiles can interact through lateral exchange processes. The tiling approach thus allows to simulate subgrid-scale heterogeneities and lateral fluxes in macro-scale models like LSMs/ESMs, without discretizing extensive landscape domains on a high-resolution mesh, thereby keeping computational costs at a reasonable level.

In addition to these modifications in the main text, we think that the new Figure 1 as well as the revised Figure 2 facilitate a more intuitive understanding of our modelling approach.

L77. "To quantify the sensitivity". I am not sure how the sensitivity was addressed?

We understand that this formulation is imprecise. We wanted to express that our objective was to investigate how the projected permafrost degradation is affected by different representations of subgrid-heterogeneities. We rephrased the objectives which now read as follows:

Specifically, we address the following objectives:

1. To investigate the transient evolution of ice-rich permafrost landscapes in response to climate warming using different representations of micro- and meso-scale heterogeneity, thereby identifying degradation pathways and feedback processes associated with lateral fluxes on these different spatial scales.

2. To quantify the sensitivity of projected permafrost thaw and ground subsidence to different representations of micro- and meso-scale heterogeneity.

L79. What type of sources of uncertainty? See my main comments. By making super refined models, we can introduce many small uncertainties which will superposition at the end. The question is, where is the golden ratio?

We understand that this formulation is imprecise. In a preceding study, Nitzbon et al. (2020) found that different hydrological boundary conditions can lead to large deviations in the projections of ice-rich permafrost thaw under the RCP8.5 warming scenario. Here, we wanted to investigate whether these deviations are reduced, when heterogeneities on the meso-scale are represented in the model. We removed this sentence from the manuscript as might be misleading, and modified the respective paragraph which now reads:

Overall, our goal is to provide a scalable framework for exploring the evolution of permafrost landscapes in response to a warming climate, which could potentially be incorporated into LSMs to allow more robust projections of permafrost loss in response to climate change. The presented simulations should thus be considered as numerical experiments to identify important scales and controls of permafrost degradation, instead of providing accurate site-specific projections.

Table 1. should it be m2? Are we talking about the gridcell resolution?

The table is only intended to introduce the terminology with respect to spatial scales. The length scales are given in the unit [m] and the numbers reflect the order of magnitude of landscape features on the respective length scale. The numbers do not correspond to a grid cell resolution of the employed model. We hope that this becomes more clear with the extended description of the tiling approach.

Figure 1. Are you simulating the transect or an entire area? If you model an entire area, then how that area is going to look under different resolutions? In an ideal case, we should be able to take any area and then apply a deferent resolution to it (zooming in and out). The different surface features will be more/less pronounced based on spatial resolution. Then we can model future changes under different resolutions and the difference between modeling results should tell us how fine we should go. This way sounds more straight forward to me...

The figure is primarily intended to give an impression of the various types and scales of landscape heterogeneities in permafrost lowlands. The transects are intended to give an impression of the variability of the ice-wedge polygons along low-gradient slopes. In light of the critique of reviewer #1 regarding this figure (and the entire section on the study area), we understand that the figure might raise wrong expectations regarding the scope of our modelling work. In fact we do not intend to simulate the landscape evolution for Samoylov Island, but rather consider generic test cases in which we vary the representation of micro- and meso-scale heterogeneities. For this, we use the tile-based modelling approach which involves some assumptions on the geometry of the modelled landscape. To clarify the scope of our study, we replaced Figure 1 with a schematic illustration of a generic ice-rich lowland landscape. In this new figure, we also indicated how the different tiles represent different parts of the overall landscape. The revised Figure 1 (in addition with the revised Figure 2), should make the scope of the study and the modelling approach more intuitive to understand. We moved the Figure of Samoylov Island to the Appendix (Figure A1), as it provides a real-world example for the simulated permafrost lowlands.

L127 what is field capacity?

Field capacity refers to the capacity of the soil to „hold" water after drainage of excess water. In CryoGrid3, there is a parameter which specifies the volumetric water content which the soil takes upon infiltration. We modified the respective formulation:

Infiltrating water is instantaneously routed downwards through unfrozen soil layers, whose water content is set equal to the field capacity parameter (i.e., the water holding capacity): [...]

If I understand it correctly, the θ_i is initialized? I suggest to rename Δp to Δd_ice. Typically, p represent pressure. So, ice thickness is initialized too? Does the model start from the initialized ice thickness or there is a steady-state run? So, the second term in equation 2 should be less than or equal to 1? Otherwise subsidence could be greater than the ice thickness. Can that be the case? I did not understand the denominator. What is 1-phi_{nat} mean?

The volumetric ice contents θ_i are initialized as presented in Table 2. The excess ice content and the depth of the excess ice bearing layers can vary between different tiles (e.g. between polygon centres, rims and troughs). We did not conduct steady-state runs for the excess ice distribution, as our numerical model is not capable of simulating the accumulation of excess ice in the subsurface (see Discussion section 4.3). Instead, we based the cryostratigraphy on available measurements from the study area and previous modelling studies.

We renamed the variable Δp to Δd. Note, however, that it corresponds to the thickness of a soil cell which contains excess ice which is not to be confused with the thickness of the excess ice fraction. To clarify this point, and also to derive equation (2), we added a section in the Appendix (B), where we derive the formula and illustrate the composition of soil cells which contain excess ice (Figure B1). With these additional derivations it should be clear why the $1-\varphi_{nat}$ term is necessary, and that the θ_x is indeed bounded between 0 and 1, so that the subsidence cannot be larger than the ice thickness.

L149. Need a reference after ". . . hierarchical approach".

This terminology was introduced in our article to refer to the novel approach to apply the tiling concept on multiple scales (micro- and meso-scale) at the same time. In the revised manuscript we dropped the adjective „hierarchical" as it might be misleading. Instead, we only speak of „multi-scale tiling". As far as we know, the concept has not been applied in this form to permafrost environments so that we do not have additional references to provide.

N^{μ} is that somewhat standard notation? I had a hard time following that notation and remembering what it means. Is there a way to change it or use some other more intuitive notation? For example, use 1m2 or 1km2 notations. What is the total area modeled? Is this modeling represent a transect or a 2d area?

N^{μ} and N^{m} refer to the number of tiles which have been used to represent the heterogeneities on the micro- and meso-scale, respectively. In our simulations, these are either set to 1 or 3, as mentioned in the listing of the model setups in Section 2.2.1 as well as in Figure 2. These variables do not directly relate to an area or a grid resolution. The tiling concept which we employed in this study does not use an explicit 2D or 3D mesh, but rather a combination of multiple (coupled) 1D „submodels“, each of which is representative for a different landscape unit.

The questions of the reviewer indicate that the tile-based modelling approach was not sufficiently explained in the original article. We hence entirely revised the description of the tiling approach and added a schematic (Figure 1), which illustrates the concept in combination with the revised Figure 2.

Does homogeneous means that one tile represents the entire transect. If so, then it would be easier to say that 1 tile approach.

We renamed all model configurations for the revised manuscript and adopted the reviewer‘s suggestion to call the most simple setup „single-tile“.

What is the external reservoir? Is that water table depth?

The external reservoir can be thought of as a constant water table exterior to the model domain. If a tile which is connected to such a reservoir has a water table which exceeds that of the external reservoir, excess water can run off into the reservoir (without changing the level of the reservoir).

Figure 2, I had a hard time to understand and follow.

We revised the Figure based on suggestions of reviewer #1 who otherwise found the Figure very helpful as it provides an overview of all setups. We are confident that the revised Figure layout, together with the extended description of the tile-based modelling approach, facilitate a more intuitive understanding of our methodology.

L204 How many topological characteristics were used? Are these characteristics represent only magnitude of the lateral fluxes or something else too?

We used the following geometrical relations to characterize the tiles: area of a tile, elevation of a tile, distance between adjacent tiles, contact length between adjacent tiles. These relations have been used to calculate the magnitude of lateral fluxes as explained in the preceding studies Nitzbon et al. (2019) and Nitzbon et al. (2020). In addition, the areal proportions of the tiles have been used to calculate the area-weighted mean thaw depth and accumulated subsidence which is displayed in Figures 7 and 8.

Table 3 what is 'not a null' over 'big sigma' columns represent?

The column shows the area-weighted mean or the sum of the different characteristics of the micro-scale tiles. For the revised manuscript, we split up the column into two and labeled it with text instead of symbols.

I like the results section and was able to make more sense of it. I think that discussing the geomorphological processes as well as figure 6 diversify the message: "the importance of the tile-approach adoption by the ESM type models." I guess, it is important to focus on that message instead of diving into the concepts and pathways of the polygonal tundra geomorphological evolution.
For this type of paper, I would like to see a more in-depth mathematical analysis of the difference between different spatial resolutions as well as discussion of the corresponding uncertainties. I understand that this might lead to way too much work and may not be feasible in this paper. Then I suggest to exclude the ESM modeling discussion from the article and give it a different angle from the beginning.

We are happy that the reviewer could follow the results section despite the lack of clarity in our method description. We agree with the reviewer that the adoption of our tile-based model-setup in coarse-scale model frameworks would deserve a more in-depth mathematical analysis and framing, which is, however, beyond the scope of our study. We still see several important implications of our work for the LSM/ESM community which we discuss in Section 4.2 of the revised manuscript.
In addition to this, we think that it is an important contribution of our work, that manifold pathways of landscape evolution can be simulated and that these involve various feedback processes which influence permafrost degradation in response to a warming climate. In the revised manuscript, we explain these (geomorphological) processes and feedbacks directly along with the results (following the suggestion of reviewer #1), and discuss them in Section 4.1. The extended discussion of the geomorphological evolution of ice-wedge polygons and thaw lakes in Section 4.3 is intended to highlight links between our modelling work and the efforts of field researchers to develop conceptual models of these processes and landforms.

Consider bringing Figure 6 into the methods or introduction. Then it will setup the stage for the follow-up story.

We thank the reviewer for the suggestion. However, we decided to keep this figure in the discussion section, as it illustrates not only the capacities of our modeling scheme but also its limitations.

References (not contained in the original manuscript)

Schneider von Deimling, T., Lee, H., Ingeman-Nielsen, T., Westermann, S., Romanovsky, V., Lamoureux, S., Walker, D. A., Chadburn, S., Cai, L., Trochim, E., Nitzbon, J., Jacobi, S., & Langer, M. (2020). Consequences of permafrost degradation for Arctic infrastructure – bridging the model gap between regional and engineering scales. The Cryosphere Discussions, 1–31. https://doi.org/10.5194/tc-2020-192