Dear Reviewer,

we thank you very much for your comments which will help to increase the quality of our manuscript. Please find in the following a point-by-point reply to your review. We furthermore provide the revised version of our manuscript as well as an “change-tracked” version in which individual changes with respect to the submitted manuscript are highlighted.

Your comments are in **bold** and extracts from the manuscript are in *italics*, changes to the manuscript are highlighted *yellow*.

> **Nitzbon et al. develop and test a new ice-wedge polygon model to represent thermokarst in polygonal tundra. The paper is very well written, convincingly argued, and balanced. I also think the topic is very important (permafrost degradation) and relevant to ongoing analyses in many groups. I appreciate their creative approach to representing a very spatially heterogeneous system with a geometric scaling approach and their analyses of the sensitivity of their results to the assumptions of their approach.**
> **However, I think the authors could strengthen the paper by considering the following suggestions:**
>
> **1. The validation runs are worthwhile, and do not appear to result in large discrepancies in moisture or temperature. However, I expect larger variation from uncertainty in soil parameters, and so suggest that such a sensitivity analysis be performed. Parameters in Tables 2, 3, and D1 are all uncertain, so I would like to see an analysis of which are dominant for the system responses you are studying, ...**
>
> We agree that many of the model parameters are uncertain and that it is hence worthwhile to study the sensitivity of the modelled quantities (ALT, WT, SD, SEB, soil temperature, soil moisture) against variations in those parameters. In our initial manuscript we presented a sensitivity analysis for parameters related to the topology, micro-topography and the snow density of our model set-up. However, we just presented the overall spread of all simulations which did not allow to infer the influence of individual parameters. In our revised manuscript we reworked the model validation and sensitivity analysis (section 3). For this, we (i) conducted additional simulations for variations of the field capacity ($\theta_{fc}$) and hydraulic conductivity ($K$), and (ii) compared the modelled quantities with measurements for each parameter variation individually. In contrast to the initial manuscript, we refrained from presenting all possible permutations of the varied parameters, and instead present only runs which differ in one parameter from a set of “default” parameters. The reworked figures (Fig. 5, 6, 7, 8, E1, E2) allow to infer the influence of certain parameters and give an idea of the overall sensitivity of the model to parameter variations. Since the model validation was completely revised, we modified some formulations in sections 3.1 and 3.2 which are highlighted in the “change-tracked” manuscript provided together with the revised manuscript.
>
> **… and then an uncertainty quantification of your main results associated with variation in the dominant parameters.**
>
> Our main objective was to investigate the influence of different hydrological conditions (reflected in the parameter $e_{res}$) on the degradation of ice-wedges. The focus in this respect was on highlighting the qualitatively different degradation pathways which were simulated within our model framework. While it is likely that other parameters than $e_{res}$ will also influence the timing (and speed) of the degradation, an in-depth analysis of

such factors would go beyond the scope of our study. Hence we did not conduct further "long-term" simulations for our main results section. However, we pointed out the influence of further factors on the modelled ice-wedge degradation in section 5.1 of the discussion:

*It should be noted that – apart from the hydrological conditions reflected in $e_{res}$ – other parameters of the model, including snow properties, the soil stratigraphy, and the depth and amount of excess ice, are likely to affect the timing of the onset of ice-wedge degradation.*

We hope that our revised manuscript (section 3) allows the readers to get a better understanding of the sensitivity of the model to the most uncertain parameters. Given this understanding and the description of the model setup in section 2.3, we are confident that the main results can be interpreted correctly.

**2. The hydrology model structure described in Appendix A is somewhat disappointing, given advances made over the past few decades in implementing more sophisticated approaches. However, the proof's in the pudding, and Figure E2 appears to show good comparisons.**
**It is probably worth mentioning in the main text that the model systematically underestimates water content in the rims. That problem may be from setting the porosity to 0.5, but it's not easy to tell.**

We added the following sentences to section 3.2.2 of the manuscript:

*There was a good agreement between modelled and measured soil moisture levels for the mostly water-saturated center, while the model underestimates soil moisture by about 10% in the dry rim profile. The latter can be attributed to the field capacity parameter ($\theta_{fc}$, Table 2), which is poorly constrained by field measurements.*

**a. Are there no observations at other depths, for both moisture and temperature? Report R2 against observations for temperature and moisture.**

We added comparisons of soil temperature and soil moisture for both polygon centres and troughs at a depth of about 0.40 m to the appendix of the revised manuscript (Figures E1 and E2).

**b. You should describe the model time step and numerical methods for solution.**

We added the following sentence to section 2.2.2 of the manuscript.

*CryoGrid3 uses a first-order forward Euler algorithm with adaptive time step for the numerical integration of the heat conduction equation (see (Westermann et al., 2016) for details).*

**c. Discuss in the main text motivation for your choice of using a simple hydrology model, and what possible implications are.**

We agree that the employed hydrology scheme is rather simplistic and that more sophisticated approaches are available and have already been applied to permafrost settings. However, it constitutes a significant improvement compared to the previous version of CryoGrid3 Westermann et al. (2016), which did not take into account variable water contents in the active layer at all. The instantaneous infiltration scheme we used, turned out to be sufficiently suited to fulfill the purpose of reflecting the spatial heterogeneity of the ground hydrological regime of polygonal tundra. Moreover, Zhang et al. (2010) showed that instantaneous infiltration algorithms do not necessarily perform worse than more sophisticated schemes. To further justify our methodology, we changed the following formulations in section 2.2.3.

Changed:

*This is a valid assumption for the upper soil layers of tundra wetlands, which are typically characterized by large hydraulic conductivities (Boike et al., 2008) and in which infiltration into the active layer is mainly controlled by thaw depth (Zhang et al., 2010).*

Added:

*We note that the employed hydrology scheme is rather simplistic compared to other schemes available (e.g., Painter et al. (2016)). We confirmed however, that the employed scheme, in combination with the lateral water transport scheme detailed in Sect. 2.2.4, was sufficiently suited to reflect the spatial heterogeneity of the subsurface hydrological regime of polygonal tundra (see Sect. 3).*

**d. Discuss the role of vegetation changes that might be expected during degradation. Currently you set the vegetation parameters at the beginning of the simulation, and I think they remain constant. But, e.g., a drying system should expect to see a transition to plants less adapted to saturated conditions, and that will affect ET.**

It is correct that the vegetation parameters (the root depth $d_R$ and the catch height of snow) remain constatnt throughout the simulations. We agree that the lack of a dynamical vegetation scheme is a limitation of our model, particularly for long-term (decadal or centennial) simulations. We mention this fact in section 5.4 and discussed it further in the revised manuscript.

Added in section 5.4:
*In reality, the change of the subsurface hydrological regimes resulting from ice-wedge degradation (e.g. the drying of polygon centres; Fig. 9), would imply also an adaption of the vegetation (Wolter et al., 2016). This in turn would affect the surface energy balance through changes to the evapotranspiration in a non-trivial way. The development of aquatic vegetation which is also not represented, would have an isolating and thus stabilizing effect on ice-wedges (Kanevskiy et al., 2017).*

**e. It's difficult to see how well the model is doing in Figure 6. Change the y-axis range to -0.2 to 0.3, and report R2 from the average of the 8 simulations, or some combination of those simulations.**

We changed the axis of the figure showing the evolution of WT. Instead of taking an average of the model simulations we display each of the model runs with an individually coloured line and indicate which parameter was varied in each run compared to the default parameter values.

**3. I am confused about what is being compared in Figure 7. How can an ECOR measurement separate out centers and rims (wet and dry)? Seems impossible, so it’s not clear what is being compared.**

The eddy-covariance measurements taken from Langer et al. (2011) were conducted at several sites on Samoylov island which had variable areal coverages of wet and dry tundra. Langer et al. (2011) combined these measurements to estimate the contributions of fluxes from wet and dry tundra based on “fractional unmixig”. Since this techniques only gives rough estimates, the given uncertainties are accordingly large. We mention this detail about the measurement data in section 3.1 of the revised manuscript:

*These data include a separation of “wet tundra” and “dry tundra” surface energy fluxes which was based on a linear decomposition of measurements conducted in different parts of Samoylov island with variable areal coverages of wet and dry tundra (see Langer et al. (2011b) for details).*

**a. Put a ’wet tundra’ label above the gray part, and a ’dry tundra’ label above the RHS part. And, describe what these terms mean in the context of an ECOR measurement.**

We added appropriate labels to the figure. For the distinction of wet and dry tundra in the measurements, see answer above.

**4. Line 15 of Page 19, where you use the word “realistic” for your ice-wedge degradation approach. I think you should move the text from lines 12-17 on page 20 up here to show that your results are reasonable, even though you have not made any direct comparisons with degradation. Otherwise, as written, on page 19 I did not see how your representation was reasonable for degradation.**

We agree that the lack of quantitative comparisons with the degradation at our study site does not support the statement made on p.19 l. 15f. With our statement we intended to refer to the reflection of the general process of ice-wedge degradation, independent from our study site. To make this more clear, we moved the comparison with measured subsidence rates to the beginning of the section and modified the original statement about “realistic degradation” such that it refers to the qualitative landscape evolution.

*There is a lack of reliable, long-term measurements of ground subsidence for the different micro-topographic units of polygonal tundra in our study area, which makes quantitative comparisons with the modelled landscape evolution unfeasible. However, Boike et al. (2018) reported recent (2013 to 2017) subsidence rates on Samoylov Island to be in the order of 0.04 m a−1 for polygon rims and < 0.01 m a−1 for polygon centres. These figures are in agreement with the modelled subsidence characteristics, with rates of about 0.02 m a−1 for the rim tiles and no subsidence for the centre tiles (see Figs. 9 and 11). While the modelled ground subsidence seems to be reasonable, the available measurements did not allow for a quantitative comparison of the degradation rates of ice-wedges underneath the troughs. The long-term (60-year) runs with variable hydrological conditions demonstrated, however, that our model framework is able to reflect the process of ice-wedge degradation and the associated changes to the micro-topography of polygonal tundra as described in other studies (e.g., Liljedahl et al. (2016)) in a qualitatively realistic way.*

**5. In Figure 2, label the colors of the features with a legend.**

We changed Figure 2 accordingly.

We hope that we were able to address all questions and comments raised in your review to your satisfaction, and that our revised manuscript is in an adequate state for publication.

Yours sincerely,

Jan Nitzbon (on behalf of the authors)

## References

Westermann, S., Langer, M., Boike, J., Heikenfeld, M., Peter, M., Etzelmüller, B., & Krinner, G. (2016). Simulating the thermal regime and thaw processes of ice-rich permafrost ground with the land-surface model CryoGrid 3. *Geosci. Model Dev.*, 9(2), 523–546. https://doi.org/10.5194/gmd-9-523-2016

Zhang, Y., Carey, S. K., Quinton, W. L., Janowicz, J. R., Pomeroy, J. W., & Flerchinger, G. N. (2010). Comparison of algorithms and parameterisations for infiltration into organic-covered permafrost soils. *Hydrology and Earth System Sciences*, *14*(5), 729–750. https://doi.org/10.5194/hess-14-729-2010