Response to anonymous referee 2

We would like to thank the reviewer for evaluating our manuscript and for the useful comments, which helped to improve it. In the following, we provide a reply to the points discussed by the reviewer as well as changes in the manuscript.

The comments of the reviewer are written in **bold**, the extracts of the manuscript in *italics* with changes highlighted in blue and line numbers referring to the revised manuscript.

Juditha Aga et al presented a new modeling capability added to CryoGrid model to simulate ice segregation and thaw consolidation considering proof-of-concept scenarios. While the capability is important and of interest to the modeling community of permafrost regions, especially considering ground heave and subsidence evolution under a changing climate, I have a few concerns that are listed below.

Major comments

A better motivation for the study is needed in the Introduction section; why should one care about ice segregation; where in the Arctic they form; what observations show us; what has been done already to address ice segregation (not Earth system models but small-scale models). Field imagery (and/or a schematic) of ground heave and subsidence can help set the stage as ground heave and subsidence are linked with ice segregation and thaw consolidation.

We added the suggestions of the reviewer in the introduction of our manuscript.

We have one paragraph in the introduction, which discusses effects of segregated ice upon thawing, such as geomorphological changes (thermokarst and ground subsidence), a potential contribution to the permafrost carbon feedback and effects on the stability of the ground. We rephrased some parts to improve the paragraph. Furthermore, we included the effect of a delayed permafrost degradation due to ice-enriched soil layers. We mention this also as a missing element in previous model studies.

Line 50: *The ground ice content and its distribution strongly determines the sensitivity of permafrost to thaw (Jorgenson et al., 2010; Nitzbon et al., 2019). Ice-rich layers in the soil can delay permafrost degradation as energy is consumed upon melting of the ground ice, which is consequently not available for the warming of the ground. In addition, ice segregation can continue even under thawing conditions, forming ice layers at the top of the permafrost, continuously delaying the warming process. However, if excess ice is melted, it can result in thermokarst and ground subsidence (Farquharson et al., 2019; Kokelj and Jorgenson, 2013; Nitzbon et al., 2019). In consequence, substantial geomorphological changes reshape the landscape, manifested in the formation of lakes, thaw slumps, gullies and the transformation of low-centered to high-centered polygons (Kokelj and Jorgenson, 2013; Liljedahl et al., 2016; Nitzbon et al., 2019). These processes could contribute to accelerating the* *mobilization of permafrost carbon, which may further increase atmospheric carbon concentrations, a process known as the permafrost carbon feedback (Miner et al., 2022; Schuur et al., 2008). Furthermore, ground ice controls the hydrological and mechanical properties of the soil by reducing permeability and increasing the mechanical strength (Painter and Karra, 2014). These parameters control the structural stability of the ground. Upon thawing, mass movements along slopes might increase and together with ground subsidence, the reduced stability can endanger human infrastructure and settlements (Hjort et al., 2022; Schneider von Deimling et al., 2021).*

Line 75: *As a consequence, previous models targeting ground ice thaw and thermokarst require excess ice distributions prescribed from field observations, which makes applications at sites without ground ice data challenging. Furthermore, as ice segregation is not implemented, they neglect the delay in permafrost warming through formation of new segregated ice layers during the simulation period.*

We added a new paragraph the occurrence of segregated ice.

Line 41: *Layers of segregated ice in the ground are widespread in permafrost environments, especially in fine-grained sediments, which are susceptible to ice segregation (French and Shur, 2010). Cryostratigraphic mapping has been performed in numerous studies, documenting segregated ice especially in Siberia (Andreev et al., 2009; Meyer et al., 2002; Schirrmeister et al., 2008; Siegert and Babiy, 2002) and North America (French et al., 1986; Heginbottom, 1995; Kanevskiy et al., 2013; Shur and Jorgenson, 1998). O'Neill et al. (2019) modelled the occurrence of segregated ice in Canada, showing abundance in fine-grained lacustrine sediments, raised peat plateaus and uplifted marine sediments.*

We added a new paragraph in the manuscript about what observations can show us.

Line 46: *Observations of the cryostructure, including segregated ice, can reveal information about the evolution of the ground, e.g. if the permafrost was formed syngenetically or epigenetically. Furthermore, thaw unconformities, such as former active layers, can be detected by changes in the ice content with depth and the isotopic signature of the ground ice (French and Shur, 2010).*

We extended the references that we already mentioned with a very interesting study that demonstrates palsa formation through ice segregation.

Line 69: *Thaw consolidation has been the focus of model development for many decades (Morgenstern and Nixon, 1971; Nixon and Morgenstern, 1973; Sykes et al., 1974; Foriero and Ladanyi, 1995; Dumais and Konrad, 2018), and different approaches have been presented for modelling ice segregation (Fu et al., 2022; Fisher et al., 2020; Lacelle et al., 2022). An example is the approach of An and Allard (1995), who successfully demonstrated palsa formation through accumulation of* *segregation ice. However, these processes are typically not implemented in land surface models and simulating the long-term evolution of segregated ice has not been performed yet (Fu et al., 2022).*

We added a schematic of ice segregation and ground heave in the introduction to visualize the process for the reader.

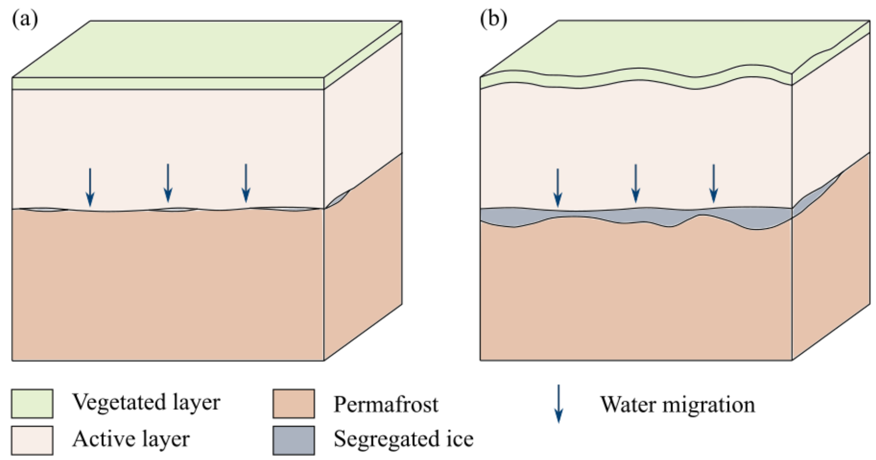


Illustration of ground heave through ice segregation at the base of the active layer. (a) Lenses of segregated ice are formed at the top of the permafrost. (b) If the segregated ice is preserved over a long time period, layers with segregated (excess) ice are forming, causing heave of the ground surface. Figure modified after Fu et al. (2022).

Lots of details have been provided about the model in the Methods but they are hard to follow, especially the description of CryoGrid class (modules) needs a better workflow; a schematic would be helpful to understand what parts of the existing CryoGrid have been modified and what new parts are added.

We rephrased extensive parts of the description of CryoGrid to improve the clarity. Furthermore, we added a table with an overview with the model components and how they have been changed.

Line 91: *In this work, we extend the capabilities of the CryoGrid community model (Westermann et al., 2023) with a representation of soil mechanical processes. The CryoGrid community model is a modular framework for simulating the permafrost thermal state and the water and ice balance. To set up simulations, the user can choose between different so-called "stratigraphy classes", which are characterized by specific model physics and state variables. As an example, one stratigraphy class can calculate soil water contents with a simple model, while another can account for water redistribution through Richards equation in unsaturated soils. Furthermore, there are dedicated stratigraphy classes for non-ground materials, in particular* *for the seasonal snow cover. The stratigraphy classes can be stacked vertically, so that the available classes representing snow can can be flexibly combined with a range of classes for ground materials.*

In this study, we demonstrate a new stratigraphy class denoted GROUND_freezeC_RichardsEq_seb_pressure, a fully fledged process model for soils. It is based on the already existing stratigraphy class GROUND_freezeC_RichardsEq_seb (Westermann et al., 2023) and inherits many of its functionalities. While a detailed description of the CryoGrid community model is provided in Westermann et al. (2023), we summarize the main aspects relevant for this work in Sect. 2.1, before describing the defining equations and main properties of the new stratigraphy class. The model is demonstrated and evaluated for two field sites (Sect. 2.2), for which we simulate a range of model scenarios (Sect. 2.4) with different settings for subsurface properties and model parameters (Sect. 2.3).

Line 107: *In the stratigraphy classs GROUND_freezeC_RichardsEq_seb, each model grid cell is characterized by its volumetric contents of the mineral, organic, water and ice components, which also define the porosity. The upper boundary is defined as the interface between the ground surface and the atmosphere at which the surface energy balance is applied, controlled by the exchange of short-wave and long-wave radiation, as well as latent and sensible heat fluxes. To calculate the surface energy balance, the forcing data must prescribe time series of air temperature, solid and liquid precipitation, wind speed, short-wave and long-wave radiation, specific humidity and air pressure. The lower boundary condition (set at a user-specified depth) is defined by a constant geothermal heat flux.*

Subsurface heat transfer is based on both conductive and advective fluxes. The calculation of heat conduction follows Fourier's law with the thermal conductivity of the material being the controlling factor. The heat transfer through advection is determined by vertical water fluxes.

A soil freezing characteristics describing the relationship between ground temperature and unfrozen water content (Painter and Karra, 2014) is implemented in GROUND_freezeC_RichardsEq_seb. To determine liquid water and ice contents in frozen soils, we first calculate the matric potential for unfrozen conditions ψ_0 $[\mathrm{m}]$ *and based on that, the matric potential in frozen state* ψ $[\mathrm{m}]$ *from which the water content can be inferred (assuming no residual water). A detailed description of the approach can be found in Westermann et al. (2023).*

Table 1. Different model components, their implementation in the CryoGrid community model after Westermann et al. (2023) and the additions to the model code, presented in this study.

Model component	*Base class GROUND_freezeC_RichardsEq_seb*	*Additions in GROUND_freezeC_RichardsEq_seb_pressure*
Upper boundary condition	*Surface energy balance*	-
Lower boundary condition	*Geothermal heat flux*	-
Subsurface heat transfer	*Heat conduction and heat convection*	-
Soil freezing characteristics	*after Painter and Karra (2014)*	-
Water flow in unsaturated conditions	*Richard's equation*	-
Water flow in saturated conditions	*Gravity-driven*	*Additional water flow when excess pore water pressures occur*
Lateral water transport	*Overland flow*	-
Stress conditions in the ground	-	*Calculation of total and effective stresses*
Compressibility of the soil column	-	*Calculated from the compression curve*
Excess ground ice	*Defined in the initial conditions, melting of excess ice possible*	*Formation and melt of segregated ice*
Sedimentation	-	*Material is added with user-defined properties*

The manuscript mainly focused on Samoylov (S) scenarios and not Bayelva (showed just one B-clay). If the focus is more on one site, it would make sense to drop the other to not confuse the readers. Samoylov's climate can be made synthetically warmer as the authors are performing proof-of-concept simulations and not validating the model.

We added two more model scenario for the Bayelva field site (*B-clay-sed1000* and *B-clay-sed370-2x*), so that the manuscript is no longer entirely focused on Samoylov. Furthermore, we compare the results of this model scenario with in situ observations from a site with high-quality ground ice observations on Svalbard (see comment below).

CryoGrid uses the surface energy balance as the upper boundary, so that it is not possible to simply make it warmer, as warming in reality also changes other parts of the forcing, in particular the incoming radiation and humidity. Therefore, Bayelva needs to be included in this study.

While I personally believe projections are not needed here to demonstrate the capability, since they are there how do they compare against simulation without the formation of ice segregation?

We performed a model run without formation of segregated ice and included the analysis in the supplement.

Supplement, Line 43: *S4: Control runs without formation of segregated ice*

We perform a model run based on the peat stratigraphy (Table 3) without ice segregation (S-peat-control). To do so, we suppress water flow into already saturated and frozen grid cells, while the rest of the functionality of the model stays the same. The model scenario S-peat forms segregated ice predominantly between 1861 and 2000 (Sect. 3.4). During this time period, the active layer thickness is on average 0.054 ± 0.029 m *shallower for S-peat compared to S-peat-control. As S-peat forms segregated ice at the top of the permafrost each year, this excess ice melts partly in the following year, consuming energy, which is consequently not available for warming of the ground. As S-peat-control does not contain segregated ice, the energy can be used directly to warm the ground, resulting in a deeper active layer.*

In the time period 2000 to 2010, large parts of the segregated ice are melted in S-peat (Sect. 3.4). Again, this process requires energy, even more than during 1861 to 2000, as more ground ice is melted during these years. Therefore, the difference in active layer thickness increases to 0.078 ± 0.088 m *on average, with S-peat having the shallower active layer than S-peat-control.*

From 2010 to 2100, S-peat forms less segregated ice, however, ice segregation continues at the top of the permafrost on a smaller scale (Sect. 3.4). Therefore, a shallower active layer is still simulated for S-peat compared to S-peat-control, even though the difference in active layer thickness decreases to 0.039 ± 0.053 m *on average.*

The comparison between S-peat-control and the model scenario S-peat with ice segregation shows, that the formation of segregated ice leads to shallower active layers, especially during the periods where the ice-enriched soil layer thaws. This can be explained by the energy required for ground ice melt, which is then not available for ground warming.

I have also some reservations about the reference date of August 31 when using different climates, this might be fall freeze-up time for one site but not for another site. So, the comparison that on August 31 one site showed this much ice segregation and another showed that much is not very convincing.

We checked the validity of the reference date for both field sites by comparing the vol. water and ice content at the upper 20 cm and 1 m of the permafrost of the reference date August 31 and the day with the maximum active layer thickness for each year. We calculated the following mean differences with corresponding standard deviations:
Samoylov: $8.4e-05 \pm 0.0269$ (20 cm) and -0.0012 ± 0.0107 (1 m)
Bayelva: -0.0050 ± 0.0202 (20 cm) and -0.0020 ± 0.0072 (1 m)
We are therefore confident, that we can use the results August 31 for our analysis. A further advantage of using a fixed date rather than the state at maximum thaw is that we can use the same procedure when a talik has formed. In this case, we would

need to introduce additional criteria on how the point in time was selected after a talik has formed, which makes it more complicated to compare ground ice contents throughout the figure. We added a statement in the revised version of the manuscript.

Line 407: *In the following, we present results of the model scenarios described in Sect. 2.4 and summarized in Table 4. To ensure comparability between the different sites and years, results are always provided at August 31.*

I would suggest just focusing on the historical climate (1000s of years) and showing how thick ice lenses (segregated ice) formed in the past – at least show the simulated segregated ice is in some comparison with the field observations. Showing the formation of 2-3 cm segregated ice is not very appealing given it is a modeling paper and the main focus is ice segregation and thaw consolidation, so the authors need to explore it further.

It is true, that the formation of thick segregated ice layers takes typically 1000s of years. To realistically simulate this process, we would require long-term historical forcing data. As we do not yet have such historical forcing data available at this point, we use a model spin-up by repeating a 10-years period. The model scenarios with a spin-up of 1000 years show, that the model is capable of building up thick ice lenses through ice segregation, while the model scenarios with a spin-up of 100 years investigate the sensitivity to different influencing factors. As the aim of the manuscript is to demonstrate the functionality of the model, we are confident that the chosen spin-up periods support the objectives of the study. To simulate a the formation of segregated ice at a specific field site, long-term historical forcing data would be beneficial as the reviewer suggests. Such simulations would be in principle possible with the presented model, but preparing bias-free historical forcing data is challenging and beyond the scope of this work. However, it would be an interesting aspect for future studies.

Although a comparison with field measurements is challenging due to lacking historical forcing data, we added a comparison with in situ observations of ground ice contents which indeed suggest that we can produce results in the right order of magnitude. Concretely, we added a sedimentation run with increased sedimentation rates for Samoylov (*S-clay-sed350-3x*) to compare it to Nitzbon et al. (2019). Furthermore, we set up a model scenario with increased sedimentation rates as well for Bayelva (*B-clay-sed370-2x*) and compare it to the study of Cable et al. (2018). We included the runs in the results and added a new section to the discussion, which is shown below. We are grateful to the reviewer for this valuable suggestion which in our opinion has improved the quality of the study!

Line 638: *4.4 Comparison to in situ observations of ground ice*

The primary goal of this study is to demonstrate the accumulation and melt of segregated ice in CryoGrid and explore its sensitivity to different model settings. However, it is meaningful to compare the simulated ground ice contents to in situ observations, at least for some of the model scenarios and in a semi-quantitative way, keeping the limitations of the model setup in mind. For most model scenarios without sedimentation, only thin layers of ground ice are formed during the model spin-up, making a comparison to situ observations, which generally report ice contents for thicker layers, challenging. However, we can use the sedimentation runs for a comparison with observed ground ice contents, in case the sedimentation history and rates can be sufficiently constrained.

At the Bayelva site, concurrent observations of sedimentation history and ground ice are not available, but we can compare the model results to observations from Adventdalen, Svalbard, located about 120 km to the south-east, which features broadly similar climate characteristics. Cable et al. (2018) present cryostratigraphic observations including excess ice contents from sediment cores, and we select a core (A2a) featuring Holocene deposits with high contents of fine-grained material and only low organic content, which is similar to the simplified stratigraphy in the B-clay scenarios. The deposits have been formed by aeolian sedimentation with sedimentation rates of 1.1 $\mathrm{mm\,a^{-1}}$ in the last 600 years (Cable et al., 2018), which is the same as the effective sedimentation rate in the B-clay-sed370-2x scenario. The observed volumetric moisture content in the aeolian deposits of the selected sediment core is approximately 76 % (calculated from the gravimetric moisture content of 54 %), with a volumetric excess ice contents of 18 %. These values are in a similar range as the simulated ice contents in the ice-enriched layer in B-clay-sed370-2x, with total volumetric moisture contents of up to 69 % and volumetric segregated ice contents of up to 14 %.

For the Samoylov site, a comparison to ground ice observations is challenging, as both wedge ice and segregated ice have contributed to the build-up of syngenetic permafrost during the Holocene (Schwamborn et al., 2002). Furthermore, the island has a complex sedimentation history, being located in the upper part of the vast delta of the Lena river. Radiocarbon dates of organic material in a core from Samoylov island are presented in Schwamborn et al. (2002), suggesting that approximately 5 m near the top of today's permafrost were deposited during 2000 to 2500 years. Another core is described in Schwamborn et al. (2023), showing calibrated radiocarbon ages of 1370 ky BP near the top of the permafrost (0.5 to 0.6 m depth below the surface), while the ages varied randomly between 4000 and 6000 ky BP between 1.5 and 6.5 m depth. While this shows that sedimentation has not been a continuous process on Samoylov island, with sediments likely being reworked, we can broadly estimate effective sedimentation rates between 0.3 $\mathrm{mm\,a^{-1}}$ and 2.5 $\mathrm{mm\,a^{-1}}$ for the permafrost section underlying the present-day active layer. For the Samoylov site, long-term sedimentation runs have been performed for effective rates of 0.55 $\mathrm{mm\,a^{-1}}$ (S-clay-sed1000) and 1.7 $\mathrm{mm\,a^{-1}}$ (S-clay-sed350-3x), which is well in the range that can be established from the drill cores. We compare the simulated ice contents to the cryostratigraphy of a polygon center where segregated ice indeed occurs, synthesized from observations on Samoylov island for modeling purposes (Nitzbon et al., 2019). For the ice-rich layer in the top-most permafrost (below 0.9 m depth in Nitzbon et al., 2019), a total volumetric ice content of 65 % is prescribed, with a volumetric excess ice content of 18 %. In comparison, S-clay-sed1000 features a total volumetric ice content of 69 % (volumetric excess ice content 14%), and S-clay-sed350-3x shows a volumetric ice content of 60 % (volumetric excess ice content 5%). Overall, it is encouraging that the simulations produce ice contents in the correct order of magnitude also for Samoylov island. However, we emphasize that this comparison is highly uncertain and should not be understood as a validation of the model, in particular since formation of segregated ice occurs in concert with wedge ice on Samoylov island.

Nevertheless, the comparison to observed ground ice contents suggests that our model (with the simplified stratigraphy) is capable of modelling ice segregation in the correct order of magnitude at both the Bayelva field site and Samoylov island.

However, we see three main limitations of our model setup which need to be addressed to simulate more realistic cryostratigraphies: (i) the climate data used for spin-up, (ii) the constant sedimentation regime and (iii) the hydrological regime. (i) Simulating ground ice contents at a specific field site requires long-term historical forcing data and our spin-up procedure with a repetition of a 10 years period in the 20th century is not suitable to represent the historical climate. This limitation can in principle be resolved by preparing time series of model forcing from paleo-climate model runs. (ii) We applied a constant sedimentation rate, while the sedimentation regime experiences changes over such long time periods, which are not represented in the current model setup. In practice, this problem is hard to resolve, but it may be enough to prescribe sedimentation rates for the most important layers within a cryostratigraphy. Moreover, "sedimentation" with organic material could be simulated by a dedicated subsurface carbon cycle model. (iii) The hydrological regime of polygonal ground is different to the 1D setup in our study. Previous modeling studies have shown that the seasonal dynamics of ground temperatures and water/ice contents in the polygonal tundra landscape cannot be captured with one-dimensional simulations (Nitzbon et al., 2019), as conducted in this study, which likely affects the accumulation of segregated ice. Our model setup could in principle be included in three-dimensional simulations with laterally coupled tiles (Aas et al., 2019; Nitzbon et al., 2019) without taking lateral cryosuction into account (Sect. 4.5). However, realistic long-term simulations would require the additional representation of wedge ice build-up, which is the primary process of ground ice accumulation in polygonal tundra landscapes.

Minor comments

L2. What does “very ice-rich soils” mean?

We rephrased the sentence.

Line 2: *The ground ice content in cold environments influences the permafrost thermal regime and the thaw trajectories in a warming climate, especially for soils containing excess ice.*

L24: Sibiria should be Siberia

Fair enough. We corrected it.

Line 24: *Especially in Siberia and Alaska, they are often associated with a high content in excess ground ice.*

L2531-34: the formation of segregation ice needs a clear description. Cryosuction can even happen from top to bottom.

We added the definition from the "glossary of permafrost and related ground-ice terms" (Harris et al., 1998) for a clear description. We mention in this paragraph, that the water can be drawn from the active layer towards the freezing front, i.e. from top to bottom.

Line 31: *Segregated ice is ground ice, which forms through migration of soil water to the frozen fringe (Harris et al., 1998). It occurs as discrete ice lenses or layers, which can range from less than a millimeter to more than 10 m (French and Shur, 2010; Harris et al., 1988). Segregated ice is typically associated with ground heave as the ice content in the ground increases (Fig. 1; Miller, 1972; Taber, 1929). It can build up in epigenetic permafrost when unfrozen soil water from the active layer is attracted towards the freezing front (Guodong, 1983; Mackay, 1983; Taber, 1929), accumulating ice near the permafrost table. Besides, ice segregation can occur at the base of the permafrost, as water is drawn from the soil below as permafrost is forming. These ice-enriched layers are widespread in permafrost environments (French and Shur, 2010) as for example in Canada (O'Neill et al., 2019). However, the evolution of syngenetic permafrost can result in enhanced ice formation. Accumulation of organic material as well as sedimentation in alluvial, eolian or hillslope settings can lead to a rise in the permafrost table and hence a growth of segregated ice (Guodong, 1983; French and Shur, 2010). In this context, segregated ice can form also together with syngenetic ice wedge growth, forming ice lenses within polygonal permafrost as observed in Yedoma deposits (Schirrmeister et al., 2013).*

L43: "thermokarst and subsequent ground subsidence" does it mean thermokarst happens first and ground subsides afterward? What is thermokarst?

We removed "subsequent" from the text.

Line 53: *However, if excess ice is melted, it can result in thermokarst and ~~subsequent~~ ground subsidence.*

L71: "contrasting permafrost sites" in terms of what?

We specified it in the revised manuscript.

Line 85: *To do so, we run various model scenarios with climate data from two contrasting permafrost sites, representing cold continental and relatively warm maritime permafrost conditions.*

L94: relative humidity or specific humidity? Please specify

We clarified it in the manuscript.

Line 110: *To calculate the surface energy balance, the forcing data must prescribe time series of air temperature, solid and liquid precipitation, wind speed, short-wave and long-wave radiation, specific humidity and air pressure.*

Section 2.1 Model description. A schematic here can really help understand all the processes in the model and how these processes are linked together. While the authors have explained the model in the text, it is still confusing for those who are not very familiar with the model.

We added a table in section 2.1, explaining where changes has been made to the model code of the CryoGrid community model. Furthermore, we rephrased large parts o the CryoGrid community model description for clarity. For details, see reply on the corresponding major comment.

L106: define the sub/superscripts.

We defined the sub- and superscripts.

Line 122: *The water balance in* GROUND_freezeC_RichardsEq_seb *is based on vertical water flow j_w^v [$\mathrm{m\,s^{-1}}$] controlled by the Richards equation (Richards, 1931):*

$$j_w^v = -K_w \left(\frac{\partial \psi}{\partial z} + 1 \right) \tag{1}$$

with ψ [m] being the matric potential, z [m] the vertical coordinate and K_w [$\mathrm{m\,s^{-1}}$] the hydraulic conductivity. The subscript w denotes water and the superscript v signifies vertical for model variables.

L128: "table" should be Table

We corrected it in the manuscript.

Line 149: *To do so, a set of additional state variables is necessary, which can be found in Table 2.*

L168: This is a model development paper, instead of referring the reader somewhere else please provide the definition here.

We added a statement why we used the bulk quantities in the model instead of referring only to Westermann et al. (2023).

Line 192: *with Θ_m [m^3] and Θ_o [m^3] being the bulk volumetric content of the mineral and organic components, i.e. the absolute volume in a grid cell filled by the respective component. In GROUND_freezeC_RichardsEq_seb, bulk quantities as Θ_m and Θ_o are conveniently used as state variables (Westermann et al., 2023), so that the grid cell thickness d for unsaturated grid cells can be updated with the porosity obtained from Eq. 7 (see Sect. 2.1.2 for the saturated case).*

L247: So Bayelva is warm and maritime, and Samoylov is cold and continental? The climate is mentioned for one site but not for the other.

We changed the manuscript accordingly.

Line 274: *Samoylov island (located in northern Siberia) represents a continental setting with cold continuous permafrost, while the Bayelva field site (located on Svalbard) is characterized by a maritime climate with warm, but still continuous permafrost.*

L251: "In summer, precipitation and evapotranspiration balance each other." It is not clear if this statement is only for this particular site or not. Please clarify. I am not sure if this is true across the Arctic, it depends on the climate.

We clarified, that this statement is only true for the Samoylov field site.

Line 280: *In summer, precipitation and evapotranspiration balance each other in Samoylov.*

L255-265: many abbreviations without definitions.

We added the definitions for the abbreviations.

Line 289: *We use the same forcing data set as Westermann et al. (2016), Nitzbon et al. (2019) and* Nitzbon et al. (2020) *for the long-term thaw susceptibility run in that study. The climatic forcing between 1960 and 2012 is derived from the reanalysis product CRU-NCEP combining data from the Climate Research Unit and National Centers for Environmental Prediction, see Kalnay et al., 1996; Harris et al., 2014), downscaled for Samoylov with in situ data from the automatic weather station. After 2012, the forcing data is taken from the Community Climate System Model CCSM4 outputs simulated under the* Representative

Concentration Pathway (RCP) 8.5 scenario (Meehl et al., 2012).

L267-268: 870 mm rainfall and 668 mm snow, so the total is about 1500 m annual precipitation. Is this for Samoylov? This is a lot different than what other studies have reported for example Liljedahl et al. (2016)

We thank the reviewer for pointing out this inconsistency. We checked our forcing data, and we made some mistake by calculating the annual precipitation (we took mm instead of mm/d). The forcing data itself is correct, so that the simulations are not affected. We corrected the numbers in the manuscript.

Line 297: *An increase in air temperature from -16.7* °C *to -8.1* °C*, an increase in longwave radiation from 214 to 249 W* m^{-2}*, a slight decrease in shortwave radiation from 105 to 101 W* m^{-2}*, a pronounced increase in rainfall from 157 to 218 mm and an increase in snowfall from 133 to 167 mm.*

L298: "undecomposed organic material features coarse pores" how old is this organic material? Also, is this just an assumption or the organic material is undecomposed at those sites?

The layer in the upper 15 cm represents the moss layer on top and is therefore poorly decomposed as fresh vegetation is still growing here. We rephrased the sentence to clarify this.

Line 330: *Hereby, we use the same soil properties in the upper 0.15 m as in the multi-layer stratigraphy to account for the insulating effect of the organic-rich moss layer on top. The poorly decomposed organic material features coarse pores with unknown values for the van Genuchten parameters n and α.*

L317: 1100 m: +10.2C, does this deep soil temperature also come from borehole measurements?

No, this value comes from a steady-state temperature profile corresponding to a geothermal heat flux of 50 mWm^{-2}. This is described in the manuscript.

Line 351: *For Samoylov, we used the same initial ground temperatures as Nitzbon et al. (2019), as we used the same forcing data as this study: 0 m depth: 0.0°C; 2 m: -2.0°C; 5 m: -7.0°C; 10 m: -9.0°C; 25 m: -9.0°C; 100 m: -8.0°C; 1100 m: +10.2°C. These values are based on borehole data from 2006 and a steady-state temperature profile corresponding to a geothermal heat flux of 50* mWm^{-2}.

L320: what is the total depth of the soil column, Table 2 lists properties for 0-9 m but L317 provides soil temperatures with depths of 0-1100 m.

We thank the reviewer for pointing out that this was not clear and we changed the manuscript accordingly. The entire model domain is 100 m, soil temperatures in greater depths are given for interpolation of the ground temperatures in greater depths. Table 2 lists only the characteristics of the upper 9 m, as the new stratigraphy class is applied here.

Line 317: *We use a model domain reaching from the surface to 100 m depth, which is described by a stack of two different ground classes. Below 9 m depth, soil mechanical processes are not accounted for due to frozen conditions and high total stress during the entire simulation. Therefore, we apply the stratigraphy class* GROUND_freeW_seb*, which operates without soil mechanical processes and water balance, as described in* Westermann et al. *(2023). Between the ground surface and 9 m depth, we use the new stratigraphy class GROUND_freezeC_RichardsEq_seb_pressure as presented in this study.*

Caption of table 3: *Stratigraphies used for the model scenarios in this study in the upper 9 m of the model domain.*

Line 361: *The initial vertical resolution of the grid cells increases stepwise from the surface to greater depths (0-1 m depth: 0.05 m; 1-5 m depth: 0.1 m; 5-10 m depth: 0.2 m; 10-20 m depth: 0.5 m; 20-50 m depth: 1 m; 50-100 m depth: 5 m). The fine resolution in the top layers allows us to analyze the ground temperatures in detail in the upper soil layers.*

L318: where did this geothermal value come from? Any reference?

We added the reference.

Line 353: *These values are based on borehole data from 2006 and a steady-state temperature profile corresponding to a geothermal heat flux of 50* $\mathrm{mW\,m^{-2}}$ *(Langer et al., 2013).*

Table 2: what is the residual saturation?

We assume no residual water as Westermann et al. (2023).

Line 119: *[...]we first calculate the matric potential for unfrozen conditions* ψ_0 $[\mathrm{m}]$ *from which the matric potential in a frozen state* ψ $[\mathrm{m}]$ *and finally the water content can be inferred (assuming no residual water).*

Caption of Table 3: *We assume no residual water.*

L330-335: This is totally confusing, the authors mentioned considering the undecomposed organic matter and now they started talking about peat which has totally different hydraulic and thermal properties. Please explain whether are you still using sandy soil properties for the organic layer and how this part (L330 onwards) is related to L298. I would assume most of the organic material at Samoylov is (partially)decomposed.

The upper 0.15 m represent the moss layer on top and is applied for each model run. In contrast, the organic soil in model scenario *S-peat* represents peat (0.15-9 m depth), which has different characteristics than the moss layer (Table 3). We clarified this in the manuscript.

Line 330: *Hereby, we use the same soil properties in the upper 0.15 m as in the multi-layer stratigraphy to account for the insulating effect of the organic-rich moss layer on top. The poorly decomposed organic material features coarse pores with unknown values for the van Genuchten parameters n and α. Therefore, it is phenomenologically set to the properties of coarse-grained (sandy) soil, which has a broadly similar retention characteristic. Between 0.15 and 9 m depth, we set homogeneous soil properties of one soil type, distinguishing between sand, silt, clay and peat. They feature different characteristics for volumetric fractions of mineral and organic content, saturation, initial void ratio, residual stress, compression index, permeability, α coefficient and n coefficient. The chosen values for the different settings can be found in Table 3.*

L342: "effect of different soil types" on what?

We added this in the manuscript.

Line 383: *A comparison of the model runs S-sand, S-silt, S-clay and S-peat show the effect of different soil types on ice segregation and thaw consolidation.*

Table 3. If out of 12, 11 scenarios are for the Samoylov site then what is the purpose of including Bayelva site? Both sites should be treated similarly or just remove Bayelva site from the manuscript and focus on one site.

The Bayelva site is important to analyse the difference between a maritime and continental setting. Therefore, the forcing data of Samoylov cannot be made synthetically warmer (see reply to major comment above). Instead of removing Bayelva from the manuscript, we added two more model scenarios at the Bayelva field site.

L365: Polygon centers are not representative of the entire polygon. Active layer thickness varies among troughs, rims, and centers. Although ALT varies across microtopographic locations, it would be reasonable to compare it against the average ALT (average of rims, troughs, and centers).

We explain in the revised version of the manuscript, why we only compare our model results against polygon centers.

Line 411: *For the Samoylov island site, we perform model validation focusing on polygon centers where segregated ice is normally found. In detail, we compare the validation run S-val designed to represent the typical ground stratigraphy of polygon centers at the site (Sects. 2.3, 2.4) to published in situ temperature data from 0.05 m and 0.40 m depth at a polygon center near the northern shore of Samoylov Island (Boike et al., 2013; Langer et al., 2011a, b; Westermann et al., 2016).*

L370-374: Have you looked at the comparison of the observed and simulated snow depths to support your reasoning for the mismatch? Have you compared the forcing data to any nearby climate station? Also, since you have taken the observed soil temperature data from a polygon center, was the polygon center inundated during the fall freeze-up? So it is not just the forcing (snow depth) that could affect freeze-up.

For our simulations, we limit the snow depth to 45 cm, following Westermann et al. (2016), who reasons this choice as following:

"In addition, the maximum snow depth in CryoGrid 3 is restricted to 0.45 m, the approximate height of the polygon rims above the centers. Snow depth observations of Boike et al. (2013) show that the maximum height in polygon centers rarely exceeds these values, since further snowfall is largely blown away by consistently strong winds."

We explain this in our manuscript:

Line 395: *For Samoylov, we set a threshold for the snow depth to account for observed snow ablation due to wind drift following the approach of Westermann et al. (2016). As the snow in the polygon centers is protected by the surrounding rims, we set a value of 0.45 m, which is in line with measured snow depths in polygons centers on Samoylov Island (Boike et al., 2013).*

For the validation run, we have temperature data from a polygon available, which is partially disintegrated with lower polygon rims (Boike et al., 2013), leading to overall lower snow depths there. This information is given in the manuscript:

Line 398: *For validation, we use temperature measurements in a polygon center. However, these measurements are not conducted in the middle of Samoylov island but close to the edge of the island where the wind drift is stronger and the height of the rims above the polygon center is reduced. Therefore, we reduce the maximum snow depth for the validation run to 0.20 m.*

Boike et al. (2013) does not describe inundation of the polygon centers.

L382: this is confusing again, at L317 it says -7C at 5 m depth, here is it -9C at 5 m depth. Why is that?

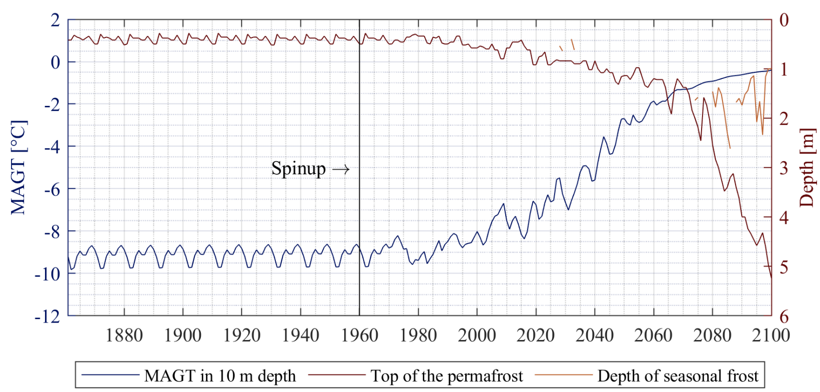
As we used the same forcing data as Nitzbon et al. (2020), we used the same initial ground temperatures to be consistent. We conduct long-term simulations and the temperature profile in the uppermost meters is independent of the initialization after a few years, so that this should not influence the model results. We state this in the mansucript:

Line 351: *For Samoylov, we used the same initial* ground temperatures as *Nitzbon et al. (2019), as we used the same forcing data as this study.*

Line 356: *Since we conduct simulations on at least centennial timescales and the temperature profile in the uppermost meters becomes independent in initialization after a few years, the initial temperatures do not affect simulation results.*

Figure 5: How do these results differ from those of Nitzbon et al. (2020)? Other than you not considering the ice-rich zone

Nitzbon et al. (2020) conducted coupled simulations, taking into account the 3D nature of the ground ice distribution in polygonal terrain, which was found to strongly influence the model results. Therefore, the results cannot be compared to 1D simulations here, so that we compare to Westermann et al. (2016), who present 1D simulations for Samoylov with the same forcing data. To do so, we changed figure 5, so that it shows mean annual ground temperatures at 10 m depth as in Westermann et al. (2016). Figure 5 was changed as following:



Mean annual ground temperatures $MAGT$ in 10 m depth, depth of the top of the permafrost and the seasonal frost in the model scenario S-clay. *The spin-up covers the years 1860 to 1960. $MAGT$ warm from around -9°C until the 1980s to values close to the thaw threshold at the end of the 21st century. The depth of the permafrost table deepens from around 0.7 m to 5 m, so that the seasonal frost doesn't reach the same depth towards the end of the century and a talik develops.*

Line 436: *The modelled changes in annual ground temperatures are in the same range as simulation results for the Samoylov field site in Westermann et al. (2016), as pointed out by the following comparison of approximate ground temperatures in 10 m depth (data of Westermann et al. (2016) shown in brackets): -9 °C (-10 °C) during the spin-up, -5 °C (-5 °C) in 2040 and -1 °C (-1 °C) in 2090.*

Figure 6: What is the focus of this figure, ice segregation or ALT or talik? The section heading says Formation of ice segregation, but the ice segregation and its formation are hard to see (visually) which makes it difficult to understand what is going on. A better comparison here would be to compare simulations with and without the formation of ice segregation and how much will it affect project ALT.

The focus of this figure should be the ice segregation, i.e. the volumetric water and ice content in the permafrost. The colours indicate, where segregated ice is formed, which is visualized with the dark blue colours at the top of the permafrost. To clarify this, we added an additional statement in the caption. The same was done for figure 9 (now figure 10) and figure 11 (now figure 12).

Caption figure 8: *Sum of volumetric water and ice content in the permafrost on the reference date August 31 for the model run S-clay. Dark blue colors indicate an increased volumetric water and ice content at the top of the permafrost, where ice segregation takes place. Values in the thawed layer are not displayed. The spin-up covers the years 1860 to 1960. The sum of volumetric water and ice content is shown as soil water can still occur below freezing temperatures dependent on the soil type. For scenario setup see Table 4.*

We agree that it is hard to see visually the ground heave due to ice segregation. We therefore added a plot with surface elevation change to the figures 8, 9 and 11 (numbering in revised version) and discussed it in the text.

Furthermore, we run a comparison with and without ice segregation. Please check the reply on the major comment in the beginning of this document.

L421-423: this contradicts the formation of more segregated ice in the future. Figure 7 does show that B-clay leads to more segregated ice.

We rephrased the paragraph to clarify, that the higher amount of segregated ice in the *B-clay* scenario can be explained by the stabilization of the permafrost table.

Line 481: *During the time period 2020 to 2100, B-clay builds up more segregated ice in deeper soil layers. This can be explained by a slower increase of the active layer thickness compared to Samoylov Island, so that the permafrost table stabilizes* *for longer time periods, enabling new formation of segregated ice. As a consequence, less ground subsidence takes place for the Bayelva run.*

In section 2.2, the authors talked about different simulations, but it is not clear what type of simulations are planned in this work. Please clearly state the set of simulations or summarize them in a table. A simulation description is lacking. It would help better follow the description if this section is split into field sites and forcing data.

Section 2.2 gives an overview about the field sites. We included the forcing data in this section for a better text flow, as the manuscript already contains lots of subsections.

The manuscript contains Tables 3 and 4, which summarize the different scenarios including details on stratigraphies, model period and model forcing. Section 2.4 describes the model scenarios. We added a link to this in section 2.2, so that the reader can see easily, where to look up information about the model scenarios.

Line 273: *To demonstrate the capabilities of the new model, we perform sensitivity studies (Sect. 2.4) for two different permafrost sites with strongly different climate conditions*

Section 3.4. Table 2 shows that the top 15 cm has identical soil properties. What is actually the S-peat scenario? How does it differ in terms of peat thickness in the soil from other scenarios (soil types)? Peat has different thermal and hydraulic properties (L344) than other soil types and leads to shallower ALT, but the thickness of the peat is important. If in summers, the peat keeps the ALT colder than in winters it helps prevent escaping heat from the soil. I am not sure if the entire soil column is peat in the S-peat scenario. If the entire 9m (or so) column has peat, is it realistic?

The *S-peat* scenario contains the moss layer on the top as all other scenarios. Below, it consists of peat until 9 m depth. The characteristics are given in Table 3. We agree with the reviewer, that the thickness of the peat is important, when representing a specific field site. However, we do not aim to represent Samoylov with the *S-peat* stratigraphy, but it is part of a sensitivity study, which should show the effect of different soil types on ice segragation and thaw consolidation. To do so, we simplify the stratigraphy to one soil type, similar with sand, silt and clay. We explain the concept of the sensitivity study in the manuscript. Furthermore, we clarify the *S-peat* scenario in the revised manuscript. The soil characteristics of peat are given in Table 2.

Line 329: *To analyze the performance and sensitivity of our model for Bayelva and Samoylov, we used simplified stratigraphies to provide standardized and comparable model setups. Hereby, we use the same soil properties in the upper 0.15 m as in the multi-layer stratigraphy from Samoylov island (see above) to account for the insulating effect of the organic-rich moss layer on top. The poorly decomposed organic material features coarse pores with unknown values for the van Genuchten parameters n and α. Therefore, it is phenomenologically set to the properties of coarse-grained (sandy) soil, which has a* *broadly similar retention characteristic. Between 0.15 and 9 m depth, we set homogeneous soil properties of one soil type, distinguishing between sand, silt, clay and peat. They feature different characteristics for volumetric fractions of mineral and organic content, saturation, initial void ratio, residual stress, compression index, permeability, α coefficient and n coefficient. The chosen values for the different settings can be found in Table 3.*

L586: “complex lateral processes”? Unless you are considering 3D simulations at a larger scale, which are mostly not feasible for permafrost regions, what complexity can lateral processes bring? My guess is even using CryoGrid with three tiles should not make it complex, but just an additional process.

We agree with the reviewer and removed "complex" from the manuscript.

Line 704: *In its present form, the model is one-dimensional and hence does not account for ~~complex~~ lateral processes, e.g. lateral water flow due to cryosuction.*

Check for typos

We checked the entire manuscript for typos and corrected them.

Map showing the location of the field sites is missing

We added a map showing the two field sites.

Figure 5: Location of Samoylov Island in the Lena River delta in northeastern Siberia and of the Bayelva climate station on Svalbard. The aerial images of the surroundings of the field sites are shown in the lower left for Samoylov Island and on the lower right of the Bayelva climate station.