(Note: The reviewer's comments are in black and the author's responses are in blue. Unless specified otherwise, the line numbers quoted in our responses are with reference to the revised manuscript with track changes turned on.)

This paper presents a useful addition to the land modeling community, with a well-described model of permafrost that can be coupled to other land models and has code and documentation available. The need for such a model I felt was well-justified and the paper provides detailed model equations and background, describes and demonstrates the model capabilities, and compares a variety of model outputs against observational data. Overall the model appears to produce reasonable results, and, with some modifications, I believe this paper will make an excellent addition to GMD.

We thank the Reviewer for the positive assessment of the manuscript.

My primary concerns are with the performance claims that I believe need additional support from model validation and with the minimal discussion of model limitations. While the model is compared to observational data, several aspects of this could use further clarification. First, it is unclear to me why the years selected are necessarily comparable to the model results and why different years are chosen for different comparisons, which may just speak to a need for some additional justification for this in the text.

We have addressed this concern by explaining why periods for analysis were chosen and by making the periods more consistent when possible. Please see a more detailed response below to the specific comment about this.

Additionally, differences between model results and data are quantified frequently in terms of RMSE but the reader has no real perspective to understand the relative significance of these RMSE values. This could be addressed by including the range or uncertainty of the observational and simulated values and by giving some measure of statistical significance.

We thank the Reviewer for this suggestion. We have removed the use of RMSE from section 4.1 (lines 352 to 366) and instead present medians and IQRs for observed and simulated variables to be consistent with the rest of the results sections. We chose to not use a measure of statistical significance because sample sizes can be increased with simulations to artificially inflate statistical significance derived from parametric approaches (Lucash et al. 2019). We now describe this in lines 346-347.

Also, some discrepancies between model results and observations seem to not really be discussed or have only limited discussion, such as the model's near-surface permafrost extent by aspect being considerably worse in the north aspect than others, later maximum annual active-layer depth or earlier freeze depth. It would be helpful to address these more and help the reader understand potential causes and the implications of these for interpreting other model results.

We now include two paragraphs in the discussion section (lines 429-451 and 452-461) that describes a number of potential reasons why at times simulated timing of maximum thaw depth and maximum snow depth diverged from observations. We also describe discrepancies between simulated and observed permafrost distributions with aspect at the landscape scale. We point to specific underpinning mechanisms that could be better represented in the model in the future.

Model limitations (such as critical assumptions or missing processes and their implications for results) are only lightly touched on in the conclusions section, and I believe the paper would benefit from further elaboration on these points, to help the reader put this model in perspective and better understand how to consider its results.

Please see response to comment above.

In the interest of putting this model in perspective, it would additionally be useful to see some comparison of iLand's performance with and without this new module to some of the observational datasets to demonstrate the value of this module beyond providing additional outputs.

We have added lines 401 – 403 to describe differences by running the model again for 300 years with the permafrost and SOL turned off. We have added such a map as figure 8.

Specific comments:

Equation 9 - Would be helpful to add units to the constant to make this conversion clearer to the reader

We have revised line 125 to include units for the constant in equation 9

L.280 Unclear to me why we expect these model simulation years to correspond to these data years. I suggest adding some clarification on this, here and in other places where model results are compared to time-specific data.

We wanted to ensure that we were capturing the landscape distribution of permafrost that was consistent over time and not locations that might be transiently considered permafrost. We have revised the window for permafrost to be for 261-300 and clarify in lines 303-305 that a 40-year window (year 261-200) was selected as a multi-decadal period that aligns with the period used to calculate SOL combustion and postfire seedling density. The time periods for calculating SOL combustion and postfire tree seedling density were chosen to ensure a sufficient sample of fires for analysis while being cognizant of the computational intensity of these calculations. We now state this in lines 318-320. We kept the period of simulation year 201-300 for evaluating the fire regime because fires in Alaska are large and infrequent leading to greater stochasticity in outcomes. We decided at least a century was necessary as it aligns with the mean historical fire return interval in interior Alaska (Johnstone et al. 2010). This is now described in lines 291-292.

L.353-354 Somewhat unclear. Would be helpful to elaborate more on the importance of aspect and mention the differences in results between north versus other aspects.

We have revised lines 377-378 as follows: “Simulated permafrost was over represented on north-facing slopes, as compared to the benchmarking product, but corresponded well on all other aspects.”

L.375 Would be useful to add comparison here to iLand results without the permafrost module

We have added lines 401 – 403 to describe differences by running the model again for 300 years with the permafrost and SOL module turned off. We have added such a map as figure 8

L.397-398 The claim “Benchmarking results demonstrate the model recreates temporal and spatial patterns consistent with observations” feels not completely supported by the text and I believe needs more justification or caveats.

We have revised lines 426-428 as follows: “With some exceptions discussed below, benchmarking results generally demonstrate the model recreates temporal and spatial patterns consistent with observations at stand to landscape scales over days to centuries.” In the discussion section, we further address caveats including the simplicity of the snow model, not representing effects of forest structure on microclimate and consequences for snow and permafrost, and the effects of tree species composition on landscape patterns of permafrost distribution in lines 429-461.

L.398-399 The claim “Our model will contribute to improving 21st -century projections of boreal forest change.” is not really supported by the current results, but would be by the addition of a figure or data showing improvements to iLand’s projections with the addition of the permafrost module.

We have added such a map as figure 8.

Figure 3. The representation of frozen and unfrozen soil in this figure is somewhat confusing, as it seems like there should be two different fills - one for observed, one for simulated - but instead there is only the one. Can this be separated out or otherwise clarified?

We thank the reviewer for pointing this out. We have revised the caption of figure 3 by adding “Grey fill represents simulated frozen soils. Blue fill represents simulated unfrozen soils.” (lines 851-852).

Figure 5. Maybe add the years for the observed dataset in this figure caption, since the simulation years are given.

We added the years for the benchmarking product (1990-2013) to the figure caption in lines 864-865.

Figure 7. Some clarification on the caption for Panel B would be helpful, as it took me a minute to figure out what was meant here.

We have revised lines 876-879 as follows: “Figure 7. A. Tree-species composition in a 61,000 ha forested landscape of interior Alaska used to initialize iLand. B. Changes in tree species dominance over 300 years of simulation. Pima (*Picea mariana*) = black spruce, Pigl (*Picea glauca*) = white spruce, Potr (*Populus tremuloides*) = trembling aspen, Bene (*Betula neoalaskana*) = Alaskan birch”

Works cited

Johnstone, J. F., F. S. Chapin, T. N. Hollingsworth, M. C. Mack, V. Romanovsky, and M. Turetsky. 2010. Fire, climate change, and forest resilience in interior Alaska. Canadian Journal of Forest Research 40:1302–1312.

Lucash, M. S., K. L. Ruckert, R. E. Nicholas, R. M. Scheller, and E. A. H. Smithwick. 2019. Complex interactions among successional trajectories and climate govern spatial resilience after severe windstorms in central Wisconsin, USA. Landscape Ecology.