We sincerely appreciate the time and effort the reviewer has dedicated to evaluating our manuscript. While we respect the review process, we respectfully disagree with the three main points raised in the evaluation that led to the rejection decision. We believe that our responses below provide further clarification to address any misunderstandings. Therefore, we kindly request that the decision be reconsidered.

Reviewer Comment:
"The study calibrates parameters using synthetic data, which is generated from the same model being calibrated. Even if the proposed method can successfully recover the set of parameters, I don't see the value in this analysis for advancing our understanding of permafrost ecosystem dynamics."

Response:
As stated in Lines 152–159 of the introduction, our study is methodological in nature, aiming to enhance the efficiency of the calibration process in a complex ecological model. This is the first time that the MADS algorithm has been applied to calibrate parameters in a terrestrial ecosystem model. Our primary objective was to demonstrate the effectiveness of the method itself, and the use of synthetic data is a standard and widely accepted approach in many scientific disciplines for this purpose. Synthetic data provides a controlled environment to validate calibration techniques before applying them to real-world data, where additional uncertainties and biases exist.

Furthermore, in response to the reviewer's request, we incorporated a calibration using real observations. This additional analysis demonstrates satisfactory calibration for over 90% of the target values (see Figure 5). As outlined in the results, this quantification confirms the validity of our calibration approach against real observations.

Finally, the objective of our manuscript was not to address potential structural issues that the calibration results may have highlighted. Instead, our approach improves calibration efficiency, enabling faster and more effective calibrations. This, in turn, allows for a broader application across diverse ecosystems, enhancing the representation of spatial heterogeneity in boreal and Arctic landscapes.

Reviewer Comment:
"If the model has structural errors that misrepresent a process, calibrating based on a biased or erroneous simulation would be meaningless."

Response:

While all models are simplifications of natural processes, this does not inherently mean they contain structural errors. Additionally, observational data itself carries significant uncertainties, as highlighted in our manuscript.

Calibration remains a valuable process because it helps identify areas where improvements are needed, both in data collection and in the mathematical representation of physical processes. Even if a model contains some level of structural uncertainty, calibration provides critical insights into parameter sensitivity and model behavior, helping to refine future versions of the model.

Reviewer Comment:
"However, if the authors can demonstrate that the method performs well on ‘ground truth’ data derived from in-situ measurements and remote sensing, I would find it more valuable. But according to the comparison between observed and calibrated results (Figure 5), the performance seems not ideal. Overall, I think this study is going in the wrong direction."

Response:
We strongly disagree with this assessment. Our study is moving in the right direction precisely because we openly acknowledge and address model-data mismatches rather than obscuring them. Furthermore, our calibration shows satisfactory results for more than 90% of our target variables, showcasing the validity of our method.

Moreover, as we discuss in our manuscript, the calibration performance may vary depending on site-specific conditions. The observed mismatches in Figure 5 do not indicate a failure of the methodology but rather highlight the inherent challenges of working with real-world data. Such challenges reinforce the need for improved calibration techniques, which is exactly what our study aims to contribute to.

Reviewer Comment:
"The authors employ a hierarchical approach to parameter calibration, but the rationale for the chosen order (e.g., calibrating vegetation parameters before soil parameters) is not clearly justified. Would reversing the calibration sequence affect the results? How sensitive is the method to the order of parameter calibration? Addressing these questions is critical for evaluating the robustness and generalizability of the proposed approach."

Response:
We addressed this concern in our initial response. To reiterate, the first step of our hierarchical calibration approach is required and consistent across applications and works well in most cases. The sequence of subsequent steps (Step 2 onward) is flexible and can be adjusted by the user. The key principle behind our approach is that calibrating all parameters simultaneously often

leads to poor results due to an excessive number of degrees of freedom. By breaking the calibration into smaller, sequential steps, we improve convergence and parameter estimation. We refer the reviewer to Section 4.7 and the final sentence of Section 4.8:

"While long-term soil parameter calibration inherently influences vegetation dynamics, the most significant changes in vegetation-related parameters typically occur during short-term model runs, resulting in minimal net changes over extended simulations."

To further clarify our rationale, we have added the following text to Section 4.7.

"Reversing the calibration sequence and starting from soil parameters is not only impractical in the context of our model, but also computationally inefficient. Vegetation-related parameters are calibrated first because vegetation carbon pools reach equilibrium significantly faster than soil carbon pools whereas soil pools require longer timescales to stabilize. Beginning with soil parameters would thus introduce unnecessary complexity and substantially increase the total computational cost of the calibration process. In addition, while the choice of calibration sequence may lead to slight variations in the final parameter estimates, our results demonstrate that the proposed "hierarchical approach" (breaking the parameter sets into smaller subsets) effectively recovers parameter values, even when for 90% parameter range variance. As we showed in this study, well-calibrated parameters exhibit a narrow range of uncertainty, reinforcing the robustness of the method."

Reviewer Comment :
"The manuscript acknowledges the challenge of equifinality but does not present a clear strategy to address it. The results indicate that the choice of initial parameter values has a greater impact on calibration outcomes when using observational data compared to synthetic data. This suggests that the proposed approach may be less reliable when applied to real-world ecosystems, where uncertainties and biases in observations are unavoidable. The study would benefit from a more rigorous discussion of how to mitigate equifinality and improve the method's performance on real-world datasets."

Response:
Equifinality is a well-documented mathematical challenge and a fundamental issue in many modeling applications, particularly in environmental and ecological modeling. It falls within the class of ill-posed problems, where multiple parameter sets can yield similar model outputs. While we do not claim to solve equifinality in our study, we explicitly acknowledge its presence and employ established techniques, such as parameter perturbation and multiple calibration runs to assess its impact.

Our approach involves conducting multiple calibration tests and evaluating convergence. If repeated calibrations converge toward a consistent parameter set, we gain confidence in the

results. Respectfully, the reviewer's expectation that equifinality can be fully eliminated suggests a misunderstanding of the inherent limitations of calibration in complex environmental models.

In summary, the statement that our method is unreliable in real-world ecosystems overlooks the fact that all models operate under uncertainty. Our study does not claim to eliminate uncertainty but rather provides a structured approach to improving parameter estimation.