Response to Referee 2

AUTHOR
Julien Vollering, on behalf of all co-authors

https://doi.org/10.5194/egusphere-2025-1046-RC2

General Comments

Comment 2.1

> This article focuses on improving peat depth mapping in two distinct peatland landscapes in Norway using a digital soil mapping approach. The study employed the Random Forest (RF) algorithm to model both peat presence and depth, using high-resolution terrain data, remotely sensed radiometric data, and polygonised peat depth data from an existing map. The RF models were calibrated using field-measured data points collected across both regions, and variable importance was analysed.

Comment 2.2

> Overall, this research addresses a significant challenge in peat depth mapping. The title is clear and likely to attract attention from readers interested in this topic. However, I have some concerns regarding the writing style. Certain sections use jargon excessively or are phrased in a way that feels less scientific, which may hinder clarity and accessibility for a broader audience.

Thank you for your positive comments on the research and title, and for your suggestions to improve the writing style. We have gone through the manuscript and revised the language to make it more formal and scientific, while aiming to retain clarity and accessibility. We have also tried to reduce jargon where possible, while retaining the necessary technical terms.

Specifics Comments

Comment 2.3

> While the title is clear, I personally find it somewhat overconfident in summarising the results. This is mainly due to the relatively low performance of the RF models and the minimal difference observed between the terrain-only and terrain-plus-radiometric models. To better support the claims, it would be helpful to include a statistical test (e.g., a t-test) to assess whether the performance differences between models are statistically significant. Furthermore, incorporating a stand-alone model that uses only radiometric data as predictors would allow for a more balanced evaluation of the variable groups and help clarify the specific contribution of radiometric inputs in predicting peat depth in Norway.

Thank you for this suggestion. To clarify: the minimal marginal improvement in the terrain-plus-radiometric models compared to the terrain-only models is in fact evidence for the terrain variables being stronger predictors. Nevertheless, you are right that adding a radiometric-only model would provide a clearer comparison of the predictive contribution of radiometric versus terrain data. We now present models with all seven possible combinations of the three variable groups. These expanded results make clearer the predictive gap that we highlight in our title.

We have also added statistical tests to assess the significance of the differences in model performance. The results of these tests are now presented in tables in the appendix. The results show that the differences in model performance are statistically significant in most cases, which supports our claims about the relative importance of terrain variables in predicting peat depth. Specifically, 8 of the 12 direct comparisons between radiometric and terrain variables (Radiometric–Terrain and RadiometricDMK–TerrainDMK configuration pairs * 3 metrics * 2 sites) are statistically significant at the 0.05 level. We now describe these results in the Results section, to better support the claim in the title.

Comment 2.4

> Figure 1: While the terrain-look map provides a general view of the study area, the relief of the regions remains unclear. Consider adding a clearer topographic representation (e.g., contour profile) to better illustrate landscape variation. Additionally, it would be helpful to explain the rationale for selecting Skrimfjella, which appears to have relatively limited peat coverage, especially given that the adjacent region seems to contain a larger peatland area.

Thank you for the suggestion to make the relief of the study areas clearer. We have added contour lines the maps of the study areas. We have not added contour profiles between specific points since we think that they may not be as useful for readers as the contour lines, but we are happy to add profiles if the contour lines are not enough.

We have also clarified the rationale for the delineation of the Skrimfjella study area, especially with respect to its limited peat coverage. In short: we were also investigating peatland extent mapping so low coverage was not disqualifying, and accessibility was also an important factor for this field work.

Comment 2.5

> It is unclear how the peat depth data are distributed, both statistically and spatially. While the sampling design is described in the text, I recommend including a figure showing the spatial distribution of the data points to help readers better understand the coverage and representativeness of the dataset. Additionally, a basic statistical summary (e.g., mean, median, range, standard deviation) of the peat depth values would be beneficial to provide context and a clearer picture of the conditions being modelled.

Thank you for this suggestion. We have added a figure showing the statistical and spatial distribution of the peat depths in the Results section. We have taken care to use the same map format as in the figure that presents land cover in the study areas, which we hope will help readers see the spatial representativeness of the dataset.

Comment 2.6

> Since the radiometric data are originally at 50 m resolution and the terrain data are at 1 m

> resolution, the method used to resample these datasets to a common 10 m resolution could influence model performance. In particular, the use of cubic spline resampling for the radiometric data may affect how well its spatial variability is represented, especially when compared to the aggregated 10 m topographic data. I would be interested to hear your thoughts on how this resampling approach might have impacted the results.

We appreciate this attention to the potential impact of radiometric resampling methodology on our results. To address this concern, we added to our code a sensitivity analysis comparing cubic spline and bilinear resampling methods for downscaling the 50 m radiometric data. Using field measurement locations from both study sites, we extracted radiometric values resampled by both methods and calculated correlations for each radiometric variable (K, Th, U, TC). The correlations between cubic spline and bilinear resampling methods were very high, with the lowest correlation being 0.995. This near-perfect correlation demonstrates that the choice of resampling method has negligible impact on the spatial representation of radiometric variability at our target resolution.

The consistently high correlations across all radiometric variables indicate that our conclusions about radiometric predictive performance would remain unchanged regardless of resampling method. Since we now mention this fact in the *Materials and methods*, we do not raise this point specifically in the *Discussion* – in the interest of brevity.

Comment 2.7

> I found the section on sampling design and peat depth measurement to be overly detailed for the main text. While this information is valuable, it may be more appropriate to move some of it to the appendix or supplementary material, as it is less central to the analysis and interpretation of results. This would help improve the flow and focus of the main manuscript.

Thank you for this helpful suggestion. We agree that this an acceptable way to reduce the length of the main text and improve its readability, especially with the added figure showing the spatial distribution of depth measurements (in response to comment 2.5). We have moved the bulk of the sections on *Peat depth sample selection* and *Depth measurements* to the appendix, while retaining brief summaries of these in the main text. This solution seems to be allowed under the journal's guidelines for manuscript composition.

Comment 2.8

> In the model interpretation section, it appears that three different methods were used to assess variable importance: FIRM, permutation importance, and Shapley values. However, the implications of using these different approaches are not clearly discussed. Each method captures different aspects of variable influence and may lead to different interpretations. Could you clarify how the results from these methods align or differ, and what their respective implications are for understanding the drivers of peat depth in your study?

Thank you for this suggestion. We have added text to the *Materials and methods* and *Results* sections to clarify how the three variable importance methods relate to each other. We have also added rank correlations to our figure on variable importance, to highlight the similarities and differences. In short, we find that Shapley and permutation importance are very similar for our application, while

FIRM values capture a different aspect of variable importance.

Since the different methods are broadly congruent and their differences do not affect our interpretation much – and in the interest of brevity – we do not go into this point in our *Discussion*.

Comment 2.9

> I recommend adding a dedicated section to summarise the key results and findings. This would help readers clearly see how the study addresses its original research aims and questions. A concise summary at the end of the results or discussion section would also improve the overall structure and coherence of the paper.

Thank you for this helpful recommendation. We have added a *Conclusions* section to summarize the key findings of the study. This summary highlights the main results and their implications, which we hope will help readers more easily see how the study addresses its research aims of quantifying predictive accuracy and identifying key predictors.

Technical Corrections

> Table 2: I think you need to move this table somewhere after it is mentioned in the main text. At first, I was quite confused with this table. What does that percentage stand for?

We have moved the table after its first mention and adjusted the table caption to clarify that the percentages represent the proportion of the total number of 10 m cells across the stratifications. We have also added standard deviations to the mean depths, in accordance with comment 2.5.

> Line 394-397: I don’t quite get the point of this section. To my understanding, you compared between the quartile predictions and the observed data to see the prediction interval coverage probability. If so, I think it would be better to plot this prediction interval together with the observed data in scatter plot, like in Figure 3.

Thank you for identifying this confusion. We have revised the text to clarify that we are comparing the quartile predictions to the observed data, to follow best practices in digital soil mapping.

> Line 413-414: which curve? The information inside the parentheses is not clear to me.

Thank you. Revised for clarity.

> Figure 4: is the null derived from the calculation between observations and assumptions of 30 cm peat depth? Where is the standard error come from?

We have removed this figure along with all sections about peat extent mapping. The null did reflect the error between observations and assuming 30 cm depth. The stadard errors on the other error estimates were from the cross-validation folds.

> Line 464: “[illegible]”?

> Line 464: "low hanging fruit" ?

Revised.

> Line 515: This sentence is unclear to me.

Revised.

> Line 531: “0% in prediction” what does this sentence mean?

Removed along with all sections about peat extent (occurrence) mapping.

> Line 614: What are the ‘both indicators variables’ referring to?

Revised for clarity.