Dear Editor,

We thank Reviewer #1 for their comments. We reproduce their comments in full below and provide our replies in blue font.

—

Reviewer #1:

The authors present PeatDepth-ML, a machine-learning framework for predicting global peat depth using a large compilation of peat depth measurements and environmental covariates. They extend existing peatland mapping approaches by incorporating additional predictors, revised spatial cross-validation, a custom metric targeting deep peat, and a bootstrapping strategy to assess sensitivity to sampling bias. Model performance is evaluated with blocked leave-one-out validation, and the resulting global peat depth map is used to estimate global peat carbon stocks, which are found to be consistent with previous studies.

I think the work is relevant for the journal and generally well-executed, though I think some revisions are in order prior to publication. I will give detailed list of comments in the following. Thank you for your work.

Thank you for your positive assessment of our study.

Detailed comments:

Lines 49 and 66: "machine learning" --> use abbreviation "ML".

Changed

Line 92, Figure A1: I think Figure A1 is quite important, presenting the peat data distributions. Why not include it in main text instead of in appendix?

Originally we felt that 13 figures in the main text was already quite a few. However, we have now moved it into the main text.

Line 97: "However, grid cells with zero peat depth consistently dominate..." --> explicitly state the percentage of zero peat depth as it is the substantial majority of the data. I think it is good to state as the data is quite, though naturally, imbalanced.

To remind the reader that there is a large amount of zero peat depth training data, there is a reference to Figure 2c (formerly Appendix Figure 1c) which shows the proportion of training data by depth on line 96. We prefer to not add further text on line 97 because on lines 106 to 113 we discuss how desert regions were dealt with, which we feel underscores the

influence of zero peat training data and gives it a more fulsome discussion than would make sense on line 97.

Line 185: "machine learning" --> "ML"

Changed

Line 189: What were the hyperparameters which were optimized? I did not see them listed.

We have now fixed this oversight by adding new text and a table to the appendix. We have added “The hyperparameters that were optimized for Peat-ML were also optimized for PeatDepth-ML (Melton et al., 2022). We performed our Bayesian hyperparameter optimization over 1500 model iterations and chose the values that produced the best results. These hyperparameters and their values can be found in Table A1 and are the same for every bootstrap run.”

Table A1. The final values of the LightGBM hyperparameters chosen to undergo Bayesian optimization. These are the same hyperparameters chosen for optimization as Peat-ML (Melton et al., 2022).

Name	Range	Optimized Value
boosting_type	gbdt, dart, goss	gbdt
num_leaves	10-50	44
n_estimators	50-300	235
learning_rate	0.005-0.4	0.05322234284123687
max_bin	25-300	80
max_depth	-1-15	7
subsample_for_bin	20 000-300 000	200 000
min_child_samples	5-60	50
reg_alpha	0-1	0.6886021585325435
reg_lambda	0-1	0.23338168719820718
colsample_bytree	0.5-1.0	0.7879219157568039

Line 192: "cross validation" --> "cross-validation"

Changed

Line 205: "don't" --> "do not"

Changed

Line 209: Add reference for LightGBM, maybe also fully open up the term. Lets not assume reader knows all the abbreviations by default.

Thanks, added reference. It seems like LightGBM should be an acronym, but in the original paper they don't define it as such. So we leave it as LightGBM, but now correct the missing reference.

Line 247: Did you mention somewhere how many predictors you had in total available for the ML runs? I would be curious to know this.

We give the number of selected predictors from the VIF stage, but had not previously given the total number. We have now added the number to the Table 1 caption as 'Environmental variable datasets provided as 397 potential peat depth predictors to PeatDepth-ML…"

Figure 8 and A1: I am not used to horizontal histograms or distributions being presented. Was there a particular reason for this? If not, why not use standard orientation in visualization (vertical bars), which, to my experience, is more common.

We chose horizontal so it can be easily mentally oriented with a soil profile.

Figure A1 caption: extra whitespace before ".", "...desert data ."

Corrected

Line 357: Open up the abbreviations, although well-known, the RMSE, MBE, NME. They are mentioned also in appendix more specifically, but good the clarify the abbreviations, once introduced.

We agree and have added the full names.

Line 362: Could you please elaborate on the null models a bit. Do you mean baseline models? Also on same line, notice extra period ". ."

The two null models are created by:

1) Taking the mean of the observations. This is indicated by the blue dashed line in Figure 12. The 'observed mean null model score' has an NME score of 1.0 by definition.
2) As in Kelley et al. (2013), the random null models are datasets equal in size to the results produced by PeatDepth-ML filled with randomly sampled (with replacement) values from the PeatDepth-ML training and testing data. We made 1000 of these null models using the version of the training and testing data that did not undergo any bootstrapping within its grid cells. The NME of each random null model was calculated using the formula in Table A2 (formerly Table A1) with a resulting mean score of 1.036 (Figure 12).

Taken together these provide NME scores that can place our bootstrap models NME score in context. Figure 12 shows our bootstrapped NME scores cluster around 0.6 indicating our

model framework is more skillful than either simply taking the mean of the observations or allowing a random sampling.

Line 370: "BLOOCV" Did you define this abbreviation, even though clear to myself. But still, define it earlier in the text when you mention cross-validation.

Yes, it was defined on line 189. It was then discussed in detail on lines 191 to 194 and also discussed the block sizes in the paragraph starting on line 198.

Figure 9: The legend is little bit unclear for me. What is "bootstrap results", what results? Maybe rephrase more clearly, if possible.

We have modified the labels to make them more clear:

1) ‘PeatDepth-ML all peat-present bootstrap results’ to ‘PeatDepth-ML predicted peat depths for all peat-present grid cells’
2) ‘PeatDepth-ML bootstrap results over training grid cells only’ to ‘PeatDepth-ML predicted peat depths over grid cells with training data’
3) ‘PeatDepth-ML bootstrap results over peat-present training grid cells only’ to ‘PeatDepth-ML predicted peat depths over grid cells with peat-present training data’