RESPONSE TO REVIEWER #1

The manuscript proposes a high-resolution forest-specific mapping approach for predicting soil bulk density and pH across China. It presents a substantial body of work and addresses a topic of interest, which has the potential to contribute to the field. However, in my opinion, the current manuscript requires major revision before it can be considered for publication.

Overall, the manuscript is informative and holds value but requires further refinement. The authors are encouraged to more clearly emphasize the importance and novelty of their work, revise redundant descriptions in Results while focusing on demonstrating statistical significance. With careful revision, this manuscript has considerable potential to make a meaningful contribution to the field.

Response

Dear reviewer #1,

We sincerely thank you for your insightful and comprehensive comments, which have been very helpful in improving the quality and clarity of this manuscript. In response to your suggestions, we have undertaken substantial revisions.

Specifically, we revised the Introduction to more clearly articulate the scientific motivation, urgency, and novelty of forest-specific, high-resolution mapping of soil bulk density and pH across China. We also carefully refined the Results section to reduce redundancy and to emphasize quantitative interpretation supported by appropriate statistical evidence, avoiding subjective or over-interpreted statements.

We believe that these revisions substantially enhance the rigor, transparency, and interpretability of the manuscript. Our point-by-point responses to all comments are provided below, and all corresponding revisions are marked in **blue** in the revised manuscript.

Best regards,

Jizhen Chen

Major comment 1

First of all, after reading the Introduction, I wasn't fully convinced of the necessity and urgency of this study. The Introduction section begins with very basic background information on forest soil, which is too general to establish a compelling rationale. The excessive introduction about methodology doesn't effectively build a case for the study's significance, either. For instance, the entire second paragraph is basically saying "a lot of people have done this", which may justify methodological reliability but not why this work is needed. The fourth paragraph focuses on the historical development of methodologies, which isn't the main goal of an Introduction. While building a nationwide forest soil profile database is potentially valuable, the current Introduction does not sufficiently highlight how this study advances beyond simply extracting forest-covered data from existing maps.

Response

We thank the reviewer for this constructive and important comment. We agree that the original Introduction did not sufficiently establish the necessity and urgency of this study.

In response, we have substantially revised the Introduction to adopt a more problem-driven

structure. General background information on forest soils and the historical development of digital soil mapping methodologies has been condensed (***Lines 37-60***). The revised text now explicitly emphasizes the limitations of existing national and global soil bulk density and pH products, which are largely derived from mixed-ecosystem samples and therefore fail to capture the distinct spatial heterogeneity and vertical structure of forest soils (***Lines 61-73***).

Importantly, we now clarify that simply extracting forest-covered pixels from existing soil maps is insufficient. Instead, we highlight the need for forest-specific modeling frameworks that explicitly account for ecosystem-specific processes and depth-dependent variability. This rationale is clearly articulated in ***Lines 79-80***, where we emphasize the ecological importance, spatial complexity, and current lack of high-resolution forest soil BD and pH estimates across China.

Major comment 2

Some findings are presented without statistical validation and therefore unconvincing. For example, L255 "BD prediction accuracy...peaking at intermediate depths (15–30 cm: MEC = 0.657) with lower accuracy in surface layers (0–5cm: MEC = 0.598) and deep layers (60–100 cm: MEC = 0.656)". Without testing for statistical significance, how can 0.656 represent "lower accuracy" compared to 0.657? Similarly, statements such as "all predictions maintained negligible bias ($|ME| \leq 0.019$) across depth intervals" lack a defined threshold for "negligible". Descriptions like "Conversely, pH predictions demonstrated superior accuracy: CV maintained strong performance across depths" appear subjective, without definition for "superior" or "strong".

Response

Thank you for for your important comment regarding the interpretation of model performance metrics. In response, we have revised the manuscript to avoid over-interpretation of small numerical differences in performance indicators and to remove subjective descriptors that were not supported by formal statistical testing.

Specifically, statements comparing prediction accuracy across soil depths (e.g., "higher" or "lower" accuracy) have been removed, as differences in MEC values such as 0.656 versus 0.657 are not statistically meaningful. Similarly, qualitative terms such as "superior," "strong," and "negligible" have been replaced with objective descriptions based on the reported ranges of MEC, RMSE, and ME values.

The revised text now focuses on presenting model performance in a descriptive and transparent manner, emphasizing the overall consistency between cross-validation and independent validation results, as well as the absence of systematic bias indicated by ME values close to zero.

These revisions can be found in ***Section 3.2 (Lines 275–288)***.

Major comment 3

Similarly, in the Result section, the authors keep emphasizing that their "patterns align with former maps", which further raises questions about the novelty and importance of this work.

Response

We thank the reviewer for this insightful comment regarding the interpretation of similarities between our results and existing soil datasets. In response, we have revised the manuscript to avoid overemphasizing pattern agreement with previous products in the Results section.

Specifically, statements such as "Macroscale patterns align with existing maps" have been removed from the Results, as simple visual consistency alone does not sufficiently reflect the novelty or contribution of this work. Instead, we have substantially revised ***Section 4.1 (Lines 451–511)*** to provide a more detailed and quantitative comparison with existing datasets (CSDLv2, ChinaSoilInfoGrids, and SoilGrids 2.0).

This revised discussion focuses on ecosystem-specific differences in both the vertical distribution and magnitude of forest soil BD and pH, highlighting discrepancies that are not captured by generalized soil products. In particular, we demonstrate that existing datasets fail to fully represent the non-linear depth-dependent pattern of forest soil BD and systematically predict higher BD values in deeper layers (60–100 cm), which may lead to overestimation of forest soil carbon stocks. Similarly, our results indicate consistently lower forest soil pH values compared to existing datasets, suggesting that ecosystem-specific acidification processes in forest soils are underestimated in generalized products.

These revisions clarify that the contribution of this study lies not in reproducing existing spatial patterns, but in providing forest-specific, high-resolution estimates that improve the representation of soil properties and associated ecological processes.

Major comment 4

Many descriptions in the Results section are excessive or repetitive (e.g., L268–270, L274–279), and some qualitative statements regarding spatial gradients (e.g., "BD values increase from the coast inland" in L271) are unclear.

Response

We thank the reviewer for pointing out that parts of the Results section were overly descriptive and repetitive, and that some qualitative statements regarding spatial gradients lacked clarity. In response, we have substantially revised the Results section to improve conciseness and clarity, and to shift from generalized qualitative descriptions toward a more quantitative and statistically supported characterization of spatial patterns.

Specifically, repetitive regional descriptions have been removed, and ambiguous statements such as "BD values increase from the coast inland" have been eliminated. Instead, we now quantify spatial patterns using latitudinal and longitudinal gradients, together with regional statistical summaries (boxplots), which provide a clearer and more objective representation of spatial variability.

To support this revision, Figures 3 and 4 have been redesigned (now ***Figures 4 and 5***) to explicitly illustrate depth-specific latitudinal and longitudinal trends, as well as regional differences in BD and pH. The revised Results section (***Section 3.3, Lines 290–345***) now presents the main spatial features more concisely, while preserving the key information needed to interpret large-scale patterns.

Major comment 5

Why is FRFS introduced in the Introduction section but QRF in the Method?

Response

We thank the reviewer for this comment regarding the structure and consistency of the

methodological description. We agree that the initial presentation may have given the impression that FRFS and QRF were introduced at different conceptual levels.

In response, we have revised the Introduction to explicitly present FRFS and QRF as complementary components of a unified modeling framework. Specifically, FRFS is introduced as a feature selection strategy designed to reduce dimensionality and improve model parsimony and interpretability, while QRF is described as the core predictive algorithm used to model soil BD and pH and to quantify prediction uncertainty.

This dual-focused strategy is now clearly outlined in the Introduction ***(Lines 46–60)***. The Methods section then follows this conceptual structure, first detailing the covariate selection procedure based on FRFS ***(Section 2.2.1)***, and subsequently describing the implementation of the QRF model as the primary predictive tool ***(Section 2.2.2)***. This revision ensures consistency between the Introduction and Methods and clarifies the distinct but integrated roles of FRFS and QRF within the overall DSM framework.

Major comment 6

Table 1 may be presented more clearly as a figure, and currently has a confusing caption.

Response

We thank the reviewer for this helpful suggestion regarding the presentation of Table 1. In response, we have revised the manuscript to improve the clarity and interpretability of this information. Specifically, Table 1 has been converted into a violin plot (now ***Fig. 3***), which more effectively illustrates the distribution of soil samples across depth intervals and highlights differences between layers.

In addition, statistical tests have been applied to assess the significance of differences among soil depths, and the corresponding results are now explicitly shown in the figure. The original table has been moved to the Supplementary Information ***(Table S5)*** for reference. The Results section has been updated accordingly to reflect the revised figure and the additional statistical information ***(Section 3.1, Lines 248–274)***. These changes improve the clarity of data presentation and provide a more informative summary of the sampling structure.

Major comment 7

Figure 6 might benefit from an overall analysis across depths, and consider adding relationships between BD and MAP (or other key covariates) in supplementary materials.

Response

We thank the reviewer for this constructive suggestion regarding the analysis of Figure 6 and the relationships between soil properties and key covariates.We agree that an overall comparison across soil depths, together with a clearer interpretation of the relationships between BD/pH and major environmental drivers, would substantially strengthen the manuscript. In response, we have extended the methodological framework by introducing SHAP (SHapley Additive exPlanations) analysis ***(Section 2.3, Lines 222–229)***.

This addition addresses a key limitation of the relative variable importance measures previously derived from the QRF models, which reflect covariate importance only in a relative sense within each depth-specific model and depend on the selected feature set.

Because FRFS yields different covariate subsets for different soil layers, these relative importance values are not directly comparable across depths. SHAP provides independent, additive contribution scores for each predictor, enabling consistent cross-depth comparison and allowing both the magnitude and direction of covariate effects to be quantitatively interpreted. Based on this approach, Figures 6 and 7 have been revised and are now presented as ***Figures 7 and 8***, illustrating depth-consistent importance patterns and the relationships between BD/pH and key covariates.

Accordingly, the previous Results subsection on variable importance has been fully revised and replaced by ***Section 3.5 (Lines 377–453)***, which now presents the SHAP-based analysis and interpretation in place of the original QRF relative importance results.

Major comment 8

L85 & 91, QRF should be explained upon its first mention.

Response

We thank the reviewer for pointing out this issue. In response, we have revised the manuscript to ensure that Quantile Regression Forest (QRF) is fully explained at its first mention ***(Lines 51)***. In addition, we conducted a systematic check of abbreviations throughout the manuscript to ensure consistent definition and usage upon first appearance.

Major comment 9

Abbreviations (including BD, SD and the abbreviations of models) in Tables and Figures should be clearly defined in their captions to make them self-explanatory.

Response

We thank the reviewer for this helpful comment. In response, we have revised all figure and table captions to ensure that abbreviations (including BD, SD, and model abbreviations) are clearly defined upon first appearance, making the tables and figures self-explanatory. This change has been applied consistently throughout the manuscript.

Major comment 10

L111 is redundant with L108.

Response

We thank the reviewer for carefully identifying this redundancy. In response, we have revised the manuscript to improve clarity and conciseness by removing the sentence.

Major comment 11

L251, rephrase “conversely”.

Response

We thank the reviewer for this helpful suggestion regarding wording. In response, we have revised the relevant sentence and removed the use of “conversely,” which was no longer appropriate given the revised structure and interpretation of the Results section. Following the substantial

revision of ***Section 3.2***, the description of pH model performance has been rewritten to avoid subjective or contrastive wording and to present the results in a more neutral and consistent manner based on the reported performance metrics.