Response to RC1:

The estimation of deep soil carbon using soil β is a well-established methodology. While the authors' global spatial prediction of soil β demonstrates scientific merit, the following methodological and analytical issues require explicit clarification and resolution prior to publication.

Response: Thank you for your constructive feedback. We appreciate your recognition of the scientific merit of our study. We have addressed the methodological and analytical issues you raised and have provided the necessary explanations and clarifications in the revised manuscript. Below are our detailed responses to each of your comments.

1. The abstract is too lengthy and should be condensed.

Response: Thank you for your comment. We have revised the abstract to make it more concise while retaining the key points of our study.

2. In line 23, the claim of 17,984 soil profiles appears inconsistent with the Rawdata.xlsx file containing only 6,817 observational records. This discrepancy requires explicit clarification and public disclosure of the methodological framework governing profile identification and data aggregation protocols. Without clarification of this discrepancy and explicit documentation of the data sources and methodological transparency regarding profile identification criteria, I cannot recommend this manuscript for publication.

Response: Thank you for your valuable feedback. The discrepancy between the 17,984 soil profiles and the 6,817 observational records in the Rawdata.xlsx file arises because the Rawdata.xlsx file contains original data extracted from the literature, while the remaining profiles come from the publicly accessible WoSIS Soil Profile Database (https://www.isric.org/explore/wosis). We have already clarified this in the manuscript and will further emphasize the profile identification and data aggregation methodology in the revised version to ensure methodological transparency.

3. Lines 96-97: The manuscript states that 17,984 soil profiles were sourced from 14,535 sites. This raises a critical methodological question: Can multiple soil profiles coexist at a single geographic site? If such spatial clustering of profiles exists, it is imperative to clarify the criteria for profile differentiation (e.g., vertical/horizontal

sampling intervals, land-use stratification, or temporal replication) and ensure these distinctions are systematically annotated in the publicly available Excel dataset.

Response: Thank you very much for your insightful comment. After careful verification and review, we confirm that multiple soil profiles can indeed coexist at a single geographic site. The criteria for profile differentiation primarily stem from differences in sampling based on fertilizer treatments, crop cultivation systems, tree age, species, sampling time, vertical sampling intervals, and soil profiles ID. We have made sure to clearly explain and organize this information in the publicly available dataset. These distinctions have been systematically annotated to ensure clarity. We hope this addresses your concern.

4. Line 104: Until which month in 2022 was the literature search conducted?

Response: Thank you for your question. After careful verification, our literature search was conducted until January 2023.

5. Lines 106-107: Please search and compare the results for: 1. "Soil organic carbon" AND "subsoil" AND "Soil profile"; 2. "Soil organic carbon" AND "Deep soil" AND "Soil profile"; 3. "Soil organic carbon" AND "Soil profile".

Response: Thank you for your valuable suggestions and comments on our paper. In response to the retrieval strategy and results you mentioned, we conducted a detailed analysis and comparison, as outlined below:

Search terms and result comparison:

Term 1: Using the keywords "Soil organic carbon" AND "subsoil" AND "Soil profile", we retrieved 818 relevant results.

Term 2: Using the keywords "Soil organic carbon" "Deep soil" and "Soil profile", we retrieved 2,038 relevant results.

Term 3: Using only the keywords "Soil organic carbon" and "Soil profile," we retrieved 13,972 relevant results.

From the above results, it can be seen that adding "subsoil" or "Deep soil" as search terms significantly reduced the number of retrieved results. However, due to our strict literature selection criteria, which only include studies with more than three measurements of organic carbon in the first meter of the soil profile, the number of relevant articles meeting these criteria is relatively low and similar to the amount we have currently obtained. Additionally, after careful verification, we identified a language expression error in the original search term. The correct term should be: "Soil organic carbon" AND "Soil profile" OR "Subsoil" OR "Deep soil". Thank you again for your careful observation and feedback.

6. Line 210: What is the rationale for the standard deviation being presented as 10% of the mean? Line 210:

Response: Thank you for your question. In Monte Carlo simulations, to reduce computational load and improve simulation efficiency, we simplify the distribution of input parameters. The choice to set the standard deviation as 10% of the mean is based on common assumptions to reflect the uncertainty of the soil β value estimation and the input parameters in the Random Forest (RF) model (Liu et al., 2024； Xu et al., 2023; Vande et al., 2004).

Liu, Y., Zhuang, M., Liang, X., Lam, S. K., Chen, D., Malik, A., Li, M., Lenzen, M., Zhang, L., Zhang, R., Zhang, L., and Hao, Y.: Localized nitrogen management strategies can halve fertilizer use in Chinese staple crop production, Nat. Food, https://doi.org/10.1038/s43016-024-01057-z, 2024.

Xu, Y., Xu, X., Li, J., Guo, X., Gong, H., Ouyang, Z., Zhang, L., and Mathijs, E.: Excessive synthetic fertilizers elevate greenhouse gas emissions of smallholder-scale staple grain production in China, *J. Cleaner Prod.*, https://doi.org/10.1016/j.jclepro.2024.128671, 2023.

Vanden B., A. J., Gregorich, E. G., Angers, D. A., & Stoklas, U. F. Uncertainty analysis of soil organic carbon stock change in Canadian cropland from 1991 to 2001. Glob. Chang. Biol.., 10(7), 983-994. https://doi.org/10.1111/j.1365-2486.2004.00780, 2004.

7. Figure 1 requires modification. It is recommended to relocate Figure S1 to the main text, combine it with the existing Figure 1 as a panel figure, and select several representative soil profiles to graphically demonstrate the variations in soil β across different ecosystem types.

Response: Thank you for the great suggestion. We have moved Figure S1 to the main text and merged it with the existing Figure 1 into a panel. We analyzed the variations in soil β values under different soil textures. Soil β values exhibited significant differences among sandy soil, loam, clay loam, and clay soil. Cropland and grassland ecosystems exhibited the highest β values in sandy soil, while forest ecosystems showed the highest β values in clay soil.

8. The Random Forest modeling data, including response variables and predictor variables, should be made publicly accessible. The critical code is also recommended to be publicly available.

Response: We have uploaded the dataset and key code to a publicly accessible repository are in the process of making the dataset and code available. The data and code can be accessed at https://doi.org/10.5281/zenodo.15019078.