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.

1) The abstract is too lengthy and should be condensed.
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.
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.
4) Line 104: Until which month in 2022 was the literature search conducted?
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".
6) Line 210: What is the rationale for the standard deviation being presented as 10% of the mean?
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.
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.