

Response to the Reviewer

Dear Reviewer,

We sincerely thank the reviewer for emphasizing the importance of data transparency, reproducibility, and reusability. We fully agree that open data and executable code are fundamental to the scientific value of a data-oriented study, particularly for ESSD. In response to this comment, we have made substantial efforts to improve the accessibility and reproducibility of our study.

First, regarding code availability, we have reorganized and released the complete and executable codebase used in this study. The codebase has been archived in Zenodo and assigned a DOI: <https://doi.org/10.5281/zenodo.20422721>. The archived repository includes scripts for data preprocessing, feature selection, model training, and spatial prediction. We have also provided documentation describing the workflow, required input data structure, software environment, script sequence, and expected outputs. Because the original soil profile observations cannot be redistributed, the repository is designed to reproduce the full analytical workflow using user-provided input data with the same structure as described in the metadata. This enables methodological reuse while respecting the legal and institutional restrictions on the original sampling dataset.

Second, regarding the underlying soil profile measurements, we fully understand the reviewer's concern. However, we would like to clarify that the original soil profile data are not owned by the individual authors. These data were collected under a national science and technology basic resources investigation project, and the authorization for public release belongs to the Ministry of Science and Technology of China. After we received the Editor's and reviewers' comments, we treated this issue as a matter of high priority and made substantial efforts to seek permission for data release. We promptly consulted the relevant project management and data authority, explained the data-sharing requirements of ESSD, and explored several possible options, including the public release of the full raw dataset, a partial subset of the sampling records, and an anonymized validation subset. However, after these formal consultations, we were informed that the original soil profile measurements and associated sampling records remain subject to project-level data management regulations, and permission to publicly release the raw sampling data, either in full or in part, could not be granted. Therefore, we are not legally or institutionally authorized to make the raw sampling records, or any subset of them, publicly available.

Nevertheless, to mitigate this concern to the extent possible, we can furnish an anonymized 20% subset of the independent validation data as confidential review material for the Editor's and Reviewers' assessment. This subset excludes exact XY coordinates, sampling-site names, and any other detailed locational information that could identify the original sampling sites. It includes newly assigned sample IDs, soil depth intervals, observed values, predicted values, and validation-related information, thereby allowing the reported validation results to be verified within the permitted scope.

We have also updated the Code Availability section of the manuscript with the new DOI. Regarding data availability, the derived gridded products are openly accessible, while the raw soil profile data remain restricted as explained above. In addition, an anonymized 20% validation subset has been prepared as confidential review material and will be provided solely for editorial and reviewer assessment. We fully acknowledge that the restricted availability of the original sampling records precludes direct reproduction of the calibration dataset. Nonetheless, this

restriction arises from data ownership and project-level data management regulations rather than from any unwillingness on the part of the authors to share the data. We have therefore exerted every possible effort to enhance transparency, reproducibility, and methodological reusability within the bounds of what is permitted. We hope that these revisions and the accompanying materials adequately address the Reviewer's concerns.

Sincerely,
Jizhen Chen
on behalf of all authors

Response to the Editor

Dear Editor,

We sincerely appreciate your thoughtful compromise and constructive suggestion that a randomly selected 20% subset of the raw data be made publicly available alongside the complete and executable code. We fully understand and share the reviewer's concern regarding data accessibility and reproducibility, and we have made every effort to comply with this request to the greatest extent possible.

However, we would like to respectfully clarify that the original soil profile measurements, including the exact geographic coordinates and raw sampling records, are not owned by the individual authors. These data were collected under a national science and technology basic resources investigation project, and the legal authority to approve any form of public release rests with the Ministry of Science and Technology of China. In response to the reviewer's comments and your recommendation, we formally consulted the relevant data authority and applied for permission to release the raw sampling data. Regrettably, this permission was not granted. As a result, we are not legally or institutionally authorized to make any subset of the original sampling records publicly accessible. We sincerely regret that we are unable to fulfill this particular expectation due to these data governance restrictions.

Nevertheless, in order to address your recommendation as far as our constraints allow, we are able to provide an anonymized 20% subset of the independent validation data as confidential review material, to be used exclusively for editorial and reviewer assessment. This subset does not include exact XY coordinates, sampling-site names, or other detailed location information that could identify the original sampling sites. The anonymized validation subset includes newly assigned sample IDs, soil depth intervals, observed values, predicted values, and validation-related information, thereby allowing the reported validation results to be checked within the permitted scope.

In addition, we have fully addressed the suggestion regarding code availability. The complete and executable codebase has been released through GitHub and archived in Zenodo with a DOI:

<https://doi.org/10.5281/zenodo.20422721>

The archived repository includes scripts for data preprocessing, feature selection, model training, validation, and spatial prediction. Because the original soil profile observations cannot be redistributed, the repository is structured to reproduce the entire analytical workflow using user-provided input data that follow the same format as described in the metadata. This design enables methodological reuse while respecting the legal and institutional restrictions on the original sampling dataset.

We hope that the combination of the openly available codebase and the confidential anonymized validation subset will be considered a workable compromise, and that this will reassure the reviewer of the transparency and reliability of our work. We are grateful for your guidance and patience, and we kindly ask for your understanding given the data governance restrictions beyond our control.

Sincerely,
Jizhen Chen
on behalf of all authors