Science Bulletin 70 (2025) 2544–2546

Contents lists available at ScienceDirect

Science Bulletin

journal homepage: www.elsevier.com/locate/scib

News & Views

Digging deeper towards accurate mapping of deep soil organic carbon in grasslands

Guocheng Wang [a,*], Zhongkui Luo [b,*], Changqing Song [a], Haojun Zheng [a], Mingming Wang [b], Shuai Zhang [b], Huirong Li [c], Haigang Li [d], Liujun Xiao [e]

[a] State Key Laboratory of Earth Surface Processes and Disaster Risk Reduction, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[b] State Key Laboratory of Soil Pollution Control and Safety, College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China
[c] Xilinhot National Climatological Observatory, Xilinhot 026000, China
[d] Inner Mongolia Key Laboratory of Soil Quality and Nutrient Resources/Key Laboratory of Agricultural Ecological Security and Green Development at Universities of Inner Mongolia Autonomous Region, Inner Mongolia Agricultural University, Hohhot 010018, China
[e] National Engineering and Technology Center for Information Agriculture, Nanjing Agricultural University, Nanjing 210095, China

Soil, as the largest terrestrial organic carbon reservoir, contains approximately 1500 petagrams of organic carbon in its top meter. This amount is roughly twice the carbon stored in the atmosphere and three times that in vegetation [1]. Enhancing soil organic carbon (SOC) storage is an effective strategy for mitigating climate change. Moreover, SOC is a key indicator of soil quality and ecological health and plays a central role in the global carbon cycle. Given increasing concerns over global warming, food security, and environmental sustainability, accurately quantifying SOC at large scales with high spatial resolution and precision is essential.

While much research has focused on surface soil, where carbon accumulation and loss are most pronounced, deeper layers also store substantial amounts of carbon [2]. Although SOC in deeper layers is generally more stable and older, it is still sensitive to global change [3]. Investigating whole-profile SOC is vital, despite the economic and technical challenges associated with deep sampling.

Over recent decades, several global and regional soil datasets have been constructed to meet different soil information requirements, while the SOC stocks (SOCs) vary substantially among different datasets [4]. The accuracy of these products, including those generated using widely applied machine learning models like SoilGrids250m, remains uncertain due to limited validation against comprehensive field measurements, particularly in deeper layers. Model accuracy may decrease with depth due to the limited relevance of surface-focused covariates for representing deeper soil carbon dynamics [5]. Therefore, rigorous evaluation of machine learning and big data-based mapping products across large spatial scales and soil profiles is needed. This study aims to evaluate the performance of commonly used global mapping models in capturing consistently sampled whole-profile SOCs. We then investigate the reasons for their varying performance across soil layers. Based on these results, we propose several research needs to improve the performance of global mapping models and reduce uncertainties in existing global-scale soil carbon datasets.

We developed a random forest (RF) model to predict SOCs using 5868 measurements from 1109 soil profiles at depths of 0–200 cm across global temperate grasslands (Fig. 1a, b). These data were sourced from the 2023 World Soil Information Service (WoSIS) database snapshot, which applies standardized quality control procedures [6]. We focused on temperate grasslands, covering 36% of Earth's land surface, due to their broad global distribution [7]. The RF model included several readily available predictors: climatic variables (mean annual temperature (MAT) and mean annual precipitation (MAP)), soil properties indicating nutrient conditions (total nitrogen (TN)), indicators of physical protection of soil carbon (clay and silt content (CS)), and soil depth intervals (top and bottom depths). All predictor variables were obtained from WoSIS. A full description of the predictors and the model training and testing procedures is provided in Text S1 (online).

The trained model explained approximately 72% of the variance in SOCs within the independent validation subset of the WoSIS dataset (Fig. 1b). Given the large sample size and the use of only four easily accessible predictors, this performance is reasonable. Typically, such a model would be used to map SOC and other soil attributes at regional and global scales [8]. However, further validation against an independent dataset obtained using consistent sampling and measurement procedures is important. To this end, we focused on the temperate grasslands of Nei Mongol, China (Fig. 1c), which account for 22% of China's total grasslands and represent one of the most continuous, diverse, and representative temperate ecosystems globally [9].

In the summer of 2020, a soil sampling campaign was conducted across 17 sites in the grasslands of Nei Mongol, China (Fig. 1c). We collected soil cores at seven depth intervals down to 200 cm and composited samples by site for analysis (Text S1 online). Laboratory procedures included cleaning, sieving, and analysis of SOC content and stocks. We found that SOCs varied considerably with depth (Fig. 1d). The topsoil (0–30 cm) had the

* Corresponding authors.
E-mail addresses: wanggc@bnu.edu.cn (G. Wang), luozk@zju.edu.cn (Z. Luo).

https://doi.org/10.1016/j.scib.2025.06.006
2095-9273/© 2025 Science China Press. Published by Elsevier B.V. and Science China Press. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

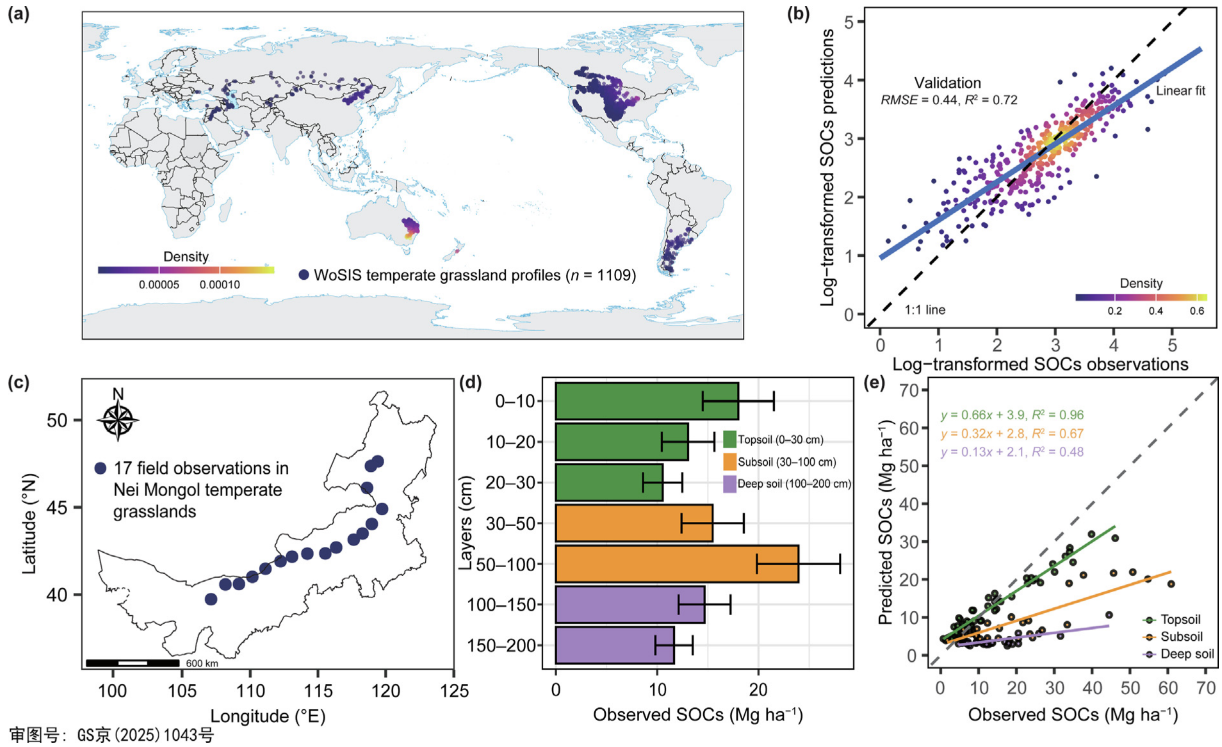

Fig. 1. Global and local SOC data, machine-learning model performance, and the comparison between predicted and measured SOC. (a) Solid circles are soil organic carbon stocks (SOCs) measurements in global temperate grasslands recorded in WoSIS dataset. (b) The performance of a random forest (RF) model developed using the global WoSIS SOCs measurements. (c) Blue solid circles are SOCs measurements in the grasslands of Nei Mongol conducted in summer of 2020, SOCs were sampled and measured in seven layers, i.e., 0–10 cm, 10–20 cm, 20–30 cm, 30–50 cm, 50–100 cm, 100–150 cm, and 150–200 cm. (d) Depth distribution of measured SOCs in the grasslands of Nei Mongol, boxes with bars show average values with standard errors across different sites. (e) Comparison of the RF-predicted SOCs and those measured in the grasslands of Nei Mongol.

highest SOC density, with average values of 18.0, 13.0, and 10.5 Mg ha^{-1} for the 0–10 cm, 10–20 cm, and 20–30 cm layers, respectively. These three layers accounted for approximately 39% of total SOC stocks within the 0–200 cm profile. The subsoil (30–50 cm and 50–100 cm) and deep soil (100–150 cm and 150–200 cm) layers held around 37% and 24% of the total SOCs, respectively. This vertical distribution is consistent with estimates reported by Jobbágy and Jackson [2].

To evaluate the generalizability of the global-scale machine learning model, we used it to predict SOCs across the seven soil depth intervals at the 17 sites of Nei Mongol, China (Fig. 1b and c). The model performed well for topsoil (0–30 cm), achieving a coefficient of determination (R^2) of 0.96. However, performance decreased with depth, with R^2 values of 0.67 and 0.48 for the subsoil (30–100 cm) and deep soil (100–200 cm), respectively (Fig. 1e). We observed a general underestimation of SOCs, particularly in the subsoil and deep soil layers. These results indicate depth-dependent model performance: the model captures SOC variability in surface soils but performs poorly at greater depths.

A partial correlation analysis revealed large differences in the relationships between SOCs and potential predictors across soil depths (Fig. S1 online). In the topsoil (0–30 cm), climate (MAT and MAP), soil nutrients (TN), microbial factors (microbial biomass carbon (MBC) and microbial biomass nitrogen (MBN)), and physical protection-related attributes (Fe, Al, CS) were dominant and interacting factors. TN exhibited strong and consistent correlations with SOCs across all depths (Fig. S1 online), reflecting carbon–nitrogen coupling and shared stabilization mechanisms. Fe and Al were also positively correlated with SOCs, likely due to their role in forming soil aggregates and physically protecting SOC.

In deeper layers, the relationships of SOC with several environmental predictors weakened or disappeared, potentially explaining the decline in model performance with depth (Fig. 1e). Climate variables had less impact, likely due to their indirect impact via plant carbon inputs, which are concentrated in the topsoil [8,10]. Similarly, microbial community effects diminished with depth (Fig. S1 online), aligning with previous studies [11]. SOC stabilization by organo-mineral complexes, involving binding to fine particles represented by silt and clay content, underpins the strong positive correlation between CS and SOCs across all depths, though this correlation weakened in deeper layers (Fig. S1 online). Applying globally trained models to local SOC prediction is challenging due to scale differences and site-specific variation. While this study expands spatial observations and addresses the limitations of global models in predicting whole-profile SOC, local-scale factors may still shape SOC patterns and dynamics. Therefore, future research should refine and validate models by incorporating detailed local data and accounting for biogeochemical processes specific to different regions and soil depths.

To improve model performance and support precise whole-profile soil mapping, several research directions are proposed. First, increasing the depth of future sampling (e.g., beyond 100 cm) would enable more accurate characterization of deep SOCs and improve model predictions. Second, expanding global soil databases (e.g., WoSIS) to include attributes related to the microbial community (MBC and MBN) and physical protection (Fe and Al) would support the development of depth-specific models depending on depth-dependent drivers (Fig. S1 online). These efforts are resource-intensive and require substantial time, funding, and logistical support. Finally, given the large disparities among existing global SOC datasets [4], we encourage producers to include more observational data and explore alternative modeling approaches to improve map accuracy. Addressing these research needs will improve our understanding of SOC dynamics and support more accurate modeling and management of this key resource.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was jointly supported by the National Natural Science Foundation of China (42375116), the Scientific Innovation Projects of Meteorological Bureau in Nei Mongol Autonomous Region (nmqxkjcx202422), and the Fundamental Research Funds for the Central Universities.

Author contributions

Guocheng Wang conceived the study and wrote the manuscript. Zhongkui Luo designed the field sampling campaign and laboratory measurements. Changqing Song and Haojun Zheng analyzed the data. Mingming Wang and Shuai Zhang performed field sampling and laboratory measurements. All authors contributed to the interpretation of the results, as well as the writing and revision of the manuscript.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scib.2025.06.006.

References

[1] Friedlingstein P, Jones MW, O'Sullivan M, et al. Global carbon budget 2021. Earth Syst Sci Data 2022;4:1917–2005.
[2] Jobbágy EG, Jackson RB. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecol Appl 2000;2:423–36.
[3] Pries CEH, Castanha C, Porras RC, et al. The whole-soil carbon flux in response to warming. Science 2017;6332:1420–2.
[4] Lin Z, Dai Y, Mishra U, et al. Global and regional soil organic carbon estimates: magnitudes and uncertainties. Pedosphere 2024;4:685–98.
[5] Liu F, Wu H, Zhao Y, et al. Mapping high resolution National Soil Information Grids of China. Sci Bull 2022;3:328–40.
[6] Batjes NH, Ribeiro E, van Oostrum A. Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019). Earth Syst Sci Data 2020;1:299–320.
[7] Sala OE. Temperate grasslands. In: Chapin FS, Sala OE, Huber-Sannwald E, editors. Global biodiversity in a changing environment: scenarios for the 21st century. New York: Springer; 2001. p. 121–37.
[8] Wang G, Xiao L, Lin Z, et al. Most root-derived carbon inputs do not contribute to long-term global soil carbon storage. Sci China: Earth Sci 2023;5:1072–86.
[9] Zheng H, Huang Y, Zhang W, et al. The implementation of ecological protection in Inner Mongolia has slowed down grassland degradation. Fundam Res 2024. http://doi.org/10.1016/j.fmre.2024.10.006.
[10] Chen L, Fang K, Wei B, et al. Soil carbon persistence governed by plant input and mineral protection at regional and global scales. Ecol Lett 2021;5:1018–28.
[11] Li J, Ding J, Yang S, et al. Depth-dependent driver of global soil carbon turnover times. Soil Biol Biochem 2023;185:109149.