

<https://doi.org/10.1038/s43247-025-02605-6>

# Spatio-temporal mapping reveals changes in soil organic carbon stocks across the contiguous United States since 1955



A list of authors and their affiliations appears at the end of the paper

Understanding spatio-temporal patterns of soil organic carbon (SOC) is critical for global climate change mitigation and sustainable soil management. However, information on long term dynamics of SOC over large area is lacking. Supported by soil samples collected over years and environmental covariates, space and time digital soil mapping (ST-DSM) has become an important and effective method to reveal the spatio-temporal changes of SOC. The contiguous United States (CONUS) has abundant and well-documented soil samples with time labels, which lays the groundwork for us to estimate the long-term SOC dynamics in multiple soil layers over that region with high resolution. Specifically, we propose leveraging time-series soil data from World Soil Information Service (WoSIS) and International Soil Carbon Network (ISCN) to build ST-DSM models at different soil depths based on matching environmental covariates and machine learning techniques (random forest framework). Then, multi-depth ST-DSM models are employed to generate spatial prediction of SOC in different layers (0–15 cm, 15–30 cm, 30–60 cm and 60–100 cm) from 1955 to 2014 at 250 m resolution and 5-year intervals (1955–1959, 1960–1964, ..., 2010–2014). Meanwhile, predictive uncertainties are quantified via Quantile Regression Forest (QRF). Furthermore, we analyze the dynamic trends in SOC stocks across various depths and land uses. The results indicate that over 60 years, overall SOC stocks in 0–100 cm demonstrate a multi-stage change of “rising-fluctuating”. SOC stocks in 0–100 cm rose from 68.40 Pg (1 Pg =  $10^{15}$  g) to 70.33 Pg with an increase rate of 32.25 Tg (1 Tg =  $10^{12}$  g) per year. SOC stocks in 0–15 cm layer declined slightly before 1970 and increased thereafter; the 15–100 cm layer fluctuated with an overall rising trend; notably, the 60–100 cm layer exhibited a steady upward trend with minimal fluctuations since 1980s, suggesting a greater SOC accumulation potential in soil beneath 15 cm. Across different land uses, the topsoil of cropland may be a source of carbon, while forests are important carbon sinks. SOC stocks in surface soil (0–15 cm) in cropland areas notably decreased, and subsurface and deep soil in cropland (15–30 cm, 30–60 cm and 60–100 cm) showed a fluctuating rising trend. SOC stocks in forest show an overall fluctuating rising trend. Pastureland, unmanaged natural grasslands and sparse/no vegetation areas once showed some carbon sequestration capacity in the past, but they have transitioned to slower accumulation. This research enhances our understanding of soil carbon dynamics at a national scale and provides references for the development of effective soil management and climate mitigation strategies.

Soil organic carbon (SOC) plays a pivotal role in providing essential ecosystem services, including climate regulation, plant growth, and water retention<sup>1,2</sup>. The upper 1 meter of soil is estimated to hold 1550 Pg (1 Pg =  $10^{15}$  g) of SOC stocks at the global scale<sup>3</sup>, representing about two-

thirds of the world's terrestrial carbon stocks<sup>4,5</sup>. Even a small change in SOC can have a notable impact on atmospheric CO<sub>2</sub> concentrations. In the context of global climate change, the fate of SOC is of special concern and its role in natural climate solutions is being increasingly recognized<sup>6,7</sup>. SOC

✉ e-mail: [yanglin@nju.edu.cn](mailto:yanglin@nju.edu.cn)

dynamics depend on multiple natural and anthropogenic factors that influence carbon input and output. Due to the spatial and temporal variation of influencing factors, the distribution of SOC over regions is highly heterogeneous and can change with time, posing challenges to accurately predicting future climate change and developing management strategies for carbon sequestration<sup>8–11</sup>. Understanding the spatio-temporal variation of SOC at regional or national scales is therefore essential for global climate change mitigation and sustainable soil management<sup>12</sup>.

Attempts have been made to map the distribution of SOC on a large scale. Some publicly available global and national soil databases, such as SoilGrids<sup>13</sup>, the Harmonized World Soil Database (HWSD), the State Soil Geographic (STATSGO) database, the Soil Survey Geographic (SSURGO) database, and the Unified North American Soil Map (UNASM), provide maps of soil properties at global and national scales produced with conventional or digital soil mapping methods<sup>14</sup>. However, these maps were all created for one single time point and cannot support time series analysis. Information on long-term SOC dynamics at large scale is still lacking.

Insights on spatio-temporal changes of SOC can be gained through revisited temporal sampling<sup>15</sup>. However, revisited soil samples are sparse because soil sampling is costly and there is no systematic soil monitoring network<sup>2,15</sup>. This lack of revisited temporal sampling data poses a notable obstacle to quantifying spatio-temporal change of SOC<sup>16,17</sup>. Process-based modeling is another way to predict the temporal change of soil properties from site-level to regional scale<sup>18,19</sup>, as they simulate the underlying physicochemical processes of the soil carbon cycle<sup>20,21</sup>. These models typically establish a static baseline SOC map and then estimate SOC changes by incorporating factors such as climate, land use, management practices, and carbon inputs<sup>18,19</sup>. However, achieving high-resolution prediction of SOC through process-based modeling remains challenging. For one thing, spatially explicit input data required for process-based models, such as detailed information on agricultural management practices over large areas, are often difficult to obtain<sup>22,23</sup>. For another, parameterizing models for large-scale applications is difficult, as it is unrealistic to assume that parameters are spatially invariant in the case of large-scale applications<sup>18</sup>.

Digital soil mapping (DSM) serves as an efficient and cost-effective method to predict the spatio-temporal pattern of SOC. The basic process of DSM is to construct the relationship between the target soil property and predictors/environmental covariates using soil sample data, and then make prediction based on environmental data<sup>24–26</sup>. Space and time DSM (ST-DSM) has recently been developed to model changes in soil properties<sup>2,18,27,28</sup>. With data from soil samples collected at certain time and matching environmental covariates, prediction can be made for each relevant time point<sup>16</sup>. ST-DSM can effectively quantify and capture the spatio-temporal dynamics of SOC stocks and provide estimated prediction uncertainties<sup>18,29</sup>. The reliability of results generated using ST-DSM relies on good spatial and temporal representativeness of samples<sup>30</sup>. As a result, ST-DSM has only been empirically tested in a limited number of regions, and its feasibility at larger scales over long term remains unvalidated. Furthermore, most ST-DSM studies have primarily concentrated on surface soils and single depth<sup>21,31</sup>. Because of the lack of sufficient sample data support or reasonable data integration across different databases, comprehensive research with large-scale, long-time-series, high-resolution and multi-depth data is often rare.

Analyses of the global carbon cycle indicate that the SOC of North America plays a notable role in the global carbon budget, acting as a potential biospheric sink for atmospheric CO<sub>2</sub><sup>32,33</sup>. The contiguous United States (CONUS) is an important study area for investigating SOC dynamics due to its extensive diversity in climatic conditions, land cover types, and land use practices. Thanks to systematic soil surveys, soil sample data from the U.S. is relatively abundant and well-documented<sup>34–36</sup>. Some studies have estimated SOC stocks using the average SOC values of each land resource region based on the STATSGO and SSURGO databases, such as Guo et al.<sup>36</sup>, Franzmeier et al.<sup>34</sup> and Sims and Nielsen<sup>35</sup>. However, this approach does not fully account for the heterogeneity of soil and environmental variables within each stratum<sup>29,37</sup>. Process-based modeling has also been used to

estimate SOC stocks in the CONUS<sup>32,38</sup>, but these studies may lack robust precision validation using widely distributed measured SOC samples.

DSM-based approaches have also been adopted for estimating SOC stocks in the U.S. For instance, Guevara et al.<sup>39</sup> used a simulated annealing regression framework to predict the spatial distribution of surface SOC density of 0–30 cm soil (with 10,385 samples) across CONUS from 1991 to 2010 at a 250 m spatial resolution, finding that SOC stocks were ~30% lower than estimates from SoilGrids and HWSD. Wang et al.<sup>40</sup> applied unsupervised multivariate geographic clustering (MGC) to divide CONUS into 20 SOC regions based on environmental similarity, and employed supervised random forest (RF) regression on more than 20,000 soil samples at depths of 0–30 cm and 30–100 cm to model each region separately. Gonçalves et al.<sup>41</sup> employed geographically weighted regression (GWR) to estimate SOC stocks at 0–100 cm based on 4559 soil samples across 21 ecoregions in CONUS and projected SOC stocks changes from 2030 to 2100. Despite these advancements, few studies have fully leveraged historical soil sampling data from CONUS, with well-defined sampling times, to capture the spatial and temporal dynamics of SOC stocks over several decades. Multi-depth, long-time-series, high-resolution dynamic SOC data for the CONUS region remain limited.

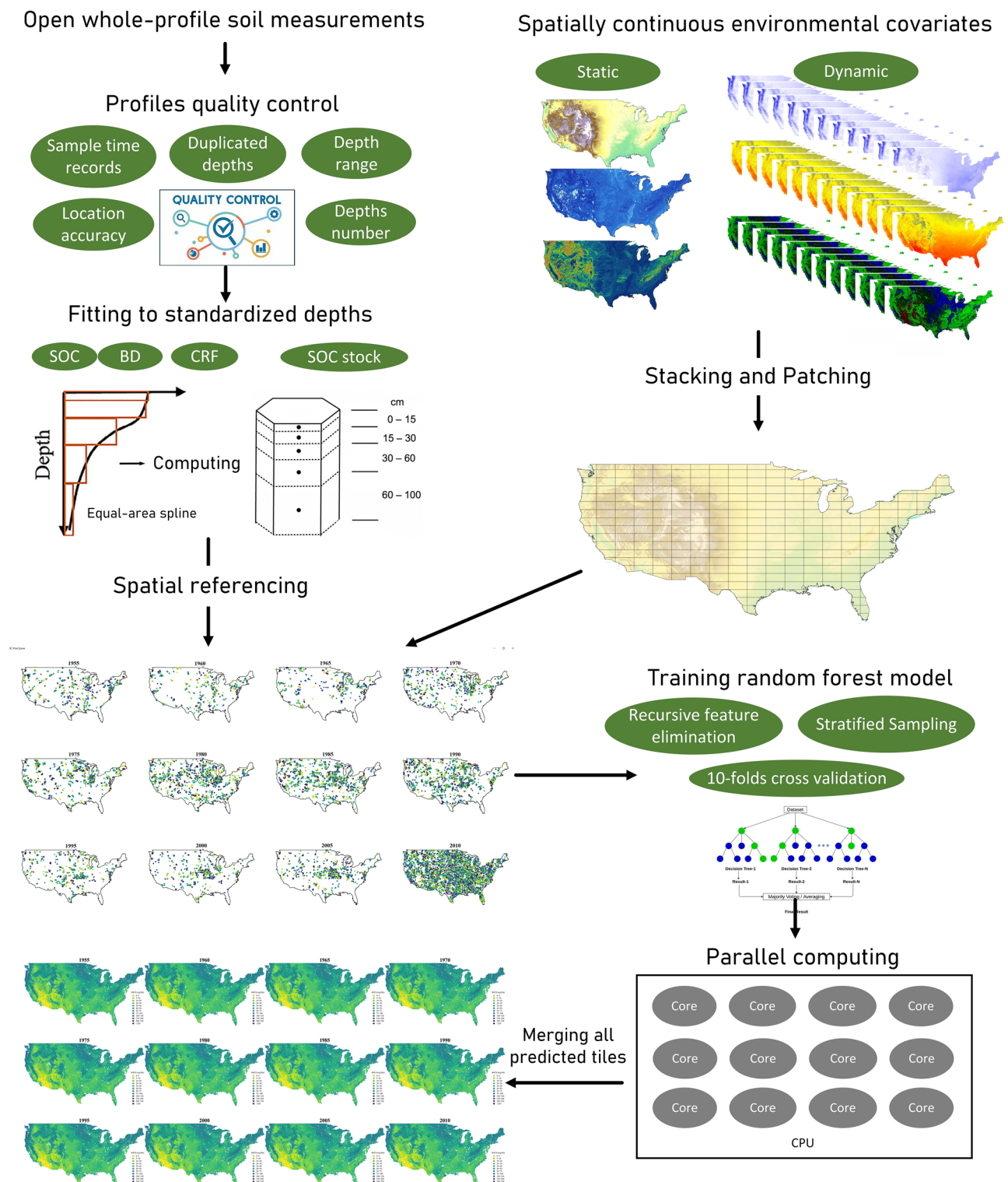
To address this, we propose leveraging time-series soil sample point data from World Soil Information Service (WoSIS) and International Soil Carbon Network (ISCN) to build ST-DSM model based on environmental covariates and machine learning techniques, specifically using a RF framework. Our goal is to bridge the gap in long-term SOC density distribution over the CONUS from 1955 to 2014. The generated data include multi-depth (0–15, 15–30, 30–60 and 60–100 cm) SOC density maps for CONUS, at 250 m spatial resolution and 5-year temporal intervals (1955–1959, 1960–1964, ..., 2010–2014), with predictive uncertainty quantified via Quantile Regression Forest (QRF). We analyze dynamic trends in SOC stocks across various depths and land uses. Methodologically, we demonstrate the feasibility of ST-DSM for large-scale, long-time-series, high-resolution, and multi-depth DSM applications. This research offers a detailed view of the spatial and temporal patterns of SOC stocks across CONUS over more than half a century, serving as valuable foundation for understanding global carbon cycle and climate mitigation strategy-making. Figure 1 shows the methodological framework of this study.

## Results

### Model training, accuracy assessment and environmental covariate importance

The hyperparameters and the selected covariates for models of different depths are presented in Table 1. Except for the surface layer (0–15 cm), the RFE selection process removed relatively few covariates from the complete set of 59. Most covariates were retained, indicating that while fewer covariates were initially used in this study compared to SoilGrids<sup>13</sup>, the final number of covariates post-RFE selection was comparable to those used in SoilGrids. The model reached stable accuracy with ~15 covariates; any additional covariates contributed to only marginal improvements in prediction accuracy. In this study, the number of trees (ntree) ranged from 100 to 200, and the mtry was set to 1, which aligns with the default settings commonly used in RF models.

The cross-validation results for each depth throughout the whole period (60 years) are summarized in Fig. 2 (displaying LCCC, R<sup>2</sup>, and RMSE). Additionally, the cross-validation results and scatter plots for each year at different depths can be found in Supplementary Fig. 1. The LCCC values were 0.63, 0.60, 0.54, and 0.51 for the depth ranges of 0–15 cm, 15–30 cm, 30–60 cm and 60–100 cm, respectively. Correspondingly, the R<sup>2</sup> values were 0.48, 0.46, 0.41, and 0.38, respectively. These accuracy metrics are comparable to those of other large scale DSM products, such as SoilGrids<sup>13,42</sup>. Overall, a decreasing trend in accuracy with increasing depth was observed, which aligns with the findings of most DSM studies<sup>42,43</sup>. This decline may be attributed to the diminishing relationship between environmental variables and the properties of deeper soil layers.



**Fig. 1** | Methodological framework of this study.

Variable importance analysis from the RF modeling highlights soil-related covariates (i.e., CEC, clay) as the most influential predictors, especially for deeper soils (Supplementary Fig. 2). Among the dynamic climate variables, precipitation-related covariates (i.e., `cm_prec_mean` and `cm_prec_mean_season1-4`) consistently rank highest in importance across all soil depths, underscoring the key role of precipitation in driving the temporal dynamics of SOC. Moreover, many topographic covariates (i.e., elevation, texture, topographic diversity) also play an important role.

Notably, for the 60–100 cm depth layer, ecoregions Level 4 (eco L4) was one of the most important covariates after soil-related covariates.

### The predicted SOC density maps

Figure 3 shows the mapping results of SOC density across different soil depths at the first period (1955–1959) and last period (2010–2014). The mapping results of other periods are shown in Supplementary Fig. 3. The 0–15 cm layer exhibits the highest SOC density, while the 60–100 cm layer

has the lowest. Different soil layers show distinct patterns of spatial and temporal variation. In the surface layer, the highest SOC densities appear in northern and western forested regions, while the arid and semi-arid areas in the southwestern United States, with sparse vegetation, exhibit lower SOC densities. As for deeper depths, besides the southwestern areas, southeastern forested areas also have low SOC density, creating a distinct west-high, east-low spatial distribution. Changes in spatial distribution patterns of SOC density are small over every five years, and similar patterns are observed across each soil layer.

### The SOC dynamics in CONUS from 1955 to 2014

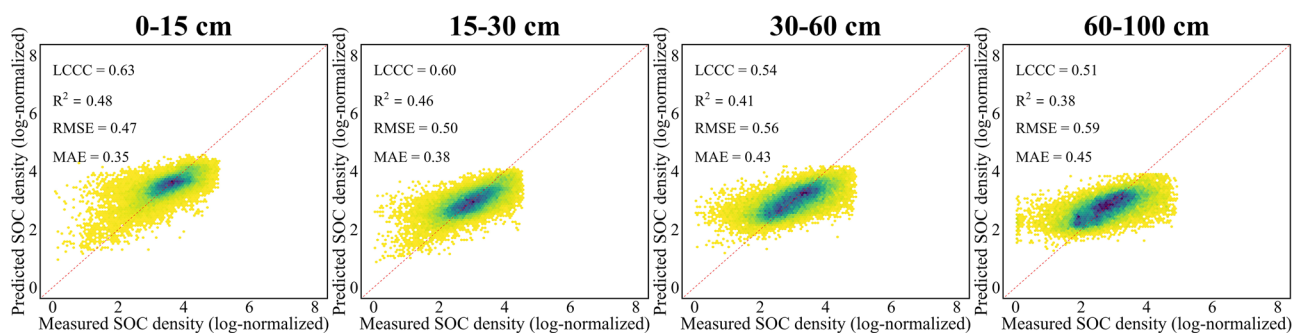
The results indicate that the relative change rate (%) of SOC density comparing 2010–2014 with 1955–1959 varies notably with soil depth and displays distinct spatial heterogeneity (Fig. 4, Table 2). In surface (0–15 cm) and subsurface layers (15–30 cm), SOC shows a “declined-then-rose”

pattern over sixty years with net increases in the central-western, northern and southeastern regions when comparing the last period (2010–2014) to the first period (1955–1959). Conversely, large areas in the southwestern and parts of the central-north and central-east regions exhibit SOC losses exceeding 10% compared to the first period, predominantly following a “rose-then-declined” trend over sixty years. The spatial distribution of SOC changes in deeper layers (30–60 and 60–100 cm) show some differences, with more pronounced and extensive increasing trends in the north-central and southeastern regions, mostly following a “declined-then-rose” or “fluctuating change” pattern, though losses still occurred in the southern-most parts of the central regions. These findings highlight the depth-dependent SOC dynamics across regions. These findings highlight the notable depth dependency and spatial heterogeneity of SOC dynamics.

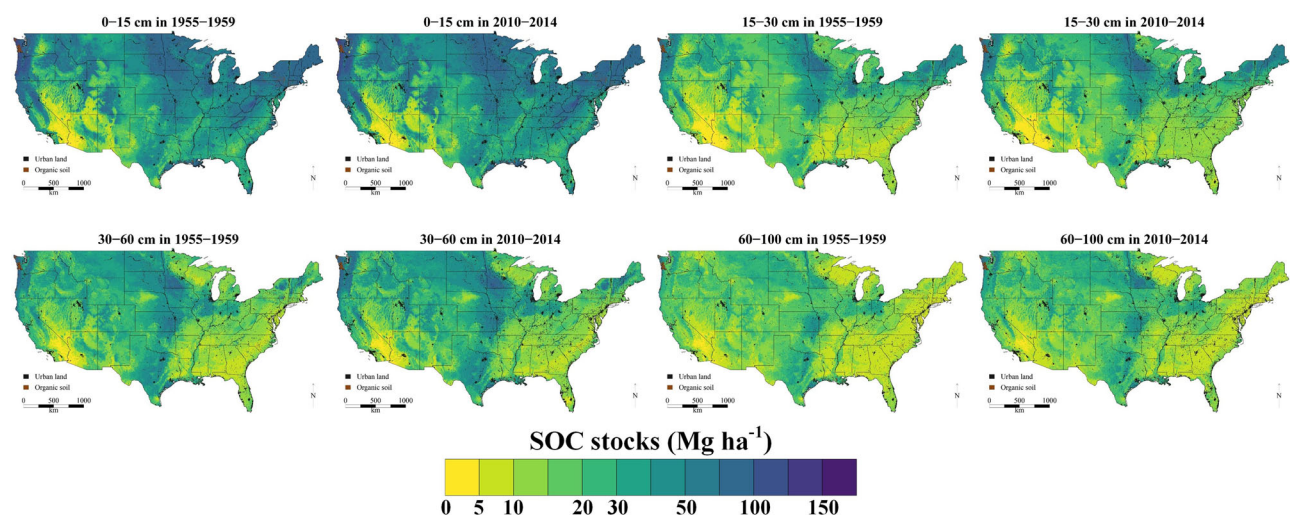
The overall SOC stocks above 1 m demonstrate a multi-stage change over 60 years (Fig. 5), with a gradual increase before 1980 followed by fluctuating changes thereafter. In terms of change between 1955–1959 and 2010–2014, total SOC stocks above 1 m rose from 68.40 Pg to 70.33 Pg, with an overall average annual change rate of 32.25 Tg per year (1 Tg =  $10^{12}$  g, Fig. 5). Temporal trends in SOC stocks also varied in different layers. The surface 0–15 cm SOC stocks declined slightly before 1970 and exhibited a fluctuating increasing trend thereafter. The deeper three soil layers showed similar obvious rising trends until 1980, but declined since 1985, and fluctuating increasing afterwards. Notably, the 60–100 cm exhibited the minimal temporal fluctuations after 1980 and the strongest rising trend over six decades, highlighting its importance as a long-term SOC reservoir potential. The overall average annual change rate for 0–15 cm, 15–30 cm, 30–60 cm

**Table 1 | Number of covariates and hyper-parameters for each depth model**

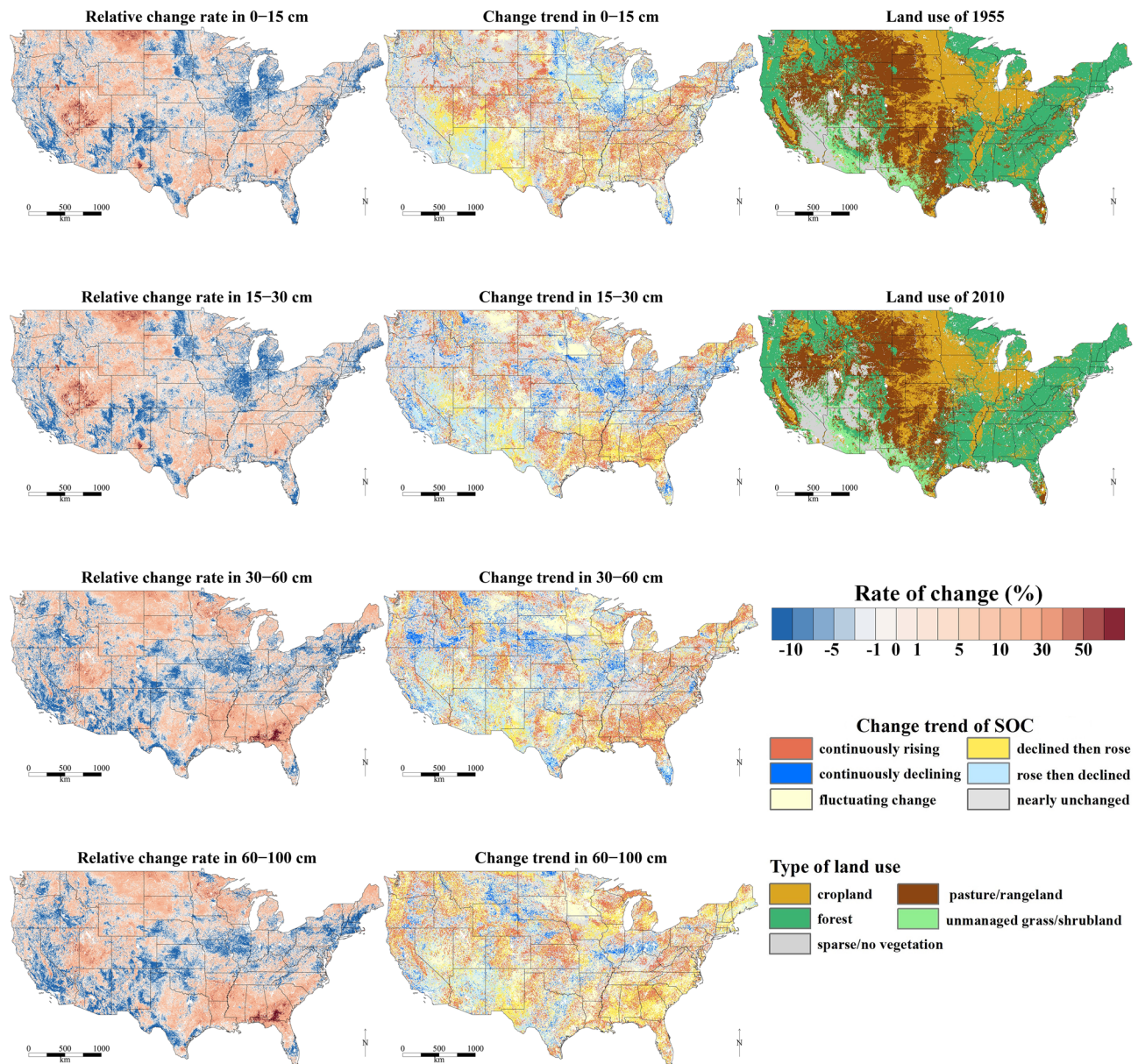
Depth	Number of covariates	ntree	mtry
0–15 cm	23	200	4
15–30 cm	47	150	6
30–60 cm	54	200	7
60–100 cm	49	200	7



**Fig. 2 | The accuracy and scatterplot of measured soil organic carbon (SOC) density (normalized) and predicted SOC density (normalized) for each depth throughout the whole period.** The red line represents the 1:1 line. The scatterplot is a density scatterplot, where colors range from light yellow to dark green, with darker colors indicating higher point density.



**Fig. 3 | The spatial distribution of soil organic carbon (SOC) density for depths 0–15 cm, 15–30 cm, 30–60 cm and 60–100 cm in 1955–1959 and 2010–2014.** Subtitles indicate the soil depth and time period represented by each panel. Colors range from light yellow to dark green, with darker colors indicating higher SOC density.



**Fig. 4 | Spatial distribution of soil organic carbon (SOC) density in 1955–1959 (first period) and 2010–2014 (last period), relative change rate between 1955–1959 and 2010–2014, and change trend in 0–15 cm, 15–30 cm, 30–60 cm and 60–100 cm.** Subtitles indicate the different meanings of each panel. In the relative change rate maps, colors range from dark blue to dark red, representing a

shift from SOC decrease to increase. In the change trend maps, different colors indicate different types of SOC change trends. The relative change rate between 1955–1959 and 2010–2014 at different depths in different periods was shown in Supplementary Fig. 4.

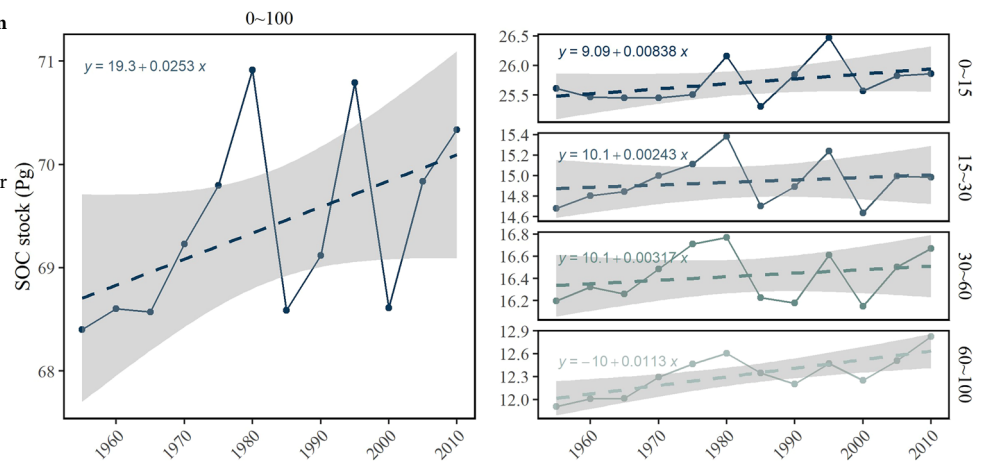
**Table 2 | Percentage of various types of changes at different depths for 60 years**

Depth	Nearly unchanged	Continuously rising	Continuously declining	Declined then rose	Rose then declined	Fluctuating change
0–15 cm	35.51%	16.90%	4.08%	14.50%	8.79%	20.22%
15–30 cm	36.69%	15.12%	7.39%	12.03%	8.57%	20.20%
30–60 cm	33.90%	15.98%	6.96%	12.09%	8.49%	22.58%
60–100 cm	25.91%	18.28%	4.24%	18.01%	5.70%	27.86%

and 60–100 cm are 4.04 Tg, 5.06 Tg, 7.88 Tg and 15.28 Tg per year, respectively. Deeper soil has a higher rate of increase than surface soil. These trends suggest that SOC accumulation is not only restricted to surface layers but also occurs notably in deeper layers, highlighting the importance of including deeper soil carbon when assessing long-term carbon storage potential.

We also evaluated changes in SOC stocks for each land use type, distinguishing between stable (i.e., unchanged over six decades) and changed land use categories, from the first period (1955–1959) to the last period (2010–2014) (Table 3, Fig. 6). Land use change is obtained by comparing the 1955 and 2010 land use maps. The SOC stocks changes mainly occurred in areas where land use remained

**Fig. 5 | The temporal trends of soil organic carbon (SOC) stocks in the Contiguous United States (CONUS) across different soil depths from 1955 to 2014.** Each panel represents a specific depth interval: 0–15 cm, 15–30 cm, 30–60 cm, and 60–100 cm. The gray area around each trend line represents the 95% confidence interval of the linear regression model.



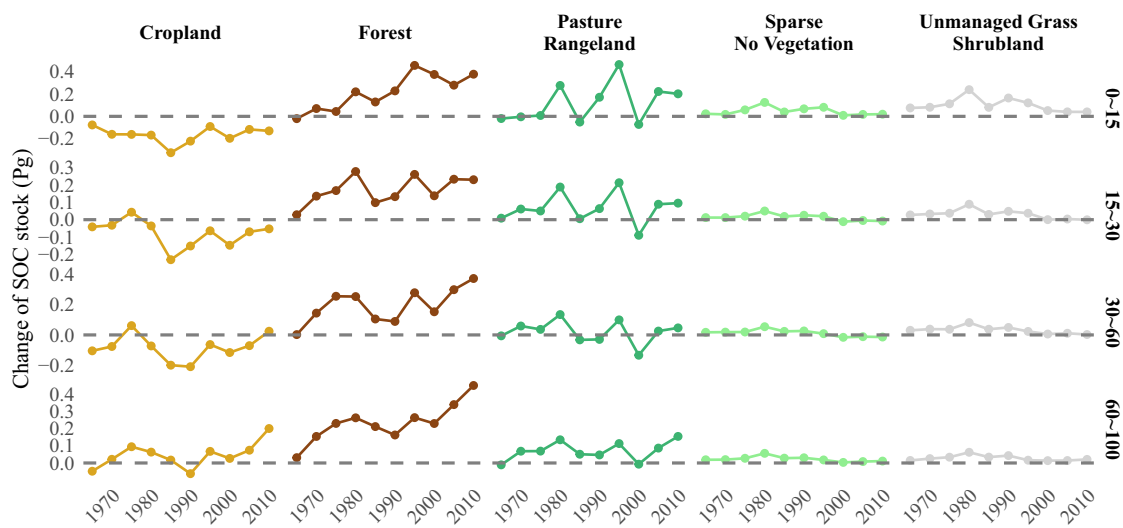
**Table 3 | Soil organic carbon (SOC) stocks change between the first period (1955–1959) and the last period (2010–2014) in each land use transition type**

	0–15 cm			15–30 cm			30–60 cm			60–100 cm			0–100 cm		
Area (10 <sup>4</sup> km <sup>2</sup> )	1955	2010	Δ (Tg)	1955	2010	Δ (Tg)	1955	2010	Δ (Tg)	1955	2010	Δ (Tg)	1955	2010	Δ (Tg)
Stable F <sup>a</sup>	235.79	9.92	10.17	242.29	4.89	5.12	238.47	4.39	4.78	392.41	2.95	3.42	461.29	22.15	23.49
F→C	1.77	0.07	0.07	0.21	0.03	0.03	1.97	0.03	0.03	2.40	0.02	0.02	3.01	0.15	0.15
F→P	2.56	0.10	0.11	3.62	0.05	0.06	2.97	0.05	0.06	4.51	0.04	0.04	5.47	0.24	0.26
F→G	2.85	0.12	0.12	3.14	0.06	0.07	3.74	0.06	0.07	5.66	0.05	0.05	5.89	0.29	0.31
F→N	0.44	0.02	0.02	0.66	0.01	0.01	0.92	0.01	0.02	2.34	0.01	0.01	2.00	0.06	0.06
Stable P	178.73	5.29	5.55	259.79	3.32	3.43	101.15	4.40	4.46	55.28	3.39	3.54	153.31	16.41	16.98
P→C	15.27	0.49	0.51	21.52	0.31	0.32	7.00	0.40	0.40	5.11	0.30	0.32	13.73	1.49	1.54
P→F	25.05	0.91	0.95	40.94	0.48	0.51	23.54	0.53	0.56	29.75	0.41	0.44	39.66	2.33	2.47
P→G	6.29	0.20	0.21	7.27	0.12	0.12	3.00	0.15	0.16	1.31	0.12	0.12	5.86	0.60	0.61
P→N	7.01	0.18	0.19	11.40	0.11	0.12	6.04	0.14	0.15	8.52	0.11	0.12	11.99	0.54	0.58
Stable C	145.71	5.71	5.65	−63.68	3.61	3.72	105.30	3.83	4.01	172.75	2.76	3.05	296.32	15.92	16.43
C→P	5.69	0.20	0.21	5.09	0.13	0.13	3.84	0.15	0.16	2.76	0.11	0.12	7.08	0.59	0.61
C→F	10.94	0.39	0.39	−1.47	0.21	0.21	8.59	0.19	0.20	11.68	0.15	0.17	18.63	0.93	0.96
C→G	3.28	0.12	0.12	2.54	0.07	0.08	2.43	0.08	0.08	1.73	0.06	0.07	4.83	0.34	0.35
C→N	0.21	0.01	0.01	0.03	0.00	0.00	0.04	0.00	0.00	−0.01	0.00	0.00	0.23	0.02	0.02
Stable G	43.89	1.03	1.06	29.76	0.68	0.69	6.39	0.90	0.91	6.79	0.69	0.71	25.42	3.29	3.36
G→C	0.94	0.02	0.03	1.61	0.02	0.02	0.54	0.02	0.02	0.60	0.02	0.02	0.96	0.07	0.08
G→P	5.83	0.17	0.17	7.38	0.11	0.11	2.94	0.14	0.14	1.11	0.11	0.12	6.75	0.52	0.54
G→F	4.26	0.16	0.17	6.40	0.09	0.09	4.10	0.10	0.11	4.02	0.08	0.08	6.02	0.43	0.45
G→N	3.33	0.04	0.04	1.78	0.03	0.03	0.15	0.05	0.05	−0.80	0.04	0.04	1.69	0.16	0.16
Stable N	44.9	0.44	0.47	26.12	0.34	0.34	3.67	0.50	0.50	3.94	0.44	0.47	26.15	1.72	1.78
N→C	0.06	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.08	0.00	0.00	0.03	0.00	0.00
N→P	0.41	0.01	0.01	0.27	0.00	0.00	0.02	0.01	0.01	0.24	0.01	0.01	0.31	0.02	0.02
N→F	0.02	0.00	0.00	0.03	0.00	0.00	0.02	0.00	0.00	0.05	0.00	0.00	0.04	0.00	0.00
N→G	2.81	0.03	0.03	1.16	0.02	0.02	−1.25	0.03	0.03	0.69	0.03	0.03	1.10	0.11	0.11

<sup>a</sup>Forest (F), pasture/rangeland (P), cropland (C), unmanaged grass/shrubland (G), sparse/no vegetation (N)

unchanged over time due to their extensive coverage. The largest SOC increase (1.334 Pg) within the 0–100 cm depth was observed in stable forest, followed by pasture (0.570 Pg), cropland (0.510 Pg), unmanaged grass/shrubland (0.068 Pg) and sparse/no vegetation area (0.060 Pg). Carbon sinks occurred in almost all depths of these stable land uses, especially in 60–100 cm. In contrast, stable cropland displayed SOC losses in the top 0–15 cm layer, possibly due to intensive agricultural practices.

Further analysis of relative change dynamics reveals complex patterns that vary by both land use types and soil depth (Fig. 6). In forests, SOC stocks decreased in the 0–15 cm layer until around the 1990s and then kept increasing. In the deeper layers, SOC stocks exhibited a fluctuating upward trend, with greater increases observed at deeper depths, suggesting that forest ecosystems may be gradually enhancing carbon storage in subsoil layers over time. Pasture showed minor increases in the 60–100 cm layer, while the upper three layers displayed continuous fluctuations with no clear



**Fig. 6 | Relative change of soil organic carbon (SOC) stocks (Pg) in each 5-year interval.** The relative change was calculated based on comparing to mean SOC stocks of different land use types in 1955–1959.

pattern until the 2000s, then increased slightly. This indicates that pastures have a limited capacity for SOC accumulation in all layers under historic management practices, but situations may have changed recently. SOC stocks in the upper three layers of croplands all went through continuous losses until the 1980s, followed by a gradual recovery. For the layer of 60–100 cm, we found no loss before the 1980s and similar increases as in the upper layers. In contrast, in sparse/no vegetation areas, unmanaged grassland and shrubland, SOC stocks increased consistently until the 1980s, after which their relative change rate started to decline, suggesting that these areas acted as carbon sinks initially, but carbon accumulation has slowed down in recent decades. These findings highlight the unique SOC dynamics across land use types and underscore the importance of land management strategies tailored for specific land uses and soil depths to optimize SOC storage across the landscape.

Land use changes caused minor SOC stocks changes, but some of them contributed to large changes in SOC stocks (Supplementary Table 1), revealing meaningful patterns in SOC dynamics for land use management. Afforestation resulted in greater SOC gains than those caused by converting forest to other land use types. In particular, the conversion of pasture into forest led to the largest increase (0.134 Pg). The shift from sparse/no vegetation area (N) to forest was accompanied by a larger increase than that happening in stable forests. These highlight the contribution of afforestation to carbon sequestration over the past 60 years. In addition, the two-way transformations between pasture and cropland (P→C and C→P), transforming cropland to unmanaged grass/shrubland (C→G) and transforming pasture to sparse/no vegetation area (P→N) resulted in SOC gains, mainly in 60–100 cm, suggesting that such land use changes may enhance carbon storage in deeper soil layers. Conversion from sparse/no vegetation area to any land use types (N→others), from cropland or unmanaged grass/shrubland to sparse/no vegetation (C→N, G→N) and from unmanaged grass/shrubland to cropland (G→C) showed almost no SOC changes, indicating that such transitions need to be optimized for greater carbon benefits.

### The uncertainty of SOC stocks predictions

Figure 7 shows the PICP validation for each depth throughout the whole period, and Fig. 8 shows the RWPIR mapping results of SOC stocks in 2010–2014 for each soil layer as an example. The PICP values closely align with the 1:1 line at all depths, indicating strong predictive performance of the QRF model regarding uncertainty. However, the PICP validation results for the periods 1965–1969 and 1970–1974, 2000–2004, as shown in Supplementary Fig. 8, are slightly over the 1:1 line across all depths, suggesting

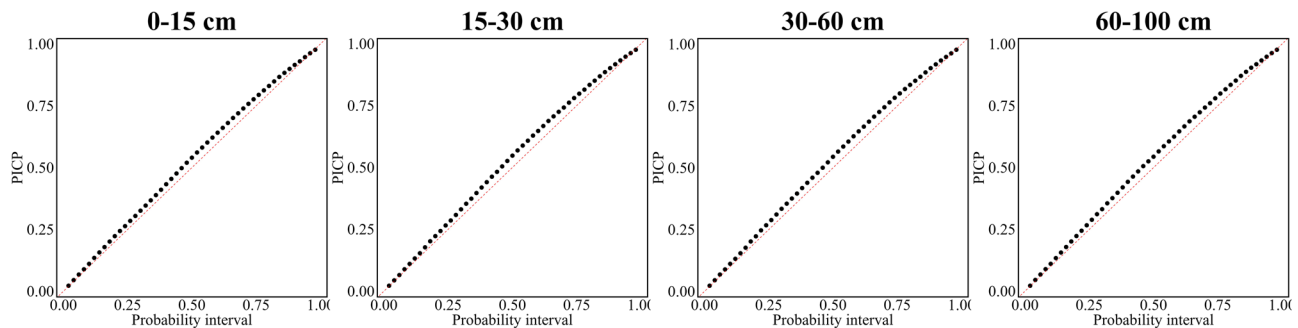
that the uncertainty estimates for these periods may be overly optimistic. The spatial distribution of RWPIR reveals that higher uncertainty is concentrated in the Western Sierra, Peninsular Florida and Great Lakes regions, likely due to the sparse distribution of soil sample sites and poor spatial representativeness in these areas (see Fig. 9). Furthermore, there is a tendency for uncertainty to increase with depth, which may be attributed to a reduced number of soil samples in deeper layers. Notably, the spatial distribution of RWPIR appears relatively consistent across different periods (Supplementary Fig. 6).

## Discussions

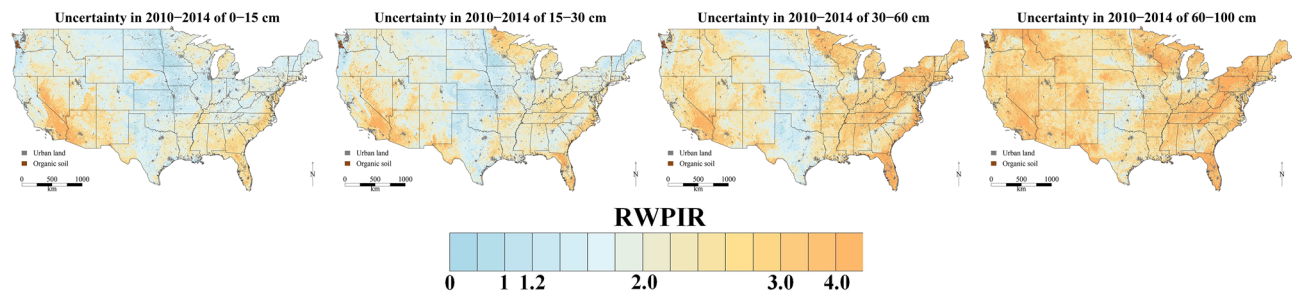
### Depth-dependent SOC changes across different periods

As one of the few sources of long-term evidence on SOC across large regions, our study reveals distinct phases of SOC change over decades using ST-DSM models. We identified shifts in SOC stocks changes in the contiguous United States at depths of 0–100 cm around 1980 and 1995. Similarly, a 32-yr period of observational study at the Walker Branch watershed in Tennessee, United States, reported fluctuations in soil carbon concentrations (%) from 1972 to 2004, with a decline in SOC observed between 1982 and 1993 that was later reversed in subsequent years<sup>44</sup>. Our findings suggest that this pattern of change is not isolated, as similar shifts have occurred across broader regions in the CONUS, further demonstrating that our dataset can serve as a foundational resource for soil carbon management across the CONUS. Additionally, this data provides a robust basis for further investigation into the drivers of carbon change, which are likely linked to climate change and anthropogenic impacts<sup>45,46</sup>. This underscores the significance and potential of our research in contributing to the development of effective soil management and climate mitigation strategies.

Furthermore, through whole-profile soil sample analysis, we discovered that these transition phases vary across soil depths. Our results emphasize the critical importance of examining long-term SOC dynamics at multiple depths, including the surface (0–15 cm), subsurface (15–30 cm & 30–60 cm), and deep layers (60–100 cm). Notably, unlike the relatively stable upper layers, SOC stocks in the 60–100 cm layer exhibited a consistent and notable increase in all land-use types over 60 years, contributing to ~45% of the total SOC gain in the contiguous United States. This finding suggests that SOC in deeper soil layers exhibits greater dynamics than conventionally recognized, likely governed by factors extending beyond mere chemical and physical stabilization mechanisms<sup>47,48</sup>. The reported trends of deep SOC dynamics remain contradictory across previous studies. For example, Wang et al.<sup>49</sup> and Zhou et al.<sup>50</sup> indicated that long-term cultivation can lead to notable losses of deep SOC, while Mayer et al.<sup>51</sup> and Liu



**Fig. 7 | The prediction interval coverage probability (PICP) validation for each depth throughout the whole period. The red line represents the 1:1 line.**



**Fig. 8 | Relative Width of the Posterior Interquartile Range (RWPIR) of soil organic carbon (SOC) density in 2010–2014 (as an example) for each depth. Colors range from light blue to dark yellow, with darker colors indicating higher RWPIR and greater uncertainty.**

et al.<sup>52</sup> observed an increasing trend in deep SOC at global or regional scale, respectively. Future studies should systematically examine the governing mechanisms under varying climatic regimes and anthropogenic disturbance intensities. Our findings highlight the importance of incorporating depth-specific analyses into SOC research and modeling frameworks to better capture long-term carbon dynamics<sup>53,54</sup>.

### SOC stocks change in different land uses

Vertical differentiation was observed in cropland SOC dynamics, while surface layers (0–15 cm) showed decelerated depletion after 1985, all sub-surface horizons (15–100 cm) maintained variable but persistent accumulation patterns. Changes in precipitation patterns are likely to have influenced cropland SOC dynamics<sup>55</sup>. Increasing precipitation after 1985 may have played a role in supporting biomass growth and organic matter input, thus possibly promoting SOC accumulation<sup>55,56</sup>. In addition to climate factors, human activities have also notably affected SOC trends in cropland. The observed decline in surface SOC stocks may be attributed to soil aeration and structural disruption caused by tillage practices, which accelerate the oxidation and loss of organic matter<sup>57,58</sup>. After 1985, the implementation of no-tillage subsidies and associated acidification processes may have contributed to SOC stabilization or recovery<sup>59,60</sup>.

Forest SOC stocks across all depths show an overall increasing trend, though with noticeable fluctuations. Compared to croplands, the increase in forest SOC stocks in the CONUS region is likely to be more strongly influenced by natural factors, such as precipitation and warming<sup>61</sup>. Precipitation in CONUS forest exhibited overall fluctuations, with an increasing trend prior to 1995, followed by a general decline. Meanwhile, temperature rose notably from 1975 to 2000 and remained relatively stable in the following years. Under the combined influence of precipitation and temperature, forest SOC stocks exhibited an overall increasing trend with some fluctuations. As the role of forest soil carbon sinks has gained increasing attention in recent years<sup>62,63</sup>, future management strategies may need to more carefully account for the potential impacts of climatic factors<sup>64</sup>.

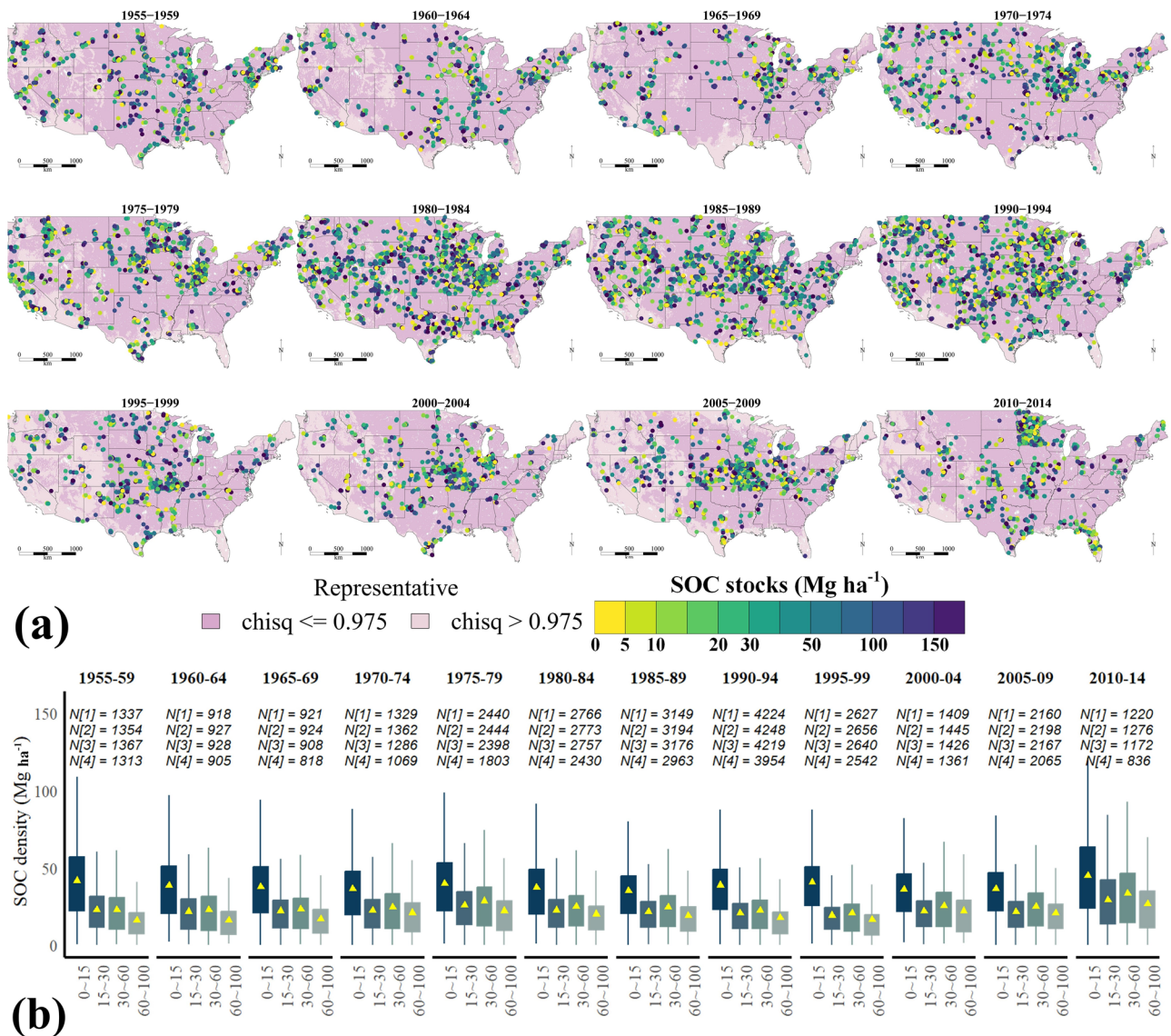
Sparse/no vegetation area and unmanaged natural grasslands have the potential to store SOC through revegetation and organic matter

enhancement<sup>9</sup>. However, it is worth noting that these two land uses are themselves small in size, so their capacity as carbon sinks may be debatable. In 1980, when the SOC stocks of these two land uses were at their highest, they stored SOC stocks close to 110% of the initial value. But these two land uses have transitioned to slower accumulation in recent decades, this may be a result of sparse vegetation cover leading to a lower upper limit of storable SOC for this land use type. Increasing vegetation cover and converting to other land use types with higher vegetation cover may be an effective way to further develop the carbon sink capacity of sparse/no vegetation and unmanaged natural grasslands<sup>65</sup>.

### Limitations and perspectives

Obtaining directly revisited temporal sampling can improve the accuracy of spatio-temporal change measurements<sup>15,66</sup>, however, this sampling of soil samples remains challenging. Compared with the revisiting temporal sampling, ST-DSM provides a practical and dependable method for modeling SOC stocks and their variability. The strong spatial representativeness of our sample distribution in different periods (~80% of sites having a chi-square value < 0.975) provides the basis for implementing a reliable ST-DSM. However, sites in 1995–1999, 2000–2004, and 2005–2009 have a chi-square value < 0.975 ~60%, which may have introduced predictive bias and higher uncertainty. This limitation is likely unavoidable when using existing open soil databases and highlights the need for more comprehensive soil databases in the future. Furthermore, given that many early sampled U.S. soil profiles did not record the O horizon even in forest, we excluded the O horizon of all profiles in the modeling to avoid estimation bias. This may lead to an underestimation of SOC stocks in forest. Future efforts to harmonize O horizon data records standardization in soil sampling could help to improve the accuracy of SOC stock assessments, particularly in organic-rich environments. In addition, incorporating the O horizon into future modeling efforts would enhance our understanding of total SOC stocks in forest ecosystems, where a notable portion of organic carbon is often stored in surface organic layers.

Vegetation-related organism covariates, such as vegetation indices derived from remote sensing, are critical in ST-DSM, as leaf litter, root litter and root exudates are essential sources of SOC. Vegetation contributes to



**Fig. 9 | Distribution of observations of soil organic carbon (SOC) density.** Spatial distribution and representativeness of soil samples in different periods at depths of 0–15 cm (a). The median, mean (yellow triangle), and interquartile range of SOC

density for samples along the vertical soil profile, note that numbers of samples from different depths are shown (i.e., N[1] represents the number of the first layer) (b).

elevated SOC stocks due to its high productivity and continuous carbon supply<sup>67,68</sup>. In this study, because of the insufficient temporal coverage of remote sensing data, we did not use remote sensing derived covariates and used potential natural vegetation and CONUS ecoregions instead (Supplementary Table 2). This choice may have led to an underestimation of the vegetation impact on SOC spatio-temporal changes.

Additionally, ST-DSM, being primarily data-driven, often overlooks the biogeochemical processes underlying SOC changes<sup>2,18</sup>. Lack of guidance on the process may introduce uncertainty in prediction results. To address this limitation, future studies could combine process-based models with ST-DSM to gain a more comprehensive understanding of SOC spatio-temporal changes and mitigate lag effects in SOC responses to environmental variables. Studies by Xie et al.<sup>21</sup> and Zhang et al.<sup>22</sup> demonstrate that integrating process-based models with ST-DSM notably improves the spatio-temporal modeling of SOC dynamics.

## Conclusion

In this study, we produced a spatially explicit SOC dataset for the CONUS spanning 1955–2014 at 5-year intervals. Using the ST-DSM approach with

59 environmental covariates, we generated predictions at 250 m spatial resolution for four depth intervals: 0–15, 15–30, 30–60, and 60–100 cm, accompanied by quantified uncertainty estimates. This work addresses a critical gap in national-scale SOC stocks data, providing both spatial predictions and temporal trends over several decades. Our data exhibit high spatio-temporal accuracy with LCCC ranging from 0.51 to 0.63 and  $R^2$  ranging from 0.38 to 0.48, underscoring both the reliability of the dataset and the feasibility of applying ST-DSM on a national scale.

Our findings indicate that over 60 years, overall SOC stocks above 1 m demonstrate a multi-stage change of “rising-fluctuating”. Total SOC stocks above 1 m rose from 68.40 Pg to 70.33 Pg with an increase rate of 32.25 Tg per year. The 0–15 cm layer declined slightly before 1970 and increased thereafter; 15–100 cm layers fluctuated with an overall rising trend; notably, the 60–100 cm layer exhibited a steady upward trend with minimal fluctuations since 1980s, suggesting a greater SOC accumulation potential with soil depth below 15 cm. Across different land uses, the topsoil of cropland may be a source of carbon, while forests are important carbon sinks. SOC stocks in surface soil (0–15 cm) in cropland areas notably decreased and subsurface and deep soil cropland SOC stocks

(15–30 cm, 30–60 cm and 60–100 cm) showed a fluctuating rising trend. Forest SOC stocks show a fluctuating rising trend overall. Pastureland and sparse/no vegetation areas and unmanaged natural grasslands once showed some carbon sequestration capacity before 1980s, but they have transitioned to slower accumulation. This research enhances our understanding of the global carbon cycle and supports the development of climate mitigation strategies.

## Materials

**Study area.** The study area for this research is the CONUS, encompassing 48 adjoining states. CONUS is one of the largest and most ecologically diverse regions in the world, covering a wide array of climatic zones from arid deserts in the Southwest to humid subtropical areas in the Southeast, as well as temperate climates in the Northeast and Midwest. Its varied topography includes vast plains, rolling hills, major mountain ranges such as the Rockies and Appalachians, and extensive river systems like the Mississippi and Missouri Rivers. This diversity in climate and topography supports a rich array of ecosystems and land uses, providing ideal settings to examine SOC dynamics across a range of environmental gradients.

**Soil observation data and quality assessment.** We collected the soil organic carbon sample data in CONUS from two soil profile datasets, WoSIS snapshot 2019 (ISRIC Data Hub) and International Soil Carbon Network version 3 Database (ISCN3). The World Soil Information Service (WoSIS)<sup>69</sup> collates the largest quality-assessed and standardized database of explicit soil profile observations across the globe by ISRIC-World Soil Information (<https://www.isric.org/>). The ISCN is an international scientific community devoted to the advancement of soil carbon research and manages an open-access soil carbon database<sup>70</sup>. We selected high-quality profiles of CONUS from the snapshot of WoSIS2019 and ISCN3 according to the following two criteria: 1) high accuracy of geographical coordinates, 2) explicit sampling year. Specifically, in the WoSIS database, we selected the sampling points with geographical coordinates error of less than 1 arc-second (~30 m). In the ISCN database, we selected geographical coordinates with a position error of less than 0.0001 arc-degree (~11 m).

To ensure the quality of the profile for the following part of digital soil mapping, we first excluded profiles with fewer than 3 layers to avoid potential biases in depth standardization caused by insufficient vertical resolution. Then, we excluded layers of those profile if their deeper depth is smaller than that of the upper depth, excluded the layers with overlapping depths within one profile (e.g. layers 0–30 and 0–15 in the same profile), and averaging the values of soil properties of layers with same upper and deeper depth. Subsequently, quality markers were assigned to each sample based on a standardized process, according to the number of soil layers in the profile, whether the top layer was an O horizon or an unnamed layer, the presence of organic soil with all O layers, and the effective recorded thickness of the profile. Finally, based on the quality makers from the original records, we excluded O horizons (the first soil layer was named O) or potential O horizons (the first layer was unnamed and with high organic content exceeding 20%) due to their inconsistent temporal distribution across records and the fact that most of the profiles didn't record this layer. In the WoSIS database, fewer than 1% of sampling sites recorded O horizons or potential O horizons, while in the ISCN database, the proportions were below 2% (before 2000) and about 4% after 2000. In addition, we removed those organic soils (defined as soils with all depths are O horizons) from our dataset due to their scarcity, accounting for less than 0.04% of the total records.

**SOC density estimation.** To calculate SOC density using soil data from WoSIS, we extracted the observed soil organic carbon content (SOCc, g/kg), bulk density (BD, g/cm<sup>3</sup>), coarse fragments (CRF %), sand (%), silt (%) and clay (%) of each soil layer. Due to the inconsistency of layer/depth division of each profile, we harmonized each soil property into four depth layers:

0–15, 15–30, 30–60 and 60–100 cm. They were taken as the standardized sample data for the following soil predictions. Specifically, equal-area quadratic splines are used to fit a continuous depth function of SOC density, BD, sand, silt and clay based on the properties measured by genetic horizons following<sup>71</sup>. As for the CRF, of which variation along a profile often includes reasonably abrupt changes, we used a numerical integration based on a rectangle rule (Hengl et al., 2017).

Among these soil profiles, only half have measurements for both BD and CRF. Following the approach of Wang et al.<sup>47</sup> and Chen et al.<sup>72</sup>, we used a machine learning-based pedo-transfer function to perform imputation for missing data. Specifically, for BD and CRF in each standard soil depth, we developed random forest models based on all measurements of the respective property (e.g., BD) using other observed soil properties, including SOCc, clay, sand and silt and sample time from the WoSIS database as the main variables for the pedo-transfer function of BD. For soil profiles that clay, sand and silt were not reported, we used the SoilGrids with 250 m spatial resolution<sup>13</sup> as a substitute at given sampling locations. After scaling and filling the above properties, the SOC density (t/ha) was calculated according to the following formulation:

$$SOC\ density(kgCm^{-2}) = SOCc(g\ kg^{-1}) \cdot BD(g\ cm^{-3}) \cdot (1 - CRF) \cdot d(m) \quad (1)$$

The ISCN3, released in December 2015, provides soil profiles with depths containing computed SOC density. There are two ways to calculate SOC density, i.e., total carbon\*bulk density, organic carbon\*bulk density. To keep consistent with the density calculation method using the WoSIS database in our study, firstly, for the soil layers calculating SOC density using total carbon (containing organic carbon and inorganic carbon), we corrected them using %CaCO<sub>3</sub> data for samples reporting both C<sub>tot</sub> and CaCO<sub>3</sub>. Specifically, we quantify soil inorganic carbon (SIC) as 12% of the % CaCO<sub>3</sub> following the convention<sup>73</sup>. Secondly, we recalculated all the SOC density by multiplying the coarse fragments (CRF %). Due to the lack of reported CRF and incomplete soil texture information (sand, silt and clay) as well, we extracted CRF from the SoilGrids dataset. Finally, we also harmonized the SOC density to four depths by summing the SOC density of depths within the depth interval, when the end point of one layer is larger than the bottom value of standard depth (e.g. 15, 30, 60, 100), we got the density proportionally assuming that the density are evenly distributed in depth.

The histograms of the two databases before the 2000s show strong alignment (Supplementary Fig. 7), with some profiles (about 73% of all profiles used from ISCN) sharing the same locations. This consistency suggests that combining the two databases has minimal impact on analyzing SOC dynamics. To further refine our data, we excluded sampling points from ISCN that shared coordinates with WoSIS, retaining unique profiles from ISCN as a complement. Ultimately, we compiled SOC density data across four depth intervals: 0–15 cm, 15–30 cm, 30–60 cm, and 60–100 cm, with sample sizes of 24,500, 24,801, 24,444, and 22,059, respectively. Given the left-skewed distribution of SOC density and the common tendency of DSM models to underestimate high values, we applied a log-normal transformation ( $\log(x + 1)$ ) during model training and subsequently back-transformed the predictions to the original scale for interpretation.

**Representativeness estimation of soil samples.** To examine the representativeness of the collected soil samples, referenced to Patoine et al.<sup>74</sup>, we calculated the environmental representativeness of the collected sample data for each five years using seven basic environmental variables: cation exchange capacity, clay, CONUS ecoregions Level IV, elevation, annual mean maximum/minimum temperature and total precipitation in each five years and parent material. Analysis of sample site representativeness showed good environmental coverage for all periods except 1995–1999, 2000–2004, and 2005–2009, with ~85% of sites having a chi-square value < 0.975 (Fig. 9a). The spatial distribution and representativeness of soil samples across periods for depths of

15–30 cm, 30–60 cm and 60–100 cm are illustrated in Supplementary Fig. 8. The histogram of SOC density in different land use types for four depths in different periods is shown in Supplementary Fig. 9.

**Environmental covariates.** We used 59 covariates as predictors that were representative of the environmental factors: s (soil), c (climate), o (organisms), r (relief), p (parent material), a (age), n (spatial location) and human<sup>24</sup>. Supplementary Table 1 provides the sources, original resolution and detailed descriptions of the environmental covariates. The covariates considered can be grouped as follows.

Soil covariates include soil type derived respectively from Gridded National Soil Survey Geographic Database (gNATSGO) and Harmonized World Soil Database (HWSD), cation exchange capacity (CEC), clay, nitrogen, phh2o from the SoilGrids. Climatic covariates include annual and seasonal averaged maximum temperature, minimum temperature, average temperature and precipitation. The 4 seasons are defined as March to May, June to August, September to November, and December to February, respectively. Monthly grids of mean maximum/minimum temperature and total precipitation at a 60 arc-second resolution (~2 km) from 1901 to 2016 were downloaded from the World Data Center for Climate (WDCC)<sup>75</sup>. This dataset was generated using ANUSPLIN by producing thin plate smoothing spline models. Organism covariates include the CONUS ecoregions (Level I, II, III and IV)<sup>76</sup> and potential natural vegetation (modified Kuchler) from the US Environmental Protection Agency (EPA). Relief covariates include elevation and factors calculated based on elevation layers by SAGA (i.e., aspect, slope, Topographic Wetness Index, Convergence Index, etc.), The original Digital Elevation Model (DEM) data at 90 m spatial resolution from the Shuttle Radar Topography Mission (SRTM), and the Continuous Heat-Insolation Load Index (CHILI), Physiographic Diversity (PD), Topographic Diversity (TD) from the USGS 3DEP 10 m National Map Seamless dataset at a resolution of 1/3 arc-second (~10 m) are downloaded through the Google Earth Engine (GEE) platform. CONUS lithological maps are used to represent parent material.

Land use from HILDA+ global land use change<sup>77</sup> and Major Land Resource Area (MLRA) from the US Department of Agriculture (USDA) are used to represent human activities. The HILDA+ reconstruction was derived from multiple openly available global, contiguous, regional, and national land use and land cover change (LUC) datasets, including remote sensing data, reconstructions, and statistics. This dataset provides a long-term global annual land use record from 1899 to 2019 at a 1 km spatial resolution, covering six categories: urban areas, cropland, pasture/rangeland, forest, unmanaged grass/shrubland, and sparse/no vegetation areas. Sampling time and latitude/longitude coordinates of soil samples are used as the age and spatial location factors.

Typically, remote sensing and vegetation data derived have been widely used as key environmental covariates in DSM studies<sup>78,79</sup>. However, while we aim to model SOC dynamics from 1955 to 2014 in this study, available remote sensing data only dates back to the 1980s at the earliest. To ensure consistency in the covariates used for ST-DSM modeling over the entire period, we chose not to include remote sensing covariates. All covariates were resampled to match the target spatial resolution of 250 m using the cubist resampling method, then cropped to the study area's extent and projected to a common coordinate reference system (WGS-1984).

## Methods

### Space and time DSM (ST-DSM)

In space-time DSM modeling, both static and dynamic covariates are used together to build the SOC-environment relationships over time. Dynamic covariates typically include climate, vegetation, and land cover, while static covariates often include topography and intrinsic soil properties<sup>21,80,81</sup>. In this study, climate, land use, and sampling time are used as dynamic covariates, while soil type, topography, biotic factors, and parent material are treated as static covariates (Supplementary Table 1). For modeling, we pooled sample points from all years to train a unified model for each soil depth. Then the dynamic covariates for each period were taken as input into

the model to generate SOC density predictions correspondingly. During the prediction phase, pixels located in urban areas and water bodies were excluded. After predicting the spatial distribution of SOC density ( $\text{Mg ha}^{-1}$ ), SOC stocks (Pg or Tg) are obtained by multiplying SOC density by area<sup>82</sup>.

The Random Forest (RF) model was employed for modeling and prediction due to its robustness, high tolerance to outliers and noise, and reduced susceptibility to overfitting<sup>83</sup>. RF has been extensively applied in previous DSM studies and in many cases demonstrated superior predictive accuracy compared to other models<sup>84,85</sup>. In ST-DSM, RF is also one of the most frequently used models<sup>18,29</sup>. The model tuning process involved two main steps: covariate selection using recursive feature elimination (RFE) and hyperparameter optimization.

Selecting an optimal set of environmental covariates is essential in DSM to reduce redundancy, streamline the model, improve computational efficiency, and minimize overfitting risk<sup>86,87</sup>. We used RFE to identify the best combination of covariates. RFE has proven effective and robust in optimizing covariate selection within the RF model<sup>88,89</sup>. This algorithm initially fits the model with all covariates, assesses model performance, and ranks covariate importance. It then iteratively removes the least important covariates, refits and reevaluates the model until all covariates have been tested. Finally, it outputs the covariate combination with the highest cross-validation accuracy<sup>13</sup>. During RFE, model hyperparameters remained at default values (500 trees for ntree and the square root of the covariate count for mtry). R-squared was chosen as the loss function, and RFE was implemented using the “caret” package in R.

Following RFE-based covariate selection, the next step was to optimize the hyperparameters of the RF model: the number of trees in the forest (ntree) and the number of predictors randomly selected for each tree (mtry). Previous studies have indicated that excessively complex hyperparameters do not notably improve model accuracy and can substantially increase computational demands<sup>13</sup>. To prevent overfitting and maintain computational efficiency, we tested ntree values of 100, 150, and 200, and mtry values of 1, 1.5, 2, and 3 times the square root of the covariates count. Each ntree and mtry combination was evaluated using 10-fold nearest-neighbor distance matching (NNDM) leave-one-out (LOO) cross-validation method<sup>90</sup>, with the configuration yielding the highest R-squared across folds selected as optimal.

The 10-fold NNDM LOO cross-validation method was applied to evaluate the prediction accuracy<sup>90</sup>. In each iteration, nine folds were used as the training dataset and one as the validation dataset. The average of the cross-validation results was taken as the final accuracy metric. To minimize bias in the splitting of training and validation sets, soil samples were stratified by sampling year. The median (0.50 quantile) of the Quantile Regression Forest (QRF) was used as the primary metric for model accuracy. Four commonly used accuracy verification indices, the Lin's concordance correlation coefficient (LCCC)<sup>91,92</sup>, model efficiency coefficient ( $R^2$ ), root mean square error (RMSE) and mean absolute error (MAE) to calibrate the mapping results. The formulas of LCCC,  $R^2$  and RMSE are demonstrated as follows:

$$LCCC = \frac{2r\sigma_y\sigma_{\hat{y}}}{\sigma_y^2 + \sigma_{\hat{y}}^2 + (\bar{y} - \bar{\hat{y}})^2} \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

$$RMSE = \frac{1}{\sqrt{n}} \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (5)$$

where  $n$  is the sample size of validation points,  $y_i$  and  $\hat{y}_i$  are observed and predicted SOM value at the corresponding validation points,  $r$  is the correlation coefficient between  $y_i$  and  $\hat{y}_i$ ,  $\bar{y}$  and  $\bar{\hat{y}}$ ,  $\sigma_y$  and  $\sigma_{\hat{y}}$  are the average and variance of  $y_i$  (observed SOM) and  $\hat{y}_i$  (predicted SOM).

We used the QRF to estimate the prediction uncertainty<sup>93</sup>. Unlike the RF which keeps only the average of the observations that fall into that node and ignores all other information, the QRF keeps the value of all the observations in this node. This approach calculates the quantiles of the conditional probability distribution for each prediction location<sup>78,94</sup>. The QRF model was constructed using the ranger package in R with the “quantreg” parameter set to 0.05, 0.95. This way it will output the 0.05 quantile and the 0.95 quantile, representing the lower and upper limits of a 90% prediction interval (PI90), as described in the GlobalSoilMap specifications<sup>13,95</sup>. The relative width of the posterior inter quantile range (RWPIR) is obtained by dividing PI90 by the median (0.50 quantile) to display the spatial distribution of uncertainty<sup>78,96</sup>. The percentage of cross-validated observations included in the PI90 (prediction interval coverage probability, PICP) is calculated to assess the reliability of uncertainty estimates. The basic idea of PICP is evaluating the soil samples from the test set lies in the  $\tau \cdot 100$  percent PIs<sup>97</sup>:

$$PICP(\tau) = \frac{1}{n} \sum_i^n \delta(l_i \leq y_i \leq u_i) \cdot 100 \quad (6)$$

where  $\delta$  is an indicator function, with a Boolean argument:

$$\delta(t) = \begin{cases} 1, & \text{if } t \text{ is TRUE} \\ 0, & \text{else} \end{cases} \quad (7)$$

where  $l_i$  and  $u_i$  are the lower and upper bounding quantiles that together define a  $\tau \cdot 100$  percent PI. Multiple PICPs are usually calculated for different PI levels ( $\tau$  values), and PICP measures the reliability of the entire uncertainty prediction distribution. In cartography, PI levels were plotted on the x-axis and PICP on the y-axis to obtain reliability scatter plots, which were plotted against a 1:1 line to visually assess the reliability of the uncertainty predictions. Values being closer to the 1:1 line means that uncertainty predictions are more reliable, and being below and above the 1:1 line indicate over-pessimistic or over-optimistic PIs.

### Analysis of the spatio-temporal dynamics of SOC

Based on mapping results, we performed the analysis of spatio-temporal dynamics of SOC. The relative change rate of SOC between 1955–1959 and 2010–2014 was calculated by subtracting the start period result (i.e., 1955–1959) from final period result (i.e., 2010–2014) and dividing it by the start period result. Referring to Ray et al.<sup>98</sup>, based on the results of SOC density mapping for each period, we categorized the changing trend of SOC density into six types: continuously rising, declined then rose, continuously declining, rose then declined, fluctuating change and nearly unchanged. Specifically, we fitted intercept-only, linear, quadratic, and cubic models on each image element of the mapping results in each period, and used the Akaike Information Criterion to decide which model fitted the observed data the best. If the intercept-only model is best, it indicates that SOC density is nearly unchanged; if the linear model is best, the positive slope indicates that SOC stocks is continuously rising, negative slope indicates that SOC density is continuously declining; if the quadratic model is best, the sign of the positive quadratic term indicated that SOC density declined then rose, the negative indicated that SOC density rose then declined; if cubic model is best, it indicated that SOC density is fluctuating.

### Data availability

The predicted SOC density maps can be readily downloaded at <https://zenodo.org/records/15454655>.

### Code availability

No custom scripts or code were developed; all data processing procedures are fully documented in the Materials and Methods section.

Received: 19 January 2025; Accepted: 21 July 2025;

Published online: 01 August 2025

### References

- Wang, X. et al. Prediction of soil organic matter using VNIR spectral parameters extracted from shape characteristics. *Soil Tillage Res.* **216**, 105241 (2022).
- Ugbemuna Ugbaje, S. et al. Space-time mapping of soil organic carbon stock and its local drivers: potential for use in carbon accounting. *Geoderma* **441**, 116771 (2024).
- Köchy, M., Hiederer, R. & Freibauer, A. Global distribution of soil organic carbon – Part 1: masses and frequency distributions of SOC stocks for the tropics, permafrost regions, wetlands, and the world. *SOIL* **1**, 351–365 (2015).
- Do, M.-T. T. et al. National variability in soil organic carbon stock predictions: impact of bulk density pedotransfer functions. *Int. Soil Water Conserv. Res.* **12**, 868–884 (2024).
- Scharlemann, J. P. W., Tanner, E. V. J., Hiederer, R. & Kapos, V. Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Manage* **5**, 81–91 (2014).
- Crowther, T. W. et al. Sensitivity of global soil carbon stocks to combined nutrient enrichment. *Ecol. Lett.* **22**, 936–945 (2019).
- Gang, C. et al. Uncertainty in land use obscures global soil organic carbon stock estimates. *Agric. For. Meteorol.* **339**, 109585 (2023).
- Lal, R. Food security impacts of the “4 per Thousand” initiative. *Geoderma* **374**, 114427 (2020).
- Lal, R. Soil carbon sequestration impacts on global climate change and food security. *Science* **304**, 1623–1627 (2004).
- Jiang, M., Li, H., Zhang, W., Liu, J. & Zhang, Q. Effects of climate change and grazing on the soil organic carbon stock of alpine wetlands on the Tibetan Plateau from 2000 to 2018. *CATENA* **238**, 107870 (2024).
- Piao, S. et al. Responses and feedback of the Tibetan Plateau’s alpine ecosystem to climate change. *Chin. Sci. Bull.* **64**, 2842–2855 (2019).
- Ramesh, T. et al. in *Advances in Agronomy* Vol. 156 (ed. Sparks, D.L.) 1–107 (Academic Press, 2019).
- Poggio, L. et al. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL* **7**, 217–240 (2021).
- Mikhailova, E. A. et al. Comparing soil carbon estimates in glaciated soils at a farm scale using geospatial analysis of field and SSURGO data. *Geoderma* **281**, 119–126 (2016).
- Smith, P. et al. How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal. *Global Change Biol* **26**, 219–241 (2020).
- Yang, R.-M. et al. Exploring the likely relationship between soil carbon change and environmental controls using nonrevisited temporal data sets: mapping soil carbon dynamics across China. *Sci. Total Environ.* **800**, 149312 (2021).
- Minasny, B., McBratney, A. B., Malone, B. P. & Wheeler, I. in *Advances in Agronomy*, Vol. 118 (ed. Sparks, D.L.) 1–47 (Academic Press, 2013).
- Heuvelink, G. B. M. et al. Machine learning in space and time for modelling soil organic carbon change. *Eur. J. Soil Sci.* **72**, 1607–1623 (2021).
- Abramoff, R. et al. The Millennial model: in search of measurable pools and transformations for modeling soil carbon in the new century. *Biogeochemistry* **137**, 51–71 (2018).
- Li, S. et al. Testing the RothC and DNDC models against long-term dynamics of soil organic carbon stock observed at cropping field soils in North China. *Soil Tillage Res.* **163**, 290–297 (2016).
- Xie, E. et al. Integration of a process-based model into the digital soil mapping improves the space-time soil organic carbon modelling in intensively human-impacted area. *Geoderma* **409**, 115599 (2022).

22. Zhang, X. et al. Modelling the spatiotemporal dynamics of cropland soil organic carbon by integrating process-based models differing in structures with machine learning. *J. Soils Sediments* **23**, 2816–2831 (2023).
23. Ellili, Y., Walter, C., Michot, D., Pichelin, P. & Lemerrier, B. Mapping soil organic carbon stock change by soil monitoring and digital soil mapping at the landscape scale. *Geoderma* **351**, 1–8 (2019).
24. McBratney, A., Mendonça Santos, M. & Minasny, B. On digital soil mapping. *Geoderma* **117**, 3–52 (2003).
25. Huang, H. et al. A review on digital mapping of soil carbon in cropland: progress, challenge, and prospect. *Environ. Res. Lett.* **17**, 123004 (2022).
26. Arrouays, D., Mulder, V. L. & Richer-de-Forges, A. C. Soil mapping, digital soil mapping and soil monitoring over large areas and the dimensions of soil security – a review. *Soil Security* **5**, 100018 (2021).
27. Helfenstein, A., Mulder, V. L., Heuvelink, G. B. M. & Hack-ten Broeke, M. J. D. Three-dimensional space and time mapping reveals soil organic matter decreases across anthropogenic landscapes in the Netherlands. *Commun. Earth Environ.* **5**, 130 (2024).
28. Szatmári, G. et al. Space-time modelling of soil organic carbon stock change at multiple scales: case study from Hungary. *Geoderma* **451**, 117067 (2024).
29. Li, N., Zhou, S. & Margenot, A. J. From prairie to crop: spatiotemporal dynamics of surface soil organic carbon stocks over 167 years in Illinois, U.S.A. *Sci. Total Environ.* **857**, 159038 (2023).
30. Yang, L. et al. Evaluation of conditioned Latin hypercube sampling for soil mapping based on a machine learning method. *Geoderma* **369**, 114337 (2020).
31. Schillaci, C. et al. Spatio-temporal topsoil organic carbon mapping of a semi-arid Mediterranean region: the role of land use, soil texture, topographic indices and the influence of remote sensing data to modelling. *Sci. Total Environ.* **601–602**, 821–832 (2017).
32. Liu, S. et al. The Unified North American Soil Map and its implication on the soil organic carbon stock in North America. *Biogeosciences* **10**, 2915–2930 (2013).
33. Gurney, K. R. et al. Towards robust regional estimates of CO<sub>2</sub> sources and sinks using atmospheric transport models. *Nature* **415**, 626–630 (2002).
34. Franzmeier, D. P., Lemme, G. D. & Miles, R. J. Organic carbon in soils of north central United States. *Soil Sci. Soc. Am. J.* **49**, 702–708 (1985).
35. Sims, Z. R. & Nielsen, G. A. Organic carbon in montana soils as related to clay content and climate. *Soil Sci. Soc. Am. J.* **50**, 1269–1271 (1986).
36. Guo, Y., Amundson, R., Gong, P. & Yu, Q. Quantity and spatial variability of soil carbon in the conterminous United States. *Soil Sci. Soc. Am. J.* **70**, 590–600 (2006).
37. Mishra, U., Lai, R., Liu, D. & Van Meirvenne, M. Predicting the spatial variation of the soil organic carbon pool at a regional scale. *Soil Sci. Soc. Am. J.* **74**, 906–914 (2010).
38. Ogle, S. M. et al. Counterfactual scenarios reveal historical impact of cropland management on soil organic carbon stocks in the United States. *Sci. Rep.* **13**, 14564 (2023).
39. Guevara, M. et al. Soil organic carbon across Mexico and the Conterminous United States (1991–2010). *Global Biogeochem. Cycles* **34**, e2019GB006219 (2020).
40. Wang, Z. et al. Upscaling soil organic carbon measurements at the continental scale using multivariate clustering analysis and machine learning. *J. Geophys. Res. Biogeosci.* **129**, e2023JG007702 (2024).
41. Gonçalves, D. R. P., Mishra, U., Wills, S. & Gautam, S. Regional environmental controllers influence continental scale soil carbon stocks and future carbon dynamics. *Sci. Rep.* **11**, 6474 (2021).
42. Hengl, T. et al. SoilGrids250m: global gridded soil information based on machine learning. *PLoS One* **12**, e0169748 (2017).
43. Keskin, H. & Grunwald, S. Regression kriging as a workhorse in the digital soil mapper’s toolbox. *Geoderma* **326**, 22–41 (2018).
44. Johnson, D. W., Todd Jr., D. E., Trettin, C. F. & Sedinger, J. S. Soil carbon and nitrogen changes in forests of walker branch watershed, 1972 to 2004. *Forest Range Wildland Soils* **71**, 1639–1646 (2007).
45. García-Palacios, P. et al. Evidence for large microbial-mediated losses of soil carbon under anthropogenic warming. *Nat. Rev. Earth Environ.* **2**, 507–517 (2021).
46. Sáez-Sandino, T. et al. Increasing numbers of global change stressors reduce soil carbon worldwide. *Nat. Clim. Change* **14**, 740–745 (2024).
47. Wang, M. et al. Global soil profiles indicate depth-dependent soil carbon losses under a warmer climate. *Nat. Commun.* **13**, 5514 (2022).
48. Luo, Z., Wang, G. & Wang, E. Global subsoil organic carbon turnover times dominantly controlled by soil properties rather than climate. *Nat. Commun.* **10**, 3688 (2019).
49. Wang, W. et al. Long-term cereal-legume intercropping accelerates soil organic carbon loss in subsoil of dryland. *Resour. Conserv. Recycl.* **211**, 107898 (2024).
50. Zhou, Z. et al. Carbon gain in upper but loss in deeper cropland soils across China over the last four decades. *Proc. Natl Acad. Sci.* **122**, e2422371122 (2025).
51. Mayer, S. et al. Soil organic carbon sequestration in temperate agroforestry systems – a meta-analysis. *Agric. Ecosyst. Environ.* **323**, 107689 (2022).
52. Liu, Y., Dang, Z.-Q., Tian, F.-P., Wang, D. & Wu, G.-L. Soil organic carbon and inorganic carbon accumulation along a 30-year grassland restoration chronosequence in semi-arid regions (China). *Land Degrad. Dev.* **28**, 189–198 (2017).
53. Liu, C. et al. Molecular insight into the vertical migration and degradation of dissolved organic matter in riparian soil profiles. *Environ. Res.* **245**, 118013 (2024).
54. Luo, Z., Viscarra Rossel, R. A. & Shi, Z. Distinct controls over the temporal dynamics of soil carbon fractions after land use change. *Global Change Biol* **26**, 4614–4625 (2020).
55. Chen, F. et al. Cropland carbon stocks driven by soil characteristics, rainfall and elevation. *Sci. Total Environ.* **862**, 160602 (2023).
56. Núñez, A., Cotrufo, M. F. & Schipanski, M. Irrigation effects on the formation of soil organic matter from aboveground plant litter inputs in semiarid agricultural systems. *Geoderma* **416**, 115804 (2022).
57. Chenu, C. et al. Increasing organic stocks in agricultural soils: knowledge gaps and potential innovations. *Soil Tillage Res* **188**, 41–52 (2019).
58. Shakoor, A. et al. A global meta-analysis of greenhouse gases emission and crop yield under no-tillage as compared to conventional tillage. *Sci. Total Environ.* **750**, 142299 (2021).
59. De Rosa, D. et al. Soil organic carbon stocks in European croplands and grasslands: How much have we lost in the past decade? *Global Change Biol* **30**, e16992 (2024).
60. Herzfeld, T., Heinke, J., Rolinski, S. & Müller, C. Soil organic carbon dynamics from agricultural management practices under climate change. *Earth Syst. Dyn.* **12**, 1037–1055 (2021).
61. Yang, Y. et al. Terrestrial carbon sinks in China and around the world and their contribution to carbon neutrality. *Sci. Sin. Vitae* **52**, 534–574 (2022).
62. Zhou, G. et al. Old-growth forests can accumulate carbon in soils. *Science* **314**, 1417–1417 (2006).
63. TANG, J., BOLSTAD, P. V. & MARTIN, J. G. Soil carbon fluxes and stocks in a Great Lakes forest chronosequence. *Global Change Biol* **15**, 145–155 (2009).
64. Kaarakka, L., Cornett, M., Domke, G., Ontl, T. & Dee, L. E. Improved forest management as a natural climate solution: a review. *Ecol. Solutions Evid.* **2**, e12090 (2021).
65. Post, W. M. et al. Enhancement of carbon sequestration in US soils. *BioScience* **54**, 895–908 (2004).
66. Bellamy, P. H., Loveland, P. J., Bradley, R. I., Lark, R. M. & Kirk, G. J. D. Carbon losses from all soils across England and Wales 1978–2003. *Nature* **437**, 245–248 (2005).

67. Guo, M. et al. Hydrology, vegetation, and soil properties as key drivers of soil organic carbon in coastal wetlands: a high-resolution study. *Environ. Sci. Ecotechnol.* **23**, 100482 (2025).
68. Sun, X.-L. et al. Spatiotemporal modelling of soil organic matter changes in Jiangsu, China between 1980 and 2006 using INLA-SPDE. *Geoderma* **384**, 114808 (2021).
69. Batjes, N. H., Calisto, L. & de Sousa, L. M. Providing quality-assessed and standardised soil data to support global mapping and modelling (WoSIS snapshot 2023). *Earth Syst. Sci. Data* **16**, 4735–4765 (2024).
70. Nave, L. et al. International Soil Carbon Network version 3 Database (ISCN3) ver 1. (ed Environmental Data Initiative) (2022).
71. Bishop, T. F. A., McBratney, A. B. & Laslett, G. M. Modelling soil attribute depth functions with equal-area quadratic smoothing splines. *Geoderma* **91**, 27–45 (1999).
72. Chen, S. et al. European topsoil bulk density and organic carbon stock database (0–20 cm) using machine-learning-based pedotransfer functions. *Earth Syst. Sci. Data* **16**, 2367–2383 (2024).
73. Nelson, D. W. & Sommers, L. E. Total Carbon, Organic Carbon, and Organic Matter. In *methods of soil analysis*. 961–1010 (1996).
74. Patoine, G. et al. Drivers and trends of global soil microbial carbon over two decades. *Nat. Commun.* **13**, 4195 (2022).
75. MacDonald, H. et al. North American historical monthly spatial climate dataset, 1901–2016. *Sci. Data* **7**, 411 (2020).
76. Omernik, J. M. & Griffith, G. E. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. *Environ. Manag.* **54**, 1249–1266 (2014).
77. Winkler, K., Fuchs, R., Rounsevell, M. & Herold, M. Global land use changes are four times greater than previously estimated. *Nat. Commun.* **12**, 2501 (2021).
78. Yang, C., Yang, L., Zhang, L. & Zhou, C. Soil organic matter mapping using INLA-SPDE with remote sensing based soil moisture indices and Fourier transforms decomposed variables. *Geoderma* **437**, 116571 (2023).
79. Yang, L. et al. Predicting soil organic carbon content in croplands using crop rotation and Fourier transform decomposed variables. *Geoderma* **340**, 289–302 (2019).
80. Yang, R.-M. et al. The effectiveness of digital soil mapping with temporal variables in modeling soil organic carbon changes. *Geoderma* **405**, 115407 (2022).
81. Szatmári, G., Pásztor, L. & Heuvelink, G. B. M. Estimating soil organic carbon stock change at multiple scales using machine learning and multivariate geostatistics. *Geoderma* **403**, 115356 (2021).
82. Hengl, T. & MacMillan, R. A. *Predictive soil mapping with R* (OpenGeoHub foundation, 2019).
83. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
84. Camera, C. et al. A high resolution map of soil types and physical properties for Cyprus: a digital soil mapping optimization. *Geoderma* **285**, 35–49 (2017).
85. Pahlavan-Rad, M. R. & Akbarimoghaddam, A. Spatial variability of soil texture fractions and pH in a flood plain (case study from eastern Iran). *CATENA* **160**, 275–281 (2018).
86. Beguin, J., Fuglstad, G.-A., Mansuy, N. & Paré, D. Predicting soil properties in the Canadian boreal forest with limited data: comparison of spatial and non-spatial statistical approaches. *Geoderma* **306**, 195–205 (2017).
87. Gomes, L. C. et al. Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma* **340**, 337–350 (2019).
88. Guyon, I., Weston, J., Barnhill, S. & Vapnik, V. Gene selection for cancer classification using support vector machines. *Mach. Learn.* **46**, 389–422 (2002).
89. Nussbaum, M. et al. Evaluation of digital soil mapping approaches with large sets of environmental covariates. *SOIL* **4**, 1–22 (2018).
90. Linnenbrink, J., Milà, C., Ludwig, M. & Meyer, H. kNNDM CV: k-fold nearest-neighbour distance matching cross-validation for map accuracy estimation. *Geosci. Model Dev.* **17**, 5897–5912 (2024).
91. Carrasco, J. L., Phillips, B. R., Puig-Martinez, J., King, T. S. & Chinchilli, V. M. Estimation of the concordance correlation coefficient for repeated measures using SAS and R. *Comput. Methods Programs Biomed.* **109**, 293–304 (2013).
92. Lawrence, I. K. L. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **45**, 255–268 (1989).
93. Meinshausen, N. Quantile regression forests. *J. Mach. Learn. Res.* **7**, 983–999 (2006).
94. Takoutsing, B. & Heuvelink, G. B. M. Comparing the prediction performance, uncertainty quantification and extrapolation potential of regression kriging and random forest while accounting for soil measurement errors. *Geoderma* **428**, 116192 (2022).
95. Arrouays, D. et al. in *Advances in Agronomy*. Vol. 125 (ed. Sparks, D. L.) 93–134 (Academic Press, 2014).
96. Yuan, Y. et al. Point process models for spatio-temporal distance sampling data from a large-scale survey of blue whales. *Ann. Appl. Stat.* **11**, 2270–2297 (2017).
97. Schmidinger, J. & Heuvelink, G. B. M. Validation of uncertainty predictions in digital soil mapping. *Geoderma* **437**, 116585 (2023).
98. Ray, D. K., Ramankutty, N., Mueller, N. D., West, P. C. & Foley, J. A. Recent patterns of crop yield growth and stagnation. *Nat. Commun.* **3**, 1293 (2012).

## Acknowledgements

This study is supported by the National Natural Science Foundation of China (Project No. 42471468), Fundamental Research Funds for the Central Universities (Grant No. 0209-14380115) and the Special Fund of Jiangsu Province Carbon Peak and Carbon Neutral Technology Innovation (Grant No. BK20220037).

## Author contributions

Chenconghai Yang: Writing – original draft, Software, Resources, Methodology, Investigation. Feixue Shen: Writing – original draft, Methodology. Xiang Li: Writing – review & editing. Wenkai Cui: Methodology. Lei Zhang: Writing – review & editing. Lin Yang: Writing – review & editing, Supervision, Resources, Methodology, Investigation, Formal analysis, Data curation. Chenghu Zhou: Supervision, Funding acquisition.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s43247-025-02605-6>.

**Correspondence** and requests for materials should be addressed to Lin Yang.

**Peer review information** *Communications Earth & Environment* thanks Qiuxiang Tian and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Mengze Li, Somaparna Ghosh A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025

---

**Chenconghai Yang<sup>1,5</sup>, Feixue Shen<sup>1,5</sup>, Xiang Li<sup>1</sup>, Wenkai Cui<sup>1</sup>, Lei Zhang<sup>2</sup>, Lin Yang<sup>3,4</sup>✉ & Chenghu Zhou<sup>1,4</sup>**

---

<sup>1</sup>School of Geography and Ocean Science, Nanjing University, Nanjing, China. <sup>2</sup>Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. <sup>3</sup>Frontiers Science Center for Critical Earth Material Cycling, Nanjing University, Nanjing, China. <sup>4</sup>State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China. <sup>5</sup>These authors contributed equally: Chenconghai Yang, Feixue Shen. ✉e-mail: [yanglin@nju.edu.cn](mailto:yanglin@nju.edu.cn)