

<https://doi.org/10.1038/s43247-024-01299-6>

Climate and land use changes explain variation in the A horizon and soil thickness in the United States

Check for updates

Yakun Zhang ^{1,2}✉, Alfred E. Hartemink ¹, Tom Vanwalleghem ³, Benito Roberto Bonfatti ⁴ & Steven Moen ⁵

Understanding spatio-temporal changes in soil thickness and their natural and anthropogenic driving factors are essential for earth system modeling and natural resource conservation. It remains challenging to accurately quantify the spatial pattern of soil thickness, and there is no assessment of its temporal changes at the national scale across eco-climatic zones. Here we compiled a long-term (1950–2018), large-scale (conterminous United States of America) topsoil (A horizon, $n = 37,712$) and solum (22,409) thickness data to quantify their spatial and temporal variations using generalized additive models and selected chronosequences in land resource regions. Climate was found associated with the spatial distribution of soil thickness, and land use and erosion associated with its temporal variation. The A horizon and solum thickness displayed strong longitudinal patterns, correlated with soil moisture and temperature, respectively. Temporal changes in the thickness varied across land resource regions, affected by topography, land use, and erosion. Severe A horizon loss primarily occurred in Mollisols of the Central Great Plains, Alfisols on steep slopes, and soils under cropping. These findings enhanced our fundamental understanding of soil formation and biogeochemical cycles during the Anthropocene across scales and identified regions for conservation practices to reduce further topsoil loss.

Soil thickness plays an important role in global hydrological and ecological processes¹. The thickness of topsoil, the organic matter and nutrient-rich and biologically active layer influences plant growth and crop yield^{2,3}, carbon storage⁴, and biogeochemical cycles⁵. The topsoil (A horizon) is a mix of living and decomposed organic matter from plants and animals and mineral particles and its formation is controlled by biotic and abiotic factors⁶. The solum thickness is determined by the balance between soil production from weathering, additions by sedimentation and atmospheric deposition, and soil loss by erosion⁷, and the thickness is largely controlled by soil forming factors (e.g., parent material, climate, topography, vegetation)⁸ and water, wind, and tillage^{9,10}. As a result, both the A-horizon and soil thickness vary spatially and temporally.

Accurately representing the spatial variation of soil thickness has recently received more attention in earth system models, which often use a constant soil depth value and cannot represent real world conditions¹¹. The

spatial distribution of soil thickness has been quantified through mechanistic¹² and empirical¹³ models. Mechanistic models assume the long-term equilibrium state of soil production and losses and quantify them using soil production functions and sediment transport models, respectively¹⁴. The equilibrium state may be rarely achieved, and variation of shallow and deep soils may contradict the concept that soil weathering rate can self-regulate with soil thickness changes¹⁵. Additionally, short-term changes in soil weathering are small compared to the whole soil regolith which may limit its applicability in investigating the soil thickness change at a decadal scale. Empirical models resolve soil thickness as a function of environmental variables that regulate soil formation, and different statistical or machine learning models have been developed¹³ (Supplementary Table 1). This has been used to map the spatial distribution of soil thickness and soil horizon thickness from local to global scales^{11,13,16,17}. Due to the high variation of soil thickness at short distances and difficulty in the thickness measurements

¹Department of Soil Science, University of Wisconsin-Madison, FD Hole Soils Lab, 1525 Observatory Drive, Madison, WI 53706, USA. ²Department of Crop and Soil Science, Oregon State University, 2750 SW Campus Way, Corvallis, OR 97331, USA. ³Department of Agronomy, University of Córdoba, Da Vinci building, Ctra. Madrid km 396, 14071 Córdoba, Spain. ⁴State University of Minas Gerais, 700 Colorado Street, Passos, Minas Gerais 37902-092, Brazil. ⁵Department of Statistics, University of Wisconsin-Madison, Medical Sciences Center, 1300 University Avenue, Madison, WI 53706, USA. ✉e-mail: yakun.zhang@oregonstate.edu

and the lack of data, it remains a challenge to accurately predict soil thickness and understand its distribution pattern¹⁸.

Soil formation is slow, and it is widely recognized that it takes 1000 years to form about 2.5 cm of soil^{19,20}. It can be faster in many regions, e.g., 250 cm ky⁻¹ in the Southern Alps of New Zealand²¹. The time needed to form a Mollic epipedon ranged from less than 200 years to as fast as 30 to 60 years in different parts of the US²². Soil loss has been accelerated by climate change and human activities²³, and the available data suggest that soil erosion rates are an order of magnitude higher (394 cm ky⁻¹) than soil formation rates (3.6 cm ky⁻¹)^{24,25}. In the US, modeled nationwide water and wind erosion rates on cropland varied from 18 to 12 ton ha⁻¹ year⁻¹ between 1982 and 2007²⁶. In the US Corn Belt, one-third of the cultivated soils have lost their A horizon²⁷, and soil thickness has been reduced by 4 to 69 cm in croplands compared to adjacent prairies²⁸. Conservation practices can reduce soil erosion whereas soil erosion can cause soil deposition in downslope areas. Tillage may mix the topsoil and subsoil and affect topsoil thickness²⁹. To the best of our knowledge, there is no assessment of temporal changes in soil thickness at a national scale across diverse eco-climatic zones. Such information is, however, essential for our understanding of soil losses under climate change and intensified human-induced activities and the consequences of conservation efforts.

Here, we used a long-term, large-scale, in-situ soil survey dataset to quantify the spatial and temporal variations of A horizon and solum thickness across the conterminous US (CONUS) over 69 years. The objectives of this study are: (1) to study the spatial distribution of A horizon and solum thickness across the CONUS and in land resource regions and quantify the effects of soil forming factors, and (2) to investigate the temporal variations of A horizon and solum thickness using selected chronosequences in land resource regions and understand their driving factors. We hypothesize that (1) the national-scale spatial variations of A horizon and solum thickness are mainly controlled by natural soil forming factors, in which climate conditions have a more significant impact on soil formation and soil thickness, followed by topography, and land cover types, and the influence of climate (precipitation and temperature) on soil thickness is strongest in arid and hot regions. (2) Temporal variations of A horizon and solum thickness are mainly driven by human activities (e.g., land cover and land use change, tillage).

Results

Spatial pattern of A horizon and solum thickness

In the CONUS, the A horizon was the shallowest in the west including the desert and Rocky Mountain regions, and shallower in the east along the Great Lakes and Appalachian Mountains, but deeper along the Mississippi River Basin and the West Coast (Fig. 1a). The solum thickness was the shallowest in southwest desert and Nebraska Sandhills, and deeper in the southeast along the Gulf of Mexico (Fig. 1a). This was similar to the soil thickness map (0–2 m) on uplands¹², while the censored global map of depth to bedrock (0–2 m) displayed uniform pattern across the CONUS¹¹. The A horizon and solum thickness displayed a strong longitudinal pattern (Figs. 1a, 2a). From west to east, the A horizon thickness decreased first and then increased. The solum thickness varied similarly to the A horizon thickness in the west but it decreased on the East Coast (Fig. 2a). Along the latitude, A horizon showed an opposite pattern to the solum thickness. The A horizon was shallow in the south but thicker in the north, whereas the solum thickness increased slightly and then decreased continuously towards the north (Fig. 2b).

In some soils, the A horizon thickness equaled the solum thickness (Fig. 1d), which indicated that the pedons had A horizons directly over C horizons and no B horizons. A deeper solum did not coincide with a thicker A horizon (Fig. 1d, Pearson correlation coefficient, $r = 0.16$), but such relationships varied by soil order. Deeper Mollisols, Entisols, and Inceptisols tended to have thicker A horizons, whereas, in Aridisols, Spodosols, and Ultisols, A horizons were shallow irrespective of solum thickness (Fig. 1d). A study conducted on a 60,000-ha forest land with mostly Entisols and

Inceptisols showed that solum thickness was weakly correlated with A horizon thickness ($r = 0.16$)³⁰.

Environmental controls on the spatial variation of A horizon and solum thickness

The longitudinal and latitudinal distribution of soil thickness was significantly correlated with climatic and topographic variables. For the longitudinal pattern, the A horizon thickness was correlated with moisture (pr, ro, sm, $r = 0.47$ – 0.49) but not with temperature, whereas the solum thickness was strongly correlated with temperature (tmmn, tmmx, aet, $r = 0.63$ – 0.74) (Fig. 2). Elevation showed opposite patterns to soil thickness ($r = -0.54$ to -0.36) and temperature and moisture along the longitude (Fig. 2), where the high elevation in the west (Colorado Plateau and Rocky Mountains) matched well with low moisture and temperature regions and shallow soils (Fig. 2). Along the latitude, temperature, and moisture affected differently on A horizon and solum thickness (Fig. 2). Temperature negatively correlated with A horizon thickness (tmmn, tmmx, $r = -0.38$ to -0.37), but positively correlated with solum thickness (tmmn, tmmx, $r = 0.41$ – 0.49). The solum thickness positively correlated with moisture deficits (vpd, def, pet, srad, $r = 0.45$ – 0.56), while the A horizon thickness negatively correlated with moisture deficits ($r = -0.44$ to -0.32). The increasing elevation from south to north negatively correlated with solum thickness ($r = -0.42$). Profile curvature negatively affected A horizon thickness along the latitude ($r = -0.26$), where positive curvature (convex terrain) matched with thin soils and negative curvature (concave terrain) matched with thick soils.

We used an empirical model (generalized additive model, GAM) to analyze the environmental controls on soil thickness at the national scale (Fig. 3) and the modeling performances were acceptable compared to other studies (Supplementary Table 1). Soil thickness decreased exponentially with slope until 30°. The data beyond 30° were too few ($n = 292$ and 223 for A horizon and solum respectively). The young soils (Entisols, Inceptisols) had a shallow solum, while intensely weathered soils (Ultisols) had a deep solum. Mollisols had the thickest A horizon (Fig. 3), while shallow A horizons occurred in Spodosols, soils with aridic conditions with limited vegetation growth (Aridisols), and the highly weathered soils of hot and humid regions (Ultisols). The soil was the deepest in hot regions (Hyperthermic, Isomesic), whereas the A horizon was the thickest in moderate-temperature regions (Isomesic). In dryer regions (Aridic, Xeric), the A horizon and solum were significantly thinner. Soils developed in alluvium and coastal sediments (West and East Coast, Mississippi River Basin), eolian sediment (Central), and glacial sediments (lake, outwash, till in Midwest, Great Lakes, and Northeast) had a thicker A horizon and solum. The A horizon thickness showed a larger variation across land-use types than the solum thickness. At the national scale, the A horizon thickness was thicker in cropland, pasture, developed, wetland, and grassland than under barren, forest, and shrubland. The barren, grassland, shrubland, and wetland had shallow soils.

We developed regional GAMs to specifically investigate the regional controlling factors on soil thickness (Fig. 3) and the modeling performances varied in different regions (Supplementary Fig. 7). The effects of slope varied in different regions (Supplementary Table 3). The negative control of slope on A horizon thickness primarily occurred in the eastern half of CONUS, where the soil moisture content was high with Udic moisture regime (Supplementary Figs. 1 and 2) which may increase runoff and erosion. At regional scales, the soil order, parent materials, and land uses played a more important role than climate variables (Fig. 3), where the land use greatly affected the A horizon thickness, while the parent material greatly affected the solum thickness. Some environmental variables showed contrasting effects in different regions. For example, the Histosols were deeper in region C (drier), but shallower in region B (wetter) than Alfisols (Fig. 3). Soils developed in glacial outwash were shallower in region D but deeper in region A and E. The A horizons of cropland soils were thicker in region D, but shallower in region F.

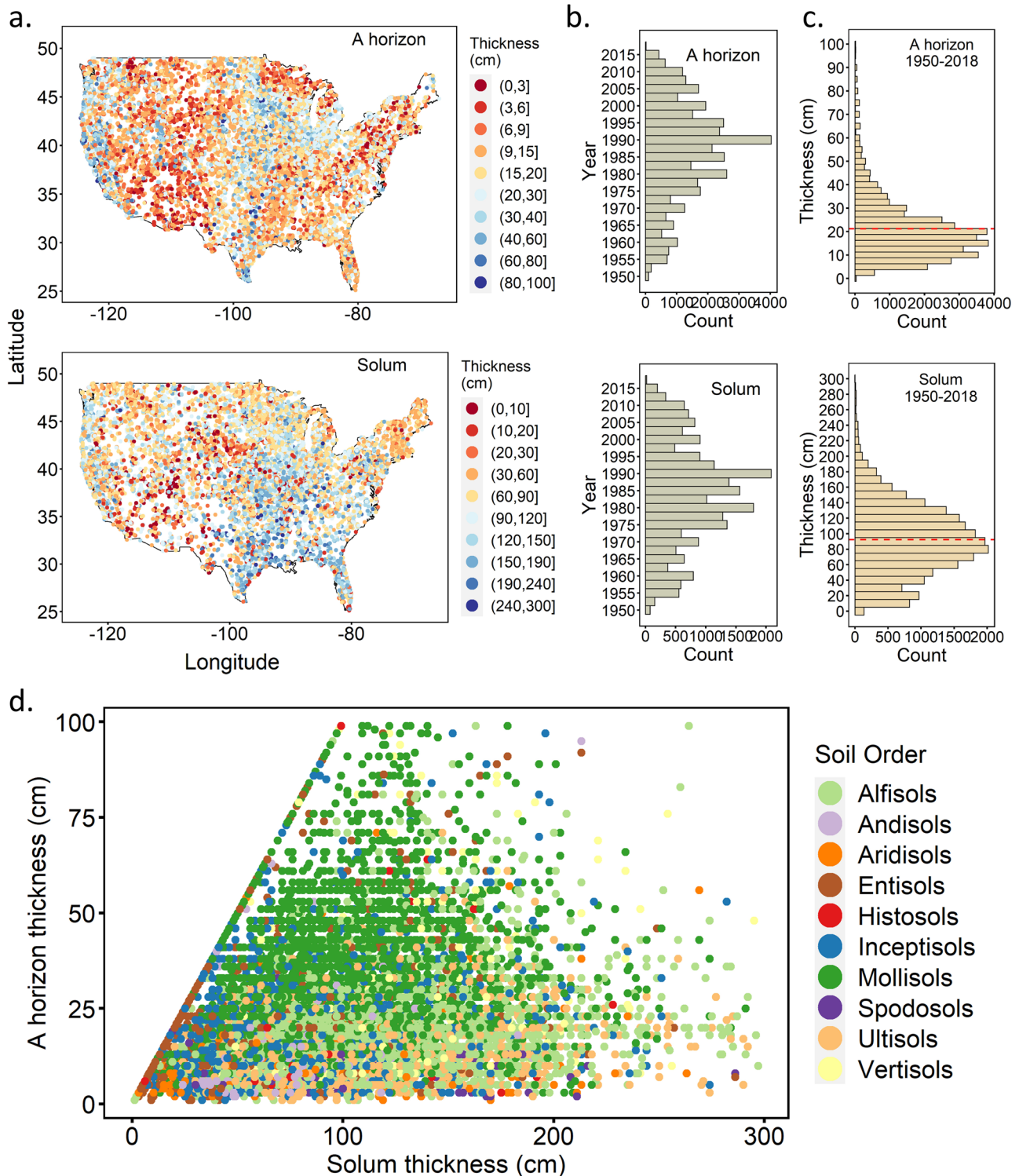


Fig. 1 | Distribution of A horizon and solum thickness. **a** The spatial distribution of measured A horizon thickness ($n = 37,712$) and solum thickness ($n = 22,409$) across the conterminous US from 1950 to 2018. Blue circles indicate deeper soils, while red circles indicate shallower soils. **b** The temporal distribution of collected samples

from 1950 to 2018. **c** The distribution of measured thickness of A horizon and solum with dashed red lines indicating the mean thickness values. **d** The relationship between measured A horizon and solum thickness for different soil orders represented by different colors.

Temporal pattern of A horizon and solum thickness and driving factors

Chronosequences were selected at the regional scale to investigate the temporal distribution of A horizon and solum thickness (Fig. 4, Supplementary Tables 4 and 5). The Mollisols under cropland in region H (Central Great Plains) had lost the highest amount of A horizon soils at an average

rate of 0.44 cm yr^{-1} . Under forest, the Alfisols with a steep slope (13°) in region C (California) and Inceptisols in region R (Northeast) also lost a high amount of soil (0.26 cm yr^{-1} and 0.20 cm yr^{-1}). In region M (Midwest), a decreasing A horizon thickness was observed in most of the chronosequences with an average decreasing rate of 0.12 cm yr^{-1} . The cropping land use has resulted in greater A horizon soil loss (0.35 cm yr^{-1}) than other

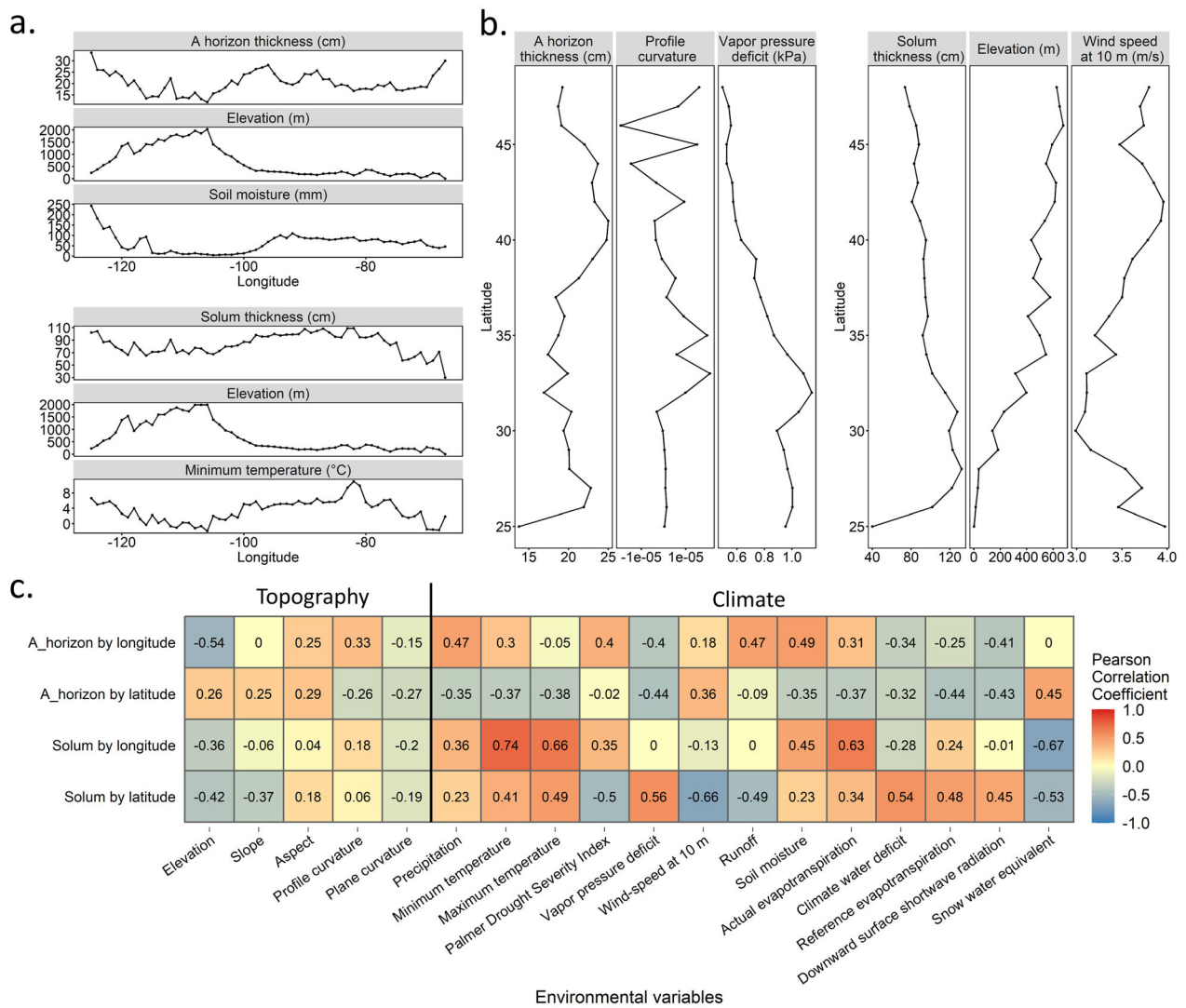


Fig. 2 | Longitudinal and latitudinal distribution of A horizon and solum thickness and their environmental controls. **a** The distribution of the longitudinal zonal means of A horizon ($n = 37,712$) and solum ($n = 22,409$) thickness (cm) and selected topographic and climatic variables. The subtle difference in the longitudinal distributions of elevation is due to the different sample sizes of soil A horizon and solum thickness measurements. **b** The distribution of the latitudinal zonal means of A horizon and solum thickness and selected topographic and climatic variables. **c** Pearson correlation coefficients of the longitudinal and latitudinal zonal means of

A horizon and solum thickness and topographic and climatic variables. Positive correlations are shown in red, while negative correlations are shown in blue. The topographic and climatic variables in (a, b) were selected due to their strongest Pearson correlations in (c), except for the A horizon by latitude, where profile curvature was selected because (1) none of the topographic variables have strong relationship ($-0.3 < r < 0.3$) with A horizon thickness along the latitude and (2) profile curvature showed clearer pattern than other topographic variables.

land use types in Alfisols. For Mollisols, the cropland formed in alluvium and coastal sediment and glacial lake sediment had thicker A horizons (41 cm and 38 cm) than that formed in eolian and glacial till (35 cm) (Supplementary Table 4). The Mollisols formed in glacial lake sediment had the highest A horizon soil loss (0.66 cm yr^{-1}). The Mollisols formed in glacial till under cropping land use also lost a significant amount of A horizon soil (0.45 cm yr^{-1}). The soils in region N (Appalachian Mountains) were primarily formed in carbonate or non-carbonate parent materials, and the land uses were forest or pasture. Most of these chronosequences showed an increase in A horizon thickness. There was no significant difference observed for soils between north-facing and south-facing landscapes. As for solum, the Alfisols under the forest in region K (Northern Lake States) have lost a high amount of soil (1.05 cm yr^{-1}). In region M (Midwest), the Alfisols under cropland have lost more soil than that under forest or pasture. Overall, temporal changes in soil thickness varied across different land resource regions as affected by topography, land use, and erosional

processes. Severe A horizon soil loss primarily occurred in Mollisols in Central Great Plain, Alfisols on steep slopes, and soils under cropping or cropland land use in the Midwest. Soils under forest and pasture showed an increase in A horizon thickness.

Discussion

The national-scale distribution of soil thickness was primarily determined by the climatic variables. Climate (precipitation and temperature) determines soil weathering, erosion, leaching, vegetation growth, organic matter accumulation and decomposition, and hence soil thickness⁸. Our results showed that two mechanisms were observed for the effects of moisture and temperature on A horizon and solum thickness. Soil moisture primarily regulates the formation and thickening of A horizon, while temperature controls the development of solum (Figs. 2 and 3). A higher temperature often leads to more soil development and hence deeper solum. However, A horizon was thicker in moderate temperature regions (Fig. 3), and the hotter

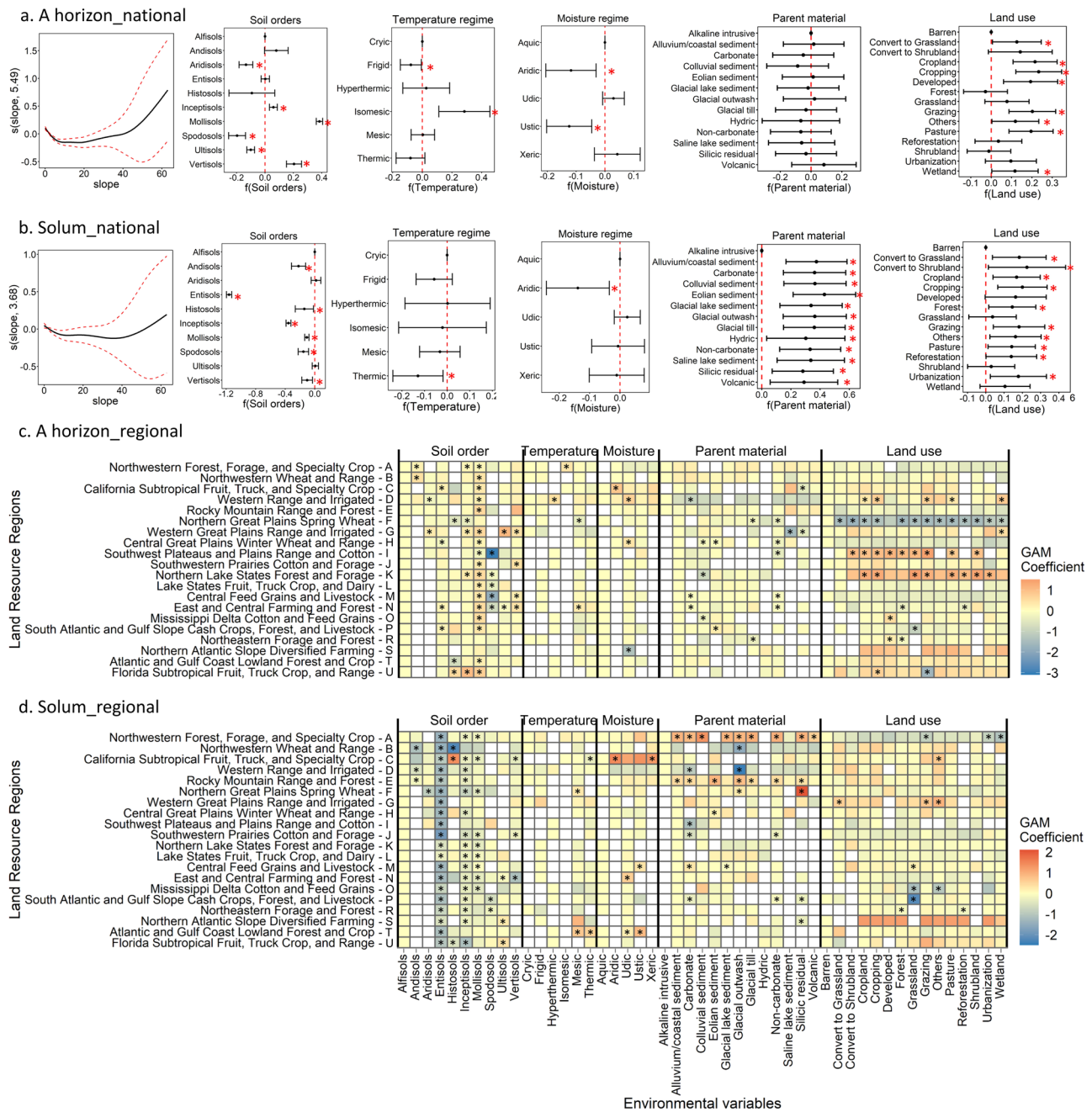


Fig. 3 | The environmental controls on the spatial variation of A horizon and solum thickness at the national scale and in land resource regions determined by generalized additive models (GAMs). a, b The effect of slope, soil order, soil temperature regime, soil moisture regime, parent material, and land use on the log-transformed A horizon and solum thickness (cm) across the conterminous US. The red dashed lines in the slope plots indicate the 95% confidence intervals of the coefficients. The red dashed lines indicate the reference type (i.e., Alfisols, Cryic, Aquic, Alkaline intrusive, and Barren) and the red stars indicate the significant difference from the reference type in each covariate (soil order, temperature regime,

moisture regime, parent material, land use). c, d The effect of soil order, soil temperature regime, soil moisture regime, parent material, and land use on the log-transformed A horizon and solum thickness in each land resource region. The white pixels indicate that specific types do not exist in certain regions. The stars indicate the significant difference from the reference type in each covariate in each region. Positive coefficients are shown in red, while negative coefficients are shown in blue. A detailed description of environmental variables is shown in Supplementary Table 1. The distribution of land resource regions is shown in Fig. 4.

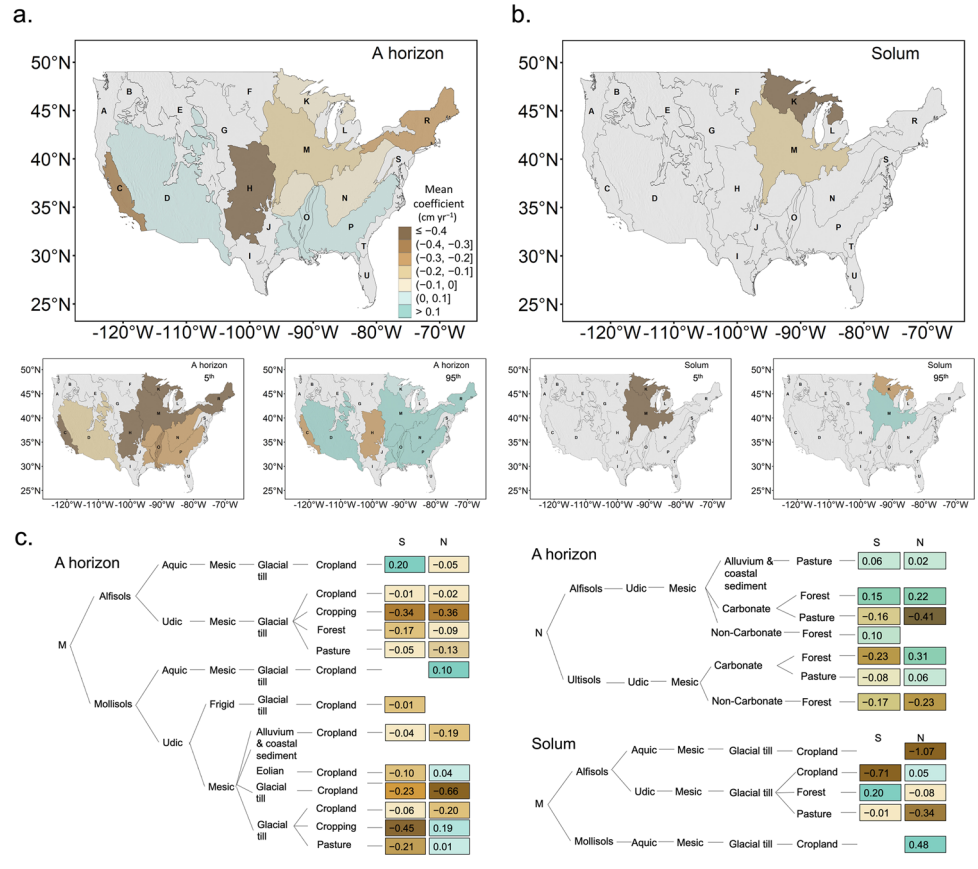
temperature in the south led to a shallower A horizon (Fig. 2). A low temperature limits vegetation growth and leads to a low organic matter input into the A horizon^{31,32}, and a high temperature expedites decomposition of organic matter in A horizons³³. Thus, the solum thickness had a stronger positive correlation with temperature than A horizon thickness. In dry regions, moisture deficit often limits soil development. However, the moisture deficits (vpd, def, pet, srad) positively correlated with the solum thickness but negatively correlated with the A horizon thickness along the

latitude. This may indicate that low temperature instead of moisture is the limiting factor for solum formation in the north and moist deficit has less impact on solum thickness than temperature (Fig. 2). The moisture effect on soil development was stronger for A horizon than for solum. Similarly, greater soil development and thicker A horizons were observed in humid climate than in dry climate in Brazil³⁴.

Topography affects soil erosion and redistribution, with thicker soils on the summit position or depositional landscapes and thinner soils at

Fig. 4 | Temporal change rates of A horizon and solum thickness in land resource regions.

a, b The mean temporal change rates (slope coefficients of linear regression models, cm yr^{-1}) of A horizon and solum thickness calculated for each land resource region ($n = 9$ regions for A horizon, $n = 2$ regions for solum) and their 5th and 95th percentile. Positive values (in blue) indicate increases of thickness, while negative values (in brown) indicate decrease of thickness. Note: the regions in gray have no data and are not calculated. **c** The anisotropic (north-facing – N or south-facing – S chronosequences) temporal change rates (slope coefficients of linear regression models) calculated within the specified soil order, moisture and temperature regime, parent material, and land use for A horizon (M and N regions only) and solum (M region only). The colors of the model coefficients indicated the change rate (cm yr^{-1}) and are on the same scale as Figs. 4a and 4b. Details about selected chronosequences and fitted linear models are shown in Supplementary Tables 4 and 5.



slopes and shoulders³⁵. Although the effect of topography is more prominent at the landscape scales by redistributing soil, affecting water flow, regulating local climate, and controlling vegetation types³⁶, it also influences heat and water distribution at the continental scale as shown in our study. The high elevation in the western CONUS (Colorado Plateau and Rocky Mountains) led to a low moisture and temperature and hence shallow soils. The northern CONUS had more negative profile curvature which corresponded to thicker topsoils. A soil order was defined by multiple factors, including climate, parent material, and soil development. As solum thickness was determined by soil development stages, more intensely weathered soils had a deeper solum. The A horizon thickness was determined by many environmental factors, and Mollisols had deeper A horizons.

Parent material affects soil fertility, salinity, sodicity, texture, structure, shrink/swell ability, erodibility, and soil thickness³⁷. It was found that carbonate rich parent materials had the shallowest soils followed by siliceous parent materials, while coarse-grained mafic parent materials led to the deepest soils³⁷. Moreover, the age of glacial deposits affected soil formation and horizon thickness³⁸. For example, the soils developed from Pliocene to early Pleistocene sediments, Illinoian sediments, late Illinoian and early Wisconsin sediments (older to newer) had argillic horizon (Bt) thickness of 250 cm, 51–54 cm, 34–92 cm, respectively, while the soils developed from the newest glacial sediments (late Wisconsin) had only Bw or Bk horizons³⁸. In our study, the effect of parent material was more prominent at the regional scale than at the national scale.

Land use may be constant or altered by human activities. Land use affected more on A horizons, whereas the B and C horizon thickness was primarily affected by soil-landscape processes instead of land use³⁹. Land use affects vegetation types (e.g., forest and grassland⁴⁰) and their root systems, thus controlling organic matter accumulation and A horizon thickness. Agricultural activities, especially tillage, may increase soil erosion and reduce topsoil⁴¹, but it may also thicken it by mixing subsoil and topsoil, and cultivation often includes input of fertilizer, irrigation water, and lime⁴, or addition of organic-rich materials⁴². Most of the Midwest is cropland and A

horizons were thicker, whereas shrubland and grassland were distributed in western CONUS and forest was distributed in northwestern and eastern CONUS with thinner A horizons (Fig. 1, Supplementary Fig. 3). The thicker A horizon in cropland may be due to the mixing effect of tillage and the continuous input of fertilizer and irrigation water leading to higher biomass production. Additionally, soils with thicker A horizons were prone to be used for crop production which may lead to the spatial pattern that cropland had thicker A horizons. Land use conversion has occurred in 16% of the land area at the national scale from 1950 to 2018. Cropping, grazing, and conversion to grassland occurred mainly in the Midwest, along the Mississippi River, at the East Coast and Great Plains, and corresponded with thicker A horizons, whereas reforestation occurred in the east (Fig. 3, Supplementary Fig. 3). Land uses may also affect local-scale variation of soil thickness, such as tree stump or tree overturn in the forest^{43,44}, but variation at that scale was not explored here.

As weathering is a relatively slow process, the temporal variation of soil thickness was dominantly controlled by soil loss, which was expedited due to climate change and intensified human activities⁴⁵. Land use conversion occurred in 16% of the land area (Supplementary Figs. 10 and 11). Specifically, cropping occurred primarily in regions B (Northwest), F (North Great Plains), M (Midwest), and O (Mississippi Delta). Reforestation occurred in the regions K (Northern Lake States), L (Lake States), N (Appalachians Mountains), P (Southern Atlantic), and R (Northeast). Urbanization occurred in the regions C (California), D (Southwest), S (Northern Atlantic), and U (Florida Subtropical). In many regions, such as E (Rocky Mountains), K (Northern Lake States), P (South Atlantic), and T (Atlantic and Gulf Coast), the land has been converted multiple times. As for climate, soil moisture significantly increased in regions K (Northern Lake State), L (Lake States), M (Midwest), N (Appalachian Mountains), O (Mississippi Delta), and R (Northeast) over the past 60 years (Supplementary Fig. 12). Although it was non-significant, soil moisture decreased in many regions in western CONUS. Temperature significantly increased at a rate of 0.02–0.03 yr^{-1} across the CONUS (Supplementary Fig. 13).

The large loss of A horizon in forest soils in region C was associated with the steep landscape (slope = 13°). Moreover, cropland in regions H and M lost a great amount of A horizon soil, whereas forest and pasture in region N did not show significant soil loss. Due to the lack of data, our study only evaluated temporal change since 1950, when the conservation practices have been used⁴⁶. The adoption of conservation tillage in cropland increased from 25% in 1985⁴⁷ to about 50% in 2018⁴⁸. The increased adoption of conservation practices reduced soil erosion, which decreased from 3.4 billion tons in 1982 to 2 billion tons in the 1990s⁴⁹. Although increased precipitation tended to increase soil erosion⁵⁰, the high adoption rate of conservation tillage in the Midwest helped to reduce erosion⁴⁸. It was predicted that the increase in rainfall and runoff erosivity due to climate change was significantly smaller in the Midwest than that in the northeastern and northwestern US⁵¹.

Although the spatiotemporal patterns in A horizon and solum thickness are significant and solid, there are some limitations of this study. (1) About 30–40% variation in A horizon and solum thickness in CONUS can be explained by spatial term and soil and environmental driving factors. The unexplained variation is due to the local-scale variation, which was not explored in this study. (2) The soil displayed significant temporal changes in the past seven decades. However, our study only evaluated temporal variation of soil thickness since widespread conservation practices have been adopted, which may underestimate the decrease in soil thickness caused by conventional agricultural activities. (3) The uneven number of observations per year (e.g., from 2 to 1460 for A horizon thickness) and inherently different soil thickness in different geographic locations and time may lead to randomness in the spatial and temporal variation of soil thickness. (4) We hypothesized that soil thickness changes are due to either erosion or soil formation, which may overlook the effects of tillage. Deep tillage may introduce organic matter into subsoils and lead to deeper Ap horizons, while conservation agriculture may not deepen the Ap horizon but reduce soil erosion. Soil organic matter (SOM) changes lead to changes in porosity/bulk density/soil structure and agricultural management practices that change SOM, affect soil thickness⁵². The relationship between soil thickness and changes of other soil properties (e.g., SOM, bulk density, soil structure) should be further explored⁵³. (5) O horizon is an important organic matter accumulative layer at the soil surface which contributes to carbon sequestration. But the O horizon thickness was not further investigated in this study (Supplementary Fig. 14).

Conclusions

The spatial and temporal variations of A horizon and solum thickness were investigated across the CONUS from 1950 to 2018 and natural and human driving factors were quantified for different land resource regions. We found that climatic variables regulated soil thickness at the national scale, in which moisture was associated with the formation of the A horizon, while temperature was associated with the development of solum. Elevation and climate affected the longitudinal pattern of soil thickness, while parent material affected soil thickness at local scales. The A horizon thickness was more influenced by land use than solum thickness. Land use and erosion process contributed to the temporal variation of soil thickness. Overall, this study provided an overview of soil thickness variation with respect to soil development and interactions with environmental and human factors across the CONUS. Regional soil thickness changes exploring land use changes or specific management practices should be studied.

Methods

The thickness dataset

The National Cooperative Soil Survey (NCSS) Soil Characterization Database collected by USDA-Natural Resources Conservation Service (NRCS) was downloaded and used in this study⁵⁴. This dataset contains soil characterization and analytical data with profile descriptions for pedons sampled since the 1900s across the US. The sampling locations were selected to best present the mapping units of the SSURGO map⁵⁵. Until now, about half of the mapping units have been sampled, and in some mapping units, more

than one pedon has been sampled. At each sampling location, the pedon was excavated by hand or with a backhoe to a maximum depth of 200 cm or bedrock. The pedons were described in the field with horizon designation and thickness recorded and samples were collected from delineated horizons for standard laboratory analysis⁵⁶.

Not all pedon data were included in the analysis. The following criteria were used to remove pedons: (1) The pedons that are outside of CONUS (48 states) or have no record of longitude and latitude coordinates; (2) The pedons that have no information on sampling year or were sampled before 1950; (3) The pedons that do not have horizon designation information or have reporting layers, D horizons, vesicular horizons, limnic horizons, or bi-sequences; (4) The pedons that have missing horizons, discontinuous horizons, or only one horizon recorded for topsoil or a fixed depth; (5) The pedons that were developed on human-transported materials; (6) After extracting soil order, soil temperature regime, soil moisture regime, parent material, and land use type to each pedon (section 2.2), the pedons in a specific type (e.g., Gelisols) that contains less than 10 pedons – the dataset is too small which may result in invalid statistical inference; (7) The pedons that are classified as water, mining, or mechanically disturbed in land use type from 1938 to 2018. The details about the pedon selection procedures are provided in Supplementary Note 1.

Two types of soil thickness were investigated in this study: (1) A horizon thickness, and (2) solum thickness. The A horizon is the organic matter-rich and biologically active layer at the upper part of the profile which is affected by environmental factors and human disturbance. The A horizons defined in this study contained multiple suffixes, such as Ap, Ab, Ah, Ak, Ag, Ax, An, Ass, Au, Av, and Ay, but the transitional horizons (e.g., AB, AE, AC, and BA) were not considered as A horizons. The A horizon thickness in this study was defined as the thickness of all A horizons in the soil profile which extended from the soil surface or bottom of O horizons to the bottom of A horizons. It was calculated by summing up the horizon thickness (measured from the bottom of the previous horizon to the bottom of current horizon) of all the A horizons in each pedon. The solum is the soil that has gone through notable soil forming processes and has been significantly modified from the parent material, including both surface and sub-surface horizons. The solum as defined in this study included all horizons with an O, A, E, and B horizons above C or R horizons. Transitional horizons (e.g., BC) were not included. The solum extended from the soil surface to the top of the parent material (e.g., C, R, BC, CR horizon). The Solum thickness was calculated as the accumulated thickness of all the horizons above C or R horizons in each pedon. If a pedon did not contain a designation of C or R horizon, it was considered as not reaching the parent material and the pedon was removed in the calculation of solum thickness. The pedons that had A horizon thickness greater than 100 cm ($n = 190$, 0.5% of the total sample size) or solum thickness greater than 300 cm ($n = 82$, 0.4% of the total sample size) were removed to reduce the skewness of the data.

The final dataset contained 37,712 measurements of A horizon thickness and 22,409 measurements of solum thickness obtained between 1950 and 2018 across the CONUS (Fig. 1a). The number of observations per year ranged from 2 to 1460 for A horizon thickness, and from 2 to 777 for solum thickness with more pedons sampled in the 1980s and 1990s (Fig. 1b). The A horizon and solum thickness data were slightly skewed with mean values of 21 cm and 92 cm respectively (Fig. 1c). This is an extensive dataset that is measured by experienced soil scientists on excavated pedons, spans ten soil orders, diverse landscapes, and climatic conditions of CONUS, and encompasses continuous and yearly measurements in the period when human activity has dominantly influenced the climate and environment.

Environment variables

Different soil and environmental variables that affect soil formation and processes were used to quantify the spatial distribution of A horizon and solum thickness (Supplementary Table 2 and Figs. 1, 2, 3). The variables were coupled to the pedon locations.

Each pedon was classified by soil survey staff in the field at the time of sampling. If soil classification was missing, the soil order was obtained from the dominant soil order of the mapping unit in the gSSURGO map or the STATSGO soil map. The A horizon and solum thickness dataset of CONUS contained ten soil orders: Alfisols, Andisols, Aridisols, Entisols, Histosols, Inceptisols, Mollisols, Spodosols, Ultisols, and Vertisols (Supplementary Fig. 1).

Soil climatic variables used in this study include soil temperature and moisture regimes, which were obtained from USDA-NRCS (Supplementary Fig. 1). The soil temperature regime map uses Albers Equal Area Projection NAD 1927 Clarke 1866 Spheroid projection system at a scale of 1:7,500,000 and contains six classes: cryic, frigid, hyperthermic, isomesic, mesic, and thermic, after removing the classes that have less than 10 samples. The soil moisture regime map uses the same projection system at a scale of 1:9,000,000 and contains five classes: aquic, aridic, udic, ustic, and xeric.

Climatic variables were obtained from TerraClimate which provides land surface monthly climate data and water balance since 1958 at the global scale with a 4.6-km spatial resolution⁵⁷. Twelve climatic variables were downloaded using Google Earth Engine for every month from 1958 to 2018, including precipitation accumulation (pr), minimum temperature (tmmn), maximum temperature (tmmx), Palmer Drought Severity Index (pdsi), vapor pressure deficit (vpd), wind-speed at 10 m (vs), runoff (ro), soil moisture (sm), actual evapotranspiration (aet), reference evapotranspiration (pet), climate water deficit (def), downward surface shortwave radiation (srad), and snow water equivalent (swe) (Supplementary Fig. 2). The surface water balance datasets (e.g., aet, def, ro, sm, and swe) from TerraClimate were derived using a one-dimensional soil water balance model. The annual means of pdsi, tmmn, tmmx, sm, srad, vpd, and vs and annual sum of aet, def, pet, pr, ro, and swe were first calculated from monthly climate data for each year and then averaged for the period of 1958–2018. We compared (1) averaging the climate data for all the years (1958–2018) and (2) averaging climate data for five years to twenty years with one year increment prior to the sampling year for each pedon, and found high correlations (Pearson correlation over 0.98) for all the averaged climatic data, so we decided to average the climatic data for 1958–2018.

The land use type (250-m resolution) was obtained from USGS Conterminous United States Projected Land-Use/Land-Cover Mosaics 1938–1992⁵⁸ and USGS LandCarbon Conterminous United States Land-Use/Land-Cover Mosaics 1992–2100⁵⁹. The yearly land use type was first extracted for each pedon from 1938 to the year of sampling. If the land use type has not been changed during that period, it was considered a constant land use type, and there were eight classes of constant land use: developed, barren, cropland, forest, grassland, pasture, shrubland, and wetland. If land use had changed, the frequencies and types of changes were recorded. We identified seven land use changes: convert to grassland (changing from any land use type to grassland), convert to shrubland (changing from any land use type to shrubland), cropping (changing from any land use type to cropland), grazing (changing from any land use type to pasture), reforestation (changing from any land use type to any type of forest), urbanization (changing from any land use type to urban land use), and others (the pedons that have gone through more than two times of land use change) (Supplementary Fig. 3).

The high-resolution (10-m) seamless 3D EP DEM dataset was used to obtain the elevation map and calculate topographic variables (slope, aspect, profile curvature, and plane curvature) through the Google Earth Engine⁶⁰. The topographic variables were extracted to all the sampling locations (Supplementary Fig. 1).

The parent material of soils was obtained from the Conservation Science Partners (CSP) Ecologically Relevant Geomorphology (ERGo) datasets, which contain landforms and physiographic patterns of CONUS at a 90-m spatial resolution^{61,62}. The dataset was downloaded from Google Earth Engine and extracted to the pedon sampling locations. The dataset was

aggregated into 14 classes: alkaline intrusive, alluvium and coastal sediment, carbonate, colluvial sediment, eolian sediment, glacial lake sediment, glacial outwash, glacial till, hydric, non-carbonate, saline lake sediment, silicic residual, and volcanic (Supplementary Fig. 1).

Spatial analysis

The A horizon and solum thickness were averaged to every degree from 125° W to 67° W along the longitude and from 25° N to 49° N along the latitude and considered as longitude and latitude zonal means. The zonal means were also calculated for topographic variables (elevation, slope, aspect, profile curvature, plane curvature) and climatic variables (pr, tmmn, tmmx, pdsi, vpd, vs, ro, sm, aet, def, pet, srad, swe). Pearson correlation coefficients between the zonal means of A horizon and solum thickness and corresponding topographic variables and climatic variables were calculated using the *cor* function in R version 4.1.0⁶³.

Because the assumptions of mechanistic models may not be valid at the national scale for a decadal dataset, a widely used data-driven approach was selected. Generalized additive models (GAMs) were used to explore the relationship between soil and environmental variables and A horizon and solum thickness. GAM is extended from the generalized linear model (GLM) in which linear terms are replaced by a set of smooth functions to account for nonlinear relationships⁶⁴. It increases the model's flexibility but does not sacrifice its interpretability. GAMs have been widely used in spatial-temporal modeling and mapping and understanding the controlling relationships of environmental variables⁶⁵.

The decision to use GAMs was supported by a supplementary analysis that showed how GAMs outperformed ordinary linear models, linear models with the LASSO penalty, and linear model with the ridge penalty. To do this, we took the national-level dataset and split it into training, validation, and test sets with 60%, 20%, and 20% of the dataset respectively. On the training set for linear models, all possible combinations of a large class of variables were considered. For example, if there were 5 variables used, there were $2^5 - 1 = 31$ models fit on the training set. For the linear models with regularization (LASSO⁶⁶ and Ridge⁶⁷), a granular sequence of penalty parameters was considered. As the penalty term grows for LASSO, fewer and fewer parameters are included in the model, focusing on the most important variables. For Ridge, the absolute value of the regression coefficient gets smaller and smaller for the variables. The ridge and LASSO methods were both implemented using the *glmnet* package⁶⁸ in R.

All the models fit on the training set were passed over to the validation set, and the best model from each class (one from linear models, one from ridge, and one from LASSO) was passed to the test set. After being fit on the training set, the GAM outperformed all of these models on the test set *without* the benefit of variable selection. That is to say, a GAM with all the variables fed into the other models was sent directly to the test set and *still* outperformed the best model from each of the other three classes. This shows that a GAM is the right tool for analyzing this dataset and is a reasonable choice since more “classical” statistical tools perform materially worse than GAMs.

The GAMs were developed using *bam* function with *fREML* method in the *mgcv* package⁶⁹ in R version 4.1.0⁶³. In the full model, longitude and latitude coordinates, soil order, temperature regime, moisture regime, parent material, land use, topographic variables (elevation, slope, profile curvature), and climatic variables (pr, tmmn, tmmx, pdsi, vpd, vs, ro, sm, aet, pet, def, srad, swe) were used as predictors. The results displayed high concavity (i.e., strong non-linear relationship) between longitude and latitude coordinates and all the climatic variables and many topographic variables, and therefore these variables were removed in the final model fitting. The final model (Eq. 1) included a spatial term ($s(X, Y)$) fitted with a thin plate spline with a basis size of 800, categorical variables (soil order, temperature regime, moisture regime, parent material, land use), and a numeric term ($s(slope)$). The A horizon and solum thickness were log-transformed in

the model fitting.

$$\begin{aligned} \log(\text{thickness}) \sim & s(X, Y, bs = \text{tp}, k = 800) + \text{Order} \\ & + \text{Temperature.regime} + \text{Moisture.regime} \quad (1) \\ & + \text{Parent.material} + \text{Land.use} + s(\text{slope}) \end{aligned}$$

The datasets of A horizon and solum thickness were randomly split into 70% for model training ($n = 26,398$ for A horizon and $n = 15,686$ for solum) and 30% for model testing ($n = 11,314$ for A horizon and $n = 6,723$ for solum)⁷⁰. The spatial and temporal coverage and the distribution of soil thickness in the training and testing datasets were visually compared to ensure the representativeness of these two datasets (Supplementary Fig. 4). The coefficient of determination (R^2), Lin's concordance correlation coefficient (ρ_C), mean error (ME), and root mean squared error (RMSE) were calculated using predicted and measured values for both the training and testing datasets and using log-transformed and original scales (Supplementary Figs. 5 and 6).

To understand different environmental controlling factors on soil thickness at regional scale, we developed regional GAMs across the CONUS. The Land Resource Region (LRR) segments the US into 20 regions across the CONUS and 8 regions in Alaska, Hawaii, the Caribbeans, and the Pacific Basin Islands⁷¹. It synthesized the knowledge of climate, geology, physiography, soil, water resources, biological resources, and land use, and was primarily designed for soil and water conservation⁷². So, we selected LRR for regional segmentation, and GAMs were developed for 20 LRR regions across the CONUS. Equation 1 was used to fit the GAM in each region, but if the slope was non-significant in a specific region, it was removed in the final model to reduce the model complexity. The basis size of the spatial term was reduced from 800 to 100, 200, 300, 400 (whichever resulted in higher deviance explained) to reduce the model complexity (Supplementary Fig. 7 and Table 3).

Temporal analysis

As the soil thickness data were not repeatedly and continuously collected at the same locations across the time, we used spatial data which were collected at different times to infer the temporal trend. However, the use of spatial data to infer the temporal trend is not flawless. First, other spatially varying environmental variables also affect the soil thickness besides time. Second, the samples were not evenly collected over time, and the sample size varied in different land uses with time (Supplementary Fig. 8). For example, sampling was primarily focused on cropland from 1950 to 1990, while since 1990 more samples were collected from forest and other less managed systems (Supplementary Fig. 8). To remove such confounding factors and solely investigate the effect of time, we selected chronosequences in each LRR by controlling other environmental factors consistent.

We used the following criteria to select chronosequences: (1) within each LRR, we split the dataset into subsets that have the same soil order, soil temperature regime, soil moisture regime, parent material, land use, and similar topographic features (elevation, slope, and aspect). To select samples with similar elevation, we first calculated the median value of elevation within this subset and then kept the samples that are within 250-m elevation difference from the median value. To select samples with similar slopes, we first calculated the median value of slope within this subset and then kept the samples that are within 7.5° difference from the median slope. For aspect, we split the subset into two categories – north facing (0–90° and 270–360°) and south facing (90–270°). (2) The subsets that had fewer than 50 samples were removed. (3) In the subset, there should be at least one sample available in each decade. The unsatisfied subset was removed. This was to reduce the uncertainty caused by missing temporal data. In total, 49 chronosequences were selected from 9 LRRs for A horizon thickness with sample size ranging from 51 to 579 in each chronosequence (Supplementary Fig. 9). For solum thickness, 9 chronosequences were selected from 2 LRRs with a sample size ranging from 51 to 219. Although the soil thickness data were not continuous and repeated measurements with time, the selection of

chronosequences largely reduced the effects of other spatial and environmental factors on the temporal analysis.

For each chronosequence, to reduce the effect of uneven sample size with time, we used sampling with replacement to randomly select one sample per decade and this procedure was repeated 100 times. This resulted in 100 sets of time-series data for each chronosequence. We developed simple linear regression models for each time series dataset (Eq. 2). The coefficient b (cm yr^{-1}) was used to evaluate the rate of thickness change with time. The median, 5%, and 95% quantiles of the coefficient b were calculated from the 100 linear models for each chronosequence and considered as the mean temporal change rates and its 90% confidence interval. The median and 90% of the R^2 and p value of the linear models were also calculated (Supplementary Tables 4 and 5). The simple linear regression was conducted using *lm* function in R version 4.1.0⁶³.

$$\text{Thickness} \sim a + b \times \text{Time} \quad (2)$$

Data availability

The data used for data analysis in this study are openly available in Zenodo at <https://doi.org/10.5281/zenodo.10655823>.

Code availability

The R code used for data analysis and generating main figures in this study are openly available in Zenodo at <https://doi.org/10.5281/zenodo.10655823>.

Received: 8 October 2023; Accepted: 1 March 2024;

Published online: 15 March 2024

References

1. Fu, Z. et al. Soil thickness effect on hydrological and erosion characteristics under sloping lands: a hydrogeological perspective. *Geoderma* **167**, 41–53 (2011).
2. Yang, J., Blanchar, R. W., Hammer, R. D. & Thompson, A. L. Soybean growth and rhizosphere pH as influenced by A Horizon thickness. *Soil Sc. Soc. Am. J.* **60**, 1901–1907 (1996).
3. Zhang, Y., Desai, A. R., Xiao, J. & Hartemink, A. E. Deeper topsoils enhance ecosystem productivity and climate resilience in arid regions, but not in humid regions. *Glob. Change Biol.* **00**, 1–18 (2023).
4. Adhikari, K. & Hartemink, A. E. Soil organic carbon increases under intensive agriculture in the Central Sands, Wisconsin, USA. *Geoderma Region.* **10**, 115–125 (2017).
5. Berhe, A. A., Barnes, R. T., Six, J. & Marin-Spiotta, E. Role of soil erosion in biogeochemical cycling of essential elements: carbon, nitrogen, and phosphorus. *Annu. Rev. Earth Planet. Sci.* **46**, 521–548 (2018).
6. Brown, L. R. The global loss of topsoil. *J. Soil Water Conserv.* **39**, 162–165 (1984).
7. Heimsath, A. M., Dietrich, W. E., Nishiizumi, K. & Finkel, R. C. The soil production function and landscape equilibrium. *Nature* **388**, 358–361 (1997).
8. Jenny, H. *Factors of soil formation: a system of quantitative pedology*, (McGraw-Hill, 1941).
9. Toy, T. J., Foster, G. R. & Renard, K. G. *Soil erosion: processes, prediction, measurement, and control*, (John Wiley & Sons, 2002).
10. Van Oost, K., Govers, G., De Alba, S. & Quine, T. Tillage erosion: a review of controlling factors and implications for soil quality. *Prog. Phys. Geogr.* **30**, 443–466 (2006).
11. Shangguan, W., Hengl, T., Mendes de Jesus, J., Yuan, H. & Dai, Y. Mapping the global depth to bedrock for land surface modeling. *J. Adv. Model. Earth Syst.* **9**, 65–88 (2017).
12. Pelletier, J. D. et al. A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling. *J. Adv. Model. Earth Syst.* **8**, 41–65 (2016).

13. Wilford, J., Searle, R., Thomas, M., Pagendam, D. & Grundy, M. A regolith depth map of the Australian continent. *Geoderma* **266**, 1–13 (2016).
14. Pelletier, J. D. & Rasmussen, C. Geomorphically based predictive mapping of soil thickness in upland watersheds. *Water Resour. Res.* **45**, W09417 (2009).
15. Phillips, J. D. The convenient fiction of steady-state soil thickness. *Geoderma* **156**, 389–398 (2010).
16. Chen, S. et al. Probability mapping of soil thickness by random survival forest at a national scale. *Geoderma* **344**, 184–194 (2019).
17. Vanwallegem, T., Poesen, J., McBratney, A. & Deckers, J. Spatial variability of soil horizon depth in natural loess-derived soils. *Geoderma* **157**, 37–45 (2010).
18. Lacoste, M., Mulder, V. L., Richer-de-Forges, A. C., Martin, M. P. & Arrouays, D. Evaluating large-extent spatial modeling approaches: a case study for soil depth for France. *Geoderma Region.* **7**, 137–152 (2016).
19. Chamberlin, T. C. Soil wastage. Proc. Conf. of Governors in the White House. US Congress 60th, 2nd Session, House Doc. 75–83 (1909).
20. Stockmann, U., Minasny, B. & McBratney, A. B. How fast does soil grow? *Geoderma* **216**, 48–61 (2014).
21. Larsen, I. J. et al. Rapid soil production and weathering in the Southern Alps, New Zealand. *Science* **343**, 637–640 (2014).
22. Hall, G., Daniels, R. & Foss, J. Rate of soil formation and renewal in the USA. *Determinants Soil Loss Toler.* **45**, 23–39 (1982).
23. Dudal, R. The sixth factor of soil formation. *Eurasian Soil Sci. C/C of Pochvovedenie* **38**, S60 (2005).
24. Minasny, B., Finke, P., Stockmann, U., Vanwallegem, T. & McBratney, A. B. Resolving the integral connection between pedogenesis and landscape evolution. *Earth Sci. Rev.* **150**, 102–120 (2015).
25. Montgomery, D. R. Soil erosion and agricultural sustainability. *Proc. Natl. Acad. Sci.* **104**, 13268–13272 (2007).
26. U.S. Department of Agriculture. Summary Report: 2007 National Resources Inventory, Natural Resources Conservation Service, Washington, DC, and Center for Survey Statistics and Methodology, Iowa State University, Ames, Iowa. 123 pages (2009). http://www.nrcs.usda.gov/technical/NRI/2007/2007_NRI_Summary.pdf.
27. Thaler, E. A., Larsen, I. J. & Yu, Q. The extent of soil loss across the US Corn Belt. *Proc. Natl. Acad. Sci.* **118**, e1922375118 (2021).
28. Thaler, E. A., Kwang, J. S., Quirk, B. J., Quarrier, C. L. & Larsen, I. J. Rates of historical anthropogenic soil erosion in the Midwestern United States. *Earth's Future* **10**, e2021EF002396 (2022).
29. Schneider, F., Don, A., Hennings, I., Schmittmann, O. & Seidel, S. J. The effect of deep tillage on crop yield – What do we really know? *Soil Tillage Res.* **174**, 193–204 (2017).
30. Tsai, C. C., Chen, Z. S., Duh, C. T. & Horng, F. W. Prediction of soil depth using a soil-landscape regression model: a case study on forest soils in southern Taiwan. *Proc. Natl. Sci. Council, Republic China Part B, Life Sci.* **25**, 34–39 (2001).
31. dos Santos, J. C. B. et al. Saprolithology applied to pedology: Integrated study of soil and saprolite derived from crystalline rocks to better understand properties of whole regoliths along a climate gradient (NE Brazil). *Geoderma* **409**, 115602 (2022).
32. Koop, A. N., Hirmas, D. R., Sullivan, P. L. & Mohammed, A. K. A generalizable index of soil development. *Geoderma* **360**, 113898 (2020).
33. Cleveland, C. C., Reed, S. C. & Townsend, A. R. Nutrient regulation of organic matter decomposition in a tropical rain forest. *Ecology* **87**, 492–503 (2006).
34. da Silva, Y. J. A. B. et al. Weathering rates and carbon storage along a climosequence of soils developed from contrasting granites in northeast Brazil. *Geoderma* **284**, 1–12 (2016).
35. Pennock, D. J. & Jong, E. D. The influence of slope curvature on soil erosion and deposition in hummock terrain. *Soil Science* **144**, 209–217 (1987).
36. Dahlgren, R. A., Boettinger, J. L., Huntington, G. L. & Amundson, R. G. Soil development along an elevational transect in the western Sierra Nevada, California. *Geoderma* **78**, 207–236 (1997).
37. Gray, J. & Murphy, B. Parent material and world soil distribution. 17th World Congress of Soil Science, Bangkok, Thailand, 1–14 (2002).
38. Karlstrom, E. T. Use of soils to identify glacial deposits of various ages east of Glacier National Park, Montana, U.S.A. *Arct. Antarct. Alpine Res.* **32**, 179–188 (2000).
39. Zebarth, B., Rees, H., Walsh, J., Chow, L. & Pennock, D. Soil variation within a hummocky podzolic landscape under intensive potato production. *Geoderma* **110**, 19–33 (2002).
40. Jobbágy, E. G. & Jackson, R. B. The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecol. Appl.* **10**, 423–436 (2000).
41. Reganold, J. P., Elliott, L. F. & Unger, Y. L. Long-term effects of organic and conventional farming on soil erosion. *Nature* **330**, 370–372 (1987).
42. Acksel, A., Kappenberg, A., Kühn, P. & Leinweber, P. Human activity formed deep, dark topsoils around the Baltic Sea. *Geoderma Region.* **10**, 93–101 (2017).
43. Schaeztl, R. J. & Follmer, L. R. Longevity of treethrow microtopography: implications for mass wasting. *Geomorphology* **3**, 113–123 (1990).
44. Shouse, M. & Phillips, J. Soil deepening by trees and the effects of parent material. *Geomorphology* **269**, 1–7 (2016).
45. Li, Z. & Fang, H. Impacts of climate change on water erosion: a review. *Earth-Sci. Rev.* **163**, 94–117 (2016).
46. Steiner, F. Soil conservation policy in the United States. *Environ. Manag.* **11**, 209–223 (1987).
47. Allmaras, R. & Dowdy, R. Conservation tillage systems and their adoption in the United States. *Soil Tillage Res.* **5**, 197–222 (1985).
48. Claassen, R., Bowman, M., McFadden, J., Smith, D. & Wallander, S. Tillage Intensity and Conservation Cropping in the United States, EIB-197, U.S. Department of Agriculture, Economic Research Service (2018).
49. Uri, N. D. & Lewis, J. A. The dynamics of soil erosion in US agriculture. *Sci. Total Environ.* **218**, 45–58 (1998).
50. Nearing, M., Pruski, F. F. & O'Neal, M. R. Expected climate change impacts on soil erosion rates: a review. *J. Soil Water Conserv.* **59**, 43–50 (2004).
51. Segura, C., Sun, G., McNulty, S. & Zhang, Y. Potential impacts of climate change on soil erosion vulnerability across the conterminous United States. *J. Soil Water Conserv.* **69**, 171–181 (2014).
52. Robinson, D. A. et al. Analytical modelling of soil porosity and bulk density across the soil organic matter and land-use continuum. *Sci. Rep.* **12**, 7085 (2022).
53. Mohammed, A. K., Hirmas, D. R., Nemes, A. & Giménez, D. Exogenous and endogenous controls on the development of soil structure. *Geoderma* **357**, 113945 (2020).
54. National Cooperative Soil Survey, National Cooperative Soil Survey Soil Characterization Database. accessed 10 Nov (2020); <http://ncsslabsdatamart.sc.egov.usda.gov/>.
55. Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database. Available online. accessed 01 Nov 2023.
56. Schoeneberger, P. J., Wysocki, D. A., Benham, E. C. *Field book for describing and sampling soils, version 3.0*: USDA Natural Resources Conservation Service, National Soil Survey Center, Lincoln, NE. (2012).
57. Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A. & Hegewisch, K. C. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Sci. Data* **5**, 1–12 (2018).

58. Sohl, T. et al. Modeled historical land use and land cover for the conterminous United States. *J. Land Use Sci.* **11**, 476–499 (2016).
59. Sohl, T. L. et al. Spatially explicit modeling of 1992–2100 land cover and forest stand age for the conterminous United States. *Ecol. Appl.* **24**, 1015–1036 (2014).
60. United States Geological Survey (USGS), 3D Elevation Program 10-Meter Resolution Digital Elevation Model.
61. Soller, D. R., Reheis, M. C., Garrity, C. P. & Van Sistine, D. R. Map database for surficial materials in the conterminous United States. *US Geological Survey Data Series* **425**, (2009).
62. Theobald, D. M., Harrison-Atlas, D., Monahan, W. B. & Albano, C. M. Ecologically-relevant maps of landforms and physiographic diversity for climate adaptation planning. *PLoS one* **10**, e0143619 (2015).
63. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2021). <https://www.R-project.org/>.
64. Hastie, T. J. & Tibshirani, R. J. *Generalized additive models*, (Routledge, 2017).
65. Yanosky, J. D. et al. Spatio-temporal modeling of particulate air pollution in the conterminous United States using geographic and meteorological predictors. *Environ. Health* **13**, 1–15 (2014).
66. Tibshirani, R. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. Series B* **58**, 267–288 (1996).
67. Hoerl, A. E. & Kennard, R. W. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* **12**, 55–67 (1970).
68. Friedman, J. H., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Software* **33**, 1–22 (2010).
69. Wood, S. N. *Generalized additive models: an introduction with R*, (Chapman and Hall/CRC, 2006).
70. Wadoux, A. M. J.-C., Heuvelink, G. B., De Bruin, S. & Brus, D. J. Spatial cross-validation is not the right way to evaluate map accuracy. *Ecol. Modell.* **457**, 109692 (2021).
71. United States Department of Agriculture-Natural Resources Conservation Service. Land resource regions and major land resource areas of the United States, the Caribbean, and the Pacific Basin. U.S. Department of Agriculture, Agriculture Handbook 296 (2022). <https://www.nrcs.usda.gov/resources/data-and-reports/major-land-resource-area-mlra>.
72. Austin, M. E. Land resource regions and major land resource areas of the United States. U.S. Department of Agriculture Handbook 296 (1965).

Acknowledgements

We are grateful for Marvin Beatty for creating a professorship and providing us funding to do this research. We also thank the financial support from Department of Crop and Soil Science, Oregon State University. We are grateful to USDA-NRCS for collecting and maintaining the large and precious open-access soil data. We also want to thank Rich Ferguson, Steve

Monteith, Scarlett Murphy for answering us any questions we asked about the dataset. We would like to thank Paul Reich for sending us soil temperature regime and soil moisture regime maps. Additionally, we thank Maria Kamenetsky for suggesting us to use GAMs for spatial-temporal analysis and quantifying effects of environmental driving factors and answering any questions about GAMs during model development.

Author contributions

Y.Z. and A.E.H. conceived the work. Y.Z. performed the data analysis, interpreted the results, and wrote the draft. A.E.H., T.V., B.R.B., and S.M. contributed to the interpretation of the results and writing.

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/s43247-024-01299-6>.

Correspondence and requests for materials should be addressed to Yakun Zhang.

Peer review information *Communications Earth & Environment* thanks Gbadamassi G.O. Dossa and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Aliénor Lavergne. Peer reviewer reports are 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 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

© The Author(s) 2024