Key predictors of soil organic matter vulnerability to mineralization differ with depth at a continental scale

Tyler L. Weiglein · Brian D. Strahm · Maggie M. Bowman · Adrian C. Gallo ·
Jeff A. Hatten · Katherine A. Heckman · Lauren M. Matosziuk ·
Lucas E. Nave · Angela R. Possinger · Michael D. SanClements ·
Christopher W. Swanston

Abstract Soil organic matter (SOM) is the largest terrestrial pool of organic carbon, and potential carbon-climate feedbacks involving SOM decomposition could exacerbate anthropogenic climate change. However, our understanding of the controls on SOM mineralization is still incomplete, and as such, our ability to predict carbon-climate feedbacks is limited. To improve our understanding of controls on SOM decomposition, A and upper B horizon soil samples from 26 National Ecological Observatory Network (NEON) sites spanning the conterminous U.S. were incubated for 52 weeks under conditions representing site-specific mean summer temperature and sample-specific field capacity (-33 kPa) water potential. Cumulative carbon dioxide respired was periodically measured and normalized by soil organic C content to calculate cumulative specific respiration (CSR), a metric of SOM vulnerability to mineralization. The Boruta algorithm, a feature selection algorithm, was used to select important predictors of CSR from 159 variables. A diverse suite of predictors was selected (12 for A horizons, 7 for B horizons) with predictors falling into three categories corresponding to SOM chemistry, reactive Fe and Al phases, and site moisture availability. The relationship between SOM chemistry predictors and CSR was complex, while sites that had greater concentrations of reactive Fe and Al phases or were wetter had lower CSR. Only three predictors were selected for both horizon types, suggesting dominant controls on SOM decomposition differ by horizon. Our findings contribute to the emerging consensus that a broad array of controls regulates SOM decomposition at large scales and highlight the need to consider changing controls with depth.

T. L. Weiglein · B. D. Strahm (✉) · A. R. Possinger
Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg,
VA, USA
e-mail: brian.strahm@vt.edu

M. M. Bowman
Environmental Studies Program, University of Colorado Boulder, Boulder, CO, USA

M. M. Bowman · M. D. SanClements
Institute of Arctic and Alpine Research (INSTAAR), University of Colorado Boulder, Boulder,
CO, USA

A. C. Gallo · J. A. Hatten · L. M. Matosziuk
Department of Forest Engineering, Resources & Management, Oregon State University, Corvallis,
OR, USA

Introduction

Soils contain the largest actively cycled pool of terrestrial organic carbon (C) in the form of soil organic matter (SOM). Estimates of soil organic C (SOC) stocks in the upper 1 m of soil range from 1325 to 1502 Pg C, while SOC stocks to a depth of 3 m or greater are estimated to range from 2344 Pg C to approximately 3000 Pg C (Jobbagy and Jackson 2000; Tarnocai et al. 2009; Köchy et al. 2015). Despite the size of this C reservoir and its potential to contribute to an amplifying carbon-climate feedback, we still lack a thorough understanding of how SOM will respond to global change drivers (Bradford et al. 2016). This knowledge gap limits society's ability to accurately predict the effectiveness of climate change mitigation strategies.

SOM consists of a complex mixture of plant residues and microbial products in varying states of degradation. Historically, it was believed that stable SOM, which has been variably defined as SOC that remains in the soil for decades to millennia, was formed from either inherently recalcitrant plant residues or chemically complex humic substances (Schmidt et al. 2011; Lehmann and Kleber 2015). However, the paradigm that has emerged over the past decade places less emphasis on the chemical properties of SOM in mineral soils. Instead, it posits that soil physicochemical properties (e.g., pH, abundance of reactive metals and divalent cations) and other ecosystem properties (e.g., vegetation, climate) are the dominant factors controlling SOM stabilization (Rasmussen et al. 2018; Kramer and Chadwick 2018; von Fromm et al. 2021; Yu et al. 2021). With this

paradigm shift, a broad suite of stabilization mechanisms has been identified, including: (1) occlusion within aggregates, (2) mineral-organic matter interactions, (3) formation of pyrogenic C, (4) biotic suppression due to climatic factors (e.g., xeric conditions, extreme temperatures), and (5) biotic suppression due to local environmental conditions (e.g., low O_2 levels) (von Lützow et al. 2006; Schmidt et al. 2011; Rasmussen et al. 2018; Wiesmeier et al. 2019). Additionally, the critical role of microbes as agents of both SOM formation (Kallenbach et al. 2016) and mineralization (Hagerty et al. 2014; Creamer et al. 2015) has gained increased attention. However, the dynamic nature of microbial communities presents a major challenge to incorporating microbial community characteristics into analyses of controls on SOM dynamics.

Given the broad suite of factors controlling SOM stabilization, it has been recognized that there is widespread spatial variability both vertically and horizontally in the dominant mechanism(s) controlling SOM stability (Wiesmeier et al. 2019; Viscarra Rossel et al. 2019). Several recent large-scale studies, however, have focused on the role of mineral-organic matter interactions, with a specific emphasis on the role of divalent cations (Ca^{2+} and Mg^{2+}) and reactive (Fe) and aluminum (Al) (Rasmussen et al. 2018; Kramer and Chadwick 2018; von Fromm et al. 2021; Yu et al. 2021). This reflects an implicit assumption that mineral-organic matter interactions are the primary stabilization mechanism in mineral soils across different ecosystems. Likewise, it is assumed that reactive Fe and Al play a more important role in stabilizing SOM with increasing depth (von Lützow et al. 2006; Kögel-Knabner et al. 2008; Rumpel and Kögel-Knabner 2011), although recent findings challenge this assumption (Yu et al. 2021).

Most recent large-scale studies of SOM have focused on measuring SOC concentration (Rasmussen et al. 2018; von Fromm et al. 2021; Quesada et al. 2020; Yu et al. 2021), SOC stocks (Nave et al. 2021), or certain SOC pools (Kramer and Chadwick 2018; Viscarra Rossel et al. 2019). Although these measures provide valuable insights into soil C cycling, they provide an incomplete picture of SOM dynamics. For example, recent work using radiocarbon (^{14}C) suggests that controls on C persistence may differ from controls on SOC concentration (Heckman et al. 2021). As such, it is necessary to use a complementary suite

K. A. Heckman · C. W. Swanston
Northern Research Station, USDA Forest Service, Houghton, MI, USA

L. E. Nave
Biological Station, University of Michigan, Pellston, MI, USA

L. E. Nave
Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA

M. D. SanClements
Battelle, National Ecological Observatory Network (NEON), Boulder, CO, USA

of measures to gain a comprehensive understanding of SOM dynamics. One such measure lacking from most recent large-scale studies of SOM (but see Doetterl et al. 2015) is the decomposability of SOM as measured using laboratory incubations. Incubations have a number of acknowledged caveats due to artificial temperature and moisture regimes, disturbed soil structure, and the exclusion of plant inputs (Torn et al. 2009). However, when viewed as a measure of potential SOM vulnerability to mineralization as opposed to a representation of in situ SOM dynamics, incubations are a valuable tool for comparative studies of SOM. As such, there is an extensive literature in which they have been used to estimate a number of SOM properties, such as mean residence time and stability (Whalen et al. 2000; Swanston et al. 2002; Torn et al. 2005; Paul et al. 2006).

For this study, we conducted a 52-week laboratory incubation using a continental-scale soil sample set from the National Ecological Observatory Network (NEON) to address two questions. First, what are the key predictors of SOM vulnerability to mineralization, which we define as the percentage of initial SOC respired, across the conterminous U.S.? Second, how do these key predictors, and thus dominant controls, differ between A and B horizons? Given that our samples were from upland sites with mineral soils, we hypothesized that mineral-organic matter interactions would be the dominant control in both A and B horizons as indicated by the selection of proxies of reactive Fe and Al and exchangeable Ca^{2+} and Mg^{2+}.

Materials and methods

Sampling sites

Between 2015 and 2017, soil samples were taken from 26 NEON sites across the conterminous United States spanning continental-scale gradients of climate, soil types, and vegetation cover (Fig. 1; Table 1). At each site except WREF (see Table 1 for Site IDs), 10 4.5-cm diameter soil cores were taken up to a depth of 2 m depending on depth to bedrock (two cores from each of the five soil array plots in the footprint of the site's micrometeorological tower) via Giddings probe

(Giddings Machine Company, Windsor, CO, USA). At WREF, 10 3.45-cm diameter soil cores were taken to a depth of 0.5 m following the same approach via handheld AMS corer (AMS Inc., American Falls, ID, USA). All cores were cooled with ice packs and shipped in coolers to Oregon State University for processing.

Sample processing

Of the 10 cores collected per site, five were frozen and archived for future use, while the remaining five were refrigerated at 3 °C until processing, typically within 2 weeks of arrival. After cores' liners were opened longitudinally, the cores were described according to U.S. Department of Agriculture Natural Resources Conservation Service (USDA NRCS) protocols (Schoeneberger et al. 2012). As part of the description process, sites that may have had carbonates present were identified based on NEON megapit or USDA NRCS Web Soil Survey soil descriptions. For this subset of sites, carbonates were tested for by applying 10% v/v hydrochloric acid (HCl) to a subsample of each horizon from each core. Each core was then split by genetic horizon. Splits corresponding to the same genetic horizon from different cores were then composited, air-dried, and passed through a 2 mm sieve to remove rock fragments and any coarse particulate matter. Subsamples of A and upper B horizons (henceforth simply referred to as B horizons) were split using a riffle box (Model CL-244; Soiltest, Inc., Evanston, IL, USA) to ensure homogeneous samples and then sent to Virginia Tech for a 52-week laboratory incubation.

Laboratory incubation

For the laboratory incubation, 20 g of air-dry soil were placed in 118 mL polypropylene specimen cups with one replicate per site-horizon type combination and a total of 26 samples per horizon type. Samples were tamped down as needed until approximately representative of bulk densities in the field (1.0 g cm^{-3} for A horizons and 1.2 g cm^{-3} for B horizons). Samples were then placed into quart (946 mL) mason jars with lids fitted with septa. A small amount (approximately
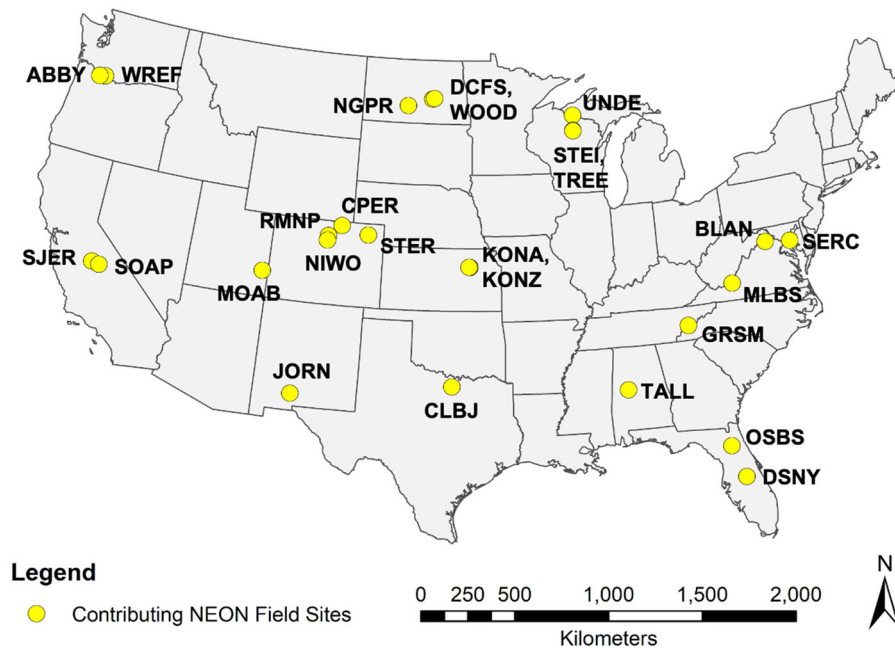
Fig. 1 Location of 26 NEON sites from which soil samples were taken. Note that the symbols for the following pairs of sites (nearly) overlap: KONA and KONZ, STEI and TREE, DCFS and WOOD (see Table 1 for full site names)

10 mL) of water was added to the bottom of the jars to maintain humidity in the jar headspace and minimize water loss via evaporation from the sample. Samples were brought to field capacity (-33 kPa) by adding an amount of DI water estimated using a pedotransfer function developed by Saxton et al. (1986) and site-specific soil texture data. Soil texture data were obtained for all sites from NEON megapit data (NEON 2019) with the exception of SJER and SOAP, for which soil texture data were obtained from the USDA NRCS Web Soil Survey (Soil Survey Staff 2017). Given the wide range in site temperatures (Table 1) and the effect of temperature on microbial community structure and function, sites were binned by mean summer temperature (PRISM Climate Group 2018), and samples were placed in a 566 L incubator (VWR Model 10753-894; VWR International, LLC, Radnor, PA, USA) set at the corresponding group mean summer temperature.

Over the course of the 52-week incubation, jar headspace samples were taken periodically (generally between every 1 to 28 days, depending on respiration rate, to avoid exceeding headspace CO_2 concentrations of 10,000 ppm) using a syringe and injected in evacuated 10 mL gas chromatograph (GC) vials with 20 mm butyl septa and aluminum crimp seals (Catalog No. 20-1100, 20-0020, and 20-0000AS, respectively; DWK Life Sciences, LLC, Millville, NJ, USA). After sampling, jars were opened to allow headspace gas to re-equilibrate with the atmosphere, and any water lost from the samples via evaporation as determined by weighing the samples was replenished by adding DI water. Gas samples for initial ambient CO_2 concentration were collected prior to resealing the jars and placing them back in their respective incubator. Headspace samples were analyzed for CO_2 concentration using a GC outfitted with a methanizer and flame ionization detector (FID) (range 370–10,000 ppm CO_2, relative standard deviation (RSD) $< 5\%$; GC-2010; Shimadzu Scientific Instruments, Columbia, MD, USA). Four CO_2 standards ranging from ~ 370 to $\sim 10,000$ ppm CO_2 were used to generate a calibration curve for each GC run. Any decrease in sample concentration due to leakage during storage was corrected for using a sample loss curve based on measurements of standards after varying lengths of storage.

Mass of CO_2-C respired was calculated from the CO_2 concentration data according to the ideal gas law. For samples with substantial amounts of carbonates (as indicated by a "k" horizon suffix), an isotope mixing model was used to partition CO_2 produced

Table 1 Site-specific location, soil order, and climate data for samples from 26 NEON sites

Site ID	Site name	Latitude	Longitude	Elevation (m)	Dominant soil order	MAT (°C)	MAP (mm)	MAP − E_{ref} (mm)	MST (°C)	Incubation Temp. (°C)
ABBY	Abby Road	45.762	−122.330	365	Inceptisol	9.4	2657	1859	17.2	17.5
BLAN	Blandy Experimental Farm	39.060	−78.072	183	Alfisol	12.2	954	−112	23.0	22.5
CLBJ	LBJ National Grassland	33.401	−97.570	272	Alfisol	17.7	870	−524	27.5	27.5
CPER	Central Plains Experimental Range	40.816	−104.746	1653	Mollisol	8.2	348	−626	20.0	20.0
DCFS	Dakota Coteau Field School	47.162	−99.107	574	Mollisol	4.3	427	−303	19.4	20.0
DSNY	Disney Wilderness Preserve	28.125	−81.436	21	Spodosol	22.2	1282	−269	27.7	27.5
GRSM	Great Smoky Mountains National Park, Twin Creeks	35.689	−83.502	576	Inceptisol	12.1	1471	307	21.7	22.5
JORN	Jornada LTER	32.591	−106.843	1321	Aridisol	15.3	266	−1270	25.5	25.0
KONA	Konza Prairie Biological Station—Relocatable (Agriculture)	39.110	−96.613	324	Mollisol	12.5	854	−181	25.0	25.0
KONZ	Konza Prairie Biological Station—Core	39.101	−96.563	415	Mollisol	12.3	863	−157	24.4	25.0
MLBS	Mountain Lake Biological Station	37.378	−80.525	1170	Entisol	8.9	1174	309	18.4	17.5
MOAB	Moab	38.248	−109.388	1803	Aridisol	9.4	301	−755	21.9	22.5
NGPR	Northern Great Plains Research Laboratory	46.770	−100.915	588	Mollisol	5.0	406	−341	19.9	20.0
NIWO	Niwot Ridge Mountain Research Station	40.054	−105.582	3483	Mollisol	0.1	866	339	9.2	10.0
OSBS	Ordway-Swisher Biological Station	29.689	−81.993	46	Entisol	20.5	1368	−192	27.2	27.5
RMNP	Rocky Mountain National Park	40.276	−105.546	2744	Inceptisol	2.4	704	55	13.7	15.0
SERC	Smithsonian Environmental Research Center	38.890	−76.560	10	Ultisol	13.5	1076	34	24.2	25.0
STEI	Steigerwaldt Land Services	45.509	−89.586	476	Spodosol	4.7	806	104	17.9	17.5
SJER	San Joaquin Experimental Range	37.109	−119.732	399	Alfisol	16.1	469	−984	25.6	25.0
SOAP	Soaproot Saddle	37.033	−119.262	1210	Alfisol	12.7	767	−408	22.3	22.5
STER	North Sterling	40.462	−103.029	1365	Mollisol	9.1	430	−561	21.6	22.5
TALL	Talladega National Forest	32.950	−87.393	163	Ultisol	16.9	1431	70	26.2	25.0
TREE	Treehaven	45.494	−89.586	466	Spodosol	4.7	807	105	17.9	17.5
UNDE	UNDERC	46.234	−89.537	520	Spodosol	4.3	815	130	17.4	17.5
WOOD	Woodworth	47.128	−99.241	590	Mollisol	4.3	420	−312	19.5	20.0
WREF	Wind River Experimental Forest	45.820	−121.952	368	Andisol	8.9	2468	1630	17.5	17.5

MAT mean annual temperature, *MAP* mean annual precipitation, E_{ref} Hargreaves reference evaporation, *MST* mean summer temperature

from heterotrophic respiration from CO_2 produced by the dissolution of carbonates using gas and solid phase $\delta^{13}C$ measurements following the methods of Tamir et al. (2011). Solid phase $\delta^{13}C$ as well as percent C and N were measured using an EA-IRMS (IsoPrime 100 EA-IRMS; Isoprime Ltd, Cheadle, UK) with samples containing substantial amounts of carbonates (those having a "k" horizon suffix) being run both with and without acidification by HCl fumigation prior to analysis (Harris et al. 2001). Gas phase $\delta^{13}C$ ($^{13}CO_2$) was measured by taking headspace samples at four intervals throughout the incubation for horizons with substantial amounts of carbonates. These samples were analyzed at the Cornell Isotope Laboratory using an IRMS (Delta V Advantage IRMS; ThermoFisher Scientific, Waltham, MA, USA). Cumulative specific respiration (CSR) was then calculated by summing the mass of CO_2-C respired and normalizing by the initial mass of SOC in the sample.

Predictor variables

To evaluate a broad suite of potential predictor variables, 159 variables (Supplementary Table S1) were compiled for use in the statistical analysis.

Site-level data

Site-level data, in addition to incubation temperature, included estimates of net primary production (NPP) from MODIS data (Running et al. 2015; ORNL DAAC 2018) and various climate-related data, such as measures of site temperature, precipitation, and moisture availability (Supplementary Table S1) (Wang et al. 2012). Specific site moisture availability predictors used include MAP − Hargreaves reference evaporation (Hargreaves and Samani 1982; Wang et al. 2012), annual heat:moisture index (calculated as (mean annual temperature (°C) + 10)/(mean annual precipitation (mm)/1000)), summer heat:moisture index (calculated as mean warmest month temperature (°C)/(mean summer precipitation (mm)/1000)), and Hargreaves climatic moisture deficit, which is a cumulative measure of the monthly difference between reference evaporation and precipitation.

Sample density fractionation

The analyses described in following sections were conducted on air-dried, sieved bulk samples from the composited horizons, and a subset of these analyses were also conducted on density fractions (refer to Supplementary Table S1 for specific data for each fraction used in the statistical analysis). The three density fractions isolated in our sample set were the free light fraction, occluded fraction, and heavy fraction. These fractions were separated using sodium polytungstate solution adjusted to a density of 1.65 g cm^{-3} and sonication applied at 750 J mL^{-1} following standard procedures (Strickland and Sollins 1987; Golchin et al. 1994; Swanston et al. 2005).

Physical and chemical characterization

Bulk soil samples were characterized physically and chemically. Physical characteristics measured include midpoint depth of composite horizon, bulk density, fine and coarse fraction content, and soil texture as determined by the laser diffraction particle size method used by Yang et al. (2019) (CILAS 1190; CPS US, Inc., Fitchburg, WI, USA). The laser diffraction method is known to systematically over-estimate silts and underestimate clay (Fisher et al. 2017; Yang et al. 2019). To correct for the bias and align with the more commonly used pipette method, the particle diameter cutoff of clay was increased from < 2.0 to < 6.0 μm (Yang et al. 2019). Magnetic susceptibility, a general proxy for iron (Fe) content and crystallinity (Mullins 1977; Shang and Tiessen 2000; von Lützow et al. 2007), was measured using a magnetic susceptibility meter and dual frequency sensor (MS2 and MS2B, respectively; Bartington Instruments Ltd, Witney, UK) on 10 g of air-dried soil in a 10 mL polystyrene sample jar. Soil pH was measured following standard procedures using a 2:1 DI water-soil mixture. Ammonium acetate extractions of calcium (Ca), potassium (K), magnesium (Mg), and sodium (Na) as well as potassium chloride extractions of Fe and aluminum (Al) were conducted at the Oregon State University Central Analytical Laboratory following standard procedures (Soil Survey Staff 2014). Selective dissolution analyses (sodium pyrophosphate, ammonium oxalate, and dithionite-citrate) to measure various phases of Fe, Al, manganese (Mn), and silicon (Si) were conducted

following standard USDA NRCS protocols (Soil Survey Staff 2014).

Pyrogenic C quantification

Pyrogenic C was quantified using the benzene poly-carboxylic acid (BPCA) method as outlined in Matosziuk et al. (2019, 2020), which uses nitric acid to oxidize the extended aromatic sheets characteristic of pyrogenic C into individual carboxylated benzene rings that can be isolated and quantified using high performance liquid chromatography (HPLC). Briefly, ground soil samples containing ~ 2 mg C were digested in 5 mL nitric acid (HNO_3) at 170 °C for 8 h using pressurized microwave vessels (MASS 6; CEM, Mathews, NC, USA). Samples were filtered through glass fiber filters (Whatman, GF/A), and the remaining solids were washed with 5 mL of 1 M sodium hydroxide (NaOH). Samples were diluted to 50 mL with deionized water, flash frozen with liquid nitrogen, and freeze-dried (FreeZone Plus; Labconco, Kansas City, MO, USA). The remaining residue was dissolved in 2 mL 1 M NaOH and filtered using 0.45 μm nylon syringe filters (Whatman). A 1 mL aliquot of sample was transferred to a clean vial, spiked with 600 μL of 2 M HCl, and analyzed by HPLC (Shimadzu LC-10AD equipped with a SPD-M20A photodiode array capable of measuring wavelengths between 190 and 400 nm; Shimadzu Scientific Instruments, Columbia, MD, USA). An Agilent Poroshell 120 SB-C18 column (Agilent, Santa Clara, CA, USA) was used with a mobile phase consisting of a binary gradient of phosphoric acid (H_3PO_4) (2% in water) and acetonitrile (Wiedemeier et al. 2013). In addition to quantifying total BPCA, this method provides additional information on the structure of pyrogenic C based on the fragmentation pattern produced by the acid digestion (Matosziuk et al. 2020), and as such, the benzene hexacarboxylic acid:total BPCA ratio can serve as an aromatic condensation index with higher values indicating a greater degree of condensation.

Cupric oxide oxidation

Whole soils and density fractions were analyzed by alkaline cupric oxide (CuO) oxidation (Goñi and Montgomery 2000; Hatten et al. 2012; Hatten and

Goñi 2016) to obtain the yields of lignin (vanillyl, syringyl, and cinamyl phenols), para-hydroxy benzenes, diacids, fatty acids, and benzoic acids (hydroxy, mono, di, and tri). Lignin is uniquely produced by plants, while many of these other biomolecules are produced by plants, microbes, fungi, algae, and animals. To isolate the non-plant signal in the CuO oxidation products, lignin was used to normalize para-hydroxy benzenes, diacids, and fatty acids. The benzene tri-carboxylic acids:lignin (BTCA:lignin) ratio was used as an indicator of pyrogenic C, following Hatten and Goñi (2016).

Dissolved organic matter (DOM) water extraction and fluorescence analysis

Dissolved organic matter (DOM) was sequentially extracted from bulk soils using ultrapure DI water at two temperatures. Sequential water extracts were collected by extracting 1 g bulk soil with 30 mL of ultrapure DI water in a combusted (450 °C for 5 h) borosilicate glass centrifuge tube. To extract cold water extractable organic matter (C-WEOM), a proxy of DOM (von Lützow et al. 2007), samples were shaken at 20 °C for 2 h in a water shaking bath (Fisherbrand Isotemp Shaking Water Bath; Thermo Fisher Scientific, Pittsburgh, PA, USA). After 2 h, samples were centrifuged (6000 rpm for 8 min), and C-WEOM in supernatant was collected, leaving the soil behind. The C-WEOM solution was filtered using a combusted fritted glass filter assembly (DWK Life Sciences LLC, Millville, NJ, USA) and combusted glass fiber filters size F (0.7 μm, Whatman GF/F, 450 °C for 5 h). For the hot water extractable organic matter (H-WEOM), which serves as a proxy of labile SOM (von Lützow et al. 2007), 30 mL of ultrapure DI water was added to the extracted soils and allowed to shake in a hot water bath at 80 °C for 16 h. After 16 h, the samples underwent the same sample processing as C-WEOM. Samples were diluted sixfold to ensure that the concentration would not interfere with analysis and stored at 4 °C until analyzed.

WEOM samples were analyzed for UV absorbance using a UV spectrophotometer (Agilent 8453; Agilent, Santa Clara, CA, USA) with a 1 cm quartz cuvette over the wavelength range 190 nm to 1100 nm. Samples with absorbances greater than 0.3 at 254 nm (Abs254) were diluted with ultrapure DI

water until Abs254 values were within the range of 0.1 to 0.2 to limit inner-filter effects during collection of excitation emission matrices (EEMs) (Ohno 2002; Miller and McKnight 2010; SanClements et al. 2012). EEMs were collected using a fluorometer (Fluoromax-3; HORIBA Jobin Yvon Inc., Edison, NJ, USA) over an excitation range of 240–450 nm in 10 nm increments, while emission was monitored from 300 to 550 nm in 2 nm increments. All scans were corrected for instrument specific bias using manufacturer supplied correction factors for both excitation and emission wavelengths, corrected for inner-filter effects using UV-VIS absorbance scans, normalized to the area under the instrument corrected Raman curve at emission 350 nm, and blank subtracted using corrected EEMs of daily ultrapure DI blanks analyzed over the same wavelengths as samples (Cory and McKnight 2005; Cory et al. 2010; Miller and McKnight 2010).

The fluorescence index (FI) was calculated as the ratio of emission intensity wavelength of 470 nm to the emission intensity at 520 nm at an excitation wavelength of 370 nm. The FI is primarily a measure of DOM source (i.e., plant versus microbial) and is generally correlated with the aromaticity of humic materials (McKnight et al. 2001; Cory and McKnight 2005). As DOM becomes increasingly microbial, the FI value increases. The humification index, an index commonly used in fluorometric studies of DOM, was calculated as the ratio of the peak area under emission at 435–485 nm and peak area under emission at 300–345 nm at excitation of 254 nm (Zsolnay et al. 1999; Wilson and Xenopoulos 2009; Huguet et al. 2009). Higher humification index values are indicative of lower H/C ratios, shifting the emission to longer wavelengths (Zsolnay et al. 1999). The biological index was calculated as the ratio of maximum intensity at emission at 380 nm divided by the maximum intensity between emission at 420 and 435 nm at an excitation of 310 nm. The biological index indicates the proportion of recently produced (likely microbial) organic matter to older, more degraded organic matter.

DOM sequential extraction and mass spectrometry analysis

Sequential water, methanol, and chloroform extractions were completed at the Environmental Molecular Sciences Lab (EMSL), a U.S. Department of Energy

(DOE) national user facility, using a modified method from Tfaily et al. (2017). Water extracts are a proxy for labile DOM, chloroform extracts are a proxy for lipids associated with minerals and cellular membranes, and methanol extracts have characteristics intermediate between water and chloroform extracts (Tfaily et al. 2015, 2017; Graham et al. 2017). Soil extracts were prepared in triplicate by adding 5 mL of solvent (water, methanol, or chloroform) to 1 g air-dried, sieved bulk soil from each composited horizon in 15 mL polypropylene centrifuge tubes (Olympus 50 mL Centrifuge Tubes; Genesee Scientific, San Diego, CA, USA) and shaken on an orbital shaker (150 rpm for 2 h). Soils were allowed to settle before spinning down and collection of supernatant. This process was repeated adding methanol and chloroform sequentially to the extracted soils. The water and methanol extracted organic matter underwent solid phase extraction (SPE) to desalt samples and improve signal. Prior to SPE, methanol extracts were diluted with ultrapure DI water, and both water and methanol extracts were acidified to a pH of 2 with H_3PO_4. SPE cartridges (Bond Elut PPL, 100 mg bed mass, 3 mL volume, non-polar; Agilent, Santa Clara, CA, USA) were primed with 1 mL of methanol prior to adding acidified extract. Once the acidified extract had completely run through, the SPE cartridge was rinsed with 15 mL of 10 mM HCl to remove contaminants. The SPE cartridge was dried using lab air, and after the cartridge was completely dry, the sample was eluted off using 1.5 mL of methanol. Extracts were stored in 2 mL MicroSolv vials at −80 °C until analysis.

High resolution mass spectra of DOM extracts were collected using a 12 T Bruker SolariX Fourier transform ion cyclotron resonance mass spectrometer (FTICR MS) located at EMSL. Chloroform extracts were diluted with methanol to improve ionization prior to analysis. The instrument was regularly calibrated as outlined in Tfaily et al. (2017). Ion accumulation time was varied to account for concentration extraction efficiency and ranged from 0.1 to 1 s to minimize biases from differing organic matter composition and allow for cross-soil comparison. Three hundred individual spectra were averaged for each sample and internally calibrated with the mass measurement accuracy being better than 1 ppm for single charged ions.

All sample peaks listed were aligned to each other prior to formula assignment to eliminate possible mass

shifts that would impact formula assignment. Formulas were assigned using Formularity (Tolić et al. 2017) that uses the compound identification algorithm (CIA). Chemical formulas were assigned using the following criteria: (1) signal-to-noise ratio $(S/N) > 7$; (2) mass measurement error < 1 ppm; (3) include only C, H, O, N, S, and P and exclude all other elements. When there were multiple possible formulas, the formula that was assigned was chosen through the propagation of CH_2, H_2, and O homologous series. To ensure the consistent choice of formula when multiple options were available, the following rules were used: (1) the formula with the lowest error and lowest number of heteroatoms was chosen, and (2) the assignment of one phosphorus atom required four oxygen atoms.

The assigned formulas were plotted on Van Krevelen diagrams using the molar H:C and O:C ratios and assigned to the major biochemical classes using the *ftmsRanalysis* package in R (Bramer et al. 2020). Based on the assigned molecular formula, Cox Gibbs free energy (GFE), double bond equivalents (DBE), double bond equivalents minus oxygen (DBE-O), aromaticity index, H:C ratio, and O:C ratio were computed. To compare composition across samples, the means of the stated metrics were calculated by taking the mean of the numeric output for each assigned formula. DBE is the sum of double bonds or rings in each molecule and is used to determine the degree of unsaturation or the density of the C–C double bonds (Koch and Dittmar 2006, 2016). With decreasing hydrogen atoms, the degree of unsaturation increases, and the DBE increases. DBE-O is a measurement of the unsaturation but includes oxygen in the calculation since most oxygen atoms in DOM molecules are part of a carboxyl group that is counted as one DBE.

Statistical analysis

All statistical analyses were performed with R version 3.6.1 (R Core Team 2019) in RStudio 1.2.1335 (RStudio Team 2018). To compare CSR and predictors in A and B horizons, paired sample tests (paired t-test, Wilcoxon signed-rank test, and Brunner-Munzel rank-order test, depending on whether normality and homoscedasticity assumptions were met as determined using the Shapiro–Wilk test and Levene's test) were used. The significance level (α) chosen was 0.05.

Where appropriate, the Bonferroni correction was used to account for multiple comparisons.

To determine which predictors were most important for predicting CSR over 52 weeks, the *Boruta* package in R was used (Kursa and Rudnicki 2010). The *Boruta* package implements the Boruta algorithm, which is a feature selection algorithm that uses random forest classification to find all relevant predictors. This approach differs from traditional "minimal-optimal" feature selection approaches in that *all* relevant predictors, regardless of correlations, are selected. As such, this approach potentially allows for greater insights into underlying mechanisms by not discarding predictors; however, this characteristic also means the output of the Boruta algorithm cannot be directly used to create a predictive model. Additionally, like other machine learning feature selection algorithms, the Boruta algorithm is affected by effect size (Degenhardt et al. 2019); consequently, it is possible for a significantly correlated variable to be excluded if its effect size is not large enough. Furthermore, the Boruta algorithm, like other classification and regression tree methods, is robust to outliers (Murphy 2012; Härdle and Simar 2019), but as a "low bias, high variance" method, it can be sensitive to the inputs used, although random forests average across many estimates to reduce this effect (Murphy 2012). Excluding one or more rows (i.e., sites in our data set) may change the predictors selected, but this may be due to simply changing the input data set versus the effect of excluding one or more outliers per se. Consequently, we present our analysis using all of our samples despite the presence of outliers.

Additionally, a subset of predictors had missing values, typically due to not being able to meet the C mass requirement for certain analyses for samples with very low C concentrations. Because the Boruta algorithm cannot be used on data sets with missing values, chained-equation based multiple imputation was conducted using the *mice* (multivariate imputation via chained equations) package in R to produce five independently imputed data sets to address any missing values (van Buuren and Groothuis-Oudshoorn 2011). The Boruta algorithm was used on all five imputed data sets with default settings (confidence level = 0.01, multiple comparisons adjustment used) except for the maximum number of runs (maximum number of runs = 1000). Those variables that were

selected in at least three of the five data sets were considered to be important predictors.

Because the Boruta algorithm uses the all-relevant approach in which all related features for predicting a given response are selected, selected predictors were clustered based on distance calculated using mean Spearman correlation coefficients for the five imputed data sets (specifically, Spearman correlation distance = $1 - |\rho_{mean}|$). A cluster dendrogram was created using the "complete" method, and the number of clusters was determined using the "elbow method" to visualize broad categories of controls on SOM vulnerability (Hennig et al. 2015). The sign of the effect of the predictors was determined by calculating the Spearman correlation coefficient between a given predictor and CSR along with the corresponding p-value using the original non-imputed data set. Using the Spearman correlation coefficient instead of the Pearson correlation coefficient takes into account potential non-linear relationships between predictors and CSR.

Results

Cumulative specific respiration (CSR)

Cumulative specific respiration (CSR) after 52 weeks ranged from 2.5 to 17.0% of the initial SOC content for A horizons and from 1.3 to 23.0% for B horizons (Fig. 2 and Supplementary Table S2). When accounting for multiple comparisons with the Bonferroni correction ($\alpha = 0.017$), CSR from 0 to 52 weeks in A horizons was significantly different (p-value = 0.013) from that in B horizons (Fig. 2). Likewise, during the first half of the incubation (0 to 26 weeks), CSR was significantly higher in A horizons than B horizons (p-value = 0.002). However, during the second half of the incubation (26 to 52 weeks), there was no significant difference between horizon types. A comparison of initial SOC content by horizon type can be found in Supplementary Fig. S1.

Predictors of CSR

Twelve predictors were selected by the Boruta algorithm for A horizon CSR (Table 2, Supplementary Table S3, and Figs. 3, 4), while only seven predictors were selected for CSR in B horizons (Table 3,

Supplementary Table S4 and Figs. 5, 6). Relationships between CSR and individual predictors are shown in Figs. 3 and 5. Based on the cluster analysis using mean Spearman correlation distance, the predictors for A horizons can be clustered into three broad categories of controls on SOM vulnerability: SOM chemistry, reactive Fe and Al phases, and site moisture availability (Fig. 4). Correlations among the predictors in these categories are shown in Supplementary Fig. S2. In contrast, the predictors for B horizons can be clustered into four groups corresponding to two broad categories of controls: SOM chemistry and site moisture availability (Fig. 6). Correlations among these predictors are shown in Supplementary Fig. S3. The signs of the effect of each predictor on CSR are summarized in Table 4 through Spearman's correlation coefficients and associated p-values. Note that these correlations and associated p-values were calculated only for observed values (i.e., using non-imputed data) and thus demonstrate that the link between selected predictors and CSR is not an artifact of imputation.

Discussion

SOM vulnerability to mineralization predicted by diverse suite of variables

In our study, a diverse suite of variables was selected for predicting SOM vulnerability to mineralization. Consistent with previous large-scale studies of SOC (Doetterl et al. 2015; Rasmussen et al. 2018; Viscarra Rossel et al. 2019; von Fromm et al. 2021; Yu et al. 2021), this varied group of predictors illustrates how numerous factors control SOM dynamics at large (i.e., continental) scales. The variables selected generally fall into three broad categories related to (1) SOM chemistry, (2) reactive Fe and Al phases, and (3) site moisture availability (Table 4; Figs. 4 and 6).

SOM chemistry

SOM chemistry likely affects SOM vulnerability to mineralization by a variety of mechanisms. Pyrogenic C reduced vulnerability, as indicated by the negative correlation between BTCA:lignin ratio and CSR, which could be due to sorption of otherwise labile DOM onto pyrogenic C or "dilution" of more
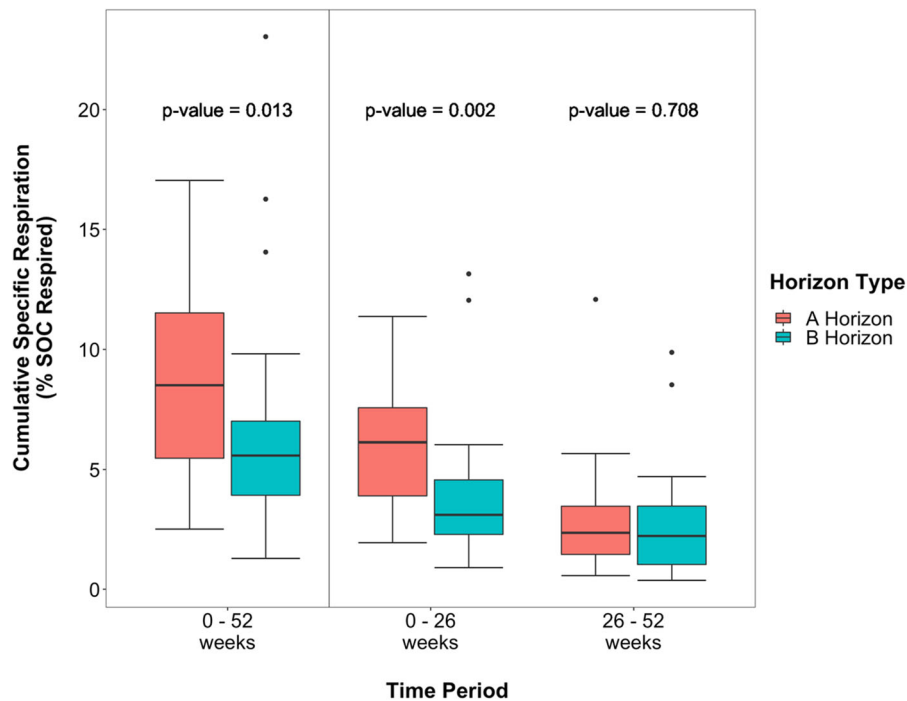
Fig. 2 Box plot of cumulative specific respiration (CSR) expressed as % initial soil organic carbon (SOC) respired by incubation time period and horizon (n = 26 for each horizon type, 1 replicate per site-horizon combination). Samples collected from 26 NEON sites spanning the conterminous U.S. and incubated at site-specific mean summer temperature and field capacity (−33 kPa) water potential. Lower and upper box edges represent 25th and 75th percentiles, respectively, whiskers are drawn to the most distant data points within 1.5 times the interquartile range (IQR), and dots indicate outliers beyond 1.5 times the IQR. Given p-values correspond to the Wilcoxon signed-rank test between paired samples from the two horizon types for a specified time interval

mineralizable substrates (DeCiucies et al. 2018). Selected degradation indices (syringyl phenols acid:aldehyde ratio and hot water extract biological index) had conflicting correlations with vulnerability. In A horizons, more degraded lignin corresponded to lower CSR, while in B horizons, more degraded DOM was associated with increased CSR. These results suggest that the relationship between SOM degradation state and vulnerability is complex and may depend on specific molecular characteristics, such as those that control a molecule's affinity for sorbing to a mineral surface (Kaiser et al. 2004; Hernes et al. 2007). Higher C:N ratios had a negative correlation with CSR, while increasing values of a proxy of the microbial contribution to SOC (chloroform extract mean DBE) had a positive correlation with CSR. These two relationships could indicate a direct effect of substrate stoichiometry on mineralization. The predictors within the SOM chemistry category underscore the importance of substrate chemistry to

decomposition. However, given seemingly contradictory relationships between CSR and different measures of substrate quality, this relationship appears to be mechanism dependent with at least some of the effects of substrate chemistry on CSR potentially being indirect (e.g., enhanced sorption to mineral surfaces).

Reactive Fe and Al phases

In contrast to the predictors in the SOM chemistry category, the four predictors within the reactive Fe and Al phases category had a consistent effect on CSR, suggesting a simpler linkage between reactive Fe and Al phases and SOM vulnerability. Specifically, the four selected selective dissolution predictors are proxies for the following: pyrophosphate extractable Fe represents organically complexed Fe (Wada 1989); ammonium oxalate extractable Fe and Al represent organo-metal complexes and short-range-

Table 2 Predictors selected by Boruta algorithm as important for predicting cumulative specific respiration (CSR) in A horizons for a majority of the imputed data sets. Predictors are associated with bulk soil samples unless specified otherwise. Times selected (%) indicates the percentage of imputed data sets (n = 5) for which a given predictor was selected. Mean importance rank indicates the relative predictive importance as determined by the Boruta algorithm (e.g., lower values indicate higher importance). Imputed (%) states the percentage of data for which missing data were imputed using multivariate imputation by chained equations for use with the Boruta algorithm

| Predictor	Times selected (%)	Mean importance rank	Imputed (%)
Hargreaves climatic moisture deficit	100	1.4	0
BTCA:lignin ratio (heavy fraction)	100	1.8	27
Ammonium oxalate extractable Fe	100	4.0	0
Ammonium oxalate extractable Al	100	5.4	0
MAP − Hargreaves reference evaporation	100	6.2	0
Sodium pyrophosphate extractable Fe	100	6.4	0
Annual heat:moisture index	100	8.6	0
Potassium chloride extractable Al	80	11.2	0
Syringyl phenols acid:aldehyde ratio	60	11.2	12
Summer heat:moisture index	80	11.6	0
BTCA:lignin ratio (bulk soil)	60	12.6	12
Dithionite-citrate extractable Al	60	13.0	0

BTCA benzene tri-carboxylic acids products, *MAP* mean annual precipitation

order oxyhydroxide phases (Wada 1989); and dithionite-citrate extractable Al mainly represents organically complexed Al, Al from short-range-order oxyhydroxide phases, and Al from isomorphous substitution into Fe oxides (Wada 1989; Pansu and Gautheyrou 2006). Increased concentrations of the four selected proxies of reactive Fe and Al were negatively correlated with CSR, which is consistent with previous studies demonstrating the importance of reactive Fe and Al phases in controlling SOM dynamics (Torn et al. 1997; Rasmussen et al. 2006, 2018; Doetterl et al. 2015; Kramer and Chadwick 2018; Yu et al. 2021). It is worth noting that exchangeable Ca^{2+} and Mg^{2+} were not selected, which suggests the protective capacity of divalent cations is reduced at higher moisture levels as used in the incubation. Furthermore, it should be noted % clay was not selected, underscoring the importance of surface chemistry in controlling SOM dynamics (Rasmussen et al. 2018).

Site moisture availability

The four predictors exclusively linked to climate in the site moisture availability group (MAP − Hargreaves reference evaporation, annual heat:moisture index, summer heat:moisture index, and Hargreaves climatic moisture deficit) had a consistent effect on CSR, with drier sites having higher CSR values. The other predictor clustered in this group was potassium chloride extractable Al, which is a measure of exchangeable Al and in soils with a pH < 5.5, a measure of active acidity (Soil Survey Staff 2014). In our sample set, potassium chloride extractable Al was dependent on a site's moisture availability (Supplementary Fig. S2), which explains it being clustered with predictors of site moisture availability. Exchangeable Al toxicity has been observed under acidic conditions (Illmer et al. 1995, 2003; Kunito et al. 2016), and consequently, Al toxicity is one potential mechanism by which SOM at wetter, more acidic sites is made less vulnerable to mineralization. Likewise, concentrations of reactive Fe and Al phases are positively correlated with site moisture availability (Supplementary Fig. S2), and thus interactions between SOM and reactive Fe and Al phases are another mechanism by which SOM at wetter sites could be protected (Rasmussen et al. 2018; Kramer and Chadwick 2018). Site moisture availability also causes changes in plant productivity and litter chemistry (Schuur and Matson 2001; Schuur 2003; Santiago
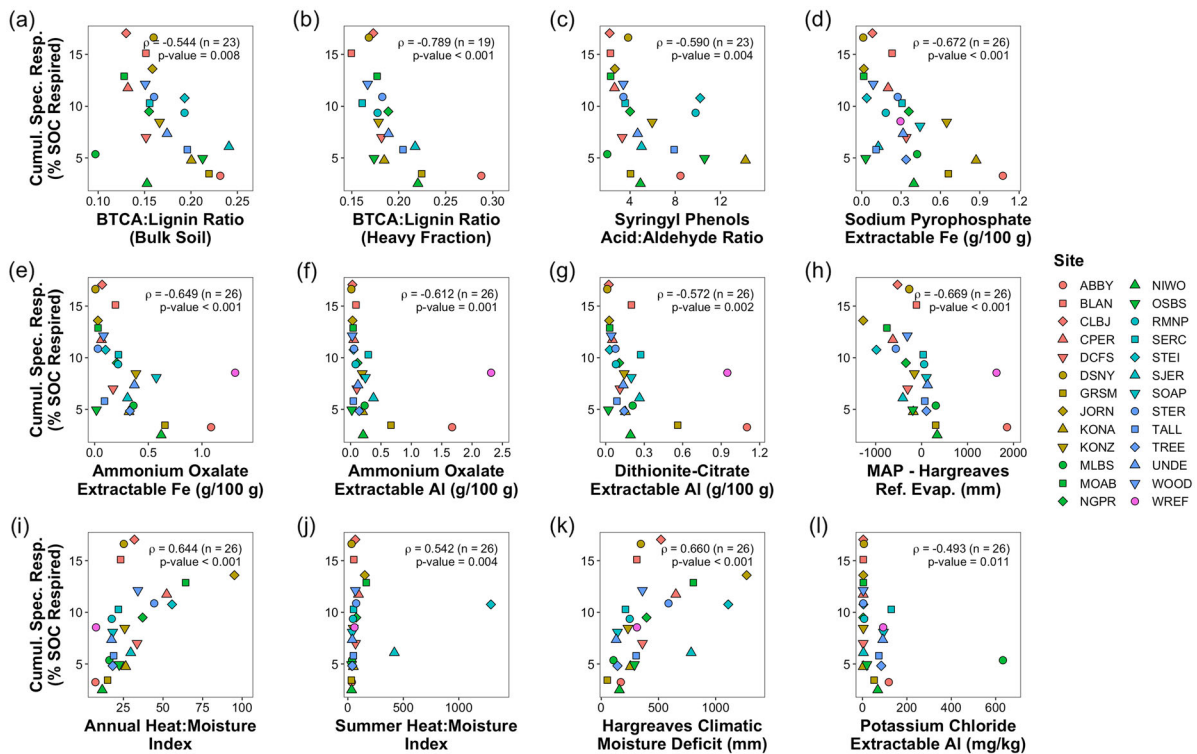
Fig. 3 Cumulative specific respiration (CSR) expressed as % initial soil organic carbon (SOC) respired versus 12 predictors selected by Boruta algorithm for A horizons. Spearman correlation coefficient (rho) and p-value were calculated using non-imputed data with number of samples for a specific predictor given in parentheses. Samples collected from 26 NEON sites spanning the conterminous U.S. and incubated at site-specific mean summer temperature and field capacity (-33 kPa) water potential. Panels correspond to the following: (**a**) benzene tri-carboxylic acids products (BTCA):lignin ratio (bulk soil), (**b**) BTCA:lignin ratio (heavy fraction), (**c**) syringyl phenols acid:aldehyde ratio, (**d**) sodium pyrophosphate extractable Fe, (**e**) ammonium oxalate extractable Fe, (**f**) ammonium oxalate extractable Al, (**g**) dithionite-citrate extractable Al, (**h**) mean annual precipitation (MAP) $-$ Hargreaves reference evaporation, (**i**) annual heat:moisture index, (**j**) summer heat:moisture index, (**k**) Hargreaves climatic moisture deficit, and (**l**) potassium chloride extractable Al

et al. 2005; Luo et al. 2017) as well as microbial community structure and function (Drenovsky et al. 2010; Brockett et al. 2012; Guenet et al. 2012; Tatsumi et al. 2019), providing additional mechanisms that may affect SOM vulnerability.

Given correlations among site moisture availability and various potential SOM stabilization mechanisms, it is challenging to distinguish direct effects from indirect effects of site moisture availability. However, because indices of site moisture availability, such as MAP $-$ Hargreaves reference evaporation, are integrative measures of several controls on SOM decomposition, such variables are potentially powerful predictors of SOM vulnerability to decomposition, and given their ability to be accurately estimated at broad scales, they are potentially useful for scaling predictions of SOM vulnerability.

Differences in key predictors with depth driven by several factors

Although there were several predictors selected for A and B horizons that indicate some controls on SOM vulnerability common to both surface and subsurface horizons (i.e., pyrogenic C, site moisture availability), most important predictors of SOM vulnerability to mineralization were unique to only one horizon type (Table 4), suggesting dominant controls on C turnover generally change with depth. We hypothesize that the differences in dominant controls are driven by changes in the soil environment; SOM concentration, chemistry, and accessibility; and microbial community composition and function with depth.

By being at or close to the surface, A horizons are more affected by surficial influences and have
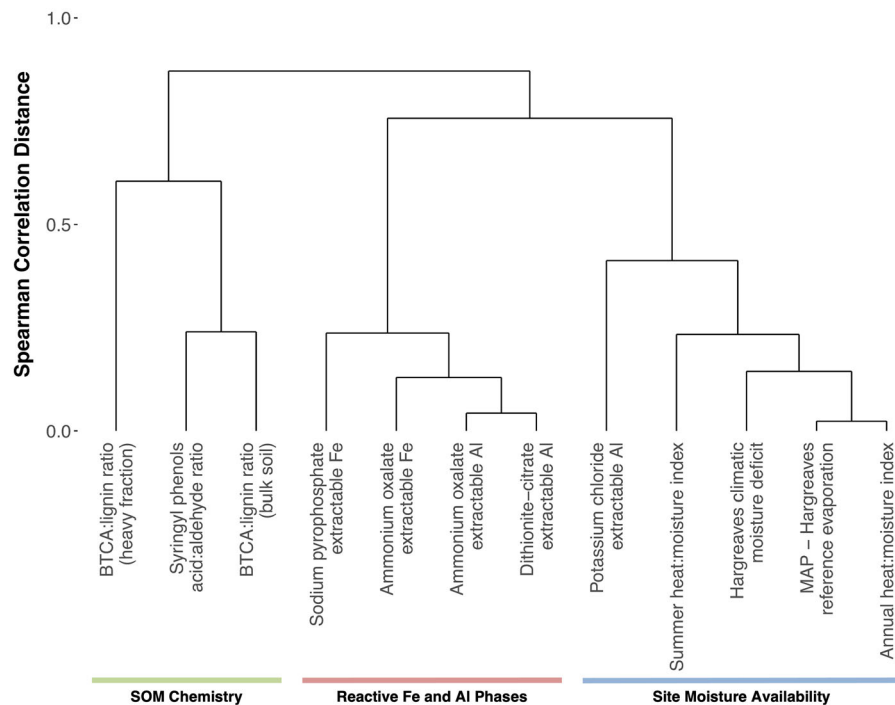
Fig. 4 Cluster dendrogram based on mean Spearman correlation distance for imputed predictors selected by Boruta algorithm for cumulative specific respiration (CSR) in A horizons. Predictors are associated with bulk soil samples unless specified otherwise. Clusters are indicated by colored lines underneath dendrogram with labels corresponding to associated predictor category (SOM chemistry, reactive Fe and Al phases, and site moisture availability). BTCA, benzene tri-carboxylic acids products; MAP, mean annual precipitation

Table 3 Predictors selected by Boruta algorithm as important for predicting cumulative specific respiration (CSR) in B horizons for a majority of the imputed data sets. Predictors are associated with bulk soil samples unless specified otherwise. Times selected (%) indicates the percentage of imputed data sets (n = 5) for which a given predictor was selected. Mean importance rank indicates the relative predictive importance as determined by the Boruta algorithm (e.g., lower values indicate higher importance). Imputed (%) states the percentage of data for which missing data were imputed using multivariate imputation by chained equations for use with the Boruta algorithm

Predictor	Times selected (%)	Mean importance rank	Imputed (%)
Hot water extract biological index	100	2.2	8
Annual heat:moisture index	100	3.2	0
MAP − Hargreaves reference evaporation	100	3.8	0
Chloroform extract mean DBE	100	4.8	0
C:N ratio (bulk soil)	80	5.2	0
BTCA:lignin ratio	80	7.0	42
C:N ratio (heavy fraction)	60	11.2	15

MAP mean annual precipitation, *DBE* double bond equivalents, *BTCA* benzene tri-carboxylic acids products

characteristics that suggest some similarities with litter layers. The selection of certain predictors exclusively for A horizons, such as syringyl phenols acid:aldehyde ratio, which is linked to lignin inputs from angiosperms, and summer heat:moisture index, which is a measure of site moisture availability for
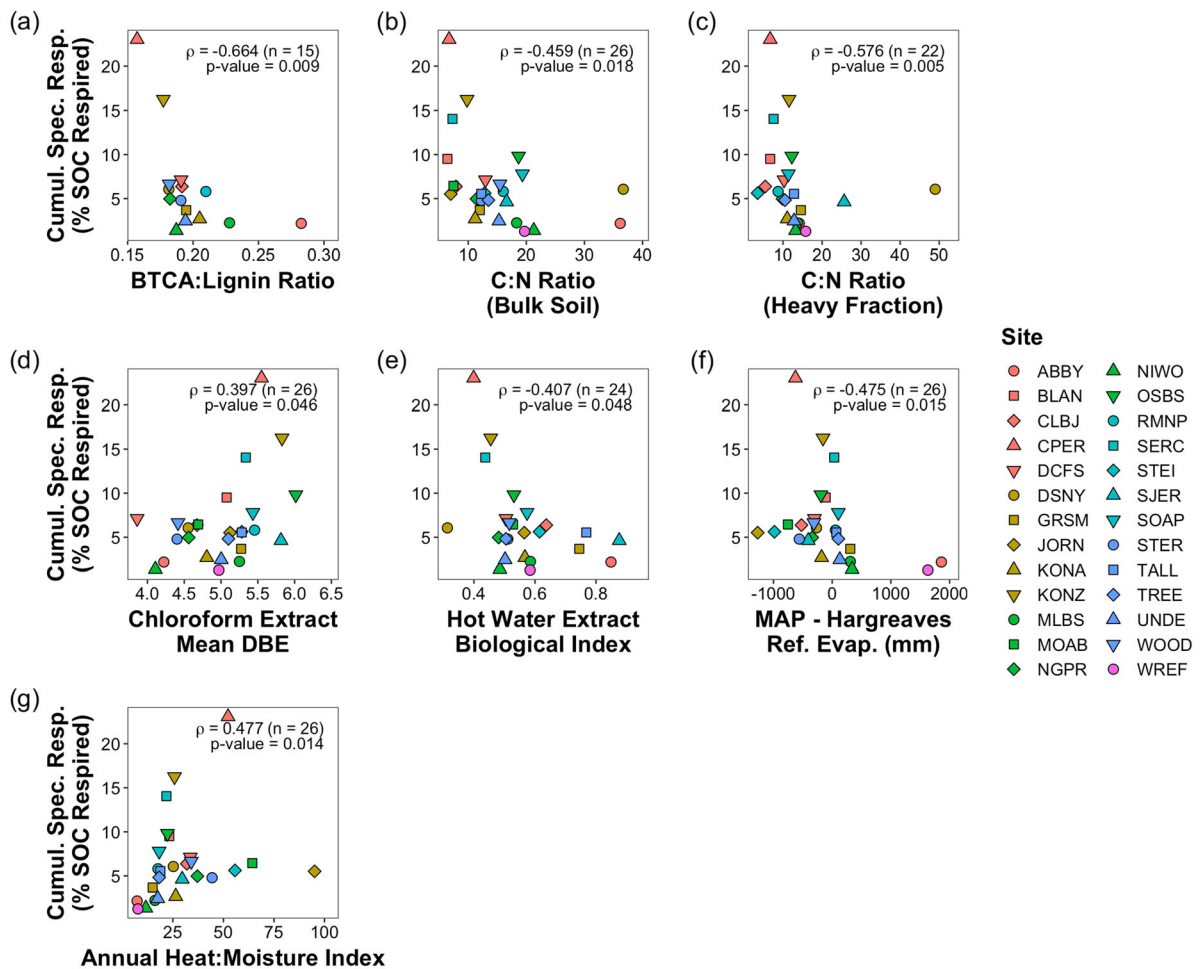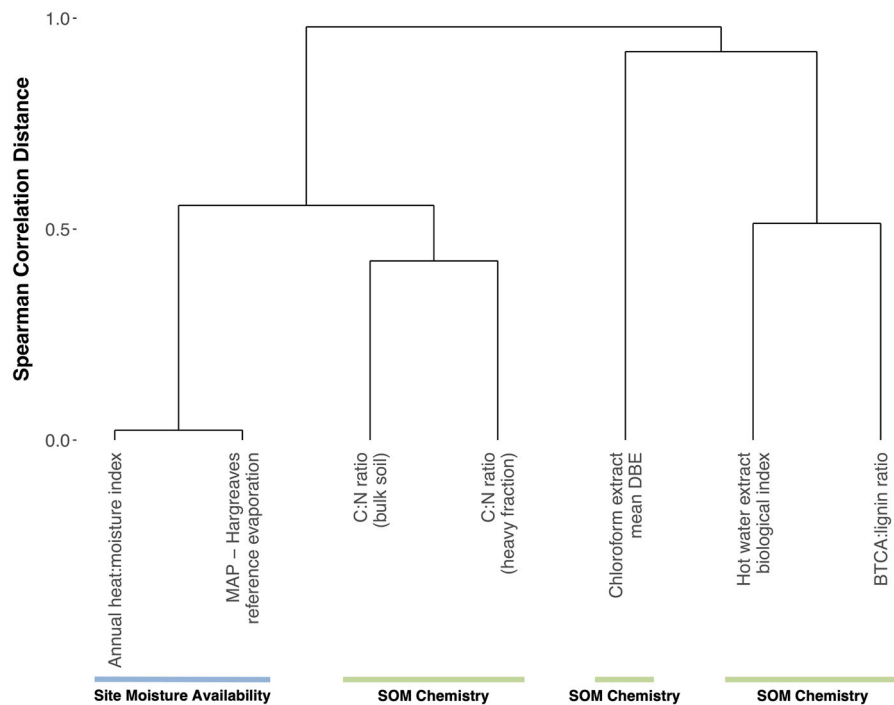
Fig. 5 Cumulative specific respiration (CSR) expressed as % initial soil organic carbon (SOC) respired versus 7 predictors selected by Boruta algorithm for B horizons. Spearman correlation coefficient (rho) and p-value were calculated using non-imputed data with number of samples for a specific predictor given in parentheses. Samples collected from 26 NEON sites spanning the conterminous U.S. and incubated at site-specific mean summer temperature and field capacity (-33 kPa) water potential. Panels correspond to the following: (**a**) benzene tri-carboxylic acids products (BTCA):lignin ratio, (**b**) C:N ratio (bulk soil), (**c**) C:N ratio (heavy fraction), (**d**) chloroform extract mean double bond equivalents (DBE), (**e**) hot water extract biological index, (**f**) mean annual precipitation (MAP) $-$ Hargreaves reference evaporation, and (**g**) annual heat:moisture index

only a fraction of the year, is consistent with our understanding of surficial influences having the greatest effect at or near the surface in soils, especially over shorter timescales (Jobbagy and Jackson 2000; Gray et al. 2015). Furthermore, differences in SOM chemistry between A and B horizons (Supplementary Fig. S4) highlight the different degrees of transformation of surficial inputs from plants. Proxies of lignin degradation (syringyl phenols acid:aldehyde ratio) and organic matter degradation (3,5-di-hydroxy benzoic acid:lignin ratio, fatty acids:lignin ratio, and di-

acids:lignin ratio) show that SOM in A horizons is less degraded than SOM in B horizons and consequently more similar to the original plant inputs. Consistent with this, A horizons also have a higher proportion of C in the free light fraction and greater concentrations of sodium pyrophosphate extractable Mn (Supplementary Fig. S4), which is a proxy for bioavailable Mn and has been shown to be important for forest floor litter decomposition (Berg et al. 2007; Keiluweit et al. 2015). These various lines of evidence point to the conclusion that SOM in A

Fig. 6 Cluster dendrogram based on mean Spearman correlation distance for imputed predictors selected by Boruta algorithm for cumulative specific respiration (CSR) in B horizons. Predictors are associated with bulk soil samples unless specified otherwise. Clusters are indicated by colored lines underneath dendrogram with labels corresponding to associated predictor category (SOM chemistry and site moisture availability). MAP, mean annual precipitation; DBE, double bond equivalents; BTCA, benzene tri-carboxylic acids products

horizons has characteristics intermediate between organic matter in litter layers and SOM in deeper mineral horizons.

B horizons are less susceptible to surficial influences and have a different set of characteristics from A horizons. B horizons have higher bulk density values than A horizons (Supplementary Fig. S4), which has implications for the movement of water, gases, DOM, and heat. Furthermore, concentrations of dithionite-citrate extractable Fe, a proxy of pedogenically active Fe (Strahm and Harrison 2008), are higher in B horizons than A horizons (Supplementary Fig. S4), suggesting a different weathering and/or leaching environment from surface horizons. Furthermore, the concentration of organic C is lower in B horizons, SOM is more chemically degraded, and a smaller proportion of C is in the free light fraction (Supplementary Fig. S4). Together, these factors may lead to a weaker relationship between reactive Fe and Al phases and vulnerability in the subsoil, which is demonstrated by a substantially lower R^2 value for B horizons when CSR is regressed against the selected proxies of reactive Fe and Al phases by horizon type (except for ammonium oxalate extractable Al, Supplementary Fig. S5). This finding is consistent with the findings of Yu et al. (2021) who found that the importance of geochemical predictors of SOC concentration did not increase with depth as typically assumed. Consequently, this is not to say that reactive Fe and Al phases play no role in controlling SOM vulnerability in B horizons per se. Rather, their effect seems to be modulated by the different physicochemical environment in the subsoil compared to surficial soils.

In addition to the proxies of the soil environment and SOM characteristics that we did measure, there are likely substantial differences in microbial community composition and function between horizon types. Microbial biomass and diversity are known, for example, to decrease with depth (Fierer et al. 2003; Eilers et al. 2012; Stone et al. 2014; Brewer et al. 2019). Consequently, although we did not measure microbial community characteristics, we can conjecture that changes in the microbial community, such as a decrease in the fungal:bacterial ratio with depth

Table 4 Spearman correlation coefficients and p-values for predictors selected by Boruta algorithm in relation to cumulative specific respiration (CSR). "Not selected" indicates that a given predictor was not selected as important by the Boruta algorithm for that horizon type. Note that correlation coefficients and p-values were calculated for only observed values (i.e., using non-imputed data)

Predictor group	Predictor	Spearman correlation coefficient (p-value)	
		A Horizons	B Horizons
SOM chemistry	BTCA:lignin ratio (bulk soil)	−0.544 (0.008)	−0.664 (0.009)
	BTCA:lignin ratio (heavy fraction)	−0.789 (< 0.001)	Not selected
	Syringyl phenols acid:aldehyde ratio	−0.590 (0.004)	Not selected
	C:N ratio (bulk soil)	Not selected	−0.459 (0.018)
	C:N ratio (heavy fraction)	Not selected	−0.576 (0.005)
	Chloroform extract mean DBE	Not selected	0.397 (0.046)
	Hot water extract biological index	Not selected	−0.407 (0.048)
Reactive Fe and Al	Sodium pyrophosphate extractable Fe	−0.672 (< 0.001)	Not selected
	Ammonium oxalate extractable Fe	−0.649 (< 0.001)	Not selected
	Ammonium oxalate extractable Al	−0.612 (0.001)	Not selected
	Dithionite-citrate extractable Al	−0.572 (0.002)	Not selected
Site moisture availability	MAP − Hargreaves reference evapotranspiration	−0.669 (< 0.001)	−0.475 (0.015)
	Annual heat:moisture index	0.644 (< 0.001)	0.477 (0.014)
	Summer heat:moisture index	0.542 (0.004)	Not selected
	Hargreaves climatic moisture deficit	0.660 (< 0.001)	Not selected
	Potassium chloride extractable Al	−0.493 (0.011)	Not selected

(Fierer et al. 2003; Stone et al. 2014) or changes in extracellular enzyme activities with depth (Dove et al. 2020), have important consequences for SOM mineralization in A horizons versus B horizons. Differences in the microbial community together with differences in the soil environment and SOM characteristics contribute to SOM in A horizons generally behaving differently from SOM in B horizons. Thus, models of SOM cycling may benefit from considering how controls on SOM decomposition change with depth.

Conclusion

Using data from a 52-week laboratory incubation of soils from NEON sites spanning the conterminous United States, this study provides evidence for the dominant role of three categories of controls on SOM vulnerability at a continental scale: SOM chemistry, reactive Fe and Al phases, and site moisture availability. Of these three categories, predictors of site moisture availability, such as MAP − Hargreaves

reference evaporation, have the most potential for being used to scale estimates of SOM vulnerability given their integrative nature and ability to be accurately estimated at broad scales. Furthermore, the few key predictors common to both horizons indicates different dominant controls on C turnover between surface and subsurface horizons, which we posit is due to differences in the soil environment, SOM characteristics, and microbial communities. A horizons have SOM with characteristics intermediate between organic matter in litter layers and SOM in deeper mineral horizons, while B horizons have more degraded SOM that is potentially more likely to be sorbed to mineral surfaces. These insights underscore the importance of the vertical heterogeneity of soils in understanding SOM dynamics and the challenge this presents for modeling efforts.

cores. Additionally, we would like to thank Jason Toyoda and Rosalie Chu at EMSL and Stephanie Duston and Dave Mitchem at Virginia Tech for assistance with laboratory analyses.

Data availability A subset of the data used in this analysis is available at https://doi.org/10.6073/pasta/99d113534ecaa0 4e820850e6169be04d; the remainder of the data are available upon request.

Code availability R code for processing and analyzing data is available at https://github.com/tweiglein/key_predictors_SOM_ vulnerability.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Berg B, Steffen KT, McClaugherty C (2007) Litter decomposition rate is dependent on litter Mn concentrations. Biogeochemistry 82:29–39. https://doi.org/10.1007/s10533-006-9050-6

Bradford MA, Wieder WR, Bonan GB et al (2016) Managing uncertainty in soil carbon feedbacks to climate change. Nat Clim Chang 6:751–758. https://doi.org/10.1038/nclimate3071

Bramer LM, White AM, Stratton KG et al (2020) ftmsRanalysis: an R package for exploratory data analysis and interactive visualization of FT-MS data. PLoS Comput Biol 16:e1007654. https://doi.org/10.1371/journal.pcbi.1007654

Brewer TE, Aronson EL, Arogyaswamy K et al (2019) Ecological and genomic attributes of novel bacterial taxa that thrive in subsurface soil horizons. Mbio. https://doi.org/10.1128/mBio.01318-19

Brockett BFT, Prescott CE, Grayston SJ (2012) Soil moisture is the major factor influencing microbial community structure and enzyme activities across seven biogeoclimatic zones in western Canada. Soil Biol Biochem 44:9–20. https://doi.org/10.1016/j.soilbio.2011.09.003

Cory RM, McKnight DM (2005) Fluorescence spectroscopy reveals ubiquitous presence of oxidized and reduced quinones in dissolved organic matter. Environ Sci Technol 39:8142–8149. https://doi.org/10.1021/es0506962

Cory RM, Miller MP, McKnight DM et al (2010) Effect of instrument-specific response on the analysis of fulvic acid fluorescence spectra. Limnol Oceanogr Methods 8:67–78. https://doi.org/10.4319/lom.2010.8.67

Creamer CA, de Menezes AB, Krull ES et al (2015) Microbial community structure mediates response of soil C decomposition to litter addition and warming. Soil Biol Biochem 80:175–188. https://doi.org/10.1016/j.soilbio.2014.10.008

DeCiucies S, Whitman T, Woolf D et al (2018) Priming mechanisms with additions of pyrogenic organic matter to soil. Geochim Cosmochim Acta 238:329–342. https://doi.org/10.1016/j.gca.2018.07.004

Degenhardt F, Seifert S, Szymczak S (2019) Evaluation of variable selection methods for random forests and omics data sets. Brief Bioinform 20:492–503. https://doi.org/10.1093/bib/bbx124

Doetterl S, Stevens A, Six J et al (2015) Soil carbon storage controlled by interactions between geochemistry and climate. Nat Geosci 8:780–783. https://doi.org/10.1038/ngeo2516

Dove NC, Arogyaswamy K, Billings SA et al (2020) Continental-scale patterns of extracellular enzyme activity in the subsoil: an overlooked reservoir of microbial activity Environmental Research Letters Continental-scale patterns of extracellular enzyme activity in the subsoil: an overlooked reservoir of microbial activity. Environ Res Lett 15:1040–1041. https://doi.org/10.1088/1748-9326/abb0b3

Drenovsky RE, Steenwerth KL, Jackson LE, Scow KM (2010) Land use and climatic factors structure regional patterns in soil microbial communities. Glob Ecol Biogeogr 19:27–39. https://doi.org/10.1111/j.1466-8238.2009.00486.x

Eilers KG, Debenport S, Anderson S, Fierer N (2012) Digging deeper to find unique microbial communities: the strong effect of depth on the structure of bacterial and archaeal communities in soil. Soil Biol Biochem 50:58–65. https://doi.org/10.1016/j.soilbio.2012.03.011

Fierer N, Schimel JP, Holden PA (2003) Variations in microbial community composition through two soil depth profiles. Soil Biol Biochem 35:167–176. https://doi.org/10.1016/S0038-0717(02)00251-1

Fisher P, Aumann C, Chia K et al (2017) Adequacy of laser diffraction for soil particle size analysis. PLoS ONE 12:e0176510. https://doi.org/10.1371/journal.pone.0176510

Golchin A, Oades J, Skjemstad J, Clarke P (1994) Study of free and occluded particulate organic matter in soils by solid state 13C CP/MAS NMR spectroscopy and scanning electron microscopy. Aust J Soil Res 32:285–309. https://doi.org/10.1071/SR9940285

Goñi MA, Montgomery S (2000) Alkaline CuO oxidation with a microwave digestion system: lignin analyses of

geochemical samples. Anal Chem 72:3116–3121. https://doi.org/10.1021/ac991316w

Graham EB, Tfaily MM, Crump AR et al (2017) Carbon inputs from riparian vegetation limit oxidation of physically bound organic carbon via biochemical and thermodynamic processes. J Geophys Res Biogeosci 122:3188–3205. https://doi.org/10.1002/2017JG003967

Gray JM, Bishop TFA, Wilson BR (2015) Factors controlling soil organic carbon stocks with depth in eastern Australia. Soil Sci Soc Am J 79:1741–1751. https://doi.org/10.2136/sssaj2015.06.0224

Guenet B, Lenhart K, Leloup J et al (2012) The impact of long-term CO2 enrichment and moisture levels on soil microbial community structure and enzyme activities. Geoderma 170:331–336. https://doi.org/10.1016/j.geoderma.2011.12.002

Hagerty SB, Jan van Groenigen K, Allison SD et al (2014) Accelerated microbial turnover but constant growth efficiency with warming in soil. Nat Clim Change 4:903–906. https://doi.org/10.1038/NCLIMATE2361

Härdle WK, Simar L (2019) Applied multivariate statistical analysis, 5th edn. Springer, New York

Hargreaves GH, Samani ZA (1982) Estimating potential evapotranspiration. J Irrig Drain Div 108:225–230. https://doi.org/10.1061/JRCEA4.0001390

Harris D, Horwáth WR, van Kessel C (2001) Acid fumigation of soils to remove carbonates prior to total organic carbon or carbon-13 isotopic analysis. Soil Sci Soc Am J 65:1853–1856. https://doi.org/10.2136/sssaj2001.1853

Hatten JA, Goñi M (2016) Cupric oxide (CuO) oxidation detects pyrogenic carbon in burnt organic matter and soils. PLoS ONE 11:e0151957. https://doi.org/10.1371/journal.pone.0151957

Hatten JA, Goñi MA, Wheatcroft RA (2012) Chemical characteristics of particulate organic matter from a small, mountainous river system in the Oregon Coast Range, USA. Biogeochemistry 107:43–66. https://doi.org/10.1007/s10533-010-9529-z

Heckman KA, Nave LE, Bowman M, et al (2021) Divergent controls on carbon concentration and persistence between forests and grasslands of the conterminous US. Biogeochemistry 156:41–56. https://doi.org/10.1007/s10533-020-00725-z

Hennig C, Meila M, Murtagh F, Rocci R (2015) Handbook of cluster analysis. Chapman and Hall/CRC, New York

Hernes PJ, Robinson AC, Aufdenkampe AK (2007) Fractionation of lignin during leaching and sorption and implications for organic matter "freshness." Geophys Res Lett. https://doi.org/10.1029/2007GL031017

Huguet A, Vacher L, Relexans S et al (2009) Properties of fluorescent dissolved organic matter in the Gironde Estuary. Org Geochem 40:706–719. https://doi.org/10.1016/j.orggeochem.2009.03.002

Illmer P, Marschall K, Schinner F (1995) Influence of available aluminium on soil micro-organisms. Lett Appl Microbiol 21:393–397. https://doi.org/10.1111/j.1472-765X.1995.tb01090.x

Illmer P, Obertegger U, Schinner F (2003) Microbiological properties in acidic forest soils with special consideration of KCl extractable Al. Water Air Soil Pollut 148:3–14. https://doi.org/10.1023/A:1025422229468

Jobbagy EG, Jackson RB (2000) The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecol Appl 10:423–436. https://doi.org/10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2

Kaiser K, Guggenberger G, Haumaier L (2004) Changes in dissolved lignin-derived phenols, neutral sugars, uronic acids, and amino sugars with depth in forested Haplic Arenosols and Rendzic Leptosols. Biogeochemistry 70:135–151. https://doi.org/10.1023/B:BIOG.0000049340.77963.18

Kallenbach CM, Frey SD, Grandy AS (2016) Direct evidence for microbial-derived soil organic matter formation and its ecophysiological controls. Nat Commun. https://doi.org/10.1038/ncomms13630

Keiluweit M, Nico P, Harmon ME et al (2015) Long-term litter decomposition controlled by manganese redox cycling. Proc Natl Acad Sci USA 112:E5253–E5260. https://doi.org/10.1073/pnas.1508945112

Koch BP, Dittmar T (2006) From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. Rapid Commun Mass Spectrom 20:926–932. https://doi.org/10.1002/rcm.2386

Koch BP, Dittmar T (2016) Erratum: From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter (Rapid Communications in Mass Spectrometry (2006) 20 (926–932) https://doi.org/10.1002/rcm.2386). Rapid Commun Mass Spectrom 30:250. https://doi.org/10.1002/rcm.7433

Köchy M, Hiederer R, Freibauer A (2015) 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. https://doi.org/10.5194/soil-1-351-2015

Kögel-Knabner I, Guggenberger G, Kleber M et al (2008) Organo-mineral associations in temperate soils: integrating biology, mineralogy, and organic matter chemistry. J Plant Nutr Soil Sci 171:61–82. https://doi.org/10.1002/jpln.200700048

Kramer MG, Chadwick OA (2018) Climate-driven thresholds in reactive mineral retention of soil carbon at the global scale. Nat Clim Change 8:1104–1108. https://doi.org/10.1038/s41558-018-0341-4

Kunito T, Isomura I, Sumi H et al (2016) Aluminum and acidity suppress microbial activity and biomass in acidic forest soils. Soil Biol Biochem 97:23–30. https://doi.org/10.1016/j.soilbio.2016.02.019

Kursa MB, Rudnicki WR (2010) Feature selection with the Boruta package. J Stat Softw 36:1–13. https://doi.org/10.18637/jss.v036.i11

Lehmann J, Kleber M (2015) The contentious nature of soil organic matter. Nature 528:60–68. https://doi.org/10.1038/nature16069

Luo W, Li MH, Sardans J et al (2017) Carbon and nitrogen allocation shifts in plants and soils along aridity and fertility gradients in grasslands of China. Ecol Evol 7:6927–6934. https://doi.org/10.1002/ece3.3245

Matosziuk LM, Alleau Y, Kerns BK et al (2019) Effects of season and interval of prescribed burns on pyrogenic carbon in ponderosa pine stands in the southern Blue Mountains, Oregon, USA. Geoderma 348:1–11. https://doi.org/10.1016/j.geoderma.2019.04.009

Matosziuk LM, Gallo A, Hatten J et al (2020) Short-term effects of recent fire on the production and translocation of pyrogenic carbon in great smoky mountains National Park. Front Glob Change 3:6. https://doi.org/10.3389/ffgc.2020.00006

McKnight DM, Boyer EW, Westerhoff PK et al (2001) Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. Limnol Oceanogr 46:38–48. https://doi.org/10.4319/lo.2001.46.1.0038

Miller MP, McKnight DM (2010) Comparison of seasonal changes in fluorescent dissolved organic matter among aquatic lake and stream sites in the Green Lakes Valley. J Geophys Res. https://doi.org/10.1029/2009jg000985

Mullins CE (1977) Magnetic susceptibility of the soil and its significance in soil science: a review. J Soil Sci 28:223–246. https://doi.org/10.1111/j.1365-2389.1977.tb02232.x

Murphy KP (2012) Machine learning: a probabilistic perspective. MIT Press, Cambridge

Nave LE, Bowman M, Gallo A et al (2021) Patterns and predictors of soil organic carbon storage across a continental-scale network. Biogeochemistry 156:75–96. https://doi.org/10.1007/s10533-020-00745-9

NEON (2019) Data product: NEON.DP1.00096.001

Ohno T (2002) Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter. Environ Sci Technol 36:742–746. https://doi.org/10.1021/es0155276

ORNL DAAC (2018) Fixed sites subsetting and visualization tool

Pansu M, Gautheyrou J (2006) Mineralogical separation by selective dissolution. In: Handbook of soil analysis: mineralogical, organic and inorganic methods. Springer Science & Business Media, New York pp 167–219

Paul EA, Morris SJ, Conant RT, Plante AF (2006) Does the acid hydrolysis–incubation method measure meaningful soil organic carbon pools? Soil Sci Soc Am J 70:1023–1035. https://doi.org/10.2136/sssaj2005.0103

PRISM Climate Group (2018) PRISM climate data. http://www.prismclimate.org

Quesada CA, Paz C, Oblitas Mendoza E et al (2020) Variations in soil chemical and physical properties explain basin-wide Amazon forest soil carbon concentrations. Soil 6:53–88. https://doi.org/10.5194/soil-6-53-2020

R Core Team (2019) R: a language and environment for statistical computing

Rasmussen C, Southard RJ, Horwath WR (2006) Mineral control of organic carbon mineralization in a range of temperate conifer forest soils. Glob Change Biol 12:834–847. https://doi.org/10.1111/j.1365-2486.2006.01132.x

Rasmussen C, Heckman K, Wieder WR et al (2018) Beyond clay: towards an improved set of variables for predicting soil organic matter content. Biogeochemistry 137:297–306. https://doi.org/10.1007/s10533-018-0424-3

RStudio Team (2018) RStudio: integrated development for R

Rumpel C, Kögel-Knabner I (2011) Deep soil organic matter: a key but poorly understood component of terrestrial C cycle. Plant Soil 338:143–158. https://doi.org/10.1007/s11104-010-0391-5

Running S, Mu Q, Zhao M (2015) MOD17A3H MODIS/Terra net primary production yearly L4 global 500m SIN Grid V006. NASA EOSDIS L Process DAAC. https://doi.org/10.5067/MODIS/MOD17A3H.006

SanClements MD, Oelsner GP, McKnight DM et al (2012) New insights into the source of decadal increases of dissolved organic matter in acid-sensitive lakes of the Northeastern United States. Environ Sci Technol 46:3212–3219. https://doi.org/10.1021/es204321x

Santiago LS, Mulkey SS, Santiago LS, Mulkey SS (2005) Leaf productivity along a precipitation gradient in lowland Panama: patterns from leaf to ecosystem. Trees 19:349–356. https://doi.org/10.1007/s00468-004-0389-9

Saxton K, Rawls WJ, Romberger J, Papendick R (1986) Estimating generalized soil-water characteristics from texture. Soil Sci Soc Am J 50:1031–1036. https://doi.org/10.1017/CBO9781107415324.004

Schmidt MWI, Torn MS, Abiven S et al (2011) Persistence of soil organic matter as an ecosystem property. Nature 478:49–56. https://doi.org/10.1038/nature10386

Schoeneberger PJ, Wysocki DA, Benham EC, Soil Survey Staff (2012) Field book for describing and sampling soils, version 3.0. Natural Resources Conservation Service, National Soil Survey Center, Lincoln

Schuur EAG (2003) Productivity and global climate revisited: the sensitivity of tropical forest growth to precipitation. Ecology 84:1165–1170. https://doi.org/10.1890/0012-9658(2003)084[1165:PAGCRT]2.0.CO;2

Schuur EAG, Matson PA (2001) Net primary productivity and nutrient cycling across a mesic to wet precipitation gradient in Hawaiian montane forest. Oecologia 128:431–442

Shang C, Tiessen H (2000) Carbon turnover and carbon-13 natural abundance in organo-mineral fractions of a tropical dry forest soil under cultivation. Soil Sci Soc Am J 64:2149–2155. https://doi.org/10.2136/sssaj2000.6462149x

Soil Survey Staff (2014) Kellogg soil survey laboratory methods manual. Soil Survey Investigations Report No. 42, Version 5.0. U.S. Department of Agriculture, Natural Resources Conservation Service

Soil Survey Staff (2017) Web soil survey

Stone MM, DeForest JL, Plante AF (2014) Changes in extracellular enzyme activity and microbial community structure with soil depth at the Luquillo Critical Zone Observatory. Soil Biol Biochem 75:237–247. https://doi.org/10.1016/j.soilbio.2014.04.017

Strahm BD, Harrison RB (2008) Controls on the sorption, desorption and mineralization of low-molecular-weight organic acids in variable-charge soils. Soil Sci Soc Am J 72:1653–1664. https://doi.org/10.2136/sssaj2007.0318

Strickland TC, Sollins P (1987) Improved method for separating light- and heavy-fraction organic material from soil. Soil Sci Soc Am J 51:1390–1393. https://doi.org/10.2136/sssaj1987.03615995005100050056x

Swanston CW, Caldwell BA, Homann PS et al (2002) Carbon dynamics during a long-term incubation of separate and recombined density fractions from seven forest soils. Soil Biol Biochem 34:1121–1130. https://doi.org/10.1016/S0038-0717(02)00048-2

Swanston CW, Torn MS, Hanson PJ et al (2005) Initial characterization of processes of soil carbon stabilization using

forest stand-level radiocarbon enrichment. Geoderma 128:52–62. https://doi.org/10.1016/J.GEODERMA.2004.12.015

Tamir G, Shenker M, Heller H et al (2011) Can soil carbonate dissolution lead to overestimation of soil respiration? Soil Sci Soc Am J 75:1414–1422. https://doi.org/10.2136/sssaj2010.0396

Tarnocai C, Canadell JG, Schuur EAG et al (2009) Soil organic carbon pools in the northern circumpolar permafrost region. Glob Biogeochem Cycles 23:1–11. https://doi.org/10.1029/2008GB003327

Tatsumi C, Taniguchi T, Du S et al (2019) The steps in the soil nitrogen transformation process vary along an aridity gradient via changes in the microbial community. Biogeochemistry 144:15–29. https://doi.org/10.1007/s10533-019-00569-2

Tfaily MM, Chu RK, Tolić N et al (2015) Advanced solvent based methods for molecular characterization of soil organic matter by high-resolution mass spectrometry. Anal Chem 87:5206–5215. https://doi.org/10.1021/acs.analchem.5b00116

Tfaily MM, Chu RK, Toyoda J et al (2017) Sequential extraction protocol for organic matter from soils and sediments using high resolution mass spectrometry. Anal Chim Acta 972:54–61. https://doi.org/10.1016/j.aca.2017.03.031

Tolić N, Liu Y, Liyu A et al (2017) Formularity: software for automated formula assignment of natural and other organic matter from ultrahigh-resolution mass spectra. Anal Chem 89:12659–12665. https://doi.org/10.1021/acs.analchem.7b03318

Torn MS, Trumbore SE, Chadwick OA et al (1997) Mineral control of soil organic carbon storage and turnover. Nature 389:170–173. https://doi.org/10.1038/38260

Torn MS, Vitousek PM, Trumbore SE (2005) The influence of nutrient availability on soil organic matter turnover estimated by incubations and radiocarbon modeling. Ecosystems 8:352–372. https://doi.org/10.1007/s10021-004-0259-8

Torn MS, Swanston CW, Castanha C, Trumbore SE (2009) Storage and turnover of organic matter in soil. In: Senesi N, Xing B, Haung PM (eds) Biophysico-chemical processes involving natural nonliving organic matter in environmental systems. Wiley, Hoboken, pp 219–272

van Buuren S, Groothuis-Oudshoorn K (2011) mice: multivariate imputation by chained equations in R. J Stat Softw 45:1–67. https://doi.org/10.18637/jss.v045.i03

Viscarra Rossel RA, Lee J, Behrens T et al (2019) Continental-scale soil carbon composition and vulnerability modulated by regional environmental controls. Nat Geosci 12:547–552. https://doi.org/10.1038/s41561-019-0373-z

von Lützow M, Kögel-Knabner I, Ekschmitt K et al (2006) Stabilization of organic matter in temperate soils: mechanisms and their relevance under different soil conditions: a review. Eur J Soil Sci 57:426–445. https://doi.org/10.1111/j.1365-2389.2006.00809.x

von Lützow M, Kögel-Knabner I, Ekschmitt K et al (2007) SOM fractionation methods: relevance to functional pools and to stabilization mechanisms. Soil Biol Biochem 39:2183–2207. https://doi.org/10.1016/J.SOILBIO.2007.03.007

von Fromm SF, Hoyt AM, Lange M, et al (2021) Continental-scale controls on soil organic carbon across sub-Saharan Africa. SOIL 7:305–332. https://doi.org/10.5194/soil-7-305-2021

Wada K (1989) Allophane and imogolite. Minerals in soil environments. Soil Science Society of America, Madison, pp 1051–1087

Wang T, Hamann A, Spittlehouse DL, Murdock TQ (2012) ClimateWNA: high-resolution spatial climate data for western North America. J Appl Meteorol Climatol 51:16–29. https://doi.org/10.1175/JAMC-D-11-043.1

Whalen JK, Bottomley PJ, Myrold DD (2000) Carbon and nitrogen mineralization from light- and heavy-fraction additions to soil. Soil Biol Biochem 32:1345–1352. https://doi.org/10.1016/S0038-0717(00)00040-7

Wiedemeier DB, Hilf MD, Smittenberg RH et al (2013) Improved assessment of pyrogenic carbon quantity and quality in environmental samples by high-performance liquid chromatography. J Chromatogr A 1304:246–250. https://doi.org/10.1016/j.chroma.2013.06.012

Wiesmeier M, Urbanski L, Hobley E et al (2019) Soil organic carbon storage as a key function of soils: a review of drivers and indicators at various scales. Geoderma 333:149–162. https://doi.org/10.1016/j.geoderma.2018.07.026

Wilson HF, Xenopoulos MA (2009) Effects of agricultural land use on the composition of fluvial dissolved organic matter. Nat Geosci 2:37–41. https://doi.org/10.1038/ngeo391

Yang Y, Wang L, Wendroth O et al (2019) Is the laser diffraction method reliable for soil particle size distribution analysis? Soil Sci Soc Am J 83:276–287. https://doi.org/10.2136/sssaj2018.07.0252

Yu W, Weintraub SR, Hall SJ (2021) Climatic and geochemical controls on soil carbon at the continental scale: interactions and thresholds. Glob Biogeochem Cycles. https://doi.org/10.1029/2020gb006781

Zsolnay A, Baigar E, Jimenez M et al (1999) Differentiating with fluorescence spectroscopy the sources of dissolved organic matter in soils subjected to drying. Chemosphere 38:45–50. https://doi.org/10.1016/S0045-6535(98)00166-0