Biogeochemistry (2025) 168:20
https://doi.org/10.1007/s10533-025-01208-9

MANUSCRIPT

Moisture and soil depth govern relationships between soil organic carbon and oxalate-extractable metals at the global scale

Sophie F. von Fromm · Hermann F. Jungkunst · Bright Amenkhienan · Steven J. Hall · Katerina Georgiou · Caitlin Hicks Pries · Fernando Montaño-López · Carlos Alberto Quesada · Craig Rasmussen · Marion Schrumpf · Balwant Singh · Aaron Thompson · Rota Wagai · Sabine Fiedler

Received: 31 May 2024 / Accepted: 6 January 2025
© The Author(s) 2025

Abstract An important control on long-term soil organic carbon (SOC) storage is the adsorption of SOC by short-range-ordered (SRO) minerals. SRO are commonly quantified by measuring oxalate-extractable metals ($M_{ox} = Al_{ox} + ½\ Fe_{ox}$), which many studies have shown to be positively correlated with SOC. It remains uncertain if this organo-mineral relationship is robust at the global scale, or if capturing regional differences is needed to maximize model accuracy. We used a global synthesis of Al_{ox} and Fe_{ox} data to test their role in controlling SOC abundance across regions. We compiled 37,344 individual soil horizon measurements, with soil depth ranging between 0 and 200 cm, from 11,122 profiles. We used the Holdridge Life Zones, which are characterized by biotemperature, precipitation, and potential evapotranspiration, to group the soil profiles by their climatic conditions that also correlate with other important soil-forming factors. Based on linear mixed-effects models, we found a positive relationship between M_{ox} and SOC across regions and depths, accounting for 49% of the SOC variation.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10533-025-01208-9.

Responsible Editor: Justin Richardson.

S. F. von Fromm (✉) · C. H. Pries · F. Montaño-López
Dartmouth College, Hanover, NH, USA
e-mail: sfromm@dartmouth.edu

S. F. von Fromm · M. Schrumpf
Max-Planck-Institute for Biogeochemistry, Jena, Germany

H. F. Jungkunst
RPTU Kaiserslautern-Landau, Landau, Germany

B. Amenkhienan · B. Singh
The University of Sydney, Sydney, Australia

S. J. Hall
University of Wisconsin-Madison, Madison, WI, USA

K. Georgiou
Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, USA

C. A. Quesada
National Institute for Amazonian Research (INPA), Manaus, Brazil

C. Rasmussen
University of Arizona, Tucson, AZ, USA

A. Thompson
University of Georgia, Athens, GA, USA

R. Wagai
National Institute for Agro-Environmental Sciences (NIAES), Tsukuba, Ibaraki, Japan

R. Wagai
National Agriculture and Food Research Organization (NARO), Tsukuba, Ibaraki, Japan

S. Fiedler
Johannes Gutenberg University Mainz, Mainz, Germany

Published online: 07 February 2025

This relationship is strongest in wetter regions and at depths between 20 and 100 cm. Across all environmental conditions, Al_{ox} is a stronger predictor of SOC than Fe_{ox}. Our analysis suggests oxalate-extractable metals are good proxies for mineral-induced SOC protection at the global scale. However, our findings also indicate that the importance of organo-mineral interactions at the global scale varies with climatic conditions and depth. The underlying mechanisms need to be considered when incorporating these relationships as proxies for mineral sorption capacity into soil C models.

Keywords Soil organic carbon dynamics · Organo-mineral interaction · Poorly-crystalline minerals · Large-scale analysis · Holdridge Life Zones · Soil carbon modeling

Introduction

Soil carbon–climate feedbacks are a major source of uncertainty in predicting how the terrestrial biosphere will respond to climate change (Todd-Brown et al. 2013). Our understanding of soils suggests that the accuracy of soil organic carbon (SOC) predictions depends on environmental conditions and their interaction with other soil properties. In this context, the classical soil-forming factors—climate, organisms, relief, parent material and time (Dokuchaev 1883; Jenny 1941)—influence how different soil processes and their interactions control the abundance and persistence of SOC (Cotrufo & Lavallee 2022). Thus, to improve soil carbon–climate feedback models, we need to incorporate mechanistic soil understanding rather than relying solely on universal global relationships.

Many biogeochemical models rely on soil texture to modify rates of SOC turnover and CO_2 fluxes to the atmosphere (Wieder et al. 2018), without considering the dependence of soil texture on soil-forming factors. In fact, many large-scale studies have shown that other soil properties predict SOC abundance better than texture under many environmental conditions (e.g., Quesada et al. 2020; Rasmussen et al. 2018; von Fromm et al. 2021; Yu et al. 2021). In particular, aluminum (Al) and iron (Fe) species—ranging from monomeric metals to short-range ordered (SRO) and crystalline mineral phases—can play an important role in SOC protection by sorption or formation of insoluble organo-metal coprecipitates, or by promoting subsequent aggregate formation (Oades 1988; Parfitt & Childs 1988; Tisdall & Oades 1982; Wagai & Mayer 2007). These organo-mineral interactions are likely governed by a hierarchy of soil-forming factors that influence SOC persistence and turnover.

Strong positive relationships between SOC abundance and oxalate-extractable forms of Al (Al_{ox}) and Fe (Fe_{ox}) have been found at spatial scales ranging from pedons to continents (Hall & Thompson 2022; Kaiser et al. 2002; Kleber et al. 2005; Masiello et al. 2004; Percival et al. 2000; Powers & Schlesinger 2002; Rasmussen et al. 2018; von Fromm et al. 2021, 2025; Wagai et al. 2020; Yu et al. 2017). These relationships are particularly strong under wet and acidic soil conditions, with acidity and moisture serving as proxies for weathering. This suggests that climate and time are dominant drivers of organo-mineral interactions at the global scale, as they govern weathering processes and influence vegetation patterns (i.e., C inputs and quality). While parent material and relief also contribute, their impact is more challenging to quantify at the global scale, as SRO minerals and aggregates can form from a wide range of parent materials (Slessarev et al. 2022), and relief is often related to time or climate through processes like erosion and elevation changes. In addition, proxies for these soil-forming factors are generally of lower quality than climate at the global scale, which precludes their use in large-scale meta-analyses.

Despite the influence of soil-forming factors, there is additional complexity in the relationships between SOC and Al_{ox} or Fe_{ox}. Various studies have found that Al_{ox} is a stronger predictor of SOC than Fe_{ox} (Hall & Thompson 2022; Rasmussen et al. 2018; Souza et al. 2017; von Fromm et al. 2021; Yu et al. 2021). This may be due to differences in how Al and Fe phases extracted by the oxalate-ammonium method interact with SOC under different environmental conditions (Hall & Thompson 2022; Rennert 2018). In some cases, Fe_{ox} is weakly correlated or even negatively associated with SOC (Hall et al. 2015; Kahle et al. 2002; Percival et al. 2000; Powers & Schlesinger 2002). The complex interplay between monomeric organic complexes, SRO minerals, and environmental variables complicates our understanding of these relationships (Lawrence et al. 2015; Masiello et al. 2004; Wagai et al. 2020). It remains unclear whether Al and Fe phases directly drive SOC retention, or whether

their correlation with SOC reflects enhanced weathering driven by biological activities, such as organic acid production by roots and microbes (Chorover 2022; Wagai et al. 2023). This raises the possibility that SOC inputs, primarily influenced by climate, may play a larger role than soil mineral composition in controlling SOC abundance (Hall & Thompson 2022). To test this, it is important to investigate the changes in organo-mineral interactions with soil depth across regions. Globally, the proportion of C associated with soil minerals increases with soil depth, while the influence of microbial and root biomass decreases with depth (Hicks Pries et al. 2023; Jackson et al. 1996; Xu et al. 2013). This suggests that the importance of Al_{ox} and Fe_{ox} should also increase with soil depth. In contrast, von Fromm et al. 2021 and Yu et al. 2021 observed a more or less constant importance of Al_{ox} and Fe_{ox} with depth.

We propose a hierarchical, mechanistic framework to predict soil-climate feedbacks, emphasizing the role of climate and time in shaping organo-mineral interactions. To refine global predictions, we introduce the Holdridge Life Zones as a valuable grouping variable. These zones, which are determined solely by temperature and water availability, effectively capture SOC distribution and persistence across climatic conditions (Jungkunst et al. 2021; Post et al. 1982). Therefore, there is no bias toward other soil-forming factors that might be useful in representing climate, as is commonly done to highlight mountain ranges. Within the Holdridge Life Zones, mountains are specific regions, representing only the change in temperature and precipitation with altitude. However, this framework does not fully account for the role of time, particularly the age of soils, which is difficult to estimate globally. Distinguishing between glacial and loess influenced soils (~ 12,000 years old) and older soils unaffected by the last ice age may help to overcome this limitation.

Here, we investigate the relationship between SOC abundance and Al_{ox} and Fe_{ox} as a function of soil depth and climatic regions at the global scale. We propose ideas about why these relationships may differ between regions, and we discuss the applicability of Al_{ox} and Fe_{ox} as a driver for SOC turnover rates in biogeochemical models. In this work we address the following research questions (RQ):

RQ1: What is the relationship between SOC and Al_{ox} and Fe_{ox} at the global scale?

We hypothesize that Al_{ox} and Fe_{ox} are positively correlated with SOC, with Al_{ox} being the more important predictor, while Fe_{ox} will show a non-linear relationship with SOC.

RQ2: Are the Holdridge Life Zones a good disaggregation approach at the global scale to better understand the role of organo-mineral interactions?

We hypothesize that a disaggregation at the global scale will improve the prediction of SOC abundance by integrating important soil-forming factors that in turn influence soil mineral phases.

RQ3: Are there regional-specific differences in the relationship between SOC and Al_{ox} and Fe_{ox} with soil depth?

We hypothesize that the importance of Al_{ox} and Fe_{ox} in predicting SOC abundance is:

- highest under warm and wet conditions.
- smallest under cold and arid conditions.
- increasing with soil depth.

Methods

For this study, we analyzed 37,344 individual soil samples from 11,122 profiles, ranging in depth from 0 to 200 cm. We used the Holdridge Life Zones to group the soil profiles according to climatic conditions (Jungkunst et al. 2021; Post et al. 1982). Linear mixed-effects models followed by a post-hoc analysis provide statistical support for the investigated relationships between SOC abundance and Al_{ox} and Fe_{ox} across soil depth and regions. Non-linear relationships between Al_{ox}, Fe_{ox} and SOC were further tested with a random forest model.

Data compilation

We compiled globally distributed soil samples that had measurements of SOC, Al_{ox} and Fe_{ox}, as well as sampling depth and location. We included publicly available datasets, such as from the U.S. Department of Agriculture's National Cooperative Soil Survey (NCSS) Soil Characterization Database (http://ncsslabdatamart.sc.egov.usda.gov/, Rasmussen et al. 2018), National Ecological Observatory Network (NEON 2023), Land Use and Coverage Area frame Survey (LUCAS, Fernandez-Ugalde et al. 2022), African Soil Information Service (AfSIS, Vågen

et al. 2021), International Soil Radiocarbon Database (ISRaD, Lawrence et al. 2020) and from various unpublished and published studies (for a complete list of included studies, see von Fromm 2025). Histosols and all organic horizons were excluded, as well as samples with SOC content > 20 wt %. This was done to focus on mineral soils where we expect to find the strongest relationship between SOC, Al_{ox} and Fe_{ox}. In addition, we limited the bottom sampling depth to 200 cm, and grouped the samples into depth bins of 0–20, 20–50, 50–100, and 100–200 cm.

Soil data were paired with the current version of the Holdridge Life Zones, which is provided by the International Institute for Applied Systems Analyses (IIASA) in Laxenburg, Austria, and were retrieved from the FAO GeoNetwork (last updated 2008). If no Holdridge Life Zone could be extracted, the corresponding zone was assigned based on climate data that was either reported in the study or extracted from the WorldClim dataset (Fick & Hijmans 2017). The Holdridge Life Zones are defined by biotemperature, precipitation, and potential evapotranspiration. Biotemperature is based on growing season length and temperature. It is measured as the mean of all annual temperatures, with all temperatures below freezing and above 30 °C adjusted to 0 °C, as most plants are dormant at these temperatures (Lugo et al. 1999). We grouped the Holdridge Life Zones into moisture groups (based on precipitation and potential evapotranspiration): arid (including superarid, perarid and arid), semiarid, subhumid, humid and perhumid (including perhumid and superhumid), and temperature groups (based on biotemperature): (sub-)polar, boreal, cool temperate, warm temperate, subtropical and tropical (Table A1). Note that the original naming of the Holdridge Life Zones includes the dominant vegetation (Jungkunst et al. 2021; Post et al. 1982). However, in some areas this can be misleading, and we use names that are directly related to the predominant climate. For example, polar deserts have neither much evapotranspiration nor precipitation, but because of their relatively low ratio of evapotranspiration to precipitation, they fall into the humid category despite having 'desert' in the name.

Statistical analysis

We used three approaches to test for differences in the concentrations of oxalate-extractable metals and SOC, and their relationships with each other across climate regions and soil depth: i) Kruskal–Wallis test, ii) linear mixed-effects models, and iii) random forest models. Depth and climate were treated as categorical variables with 0–20, 20–50, 50–100, and 100–200 cm as levels for depth and either i) all Holdridge Life Zones (n = 34), ii) Holdridge Life Zones grouped by moisture (n = 5), or iii) Holdridge Life Zones grouped by temperature (n = 6) as levels for climate. For oxalate-extractable metals, we either used Al_{ox} and Fe_{ox} as individual variables/predictors or summed them up to $M_{ox} = Al_{ox} + ½\ Fe_{ox}$. The latter one was especially necessary for the linear mixed-effects models because Al_{ox} and Fe_{ox} are highly correlated in our data ($\rho = 0.7$, p-value < 0.0001) and thus, cannot be used together as predictors in the same model.

Kruskal–Wallis tests allowed us to test for differences in SOC, Al_{ox}, Fe_{ox}, or M_{ox} across climate groups and soil depth, followed by a post-hoc Dunn test with Bonferroni corrections for the p-values. The Bonferroni correction was made for each individual test performed based on the number of unique pairs compared. Both tests were performed using the 'kruskal_test' and 'dunn_test' function from the 'rstatix' R package, respectively (Kassambara 2023b).

Linear mixed-effects models allowed us to test whether the slope between SOC abundance and oxalate-extractable metals (Al_{ox}, Fe_{ox}, or M_{ox}) differed significantly among depth and climate groups. To account for interactions, we included two two-way interactions between oxalate-extractable metals (Al_{ox}, Fe_{ox}, or M_{ox}) by depth groups and by climate groups, and one three-way interaction between all three predictors, respectively. In total, we built seven linear mixed-effects models, three focusing on the role of moisture groups as a climate predictor, each with Al_{ox}, Fe_{ox}, of M_{ox} as a predictor, three focusing on the role of temperature groups as a climate predictor, each with Al_{ox}, Fe_{ox} or M_{ox} as a predictor, and one using all Holdridge Life Zones as a predictor with only M_{ox} as a predictor. Note that for the model with all Holdridge Life Zones, we excluded any Holdridge Life Zone that had less than 10 profiles at any given depth interval (n = 9). All linear mixed-effects modeling was performed using the 'lme' function from the 'nlme' package (Pinheiro et al. 2023).

We fit an initial model with un-transformed variables, but because residuals showed clear deviation

from normality, we log-transformed our continuous response variable (SOC) and predictors (Al_{ox}, Fe_{ox}, and M_{ox}) for all subsequent models. Besides fitting a linear model, the relationship between SOC and the predictors in the original data may not be linear. For each linear mixed-effects model, the random effects were set to allow a random slope and intercept for soil depth (continuous) within each soil profile. Thus, the models allowed for the relationship between SOC and depth to vary by soil core or pit. No spatial autocorrelation was observed in the final model (Online Appendix Fig. A1 and A2). The variation explained by each model was obtained by calculating the marginal R^2 (excluding the variation explained by the random effects) using the 'r.squaredGLMM' function from the 'MuMIn' package (Bartoń, 2024). In addition, we used the Akaike Information Criterion (AIC) to test which model performed best within each climate group to identify whether Al_{ox}, Fe_{ox}, or M_{ox} is the most important predictor. Following the linear mixed-effects modeling, we performed a post-hoc analysis to explore statistically significant interactions between any of the oxalate-extractable metals (Al_{ox}, Fe_{ox} or M_{ox}), depth and climate groups (all Holdridge Life Zones, moisture groups or temperature groups) of the estimated marginal means using the 'emmeans' package (Lenth 2024).

Random forest model was used to further explore potential non-linear relationships between SOC and Al_{ox} and Fe_{ox}, using the 'mlr3' and 'ranger' packages (Lang et al. 2019; Wright & Ziegler 2017). We included the untransformed continuous values of Al_{ox} and Fe_{ox} as well as depth and all the Holdridge Life Zones as categorical values as predictors. For the validation of the resulting model, we performed a tenfold cross-validation, ensuring that each soil profile was either fully in the training (70%) or the test (30%) dataset. Based on the cross-validation, we calculated the R^2 and root mean squared error. To further explore the model, we calculated the permutation feature importance and partial dependence plots, using the 'iml' package (Molnar & Bischl 2018). The partial dependence plots show the marginal effects of a predictor variable on the predicted outcome of the random forest model (Friedman 2001).

All statistical analyses were performed in the R environment (v4.4.1, R Core Team 2024). In addition to R packages mentioned above, the following R packages were used: 'tidyverse' (Wickham et al. 2019), 'RColorBrewer' (Neuwirth 2022), 'ggpubr' (Kassambara 2023a), 'scales' (Wickham et al. 2023). The R code to reproduce all analysis can be found in von Fromm 2025.

Results

The compiled dataset covered the entire range of global environmental conditions, ranging from superarid to superhumid and from polar to tropical (Fig. 1 and Online Appendix Table A1). We calculated the global coverage of each climate region based on the land area they covered and compared this to the coverage based on the number of unique profiles within each climate region used in this study (Fig. 1a). When grouped by the dominant moisture condition, our dataset profile distribution underrepresented arid (only 9% of the fractional distribution of arid regions) and semiarid (70% coverage), and overrepresented subhumid (112%), humid (133%), and perhumid (130%) regions. Grouped by temperature, our data covered 34% of (sub-)polar, 77% of boreal, 296% of cool temperate, 185% of warm temperate, 55% of subtropical and 33% of tropical regions, highlighting the underrepresentation of very cold and hot environments in contrast to the strong overrepresentation of cool and warm temperate environments. Geographically, there is an apparent sampling bias toward the United States of America and Western Europe. Interestingly, humid and perhumid regions were dominated by colder environments globally and in our dataset, while all other moisture groups were dominated by warmer temperatures (Fig. 1b).

Globally, we generally found higher concentrations of SOC and oxalate-extractable metals ($M_{ox} = Al_{ox} + ½\ Fe_{ox}$) under wetter conditions irrespective of soil depth and temperature based on Kruskal–Wallis pairwise comparison ($df = 33$, p-value ≤ 0.0001) followed by Dunn test (Fig. 2, Online Appendix Figs. A3 and A4). A similar pattern occurred for Al_{ox} and Fe_{ox} individually (Online Appendix Figure A5). Note that (sub-)polar subhumid, boreal semiarid, warm temperate perarid, warm temperate superhumid, subtropical arid, and tropical perhumid had less than ten profiles, and their data distribution should be interpreted with caution (Table A1). Overall, M_{ox} tended to vary more among Holdridge Life Zones than SOC, as indicated by

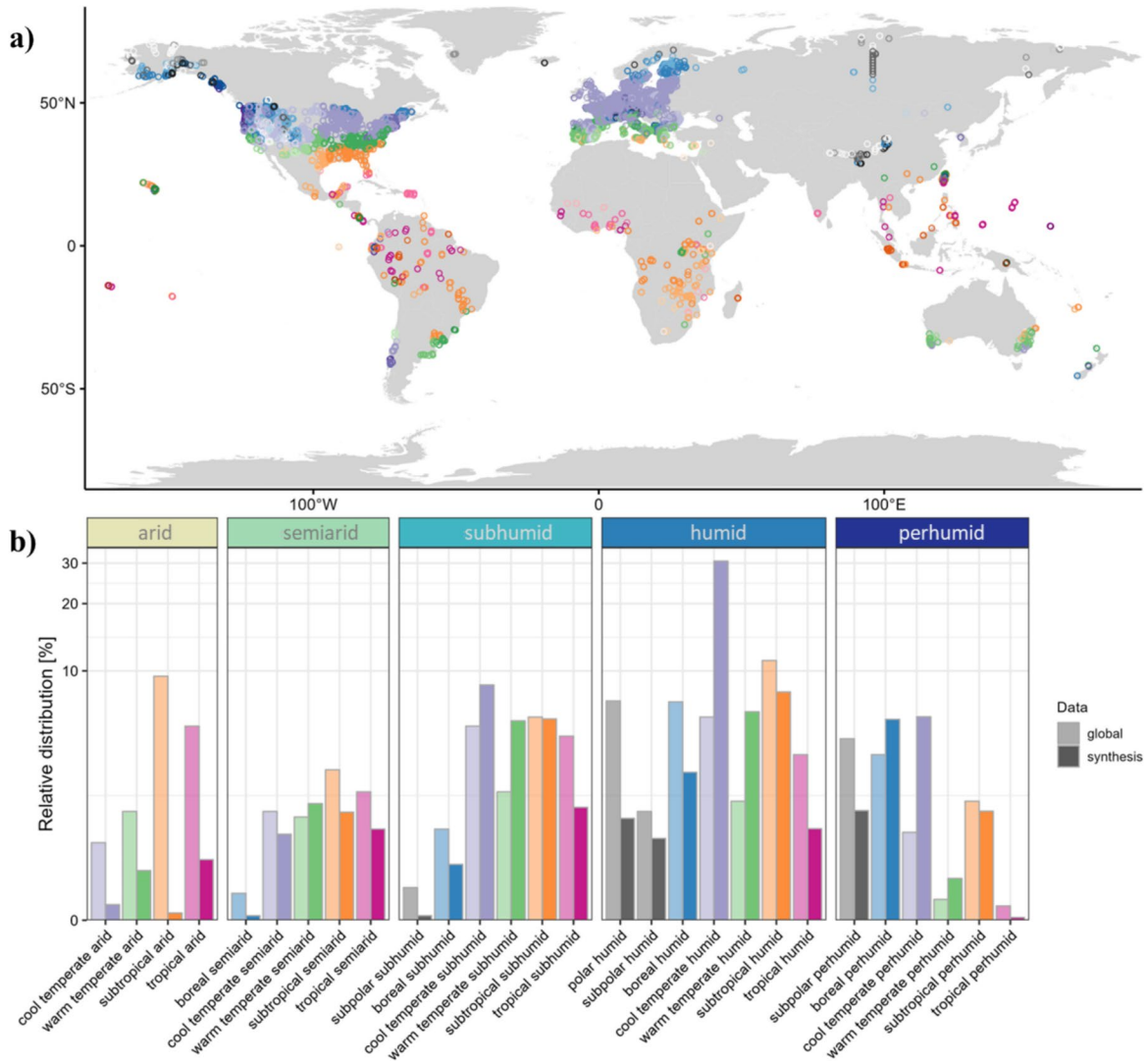

Fig. 1 (**a**) Sampling locations (lighter color within the same temperature regime refers to drier conditions) and (**b**) Relative distributions of all Holdridge Life Zones based on their global coverage (lighter colors of bars) and based on number of unique profiles used in this synthesis (darker colors of bars). Note that the y axis in panel b is on a pseudo-logarithmic scale

the lower number of significant differences between Holdridge Life Zones for SOC (n = 326; based on pairwise comparison) than for M_{ox} (n = 369; Fig. 2, Online Appendix Figs. A3 and A4).

Differences in SOC and M_{ox} were greater across moisture groups than across temperature groups (Online Appendix Fig. A6–9). All moisture groups were significantly different in M_{ox} concentrations, except for arid and semiarid (Kruskal–Wallis test: df = 4, p-value $\leq$ 0.0001 followed by Dunn test; Online Appendix Fig. A6). Within each moisture group, M_{ox} concentrations decreased with soil depth, with the smallest differences in arid regions. Concentrations in SOC were only significantly different across all moisture groups at the surface, while at 100–200 cm, only subhumid and perhumid were significantly different from the other moisture groups. Within each moisture group, SOC concentrations significantly decreased with soil

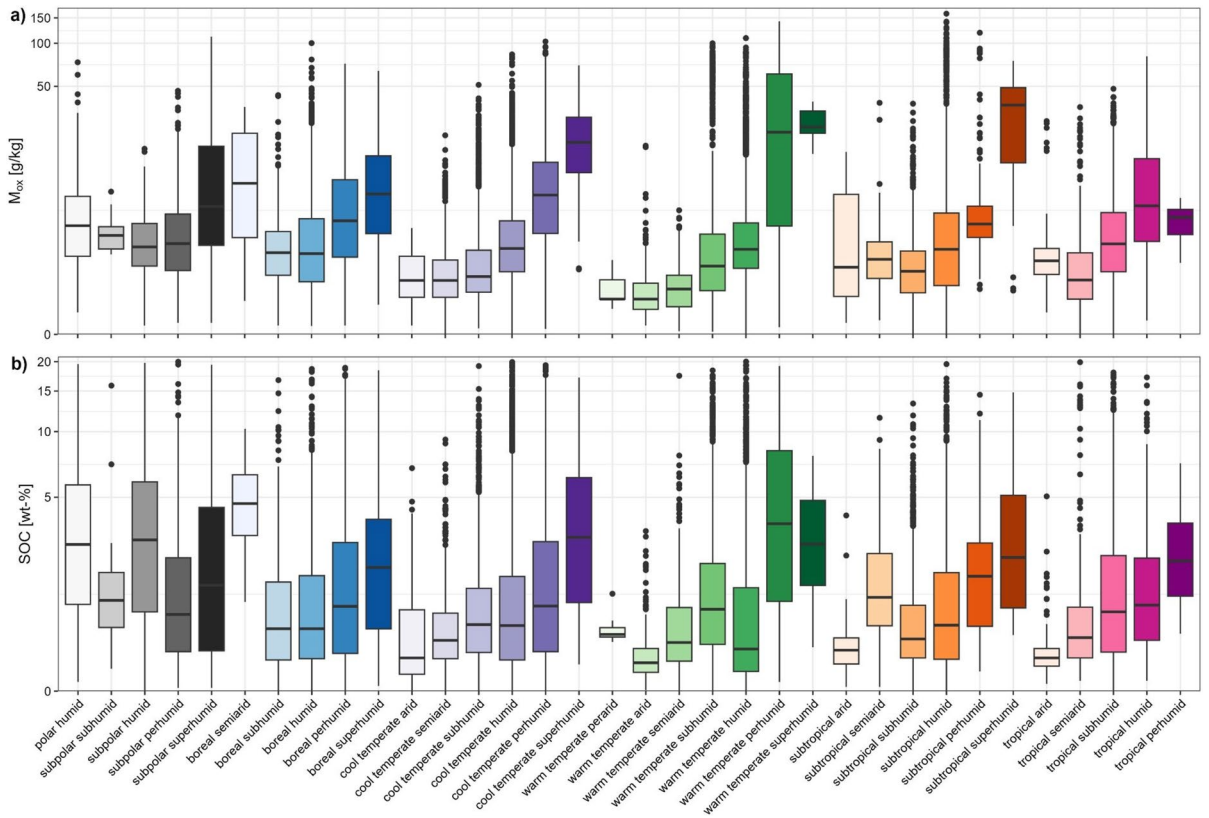

Fig. 2 Concentration of (**a**) oxalate-extractable metals ($M_{ox} = Al_{ox} + ½\ Fe_{ox}$) and (**b**) soil organic carbon (SOC) for each Holdridge Life Zone based on individual measurements across all sampling depths

depth, except for arid regions between 20–50 and 50–100 cm (Online Appendix Fig. A7).

Importance of oxalate-extractable metals across moisture groups

We found that M_{ox} was positively correlated with SOC across moisture groups and depth layers based on the linear mixed-effects models (Fig. 3, Online Appendix Fig. A10 and Online Appendix Table A2). The differences between the linear mixed-effects models with M_{ox}, $Al_{ox,}$ or Fe_{ox} as a predictor were marginal, yet the model with M_{ox} as a predictor performed the best, followed by the model with Al_{ox} as a predictor according to their AIC and F-values (Online Appendix Figs. A11 and A12, Tables A3-4). The M_{ox} model explained about 49% of the variation in SOC with a root mean squared error (RMSE) of 0.42 wt-%. Based on F-values within the model, M_{ox} was the most important predictor, followed by depth, moisture, the interaction between moisture and depth, the interaction between M_{ox} and depth, the interaction between M_{ox} and moisture, and the three-way interaction between M_{ox}, moisture and depth (Online Appendix Table A2). Although the three-way interaction was the least important predictor, we found significant differences in the slopes between M_{ox} and SOC across moisture groups and depth (Fig. 3 and Table 1).

Differences in slopes between M_{ox} and SOC across moisture groups increased with soil depth (Fig. 3 and Table 1). In the surface layer (0–20 cm), semiarid and subhumid regions had the steepest slopes (0.47 and 0.43, respectively), but the second lowest median concentrations of M_{ox} (1.46 and 1.85 g/kg, respectively) and SOC (1.17 and 1.51 wt-%, respectively) after arid regions (Table 1). Perhumid regions had the steepest slope for all other depth layers, with the steepest slope

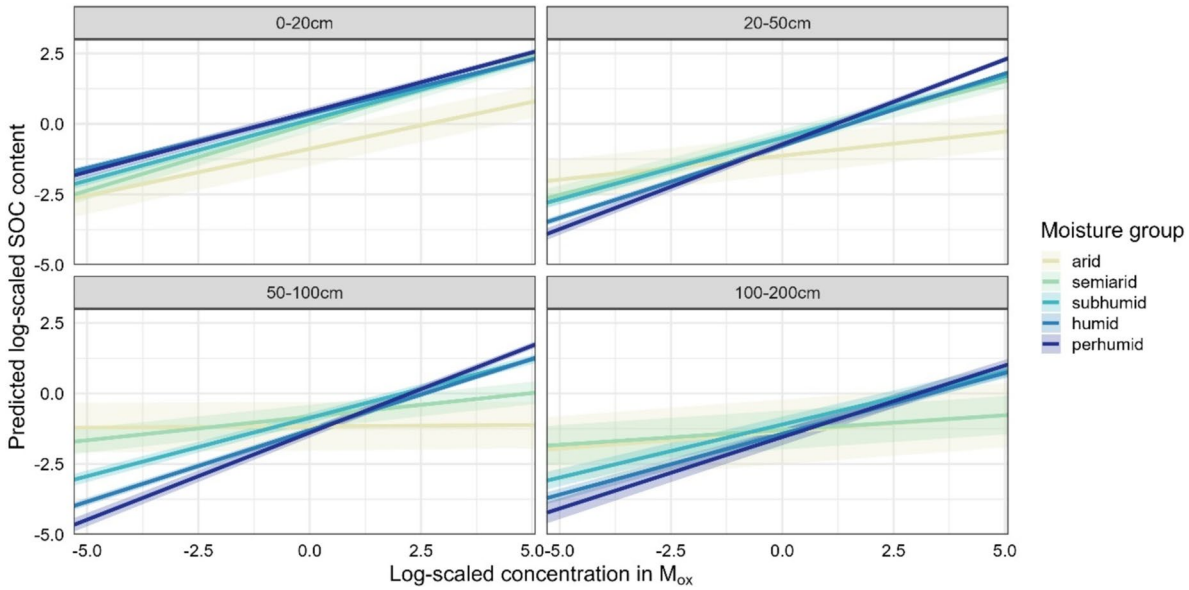

Fig. 3 Relative predicted soil organic carbon (SOC) content based on estimated marginal means for the interaction term M_{ox} (= Al_{ox} + ½ Fe_{ox}) * moisture * depth. Shaded areas indicate 95th confidence intervals. Note that all variables were log-transformed prior analysis

Table 1 Linear mixed-effects model summary statistics for the moisture model. Number of samples, linear predicted slopes, lower and upper confidence intervals (CI; 95th) based on the interaction term M_{ox} (= Al_{ox} + ½ Fe_{ox}) * moisture * depth, and range of M_{ox} (g/kg) and SOC (wt-%) with the median value in parentheses for each moisture group and depth interval, respectively

	Depth	n	Slope	lower CI	upper CI	Range of M_{ox}	Range of SOC
Arid	0–20 cm	226	0.33	0.22	0.45	0.16–25.41 (1.18)	0.05–6.86 (0.40)
	20–50 cm	175	0.17	0.04	0.30	0.16–28.41 (1.15)	0.01–2.65 (0.33)
	50–100 cm	103	0.01	-0.15	0.17	0.16–18.21 (0.86)	0.04–4.18 (0.31)
	100–200 cm	75	0.12	0.10	0.33	0.16–19.9 (0.80)	0.01–0.97 (0.21)
Semiarid	0–20 cm	824	0.47	0.42	0.53	0.09–38.26 (1.46)	0.01–19.9 (1.17)
	20–50 cm	533	0.40	0.35	0.46	0.01–35.8 (1.45)	0.01–14.26 (0.63)
	50–100 cm	292	0.17	0.09	0.25	0.01–35.71 (1.16)	0.01–13.38 (0.41)
	100–200 cm	192	0.10	-0.02	0.23	0.06–22.45 (0.90)	0.01–1.85 (0.28)
Subhumid	0–20 cm	2794	0.43	0.40	0.46	0.01–77.07 (1.85)	0.01–19.16 (1.51)
	20–50 cm	1788	0.43	0.40	0.46	0.01–87.89 (2.16)	0.01–18.00 (0.83)
	50–100 cm	1528	0.41	0.38	0.45	0.01–96.08 (1.71)	0.01–17.40 (0.53)
	100–200 cm	1215	0.38	0.32	0.43	0.01–99.89 (1.41)	0.01–16.30 (0.33)
Humid	0–20 cm	6920	0.39	0.37	0.40	0.01–92.3 (3.05)	0.01–19.98 (2.10)
	20–50 cm	5093	0.51	0.49	0.53	0.01–115.21 (3.75)	0.01–19.78 (0.85)
	50–100 cm	5022	0.51	0.48	0.53	0.01–128.91 (2.76)	0.01–20.00 (0.37)
	100–200 cm	3232	0.43	0.40	0.47	0.01–160.35 (2.11)	0.01–19.20 (0.26)
Perhumid	0–20 cm	2209	0.42	0.40	0.45	0.13–133.51 (7.01)	0.04–19.97 (3.35)
	20–50 cm	2034	0.60	0.57	0.63	0.10–141.91 (8.02)	0.03–19.18 (1.64)
	50–100 cm	1777	0.62	0.58	0.65	0.16–127.55 (5.86)	0.01–18.64 (0.58)
	100–200 cm	1057	0.51	0.45	0.56	0.16–126.16 (4.52)	0.01–9.18 (0.31)

(0.62) between 50 and 100 cm, which also corresponded to the highest median M_{ox} (5.86 g/kg) and SOC concentration (0.58 wt-%) in this depth layer. Below 50 cm, the slope of arid regions was not significantly different from 0, and the same was true for semiarid regions at 100 to 200 cm.

The random forest model and its partial dependence plots revealed that for the range of Al_{ox} and Fe_{ox} where we had most of the data (0 to 100 g/kg), the relationship between the oxalate-extractable metals (Al_{ox} and Fe_{ox}) and predicted SOC content was nearly linear (Online Appendix Fig. A13). The performance of the random forest model ($R^2 = 0.36 \pm 0.05$ and $RMSE = 1.97 \pm 0.05$ wt-%) was comparable to that of the linear mixed-effects models. Variable importance also agreed with the linear mixed-effects models, with Al_{ox} being most important, followed by depth, Fe_{ox} and Holdridge Life Zones (Online Appendix Fig. A14). However, direct comparison of model performance and variable importance should be done with caution because the variables were not transformed for the random forest model as they were for the linear mixed-effects models.

Importance of oxalate-extractable metals across temperature groups

We found a positive relationship between SOC and oxalate-extractable metals across all temperature groups and depth layers (Fig. 4 and Online Appendix Fig. A15). Differences between the models with M_{ox}, Al_{ox} or Fe_{ox} as a predictor were again marginal, and the model with M_{ox} as a predictor performed best, followed by the model with Al_{ox} as a predictor (Online Appendix Figs. A16 and A17, Tables A5-A7). Like the moisture model, the temperature model with M_{ox} as a predictor explained 49% of the variance in SOC and had an RMSE of 0.43 wt-%. Based on F-values, M_{ox} was the most important predictor, followed by depth, temperature, the interaction between M_{ox} and depth, the interaction between temperature and depth, the three-way interaction between M_{ox}, temperature and depth, and the interaction between M_{ox} and temperature (Online Appendix Table A5).

We found less consistent patterns across temperature groups and soil depth based on the slopes between SOC and M_{ox} (Fig. 4 and Table 2). Interestingly, although model performance was similar between the moisture and temperature groups, the

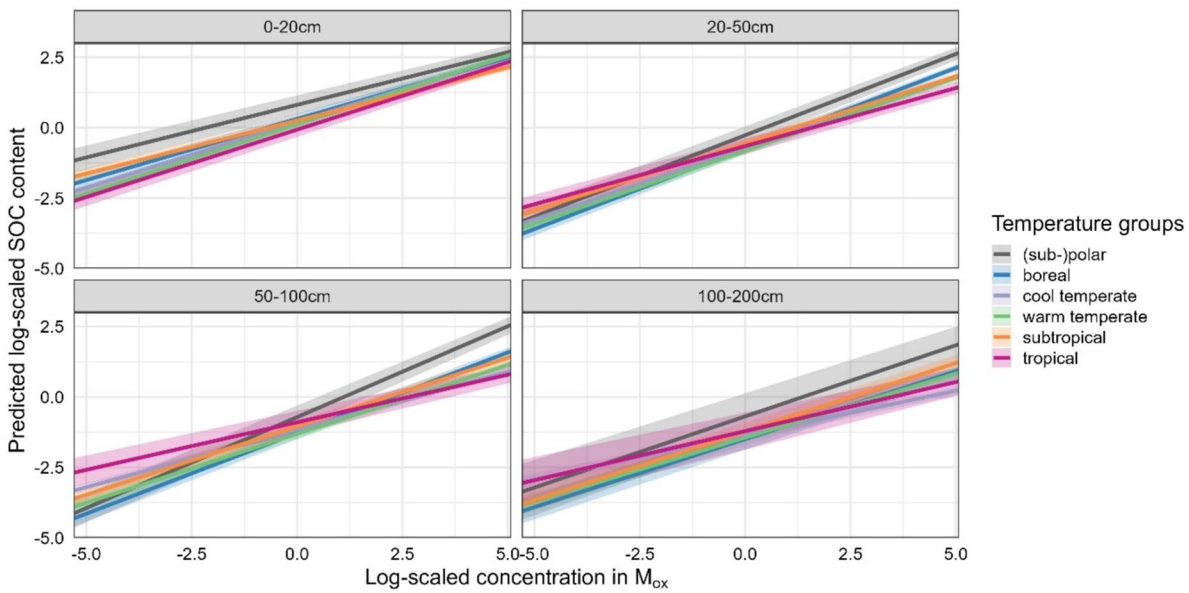

Fig. 4 Relative predicted soil organic carbon (SOC) content based on estimated marginal means for the interaction term M_{ox} ($= Al_{ox} + ½\ Fe_{ox}$) * temperature * depth. Shaded areas indicate 95th confidence intervals. Note that all variables were log-transformed prior analysis

Table 2 Linear mixed-effects model summary statistics for the temperature model. Number of samples (n), linear predicted slopes, lower and upper confidence intervals (CI; 95th) based on the interaction term M_{ox} (= Al_{ox} + ½ Fe_{ox}) * temperature * depth, and range of M_{ox} (g/kg) and SOC (wt-%) with the median value in brackets for each temperature group and depth interval, respectively

	Depth	n	Slope	lower CI	upper CI	Range of M_{ox}	Range of SOC
(Sub-)polar	0–20 cm	602	0.37	0.31	0.44	0.28–111.16 (3.81)	0.09–19.97 (3.43)
	20–50 cm	615	0.58	0.52	0.64	0.36–97.06 (4.66)	0.03–19.78 (1.98)
	50–100 cm	459	0.65	0.57	0.72	0.16–110.36 (3.95)	0.01–18.64 (1.04)
	100–200 cm	97	0.51	0.35	0.66	0.21–58.9 (2.50)	0.03–17.06 (0.46)
Boreal	0–20 cm	1409	0.43	0.40	0.47	0.16–64.1 (4.30)	0.04–18.63 (2.58)
	20–50 cm	1220	0.57	0.54	0.61	0.16–72.05 (5.38)	0.03–18.85 (1.17)
	50–100 cm	1106	0.57	0.53	0.61	0.15–100.21 (3.30)	0.01–18.64 (0.47)
	100–200 cm	545	0.48	0.41	0.55	0.16–62.25 (2.21)	0.01–9.11 (0.26)
Cool temperate	0–20 cm	6360	0.47	0.45	0.48	0.12–77.2 (2.86)	0.01–19.93 (1.98)
	20–50 cm	4250	0.51	0.49	0.53	0.09–93.65 (3.91)	0.01–18.28 (0.95)
	50–100 cm	4347	0.41	0.39	0.43	0.01–102.76 (2.76)	0.01–12.03 (0.42)
	100–200 cm	2869	0.31	0.28	0.35	0.15–84.55 (2.11)	0.01–10.68 (0.28)
Warm temperate	0–20 cm	2134	0.49	0.46	0.52	0.01–133.51 (2.37)	0.04–19.98 (1.64)
	20–50 cm	1491	0.52	0.49	0.55	0.06–141.91 (2.70)	0.01–19.38 (0.62)
	50–100 cm	1536	0.49	0.46	0.52	0.08–127.55 (2.51)	0.01–20.00 (0.32)
	100–200 cm	1154	0.45	0.40	0.50	0.06–126.16 (2.06)	0.01–19.20 (0.21)
Subtropical	0–20 cm	1800	0.38	0.35	0.41	0.01–92.30 (2.76)	0.01–19.52 (1.86)
	20–50 cm	1545	0.47	0.45	0.50	0.01–118.26 (2.83)	0.01–14.86 (0.94)
	50–100 cm	977	0.49	0.45	0.52	0.01–128.91 (2.20)	0.01–14.11 (0.46)
	100–200 cm	881	0.49	0.43	0.55	0.01–160.35 (1.75)	0.01–13.98 (0.32)
Tropical	0–20 cm	668	0.48	0.43	0.53	0.21–81.06 (3.30)	0.12–19.90 (1.86)
	20–50 cm	502	0.41	0.36	0.46	0.01–73.4 (2.92)	0.08–18.00 (0.81)
	50–100 cm	297	0.34	0.26	0.41	0.01–71.01 (3.61)	0.03–13.38 (0.63)
	100–200 cm	225	0.35	0.23	0.47	0.26–72.5 (3.16)	0.01–7.56 (0.38)

relative number of significant differences between slopes was smaller for the temperature model (43%) than for to the moisture model (64%; Figs. 3 and 4 and Tables 1 and 2). In addition, the range of observed slopes was smaller for the temperature model (0.31–0.65) compared to the moisture model (0.01–0.62; Tables 1 and 2). Among the slopes that differed significantly from each other in the temperature model, (sub-)polar regions always had the steepest slope, with the highest value at 50 to 100 cm (0.65), except in the topsoil where warm temperate and tropical regions had the steepest slope (0.49 and 0.48, respectively).

Importance of oxalate-extractable metals across all climate regions

The linear mixed-effects model that included all individual Holdridge Life Zones explained the most variance in SOC (R^2 = 0.55, RMSE = 0.42 wt %; Online Appendix Table A8). In comparison, the linear mixed-effects model that did not include any climate predictor (i.e., one global slope) explained 46% in the variance of SOC (RMSE = 0.42 wt %; Online Appendix Table A9). We compared the slopes between M_{ox} and SOC across individual Holdridge Life Zones within each depth layer (Fig. 5). The more the slope deviates from the estimated marginal means from the model without climate variables, the more important

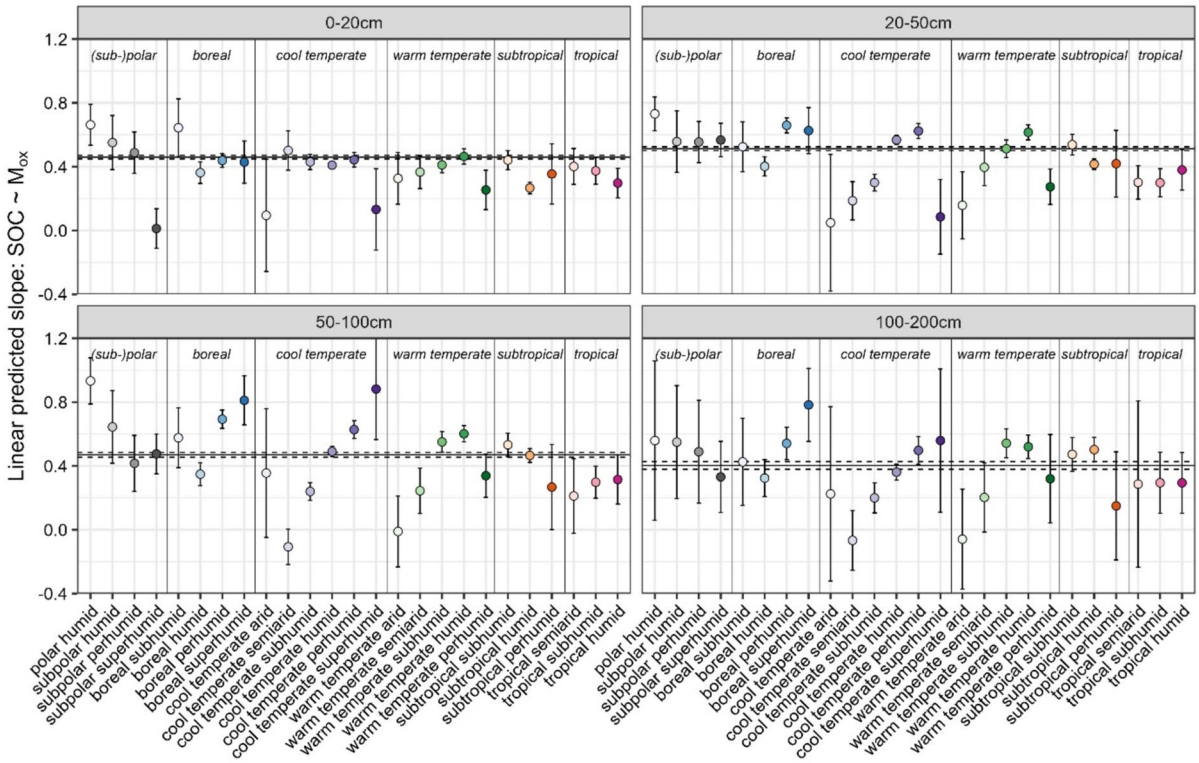

Fig. 5 Linear predicted slopes based on estimated marginal means for the interaction term M_{ox} ($=Al_{ox}+½\ Fe_{ox}$) * Holdridge Life Zone * depth. Error bars indicate 95th confidence intervals. Note that all variables were log-transformed prior analysis. Solid horizontal line refers to the estimated marginal means without the Holdridge Life Zones as a predictor and the dashed horizontal lines refer to the upper and lower 95th confidence interval, respectively. Solid vertical lines separate the Holdridge Life Zones by temperature. Within each temperature group, zones are ordered from driest (light) to wettest (dark)

it is to include climate-specific information in that region.

On average, the slopes at the surface (0–20 cm) and at the maximum depth (100–200 cm) deviated the least from the estimated marginal mean without any climate variable (Fig. 5). This observation agrees well with the model results by moisture and temperature. Below 20 cm, the Holdridge Life Zones that differed most from the marginal mean without any climate variable were typically very dry or very wet from boreal, cold, or warm temperate regions, with wetter regions usually having a larger estimated marginal mean and drier regions having a smaller estimated marginal mean. For (sub-)polar, subtropical and tropical regions, the distance to the estimated marginal mean without any climate variable was usually smaller and less distinguishable by moisture; except for polar humid which showed the largest slope (0.82; Fig. 5).

Discussion

What is the relationship between SOC and Al_{ox} and Fe_{ox} at the global scale?

Our results highlight that M_{ox}, Al_{ox} and Fe_{ox} are promising proxies for organo-mineral interactions (i.e., sorption capacity) at the global scale, albeit with notable differences between environmental conditions. This agrees well with previous large-scale studies (e.g., Hall & Thompson 2022; Quesada et al. 2020; Rasmussen et al. 2018; von Fromm et al. 2021; Yu et al. 2021). Similar to these studies, we found that Al_{ox} is a better predictor than Fe_{ox}. Reasons for this may include differences in their chemical characteristics, such as Al-bearing minerals typically dissolving at higher pH values than Fe phases, higher total Al abundance in the parent material, and biogeochemical reduction of Fe^{III} phases to Fe^{II} in anoxic microsites of unsaturated

soils—a more detailed discussion of these differences is provided in Hall & Thompson (2022). Furthermore, we also found a curvilinear relationship between SOC and M_{ox} meaning that the response of SOC to M_{ox} becomes weaker at higher levels of M_{ox}—which is also reflected in the partial dependence plots of our random forest model (Online Appendix Fig. A13; Hall and Thompson 2022; von Fromm et al. 2021; Yu et al. 2021).

Are the Holdridge Life Zones a good climate disaggregation approach at the global scale to better understand the role of organo-mineral interactions?

Grouping soil data by Holdridge Life Zones improved overall model accuracy at the global scale. This highlights that the role—and mechanisms—of oxalate-extractable metals in governing SOC abundance differ between climatic conditions. Our data suggest that the Holdridge Life Zones are an effective hierarchical grouping variable at the global scale, reflecting not only geographic patterns in SOC abundance (Jungkunst et al. 2021), but also in soil mineralogy and geochemistry. This has been shown previously at the continental scale and with different grouping variables (e.g., Rasmussen et al. 2018; von Fromm et al. 2021, 2024; Yu et al. 2021). However, we recognize that other soil-forming factors, particularly time (i.e., soil age), may also modulate the relationship between SOC and M_{ox} within a given climatic region. For example, cool temperate zones may be dominated by either "younger" soils (< 12,000 years) or "older" soils unaffected by glaciers or loess deposits (see Online Appendix and Fig. A18). Indeed, when we included soil age in the model for cool temperate regions (data not shown), it explained the same variation in SOC as the model that included only moisture as a grouping variable, but the overall importance of moisture as a predictor decreased. This suggests that the model including soil age provides different information, meaning that refining models for specific climatic regions by further disaggregation could be valuable. However, which soil formation factor to use will be different for different climatic regions. Thus, we argue that the Holdridge Life Zones are the first hierarchical grouping variable to be considered at the global scale, and for many climatic regions more data are needed to further disaggregate them within each climatic region.

Most of the slopes between SOC abundance and oxalate-extractable metals and their intercepts differed significantly among climatic regions and with soil depth; the explained variance of SOC increased by about 10% when Holdridge Life Zones were included. The intercept likely reflects the ratio of mineral-associated carbon (MAOC) to particulate-organic carbon (POC). In other words, if MAOC is close to zero when there are no or only small amounts of SRO minerals, then the intercept should be equal to POC (Kirschbaum et al. 2020). In our case, dry regions usually had the highest intercept (except for 0–20 cm), which is consistent with other studies that found higher POC under dry conditions (Cotrufo & Lavallee 2022; Haddix et al. 2020). The slopes themselves can be related to the SOC load on the SRO minerals (i.e., g C / kg mineral; or the achievable/effective capacity based on current environmental conditions). Again, we usually found the smallest slopes under dry conditions, which is consistent with the idea that at higher pH values, less SOC can be adsorbed on minerals, and that under drier conditions, less dissolved organic matter is produced, which is usually associated with higher MAOC (Cotrufo and Lavalle 2022). However, it is also important to note that under drier conditions there can be other important mechanisms, such as cation (Ca) bridging that can result in higher SOC abundance (Rasmussen et al. 2018; Rowley et al. 2018; von Fromm et al. 2021). All of this highlights the importance of capturing regional-specific differences at the global scale that are related to climate, soil mineralogy and development, independent of overall model performance. It is particularly important to correctly capture regional-specific relationships if we are interested in improving model predictions of soil change at the global scale. Our data showed that rather than making a universal global prediction with a set of defined relationships, it may be more useful from a mechanistic point of view to identify regions controlled by similar factors and make area-specific predictions. This is something that should be relatively easy to implement in global soil C models, as it does not necessarily require changing the model structure, but rather defining under which climatic conditions which factors are most important in predicting SOC abundance.

Interestingly, moisture was a more important driver of the observed patterns between SOC and oxalate-extractable metals than temperature. This underscores

the importance of moisture for soil formation in general and its fundamental role in mediating organo-mineral interactions that promote SOC protection. However, it also highlights the challenge of separating organo-metal complexation from the influence of SRO minerals on SOC abundance, as SOC content is generally higher in humid regions due to higher C inputs. Nevertheless, our data also showed that SOC abundance and oxalate-extractable metals do not always vary in the same way across Holdridge Life Zones and soil depth. Thus, differences in soil depth across distinct climatic regions may help to better disentangle whether oxalate-extractable metals drive, respond to, or covary with SOC abundance.

Are there region-specific differences in the relationship between SOC and Al_{ox} and Fe_{ox} with soil depth?

The relationship (i.e., slope) between SOC and oxalate-extractable metals varied with soil depth across environmental conditions, highlighting the importance of vertical soil processes. For a given environmental condition, the steepest slope (SOC ~ oxalate-extractable metals) was usually found between 20 and 100 cm, except for arid, semiarid, and tropical regions, which had the steepest slopes in the surface layers (0–50 cm). These results suggest soil depth is an important modulator of the soil processes that drive the relationship between SOC and oxalate-extractable metals. For example, most surface soils are strongly influenced by C inputs, which also likely results in higher concentrations of oxalate-extractable metals due to organic acid-mediated weathering (Collignon et al. 2012; Yu et al. 2017). Whereas, below the dominant rooting zone (~ 20–30 cm), C inputs are likely lower and thus, it is more likely that the positive slopes we observe are driven by oxalate-extractable metals that favor the sorption of SOC to short-range ordered minerals. Interestingly, we observed a smaller slope below 100 cm under most environmental conditions compared to 20 to 100 cm. Due to the overall low SOC abundance at this depth, little to no relationship between oxalate-extractable metals and SOC might be expected if available binding sites on mineral surfaces are not occupied (Schrumpf et al. 2013). In addition, it has been shown that the proportion of crystalline mineral phases increases with depth, which would result in less SOC being adsorbed by minerals (Chen et al. 2019). Exceptions to the observed depth pattern in our dataset are usually found in warmer regions (warm temperate, subtropical, and tropical), where we typically found steep slopes at 100 to 200 cm that were not significantly different from the depth layers above. This may be related to deeper weathering of these soil profiles and deeper rooting depth (Jackson et al. 1996; Yang et al. 2016).

It is evident from our findings that the relationship between SOC abundance and oxalate-extractable metals for certain environmental conditions deviates more from the global mean than for others. This has implications for understanding where pedo-climatic interactions are most important, and underscores the need to incorporate these interactions into models to improve predictions of soil change. For example, both dry and wet conditions in the boreal, cool and warm temperate regions tend to deviate most from the global mean below the surface layer, while (sub-)polar, subtropical, and tropical regions tend to be closer to the global mean regardless of moisture. Although the latter three show strong relationships between SOC and oxalate-extractable metals, the fact that they are either at the beginning of the weathering process (subpolar) or at the end of the weathering process (subtropical and tropical) suggests that soil weathering and the time for soil development influence organo-mineral interactions in a non-linear fashion (Torn et al. 1997). This also highlights the fact that different processes can lead to a correlation between SOC abundance and oxalate-extractable metals that differ across pedo-climatic conditions. For example, although weathering time is relatively short in (sub-)polar regions, we found a comparable importance of oxalate-extractable metals regarding SOC abundance to other regions. This may be related to the importance of these metals in arctic regions in forming mineral-organic complexes to protect SOC from decomposition (Monhonval et al. 2023; Thomas et al. 2023). However, it is also important to note that we do not have sufficient representation in our dataset for some of these Holdridge Life Zones to further disaggregate them based on other important soil-forming factors that would also allow us to better understand within-group variation. In contrast, in cool and warm temperate regions, mineral weathering has resulted in distinct soil characteristics along moisture gradients, leading to distinct relationships between SOC abundance and

oxalate-extractable metals. Therefore, it is especially important in these regions to correctly capture these internal dynamics. In the future, the identified climate regions can be further disaggregated based on other important soil-forming factors to further improve the predictions of soil-climate feedbacks.

Conclusions and outlook

Our results underline the importance of oxalate-extractable metals as predictors for organo-mineral interactions at the global scale, although there are significant differences between environmental conditions and soil depth. Oxalate-extractable metals are usually most important between 20 and 100 cm under wet conditions. Temperature alone tends to be a smaller driver of organo-mineral interactions than moisture. At the surface it is likely that the positive relationship between SOC and oxalate-extractable metals is due to organic acid-mediated weathering. All these findings, and thus the underlying mechanisms are important to consider in global soil C models. Our synthesis shows that there is enough data worldwide available to try to implement the observed patterns into biogeochemical soil models. For example, a reasonable first step might be to include an empirical relationship between oxalate-extractable metal content and the decomposition rate of slower-cycling carbon pools, and to make this empirical constant vary predictably with climatic regions (i.e., based on the Holdridge Life Zones). This hierarchical framework, starting with a climatic disaggregation, followed by further disaggregation based on other soil-forming factors within specific climatic groups, will help improve soil-climate feedbacks at the global scale. However, our data lacked the temporal component of soils to changes and did not allow us to test for non-static relationships between oxalate-extractable metals and SOC abundance–this should be the focus of future studies so that model predictions can be tested against time series.

In addition to modeling implications, our synthesized dataset also highlights geographic areas where we still lack data. These are usually arid regions, but the importance of oxalate-extractable metals in these regions is also limited. Thus, future sampling efforts should focus on boreal regions where we usually observed the largest discrepancy between the global coverage and our data, as well as low data coverage. Furthermore, subtropical and tropical regions should also be the focus of future sampling efforts due to their underrepresentation. Lastly, we showed that the grouping based on Holdridge Life Zones works well for oxalate-extractable metals and SOC abundance, future studies may further test other important predictors of SOC content such as soil age or parent material. It is important to keep in mind that climate is much easier to quantify at the global scale than other factors such as parent material, soil age, organisms, and relief. This may statistically assign a higher importance to climate than to other soil-forming factors. Future research should seek to improve global products that better represent soil mineralogy and age.

Funding Open Access funding enabled and organized by Projekt DEAL. The authors thank everyone who contributed data to this analysis – either by making their data publicly available or by sharing it directly. SvF received funding from the William H. Neukom 1964 Institute for Computational Science. CR received support from USDA NIFA Hatch/Multi-State project #ARZT-1370600-R21-189. BA and BS wish to acknowledge the Department of Agriculture, Water and the Environment, The Commonwealth of Australia, via the Soil Science Challenge project—A soil–plant nexus to maximize organic carbon sequestration in the soil (4-H4T0SA3). CHP and FML received support from the National Science Foundation, Office of Polar Programs under Grant #2031253. KG was supported by the LLNL-LDRD Program under Project No. 22-ERD-019, and under the auspices of DOE Contract DE-AC52-07NA27344.

Data availability The data that support the findings of this study are openly available on Zenodo, including all R code necessary to reproduce the analysis and figures (von Fromm 2024).

Declarations

Conflict of interest The authors declare no competing interests.

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/.

References

Bartoń K (2024) MuMIn: Multi-Model Inference (Version R package version 1.48.4) [Computer software]. https://CRAN.Rproject.org/package=MuMIn

Börker, J., Hartmann, J., Amann, T., & Romero-Mujalli, G. (2018). Global Unconsolidated Sediments Map Database v1.0 (shapefile and gridded to 0.5° spatial resolution). In *Supplement to: Börker, J et al. (2018): Terrestrial Sediments of the Earth: Development of a Global Unconsolidated Sediments Map Database (GUM). Geochemistry, Geophysics, Geosystems, 19(4), 997-1024,* https://doi.org/10.1002/2017GC007273. *PANGAEA. https://doi.org/10.1594/PANGAEA.884822*

Chen C, Barcellos D, Richter DD, Schroeder PA, Thompson A (2019) Redoximorphic Bt horizons of the Calhoun CZO soils exhibit depth-dependent iron-oxide crystallinity. J Soils Sediments 19(2):785–797. https://doi.org/10.1007/s11368-018-2068-2

Chorover, J. (2022). Microbe-Biomolecule-Mineral Interfacial Reactions. In *Multi-Scale Biogeochemical Processes in Soil Ecosystems* (pp. 117–140). https://doi.org/10.1002/9781119480419.ch5

Collignon, C., Ranger, J., & Turpault, M. P. (2012). Seasonal dynamics of Al- and Fe-bearing secondary minerals in an acid forest soil: Influence of Norway spruce roots (Picea abies (L.) Karst.). *European Journal of Soil Science*, *63*(5), 592–602. https://doi.org/10.1111/j.1365-2389.2012.01470.x

Cotrufo MF, Lavallee JM (2022) Soil organic matter formation, persistence, and functioning: a synthesis of current understanding to inform its conservation and regeneration. Adv Agron 172:1–66. https://doi.org/10.1016/bs.agron.2021.11.002

Dokuchaev, V. V. (1883). The Russian chernozem report to the free economic society. *Imperial University of St. Petersburg: St. Petersburg)[in Russian]*

Fernandez-Ugalde O, Scarpa S, Orgiazzi A, Panagos P, Van Liedekerke M, Marechal A, Jones A (2022) LUCAS 2018 soil module. Presentation of Dataset and Results. https://doi.org/10.2760/215013

Fick SE, Hijmans RJ (2017) WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. Int J Climatol 37(12):4302–4315. https://doi.org/10.1002/joc.5086

Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Statist. https://doi.org/10.1214/aos/1013203451

von Fromm SF (2025) SophievF/Global_Mox_analysis: v1.1.0 (v1.1.0). Zenodo. https://doi.org/10.5281/zenodo.14732942

Haddix ML, Gregorich EG, Helgason BL, Janzen H, Ellert BH, Francesca Cotrufo M (2020) Climate, carbon content, and soil texture control the independent formation and persistence of particulate and mineral-associated organic matter in soil. Geoderma 363:114160. https://doi.org/10.1016/j.geoderma.2019.114160

Hall SJ, Thompson A (2022) What do relationships between extractable metals and soil organic carbon concentrations mean? Soil Sci Soc Am J 86(2):195–208. https://doi.org/10.1002/saj2.20343

Hall SJ, McNicol G, Natake T, Silver WL (2015) Large fluxes and rapid turnover of mineral-associated carbon across topographic gradients in a humid tropical forest: Insights from paired 14C analysis. Biogeosciences 12(8):2471–2487. https://doi.org/10.5194/bg-12-2471-2015

Hicks Pries CE, Ryals R, Zhu B, Min K, Cooper A, Goldsmith S, Pett-Ridge J, Torn M, Berhe AA (2023) The deep soil organic carbon response to global change in annual review of ecology. Evol Systemat. https://doi.org/10.1146/annurev-ecolsys-102320-085332

Jackson RB, Canadell J, Ehleringer JR, Mooney HA, Sala OE, Schulze ED (1996) A global analysis of root distributions for terrestrial biomes. Oecologia 108(3):389–411. https://doi.org/10.1007/BF00333714

Jenny, H. (1941). *Factors of Soil Formation: A System of Quantitative Pedology*. McGraw-Hill. https://books.google.com/books?id=h-dIAAAAMAAJ

Jungkunst HF, Goepel J, Horvath T, Ott S, Brunn M (2021) New uses for old tools: reviving holdridge life zones in soil carbon persistence research. J Plant Nutr Soil Sci 184(1):5–11. https://doi.org/10.1002/jpln.202100008

Kahle M, Kleber M, Jahn R (2002) Predicting carbon content in illitic clay fractions from surface area, cation exchange capacity and dithionite-extractable iron. Eur J Soil Sci 53(4):639–644. https://doi.org/10.1046/j.1365-2389.2002.00487.x

Kaiser K, Eusterhues K, Rumpel C, Guggenberger G, Kögel-Knabner I (2002) Stabilization of organic matter by soil minerals—Investigations of density and particle-size fractions from two acid forest soils. J Plant Nutr Soil Sci 165(4):451–459. https://doi.org/10.1002/1522-2624(200208)165:4%3c451::AID-JPLN451%3e3.0.CO;2-B

Kassambara, A. (2023a). *ggpubr: "ggplot2" Based Publication Ready Plots* (Version R package version 0.6.0) [Computer software]. https://CRAN.R-project.org/package=ggpubr

Kassambara, A. (2023b). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests* (Version R package version 0.7.2) [Computer software]. https://rpkgs.datanovia.com/rstatix/

Kirschbaum MUF, Moinet GYK, Hedley CB, Beare MH, McNally SR (2020) A conceptual model of carbon stabilisation based on patterns observed in different soils. Soil Biol Biochem 141:107683. https://doi.org/10.1016/j.soilbio.2019.107683

Kleber M, Mikutta R, Torn MS, Jahn R (2005) Poorly crystalline mineral phases protect organic matter in acid subsoil horizons. Eur J Soil Sci 56(6):717–725. https://doi.org/10.1111/j.1365-2389.2005.00706.x

Lang M, Binder M, Richter J, Schratz P, Pfisterer F, Coors S, Au Q, Casalicchio G, Kotthoff L, Bischl B (2019) mlr3: a modern object-oriented machine learning framework in R. J Open Sour Softw. https://doi.org/10.21105/joss.01903

Lawrence CR, Harden JW, Xu X, Schulz MS, Trumbore SE (2015) Long-term controls on soil organic carbon with depth and time: a case study from the Cowlitz River

Chronosequence, WA USA. Geoderma 247–248:73–87. https://doi.org/10.1016/j.geoderma.2015.02.005

Lawrence CR, Beem-Miller J, Hoyt AM, Monroe G, Sierra CA, Stoner S, Heckman K, Blankinship JC, Crow SE, McNicol G, Trumbore S, Levine PA, Vindušková O, Todd-Brown K, Rasmussen C, Hicks Pries CE, Schädel C, McFarlane K, Doetterl S, Wagai R (2020) An open-source database for the synthesis of soil radiocarbon data: international soil radiocarbon database (ISRaD) version 10. Earth Syst. Sci. Data 12(1):61–76. https://doi.org/10.5194/essd-12-61-2020

Lenth, R. V. (2024). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (Version R package version 1.10.3) [Computer software]. https://CRAN.R-project.org/package=emmeans

Lugo AE, Brown SL, Dodson R, Smith TS, Shugart HH (1999) The Holdridge life zones of the conterminous United States in relation to ecosystem mapping. J Biogeogr 26(5):1025–1038. https://doi.org/10.1046/j.1365-2699.1999.00329.x

Masiello CA, Chadwick OA, Southon J, Torn MS, Harden JW (2004) Weathering controls on mechanisms of carbon storage in grassland soils. Global Biogeochem Cycles. https://doi.org/10.1029/2004GB002219

Molnar C, Bischl B (2018) iml: an R package for interpretable machine learning. J Open Sour Softw. https://doi.org/10.21105/joss.00786

Monhonval A, Mauclet E, Hirst C, Bemelmans N, Eekman E, Schuur EAG, Opfergelt S (2023) Mineral organic carbon interactions in dry versus wet tundra soils. Geoderma 436:116552. https://doi.org/10.1016/j.geoderma.2023.116552

NEON. (2023). *Soil physical and chemical properties, distributed initial characterization (dp1.10047.001)*. . https://doi.org/10.48443/KA55-2852.

Neuwirth, E. (2022). *RColorBrewer: ColorBrewer Palettes* (Version R package version 1.1–3) [Computer software]. https://CRAN.R-project.org/package=RColorBrewer

Oades JM (1988) The retention of organic matter in soils. Biogeochemistry 5:35–70. https://doi.org/10.1007/BF02180317

Parfitt RL, Childs CW (1988) Estimation of forms of Fe and Al-a review, and analysis of contrasting soils by dissolution and Mossbauer methods. Soil Res 26(1):121–144. https://doi.org/10.1071/SR9880121

Percival HJ, Parfitt RL, Scott NA (2000) Factors controlling soil carbon levels in New Zealand grasslands is clay content important? Soil Sci Soc Am J 64(5):1623–1630. https://doi.org/10.2136/sssaj2000.6451623x

Pinheiro, J., Bates, D., & R Core Team. (2023). *nlme: Linear and Nonlinear Mixed Effects Models* (Version R package version 3.1–164) [Computer software]. https://CRAN.R-project.org/package=nlme

Post WM, Emanuel WR, Zinke PJ, Stangenberger AG (1982) Soil carbon pools and world life zones. Nature 298(5870):156–159. https://doi.org/10.1038/298156a0

Powers JS, Schlesinger WH (2002) Relationships among soil carbon distributions and biophysical factors at nested spatial scales in rain forests of northeastern Costa Rica. Geoderma 109(3–4):165–190. https://doi.org/10.1016/S0016-7061(02)00147-7

Quesada CA, Paz C, Oblitas Mendoza E, Phillips OL, Saiz G, Lloyd J (2020) Variations in soil chemical and physical properties explain basin-wide Amazon forest soil carbon concentrations. Soil 6(1):53–88. https://doi.org/10.5194/soil-6-53-2020

R Core Team. (2024). *R: A Language and Environment for Statistical Computing* (Version R version 4.4.1) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rasmussen C, Heckman K, Wieder WR, Keiluweit M, Lawrence CR, Berhe AA, Blankinship JC, Crow SE, Druhan JL, Hicks Pries CE (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

Ray N, Adams JM (2001) A GIS-based Vegetation Map of the World at the Last Glacial Maximum (25,000–15,000 BP). Int Archaeol. https://doi.org/10.11141/ia.11.2

Rennert T (2018) Wet-chemical extractions to characterise pedogenic Al and Fe species–a critical review. Soil Research 57(1):1–16. https://doi.org/10.1071/SR18299

Rowley MC, Grand S, Verrecchia ÉP (2018) Calcium-mediated stabilisation of soil organic carbon. Biogeochemistry 137(1):27–49. https://doi.org/10.1007/s10533-017-0410-1

Schrumpf M, Kaiser K, Guggenberger G, Persson T, Kögel-Knabner I, Schulze E-D (2013) Storage and stability of organic carbon in soils as related to depth, occlusion within aggregates, and attachment to minerals. Biogeosciences 10(3):1675–1691. https://doi.org/10.5194/bg-10-1675-2013

Slessarev EW, Chadwick OA, Sokol NW, Nuccio EE, Pett-Ridge J (2022) Rock weathering controls the potential for soil carbon storage at a continental scale. Biogeochemistry 157(1):1–13. https://doi.org/10.1007/s10533-021-00859-8

Souza IF, Archanjo BS, Hurtarte LC, Oliveros ME, Gouvea CP, Lidizio LR, Achete CA, Schaefer CE, Silva IR (2017) Al-/Fe-(hydr) oxides–organic carbon associations in Oxisols—From ecosystems to submicron scales. CATENA 154:63–72. https://doi.org/10.1016/j.catena.2017.02.017

Thomas M, Monhonval A, Hirst C, Bröder L, Zolkos S, Vonk JE, Tank SE, Keskitalo KH, Shakil S, Kokelj SV, van der Sluijs J, Opfergelt S (2023) Evidence for preservation of organic carbon interacting with iron in material displaced from retrogressive thaw slumps: case study in Peel Plateau, western Canadian Arctic. Geoderma 433:116443. https://doi.org/10.1016/j.geoderma.2023.116443

Tisdall JM, Oades JM (1982) Organic matter and water-stable aggregates in soils. J Soil Sci 33(2):141–163. https://doi.org/10.1111/j.1365-2389.1982.tb01755.x

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

Todd-Brown KEO, Randerson JT, Post WM, Hoffman FM, Tarnocai C, Schuur EAG, Allison SD (2013) Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. Biogeosciences 10(3):1717–1736. https://doi.org/10.5194/bg-10-1717-2013

Vågen, T.-G., Winowiecki, L. A., Desta, L., Tondoh, E. J., Weullow, E., Shepherd, K. D., Sila, A. M., Dunham, S. J.,

Hernández-Allica, J., Carter, J., & McGrath, S. P. (2021). *Wet chemistry data for a subset of afsis: Phase i archived soil samples. World Agroforestry—Research Data Repository*. https://doi.org/10.34725/DVN/66BFOB

von Fromm SF, Hoyt AM, Lange M, Acquah GE, Aynekulu E, Berhe AA, Haefele SM, McGrath SP, Shepherd KD, Sila AM, Six J, Towett EK, Trumbore SE, Vågen T-G, Weullow E, Winowiecki LA, Doetterl S (2021) Continental-scale controls on soil organic carbon across sub-Saharan Africa. SOIL 7(1):305–332. https://doi.org/10.5194/soil-7-305-2021

von Fromm SF, Doetterl S, Butler BM, Aynekulu E, Berhe AA, Haefele SM, McGrath SP, Shepherd KD, Six J, Tamene L, Tondoh EJ, Vågen T-G, Winowiecki LA, Trumbore SE, Hoyt AM (2024) Controls on timescales of soil organic carbon persistence across sub-Saharan Africa. Glob Change Biol 30(1):e17089. https://doi.org/10.1111/gcb.17089

Wagai R, Mayer LM (2007) Sorptive stabilization of organic matter in soils by hydrous iron oxides. Geochim Cosmochim Acta 71(1):25–35. https://doi.org/10.1016/j.gca.2006.08.047

Wagai R, Kajiura M, Asano M (2020) Iron and aluminum association with microbially processed organic matter via meso-density aggregate formation across soils: Organometallic glue hypothesis. SOIL 6(2):597–627. https://doi.org/10.5194/soil-6-597-2020

Wagai, R., Yang, P.-T., & Kaiser, K. (2023). Interfacial reactions of microorganisms with minerals and organic matter. In M. J. Goss & M. Oliver (Eds.), *Encyclopedia of Soils in the Environment (Second Edition)* (pp. 458–469). Academic Press. https://doi.org/10.1016/B978-0-12-822974-3.00232-9

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J (2019) Welcome to the Tidyverse. J Open Sour Softw 4(43):1686. https://doi.org/10.21105/joss.01686

Wickham, H., Pedersen, T. L., & Seidel, D. (2023). *scales: Scale Functions for Visualization* (Version R package version 1.3.0) [Computer software]. https://CRAN.R-project.org/package=scales

Wieder WR, Hartman MD, Sulman BN, Wang Y-P, Koven CD, Bonan GB (2018) Carbon cycle confidence and uncertainty: exploring variation among soil biogeochemical models. Glob Change Biol 24(4):1563–1579. https://doi.org/10.1111/gcb.13979

Wright MN, Ziegler A (2017) Ranger: a fast implementation of random forests for high dimensional data in C++ and R. J Stat Softw 77(1):1–17. https://doi.org/10.18637/jss.v077.i01

Xu X, Thornton PE, Post WM (2013) A global analysis of soil microbial biomass carbon, nitrogen and phosphorus in terrestrial ecosystems. Glob Ecol Biogeogr 22(6):737–749. https://doi.org/10.1111/geb.12029

Yang Y, Donohue RJ, McVicar TR (2016) Global estimation of effective plant rooting depth: implications for hydrological modeling. Water Resour Res 52(10):8260–8276. https://doi.org/10.1002/2016WR019392

Yu G, Xiao J, Hu S, Polizzotto ML, Zhao F, McGrath SP, Li H, Ran W, Shen Q (2017) Mineral availability as a key regulator of soil carbon storage. Environ Sci Technol 51(9):4960–4969. https://doi.org/10.1021/acs.est.7b00305

Yu W, Weintraub SR, Hall SJ (2021) Climatic and geochemical controls on soil carbon at the continental scale: Interactions and thresholds. Global Biogeochem Cycles. https://doi.org/10.1029/2020GB006781

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