



# How much carbon can be added to soil by sorption?

Rose Z. Abramoff · Katerina Georgiou · Bertrand Guenet · Margaret S. Torn · Yuanyuan Huang · Haicheng Zhang · Wenting Feng · Sindhu Jagadamma · Klaus Kaiser · Dolly Kothawala · Melanie A. Mayes · Philippe Ciais

Received: 7 October 2020 / Accepted: 13 January 2021 / Published online: 26 January 2021  
© The Author(s) 2021

**Abstract** Quantifying the upper limit of stable soil carbon storage is essential for guiding policies to increase soil carbon storage. One pool of carbon considered particularly stable across climate zones and soil types is formed when dissolved organic carbon sorbs to minerals. We quantified, for the first time, the potential of mineral soils to sorb additional dissolved organic carbon (DOC) for six soil orders. We compiled 402 laboratory sorption experiments to estimate the additional DOC sorption potential, that is

the potential of excess DOC sorption in addition to the existing background level already sorbed in each soil sample. We estimated this potential using gridded climate and soil geochemical variables within a machine learning model. We find that mid- and low-latitude soils and subsoils have a greater capacity to store DOC by sorption compared to high-latitude soils and topsoils. The global additional DOC sorption potential for six soil orders is estimated to be  $107 \pm 13$  Pg C to 1 m depth. If this potential was realized, it would represent a 7% increase in the existing total carbon stock.

Responsible Editor: Stuart Grandy.

This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10533-021-00759-x>.

R. Z. Abramoff (✉) · B. Guenet · P. Ciais  
Laboratoire des Sciences du Climat et de  
l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ,  
Université Paris-Saclay, 91191 Gif-sur-Yvette, France  
e-mail: rose.abramoff@lsce.ipsl.fr

**Keywords** Sorption · Additional sorption potential · Saturation · Soil organic carbon · Mineral association

## Introduction

Carbon (C) inputs to soil, much of it in the form of DOC, are expected to increase from CO<sub>2</sub> fertilization (Drake et al. 2011; Jiang et al. 2020; Palmroth et al. 2006) and improved management practices (Cardinael et al. 2018; Maillard et al. 2017; Poeplau and Don

B. Guenet  
e-mail: bertrand.guenet@lsce.ipsl.fr

P. Ciais  
e-mail: philippe.ciais@lsce.ipsl.fr

2015) such as those outlined in the ‘4 per 1000’ initiative, which aims at removing atmospheric CO<sub>2</sub> by increasing soil C sequestration with an emphasis on long-term storage (Minasny et al. 2017). In addition, warming soils are likely to experience increased decomposition of particulate organic matter, leading to some increase in DOC production (Bengtson and Bengtsson 2007; Fu et al. 2019; Reynolds and Fenner 2001). This additional DOC can be lost through decomposition or runoff, or can sorb to mineral particles as it percolates down the soil column (Kaiser and Kalbitz 2012) thereby increasing the size of the total soil C pool with long-term storage characteristics. Currently, there is no existing global-scale estimate of how much additional C can be stored in soils through sorption of DOC. This lack of knowledge prevents an understanding of the upper limit of stable carbon storage in soils to inform policies aiming to increase soil C concentrations.

Different soil C pools and stabilization mechanisms exist, but a pool of C considered to be particularly stable is the mineral-associated fraction, measured using density or size fractionation (Cotrufo et al. 2019; Kleber et al. 2015; Poeplau et al. 2018). During the 1970s and 1980s, soil organic matter was thought to be stable to the extent that it contained chemically recalcitrant compounds (Kogel-Knabner 1986;

Kogel-Knabner et al. 1988). Over the next three decades, researchers demonstrated that those so-called recalcitrant compounds can in fact be degraded under the appropriate conditions (Gleixner et al. 2002; Rasse et al. 2006), and identified other mechanisms that stabilize soil organic matter and prevent it from being respired: interaction with minerals (Gleixner et al. 2001; Torn et al. 1997), physical protection within aggregates (Tisdall and Oades 1982), and environmental limitations of microbial activities (Schmidt et al. 2011).

One key mechanism driving interactions with minerals is sorption, the formation of chemical associations between soil minerals and organic compounds. Sorption can protect organic C from decomposition, even if some of the compounds in organic C are labile or young in age (Eusterhues et al. 2003; Kaiser et al. 2002; Kiem and Kogel-Knabner 2002; Kögel-Knabner et al. 2008; Porras et al. 2018; Schmidt et al. 1999). Sorption is now considered to be one of the dominant mechanisms leading to a mineral-associated C pool (Cotrufo et al. 2013; Schmidt et al. 2011) and its importance is beginning to be recognized and explicitly incorporated in soil C models (Ahrens et al. 2015; Robertson et al. 2019; Sulman et al. 2018).

Laboratory batch isotherm experiments measure the equilibrium partitioning of C between the solid and

---

K. Georgiou

Department of Earth System Science, Stanford University,  
Yang & Yamazaki Building, 473 Via Ortega, Stanford,  
CA 94305-4216, USA  
e-mail: georgiou@stanford.edu

B. Guenet

Laboratoire de Géologie, UMR 8538, Ecole Normale  
Supérieure, PSL Research University, CNRS, Paris,  
France

M. S. Torn

Lawrence Berkeley National Laboratory, 1 Cyclotron  
Road, Berkeley, CA, USA  
e-mail: mstorn@lbl.gov

Y. Huang

CSIRO Ocean and Atmosphere, Aspendale 3195,  
Australia  
e-mail: Yuanyuan.Huang@csiro.au

H. Zhang

Department Geoscience, Environment & Society,  
Université Libre de Bruxelles, 1050 Bruxelles, Belgium  
e-mail: haicheng.zhang@ulb.ac.be

W. Feng

Institute of Agriculture Resources and Regional Planning,  
Chinese Academy of Agricultural Sciences, 12  
Zhongguancun South Street, Beijing 10081, China  
e-mail: fengwenting@caas.cn

S. Jagadamma

Department of Biosystems Engineering and Soil Science,  
University of Tennessee, 2506 E.J. Chapman Drive,  
Knoxville, TN 37996, USA  
e-mail: sjagada1@utk.edu

K. Kaiser

Soil Science and Soil Protection, Martin Luther  
University Halle-Wittenberg, von-Seckendorff-Platz 3,  
06120 Halle (Saale), Germany  
e-mail: klaus.kaiser@landw.uni-halle.de

solution phases, controlled by solution ionic strength, chemical composition, mineral surface area, and soil properties such as pH and mineralogy. While some simulation models represent these organo-mineral associations with a linear equation (Camino-Serrano et al. 2018; Nakhavali et al. 2017), the non-linear Langmuir equation is more realistic (Abramoff et al. 2018; Ahrens et al. 2015; Robertson et al. 2019; Tang and Riley 2015; Wang et al. 2013) because of its ability to capture the asymptotic behavior observed in laboratory sorption experiments. The Langmuir equation assumes direct and reversible association of a solute, here DOC, with a surface that has a limited saturation capacity (Langmuir 1918). Given measurements of  $Q$ , the amount of sorbed DOC, this empirical equation allows the prediction of the maximum sorption potential, denoted  $Q_{max}$ , as given by:

$$Q = \frac{Q_{max} \cdot k \cdot DOC}{1 + k \cdot DOC} \quad (1)$$

where  $k$  is the equilibrium constant, which represents the relative tendency of the forward (at high values of  $k$ ) or reverse (at low values) reactions. The original Langmuir Eq. (1) assumes a zero intercept, that the amount of sorbed DOC is zero when the DOC concentration in solution is zero. This is theoretically true in a completely reversible system, but does not account for the presence of native organic matter, some of which is released into solution during batch experiments at low DOC concentrations. To correct for this observed non-zero intercept, Lilienfein et al. (2004) introduced the intercept  $b$ , representing the amount of DOC released into the solution if the DOC concentration is 0 mg/L, and which has been used by subsequent batch experiment studies on native soils (Kothawala et al. 2009; Mayes et al. 2012). If we use the parameter  $b$  to account for the amount of DOC that leaves the native soil when no DOC is added, then any new DOC sorbed ( $Q$ ) when DOC is added is additional to the organic C that already exists in the soil sample.

Therefore, the maximum sorption potential ( $Q_{max}$ ) is also additional to the organic C that already exists in the soil sample. Therefore, we rename  $Q_{max}$  as  $Q_{sp}$ , to represent the interpretation of  $Q_{sp}$  as the *additional* sorption potential that could be realized given that some mineral surface sites are already occupied in the native soil.

$$Q = \frac{Q_{sp} \cdot k \cdot DOC}{1 + k \cdot DOC} - b \quad (2)$$

In this study,  $Q_{sp}$ ,  $k$ , and  $b$  are parameters fitted to laboratory batch experiments where different amounts of DOC are added to a well-mixed soil sample.

Sorption of DOC is a nano-to-micro-scale process, but there is evidence that sorption potentials are related to macro-scale differences in climate and soil properties. For example, Mayes et al. (2012) found relationships between sorption potential measured by batch experiments and clay fraction, pH, and soil order. Kothawala et al. (2009) also found a relationship between sorption potential and clay fraction. Soil texture, especially clay and silt fraction, are a commonly-used proxy for mineral surface availability in empirical models of soil C sequestration (Angers et al. 2011; Hassink, 1997; Wiesmeier et al. 2018), as well as in process-based models of soil C cycling (Sulman et al. 2018). Changing the pH balance of soil (e.g., via root exudation) has direct effects on sorption to mineral surfaces (Keiluweit et al. 2015).

Soil order is not a measurable quantity like the clay fraction but rather is a taxonomic classification that synthesizes important functional information, such as the dominant or typical mineralogy, weathering process, environmental conditions, and vegetation cover. Sorption may also be affected by climate, either through indirect controls on mineral weathering (Gislason et al. 2009), direct control of temperature on sorption dynamics (Abramoff et al. 2019; Conant et al. 2011), or the influence of precipitation on DOC infiltration down the soil column (Kaiser and Guggenberger 2005). Climate may also have indirect effects on ecosystem properties affecting DOC production and availability such as net primary production and average decomposition rates.

In this paper, we estimate the potential for additional DOC sorption from a new compilation of equilibrium batch sorption experiments (Feng et al. 2014; Jagadamma et al. 2012; Kaiser et al. 1996;

D. Kothawala  
Department of Ecology and Genetics, Uppsala University,  
752 36 Uppsala, Sweden  
e-mail: dolly.kothawala@ebc.uu.se

M. A. Mayes  
Environmental Sciences Division & Climate Change  
Science Institute, Oak Ridge National Laboratory,  
Oak Ridge, TN 37831, USA  
e-mail: mayesma@ornl.gov

Kothawala et al. 2009; Mayes et al. 2012). We fitted the Langmuir equation parameters from those experiments to estimate the additional DOC sorption potential. We trained a machine learning model on the relationship between DOC sorption potential and widely-measured climate and geochemical variables. We then used this relationship to estimate the global DOC sorption potential ( $Q_{sp}$ ) for six soil orders (Alfisols, Entisols, Inceptisols, Mollisols, Spodosols, Ultisols), using data from 14,631 soil profiles distributed globally. Defining the potential contribution of DOC sorption to total soil organic C (SOC) stock is a first step to estimate the potential for soils to accrue stable C and to improve soil C models.

## Methods

### Compilation of batch experiment data

We compiled a database of published sorption experiments, where DOC was experimentally adsorbed to mineral soils using standard batch experiments (Feng et al. 2014; Jagadamma et al. 2012; Kaiser et al. 1996; Kothawala et al. 2009; Mayes et al. 2012). We selected these experiments based on the similarity of their methodology, as the methods of Feng et al. (2014) and Jagadamma et al. (2012) are based on Mayes et al. (2012), which is based on Kothawala et al. (2009), which is based on Kaiser et al. (1996). Thus, this database is not the result of meta-analysis, but rather represents a limited group of sorption experiments with similar methods. Soils were not fractionated but it is assumed that sorption to minerals occurs in the mineral fraction. DOC was added to mineral soil and allowed to equilibrate for 18–48 h, filtered, and analyzed for total C concentration before and after adsorption. To develop a relationship between the additional sorption potential ( $Q_{sp}$ ), equilibrium constant ( $k$ ), and soil geochemical variables, we used five studies with a total of 402 samples from 111 locations in North America and Europe, across six soil orders (Alfisols, Entisols, Inceptisols, Mollisols, Spodosols, Ultisols) representing 62% of the Earth's ice-free land surface (Feng et al. 2014; Jagadamma et al. 2012; Kaiser et al. 1996; Kothawala et al. 2009; Mayes et al. 2012). Samples from other soil orders were present in the original studies, but we included in this analysis only the soil orders for which there were more than 10

samples, Alfisols (N = 109), Entisols (N = 13), Inceptisols (N = 87), Mollisols (N = 71), Spodosols (N = 30), and Ultisols (N = 66). We also obtained reported  $Q_{sp}$  and  $k$  estimates from published literature as well as raw sorption data from the studies. We used published values where available (N = 275) and refitted values (N = 127) where only raw data were available. We refitted the available isotherm data to the Langmuir Eq. (1) (Kothawala et al. 2009; Mayes et al. 2012). Some studies adjusted for the amount of DOC released into solution when the initial DOC concentration is 0 mg/L,  $b$  from Eq. (2), measured using solution blanks, before fitting the model without  $b$  (Jagadamma et al. 2012; Mayes et al. 2012). Other studies fitted or explicitly reported  $b$ . Where  $b$  was fitted in the original study, we refitted the Langmuir equation in the same way, estimating  $Q_{sp}$ ,  $k$ , and  $b$ . We summarize reported estimates of  $b$  (N = 186) in the Results section, but we did not attempt to extrapolate this quantity for several reasons. First, this value is generally used to control for DOC release from native soil occurring at 0 mg/L. The DOC released from native soil under these conditions may not have been protected by sorption and therefore does not necessarily represent the desorption potential. Second, soil C models that use the Langmuir equation to represent sorption do not typically include this parameter. Lastly, preliminary analysis with a Random Forest model found that of the predictors in the laboratory experiment dataset  $b$  was primarily related to the organic carbon concentration rather than climatic or geochemical predictors.

The parameters  $Q_{sp}$  and  $k$  were fitted using non-linear regression in R (package *drc*; Ritz et al. 2015). Across the five studies, 140 of the reported  $Q_{sp}$  estimates and 86 of the reported  $k$  estimates had corresponding raw data that we were able to fit. Refitted values were correlated with those reported in the literature with minimal bias for  $Q_{sp}$  (slope = 0.81,  $R^2 = 0.74$ ,  $P < 0.05$ ) and  $k$  (slope = 0.96,  $R^2 = 0.83$ ,  $P < 0.05$ ; Figure S1). Differences between reported and refitted values are likely due to differences in the method used for curve fitting. During refitting, we identified and removed one outlier that was greater than three standard deviations from the mean.

In all studies but one, DOC was extracted from the soil organic horizon, leaf litter, or from stream water, while in Jagadamma et al. (2012), DOC solutions prepared using five C compounds were reacted with

each sample: glucose, L-alanine, oxalic acid, salicylic acid, and sinapyl alcohol (Jagadamma et al. 2012). For our database, we averaged the parameters derived from each compound for each sample.

Additional soil characteristics that were measured for > 90% of the samples collected were percent clay, pH (CaCl<sub>2</sub>), dithionite-extractable iron (Fe; mg/kg), and organic carbon content (mg/kg). A subset of the samples measured exchangeable calcium (Ca<sup>2+</sup>; cmol + /kg), oxalate- and pyrophosphate-extractable Fe, as well as dithionite-, oxalate-, and pyrophosphate-extractable aluminum (Al). Dithionite extraction was developed to extract Fe oxides because dithionite can reduce Fe<sup>3+</sup> to Fe<sup>2+</sup> but cannot dissolve Al oxides, and thus, any interpretation of Al dithionite extraction should consider that unlike the oxalate extraction, dithionite extraction was not developed to measure Al (McKeague and Day 1966).

We extracted mean annual temperature (MAT) and mean annual precipitation (MAP) for 1979–2000 from WorldClim Version 2 (Fick and Hijmans 2017) at 30-s resolution for each location where soil was collected for batch experiments. Some batch experiments directly reported MAT and MAP at the locations where soil was collected. Reported values for MAT were correlated with extracted values with minimal bias (slope = 0.94, R<sup>2</sup> = 0.89, N = 133), while reported values for MAP were slightly lower than extracted values (slope = 0.72, R<sup>2</sup> = 0.65, N = 131).

## Data analysis

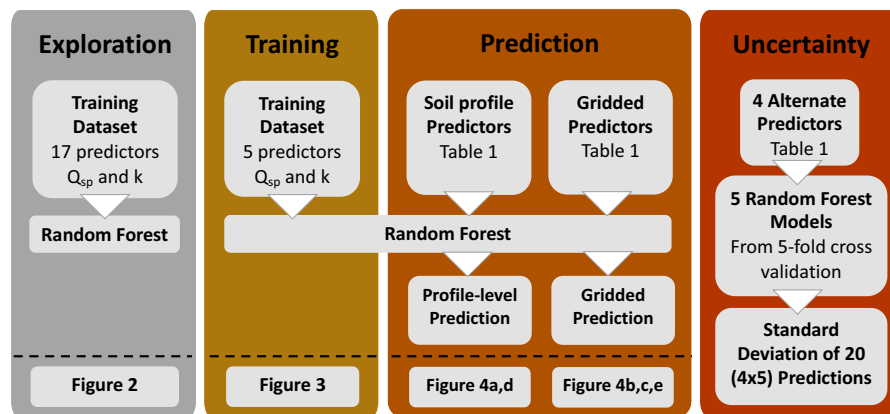
### *Data interpolation for missing values*

Many samples did not have measurements for every variable (0–86% missing; Figure S2), but many of the variables where measurements were not reported by the experiments were correlated with related variables where measurements were more complete (e.g., different types of metal extractions, different texture measurements; Figure S3). We estimated missing data using multiple imputation by chained equations. Missing data were assumed to have a multivariate normal distribution. Although a multivariate normal distribution may predict some values below zero, all of the interpolated values were positive so we did not alter this assumption. Some of the measured values used for imputation were greater than three standard deviations from the mean, but because all were

reasonable values for the measurements they represented, we did not remove or transform any data as outliers. Multiple values were drawn from this distribution for each missing datum by Markov-Chain Monte Carlo, conditional on the other data in the data set. We used the classification and regression tree (CART) method to estimate the missing value from the available data (Van Buuren and Groothuis-Oudshoorn 2011). The distributions of the estimated values were reasonably well-matched to the distributions of the measured values, though some of the most common values were over-represented (Figure S4). For the predictors carried forward into the global-scale analysis (i.e., percent clay, pH, and soil order), only 6% of data were missing.

### *Training for machine learning*

We trained Random Forest machine-learning algorithms (Liaw and Wiener 2002) to quantify the relative importance of different climate and soil characteristics in our observational dataset for predicting  $Q_{sp}$  and  $k$  for measured samples (Fig. 1: Exploration). Then, we built another Random Forest model of  $Q_{sp}$  and  $k$  using a subset of predictors (MAT, MAP, percent clay, pH, and soil order) which are available at global scales (Fig. 1: Training). For each Random Forest model, we log-transformed  $Q_{sp}$  and  $k$  to improve model performance. Model performance was evaluated using R<sup>2</sup> from permutation cross-validation, where the model is trained on 80% of the data and tested on the remaining 20%, and this procedure is repeated 99 times (package ‘rfUtilities’; Evans and Murphy 2018). Importance is defined as the mean increase in node purity (a measure minimizing the homogeneity of classes or labels) when the predictor is used to split regression trees in the model. We grew 500 trees for each model, and the number of variables tried at each split was set at the largest integer smaller than or equal to the number of predictors divided by three (Breiman 2001). We tested other values of the number of variables to try at each node, but the method above provided the best model fit. We used the ‘forestFloor’ package (Welling et al. 2016) in R Statistical Language to derive partial feature contributions for the predictors included in the  $Q_{sp}$  and  $k$  Random Forest regression models. Partial feature contributions show the partial response of the



**Fig. 1** Diagram of analyses performed. First, 17 predictors from the batch sorption experiments were used to explore the relationship between sorption parameters ( $Q_{sp}$  and  $k$ ) and environmental variables. Then, 5 predictors available at global scales were used to train a Random Forest model for global

estimation of  $Q_{sp}$  and  $k$ . Next, soil profile-level predictors from WoSIS and gridded predictors from SoilGrids were used to predict  $Q_{sp}$  and  $k$ . Lastly, we estimated the uncertainty due to model and predictor error by creating 20 new predictions using alternate datasets and models

dependent variable to changes in the independent variable across its range.

#### Estimating the DOC sorption potential at global scales

We used the model generated in the training step to estimate the spatial distribution of  $Q_{sp}$  and  $k$  at global scales using climatic and geochemical predictors from profile and gridded data in Table 1 (Fig. 1: Prediction). We estimated  $Q_{sp}$  (mg/kg) and  $k$  (L/mg) for 14,631 soil profiles from the World Soil Information Service (WoSIS) database (Batjes et al. 2017) (Table 1). We analyzed the whole soil profile from surface to 1 m depth, and repeated the analysis for 13,609 topsoil profiles (0–30 cm) and 13,993 subsoil profiles (30 cm–1 m) that reported depth-specific clay and pH values.

Both  $Q_{sp}$  and  $k$  were mapped globally using 10 km gridded fields of MAT and MAP, percentage clay, pH, soil order, bulk density, and SOC stock to 1 m (Table 1; Hengl et al. 2017). For each 10 km grid cell,  $Q_{sp}$  and  $k$  were estimated in the same way as described above for WoSIS profiles. To estimate a global  $Q_{sp}$  in Pg carbon for 6 soil orders, we converted  $Q_{sp}$  from concentrations to stocks using the bulk density and volume of coarse fragments following (Hengl et al. 2017; Tifafi et al. 2018) for each grid cell.

#### Estimating uncertainty

We estimated the uncertainty in the global-scale  $Q_{sp}$  by accounting for predictor error and model error (Fig. 1: Uncertainty). To estimate predictor error we used four additional gridded predictor datasets (Alternate MAT, MAP, %Clay and pH; Table 1). We substituted each alternate dataset for the standard predictor dataset to generate four additional predictions. To estimate model error, we generated 5 Random Forest models from fivefold cross validation. The overall error is estimated from the standard deviation of the  $Q_{sp}$  estimate from the 20 (4 datasets  $\times$  5 models) additional predictions.

## Results

### Influential variables controlling sorption in batch experiments

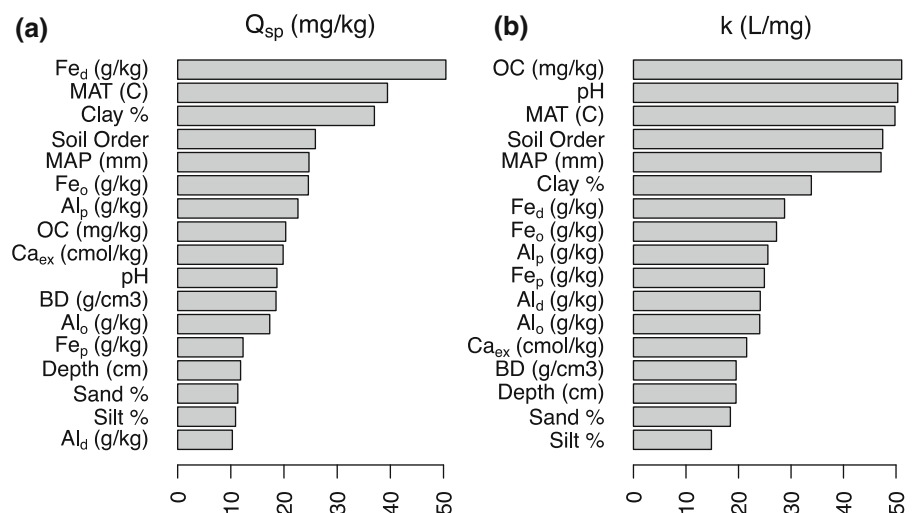
Our Random Forest models (Liaw and Wiener 2002) were used to identify the key climate and geochemical variables that influence  $Q_{sp}$  per unit soil mass (mg/kg) from the laboratory experiment measurements (Fig. 1: Exploration). According to this model, the most influential variables were dithionite-extractable Fe, mean annual temperature (MAT) and clay percentage ( $R^2 = 0.46$ , Fig. 2a). The equilibrium constant  $k$  (L/mg) was most strongly related to organic carbon

**Table 1** Description of profile and gridded data used for estimating  $Q_{sp}$  at global scales and alternate gridded data for uncertainty quantification of global totals of  $Q_{sp}$ 

Data type	Variable	Source	Unit	Resolution	References
Profile	% Clay	WoSIS	%	–	(Batjes et al. 2017)
	pH	WoSIS	–	–	(Batjes et al. 2017)
	Soil Order	WoSIS	–	–	(Batjes et al. 2017)
	Bulk Density	WoSIS	kg dm <sup>-3</sup>	–	(Batjes et al. 2017)
	SOC stock	WoSIS	t/ha	–	(Batjes et al. 2017)
Gridded	MAT (1970–2000)	WorldClim V2	°C	30 s	(Fick and Hijmans 2017)
	MAP (1970–2000)	WorldClim V2	mm	30 s	(Fick and Hijmans 2017)
	% Clay	SoilGrids	%	10 km	(Hengl et al. 2017)
	pH	SoilGrids	–	10 km	(Hengl et al. 2017)
	Soil Order	SoilGrids	–	10 km	(Hengl et al. 2017)
	Bulk Density	SoilGrids	kg dm <sup>-3</sup>	10 km	(Hengl et al. 2017)
	SOC stock	SoilGrids	t/ha	10 km	(Hengl et al. 2017)
	Alternate MAT (1973–2013)	CHELSA	°C	30 s	(Karger et al. 2017, 2018)
	Alternate MAP (1973–2013)	CHELSA	mm	30 s	(Karger et al. 2017, 2018)
	Alternate % Clay	GSDE	%	5 min	(Shangguan et al. 2014)
	Alternate pH	GSDE	–	5 min	(Shangguan et al. 2014)

The reported resolution is the original resolution of the dataset, all data were rescaled to 10 km

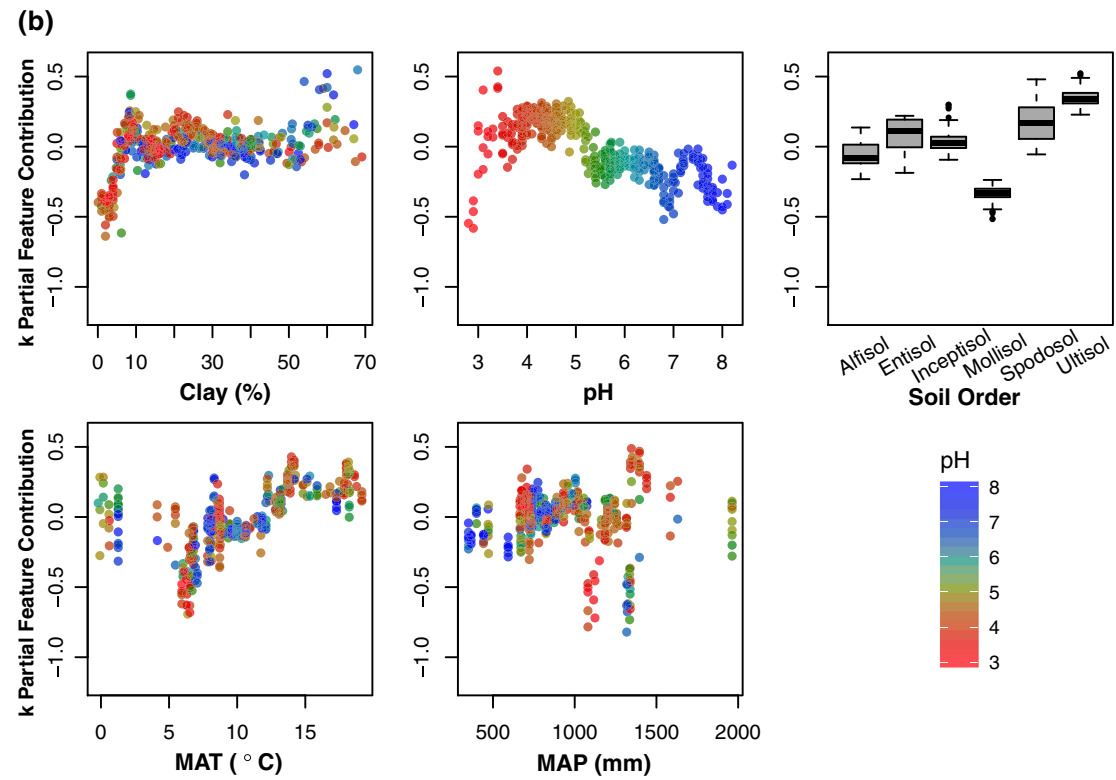
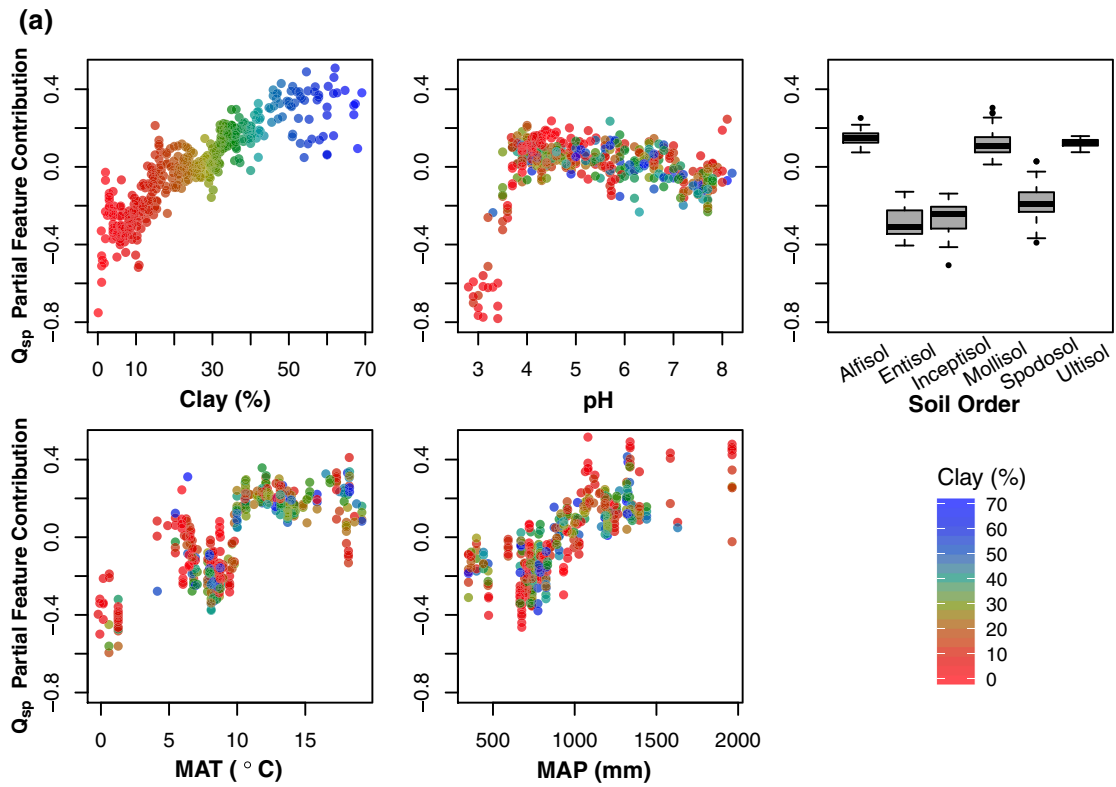
**Fig. 2** The mean increase in node purity when the predictor variable is used to split regression trees in the model for 17 predictors of **a** additional sorption potential ( $Q_{sp}$ ) in mg/kg and **b** the equilibrium constant ( $k$ ) in L/mg. The most influential variables are shown from the top to the bottom of the graph



content, pH, MAT, soil order, and mean annual precipitation (MAP) ( $R^2 = 0.32$ , Fig. 2b).

For global estimation of  $Q_{sp}$ , we built a Random Forest model with only the five predictors in Fig. 2 that were available at global scales (Fig. 1: Training). Percent clay had the highest mean increase in node purity for the globally-applied  $Q_{sp}$  model ( $R^2 = 0.38$ , Table S1), with a positive relationship between  $Q_{sp}$

and percent clay (Fig. 3a). There were also positive, mostly monotonic relationships between  $Q_{sp}$  and MAT, and between  $Q_{sp}$  and MAP.  $Q_{sp}$  had a threshold relationship with pH, with a low  $Q_{sp}$  in very acidic soils (pH < 4) and no relationship above that threshold (Fig. 3a). Soil order influenced  $Q_{sp}$ , with Alfisols, Mollisols, and Ultisols tending to have higher  $Q_{sp}$  relative to the other soil orders (Fig. 3a). The most





◀ **Fig. 3** Partial feature contributions of percent clay, pH, Soil Order, mean annual temperature (MAT), and mean annual precipitation (MAP) for predicting **a** the additional sorption potential ( $Q_{sp}$ ) and **b** the equilibrium constant ( $k$ ). Symbol colors reflect the position on the x-axis of the most important variable, **a** Clay (%) and **b** pH, respectively. For example in (a), sites with values of pH 3 are mostly red symbols, indicating that they also have low values of Clay

important parameter for estimating  $k$ , measured by mean increase in node purity, was pH ( $R^2 = 0.32$ , Table S1), which also had a threshold effect with low  $k$  in very acidic soils ( $\text{pH} < 4$ ) and a negative relationship between  $k$  and pH at  $\text{pH} > 4$  (Fig. 3b). There were non-monotonic relationships between  $k$  and MAT, MAP, and percent clay (Fig. 3b). Of the six soil orders, Spodosols and Ultisols were most likely to have higher values of  $k$  (Fig. 3b).

We could not use a measure of reactive or extractable Fe and Al to predict  $Q_{sp}$  at global scales because of the lack of sufficient gridded or multi-site data. Nevertheless, dithionite-extractable Fe was the most important predictor of  $Q_{sp}$  in our Random Forest model based on the sorption batch experiment data (Fig. 2a). Although many Fe and Al measurements in measured profiles were missing and had to be imputed, the correlations between extractable metals and  $Q_{sp}$  were more strongly positive than any other single predictor in the dataset (Figure S3). This finding, and the wealth of research about the importance of extractable Fe and Al for predicting C storage and mineral stabilization (Kaiser and Guggenberger 2000; Kalbitz and Kaiser 2008; Kramer and Chadwick 2019; Rasmussen et al. 2007; Schrumpp et al. 2013; Torn et al. 1997), suggests a strong need for a global dataset of those extractable metals.

The addition of other soil properties to the Random Forest model of  $Q_{sp}$ , especially extractable Fe and Al content, increased the variance explained by 8%. It is possible that adding other soil properties such as C, N, and other nutrient contents, bulk density, cation exchange capacity, base saturation, and soil moisture would increase the variance explained for the sorption potential. However, these variables were either not available or mostly missing in the training dataset, and would have had to have been estimated from map-based products, which may well represent the average value at large scales but not field-scale measurements. Soils in particular are very heterogeneous, and some

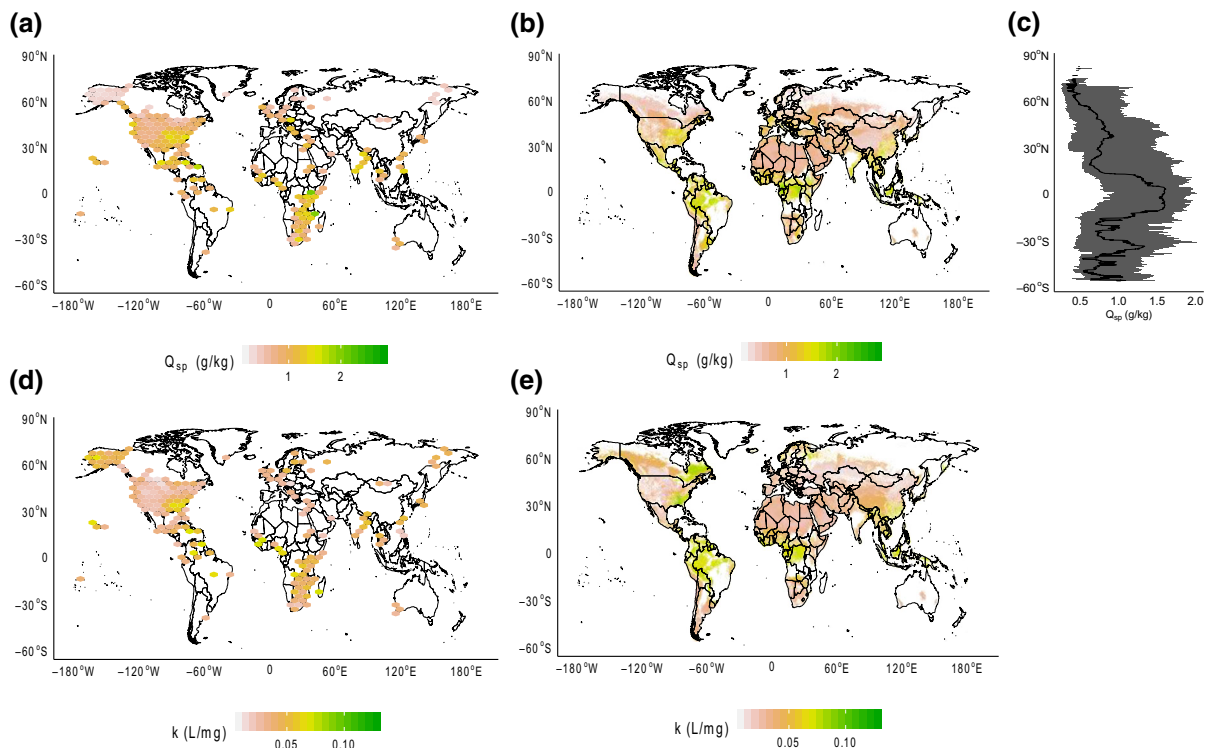
regionally or globally-applied models of soil organic C concentration ( $R^2 = 0.23$  in Hengl et al. 2014, improved to  $R^2 = 0.64$  in Hengl et al. 2017), SOC stock ( $R^2 = 0.54$  in Sanderman et al. 2017), SOC stock without land use ( $R^2 = 0.34$  in Sanderman et al. 2017), or dissolved organic C ( $R^2 = 0.36$  in Langeveld et al. 2020) explain less variance than models more related to plant processes such as mycorrhizal fungi type ( $R^2 \approx 0.5\text{--}0.8$  in Steidinger et al. 2019) or gross primary production ( $R^2 = 0.7$  in Tramontana et al. 2015). These limitations may be overcome with additional soil property measurements at the level where other measurements are collected, or a better understanding of soil spatial heterogeneity.

The amount of DOC released into solution when the initial DOC concentration is 0 mg/L,  $b$  from Eq. (2), has a median value of 0.08 (0.01, 0.77) g C kg soil<sup>-1</sup> and is positively correlated with organic carbon content ( $\rho = 0.69$ ), which is an order of magnitude lower than the amount that can be added to soil (i.e.,  $Q_{sp}$ ). Since this value is related to the amount of total C, it may include some desorbed C but may also include C released from particulate organic matter such as litter.

#### Influential variables controlling global-scale patterns of sorption

A map of carbon storage potential ( $Q_{sp}$ ) concentration in g C kg soil<sup>-1</sup> was generated by applying the global-scale Random Forest model to predictors of  $Q_{sp}$  for 14,631 soil profiles from the World Soil Information Service (WoSIS) database (Fig. 4a; Batjes et al. 2017), and for each grid cell of the globe using the SoilGrids data products (Fig. 4b; Hengl et al. 2017). Across a global range of soil profiles ( $N = 14,631$ ), the median potential of additional C storage from DOC sorption,  $Q_{sp}$ , was 1.1 g C kg soil<sup>-1</sup> (0.43 to 1.9 for the 95% Confidence Intervals; CI). DOC sorption potential values were summed to estimate a global additional sorption potential of  $107 \pm 13$  Pg C across 6 soil orders to 1 m depth, that is, a 7% increase in the current global SOC stock of these soil orders (1615 Pg; Table 2).

The additional DOC sorption potential  $Q_{sp}$  was highest in low- and mid-latitude soils with high clay content, and lowest in high-latitude soils dominated by organic matter (Fig. 4a–c). We found that high  $Q_{sp}$  values prevail in parts of eastern North America, the



**Fig. 4** **a, b** Additional sorption potential ( $Q_{sp}$ ) and **d, e** equilibrium constant ( $k$ ) across a global range of soil profiles ( $N = 14,631$ ) representing 6 soil orders. In **a, d** median values from each profile are binned hexagonally, and in **b, e** values are estimated using the SoilGrids gridded global data products.

**c** The additional sorption potential ( $Q_{sp}$ ) by latitude, estimated using SoilGrids gridded global data products. The black line is the median  $Q_{sp}$  and the gray shading represents the 5th and 95th percentile range

**Table 2** Summary of estimated quantities per volume of soil, the number of profiles used to estimate the quantity, the depth increment, and the global estimate for 6 soil orders

Quantity	Median value (95% CI) in g C/kg soil	Depth increment	Number of profiles	Global estimate (uncertainty) in Pg
$Q_{sp}$	1.1 (0.43, 1.9)	0–1 m	14,631	107 ( $\pm$ 13)
Existing SOC stock		0–1 m		1615
$k$	0.034 (0.014, 0.078)	0–1 m	14,631	–
$Q_{sp}$ topsoil	0.99 (0.43, 1.6)	0–30 cm	13,609	–
$Q_{sp}$ subsoil	1.14 (0.44, 1.98)	30 cm–1 m	13,993	–
$k$ topsoil	0.033 (0.014, 0.076)	0–30 cm	13,609	–
$k$ subsoil	0.035 (0.014, 0.076)	30 cm–1 m	13,993	–

CI confidence intervals

Amazon, central Africa, and Indonesia. These areas have a high percent clay, higher MAT, and moderately acidic pH (Figure S5a–c). High  $k$  occurred in areas where pH was between 4 and 5 (Figs. 3b, 4d, e).

Across the individual WoSIS profiles, Alfisols, Mollisols, and Ultisols had higher average  $Q_{sp}$  relative to other soil orders (Figure S6a). In the same profiles, Spodosols and Ultisols were estimated to have higher average values of  $k$  (Figure S6b). Although we

excluded Gelisols and Histosols from this analysis because they were poorly represented in the laboratory sorption experiments, high-latitude profiles from other soil orders were estimated to have consistently low  $Q_{sp}$ , implying that most mineral sites are already occupied by organic C. Treating grid cells as independent, there is a weak, negative correlation ( $\rho = 0.27$ ) between  $Q_{sp}$  in units of stock and total organic carbon stock ( $\text{kg m}^{-2}$ ), suggesting that soils with high initial organic C stocks will have less capacity to accrue more C via sorption. The negative influence of organic matter on sorption of various chemical species has been observed in previous sorption experiments (Johnson and Todd 1983; Klotzbücher et al. 2020; Redman et al. 2002).

We subset the soil profiles from the WoSIS database to look at the difference in additional DOC sorption potential between two soil depths representing topsoils (0–30 cm,  $N = 13,609$ ) and subsoils (30–100 cm,  $N = 13,993$ ). We found that  $Q_{sp}$  was higher in the subsoil (Figure S7a), while  $k$  was not different between topsoil and subsoil (Figure S7b). Subsoils therefore are found to have a greater capacity to sorb additional DOC, likely because mineral concentrations and therefore available mineral surface area increase with depth, while soil C concentrations decrease with depth (Figure S7c). We did not have sufficient data to extend our analysis to soils deeper than 1 m, but acknowledge there is substantial C stored at depths  $> 1$  m that, like soils at 30–100 cm, are likely to have higher mineral surface area relative to topsoils (Batjes et al. 1996; Jobbágy and Jackson 2000).

## Discussion

This study provides an estimate of the potential additional C that could be sorbed to minerals in soils spanning six soil orders and a range of climate conditions globally. Our best estimate of 107 Pg C should be interpreted as the maximum of additional C that can potentially be sorbed as DOC. Our estimated sorption potential for six soil orders is similar in magnitude to the 116 Pg estimated SOC stock lost from human agricultural activity over the last 12,000 years (Sanderman et al. 2017). Although we do not know the land use history of our soils, and disturbed soil may have a different DOC sorption

potential compared to undisturbed soil, this result suggests that it may be possible to compensate at least partially the loss of SOC due to soil degradation on a global scale, from the potential of soil to sorb more DOC. The inferred 7% increase in SOC stocks shows that if the DOC sorption potential could be realized for all soils, disturbed and undisturbed, it would sustain the ‘4 per 1000’ goal for 12 years.

Theoretically, some increase in soil C sequestration may be realistically achieved by increasing DOC concentration in the soil. This could be accomplished by increasing plant inputs, reducing DOC consumption by microorganisms, and changing soil moisture and water flow which affect local DOC concentration and leaching. The latter options are more complicated to manipulate, whereas land management options that increase C inputs into the soil are well-established; e.g., cover-crops, organic fertilizer amendments, moderate grazing or agroforestry (Cardinael et al. 2018; Maillard et al. 2017; Poeplau and Don 2015). These options may be particularly useful in the croplands of eastern North America and managed or converted forests of the tropics, which have high additional sorption potential (Fig. 4a, b). It would also be possible to increase soil C sequestration by changing the environmental conditions known to affect  $k$ , such as pH, keeping in mind that making modifications to pH may affect many other aspects of soil quality. Although beyond the scope of this study, the chemical composition of DOC (Jagadamma et al. 2012) and seasonality of inputs may also affect the sequestration potential. Nevertheless, DOC is generally highly reactive if not sorbed and any change in environmental conditions inducing DOC desorption may drastically increase soil  $\text{CO}_2$  emissions.

We do not argue that the sorption potential estimated in this study is fully realizable. In fact, the DOC sorption potential in field conditions is likely lower than that measured and extrapolated from batch experiments. Batch experiments, like most soil laboratory experiments, homogenize the soil before adding DOC, which disrupts soil structure and preferential flow paths that would limit the access of soil minerals to DOC in situ (Kaiser and Guggenberger 2005). Therefore, the sorption potential that we estimate here is a maximum assuming that mineral sites in soil are equally accessible to DOC, a condition which is rarely satisfied in the field.

On the other hand, batch sorption experiments may underestimate the C sequestration potential of DOC because they only consider DOC sorption which occurs on a fast timescale. Batch sorption experiments are brought to equilibrium over a period of 24–48 h. This equilibration is sufficient to associate DOC with soil minerals under saturated conditions but does not account for microbial turnover or structural changes, such as aggregation, that occur over longer periods of time. Also, soils contain a wide range of forms of organic matter, interacting with minerals along different pathways (Kaiser and Guggenberger 2007; Masiello et al. 2004; Mikutta et al. 2011; Weng et al. 2017). For example, microbial growth, consumption of DOC, and subsequent turnover into necromass can contribute to mineral-associated C as well as total SOC (Kallenbach et al. 2016). Therefore, a larger increase of SOC than that observed in the batch sorption experiments might be possible due to mechanisms that operate over longer timescales.

Though DOC sorption is a fast process, the rate of C sequestration by this mechanism is still limited by the amount of nitrogen and phosphorus required to sequester C as soil organic matter. Recent discussions have pointed out that it would require  $593 \text{ Tg N year}^{-1}$  (assuming a soil C:N of 15) and  $35\text{--}75 \text{ Tg P year}^{-1}$  to achieve the '4 per 1000' goal of sequestering  $8.9 \text{ Pg C year}^{-1}$ , the rate of anthropogenic emissions (Davies et al. 2020; Spohn 2020). The amount of N required is greater than global natural and anthropogenic nitrogen fixation combined,  $413 \text{ Tg N year}^{-1}$  (Fowler et al. 2013), and the amount of phosphorus required is a substantial proportion of global phosphate rock mined each year,  $240 \text{ Tg P year}^{-1}$  (USGS 2020). As a result, soil C sequestration that draws on coupled nutrients cycles is likely only achievable at a rate well below that of anthropogenic emissions.

Our results also have implications for models. Soil decomposition models that use Langmuir sorption to represent all interactions in the mineral fraction typically have  $k$  constants orders of magnitude higher, e.g.,  $2.5 \times 10^{-1}$  (Abramoff et al. 2018; Robertson et al. 2019) and  $9 \times 10^3 \text{ m}^3/\text{g C}$  (Ahrens et al. 2015), than the values estimated from the 402 batch sorption experiments in this study ( $10^{-4}$  to  $10^{-2} \text{ m}^3/\text{g C}$ ), suggesting that processes other than sorption contribute to C accumulation in the mineral fraction (Abramoff et al. 2018; Wang et al. 2013) and that models may over-tune sorption as a process to create C

storage. Mineral-associated C is itself operationally-defined, with many potential mechanisms of physical and chemical fractionation used (Poeplau et al. 2018). As a result, mineral-associated C may not only form upon direct sorption but also coprecipitation (Mikutta et al. 2011), organic-organic interactions (Kleber et al. 2007), or occlusion within aggregates (Rasmussen et al. 2005; Six and Paustian 2014). These additional mechanisms can reduce the amount of mineral-associated C that is in contact with aqueous DOC, and therefore reduce the amount that can be readily exchanged (Kleber et al. 2007; Zhuang et al. 2008). Models are emerging that attempt to partition these varied processes into categories that can be more directly measured (Abramoff et al. 2018; Robertson et al. 2019).

According to our results, a greater proportion of soil C can be sorbed to minerals in highly weathered soils, (e.g., Ultisols), relative to less weathered soils (e.g., Inceptisols; Figure S6a). Presumably this is due to the increased prevalence of mineral surfaces per volume of soil in highly weathered soils relative to less weathered soils. However, we did not consider how the types of clay minerals found in soils with different weathering status may affect our predictions. For example, clay minerals commonly found in highly weathered soils (e.g., kaolinite) can have low reactivity (Doetterl et al. 2018) which may reduce their effective sorption potential per volume of soil relative to less weathered soils. Of the soil orders not included in this analysis, some are expected to have relatively high sorption potential, because they are characterized by the presence of secondary minerals (Andisols and Oxisols), or by high clay content (Vertisols). Others we would hypothesize to have low sorption potential due to environmental limitations on weathering (Aridisols), or strong environmental controls on C storage (Gelisols, Histosols). Future work to expand measurements in these soil orders could answer questions about the role of weathering and clay minerals while also contributing to a complete estimate of the global sorption potential.

Our study points to important gaps in our understanding of stabilized organic C. How much of the theoretical potential of stable C storage is realizable? For example, Sanderman et al. (2017) found that human agricultural activity has created a C deficit of  $116 \text{ Pg}$  globally, but suggests that only part of this amount is recoverable through management. What is

the importance of sorption relative to other mechanisms for soils under different climatic or management regimes? Resolving these knowledge gaps will be necessary if soils are to contribute to the resolution of the global C imbalance due to burning of fossil fuels.

**Acknowledgements** RZA was supported by the French government grant “Make Our Planet Great Again” and by a Marie Skłodowska-Curie Individual Fellowship (Grant No. 834-169) from the European Union’s Horizon 2020 program. KG was supported by a US Department of Agriculture NIFA Postdoctoral Fellowship. BG was supported by French government under the ANR “Investissements d’avenir” program with the reference CLAND ANR-16-CONV-0003. MAM’s participation in this research was supported through an Early Career Award from the United States Department of Energy, Office of Science, Office of Biological and Environmental Research. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the United States Department of Energy under contract DE-AC05-00OR22725.

**Author contributions** RZA, KG, BG, MST, and PC designed the study. WF, SJ, KK, DK, and MAM provided data. RZA performed the analyses and wrote the first draft. All authors contributed extensively to writing the manuscript.

**Data and code availability** The sorption experiment database and scripts used to generate figures for this paper can be found in the Github repository [github.com/rabramoff/DOC\_potential].

#### Compliance with ethical standards

**Conflict of interest** The authors declare no conflict of interest.

**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

Abramoff R, Xu X, Hartman M, O’Brien S, Feng W, Davidson E et al (2018) The Millennial model: in search of measurable pools and transformations for modeling soil carbon in the new century. *Biogeochemistry* 137(1–2):51–71. <https://doi.org/10.1007/s10533-017-0409-7>

- Abramoff RZ, Torn MS, Georgiou K, Tang J, Riley WJ (2019) Soil organic matter temperature sensitivity cannot be directly inferred from spatial gradients. *Glob Biogeochem Cycles* 33:761–776. <https://doi.org/10.1029/2018GB006001>
- Ahrens B, Braakhekke MC, Guggenberger G, Schrumph M, Reichstein M (2015) Contribution of sorption, DOC transport and microbial interactions to the 14C age of a soil organic carbon profile: insights from a calibrated process model. *Soil Biol Biochem* 88(June):390–402. <https://doi.org/10.1016/j.soilbio.2015.06.008>
- Angers DA, Arrouays D, Saby NPA, Walter C (2011) Estimating and mapping the carbon saturation deficit of French agricultural topsoils. *Soil Use Manag* 27(December):448–452. <https://doi.org/10.1111/j.1475-2743.2011.00366.x>
- Batjes NH, Reference IS, Isric IC, Box PO, Wageningen AJ (1996) Total carbon and nitrogen in the soils of the world. *Eur J Soil Sci* 47(June):151–163
- Batjes NH, Ribeiro E, Oostrum AV, Leenaars J, Hengl T, De Jesus JM (2017) WoSIS: providing standardised soil profile data for the world. *Earth Syst Sci Data* 9:1–14. <https://doi.org/10.5194/essd-9-1-2017>
- Bengtson P, Bengtsson G (2007) Rapid turnover of DOC in temperate forests accounts for increased CO<sub>2</sub> production at elevated temperatures. *Ecol Lett* 10(9):783–790. <https://doi.org/10.1111/j.1461-0248.2007.01072.x>
- Breiman L (2001) Random forests. *Mach Learn*. <https://doi.org/10.1201/9780367816377-11>
- Camino-Serrano M, Guenet B, Luysaert S, Ciais P, Bastrikov V, De Vos B et al (2018) ORCHIDEE-SOM: modeling soil organic carbon (SOC) and dissolved organic carbon (DOC) dynamics along vertical soil profiles in Europe. *Geosci Model Dev* 11:937–957. <https://doi.org/10.5194/gmd-2017-255>
- Cardinael R, Umulisa V, Toudert A, Olivier A, Bockel L, Bernoux M (2018) Revisiting IPCC Tier 1 coefficients for soil organic and biomass carbon storage in agroforestry systems. *Environ Res Lett*. <https://doi.org/10.1088/1748-9326/aab5f>
- Conant RT, Ryan MG, Agren GI, Birge HE, Davidson EA, Eliasson PE et al (2011) Temperature and soil organic matter decomposition rates—synthesis of current knowledge and a way forward. *Glob Change Biol* 17:3392–3404. <https://doi.org/10.1111/j.1365-2486.2011.02496.x>
- Cotrufo MF, Wallenstein MD, Boot CM, Deneff K, Paul E (2013) The Microbial Efficiency-Matrix Stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter? *Glob Change Biol* 19(4):988–995. <https://doi.org/10.1111/gcb.12113>
- Cotrufo MF, Ranalli MG, Haddix ML, Six J, Lugato E (2019) Soil carbon storage informed by particulate and mineral-associated organic matter. *Nat Geosci* 12:989–994. <https://doi.org/10.1038/s41561-019-0484-6>
- Davies CA, Robertson AD, McNamara NP (2020) The importance of nitrogen for net carbon sequestration when considering natural climate solutions. *Glob Change Biol*. <https://doi.org/10.1111/gcb.15381>
- Doetterl S, Berhe AA, Arnold C, Bodé S, Fiener P, Finke P et al (2018) Links among warming, carbon and microbial

- dynamics mediated by soil mineral weathering. *Nat Geosci* 11(8):589–593. <https://doi.org/10.1038/s41561-018-0168-7>
- Drake JE, Gallet-Budynek A, Hofmockel KS, Bernhardt ES, Billings SA, Jackson RB et al (2011) Increases in the flux of carbon belowground stimulate nitrogen uptake and sustain the long-term enhancement of forest productivity under elevated CO<sub>2</sub>. *Ecol Lett* 14(4):349–357. <https://doi.org/10.1111/j.1461-0248.2011.01593.x>
- Eusterhues K, Rumpel C, Kleber M, Ko I (2003) Stabilisation of soil organic matter by interactions with minerals as revealed by mineral dissolution and oxidative degradation. *Org Geochem* 34:1591–1600. <https://doi.org/10.1016/j.orggeochem.2003.08.007>
- Evans JS, Murphy MA (2018) rfUtilities: random forests model selection and performance evaluation. R package version 2.1-3. Retrieved from <https://cran.r-project.org/web/packages/rfUtilities/>
- Feng W, Plante AF, Aufdenkampe AK, Six J (2014) Soil organic matter stability in organo-mineral complexes as a function of increasing C loading. *Soil Biol Biochem* 69:398–405. <https://doi.org/10.1016/j.soilbio.2013.11.024>
- Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* 4315(37):4302–4315. <https://doi.org/10.1002/joc.5086>
- Fowler D, Coyle M, Skiba U, Sutton MA, Cape JN, Reis S et al (2013) The global nitrogen cycle in the twenty-first century. *Philos Trans R Soc B*. <https://doi.org/10.1098/rstb.2013.0164>
- Fu J, Gasche R, Wang N, Lu H, Butterbach-Bahl K, Kiese R (2019) Dissolved organic carbon leaching from montane grasslands under contrasting climate, soil and management conditions. *Biogeochemistry* 145(1–2):47–61. <https://doi.org/10.1007/s10533-019-00589-y>
- Gislason SR, Oelkers EH, Eiriksdottir ES, Kardjilov MI, Gisladottir G, Sigfusson B et al (2009) Direct evidence of the feedback between climate and weathering. *Earth Planet Sci Lett* 277(1–2):213–222. <https://doi.org/10.1016/j.epsl.2008.10.018>
- Gleixner G, Czimeczik CJ, Kramer C, Lühker B, Schmidt MWI (2001) Plant compounds and their turnover and stabilization as soil organic matter. *Global biogeochemical cycles in the climate system*. Elsevier, Amsterdam, pp 201–215
- Gleixner G, Poirier N, Bol R, Balesdent J (2002) Molecular dynamics of organic matter in a cultivated soil. *Org Geochem* 33(3):357–366. [https://doi.org/10.1016/S0146-6380\(01\)00166-8](https://doi.org/10.1016/S0146-6380(01)00166-8)
- Hassink J (1997) The capacity of soils to preserve organic C and N by their association with clay and silt particles. *Plant Soil* 191:77–87
- Hengl T, de Jesus JM, MacMillan RA, Batjes NH, Heuvelink GB, Ribeiro E et al (2014) SoilGrids1km—global soil information based on automated mapping. *PLoS ONE* 9(8):e105992. <https://doi.org/10.1371/journal.pone.0105992>
- Hengl T, De Jesus JM, Heuvelink GBM, Ruiperez M, Kilibarda M, Blagoti A et al (2017) SoilGrids250m: global gridded soil information based on machine learning. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0169748>
- Jagadamma S, Mayes MA, Phillips JR (2012) Selective sorption of dissolved organic carbon compounds by temperate soils. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0050434>
- Jiang M, Medlyn BE, Drake JE, Duursma RA, Anderson IC, Barton CVM et al (2020) The fate of carbon in a mature forest under carbon dioxide enrichment. *Nature* 580(7802):227–231. <https://doi.org/10.1038/s41586-020-2128-9>
- Jobbágy EG, Jackson RB (2000) The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecol Appl* 10(2):423–436. [https://doi.org/10.1890/1051-0761\(2000\)010\[0423:TVDOSO\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2)
- Johnson DW, Todd DE (1983) Relationships among iron, aluminum, carbon, and sulfate in a variety of forest soils. *Soil Sci Soc Am J* 47(4):792–800
- Kaiser K, Guggenberger G (2000) The role of DOM sorption to mineral surfaces in the preservation of organic matter in soils. *Org Geochem* 31:711–725
- Kaiser K, Guggenberger G (2005) Storm flow flushing in a structured soil changes the composition of dissolved organic matter leached into the subsoil. *Geoderma* 127(3–4 SPEC. ISS.):177–187. <https://doi.org/10.1016/j.geoderma.2004.12.009>
- Kaiser K, Guggenberger G (2007) Sorptive stabilization of organic matter by microporous goethite: sorption into small pores vs. surface complexation. *Eur J Soil Sci* 58(1):45–59. <https://doi.org/10.1111/j.1365-2389.2006.00799.x>
- Kaiser K, Kalbitz K (2012) Cycling downwards—dissolved organic matter in soils. *Soil Biol Biochem* 52:29–32. <https://doi.org/10.1016/j.soilbio.2012.04.002>
- Kaiser K, Guggenberger G, Zech W (1996) Sorption of DOM and DOM fractions to forest soils. *Geoderma* 74:281–303
- Kaiser K, Eusterhues K, Rumpel C, Guggenberger G, Kögelknabner I, Bayreuth U et al (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* 165:451–459
- Kalbitz K, Kaiser K (2008) Contribution of dissolved organic matter to carbon storage in forest mineral soils. *Soil Biol Biochem* 1090:52–60. <https://doi.org/10.1002/jpln.200700043>
- Kallenbach CM, Grandy A, Frey SD (2016) Direct evidence for microbial-derived soil organic matter formation and its ecophysiological controls. *Nat Commun*. <https://doi.org/10.1038/ncomms13630>
- Karger DN, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-auza RW et al (2017) Data descriptor: climatologies at high resolution for the earth’s land surface areas. *Nature Publishing Group*. <https://doi.org/10.1038/sdata.2017.122>
- Karger D, Conrad O, Böhner J, Kawohl T, Kreft H, Soria-Auza R et al (2018) Data from: climatologies at high resolution for the earth’s land surface areas. *Dryad Digital*. <https://doi.org/10.5061/dryad.kd1d4>
- Keiluweit M, Bougoure JJ, Nico PS, Pett-Ridge J, Weber PK, Kleber M (2015) Mineral protection of soil carbon counteracted by root exudates. *Nat Clim Change* 5(6):588–595. <https://doi.org/10.1038/nclimate2580>
- Kiem R, Kögel-Knabner I (2002) Refractory organic carbon in particle-size fractions of arable soils II: organic carbon in

- relation to mineral surface area and iron oxides in fractions < 6 m m. *Org Geochem* 33:1699–1713
- Kleber M, Sollins P, Sutton R (2007) A conceptual model of organo-mineral interactions in soils: self-assembly of organic molecular fragments into zonal structures on mineral surfaces. *Biogeochemistry* 85(1):9–24. <https://doi.org/10.1007/s10533-007-9103-5>
- Kleber M, Eusterhues K, Keiluweit M, Mikutta C, Mikutta R, Nico PS (2015) Mineral-organic associations: formation, properties, and relevance in soil environments. *Advances in agronomy*, vol 130. Elsevier Ltd., Amsterdam. <https://doi.org/10.1016/bs.agron.2014.10.005>
- Klotzbücher T, Treptow C, Kaiser K, Klotzbücher A, Mikutta R (2020) Sorption competition with natural organic matter as mechanism controlling silicon mobility in soil. *Sci Rep* 10(1):1–11. <https://doi.org/10.1038/s41598-020-68042-x>
- Kogel-Knabner I (1986) Estimation and decomposition pattern of the lignin component in forest humus layers. *Soil Biol Biochem* 18(6):589–594
- Kogel-Knabner I, Wolfgang Z, Hatcher PG (1988) Chemical composition of the organic matter in forest soils: the humus layer. *Z Pflanzenernähr Bodenkd* 151:331–340
- Kögel-Knabner I, Guggenberger G, Kleber M, Kandeler E, Kalbitz K, Scheu S et al (2008) Organo-mineral associations in temperate soils: integrating biology, mineralogy, and organic matter chemistry. *J Plant Nutr Soil Sci* 171(1):61–82. <https://doi.org/10.1002/jpln.200700048>
- Kothawala D, Moore TR, Hendershoot W (2009) Soil properties controlling the adsorption of dissolved organic carbon to mineral soils. *SSSAJ*. <https://doi.org/10.2136/sssaj2008.0254>
- Kramer MG, Chadwick OA (2019) Climate-driven thresholds in reactive mineral retention of soil carbon at the global scale. *Nat Clim Change*. <https://doi.org/10.1038/s41558-018-0341-4>
- Langeveld J, Bouwman AF, Joost W, Lauriane VH, Beusen AHW, Mogollón JM, Middelburg JJ (2020) Estimating dissolved carbon concentrations in global soils: a global database and model. *SN Appl Sci*. <https://doi.org/10.1007/s42452-020-03290-0>
- Langmuir I (1918) The adsorption of gases on plane surfaces of glass, mica and platinum. *J Am Chem Soc* 40(9):1361–1403
- Liaw A, Wiener M (2002) Classification and regression by randomForest. *R News* 2(3):18–22
- Lilienfein J, Qualls RG, Uselman SM, Bridgham SD (2004) Adsorption of dissolved organic carbon and nitrogen in soils of a weathering chronosequence. *Soil Sci Soc Am J* 68(1):292–305. <https://doi.org/10.2136/sssaj2004.2920>
- Maillard É, McConkey BG, Angers DA (2017) Increased uncertainty in soil carbon stock measurement with spatial scale and sampling profile depth in world grasslands: a systematic analysis. *Agric Ecosyst Environ* 236:268–276. <https://doi.org/10.1016/j.agee.2016.11.024>
- Masiello CA, Chadwick OA, Southon J, Torn MS, Harden JW (2004) Weathering controls on mechanisms of carbon storage in grassland soils. *Glob Biogeochem Cycles* 18(4):1–9. <https://doi.org/10.1029/2004GB002219>
- Mayes MA, Heal KR, Brandt CC, Phillips JR, Jardine PM (2012) Relation between soil order and sorption of dissolved organic carbon in temperate soils. *Soil Sci Soc Am J* 76(1):61–69. <https://doi.org/10.2136/sssaj>
- McKeague J, Day J (1966) Dithionite- and oxalate-extractable Fe and Al as aids in differentiating various classes of soils. *Can J Soil Sci* 46(154):13–22
- Mikutta R, Zang U, Chorover J, Haumaier L, Kalbitz K (2011) Stabilization of extracellular polymeric substances (*Bacillus subtilis*) by adsorption to and coprecipitation with Al forms. *Geochim Cosmochim Acta* 75(11):3135–3154. <https://doi.org/10.1016/j.gca.2011.03.006>
- Minasny B, Malone BP, McBratney AB, Angers DA, Arrouays D, Chambers A et al (2017) Soil carbon 4 per mille. *Geoderma* 292:59–86. <https://doi.org/10.1016/j.geoderma.2017.01.002>
- Nakhavali M, Friedlingstein P, Lauerwald R, Tang J, Chadburn S, Camino-Serrano M et al (2017) Representation of dissolved organic carbon in the JULES land surface model (vn4.4\_JULES-DOCM). *Geosci Model Dev Discuss*. <https://doi.org/10.5194/gmd-2017-172>
- Palmroth S, Oren R, McCarthy HR, Johnsen KH, Finzi AC, Butnor JR et al (2006) Aboveground sink strength in forests controls the allocation of carbon below ground and its [CO<sub>2</sub>]-induced enhancement. *Proc Natl Acad Sci USA* 103(51):19362–19367. <https://doi.org/10.1073/pnas.0609492103>
- Poeplau C, Don A (2015) Agriculture, ecosystems and environment carbon sequestration in agricultural soils via cultivation of cover crops—a meta-analysis. *Agric Ecosyst Environ* 200:33–41. <https://doi.org/10.1016/j.agee.2014.10.024>
- Poeplau C, Don A, Six J, Kaiser M, Benbi D, Chenu C et al (2018) Isolating organic carbon fractions with varying turnover rates in temperate agricultural soils—a comprehensive method comparison. *Soil Biol Biochem* 125(July):10–26. <https://doi.org/10.1016/j.soilbio.2018.06.025>
- Porras RC, Pries CEH, Torn MS, Nico PS (2018) Synthetic iron (hydr) oxide-glucose associations in subsurface soil effects on decomposability of mineral associated carbon. *Sci Total Environ* 613–614:342–351. <https://doi.org/10.1016/j.scitotenv.2017.08.290>
- Rasmussen C, Torn MS, Southard RJ (2005) Mineral assemblage and aggregates control carbon dynamics in a California Conifer Forest. *Soil Sci Soc Am J*. <https://doi.org/10.2136/sssaj2005.0040>
- Rasmussen C, Southard RJ, Horwath WR (2007) Soil mineralogy affects conifer forest soil carbon source utilization and microbial priming. *Soil Sci Soc Am J* 71(4):1141–1150. <https://doi.org/10.2136/sssaj2006.0375>
- Rasse DP, Dignac MF, Bahri H, Rumpel C, Mariotti A, Chenu C (2006) Lignin turnover in an agricultural field: from plant residues to soil-protected fractions. *Eur J Soil Sci* 57(4):530–538. <https://doi.org/10.1111/j.1365-2389.2006.00806.x>
- Redman AD, Macalady DL, Ahmann D (2002) Natural organic matter affects Arsenic speciation and sorption onto hematite. *Environ Sci Technol* 36(13):2889–2896. <https://doi.org/10.1021/es0112801>
- Reynolds B, Fenner N (2001) Export of organic carbon from peat soils. *Nature* 412:785–786

- Ritz C, Baty F, Streibig J, Gerhard D (2015) Dose-response analysis using R. *PLoS ONE* 10(12):e0146021
- Robertson AD, Paustian K, Ogle S, Wallenstein MD, Lugato E, Cotrufo MF (2019) Unifying soil organic matter formation and persistence frameworks: the MEMS model. *Biogeosci Discuss* 16:1225–124836. <https://doi.org/10.5194/bg-2018-430>
- Sanderman J, Hengl T, Fiske GJ (2017) Soil carbon debt of 12,000 years of human land use. *Proc Natl Acad Sci USA*. <https://doi.org/10.1073/pnas.1706103114>
- Schmidt M, Rumpel C, Kogel-Knabner I (1999) Evaluation of an ultrasonic dispersion procedure to isolate primary organomineral complexes from soils. *Eur J Soil Sci* 50(March):87–94
- Schmidt MW, Torn MS, Abiven S, Dittmar T, Guggenberger G, Janssens IA, Kleber M et al (2011) Persistence of soil organic matter as an ecosystem property. *Nature* 478(7367):49–56. <https://doi.org/10.1038/nature10386>
- Schrumpf M, Kaiser K, Guggenberger G, Persson T, Kogel-Knabner I, Schulze ED (2013) Storage and stability of organic carbon in soils as related to depth, occlusion within aggregates, and attachment to minerals. *Biogeosciences* 10:1675–1691. <https://doi.org/10.5194/bg-10-1675-2013>
- Shangguan W, Dai Y, Duan Q, Liu B, Yuan H (2014) A global soil data set for earth system modeling. *J Adv Model Earth Syst* 6(1):249–263. <https://doi.org/10.1002/2013MS000293>
- Six J, Paustian K (2014) Aggregate-associated soil organic matter as an ecosystem property and a measurement tool. *Soil Biol Biochem* 68:A4–A9. <https://doi.org/10.1016/j.soilbio.2013.06.014>
- Spohn M (2020) Increasing the organic carbon stocks in mineral soils sequesters large amounts of phosphorus. *Glob Change Biol*. <https://doi.org/10.1111/gcb.15154>
- Steidinger BS, Liang J, Van Nuland ME, Werner GDA, Nabuurs G, Zhou M et al (2019) Climatic controls of decomposition drive the global biogeography of forest-tree symbioses. *Nature*. <https://doi.org/10.1038/s41586-019-1128-0>
- Sulman B, Moore J, Abramoff R, Averill C, Kivlin S, Georgiou K et al (2018) Multiple models and experiments underscore large uncertainty in soil carbon dynamics. *Biogeochemistry* 141:109–123
- Tang J, Riley WJ (2015) Weaker soil carbon–climate feedbacks resulting from microbial and abiotic interactions. *Nat Clim Change*. <https://doi.org/10.1038/nclimate2438>
- Tifafi M, Guenet B, Hatté C (2018) Large differences in global and regional total soil carbon stock estimates based on soilgrids, HWSD, and NCSCD: intercomparison and evaluation based on field data from USA, England, Wales, and France. *Glob Biogeochem Cycles*. <https://doi.org/10.1002/2017GB005678>
- Tisdall J, Oades J (1982) Organic matter and water-stable aggregates in soils. *J Soil Sci* 33:141–163. <https://doi.org/10.1111/j.1365-2389.1982.tb01755.x>
- Torn MMS, Trumbore SSE, Chadwick OAO, Vitousek PPM, Hendricks DDM (1997) Mineral control of soil organic carbon storage and turnover. *Nature* 389:3601–3603. <https://doi.org/10.1038/38260>
- Tramontana G, Ichii K, Camps-valls G, Tomelleri E, Papale D (2015) Uncertainty analysis of gross primary production upscaling using random forests, remote sensing and eddy covariance data. *Remote Sens Environ* 168:360–373. <https://doi.org/10.1016/j.rse.2015.07.015>
- USGS (2020) Phosphate Rock Data Sheet. Mineral Commodities Summary 2020. Retrieved from <https://www.usgs.gov/centers/nmic/phosphate-rock-statistics-and-information>
- Van Buuren S, Groothuis-Oudshoorn KMICE (2011) Multi-variate imputation by chained equations. *J Stat Softw* 45(3):1–67
- Wang G, Post WM, Mayes MA (2013) Development of microbial-enzyme-mediated decomposition model parameters through steady-state and dynamic analyses. *Ecol Appl* 23(1):255–272. <https://doi.org/10.1890/12-0681.1>
- Welling SH, Refsgaard HHF, Brockhoff PB, Clemmensen LKH (2016) Forest floor visualizations of random forests. *ArXiv E-Prints*.
- Weng Z, Van Zwieten L, Singh BP, Tavakkoli E, Joseph S, Macdonald LM et al (2017) Biochar built soil carbon over a decade by stabilizing rhizodeposits. *Nat Clim Change* 7(5):371–376. <https://doi.org/10.1038/nclimate3276>
- Wiesmeier M, Lungu M, Cerbari V, Boincean B, Hubner R, Kogel-Knabner I (2018) Rebuilding soil carbon in degraded steppe soils of Eastern Europe. *Land Degrad Dev*. <https://doi.org/10.1002/ldr.2902>
- Zhuang J, Mccarthy JF, Perfect E (2008) Soil water hysteresis in water-stable microaggregates as affected by organic matter. *Soil Water Manag Conserv* 72(1):212–220. <https://doi.org/10.2136/sssaj2007.0001S6>

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