Geoderma 469 (2026) 117801

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

Empirical evaluation and recommendations for equivalent soil mass adjustment of soil organic carbon stocks

Eric Potash [a,b,*], Adam C. von Haden [c], Luke Bergschneider [d], Michael S. Douglass [e], Lenarth Ferrari [e], Rosa Ibarra Lopez [e], Adriana Reconco Martinez [e], Andrew J. Margenot [a,e,*]

[a] Agroecosystem Sustainability Center, Institute for Sustainability, Energy, and Environment, University of Illinois Urbana-Champaign, Urbana, IL, USA
[b] Department of Natural Resources and Environmental Sciences, College of Agricultural, Consumer and Environmental Sciences, University of Illinois Urbana-Champaign, Urbana, IL, USA
[c] Department of Soil and Environmental Sciences, University of Wisconsin-Madison, Madison, WI, USA
[d] Institute for Sustainability, Energy, and Environment, University of Illinois Urbana-Champaign, Urbana, IL, USA
[e] Department of Crop Sciences, College of Agricultural, Consumer and Environmental Sciences, University of Illinois Urbana-Champaign, Urbana, IL, USA

ARTICLE INFO

Handling Editor: Dr. Cornelia Rumpel

Keywords:
Soil carbon
Soil carbon change
Equivalent soil mass
Bulk density
Systematic error

ABSTRACT

Measuring changes in fixed depth (FD; e.g. 0–30 cm) soil organic carbon (SOC) stocks is problematic if soils have compacted or expanded, i.e. if bulk density (BD) has changed. In theory, equivalent soil mass (ESM) remedies this. But in practice ESM is an approximation whose accuracy (systematic error) has not been evaluated empirically or compared to the FD baseline. We sampled 72 soil cores (0–75 cm) across six sites (central USA) of two regionally representative soils (Mollisols and Alfisols). SOC concentration and BD were measured at 2.5 cm depthwise intervals (dataset provided) to support this novel empirical evaluation of ESM accuracy. Under moderate BD changes (± 2.5 cm, or about 0.1 g cm^{-3}), 0–30 cm FD measurements had systematic errors up to 8% (5 Mg ha^{-1}), which would substantially distort SOC stock change estimates. Linear ESM using those single layer measurements improved accuracy somewhat in Mollisols but not Alfisols. Measuring a 30–35 cm "correction layer" substantially reduced systematic error in both soil types using linear and especially spline ESM (< 1.5% or 0.5 Mg ha^{-1}). Of practical importance, pooling the correction layer did not sacrifice accuracy while substantially reducing ESM cost. For deeper soils (60 cm), FD was less inaccurate but ESM with multiple layers further improved accuracy. We conclude with recommendations on when and how to use ESM for accurate and cost-effective SOC stock change estimation in the two major soil types of central USA. However, caution is warranted when estimating SOC changes with short timescales, more extreme BD changes (e.g. ± 5 cm), and other soil types.

1. Introduction

Accurate measurements of changes in soil organic carbon (SOC) stocks are essential for experimental research, soil survey, carbon accounting and crediting, and as inputs into biogeochemical modeling. Changes in SOC stocks are traditionally estimated using the fixed depth (FD) paradigm: soils are initially sampled at time t_0 to a selected depth (e.g. 0–30 cm) and SOC stocks are measured (by measuring SOC concentration and bulk density, BD). At a later time t_1, SOC stocks are remeasured to the same depth. Their difference is reported as the SOC stock change.

However, it has long been recognized that the FD paradigm is fallible due to changes in BD. For example, suppose that between t_0 and t_1, SOC stock does not change but that the soil is compacted. Then an FD sample at t_1, compared to t_0, will contain more soil and therefore more SOC stock, leading to the incorrect conclusion that SOC stock has increased. To address this, the equivalent soil mass (ESM) paradigm has been developed (Ellert and Bettany, 1995; Wendt and Hauser, 2013). Instead of comparing SOC stocks in a fixed *depth* of soil (e.g. 0–30 cm; Fig. 1a), the ESM paradigm compares SOC stocks in a fixed *mass* of dried mineral soil (e.g. 0–3 Gg ha^{-1}; Fig. 1b). In the above example, under ESM the same amount of soil will be sampled at both times, leading to the correct

* Corresponding authors at: Agroecosystem Sustainability Center, Institute for Sustainability, Energy, and Environment, University of Illinois Urbana-Champaign, Urbana, IL, USA.

E-mail addresses: epotash@illinois.edu (E. Potash), margenot@illinois.edu (A.J. Margenot).

https://doi.org/10.1016/j.geoderma.2026.117801
Received 15 December 2025; Received in revised form 11 March 2026; Accepted 30 March 2026
Available online 4 April 2026

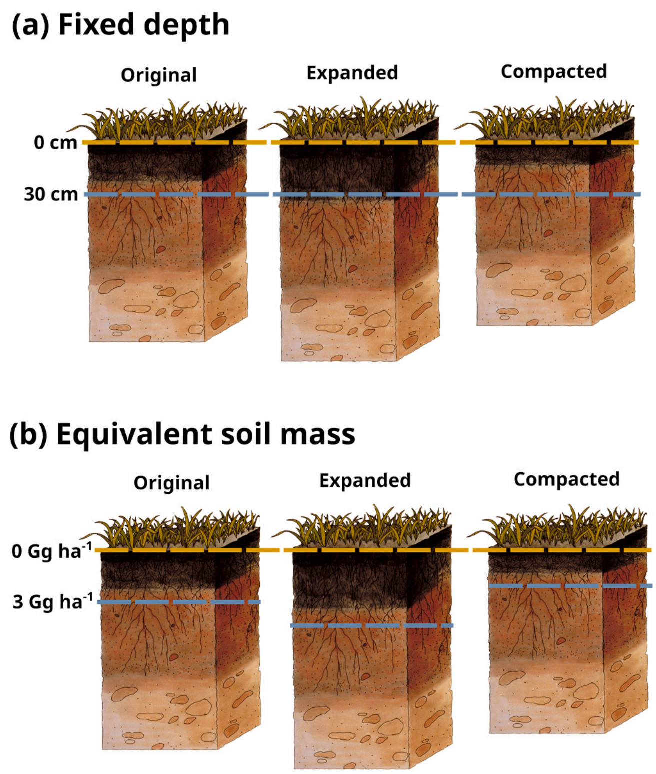

Fig. 1. Two measurement paradigms for estimating changes in soil organic carbon (SOC) stocks of (a) fixed depth (FD) and (b) equivalent soil mass (ESM) differ in how the bottom of the sample layer (blue dashed line) is defined. In FD, it is defined by a depth (cm). In ESM, it is defined by a dry mineral soil mass (Gg ha^{-1}). Modified from United States Department of Agriculture (2016).

conclusion of no change in SOC stock. Past studies have found that accounting for changes in BD using ESM leads to substantially different estimates of SOC stock change compared to FD (e.g. Peng et al., 2024; Don et al., 2011).

When re-measuring a fixed depth of soil, BD may change for a variety of reasons, in particular at surface depths (< 30 cm) and under agricultural land use. Some of these reasons, such as management changes (e.g. tillage, crop rotation) or climatic conditions (e.g. drought), may reflect real changes. Other changes are more superficial consequences of the measurement process itself. For example, when resampling a location, the sampling point moves but short-range variability implies the BD measurement will change (Poeplau et al., 2022; Potash et al, 2025). And while it is best practice for resampling to occur at the same time of year as the initial sampling, this is not always practical. For example, initial sampling may occur in the fall before tillage, but we may be forced due to weather conditions to resample after tillage. In this example, BD changes not because of a change in tillage but because of a change in the timing of sampling. However, we emphasize that both real and superficial BD changes can threaten the validity of FD measurements.

The ESM paradigm accounts for apparent BD changes by shifting the soil sampling frame from a given *depth* of soil to a *mass* of soil. ESM SOC stocks could be measured directly by drying intact cores to remove all moisture and then removing soil from the bottom depth until the target mass is reached (Wuest, 2025). However, this approach works on a whole mass basis, whereas ESM has typically been defined on a fine earth fraction (< 2 mm) and mineral soil basis, i.e. excluding coarse fragments and organic matter. Moreover, fully drying intact soil cores can take days and removing soil mass from the bottom depth of an intact core is not feasible for soils with coarse texture or insufficient structural integrity.

Since direct measurements of ESM SOC stocks are challenging and labor intensive, they are not generally performed and in practice approximated by ESM adjustments. An ESM adjustment takes FD measurements and adjusts them through interpolation and extrapolation to approximate ESM SOC stocks. A variety of such ESM adjustment methods have been proposed (Ellert and Bettany, 1995; Lee et al., 2009; Rovira et al., 2015; von Haden et al., 2020) varying primarily by the number of FD measurements employed and their approximation method such as linear or spline.

While ESM adjustments avoid the challenges of direct ESM measurement described above, evaluating adjustments remains a challenge precisely because of the challenges of direct ESM measurement. Past studies have evaluated ESM adjustments using two approaches. The most common approach is to simulate BD and SOC concentration depth distributions (Lee et al, 2009; von Haden et al, 2020; Fowler et al., 2023; Rovira et al., 2015). These evaluations are useful, but rely on synthetic rather than empirically measured soils data and typically ignore spatial variation and measurement error. The second approach is empirical but, due to a lack of ground truth ESM measurements, has not been able to assess systematic errors (i.e., accuracy) of ESM adjustments. Thus the empirical approach either evaluates agreement between adjustment methods (Lee et al, 2009; Peng et al., 2024) or evaluates random error (Boivin et al., 2025).

As a result, evaluations of ESM to date have either used simulated data or used empirical data but were unable to evaluate systematic errors. Our scientific objective was therefore to quantify the accuracy of ESM adjustments as well as FD measurements through empirical evaluation. We looked at the effect on accuracy of (1) different ESM adjustment methods (none, linear, and spline); (2) depthwise measurement designs (e.g. 0–30, 30–35 cm; Table 1); (3) soil types (Alfisols and Mollisols); and (4) initial soil depth (30 and 60 cm). Based on the results of these analyses we generated recommendations for future studies. This novel empirical evaluation of ESM systematic errors was enabled by finely sectioned measurements (2.5 cm depth intervals) of 72 soil cores (0–75 cm depth) across six sites (central USA).

2. Methods

2.1. Equivalent soil mass adjustments

It is typically not feasible to directly measure ESM SOC stocks (see introduction). Instead, we make FD measurements (albeit with error) of SOC concentration and BD. Then ESM adjustment is a prediction problem: given FD measurements and a reference mass, the adjustment seeks to predict the SOC stock at the reference mass.

We assume paired resampling, i.e. the same locations are sampled at t_0 and t_1. It is convenient to make an FD (e.g. 0–30 cm) measurement at t_0 and define the reference mass to be the mineral soil mass of this sample. Then the ESM SOC stock *change* simplifies because at t_0 the ESM SOC stock by definition simply equals the direct FD measurement. Thus only the t_1 ESM SOC stock requires ESM adjustment (equation S1). We leverage this insight to evaluate the errors of ESM adjustment for SOC stock change using measurements at a single point in time (section 2.3).

2.1.1. ESM adjustment methods

A variety of ESM adjustments have been proposed in the literature (Peng et al., 2024). In this study we consider three main techniques (Fig. 2; see supplementary text for formulas):

(1) Fixed depth (FD): measure a single layer of soil and do not adjust.
(2) Linear adjustment: assume a constant SOC concentration within each soil layer.
(3) Spline adjustment: the cumulative SOC stock curve of multiple layers is interpolated using a monotonic smoothing spline (e.g., Hyman spline) (von Haden et al., 2020).

Linear adjustment can be performed using a single soil layer, but can

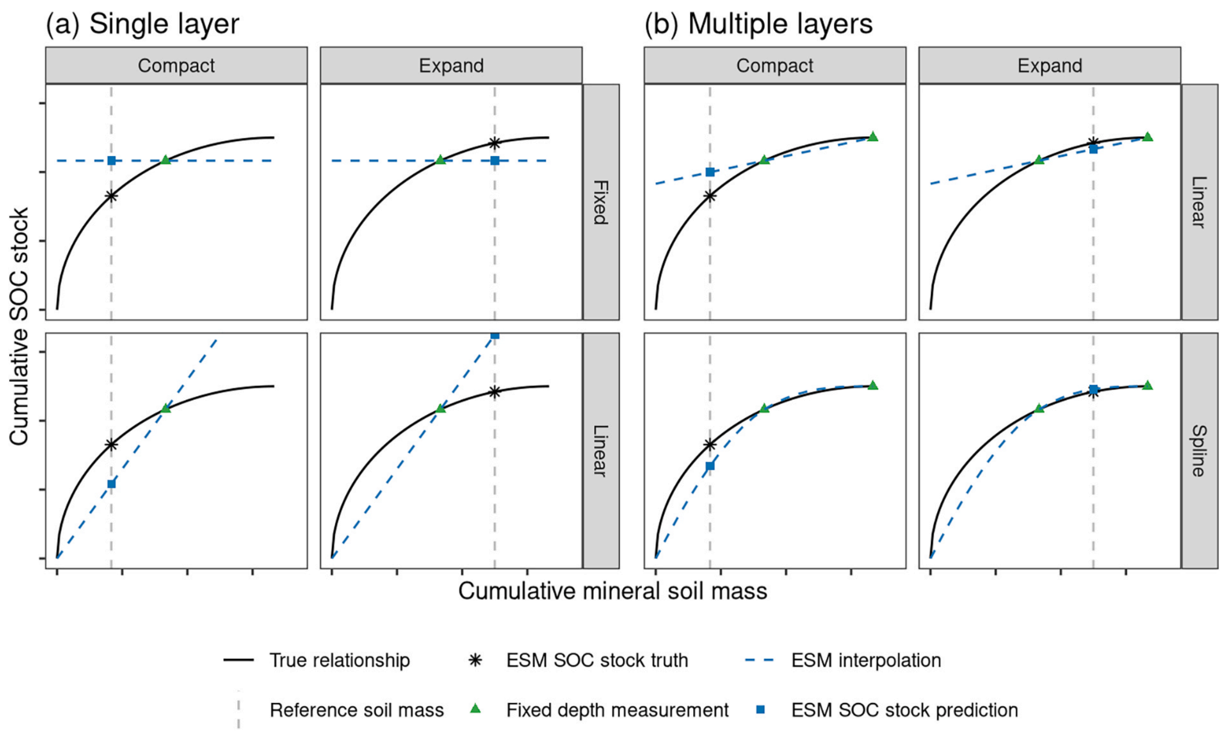

Fig. 2. Illustration of ESM adjustment: FD measurements (green triangles) are inputs to ESM adjustment prediction (blue square) of true ESM SOC stock (black star). Panels are the number of FD measurement layers used by ESM adjustment: (a) single layer, e.g. 0–30 cm; or (b) multiple layers, e.g. the correction layer design of 0–30, 30–35 cm (section 2.1.2). Columns are BD change scenarios (section 2.3.1). Rows are ESM adjustment methods (section 2.1.1). ESM adjustment error is the vertical distance between ESM adjustment (blue square) and true ESM SOC stock (black star) and is typically larger with a single layer (a) compared to multiple layers (b).

benefit by measuring multiple layers to better identify SOC near the reference mass (Fowler et al., 2023). Following Boivin et al. (2025), when multiple soil layers were sampled (e.g. 0–30, 30–35 cm) we used the deepest layer (30–35 cm) for linear adjustment. Spline adjustment requires at least two soil layers. We implemented the ESM adjustment methods in the R programming language and our code is provided.

2.1.2. Depthwise measurement designs

An important consideration for ESM adjustment is the depthwise measurement design, that is how the intact soil core is sectioned into fixed depth layers to measure BD and SOC concentration (Table 1). The design may involve one (fixed or linear adjustment) or multiple (linear or spline adjustment) soil layers (Fig. 2). For example, consider an initial sample of 0–30 cm. To estimate SOC stock change, we may simply follow up with another 0–30 cm sample (i.e. single layer design, Table 1). However, we may supplement the follow-up sampling with a measurement at 30–35 cm (correction layer design) (Boivin et al., 2025). The purpose of a correction layer is to provide information to the adjustment

Table 1
Depthwise measurement designs being evaluated. Each design consists of one or more FD soil layers. Each layer is either measured independently at each location (default; solid gray in figure), pooled across locations (composited and measured once; hatched), or averaged across locations (measured independently then averaged, as in meta-analysis; double hatched). The applicable initial sampling depth and adjustment methods for each design are indicated.

Design name	**Specification**		**Initial depth (cm)**		**Adjustment methods**		
	Figure	**Layers (cm)**	**30**	**60**	**Fixed**	**Linear**	**Spline**
Single layer		0–30	✓		✓	✓	
Single layer		0–60		✓	✓	✓	
Correction layer		0–30, 30–35	✓			✓	✓
Correction layer pooled		0–30, 30–35 pooled	✓			✓	✓
Multilayer		0–30 30–60	✓	✓		✓	✓
Multilayer summary		0–30 averaged 30–60 averaged	✓	✓		✓	✓

method (linear or spline) about SOC near the reference mass. Thus, a correction layer is expected to be more important in a soil where SOC is changing more abruptly, such as in mollic epipedon transition of Mollisols (Fig. 3).

In addition to single layer and correction layer designs, we also considered three practical alternatives. The first is to use the same 30–35 cm correction layer but to pool all of the correction layer samples across cores at a site into a single measurement. Proposed by Boivin et al. (2025), this pooled correction layer offers the practical advantage of substantially reducing laboratory costs.

The second alternative we considered is to section the cores from 0 to 30 and 30–60 cm (multilayer design). In this case, the 30–60 cm layer serves as a correction layer. While sampling to 60 cm can require more resources than sampling to 30 or 35 cm, the 0–30 and 30–60 cm design is relevant to soil surveys and research studies which already sample multiple depths but not a relatively thin correction layer.

The final alternative we considered was to operate on average measurements of the 0–30 and 30–60 cm layers (multilayer summary). This design is relevant for meta-analysis and other secondary analyses that apply ESM adjustment to FD measurements reported in the literature without having access to data at individual locations (e.g., Don et al., 2011).

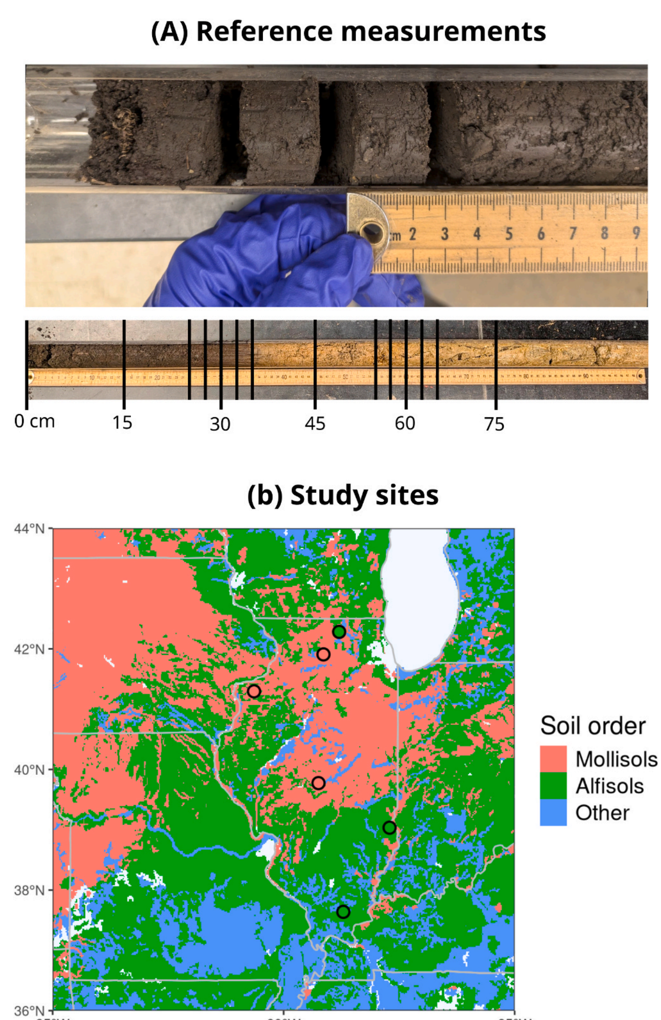

Fig. 3. Sampling designs. (a) depthwise measurement design of soil cores into 13 sections for analysis and (b) six sites in Illinois, USA including 3 Alfisols (red) and 3 Mollisols (green).

2.2. Study data

To support ESM adjustment evaluation, we sampled six agricultural sites in Illinois, USA, selected to represent the two dominant soil orders of the region: Alfisols and Mollisols (Table 2). Following crop harvest in fall 2024, we collected 12 intact soil cores (4.4 cm diameter) per site using a hydraulic probe to a target depth of 75 cm. The cores were located using a stratified random sampling design to capture spatial variability within the site.

The intact cores were first sliced in 15 cm sections down to 75 cm depth. Then, centered at each of 30 cm and 60 cm, four 2.5 cm sections were sliced. For example at 30 cm, this resulted in 25–27.5, 27.5–30, 30–32.5, and 32.5–35 cm sections (Fig. 4). In total, we analyzed 13 sections per core, slicing at the following depths: 0, 15, 25, 27.5, 30, 32.5, 35, 45, 55, 57.5, 60, 62.5, 65, and 75 cm. With 12 cores per site and 6 sites we analyzed a total of 932 layers (4 layers missing due to cores $<$ 75 cm final depth). Sections were weighed, air dried at 24°C, and gently ground by mortar and pestle to pass a 2-mm sieve and remove rocks and roots. A 10–20 g subsample was oven dried at 105°C for 16 h to constant mass to measure oven-dried BD given the known soil volume. Samples with pH $>$ 7.2 were acidified using HCl vapor to remove carbonates (Potash et al., 2023) and SOC concentration was measured by dry combustion.

2.3. Evaluation

Past empirical studies have been unable to assess ESM accuracy because they have not measured ground truth ESM SOC stocks. Measuring such stocks is labor intensive and potentially infeasible. Moreover, capturing a variety of depthwise measurement designs (section 2.1.2), soil types (with different characteristic SOC profiles) and BD change scenarios (varying degrees of soil compaction and expansion) increases this challenge further.

In this study we overcome these challenges using the following cross validation logic. If a 0–30 cm FD measurement were made at t_0 and the soil compacts for example by 2.5 cm, then at t_1 the desired ESM SOC stock is equivalent to an FD SOC stock of 0–27.5 cm. So we can evaluate a given ESM adjustment method (e.g. spline) and measurement design (e.g. 0–30, 30–35 cm) by using it to predict 0–27.5 cm SOC stock and comparing it to the measured value. In order to implement this, we would need to measure the ground truth (e.g., 0–27.5 cm SOC stock) as well as the inputs (e.g. 0–30, 30–35 cm SOC and BD). Thus by finely slicing soil cores (section 2.2), we can leverage this idea to simultaneously evaluate many possible measurement designs and compaction and expansion scenarios. Notably, this evaluation does not involve t_0 measurements, so that we can implement it using measurements at a single point in time which we interpret as t_1.

2.3.1. Evaluation implementation

Here are the details of our implementation of the above cross validation logic. At each site we evaluated ESM methods by (1) depthwise measurement design, (2) adjustment method, and (3) BD change scenario (Fig. 4). For example, we evaluated the correction layer design with linear adjustment and under a compaction scenario. To perform this evaluation, we first sampled (statistically) a target depth for each of the 12 cores at the site from the BD change scenario distribution (Fig. 5). This was used to determine the ground truth SOC stock for each core, as well as the reference mass for adjustment. Next we simulated the measurement process that would have occurred under the specified measurement design by aggregating the finely sectioned measurements. For example, we aggregated to 0–30 cm, and 30–35 cm depth. To account for measurement error, we add random errors to these measurements with standard deviation of 2% for BD and 4% for SOC (Potash et al., 2025). Finally we apply the specified adjustment method (e.g. spline) to these measurements to make an ESM prediction.

Given these ground truth and predicted ESM SOC stock values at

Table 2
Characteristics of soil sampling sites. Soil taxonomy and properties are estimated by USDA NRCS Soil Survey Geographic Database (SSURGO). Abbreviations: mean annual precipitation (MAP), mean annual temperature (MAT).

Site	Lon, Lat	Climate		USDA soil taxonomy		Soil properties	
		MAP (cm)	MAT (°C)	Order	Subgroup	Texture	A horizon thickness (cm)
Boon	−88.8, 42.3	101	9	Alfisols	Typic Hapludalfs	Silt loam	18
Chri	−89.2, 39.8	103	12	Mollisols	Typic Endoaquolls	Silty clay loam	43
Craw	−87.7, 39.0	125	13	Alfisols	Aquic Hapludalfs	Silt loam	20
Merc	−90.6, 41.3	104	10	Mollisols	Aquic Argiudolls	Silt loam	33
Ogle	−89.1, 41.9	97	9	Mollisols	Aquic Argiudolls	Silt loam	41
Sali	−88.7, 37.6	132	14	Alfisols	Typic Fragiudalfs	Silt loam	5

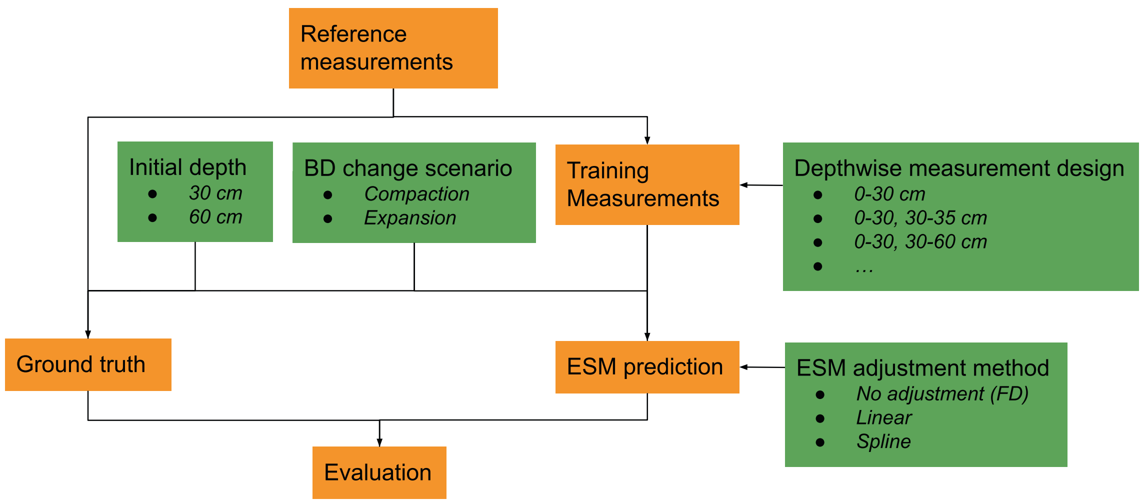

Fig. 4. ESM empirical evaluation workflow. Evaluation parameters in green, examples listed in italics.

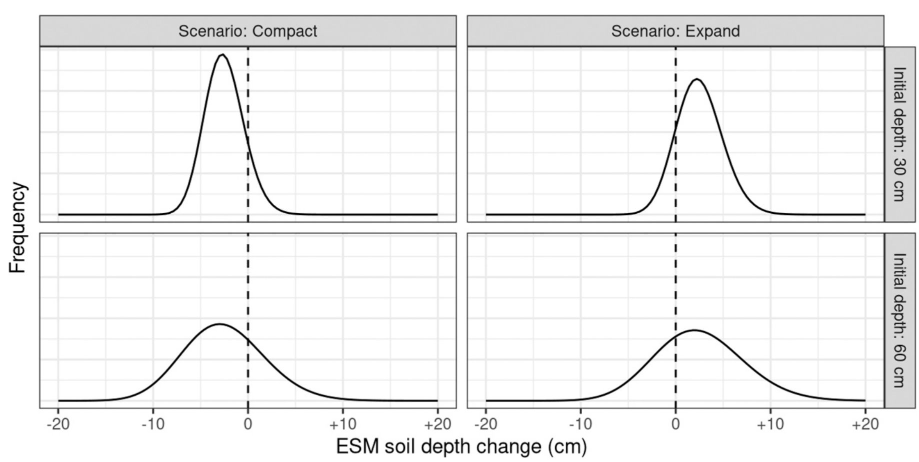

Fig. 5. Bulk density change distributions used to evaluate ESM adjustment errors. The distributions are lognormal with mean + 2.5 cm (expansion) and −2.5 cm (compaction) and standard deviation 7.5%.

each of the 12 locations at the site, we averaged across the 12 locations to obtain ground truth and predicted SOC stock averages. Their difference is the ESM prediction error. For each specification, we repeated this process 200 times to capture the BD change distribution as well as measurement variability. Finally, given these 200 errors, we summarize with the systematic error (accuracy) and random error (precision). The systematic error is the mean and the random error is the standard deviation. Random errors are less concerning than systematic errors, which do not diminish with increased sample size.

To generate the BD changes for our evaluation (section 2.2), we used lognormal distributions. The study sites were not re-measured so we relied on other data to inform these distributions (von Haden et al., 2020; Bauer and Black, 1981; Don et al., 2011). We selected our moderate expansion and compaction scenarios to have mean changes of plus and minus 2.5 cm with standard deviations of about 7.5%, i.e. 0.075 on the log scale (Fig. 5). We also considered more extreme mean changes of plus and minus 5.0 cm (Fig. S1). To use these continuous BD change distributions in our evaluation we discretized them to align with the sampling depths at our six evaluation sites (supplementary information).

3. Results

Measured SOC stocks at the six sites followed the patterns expected by their mapped soil types (Fig. 3). Specifically, the SOC stocks of Alfisols tended to be lower and have a more abrupt decline with depth than Mollisols over 0–75 cm depth (Fig. 6). This trend was driven by SOC concentration rather than BD (Fig. S2).

3.1. Initial sampling depth 30 cm

With a single 0–30 cm layer of measurements, there are two approaches: FD (no adjustment) and linear adjustment. As expected, in both the compaction and expansion scenarios, systematic errors under FD were large (2–8%, Fig. 7a; or 1–5 Mg ha^{-1}, Fig. S3). Linear adjustment reduced these errors for Mollisols (2–4%; or 1–3 Mg ha^{-1}). However, errors remained similar for Alfisols and in one case (Sali site, expansion scenario) were substantially increased.

Next we considered adding a correction layer measurement of 30–35 cm with two ESM adjustment methods: linear and spline (Fig. 7a). With linear adjustment, systematic errors were reduced below about 2% in both soil types (1 Mg ha^{-1}). Spline adjustment reduced systematic errors further, to less than about 1.5% (0.5 Mg ha^{-1}). Random errors were also reduced from 2–3% without a correction layer to 1–2% with a correction layer (Fig. S4).

Thus we have established acceptable ESM adjustment performance using both linear and spline methods with the correction layer depthwise measurement design. Next we considered three alternative depthwise measurement designs of practical interest. First, when pooling the correction layer across a site, we found almost no degradation of ESM adjustment performance (Fig. 7a), despite reducing the number of soil measurements by almost half. Moreover, random errors were also not substantially affected (Fig. S4). Second, with a 0–30, 30–60 cm multilayer design (e.g. soil surveys or research plots), spline adjustment continued to have low systematic error of < 3% or 1 Mg ha^{1}, but linear adjustment exhibited up to 5% or 3 Mg ha^{-1} systematic error (though it outperformed splines on Alfisols). Lastly, when performing ESM adjustment on summary data (site-level), performance was degraded only slightly compared to ESM adjustment on primary data (location-level).

3.2. Initial sampling depth 60 cm

We also evaluated ESM adjustment errors following an initial sampling depth of 60 cm. We considered depthwise measurement designs of both a single layer and multiple layers. Compared to 30 cm results above, FD changes for 60 cm had smaller systematic errors in both relative terms (< 3%, Fig. 7b) and absolute terms (< 2 Mg ha^{-1}, Fig. S3). This occurred because SOC density near 60 cm in these soils was smaller than SOC density near 30 cm (Fig. 6) both absolutely (g cm^{-3}) and relatively (compared to cumulative SOC stock to that depth). Another discrepancy was that linear adjustment without a correction layer had consistently higher systematic error than FD. This was due to average SOC in the 0–60 cm layer being substantially different from SOC at the adjustment depth of 60 cm (Fig. 6). Adjustment with multiple layers reduced errors to below around 1.5% (1.5 Mg ha^{-1}). As with 30 cm, 60 cm adjustment was not substantially affected by using only summary data.

4. Discussion

4.1. Summary of findings

We empirically evaluated the systematic errors of ESM adjustments for SOC stock change under moderate BD changes (± 2.5 cm) with both 0–30 cm and 0–60 cm initial measurements. For 0–30 cm initial measurements, we found that re-measuring the 0–30 cm layer with no adjustment (FD) would result in the largest systematic error of up to 8% or 5 Mg ha^{-1} (Fig. 7, S2). Linear adjustment of the single layer reduced systematic errors in Mollisols to around 3% or 2 Mg ha^{-1}, but offered no benefit for Alfisols. This discrepancy is a consequence of the fact that SOC density declines more abruptly in the soil profile of Alfisols than Mollisols (Fig. 6). Since the magnitude of annual SOC stock changes is commonly much less than 1% or 1 Mg ha^{-1} y^{-1} (e.g. Sanford et al., 2012), our results suggest that the widespread use of FD for SOC stocks, or even single-layer linear ESM adjustment, can substantially distort estimates of SOC stock change.

However, measuring a correction layer (e.g. 30–35 cm) reduced systematic errors substantially with linear and especially spline adjustment (≤ 1.5% or 0.5 Mg ha^{-1}). This would enable relatively accurate

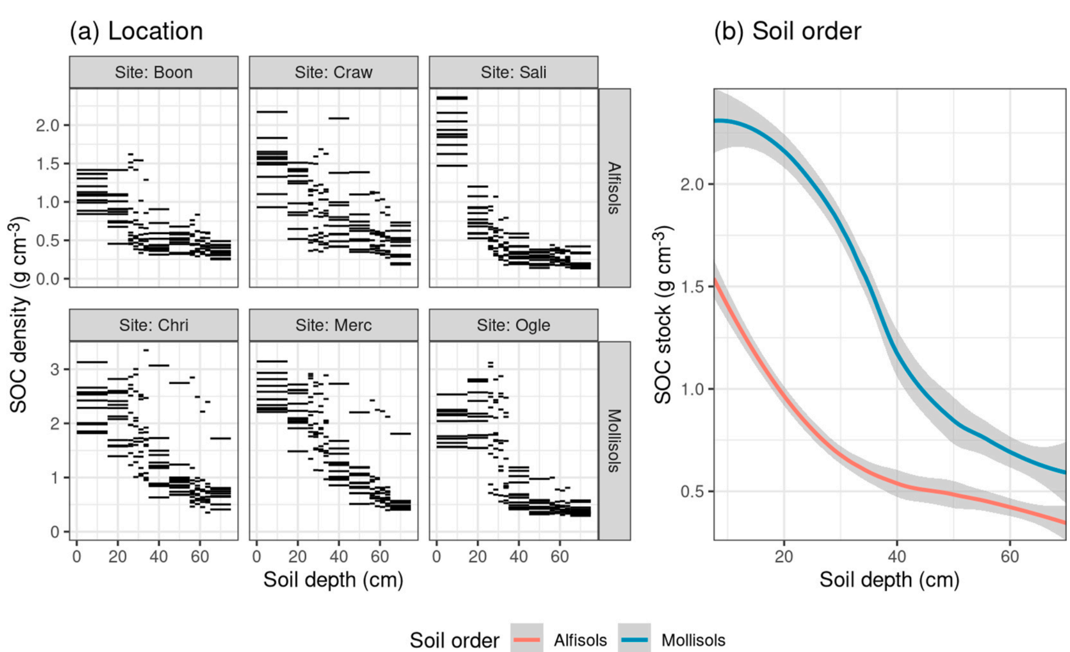

Fig. 6. Depthwise SOC density distributions (a) for each of the six sites (n = 12 cores per site to a target depth of 75 cm) and (b) summarized by soil order using LOESS. Three sites were Mollisols and three were Alfisols. See Fig. S2 for SOC density distributions by mineral soil mass instead of depth.

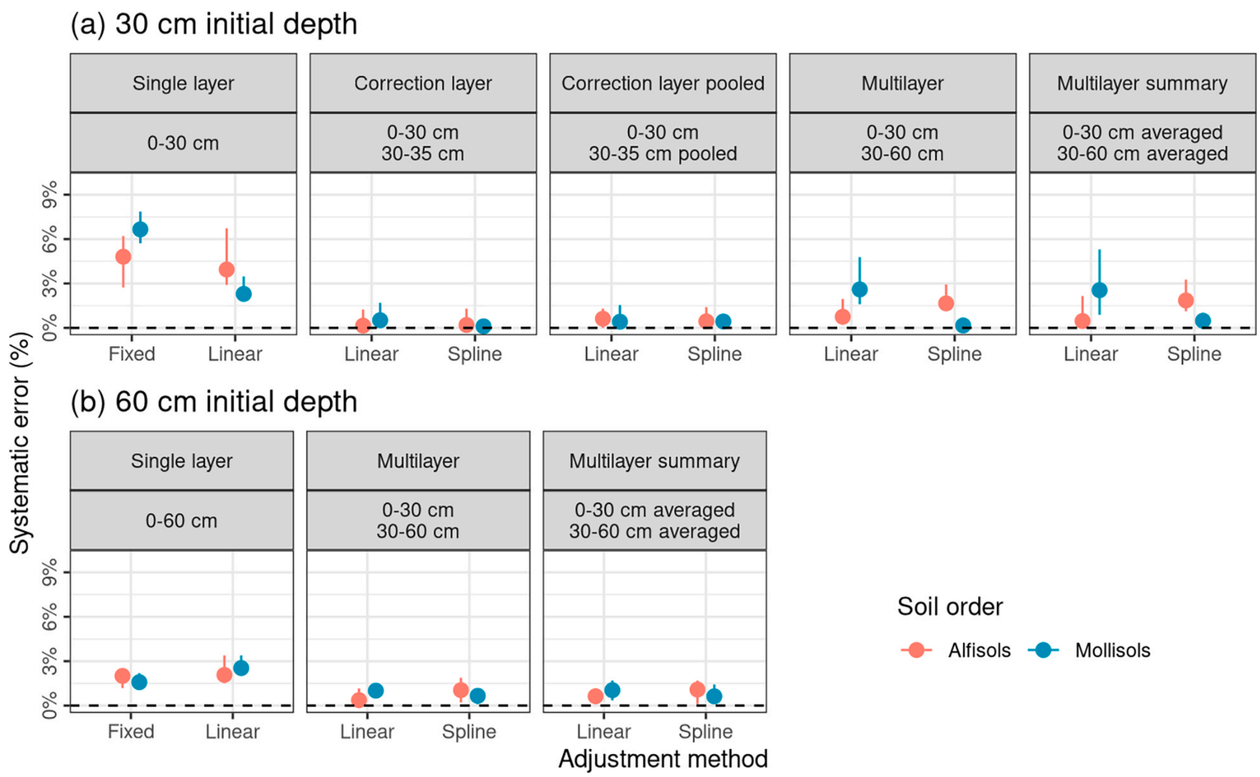

Fig. 7. Systematic errors under FD and ESM adjustment in estimating SOC stock change around (a) 30 cm and (b) 60 cm. Each panel is a depthwise measurement design. Points and bars show median and range, respectively, of magnitude of systematic errors across 6 sites and two scenarios (compaction and expansion).

longer-term (e.g. $\geq$ 5 year) monitoring relative to typical rates of SOC stock change (e.g. 5% or 5 Mg ha^{-1}). For 0–60 cm initial measurements, FD remeasurement resulted in smaller systematic errors but adjustment with multiple layers (e.g. 0–30, 30–60 cm) also reduced systematic errors at this greater depth. Under more substantial BD changes ($\pm$ 5 cm, see methods) FD systematic errors became extremely large (up to 20% or 11 Mg ha^{-1}) while spline ESM with multiple layers kept errors under 4% or 2 Mg ha^{-1} (Figs. S5-S6).

4.2. Recommendations

Based on our findings, we can make recommendations for agricultural Alfisols and Mollisols in central USA (Fig. 3) and more broadly soils with similar depthwise SOC distribution (section 4.3). We note that in the WRB taxonomy, Alfisols are generally classified as Luvisols and Mollisols are classified as Chernozems (IUSS Working Group WRB, 2022). These are major soils of the USA, where they occupy 22% and 14% of land area, respectively, as well as globally, where they occupy 7% and 10% of land area, respectively (University of Idaho, 2026).

For $\leq$ 30 cm stock measurements, we strongly recommend against FD and single layer linear ESM in favor of using multiple layers to perform linear or spline ESM. For deeper soil ($\geq$60 cm), FD may result in systematic errors that are acceptable for estimating SOC stock change on long timescales if BD changes are moderate. However, ESM with multiple layers reduces systematic errors which may be necessary for shorter timescales or more extreme changes in BD.

Our results show that pooling the correction layer across cores into a single sample did not substantially increase errors (Fig. 7a, panel 3), suggesting that relatively accurate ESM adjustment can be performed without substantially increasing lab analysis costs. Moreover, performing ESM adjustment with summary data only (e.g., meta-analysis) also did not substantially increase errors, suggesting that studies that do not perform ESM adjustment themselves can be accurately corrected post-hoc provided that BD and SOC concentration are reported.

Our results address SOC stock *changes,* but hold direct implications for SOC stock *treatment effects* (Potash et al., 2025). The treatment effect (also known as dynamic baseline change) is the difference in change under a treatment management (e.g. cover cropping) from the change under a baseline or control management (e.g. no cover cropping). The systematic error in estimating the treatment effect is equal to the difference in systematic errors for estimating change under each management. If the systematic errors for the managements are similar in both direction and magnitude, they cancel each other out and the treatment effect will be estimated more accurately than the change under each management. This optimistic scenario is plausible when BD changes are similar under both treatment and control, i.e. treatment does not affect BD and sampling conditions are consistent across sites. However, this is often unrealistic as many treatments have been found to affect BD, e.g. expansion under cover cropping (Yan & Arthur, 2025). In such cases, the systematic errors do not cancel and treatment effect estimation will be subject to systematic error, elevating the need for ESM adjustment following the above recommendations.

4.3. Limitations and future work

For soil types other than the Alfisols and Mollisols studied here, we recommend considering the depthwise distribution of SOC in the soil profile to assess the extent to which our results translate. This may be done on the basis of prior information such as soil surveys at similar sites, or initial measurements at the target site. If the soils differ in their SOC distribution substantially from ours (Fig. 6), we recommend performing a similar empirical evaluation of systematic errors for ESM adjustments for the target soil.

We observed subtle differences in systematic errors between Alfisols and Mollisols such as under linear adjustment with a single 0–30 cm soil layer (see above). Overall, however, the patterns between these soil orders were very similar. This is encouraging for the generalizability of our results to other soils, as the replicated soils of these two orders bound a diversity of vertical SOC distributions (Fig. 6). Potential outliers may include organic soils (e.g., Histosols, Gelisols) or soils with O horizons, such as Inceptisols under forest land use. Soils with a sharp discontinuity in SOC distribution at depth (e.g. plow layer, E horizon, fragipan, buried horizons) warrant caution in the application of ESM near this discontinuity.

A limitation of all of the ESM adjustments in the literature and evaluated here is that they do not quantify prediction uncertainty. Thus we must rely on studies such as the present one to provide estimates of the possible systematic and random errors. A need for future research is the development of ESM adjustments that quantify their prediction uncertainty. For example, kriging is an interpolation method that quantifies uncertainty and is widely used in soils research.

5. Conclusion

We present the first empirical evaluation of the systematic errors of FD and ESM adjusted SOC stock measurements, made possible by finely sectioned measurements (2.5 cm depth intervals) of soil cores to 75 cm depth (dataset provided). Our results show that under moderate BD changes in a variety of central USA soils (Alfisols and Mollisols), 0–30 cm FD measurements substantially distort estimates of SOC stock change with large systematic errors. We find that ESM adjustment with this single 0–30 cm layer is insufficient for reducing systematic errors to an acceptable level for SOC stock change estimation. However, ESM adjustment with multiple layers consistently and substantially reduces systematic errors, enabling accurate estimation of SOC stock changes. Moreover, we find that ESM can be made cost effective by using a pooled "correction layer" which substantially reduces costs while not degrading accuracy. For deeper soil measurements (0–60 cm), the systematic errors of FD measurements are smaller and may be less problematic, but accuracy can be further improved by ESM with multiple layers.

CRediT authorship contribution statement

Eric Potash: Writing – review & editing, Writing – original draft, Software, Methodology, Conceptualization. **Adam C. von Haden:** . **Luke Bergschneider:** Investigation. **Michael S. Douglass:** Investigation. **Lenarth Ferrari:** Investigation. **Rosa Ibarra Lopez:** Investigation. **Adriana Reconco Martinez:** Investigation. **Andrew J. Margenot:** Writing – review & editing, Conceptualization.

Funding

The authors acknowledge financial support from USDA NRCS award NR243A750023C034 (EP, AJM), NASA Acres Consortium (EP), NCR SARE LNC22-475 (ACV), NSF Smart & Connected Communities 2,125,626 (AJM), and Illinois NREC 2021–4-360731–469 (AJM).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Eric Potash reports financial support was provided by NASA Acres Consortium. Eric Potash reports financial support was provided by Natural Resources Conservation Service. Andrew Margenot reports financial support was provided by Natural Resources Conservation Service. Adam von Haden reports financial support was provided by Sustainable Agriculture Research & Education. Andrew Margenot reports financial support was provided by National Science Foundation. Andrew Margenot reports financial support was provided by Illinois Nutrient Research & Education Council. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.geoderma.2026.117801.

Data availability

The data and code that support the findings of this study are available in figshare at https://doi.org/10.6084/m9.figshare.31901941.

References

Bauer, A., Black, A.L., 1981. Soil carbon, nitrogen, and bulk density comparisons in two cropland tillage systems after 25 years and in virgin grassland. Soil Science Society of America Journal 45(6) 1166–1170. https://doi.org/10.2136/sssaj1981.03615995004500060032x.

Boivin, P., et al., 2025. The single-layer equivalent soil mass method for the evaluation of soil organic carbon stocks: sources of errors, simplification, and associated detectable change. Geoderma 456, 117279. https://doi.org/10.1016/j.geoderma.2025.117279.

Don, A., Schumacher, J., Freibauer, A., 2011. Impact of tropical land-use change on soil organic carbon stocks–a meta-analysis. Glob. Chang. Biol. 17 (4), 1658–1670. https://doi.org/10.1111/j.1365-2486.2010.02336.x.

Ellert, B.H., Bettany, J.R., 1995. Calculation of organic matter and nutrients stored in soils under contrasting management regimes. Canadian Journal of Soil Science 75(4) 529–538. https://doi.org/10.4141/cjss95-075.

Fowler, A.F., et al., 2023. A simple soil mass correction for a more accurate determination of soil carbon stock changes. Sci. Rep. 13 (1), 2242. https://doi.org/10.1038/s41598-023-29289-2.

von Haden, Adam C., Wendy H. Yang, and Evan H. DeLucia. "Soils' dirty little secret: Depth-based comparisons can be inadequate for quantifying changes in soil organic carbon and other mineral soil properties." *Global Change Biology* 26.7 (2020): 3759-3770. doi:10.1111/gcb.15124.

IUSS Working Group WRB. 2022. World Reference Base for Soil Resources. International soil classification system for naming soils and creating legends for soil maps. 4th edition. International Union of Soil Sciences (IUSS), Vienna, Austria.

Lee, J., et al., 2009. Determining soil carbon stock changes: simple bulk density corrections fail. Agriculture, Ecosystems & Environment 134(3-4) 251–256. https://doi.org/10.1016/j.agee.2009.07.006.

Peng, Y., et al., 2024. Comparison of equivalent soil mass approaches to estimate soil organic carbon stocks under long-term tillage. In: *Soil and Tillage Research* 238. https://doi.org/10.1016/j.still.2024.106021.

Poeplau, C., Prietz, R., Don, A., 2022. Plot-scale variability of organic carbon in temperate agricultural soils—Implications for soil monitoring. J. Plant Nutrition Soil Sci. 185, 403–416. https://doi.org/10.1002/jpln.202100393.

Potash, E., et al., 2023. Multi-site evaluation of stratified and balanced sampling of soil organic carbon stocks in agricultural fields. Geoderma 438, 116587. https://doi.org/10.1016/j.geoderma.2023.116587.

Potash, E., et al., 2025. Measure-and-remeasure as an economically feasible approach to crediting soil organic carbon at scale. Environ. Res. Lett. 20 (2), 024025. https://doi.org/10.1088/1748-9326/ada16c.

Rovira, P., et al., 2015. Towards sound comparisons of soil carbon stocks: a proposal ba sed on the cumulative coordinates approach. Catena 133 420–431. https://doi.org/10.1016/j.catena.2015.05.020.

Sanford, G.R., et al., 2012. Soil carbon lost from Mollisols of the North Central USA with 20 years of agricultural best management practices. Agriculture, Ecosystems & Environment 162 68–76. https://doi.org/10.1016/j.agee.2012.08.011.

United States Department of Agriculture, 2016, http://www.nrcs.usda.gov/Internet/FSE_MEDIA/nrcseprd423217.jpg. Archived at Internet Archive, accessed 11 Nov 2025.

University of Idaho, https://www.uidaho.edu/agricultural-life-sciences/soil-orders. Accessed 9 Mar 2026.

Wendt, J.W., Hauser, S., 2013. An equivalent soil mass procedure for monitoring soil organic carbon in multiple soil layers. Eur. J. Soil Sci. 64 (1), 58–65. https://doi.org/10.1111/ejss.12002.

Wuest, S., 2025. Measuring soil carbon stocks with greater simplicity, accuracy, and repeatability. Soil Sci. Soc. Am. J. 89 (1), e70012. https://doi.org/10.1002/saj2.70012.

Yan, F., Arthur, E., 2025. Cover crops alter soil physicochemical properties: a global meta-analysis. Geoderma 460, 117436. https://doi.org/10.1016/j.geoderma.2025.117436.