

A Molecular Investigation of Soil Organic Carbon Composition Across a Subalpine Catchment

Supplementary Materials

Hsiao-Tieh Hsu^{1,2}, Corey R. Lawrence³, Matthew J. Winnick⁴, John R. Bargar⁵, and Katharine Maher^{2,6*}

¹ Department of Chemistry; Stanford University, Stanford, CA 94305, USA; hthsu9@stanford.edu

² Department of Earth System Science, Stanford University, Stanford, CA 94305, USA

³ U.S. Geological Survey, Denver, CO 80215, USA; clawrence@usgs.gov

⁴ Department of Geological Sciences, Stanford University, Stanford, CA 94305, USA; mwinnick@stanford.edu

⁵ Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA; bargar@slac.stanford.edu

⁶ Rocky Mountain Biological Laboratory, Crested Butte, CO 81224, USA

* Correspondence: kmaher@stanford.edu

Keywords: soil organic carbon, FT-IR, XAS

Type of paper: Article

1. Deconvoluted FT-IR spectra.....	2
2. Deconvoluted bulk C XAS spectra.....	3
3. Sample selections for equation (8).....	4
3.1. Universal vs. site-specific vs. depth-specific fits for SOC	4
3.2. XRF.....	8
3.3. PXRD.....	11
3.4. Sample selection conclusions	12
4. Uncertainties in α_i : MW_{SOC} vs. final outcome of linear mass balance fit.....	12
5. Boundary conditions of the linear mass balance fit and slope t-tests	14
6. Elemental analysis: TIC, TOC, and TN.....	18
7. Bulk density of soils and biomass.....	19
8. Biomass FT-IR spectra	20
9. Biomass bulk C XAS spectra.....	21
10. Organic carbon composition for BCM, BCW, and RCM: α_i values	23
11. Evaluation of the SOC-fga method relative to single techniques	26
12. References.....	27

1. Deconvoluted FT-IR spectra

Figures S1-2 are examples of deconvoluted FT-IR spectra using the Igor Pro software package (version 7, WaveMetrics Inc., Portland, OR^{*}).

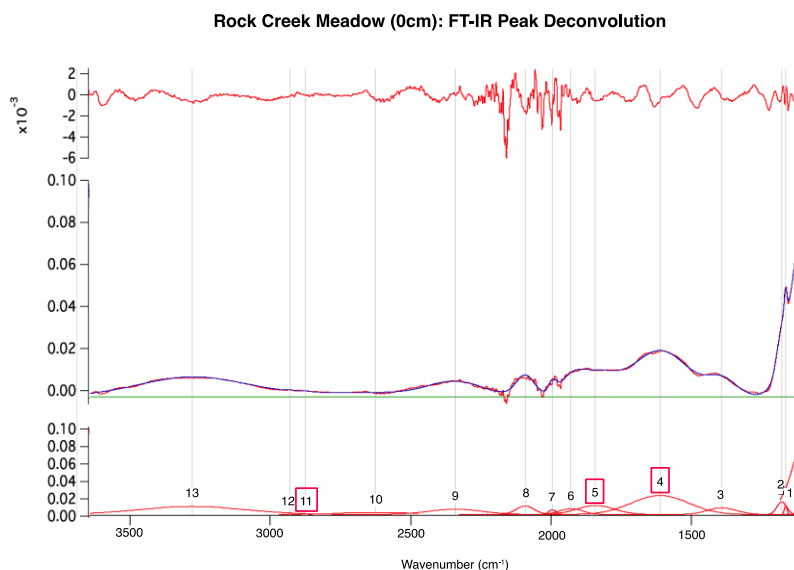


Figure S1. Example of peak deconvolution of soil FT-IR spectrum using the Igor Pro software package. Red boxes represent peaks chosen for mass balance fit (aliphatic symmetric C-H stretch, carbonyl C=O stretch, and amide I).

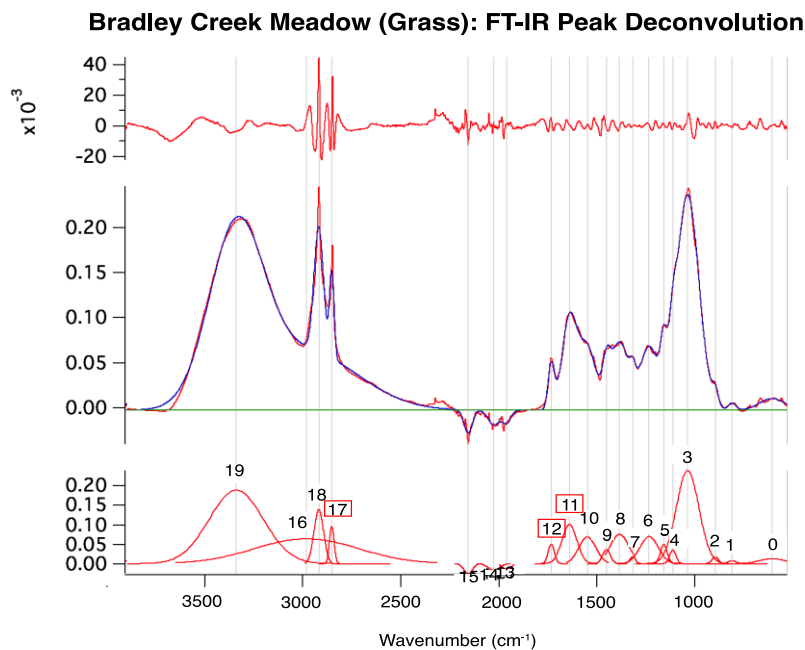


Figure S2. Example of peak deconvolution of above-ground biomass (AGB) FT-IR spectrum using the Igor Pro software package. Red boxes represent peaks chosen for mass balance fit (aliphatic symmetric C-H stretch, carbonyl C=O stretch, and amide I).

^{*} Any use of trade, firm, or product names is for descriptive purposes and does not imply endorsement by the U.S. Government.

2. Deconvoluted bulk C XAS spectra

Figures S3-4 are examples of deconvoluted bulk C XAS spectra using the Athena software package (version 0.9.25, Bruce Ravel*).

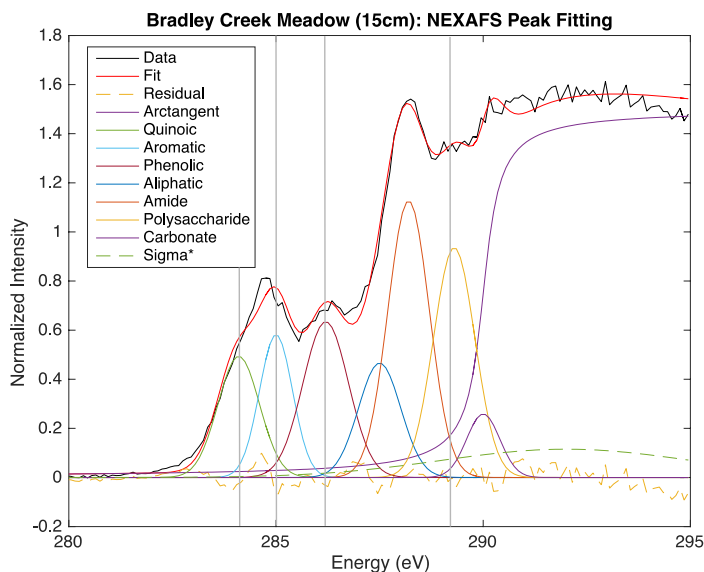


Figure S3. Example of peak deconvolution of soil bulk C XAS spectrum using the Athena software package. Vertical lines represent peaks chosen for mass balance fit (quinonic, aromatic, phenolic, and polysaccharide).

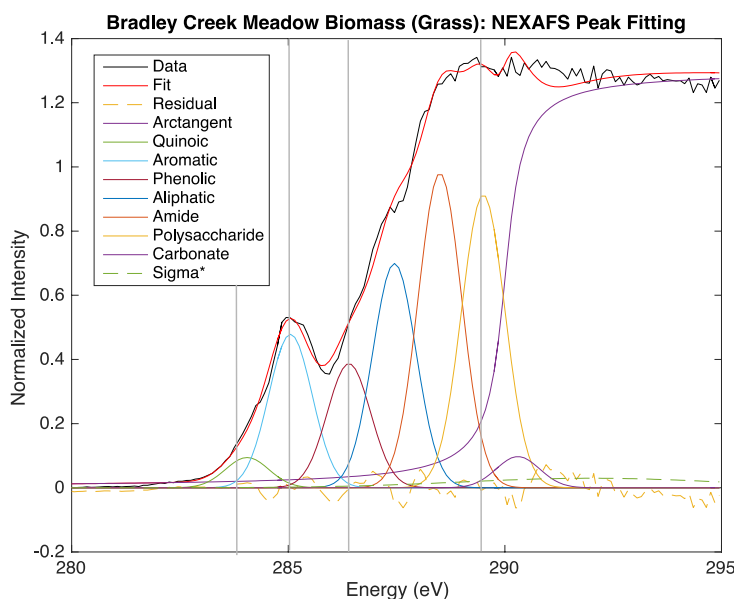


Figure S4. Example of peak deconvolution of biomass bulk C XAS spectrum using the Athena software package. Vertical lines represent peaks chosen for mass balance fit (quinonic, aromatic, phenolic, and polysaccharide).

* Any use of trade, firm, or product names is for descriptive purposes and does not imply endorsement by the U.S. Government.

3. Sample selection for equation (8)

3.1 Universal vs. site-specific vs. depth-specific fits for SOC

We tested 3 methods to apply equation (8) to soil samples to find the optimal way of applying equation (8):

Method (1): a universal fit, where all soil samples from BCM, BCW, RCM, and RCF were fitted into a single matrix using equation (8) to determine whether there is a universal set of α_i values that can be applied to all soil samples;

Method (2): a site-specific fit, where soil samples collected from the same site with various depths were fitted with a single matrix equation, assuming that depth profiles reflect similar soil matrix and SOC input;

Method (3): a depth-specific fit, where that soil samples at the same depths from different sites were fitted with the same matrix equation, assuming depth intervals associated with soil horizons are more similar across sites due to similar extents of soil development and TOC contents.

All matrix equations were optimized with the boundary condition $\alpha_i > 0$.

The fitted results for Methods (1), (2), and (3) are shown in Fig. S1-3, respectively, where the R^2 value between the optimized TOC% and experimental TOC% measured by EA is an indication of how well the fit correlates to experimental data. The calculated α_i values for all three methods are shown in Table S1. The R^2 value for Method (1) is 0.85 (Fig. S1a). However, because TOC at RCF is over ten times higher than the other three sites, when the same fit was only applied to BCM, BCW, and RCM, the R^2 value is 0.34 (Fig. S1b), and at $p < 0.05$ confidence level, the difference between the fitted data and experimental data is significant (Table S.1). Therefore, Method (1) is not an optimal approach due to potential variations in α_i across sites. Further, in Fig. S1b, while the fitted TOC% for soil samples from BCW scattered across the $y=x$ line, the fitted results for RCM are all above the $y=x$ line and those for BCM are all below the $y=x$ line, suggesting that soil samples from different sites might have different chemical properties that can result in significantly different α_i values.

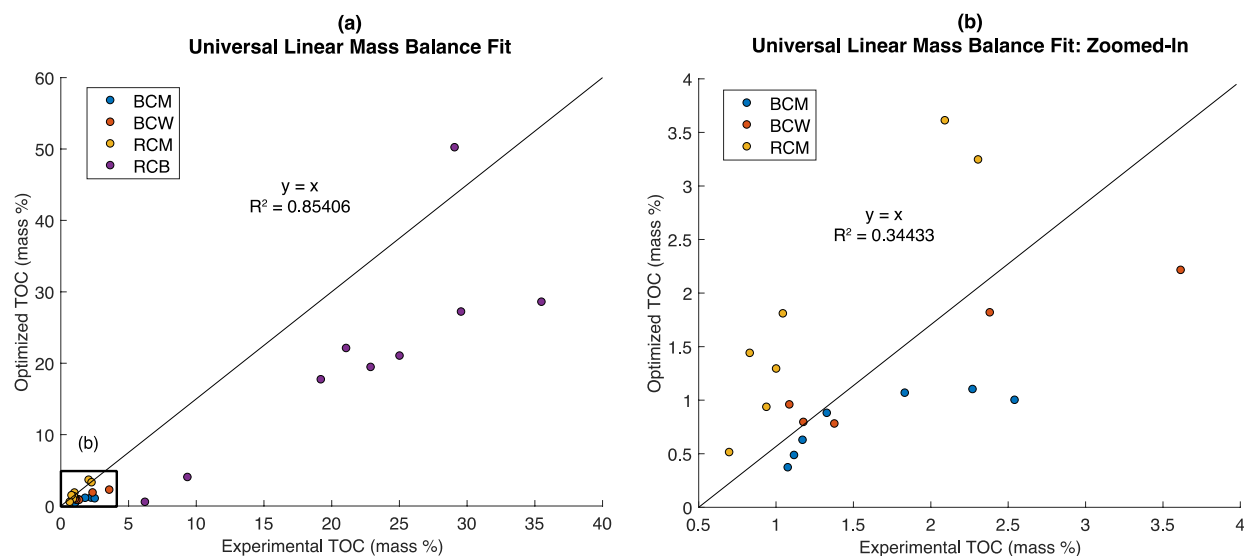


Figure S5. (a) Universal linear mass balance fit (Method (1)) for all soil samples using equation (8) ($R^2 = 0.85$, $p < 0.05$). The box shows the inset for panel (b). (b) Zoom in of universal fit for BCM, BCW, and RCM ($R^2 = 0.34$, $p > 0.05$).

Table S1. Slope t-test against $y=x$ for the universal fit ($p < 0.05$).

Unpooled error variance		Pooled error variance	
-	-	S_{Res}^2	0.589
S_{b1-b2}	0.172	S_{b1-b2}	0.203
t	2.613	t	-2.216
df	31	df	31
alpha	0.05	alpha	0.05
p-value	0.014	p-value	0.034
t-crit	2.040	t-crit	2.040
significance	yes	significance	yes

The R^2 values are higher using Method (2) ($R^2 > 0.93$) compared to Method (3) ($R^2 > 0.86$) (Fig. S2-3). Although there are variations α_i values between the two methods, there is more consistency within Method (2) (Table S2). For example, the α_i values for BCM and BCW for the same SOC species are of the same order of magnitude, which is supported by the proximity of BCM and BCW and their depositional relationship. The α_i values of RCM and RCF deviate more from those of BCM and BCW. Although RCM and RCF are also in close proximity, RCF is within a wetland area characterized by very high TOC and local inputs of carbon. In Method (3), the α_i value for aliphatic functional groups increase with depth while the α_i values for amide decrease with depth. However, the algorithm minimized more α_i values to 1E-04 for more functional groups: for soil samples deeper than 60 cm, all functional groups except for carbonyl and aliphatic are minimized, suggesting potential mathematical artefacts. Therefore, out of the three methods tested, Method (2) is assumed to be the preferable method for the application of equation (8) at our field sites. The importance of soil matrix was further supported by XRF and XRD results in the following sections.

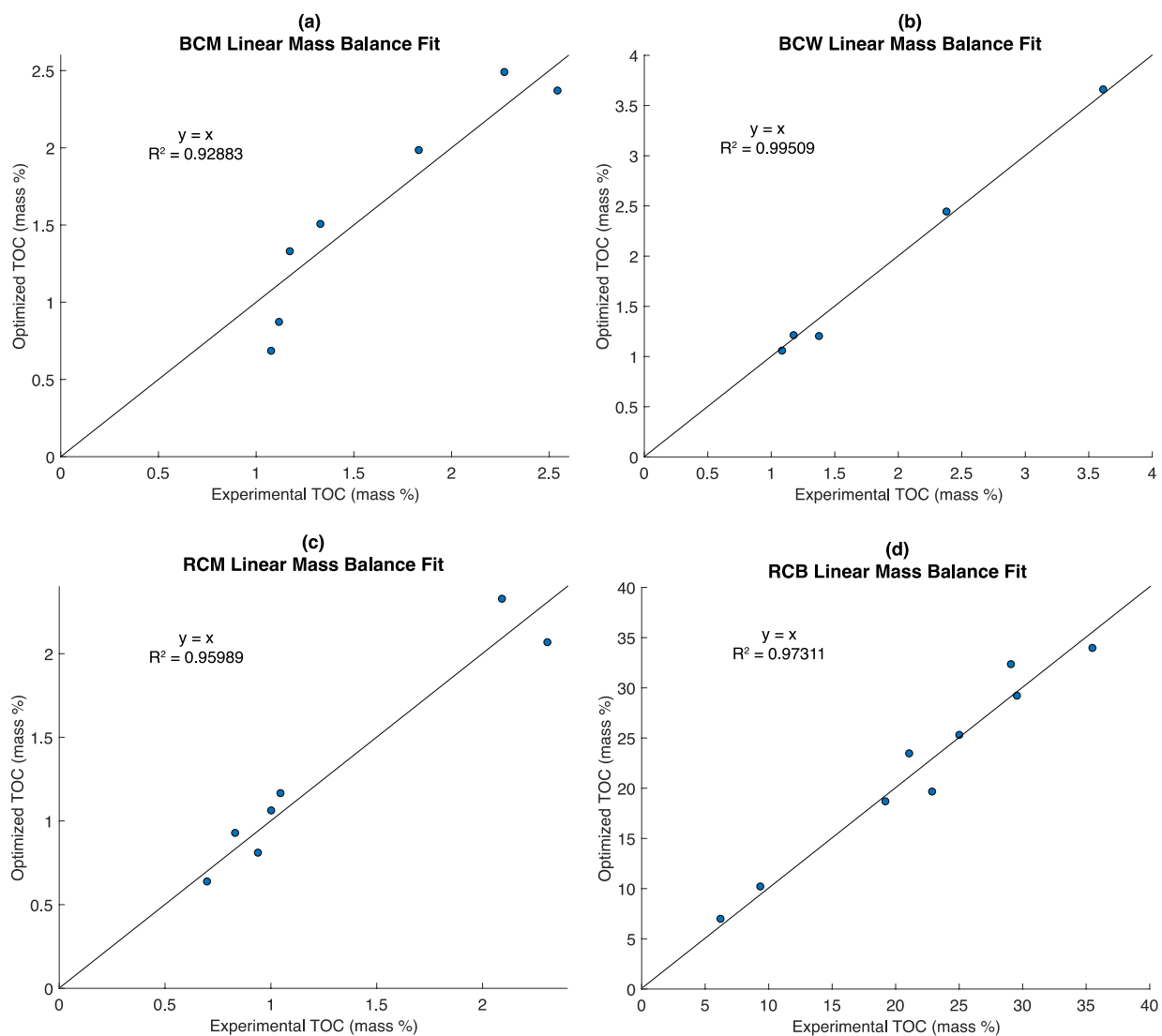


Figure S6. Site-specific linear mass balance fit (Method (2)) for soil samples at (a) BCM ($R^2 = 0.93$, $p < 0.05$); (b) BCW ($R^2 = 0.99$, $p < 0.05$); (c) RCM ($R^2 = 0.96$, $p < 0.05$); (d) RCF ($R^2 = 0.97$, $p < 0.05$) using equation (8).

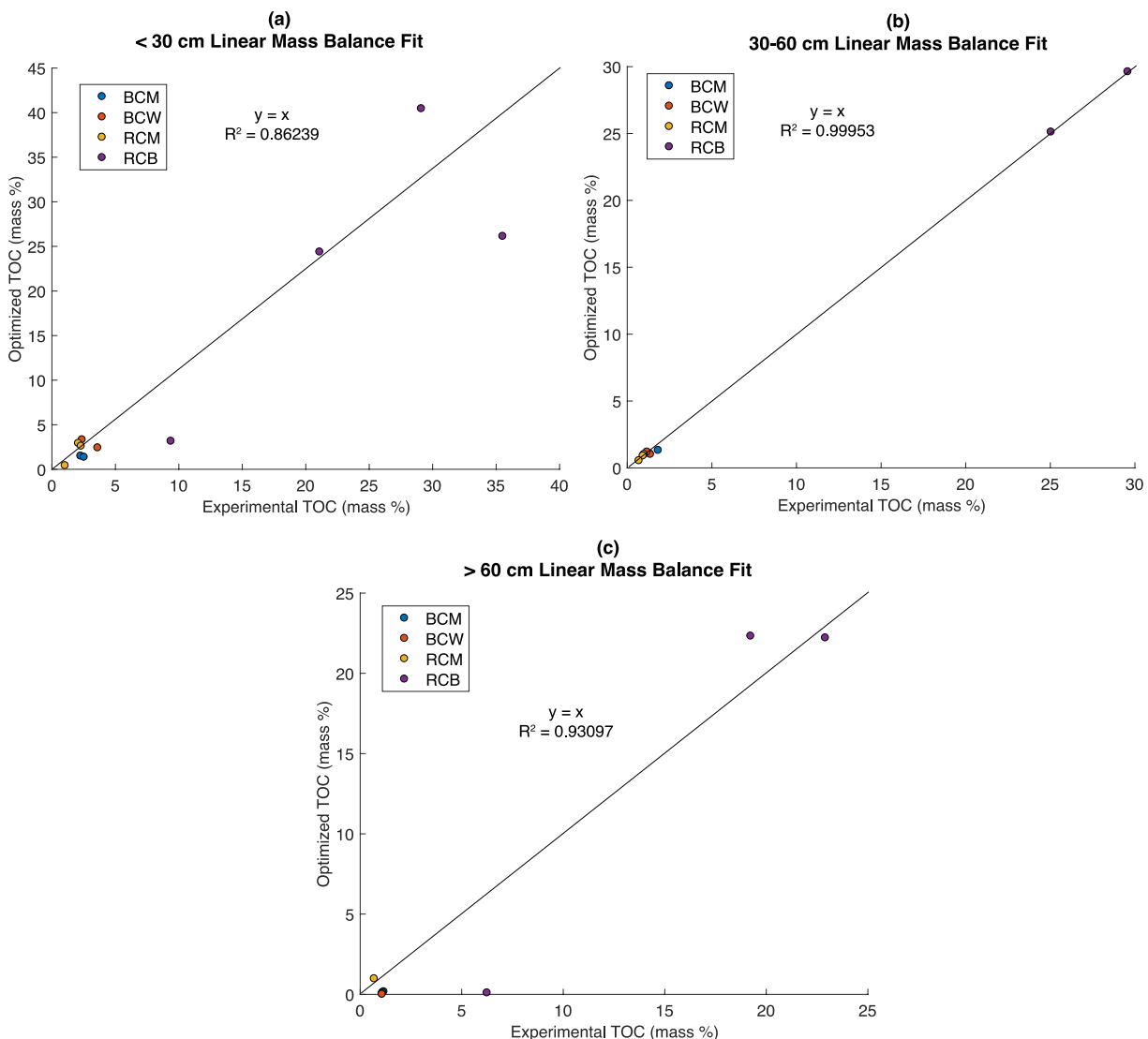


Figure S7. Depth-specific (Method (3)) linear mass balance fit for soil samples from all sites at (a) top 30 cm ($R^2 = 0.86$, $p < 0.05$); (b) 30-60cm ($R^2 = 0.99$, $p < 0.05$); and (c) > 60cm ($R^2 = 0.93$, $p < 0.05$) using equation (8).

Table S2. Fitted α_i values for SOC functional groups for Methods (2) and (3). *Carbonyl was not observed in BCM and BCW by IR, and therefore α_i cannot be solved for.

Functional Group (<i>i</i>)	Method (1)	α_i						
		Method (2)				Method (3)		
		BCM	BCW	RCM	RCF	30 cm	30-60 cm	> 60 cm
Carbonyl (IR)	1.404	*-	*-	1.188	4.400	0.586	16.810	9.696
Aliphatic (IR)	15.14	93.93	43.67	1E-04	2.714	7.231	22.088	34.084
Amide (IR)	0.228	0.494	0.156	0.459	0.177	1.157	0.154	1E-04
Quinonic (XAS)	0.391	1E-04	1E-04	1E-04	1E-04	1E-04	1.542	1E-04
Aromatic (XAS)	1E-04	1E-04	1E-04	1E-04	11.20	1E-04	1E-04	1E-04
Phenolic (XAS)	1E-04	0.264	0.448	0.702	1E-04	1E-04	1E-04	1E-04
Polysaccharide (XAS)	1E-04	0.765	0.796	1E-04	1E-04	1E-04	1E-04	1E-04

3.2. XRF

Since ε (and hence α_i) value can be affected by soil matrices, X-ray fluorescence spectroscopy (XRF) was used to study the elemental composition of soil samples from BCM, BCW, RCM and RCF to determine whether soil samples from the same site have similar characteristics. The XRF results (Fig. S4-7) indicate that elemental composition is relatively consistent throughout depth at BCM, BCW, and RCM, with the exception of an increase in Ca below 60 cm at BCM, which is consistent with the calcite reaction front identified with EA (Fig. S12). The element content also varies by depth at RCF because RCF is a deposition fen. XRF data from BCM, BCW, and RCM support using Method (2) for sample selection because overall, soil samples collected from the same site have more similar composition and thus more similar ε and α_i values. The average element contents and standard deviation for each site are summarized in Table S3. The three elements with the highest mass % (Al, Si, and Ca) vary more between BCM than BCW and RCM due to the increase in Ca at the calcite reaction front. For all trace elements measured in ppm (P, S, Mn, V, Cr, As, Se, U), BCM and BCW have values close to each other, supporting the similarity in characteristics between the two sites.

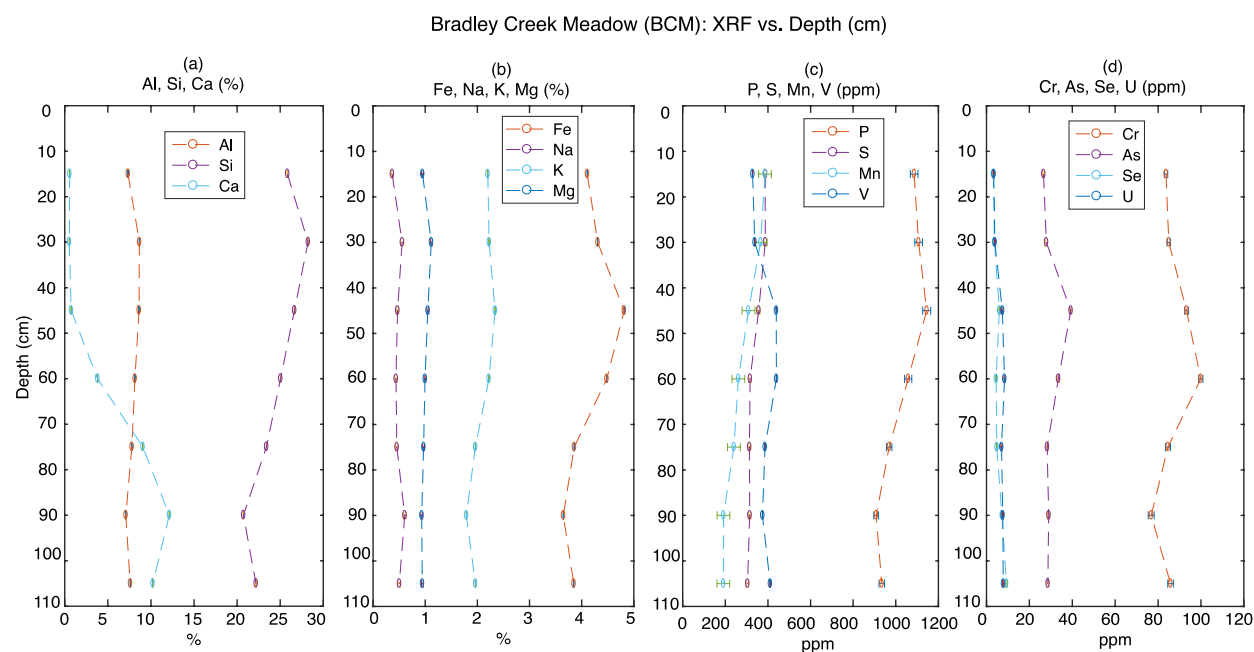


Figure S8. XRF spectra for BCM for (a) Al, Si, Ca (> 5%); (b) Fe, Na, K, Mg (< 5%); (c) P, S, Mn, V (> 120ppm); and (d) Cr, As, Se, U (< 120ppm).

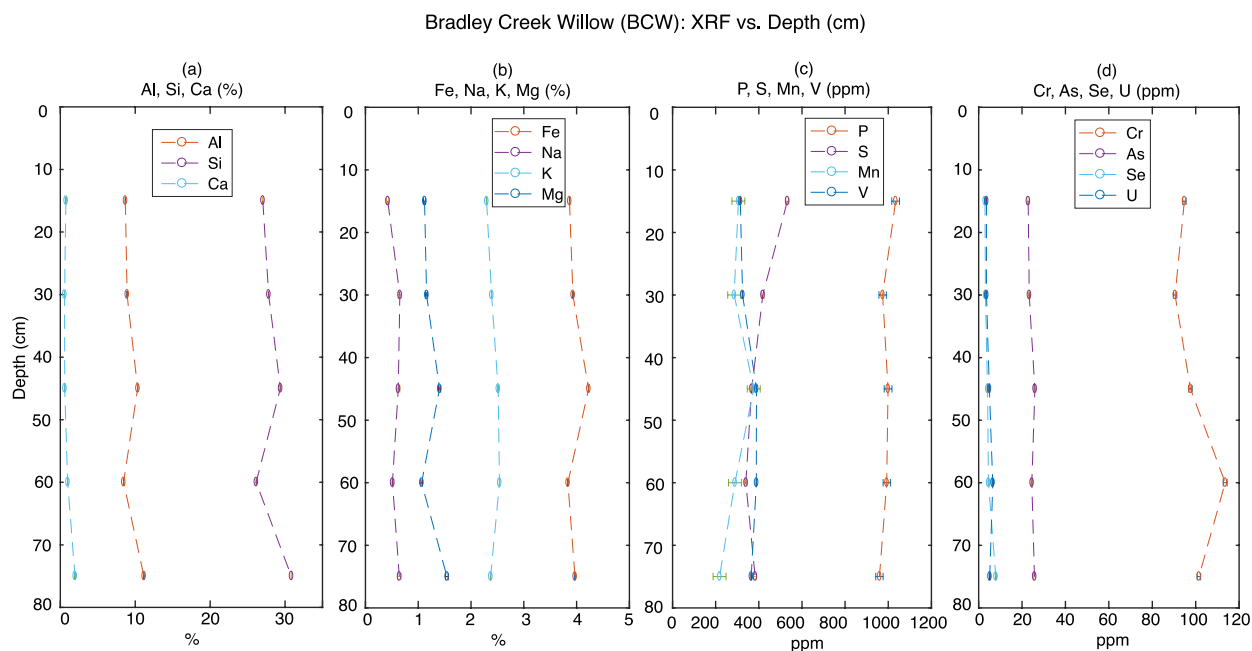


Figure S9. XRF spectra for BCW for (a) Al, Si, Ca (> 5%); (b) Fe, Na, K, Mg (< 5%); (c) P, S, Mn, V (> 120ppm); and (d) Cr, As, Se, U (< 120ppm).

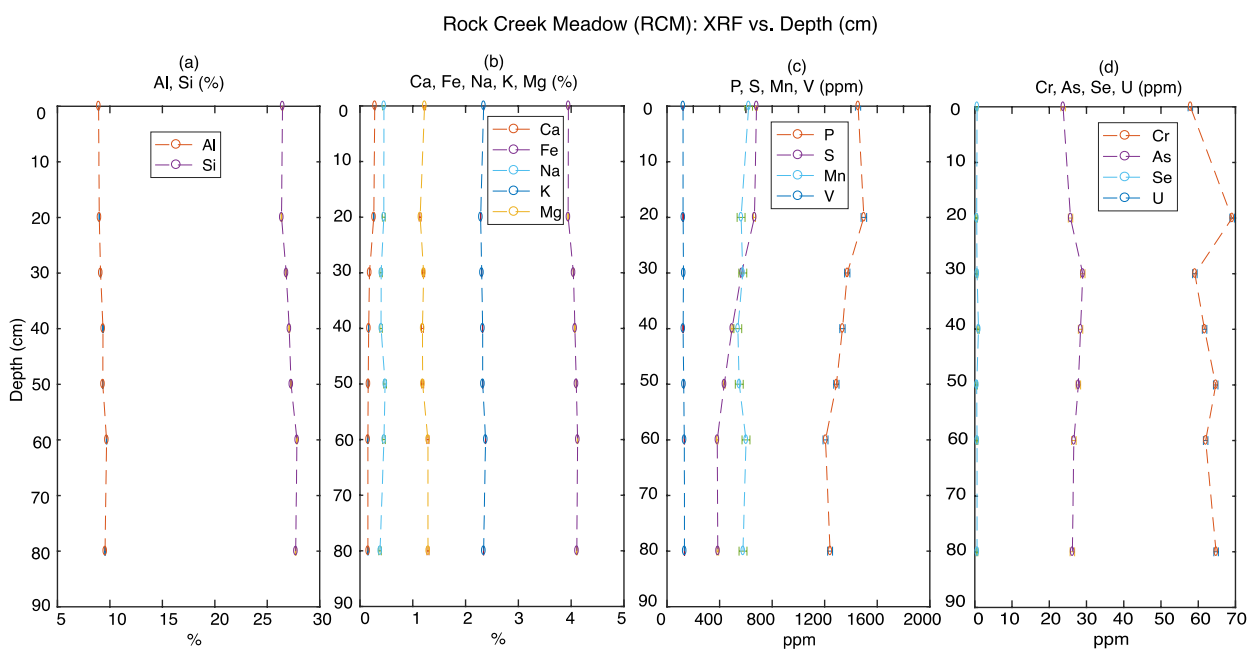


Figure S10. XRF spectra for RCM for (a) Al, Si (> 5%); (b) Ca, Fe, Na, K, Mg (< 5%); (c) P, S, Mn, V (> 70ppm); and (d) Cr, As, Se, U (< 70ppm).

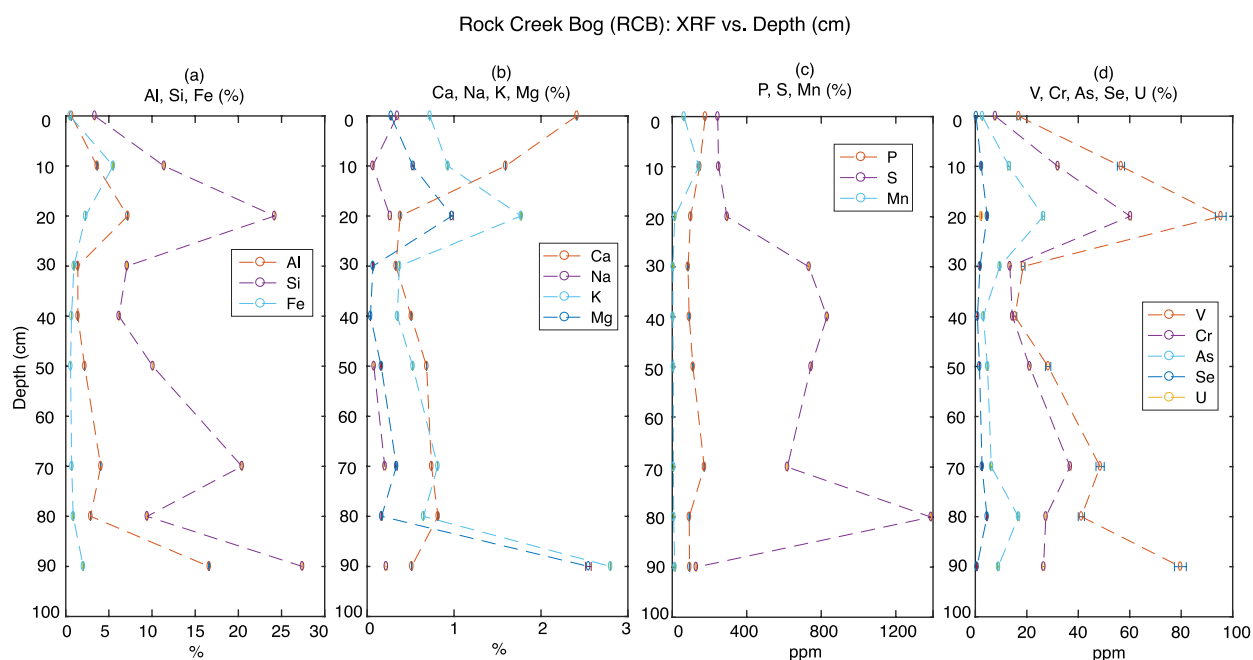


Figure S11. XRF spectra for RCM for (a) Al, Si, Fe (> 3%); (b) Ca, Na, K, Mg (< 3%); (c) P, S, Mn (> 100 ppm); and (d) V, Cr, As, Se, U (< 100ppm).

Table S3. The average element content and standard deviation of BCM, BCD, RCM, RCF measured by XRF.

Element	Average mass % (stdev.)			
	BCM	BCW	RCM	RCF
Al	7.889 (0.610)	9.503 (1.183)	9.273 (0.283)	4.442 (4.954)
Si	24.583 (2.625)	28.240 (1.879)	27.074 (0.575)	13.28 (8.552)
Ca	5.291 (5.05)	0.967 (0.576)	0.192 (0.055)	0.886 (0.684)
Fe	4.153 (0.410)	3.966 (0.156)	4.047 (0.739)	1.560 (1.610)
Na	0.484 (0.079)	0.570 (0.099)	0.431 (0.037)	0.130 (0.130)
K	2.010 (0.197)	2.422 (0.103)	2.329 (0.030)	0.991 (0.800)
Mg	0.994 (0.067)	1.258 (0.206)	1.217 (0.055)	0.566 (0.797)
Average ppm (stdev.)				
P	1030 (92.01)	992 (28.40)	1342 (107.5)	118 (36.8)
S	339.9 (36.41)	408.0 (74.95)	515.9 (125.2)	581 (40.0)
Mn	277.81 (78.76)	294.9 (56.56)	575.36 (28.57)	285.2 (465.7)
V	388.04 (44.79)	356.14 (35.19)	125.91 (4.55)	44.5 (28.61)
Cr	87.01 (7.44)	99.60 (8.80)	62.76 (3.80)	26.61 (15.71)
As	30.46 (4.46)	24.44 (1.39)	26.77 (1.81)	10.14 (7.65)
Se	5.529 (2.15)	4.50 (2.02)	0.629 (0.170)	2.12 (1.57)
U	6.50 (2.00)	4.74 (1.21)	< 0.4	< 0.4

3.3. PXRD

In addition to XRF, powder X-ray diffraction (PXRD) was used to study the structure of soil matrices from BCM, BCW, RCM, and RCF. The PXRD spectra are shown in Fig. S8. The results show that at all sites, soil matrix structure is relatively consistent through depth with the exception of additional calcite peaks at BCM after 60 cm. The major mineral compositions at BCM, BCW, and RCM are silicate and muscovite, with a small amount of clinochlore identified in BCM, and BCW. This further supports the similarity between BCM and BCW, likely due to the spatial proximity between the two sites. The major mineral composition at RCF is also silicate, but since it is a deposition fen, the minor mineral composition at each depth varies. Similar to the other three sites, muscovite is identified at 10 and 20 cm at RCF, and like RCM, clinochlore is found at 30 and 90 cm. Nontronite is identified from 40 – 80 cm at RCF.

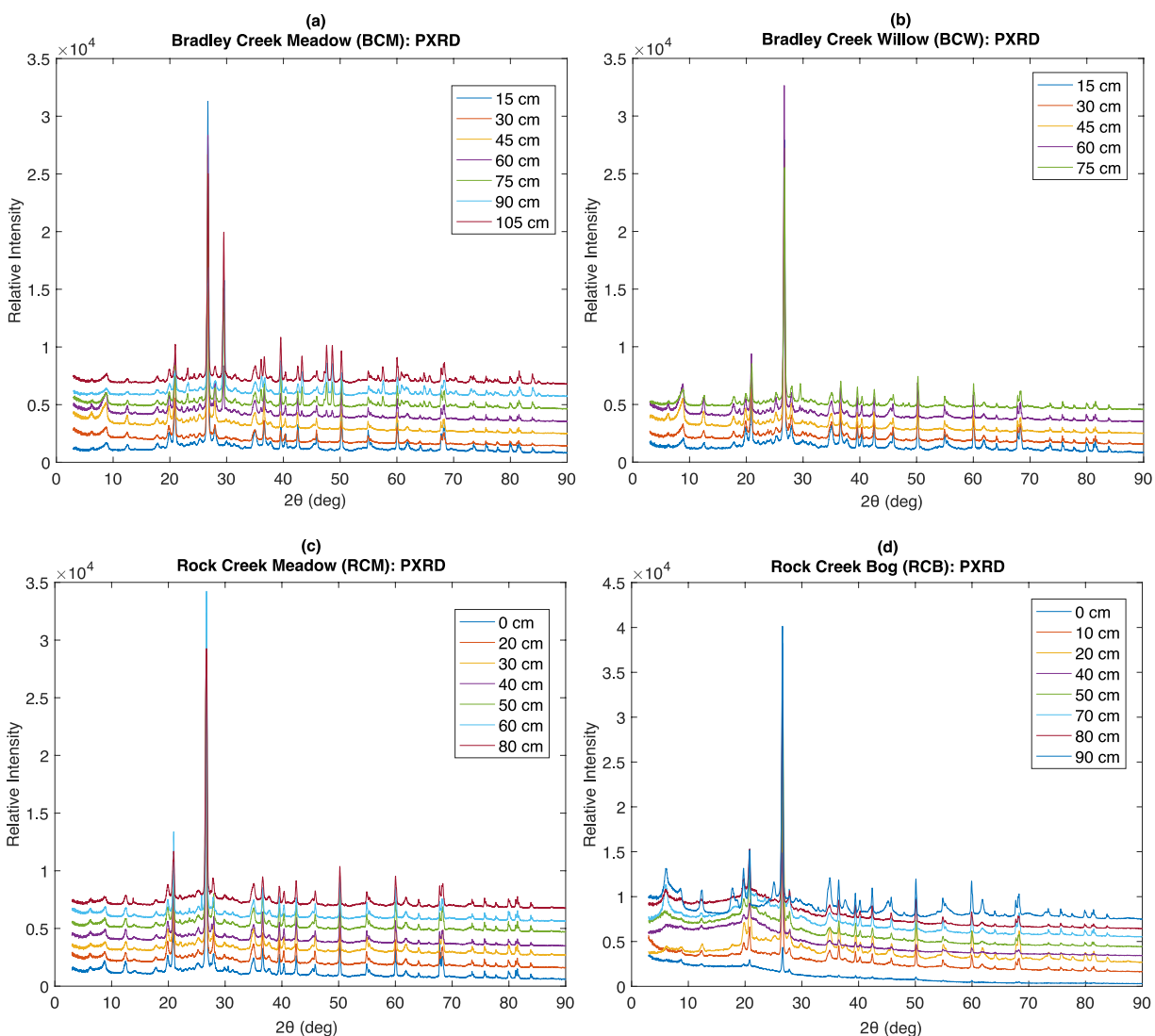


Figure S12. PXRD depth profiles for (a) BCM, (b) BCW, (c) RCM, and (d) RCF.

3.4. Sample selection conclusions

After testing Methods (1)-(3), we concluded that equation (8) can be applied to soil samples collected from any location as long as there are enough samples to constrain the optimized solution (i.e. more soil samples than the number of functional groups). However, the α_i values reported in this study cannot be directly applied to soils from locations with drastically different environmental characteristics. Site-specific fits in Method (2) yielded more reliable fitting results than depth-specific fits in Method (3). In addition, since the α_i values are relatively similar for each functional group between BCM and BCW using Method (2), possibly due to the proximity and similarity between the two sites, all samples from BCM and BCW were fitted with the same matrix to provide more constraints for the solved α_i values. The results are shown in Fig. S9 ($R^2 = 0.93$, $p < 0.05$) with a set of α_i values that can be applied to samples across the two sites.

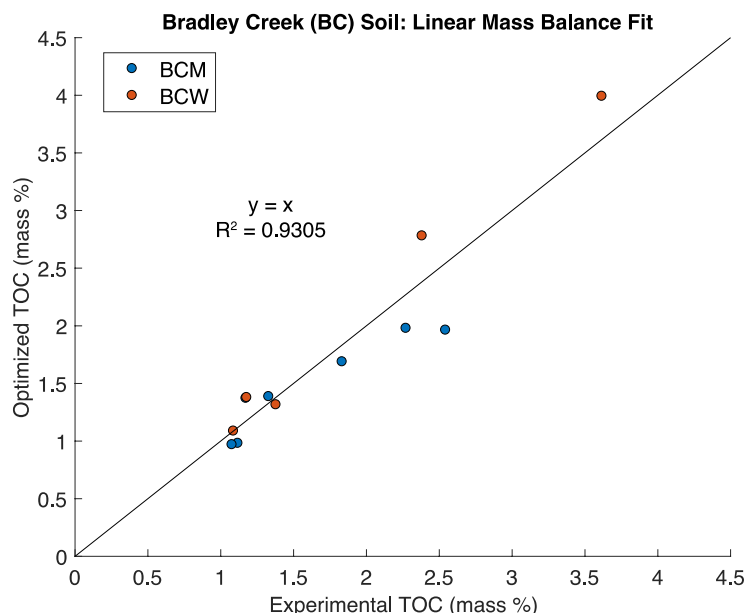


Figure S13. Linear mass balance fit for all soil samples from Bradley Creek (BCM and BCW, $R^2 = 0.93$, $p < 0.05$) using equation 8.

4. Uncertainties in α_i : MW_i vs. final outcome of linear mass balance fit

We used MW_i (i is an index representing the functional group) as an example to demonstrate how the final outcome of the linear mass balance fit will not be affected we change a value (MW_i , n , d , ε or b) that is consistent within each functional group. Table S4 shows the MW_i for all functional groups under 2 scenarios: (1) $MW_i = MW$ of only the functional groups since absorption spectroscopy is only identifying functional groups rather than molecules, or (2) $MW_i =$ average MW of the class of compound that is abundant in each functional group. Equation (7) is modified into equation (S1):

$$TOC(mass\%) = \sum_i \frac{A_i \times MW_i \times \beta_i}{density_{soil}}, \quad (S1)$$

where β_i is defined in equation (S2):

$$\beta_i = \frac{1}{n_i \times d \times \varepsilon_i \times b}, \quad (S2)$$

and equation (8) is modified into equation (S3).

$$\begin{matrix}
1 & \frac{A_{C=O}^1 \times MW_{C=O}}{density_{soil}^1} & \frac{A_{CH_3}^1 \times MW_{CH_3}}{density_{soil}^1} & \frac{A_{NC=O}^1 \times MW_{NC=O}}{density_{soil}^1} & \frac{A_{quin}^1 \times MW_{quin}}{density_{soil}^1} & \frac{A_{arom}^1 \times MW_{arom}}{density_{soil}^1} & \frac{A_{phen}^1 \times MW_{phen}}{density_{soil}^1} & \frac{A_{O-alkyl}^1 \times MW_{O-alkyl}}{density_{soil}^1} \\
2 & \frac{A_{C=O}^2 \times MW_{C=O}}{density_{soil}^2} & & & & & & \\
3 & \frac{A_{C=O}^3 \times MW_{C=O}}{density_{soil}^3} & & & & & & \\
4 & \frac{A_{C=O}^4 \times MW_{C=O}}{density_{soil}^4} & & & & & & \\
5 & \frac{A_{C=O}^5 \times MW_{C=O}}{density_{soil}^5} & & & & & & \\
6 & \frac{A_{C=O}^6 \times MW_{C=O}}{density_{soil}^6} & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
7 & \vdots & & & & & & \\
8 & \frac{A_{C=O}^8 \times MW_{C=O}}{density_{soil}^8} & & & & & & \\
9 & \frac{A_{C=O}^9 \times MW_{C=O}}{density_{soil}^9} & & & & & & \\
10 & \frac{A_{C=O}^{10} \times MW_{C=O}}{density_{soil}^{10}} & & & & & & \\
& \vdots & & & & & &
\end{matrix}
\times
\begin{bmatrix}
\beta_{C=O} \\
\beta_{CH_3} \\
\beta_{NC=O} \\
\beta_{quin} \\
\beta_{arom} \\
\beta_{phen} \\
\beta_{O-alkyl}
\end{bmatrix}
=
\begin{bmatrix}
TOC^1 \\
TOC^2 \\
TOC^3 \\
TOC^4 \\
TOC^5 \\
TOC^6 \\
TOC^7 \\
TOC^8 \\
TOC^9 \\
TOC^{10} \\
\vdots
\end{bmatrix} \quad (S3)$$

Table S4. MW_i when 1) $MW_i = MW$ of only the functional groups, and 2) $MW_i =$ average MW of the class of compound that is abundant in each functional group.

Functional Group (<i>i</i>)	Scenario	
	(1) $MW_i = MW_{i(f.g.)}$	(2) $MW_i = MW_{i(cpmd)}$
Carbonyl (IR)	28	10,000 (lignin)[1]
Aliphatic (IR)	15	500 (lipid)[2,3]
Amide (IR)	42	30,000 (protein)[4]
Quinonic (XAS)	28	10,000 (lignin)[1]
Aromatic (XAS)	24	10,000 (lignin)[1]
Phenolic (XAS)	28	10,000 (lignin)[1]
Polysaccharide (XAS)	31	50,000[5]

Using all soil samples from Bradley Creek (BCM, BCW) as an example, Fig. S10 shows the linear mass balance fit when equation (S3) under scenario 1 (Fig. 10a), and equation (S3) under scenario 2 (Fig. 10b) were applied. The respective α_i and β_i values are summarized in Table S5. The results showed that although the α_i and β_i values between the three plots are different, the fits in Fig. S10 turns out to be identical to that in Fig. S9, where equation (8) was applied to BC soils. Further, the relationship between α_i and β_i values is demonstrated in equation (S4).

$$\alpha_i = MW_{i(f.g.)} \times \beta_{i(f.g.)} = MW_{i(compd)} \times \beta_{i(compd)}, \quad (S4)$$

This shows that the different values simply result from a linear algebra operation rather than the different molecular weights. In this study, we are interested in the final output of the linear mass balance fit and we do not intend to solve for n , MW_i , d , b , and ε individually. Therefore, assuming that n and MW_i values are consistent within the same sites due to similar SOC input and soil matrix, the actual value of n , MW_i , d , b and ε does not affect the final fit.

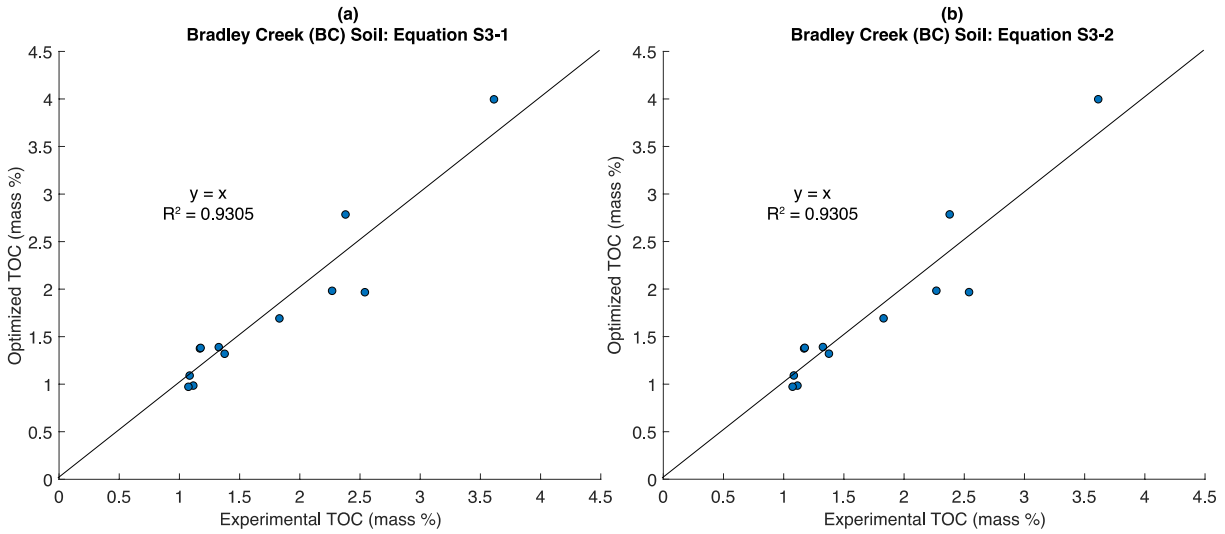


Figure S14. Linear mass balance fit for Bradley Creek soil samples when (a) equation (8), (b) equation (S3) under scenario 1, and (c) equation (S3) under scenario 2 was applied.

Table S5. α_i and β_i values for Bradley Creek soil samples when (a) equation 8, (b) equation S3 under scenario 1, and (c) equation S3 under scenario 2 was applied. *Carbonyl was not observed in BCM and BCW by IR, and therefore α and α_{MW} cannot be solved for.

Functional Group (i)	α_i (Equation 8)	β_i (Equation S3-1)	β_i (Equation S3-2)
Carbonyl (IR)	_*	_*	_*
Aliphatic (IR)	44.365	2.958	0.089
Amide (IR)	0.0338	8E-04	1E-6
Quinonic (XAS)	1E-04	4E-06	1E-10
Aromatic (XAS)	1E-04	4E-06	1E-10
Phenolic (XAS)	1E-04	4E-06	1E-10
Polysaccharide (XAS)	1.3226	0.043	3E-05

5. Boundary conditions of the linear mass balance fit and slope t-tests

In Table S2, the algorithm minimized α_i values for several functional groups ($\alpha_i = 1E-04$) when $\alpha_i > 0$ is the only boundary condition. The resulting mass% is lower than the detection limits of ATR FT-IR and XAS (tens of ppm) [6,7]. Thus, different boundary conditions were tested to correct for the fitting results. Fig. S10 shows the fitting results when the boundary conditions $\alpha_i > 0.001$, $\alpha_i > 0.01$, and $\alpha_i > 0.1$ were applied, using the soil samples from Bradley Creek (BCM, BCW) as an example. The results show that the R^2 value for the fitted and experimental TOC% decreases slightly with increasing α_i boundary, but the lowest R^2 value (0.91) is still acceptable. The α_i values are listed in Table S6. No matter which boundary condition was applied, the algorithm minimized the same three functional groups: quinonic, aromatic, and phenolic carbon. Slope t-tests against $y=x$ (Table S7) and correlation analysis of α_i values between different boundary conditions (Table S8) suggest that $\alpha_i > 0.01$ yields the best results while ensuring that the calculated mass% for each functional group is above the detection limits of IR and XAS. Therefore, $\alpha_i > 0.01$ will be used as the boundary condition for all linear mass balance fit in this work.

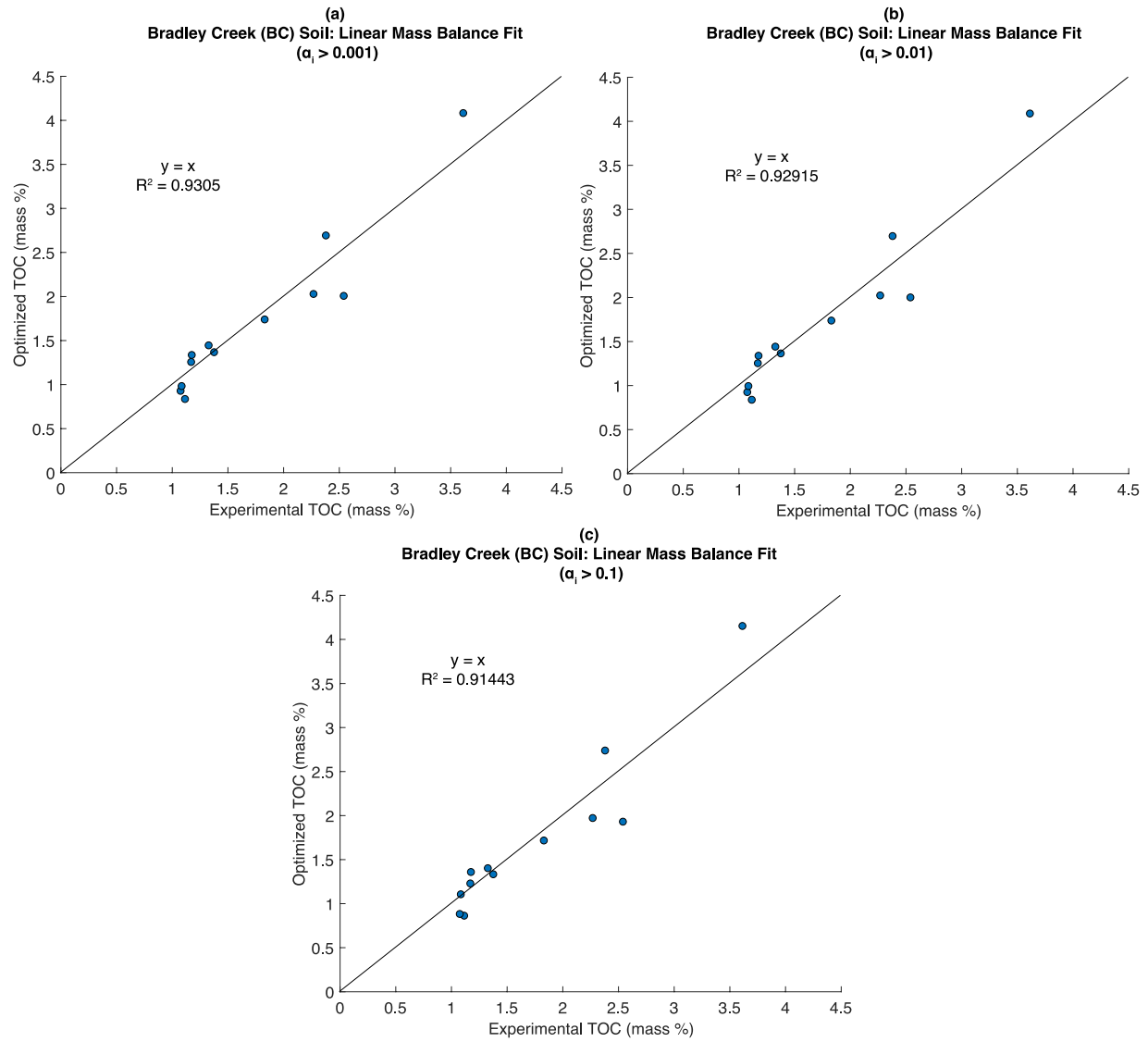


Figure S15. Linear mass balance fit for soil samples from Bradley Creek (BC) with boundary conditions (a) $\alpha_i > 0.001$ ($R^2 = 0.93$, $p < 0.05$); (b) $\alpha_i > 0.01$ ($R^2 = 0.93$, $p < 0.05$); and (c) $\alpha_i > 0.1$ ($R^2 = 0.1$, $p < 0.05$) using equation 8.

Table S6. Fitted α_i values for SOC functional groups for soil samples from Bradley Creek when different boundary conditions are applied. *Carbonyl was not observed in BCM and BCW by IR.

Functional Group (<i>i</i>)	α_i		
	Boundary Condition		
	$\alpha_i > 0.001$	$\alpha_i > 0.01$	$\alpha_i > 0.1$
Carbonyl (IR)	_*	_*	_*
Aliphatic (IR)	44.37	44.84	49.07
Amide (IR)	0.0338	0.0434	0.133
Quinonic (XAS)	1E-04	1E-02	1E-01
Aromatic (XAS)	1E-04	1E-02	1E-01
Phenolic (XAS)	1E-04	1E-02	1E-01
Polysaccharide (XAS)	1.323	1.450	1.08

Table S7. Slope t-tests against $y=x$ on fits obtained under the boundary conditions $\alpha' > 0$, $\alpha' > 0.001$, $\alpha' > 0.01$, and $\alpha' > 0.1$.

Fit		$\alpha' > 0$	$\alpha' > 0.001$	$\alpha' > 0.01$	$\alpha' > 0.1$
BC	p-value	0.424	0.467	0.469	0.497
	p-value (pooled)	0.643	0.673	0.672	0.659
	significance	no	no	no	no
RCM	p-value	0.952	0.875	0.978	0.876
	p-value (pooled)	0.966	0.914	0.984	0.915
	significance	no	no	no	no
Litter	p-value	0.722	0.724	0.722	0.876
	p-value (pooled)	0.162	0.165	0.162	0.915
	significance	no	no	no	no
Plants	p-value	0.518	0.524	0.524	0.523
	p-value (pooled)	0.090	0.093	0.093	0.098
	significance	no	no	no	no

Table S8. Correlation analysis between fits obtained from boundary conditions $\alpha' > 0$, $\alpha' > 0.001$, $\alpha' > 0.01$, and $\alpha' > 0.1$ at all sites.

R²	$\alpha' > 0$	$\alpha' > 0.001$	$\alpha' > 0.01$	$\alpha' > 0.1$
	BC			
$\alpha' > 0$	1			
$\alpha' > 0.001$	0.999	1		
$\alpha' > 0.01$	0.999	0.999	1	
$\alpha' > 0.1$	0.999	0.999	0.999	1
	RCM			
$\alpha' > 0$	1			
$\alpha' > 0.001$	0.999	1		
$\alpha' > 0.01$	0.999	0.999	1	
$\alpha' > 0.1$	0.999	0.999	0.999	1
	All soils			
$\alpha' > 0$	1			
$\alpha' > 0.001$	0.999	1		
$\alpha' > 0.01$	0.992	0.992	1	
$\alpha' > 0.1$	0.591	0.548	0.602	1
	Litter			
$\alpha' > 0$	1			
$\alpha' > 0.001$	0.999	1		
$\alpha' > 0.01$	0.999	0.999	1	
$\alpha' > 0.1$	0.999	0.999	0.999	1
	Plants			
$\alpha' > 0$	1			
$\alpha' > 0.001$	0.999	1		
$\alpha' > 0.01$	0.999	0.999	1	
$\alpha' > 0.1$	0.999	0.999	0.999	1
	All biomass			
$\alpha' > 0$	1			
$\alpha' > 0.001$	0.939	1		
$\alpha' > 0.01$	0.939	0.999	1	
$\alpha' > 0.1$	0.943	0.999	0.999	1

6. Elemental Analysis: TIC, TOC, and TN

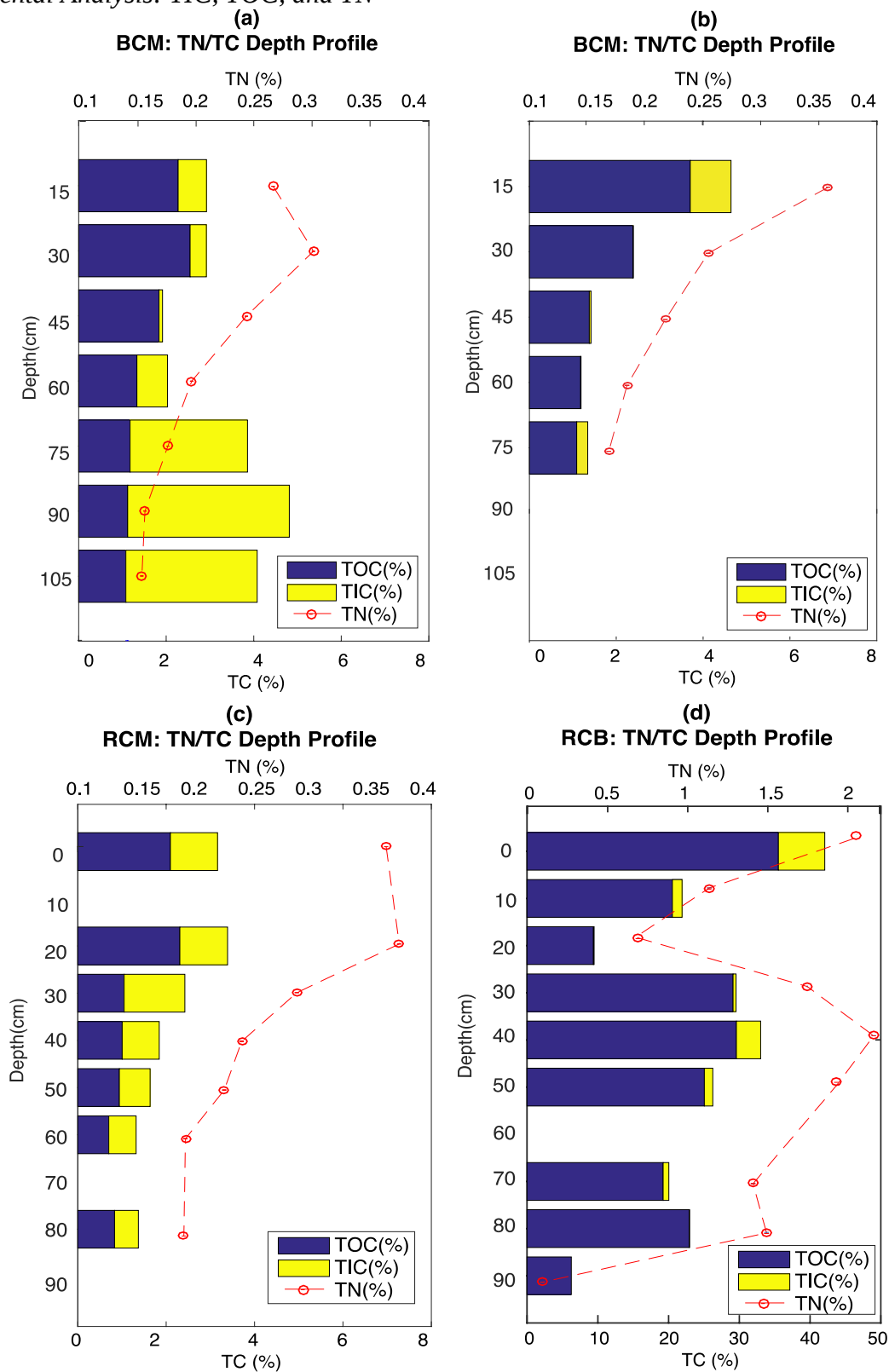


Figure S16. TIC, TOC, and TN depth profiles at (a) BCM; (b) BCW; (c) RCM; (d) RCF.

7. Bulk density of soils and biomass

The bulk densities of soils and biomasses were estimated in the laboratory. A 2 mL snap-cap Eppendorf tube was tared and filled with a soil sample or a biomass sample, and the filled tube was weighed. The mass of the sample was divided by the volume of the sample (2 mL) to obtain the estimated bulk density of the sample.

The bulk density of soil and biomass samples is shown in Figure S18. At Bradley Creek, soil density increases with depth, and BCM and BCW have identical densities from 30-60 cm (Fig. 18a). The soil density is more even through depth at RCM, while the trend of soil density is less obvious at RCF because it is a deposition fen (Fig. 18a). The litter samples have slightly higher density than plants (Fig. 18b).

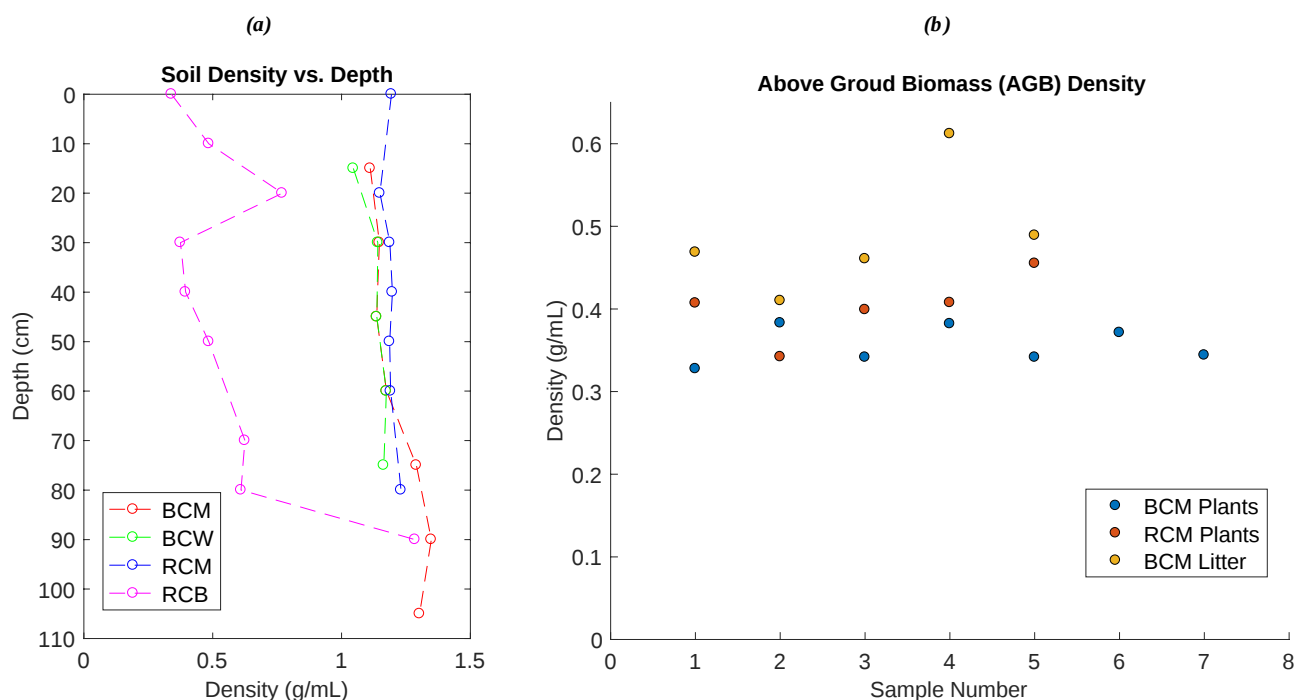


Figure S17. Bulk density of (a) soil samples and (b) AGB samples.

8. Biomass FT-IR spectra

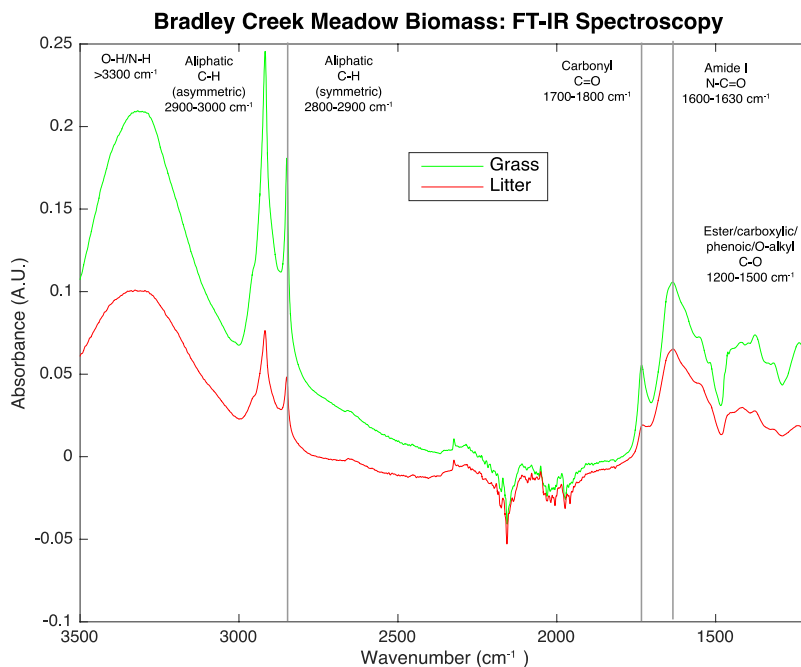


Figure S18. FT-IR spectra of Bradley Creek Meadow (BCM) biomass. Vertical lines represent peaks chosen for mass balance fit (aliphatic symmetric C-H stretch, carbonyl C=O stretch, and amide I).

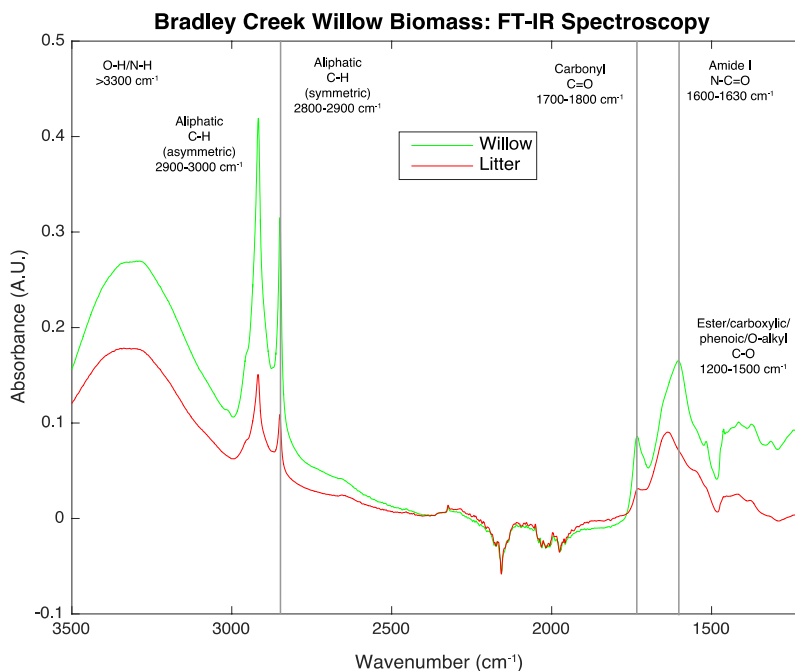


Figure S19. FT-IR spectra of Bradley Creek Willow (BCW) biomass. Vertical lines represent peaks chosen for mass balance fit (aliphatic symmetric C-H stretch, carbonyl C=O stretch, and amide I).

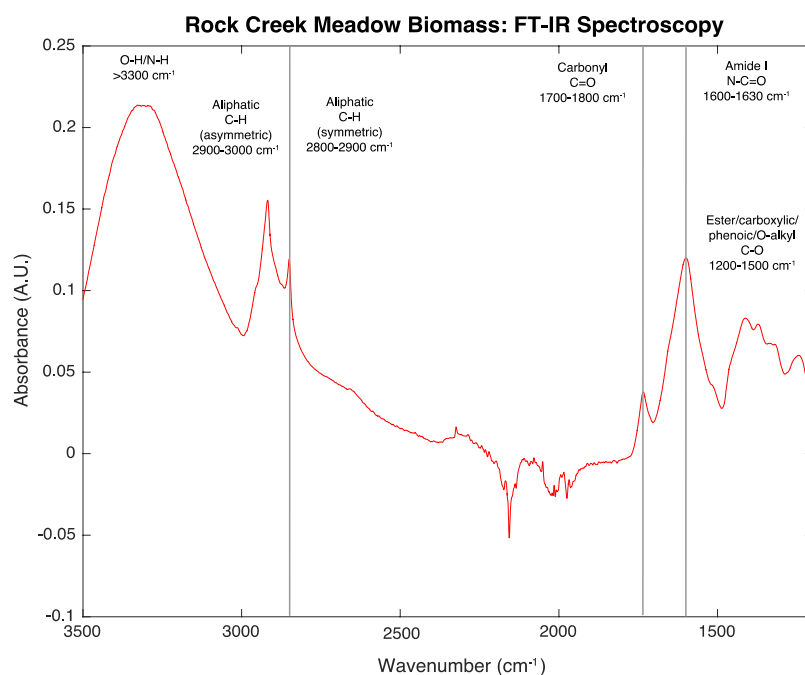


Figure S20. FT-IR spectra of Rock Creek Meadow (RCM) biomass. Vertical lines represent peaks chosen for mass balance fit (aliphatic symmetric C-H stretch, carbonyl C=O stretch, and amide I).

9. Biomass bulk C XAS spectra

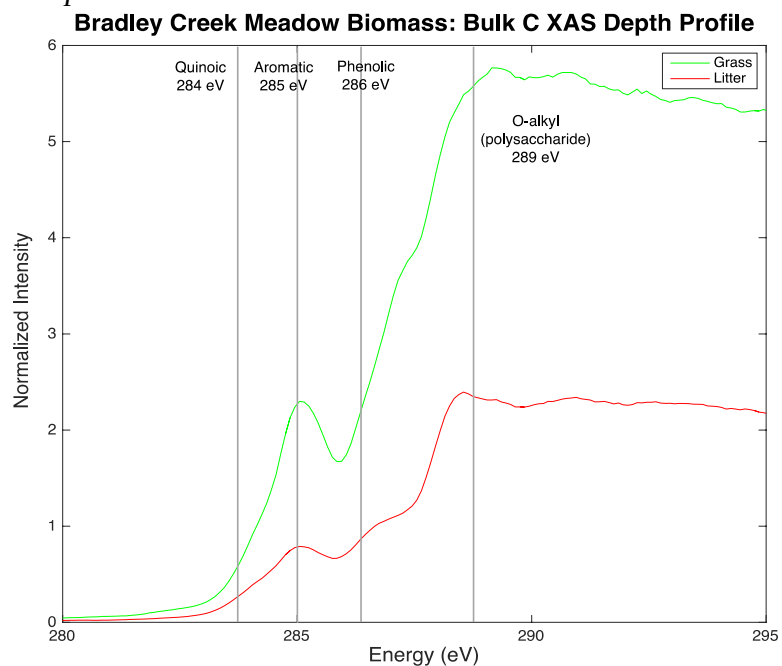


Figure S21. Bulk C XAS spectra of Bradley Creek Meadow (BCM) biomass. Vertical lines represent peaks chosen for mass balance fit (quinonic, aromatic, phenolic, and polysaccharide).

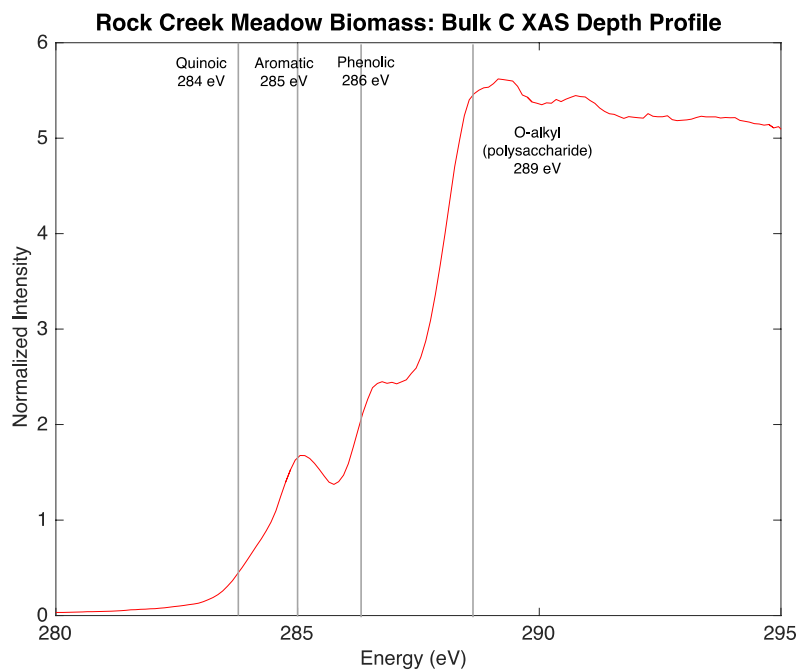


Figure S22. Bulk C XAS spectra of Rock Creek Meadow (RCM) biomass. Vertical lines represent peaks chosen for mass balance fit (quinonic, aromatic, phenolic, and polysaccharide).

10. Organic carbon composition for BCM, BCW, and RCM: α_i values

The peak areas and α_i values of linear mass balance fits for soil and biomass samples using $\alpha_i > 0.01$ as the boundary condition are summarized in Tables S10-13. The correlation between experimental and fitted TOC% for BCM and BCW are shown in Fig. S18b, and the remaining are shown in Figures S29.

Table S9. Peak area and α_i values from linear mass balance fit for Bradley Creek soils.

Bradley Creek Soils Mass Balance Fit							
Depth (cm)	Peak area						
	Carbonyl (IR)	Aliphatic (IR)	Amide (IR)	Quinonic (XAS)	Aromatic (XAS)	Phenolic (XAS)	Polysaccharide (XAS)
BCM							
15	0	0.0103	1.35	0.62	0.58	0.87	1.17
30	0	0.0105	1.27	0.55	0.77	0.7	1.2
45	0	0.0043	1.50	0.75	0.9	0.83	1.16
60	0	0	1.52	0.64	0.91	0.66	1.1
75	0	0.0060	0.60	0.61	0.89	0.64	0.89
90	0	0.0045	0.26	0.64	0.9	0.58	0.61
105	0	0.0016	0	0.65	0.9	0.6	0.76
BCW							
15	0	0.0497	1.86	0.42	0.75	0.66	1.34
30	0	0.0210	3.13	0.43	0.71	0.65	1.36
45	0	0.0008	1.13	0.62	0.79	0.81	0.99
60	0	0.0044	0.77	0.76	1.01	0.87	0.9
75	0	0	1.40	0.83	0.9	0.95	0.73
α_i	0	44.84	0.043	0.011	0.011	0.011	1.45

Table S10. Peak area and α_i values from linear mass balance fit for RCM soils.

Rock Creek Fen Soils Mass Balance Fit							
Depth	Peak area						
	Carbonyl (IR)	Aliphatic (IR)	Amide (IR)	Quinonic (XAS)	Aromatic (XAS)	Phenolic (XAS)	Polysaccharide (XAS)
0	0.556	0.222	6.00	0.31	0.66	0.72	1.19
20	0.433	0.232	8.54	0.35	0.65	0.75	1.18
30	0	0.088	1.54	0.24	0.65	0.73	1.22
40	0	0.570	9.54	0.27	0.79	0.72	1.18
50	0.208	0.320	4.24	0.31	0.8	0.78	1.13
60	0.283	0.276	5.27	0.32	0.83	0.74	1.14
80	0.309	0.321	4.34	0.31	0.77	0.67	1.13
α_i	4.40	2.71	0.177	0.011	11.2	0.011	0.011

Table S11. Peak area and α_i values from linear mass balance fit for RCF soils.

Rock Creek Meadow Soils Mass Balance Fit							
Depth	Peak area						
	Carbonyl (IR)	Aliphatic (IR)	Amide (IR)	Quinonic (XAS)	Aromatic (XAS)	Phenolic (XAS)	Polysaccharide (XAS)
0	1.09	0.0233	2.30	0.23	0.55	0.6	1.43
20	0.88	0.0257	1.97	0.22	0.53	0.6	1.43
30	0.82	0	0	0.24	0.51	0.58	1.5
40	0	0.0233	1.95	0.25	0.48	0.53	1.51
50	0	0.0205	0.94	0.3	0.61	0.75	1.41
60	0	0.0032	0.85	0.28	0.51	0.52	1.58
80	0	0.0349	1.71	0.27	0.53	0.5	1.48
α_i	1.25	18.59	0.105	0.010	0.010	0.445	0.010

Table S12. Peak area and α_i values from linear mass balance fit for plants.

Plants Mass Balance Fit							
Sample	Peak area						
	Carbonyl (IR)	Aliphatic (IR)	Amide (IR)	Quinonic (XAS)	Aromatic (XAS)	Phenolic (XAS)	Polysaccharide (XAS)
BCM1 (grass)	1.98	2.757	9.51	0.12	0.48	0.88	1.15
BCM1 (herb)	1.63	1.923	15.63	0.08	0.58	0.53	1.2
BCM2	1.57	2.410	4.50	0.12	0.65	0.35	1.22
BCM3	1.85	2.490	17.69	0.05	0.71	0.27	1.23
BCM4	1.82	2.469	16.99	0.16	0.49	0.36	1.23
BCW (willow)	3.00	4.259	21.89	0.13	0.87	0.4	1.31
RCM1	1.29	1.719	14.24	0.11	0.43	0.57	1.18
RCM2	1.58	1.307	11.61	0.15	0.58	0.56	1.14
RCM3	1.54	1.026	11.66	0.17	0.64	0.49	1.14
RCM4	1.78	1.539	12.30	0.19	0.81	0.62	1.14
RCM5	2.24	2.056	13.53	0.11	0.71	0.49	1.12
RCM6	0.783	1.969	6.85	0.11	0.65	0.4	1.31
α_i	0.011	0.011	0.182	0.011	13.11	0.81	5.79

Table S13. Peak area and α_i values from linear mass balance fit for litter.

Litter Mass Balance Fit							
Sample	Peak area						
	Carbonyl (IR)	Aliphatic (IR)	Amide (IR)	Quinonic (XAS)	Aromatic (XAS)	Phenolic (XAS)	Polysaccharide (XAS)
BCM1	0.79093	0.96018	5.9158	0.18	0.57	0.33	1.12
BCM2	0.51822	0.53189	4.1217	0.23	0.7	0.33	1.1
BCM3	0.99386	0.39068	3.7117	0.21	0.79	0.41	1.21
BCM4	0.54663	0.54618	5.5013	0.23	0.76	0.42	1.12
BCW	1.0636	1.1995	8.6385	0.27	0.71	0.37	1.06
α_i	0.011	1.51	0.011	0.011	0.011	0.011	9.55

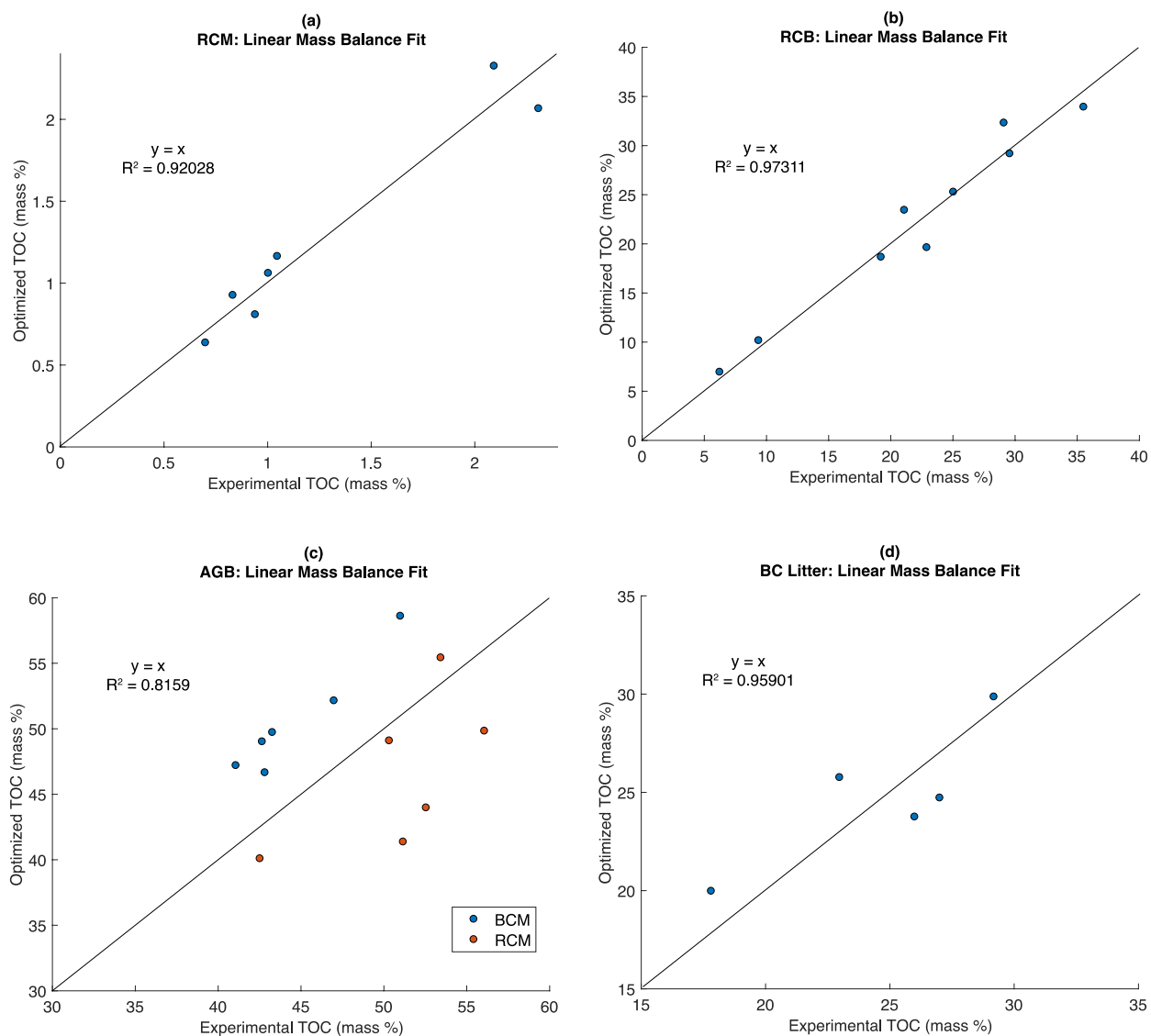


Figure S23. Linear mass balance fit for (a) RCM soils; (b) RCF soils; (c) AGB; and (d) litter.

11. Evaluation of the SOC-fga method relative to single techniques

To statistically compare the calculated SOC mass% between the SOC-fga method and individual techniques, we performed pooled error variance t-test ($p < 0.05$) on all functional groups between (1) FT-IR and bulk C XAS (Table S14), (2) the SOC-fga method vs. FT-IR (Table S15), and (3) the SOC-fga method vs. bulk C XAS (Table S16). The t-test results show that the differences in the calculated mass% for all functional groups between FT-IR and bulk C XAS are significant ($P < 0.05$) except for aromatic C (Table S14). The differences in the calculated mass% are insignificant ($p > 0.05$) for quinonic C between FT-IR and the SOC-fga method (Table S15), and aliphatic, quinonic, phenolic C, and polysaccharide (Table S16), and the differences are significant ($p < 0.05$) in the other functional groups.

Table S14. Pooled error variance t-test for calculated mass% for all functional groups between FT-IR and bulk C XAS.

	Carbonyl	Aliphatic	Amide	Quinonic	Aromatic	Phenolic	Polysaccharide
S_{Res}^2	1.554	0.673	0.667	0	0.667	0.667	0.691
S_{b1-b2}	0.260	0.021	0.398	0	0.849	0.031	0.147
t	8.004	-41.756	-3.061	0	-1.191	-30.673	-8.378
df	24	52	52	52	52	52	52
alpha	0.05	0.05	0.05	0.05	0.05	0.05	0.05
p-value	3.13E-08	1.13E-41	0.003	0	0.239	5.63E-35	3.23E-11
t-crit	2.064	2.007	2.007	2.007	2.007	2.007	2.007
significance	yes	yes	yes	yes	no	yes	yes

Table S15. Pooled error variance t-test for calculated mass% for all functional groups between the SOC-fga method and FT-IR.

	Carbonyl	Aliphatic	Amide	Quinonic	Aromatic	Phenolic	Polysaccharide
S_{Res}^2	0.355	0.878	0.669	0.667	0.667	1.076	0.688
S_{b1-b2}	0.061	0.189	0.114	55.270	0.017	1.654	0.192
t	-3.862	31.469	-9.590	-0.018	-58.646	-5.627	-6.679
df	14	52	52	52	52	52	52
alpha	0.05	0.05	0.05	0.05	0.05	0.05	0.05
p-value	1.73E-03	1.59E-35	4.33E-13	9.86E-01	3.49E-49	7.44E-07	1.61E-08
t-crit	2.145	2.007	2.007	2.007	2.007	2.007	2.007
significance	yes	yes	yes	no	yes	yes	yes

Table S16. Pooled error variance t-test for calculated mass% for all functional groups between the SOC-fga method and bulk C XAS.

	Carbonyl	Aliphatic	Amide	Quinonic	Aromatic	Phenolic	Polysaccharide
S _{Res} ²	1.587	0.677	0.667	0.667	0.667	0.667	0.669
S _{b1-b2}	0.215	0.146	0.114	55.270	0.017	1.025	0.187
t	6.757	-0.002	-7.838	0.002	-58.497	-1.008	1.896
df	24	52	52	52	52	52	52
alpha	0.05	0.05	0.05	0.05	0.05	0.05	0.05
p-value	5.47E-07	9.99E-01	2.30E-10	9.98E-01	3.98E-49	3.18E-01	6.35E-02
t-crit	2.064	2.007	2.007	2.007	2.007	2.007	2.007
significance	yes	no	yes	no	yes	no	no

12. References

1. Tolbert, A., Akinosho, H., Khunsupat, R., Naskar, A. K. & Ragauskas, A. J. Characterization and analysis of the molecular weight of lignin for biorefining studies. *Biofuels Bioprod. Biorefining* **8**, 836–856 (2014).
2. Wildfire effects on soil lipid composition in burnt eucalypt stands, in north-central Portugal. Available at: https://www.researchgate.net/publication/258783439_Wildfire_effects_on_soil_lipid_composition_in_burnt_eucalypt_stands_in_north-central_Portugal. (Accessed: 19th January 2017)
3. Membrane Assumptions. Available at: <http://cr.middlebury.edu/biology/labbook/membranes/frap/membranes/pages/MembrStruct1aside1.htm>. (Accessed: 19th January 2017)
4. Huang, H.-J., Chen, W.-Y. & Wu, J.-H. Total protein extraction for metaproteomics analysis of methane producing biofilm: the effects of detergents. *Int. J. Mol. Sci.* **15**, 10169–10184 (2014).
5. Schnitzer, M. & Khan, S. U. *Soil Organic Matter*. (Elsevier, 1975).
6. Lanzarotta, A. Approximating the detection limit of an infrared spectroscopic imaging microscope operating in an attenuated total reflection (ATR) modality: theoretical and empirical results for an instrument using a linear array detector and a 1.5 millimeter germanium hemisphere internal reflection element. *Appl. Spectrosc.* **69**, 205–214 (2015).
7. Witte, K. *et al.* Magnesium K-Edge NEXAFS Spectroscopy of Chlorophyll a in Solution. *J. Phys. Chem. B* **120**, 11619–11627 (2016).