

Syracuse University

**SURFACE**

---

Civil and Environmental Engineering

College of Engineering and Computer Science

---

4-12-2013

## Compositional Characterization of Soil Organic Matter and Hot-Water Extractable Organic Matter in Organic Horizons Using a Molecular Mixing Model

Ankit Balaria

Chris E. Johnson  
*cejohns@syr.edu*

Follow this and additional works at: <https://surface.syr.edu/cie>



Part of the [Civil and Environmental Engineering Commons](#)

---

### Recommended Citation

Balaria, A., & Johnson, C. E. (2013). Compositional characterization of soil organic matter and hot-water-extractable organic matter in organic horizons using a molecular mixing model. *Journal of Soils and Sediments*, 13(6), 1032–1042. <http://doi.org/10.1007/s11368-013-0690-6>

This Article is brought to you for free and open access by the College of Engineering and Computer Science at SURFACE. It has been accepted for inclusion in Civil and Environmental Engineering by an authorized administrator of SURFACE. For more information, please contact [surface@syr.edu](mailto:surface@syr.edu).

**Compositional characterization of soil organic matter and hot-water extractable organic matter in organic horizons using a molecular mixing model**

**Ankit Balaria • Chris E. Johnson**

A. Balaria • C. E. Johnson (✉)

Department of Civil and Environmental Engineering, 151 Link Hall, Syracuse University,  
Syracuse NY 13244, USA

e-mail: cejohns@syr.edu

(✉) **Corresponding author:**

Chris E. Johnson

Tel. 1 315 443 4425

e-mail: cejohns@syr.edu

**Abstract**

*Purpose* Microbial decomposition of soil organic matter (SOM) is generally believed to be heterogeneous, resulting in the preferential loss of labile compounds such as carbohydrates and proteins and the accumulation of recalcitrant compounds such as lipids and lignin. However, these fractions are difficult to measure directly in soils. We examined patterns in the biomolecular composition of SOM and hot-water-extractable organic matter (HWEOM) by using a molecular mixing model (MMM) to estimate the content of carbohydrates, protein, lipids, and lignin.

*Materials and methods* Organic-horizon soils from Spodosols at the Hubbard Brook Experimental Forest in New Hampshire, USA were analyzed for this study. The MMM uses data from elemental analysis (C,H,N) and <sup>13</sup>C nuclear magnetic resonance (NMR) spectroscopy with cross-polarization and magic-angle spinning (CPMAS) to estimate the percentage of total C in the various classes of biomolecules.

*Results and discussion* Carbohydrate content decreased from about 50% of the C in recent litter to approximately 35% in the bottom of the humus layer. Lipids accounted for about 18% of C in recent litter and increased to 40% in the lower humus layers. The HWEOM fraction of SOM was dominated by carbohydrates (40-70% of C). Carbohydrates and lipids in HWEOM exhibited depth patterns that were the opposite of the SOM. The results from the MMM confirmed the selective decomposition of carbohydrates and the relative accumulation of lipids during humus formation. The depth patterns in HWEOM suggest that the solubility of carbohydrates increases during decomposition, while the solubility of the lipid fraction decreases. The MMM was able to reproduce the spectral properties of SOM and HWEOM very accurately, although there were some discrepancies between the predicted and measured H:C and O:C ratios.

*Conclusions* The MMM approach is an accurate and cost-effective alternative to wet-chemical methods. Together, carbohydrates and proteins account for up to 85% of the C in HWEOM, indicating that the HWEOM fraction represents a labile source of C for microbes. Humification resulted in a decrease in carbohydrate content and an increase in lipids in SOM, consistent with investigations carried out in diverse soil environments.

**Keywords** Carbon • Decomposition • Forest soil • Modeling • Nuclear magnetic resonance spectroscopy • Soil organic matter

## **1 Introduction**

The composition of soil organic matter (SOM) influences numerous chemical and biological properties of the soil, including pH, cation exchange capacity, nutrient transport and immobilization, and microbial activity (Stevenson 1994; Schlesinger 1997). The composition of SOM varies significantly from one location to another due to variations in vegetation type, soil texture, temperature and precipitation, and microbial diversity. A number of studies have attempted to quantify the organic biomolecules present in SOM through degradative wet chemical methods. For example, Baldock et al. (1987) measured carbohydrate content using hydrolysis of SOM with sulfuric acid. Hydrolysis with hydrochloric acid has been used to estimate protein content by measuring the release of amino acids (Beavis and Mott 1996). Lignin present in organic matter has been quantified by CuO oxidation (Reeves 1995) and

tetramethylammonium hydroxide thermochemolysis (Hatcher et al. 1995). Lipids have been estimated by extraction with organic solvents (Heng and Goh 1981; Wu et al. 1995). These studies can then be combined to provide an overall composition of the SOM (Preston et al. 1997). However, it is very time consuming and expensive to carry out all of these individual analyses for routine studies involving large number of samples. More importantly, these analyses have been found to achieve only 20-80% quantification (Ogner 1985; Kögel et al. 1988; Kögel-Knabner et al. 1988), and their chemical degradative steps sometimes result in secondary reactions causing the formation of alteration byproducts, leading to ambiguous results (Preston et al. 1997; Nelson and Baldock 2005).

Nuclear magnetic resonance (NMR) spectroscopy is a well-established tool for investigating the chemical structure of organic compounds. However, its use for analyzing soils has been somewhat limited due to the complex composition of SOM. The most common NMR technique for SOM is  $^{13}\text{C}$  NMR with cross-polarization and magic-angle spinning (CPMAS), which involves determination of broad functional group categories: alkyl C, N-alkyl/methoxyl C, O-alkyl C, di-O-alkyl C, aryl C, O-aryl C, and carbonyl/amide C (e.g., Kögel-Knabner et al. 1988). The spectral intensities associated with these functional group classes can then be interpreted as indicators of organic biomolecules such as carbohydrates, proteins, lignins, and lipids. These interpretations can be supplemented using other, independent methods such as elemental analysis or mass spectroscopy (Leinweber et al. 1995; Landgraf et al. 2006; Kaal et al. 2007; Balaria et al. 2009). While compositional analysis based on NMR provides much-needed information about the nature of organic matter, it does not provide a direct, quantitative measure of the organic biomolecules that form the soil organic matter.

Although NMR does not detect different biomolecule types directly, the NMR spectrum is constrained by the biomolecular composition of the sample. Several recent studies have utilized a molecular mixing model (MMM) to quantify the biomolecules present in marine and soil organic matter using the results of non-degradative solid-state NMR spectroscopy (Nelson et al. 1999; Hedges et al. 2001, 2002; Baldock et al. 2004; Nelson and Baldock 2005; Ahmad et al. 2006; Dickens et al. 2006; Kaal et al. 2007). This model uses the NMR spectral data and C:N ratios for a sample to predict its biomolecular composition in six families of compounds: carbohydrates, proteins, lignins, lipids, char, and carbonyl. The model is a linear deconvolution routine that

estimates the fractions of each biomolecule type that yield a total spectral distribution that best fits the sample spectrum (Nelson and Baldock 2005).

While modeling SOM as six (or fewer) biomolecules is an oversimplification, these molecules are known to be responsible for most of the structural composition of SOM, and inclusion of more molecule types in the model results in a reduction in accuracy (Nelson and Baldock 2005). The MMM is the best tool available at present for predicting biomolecular composition of natural organic matter, and in most cases predicts the biomolecular distribution significantly better than other methods such as mass spectroscopy and wet chemical methods (Nelson and Baldock 2005; Kaal et al. 2007).

In this paper, we investigate the molecular composition of soil organic matter and its presumptively biodegradable fraction: hot-water-extractable organic matter (HWEOM). The HWEOM fraction has been shown to be positively correlated with microbial biomass C (Sparling et al. 1998) and bioavailable organic carbon (Chen et al. 2004). Also, Balaria (2011) observed up to 60% biodegradation of HWEOM in a six-week incubation experiment using soils from the site of this study. In a previous paper we analyzed the variation in the elemental composition and  $^{13}\text{C}$  NMR properties of SOM and HWEOM with soil depth in organic soil horizons in some acidic soils from the northeastern USA (Balaria et al. 2009). In interpreting those results, we hypothesized that the HWEOM contained a higher proportion of carbohydrate and protein structures compared to the SOM from which it was extracted. We also hypothesized that carbohydrates and proteins were depleted with soil depth in O horizons due to their selective utilization by soil microorganisms. Here we test those hypotheses by applying the MMM of Nelson and Baldock (2005) to data from organic horizons in a forested Spodosol.

## **2 Materials and methods**

### **2.1 Sampling and analysis**

This study was conducted on soils from the Hubbard Brook Experimental Forest (HBEF), New Hampshire (43°56' N, 71°45' W). The dominant forest type in the study area is the northern hardwood association, a roughly equal mixture of sugar maple (*Acer saccharum*), American beech (*Fagus grandifolia*) and yellow birch (*Betula alleghaniensis*). At the upper elevations of the study area, the northern hardwoods give way to an assemblage of red spruce (*Picea rubens*), balsam fir (*Abies balsamea*) and paper birch (*Betula papyrifera*). Soils at the HBEF are generally

classified (Soil Survey Staff 2010) as well-drained Spodosols (Typic and Lithic Haplorthods) and Inceptisols (Typic Dystrochrepts) formed on glacial till, and have an average pH of 3.4 to 4.0 in the surface organic horizons (Johnson et al. 1991). These organic horizons consist of a largely undecomposed litter layer (Oi horizon), a partially decomposed, fibrous layer with a pronounced root mat (Oe horizon), and a highly decomposed humic layer (Oa horizon).

At four replicate sites, all of which were underlain by Spodosols, one sample was collected from the Oi horizon and one from the Oe horizon. The Oa horizon was sampled in 1-cm increments to 5 cm depth. The samples were collected by exposing a soil face and cutting back with a flat trowel. The soil samples were stored in sealed plastic bags in a cold room maintained below 5 °C. Moist soil sub-samples were passed through a 5-mm sieve to remove live roots, and dried at 80 °C to constant weight to determine the moisture content. These dried sub-samples were then ground and stored for elemental and NMR analyses.

The procedures used for hot-water extraction of organic matter and for elemental analysis are described in detail in Balaria et al. (2009). Briefly, 150 mL of deionized water was added to 10 g or 30 g (dry equivalent) of Oi or Oe/Oa horizon soil, respectively, in acid-washed Nalgene® high-density polyethylene (HDPE) bottles, which were shaken for 30 minutes using a wrist-action shaker and incubated for 18 hours in a water bath maintained at 70 °C. The extracts were centrifuged at 3500 rpm for 20 minutes and filtered through Fisherbrand glass fiber prefilters, followed by 0.45-µm MF-Millipore mixed cellulose ester membrane filters. The solution was freeze-dried into solid HWEOM for elemental and NMR analyses. Carbon recovery in the freeze-drying process ranged from 99% to 116% (Balaria et al. 2009).

The elemental analysis (C, H, and N) of soils and isolated HWEOM samples was performed using a Carlo Erba EA 1108 elemental analyzer. Sulfur content was also measured on HWEOM samples using the elemental analyzer, while the S content of soil was taken from Mitchell et al. (1989). Oxygen content was determined by the subtraction of C, H, N, and S from total ash-free mass. In order to determine the ash content, soil and HWEOM samples were combusted overnight in a muffle furnace maintained at 500 °C. The resulting ash was white to light-gray, suggesting that organic combustion was complete, or very nearly so.

Solid-state  $^{13}\text{C}$  NMR spectra were obtained using a Bruker AVANCE 300 spectrometer operating at 75 MHz for  $^{13}\text{C}$ . We employed cross-polarization and magic-angle spinning (CPMAS) for signal enhancement using a 1 ms contact time, 3 sec recycle delay, 17.5 ms

acquisition time, and 7000 Hz spinning speed using 7-mm zirconia rotors with Kel-F caps. The spectra were analyzed using 50-Hz Gaussian line broadening and integrated using the chemical shift assignments in Table 1.

## **2.2 Molecular mixing model**

The molecular mixing modeling approach takes the NMR spectral data as input and calculates the fractions of the major biomolecule types present in the sample by deconvolution, based on the spectral properties of these individual biomolecules in key spectral regions (see Table 1). For example, in order to calculate the protein content of SOM, the model uses a standard spectral distribution of soil protein taken from the amino acid fraction in soils presented by Friedel and Scheller (2002). Using their data, the total signal intensity for soil protein is distributed as follows: 35.4% as alkyl C (chemical shift: 0-45 ppm), 22.6% as N-alkyl/methoxyl C (45-60 ppm), 3.5% as O-alkyl C (60-95 ppm), 0% as di-O-alkyl C (95-110 ppm), 8.9% as aryl C (110-145 ppm), 1.3% as O-aryl C (145-165 ppm), and 28.3% as carbonyl/amide C (165-215 ppm). Similarly, the spectral distribution of carbohydrate uses cellulose as a model; the lignin distribution was obtained by taking the average of spectral distributions of spruce and red alder lignin (Wilson 1987); and lipid distributions were taken from Kolattukudy (1980). The model also includes char and carbonyl as additional biomolecular components of natural organic matter. Char is used to account for thermally altered recalcitrant organic material resulting from fire; spectra for red pine wood heated to 300 °C (Baldock and Smernik 2002) are used in the model to represent its spectral distribution. The carbonyl component is included in order to account for carboxylated molecules and for the progressive oxidation that occurs as the other various biomolecules are decomposed.

We entered the integrated spectral data for soil and HWEOM samples into the MMM, which was run using five components: carbohydrates, proteins, lignins, lipids, and carbonyl. While the full MMM is a six-component model, it can be restricted by excluding any undesired components in order to improve accuracy. For example, Hedges et al. (2002) excluded the lignin component from an analysis of marine plankton because these aquatic organisms do not contain lignin. In this case, since there is no evidence that the soils at Hubbard Brook have experienced forest fire, the char component was excluded from our calculations.

The model iteratively adjusts the biomolecular composition until the sum of squared differences between predicted and measured spectral intensities is minimized. The model then uses the theoretical elemental (C, H, O, and N) compositions of the model compounds to estimate a molecular formula for the organic matter sample. The carbohydrate elemental ratio used in the model was calculated using the structure of glucose with a water molecule subtracted to account for polymerization (C:O:H = 1:0.833:1.667) (Baldock et al. 2004). The elemental ratio of protein (C:O:H:N = 1:0.155:1.101:0.27) was calculated from amino acid compositions, based on data in Friedel and Scheller (2002). Lignin stoichiometry was determined by assuming an equal mixture of syringyl and vanillyl lignin (C:O:H = 1:0.429:1.238) (Baldock et al. 2004). The elemental ratio for the lipid fraction (C:O:H = 1:0.235:1.941) was estimated based on cutin composition as an equal mixture of 10,16-dihydroxyhexadecanoic and 9,10,18-trihydroxyoctadecanoic acids (Goni and Hedges 1990). The carbonyl component was simply assumed to have the COOH structure, and hence the C:O:H was 1:2:1.

The MMM was further constrained using measured N:C ratios (from elemental analysis of the samples) in order to provide better predictions for N-containing molecules such as proteins (Kaal et al. 2007). The H:C and O:C ratios predicted by the model were then compared to measured H:C and O:C ratios to assess model performance.

### **2.3 Statistical analyses**

We performed analysis of variance (ANOVA) on each biomolecular component to determine whether significant differences existed between the composition of whole soil and hot-water extractable organic matter, and whether there was a significant increasing or decreasing pattern with soil depth. We also prepared measured vs. predicted H:C and O:C ratio plots to evaluate model performance.

## **3 Results and discussion**

### **3.1 Biomolecular composition by depth and source material**

In a previous paper we described the patterns in the NMR spectral properties of both SOM and HWEOM at Hubbard Brook (Balaria et al. 2009). Briefly, we divided the  $^{13}\text{C}$  NMR spectra into five main functional group classes to make structural interpretations: alkyl C, O-alkyl C, aryl C, O-aryl C, and amide/carbonyl C. We found that while SOM as a whole had substantial amounts



of O-alkyl C and alkyl C, the HWEOM fraction was dominated by O-alkyl C. Aryl C, O-aryl C, and amide/carbonyl C together represented no more than 15% of spectral intensity in SOM and HWEOM. In the organic horizons studied, aryl C ranged between 7-9% and 4-9% for SOM and HWEOM, respectively. O-aryl C ranged from 1.5-4.5% in SOM to 1.5-5% in HWEOM. Carbonyl C content was 5-7% of spectral intensity for the SOM and 8-11% for the HWEOM.

The NMR signal intensity derived from O-alkyl C was found to decrease with the soil depth within the organic horizon for SOM, while this fraction increased in relative importance for the HWEOM. This pattern in O-alkyl C suggested that carbohydrates are preferentially lost during decomposition in the O horizon, while at the same time representing a greater fraction of HWEOM in deeper layers (Balaria et al. 2009). Based on the spectral and elemental data, we speculated that HWEOM represents a mixture of carbohydrates, from plant residues and microbial exudates, and protein, possibly from microbial biomass. Application of the MMM provides further evidence for these observations.

Table 2 summarizes the statistically significant differences in molecular composition between SOM and HWEOM, and their variation patterns with depth. The estimated biomolecular composition was significantly different for SOM and HWEOM for every biomolecular component, even though the HWEOM was extracted from the same soil. Hot-water-extractable organic matter had significantly greater amounts of carbohydrate, protein and carbonyl than the SOM, indicating its richness in labile organic components. At the same time, the soil was significantly richer in the more recalcitrant components such as lignin and lipid moieties. Table 2 also shows the significant depth-related patterns for the various components in SOM and HWEOM. These results are discussed in greater detail in the following sections.

The percentage of each biomolecular component in SOM that was extracted by hot water is reported in Table 3. For all components, the extraction percentage was significantly greater in the litter layer (Oi horizon) than in the other horizons, which reflects the higher overall solubility of organic matter in litter (Table 4). Aside from the Oi horizon, the percentage of carbohydrates and proteins extracted by hot water was significantly greater than the percentage of lignin and lipids extracted (Table 3).

The depth patterns we observed are assumed to reflect changes in organic matter composition during decomposition, with increasing depth in the O horizon reflecting increasing time of decomposition. This stratigraphic model of decomposition assumes that annual litterfall is the

dominant organic matter source to the O horizon. The O horizon is an important rooting zone, however, so root litter within the O horizon may influence depth patterns in organic matter chemistry. Fine root tissues from various tree species have been shown to have higher contents of O-alkyl C (carbohydrates) and lower contents of alkyl C (lipids) than foliar litters from the same species (e.g., Lemma et al. 2007; Mathers et al. 2007). Our sieving procedure removed most of the live root biomass from the samples prior to analysis. The composition of SOM in this study therefore reflects the decomposition products of both surface and root litter. Fahey et al. (2005) estimated that root mortality at the HBEF resulted in an input of  $52 \text{ g C m}^{-2} \text{ yr}^{-1}$  to O horizons. This flux is approximately one-fourth of the annual C input in litterfall ( $210 \text{ g C m}^{-2} \text{ yr}^{-1}$ ), suggesting that root litter is an important, though quantitatively secondary, C source. While our results are consistent with a stratigraphic decomposition model, the influence of root litter cannot be discounted.

### **3.2 Carbohydrates**

The estimated fraction of carbohydrates in SOM decreased from more than 50% of total C in the litter layer (Oi horizon) to 35% at the bottom of the Oa horizon (Fig. 1). For both the HWEOM and SOM, the variability for carbohydrate values within each horizon/layer was very low. These SOM carbohydrate estimates are similar to the approximately 50 % of total C estimated by Kaal et al. (2007) for Spanish peat soils. The decreasing pattern for carbohydrates in SOM is probably due to their preferential consumption by soil microbes, as these are among the most easily degradable biomolecules (Rosenberg et al. 2003). Similar decreases in the modeled carbohydrate fraction with depth have also been observed by Kaal et al. (2007) for peat soils in Spain. Similarly, Baldock et al. (2004) found a decrease in the modeled carbohydrate proportion with increasing duration of decomposition for both forest and agricultural soils.

Carbohydrate molecules represented 40-70% of total C mass in the HWEOM, significantly more than the 30-55% in the SOM. This indicates that carbohydrates are a particularly soluble component of SOM. The increase in the carbohydrate fraction in HWEOM with increasing soil depth suggests that carbohydrate solubility increases as decomposition proceeds in the deeper soil horizons. Table 3 presents the overall percentage of carbohydrate in the soil that was extracted by the hot water. After decreasing from the Oi to the Oe horizon, the percentage of soil carbohydrates extracted by hot water increased with increasing soil depth from the Oe to the

bottom of the Oa horizon. The increase in the carbohydrate fraction in HWEOM with increasing soil depth could be due to physicochemical preservation of labile C through interaction with minerals (Rass et al. 2005; Ekschmitt et al. 2008). Ekschmitt et al. (2008) suggested that biodegradation of particularly labile C, especially by soil bacteria, is disrupted by reduced diffusion due to the formation of hydrophobic zones with increasing soil depth. Rass et al. (2005) suggested that labile fractions of the soil organic matter are more prone to sorption to the mineral phase in deeper soils with increasing content of Fe and Al oxides, protecting them from microbial attack. Although our study was conducted on O-horizon soils, the mineral content of the soils increased with depth, from an average of 139 g kg<sup>-1</sup> in the top 1-cm of the Oa horizon to 394 g kg<sup>-1</sup> at the bottom of the Oa horizon (Table 4). Thus, the mineral surface area available for sorption of labile C forms increases with depth.

### 3.3 Proteins

In both SOM and HWEOM, the estimated protein content was lowest in the Oi horizon and increased in lower organic horizons (Fig. 2). Baldock et al. (2004) also observed an increase in the protein proportion with time of decomposition in forest soils. Kaal et al. (2007) found a slight increase with depth in the protein fraction in upper soil horizons (from ~12% to 15%), similar in magnitude and pattern to these results. Kaal et al. (2007) also found that by constraining the MMM with measured N:C ratios, as we did, modeled protein values were closer to those estimated using other, independent methods.

Although microbial biomass contains abundant proteins, it cannot by itself account for the depth pattern since soil microbial biomass is only about 1% of the total C in the Oe and Oa horizons in Hubbard Brook soils (Bohlen et al. 2001), whereas proteins account for 11-17% of soil C (see Fig. 2a). The relative increase in the protein fraction of SOM may be a reflection of the fact that carbohydrates are consumed at a relatively faster rate than proteins, resulting in the relative accumulation of protein. It may also be due to the production of proteinaceous by-products of microbial activity.

The estimated proportion of protein molecules in HWEOM was significantly greater than in the whole soils (see Table 2; Fig. 2). The higher protein content in HWEOM may be partly the result of selective extraction of microbial biomass and microbial exudates by the hot water. However, microbial biomass carbon can account for a maximum of 40% of the carbon extracted in the

HWEOM (Balaria et al. 2009). Hot water extracts 1-5% of the proteins present in the whole soil from which it was extracted (Table 3). The percentage of proteins extracted from the Oi horizon was significantly greater than in the Oe and Oa horizons, but there were no significant differences among Oe and Oa horizons.

### **3.4 Lignin**

In both SOM and HWEOM, the estimated lignin fraction decreased significantly with soil depth (Fig. 3). This decrease corresponds to the observed decreases in aromatic and phenolic C reported in Balaria et al. (2009). The decrease in the lignin contribution to SOM signifies that although it is a less preferable energy source for microbes than carbohydrate, it nevertheless undergoes substantial decomposition in organic horizons. For these organic horizons, the lignin fraction was between 8% and 20% in SOM, slightly less than Kaal et al. (2007) observed in Spanish peat soils (~ 20-25%). With the exception of the Oi horizon, the lignin fraction was significantly lower in HWEOM than in SOM (see Fig. 3). This suggests that modification of lignin during decomposition reduces its solubility, or that the more soluble lignin monomers and by-products are preferentially decomposed. Hot water extracted an average of 4.0% of the lignin in the Oi horizon, but less than 1% of the lignin in the Oe and Oa horizons (see Table 3), further suggesting that the solubility of lignin decreases during decomposition.

### **3.5 Lipids**

The relative amount of lipids had different patterns with depth for SOM and the HWEOM (Fig. 4). In SOM, the fraction of lipids showed a substantial increase with soil depth from 18% in the Oi horizon to more than 40% in the lower Oa horizon. This pattern is consistent with the increase in alkyl groups observed in Balaria et al. (2009). This increase is most likely due to the selective preservation of highly recalcitrant, lipid-rich compounds (e.g., resins, waxes, suberin, cutin) during the course of organic matter decomposition in soil. Furthermore, the increase in the estimated lipid fraction can largely account for the estimated decreases in both carbohydrate and lignin in SOM (see Figs. 1 and 3). This pattern is also in agreement with the increased lipid percentage observed with depth by Kaal et al. (2007) and with decomposition time by Baldock et al. (2004).

In contrast, we observed a significant decrease in the lipid fraction with depth for HWEOM. Lipids are not very soluble in water, as shown by the low percentages we observed in HWEOM (see Fig. 4). In the lowest Oa layers lipids accounted for more than 35% of soil C, but only 4-5% of HWEOM C. Moreover, hot water extracted less than 0.5% of the lipids in SOM in the Oe and Oa horizons (Table 3). Our data suggest that as decomposition proceeds, the least soluble, most recalcitrant lipids are retained in the SOM, resulting in a decrease in the fraction of lipids in HWEOM.

### **3.6 Carbonyl**

The carbonyl component exhibited great variability and no specific depth pattern for SOM (Fig. 5). However, it accounted for less than 3% of C in all depths and horizons considered. Ussiri and Johnson (2003) estimated the carboxyl (COOH) fractions in Oa horizons for humic and fulvic acids, with an average value of 3.7% of total C in Hubbard Brook soils. These estimates are similar to the model predictions in this study, corroborating the accuracy of this methodology for carbonyl predictions. Low concentrations and a lack of pattern with soil depth of this carbonyl component were also observed by Kaal et al. (2007) for peat organic matter. This carbonyl fraction represents carboxylated molecules and decomposing biomolecules undergoing progressive oxidation. Thus, unlike the other modeled fractions, they do not represent a specific biomolecular family.

The carbonyl percentage was significantly higher for HWEOM (~3-11%) and had a significant decreasing pattern with depth (see Fig. 5). This elevated contribution to HWEOM indicates that these carboxylated molecules are highly soluble. It must be noted however that this fraction had the highest variability among individual samples for SOM, with the model estimating its percentage to be zero for some of the samples. This variability was also great in the estimation of the carbonyl fraction in HWEOM (see Table 3), which ranged from 3.4% to 26.5%.

### **3.7 Model performance**

The performance of the MMM was evaluated by comparing measured and model-predicted parameters such as spectral intensities and elemental ratios. The observed and predicted average integrated spectral intensities for SOM and HWEOM are presented in Figs. 6a and 6b respectively. These values were averaged over all horizons and replicates. For both the SOM and

the HWEOM, the MMM predicted spectral intensities that were almost identical to the measured values, suggesting that the model explains these spectra very well.

The MMM also computes molar H:C and O:C ratios based on the distribution and elemental composition of the model biomolecules. The predicted H:C and O:C values are compared to the measured H:C and O:C values in Figs. 7 and 8, respectively. The model overestimated the H:C ratio substantially for SOM in all layers (see Fig. 7). However, these predicted ratios fell in a very narrow range of approximately 1.4-1.7. The H:C ratios are generally higher for lipids and carbohydrates. This discrepancy therefore suggests that the lipid and carbohydrate percentage in whole soils may have been somewhat overestimated for the soil and slightly underestimated for some of the HWEOM samples. However, carbohydrates also have high O:C ratios and thus it is necessary to examine the O:C ratios as well.

The model did a much better job predicting O:C ratios in organic matter than H:C ratios. Nevertheless, the O:C ratios were slightly under-predicted for both SOM and HWEOM (see Fig. 8). These ratios depend heavily on O-containing functional groups generally present in carbohydrates and carbonyl structures. The fact that O:C ratios were generally underpredicted suggests that the organic matter may have had a greater fraction of carbohydrates, carbonyl, or both, than the MMM predictions. Alternatively, the elemental ratios for the model compounds, calculated using literature-based elemental compositions, may not be ideal for the Hubbard Brook soil system. It must also be noted that the measured O concentration was estimated by taking the difference of H, C, N, and S from total ash-free weight. While this method provides very similar values to direct measurement (Malcolm and MacCarthy 1986), it is almost certainly an overestimate because it fails to account for elements such as P and metals (Rice and MacCarthy 1991). Thus, the measured O:C ratios in Fig. 8 are slightly overestimated, causing the points to be shifted slightly to the right of the 1:1 line.

The fact that the model overpredicted the H:C ratio in SOM while the predicted O:C ratios were close to the measured ratios suggests that the discrepancy is primarily related to hydrocarbon structures. The presence of condensed aromatic rings in SOM would result in a lower measured H:C ratio, and therefore overprediction of the ratio by the model. However, aryl C only accounted for 7-10% of the spectral intensity of our SOM samples (Balaria et al. 2009). Even if all of the aryl C was in the form of phenanthrene-type structures (H:C = 0.71) as opposed to benzene (H:C = 1.0), the effect on the overall H:C ratio of the SOM would only be on the order

of 0.02 to 0.03 – not enough to explain the discrepancy in Figure 7. A more plausible explanation is undersaturation of alkyl C structures. A saturated aliphatic chain of infinite length would have an H:C ratio of 2.0. If one-third of the C-C bonds in such a chain were double bonds, the H:C ratio would be 1.5. With 25-45% of the SOM spectral intensity in the alkyl C region (Balaria et al. 2009), undersaturation of aliphatic C chains could lower the measured H:C ratio by an amount on the order of 0.2, which would explain much of the discrepancy in Figure 7.

By using a relatively small number of model compounds, the MMM provides a somewhat simplified picture of the biochemical composition of natural organic matter. For example, chitin is an important component of fungal cell walls and exoskeletons of soil arthropods (Nelson et al. 1999). Soils at Hubbard Brook contain colonies of arbuscular mycorrhizal fungi, especially associated with sugar maple trees, as well as active arthropod communities (Fisk et al. 2006; Juice et al. 2006). Thus, it is likely that chitin is a significant component of SOM at Hubbard Brook (Balaria et al. 2009). However, the  $^{13}\text{C}$  NMR spectral properties and the N:C, H:C and O:C ratios of chitin are indistinguishable from a mixture of amino acid and carbohydrate (Nelson et al. 1999; Balaria et al. 2009). Therefore, it is likely that some of the C in SOM and HWEOM that has been attributed to proteins and carbohydrates is actually in chitin.

#### **4 Conclusions**

Application of the MMM revealed significant depth patterns in the composition of SOM and HWEOM, as well as significant differences between the two. Approximately 50% of the C in SOM of the uppermost Oi soil horizon was estimated to be in carbohydrates. This fraction decreased to approximately 35% in the bottom of the Oa horizon, indicating the selective decomposition of this labile C fraction. The lipid fraction, which includes highly recalcitrant and hydrophobic substances such as waxes, resins, suberin and cutin, increased from about 18% of C in the Oi horizon to more than 40% in the lower Oa layers. The relative accumulation of this recalcitrant C is consistent with other NMR-based studies of SOM decomposition. The HWEOM fraction of SOM was dominated by carbohydrates (40-70% of C). Carbohydrates and lipids in HWEOM exhibited depth patterns that were the opposite of the whole soils, suggesting that the solubility of carbohydrates increases during decomposition, while the solubility of the lipid fraction decreases.

The quantitative information provided by the MMM is highly useful for the chemical analysis of any soil since it is fairly accurate, fast, and cost-effective compared to usual wet-chemical methods. Nevertheless, the discrepancies in the H:C ratios suggest that while the MMM is a great tool to investigate biomolecular composition of complex organic mixtures such as SOM and HWEOM, more work is needed to improve the selection of model compounds, especially for lipids and lignin.

**Acknowledgments** This work was funded by the US Department of Agriculture National Research Initiative Competitive Grants program (award no. 2005-35107-16200). Support was also provided by the National Science Foundation, through the Long-Term Ecological Research program (grant no. 1114804). Jeff Baldock and Ron Smernik provided valuable advice on the application of the molecular mixing model and interpretation of the results. Ankit Balaria was supported by a W.L. Li Fellowship while preparing this manuscript. The comments of two anonymous reviewers greatly improved the paper. This is a contribution of the Hubbard Brook Ecosystem Study. The Hubbard Brook Experimental Forest is operated by the Northern Research Station of the USDA Forest Service, Newtown Square, PA.

## References

- Ahmad R, Nelson PN, Kookana RS (2006) The molecular composition of soil organic matter as determined by  $^{13}\text{C}$  NMR and elemental analyses and correlation with pesticide sorption. *Eur J Soil Sci* 57:883-893
- Balaria A (2011) Effects of calcium addition on structure and bioavailability of soil organic matter. Dissertation, Syracuse University, Syracuse, NY, 228 pp.
- Balaria A, Johnson CE, Xu Z (2009) Molecular-scale characterization of hot-water extractable organic matter in a northeastern forest soil. *Soil Sci Soc Am J* 73:812-821
- Baldock JA, Kay BD, Schnitzer M (1987) Influence of cropping treatments on the monosaccharide content of the hydrolyzates of a soil and its aggregate fractions. *Can J Soil Sci* 67:489-499
- Baldock JA, Masiello CA, Gélinas Y, Hedges JI (2004) Cycling and composition of organic matter in terrestrial and marine ecosystems. *Marine Chem* 92:39-64



- Baldock JA, Smernik RJ (2002) Chemical composition and bioavailability of thermally altered *Pinus resinosa* (red pine) wood. *Org. Geochem.* 33:1093–1109
- Beavis J, Mott CJB (1996) Effects of land use on the amino acid composition of soils: 2. Soils from the park grass experiment and broadband wilderness, Rothamsted, England. *Geoderma* 91:173-190
- Bohlen PJ, Groffman PG, Driscoll CT, Fahey TJ, Siccama TG (2001) Plant-soil-microbial interactions in a northern hardwood forest. *Ecology* 82:965-978
- Chen CR, Xu ZH, Mathers NJ (2004) Soil carbon pools in adjacent natural and plantation forests of subtropical Australia. *Soil Sci Soc Am J* 68:282-291
- Dickens AF, Baldock JA, Smernik RJ, Wakeham SG, Arnarson TS, Gélinas Y, Hedges JI (2006) Solid-state  $^{13}\text{C}$  NMR analysis of size and density fractions of marine sediments: Insight into organic carbon sources and preservation mechanism. *Geochim Cosmochim Acta* 70:666-686
- Ekschmitt K, Kandeler E, Poll C, Brune A, Buscot F, Friedrich M, Gleixner G, Hartmann A, Kästner M, Marhan S, Miltner A, Scheu S, Wolters V (2008) Soil-carbon preservation through habitat constraints and biological limitations on decomposer activity. *J Plant Nutr Soil Sci* 171:27–35
- Fahey TJ, Siccama TG, Driscoll CT, Likens GE, Campbell J, Johnson CE, Battles JJ, Aber JD, Cole JJ, Fisk MC, Groffman PG, Hamburg SP, Holmes RT, Schwarz PA, Yanai RD (2005) The biogeochemistry of carbon at Hubbard Brook. *Biogeochemistry* 75:109-176
- Fisk MC, Kessler WR, Goodale CA, Fahey TJ, Groffman PM, Driscoll CT (2006) Landscape variation in microarthropod response to calcium addition in a northern hardwood forest. *Pedobiologia* 50:69-78
- Friedel JK, Scheller E (2002) Composition of hydrolysable amino acids in soil organic matter and soil microbial biomass. *Soil Biol Biochem* 34:315-325
- Goni MA, Hedges JI (1990) Cutin-derived cupric oxide reaction products from purified cuticles and tree leaves. *Geochim Cosmochim Acta* 54:3065-3072
- Hatcher PG, Nanny MA, Minard RD, Dible SD, Carson DM (1995) Comparison of two thermochemolytic methods for the analysis of lignin in decomposing gymnosperm wood: the CuO oxidation method and the method of thermochemolysis with tetramethylammonium hydroxide (TMAH). *Org Geochem* 23:881-888

- Hedges JI, Baldock JA, Gélinas Y, Lee C, Peterson M, Wakeham SG (2001) Evidence for non-selective preservation of organic matter in sinking marine particles. *Nature* 409:801-804
- Hedges JI, Baldock JA, Gélinas Y, Lee C, Peterson ML, Wakeham SG (2002) The biochemical and elemental compositions of marine plankton: a NMR perspective. *Marine Chem* 78:47–63
- Heng S, Goh KM (1981) A rapid method for extracting lipid components from forest litter especially adapted for ecological studies. *Commun. Soil Sci Plant Anal* 12:1283-1292
- Johnson CE, Johnson AH, Huntington TG, Siccama TG (1991) Whole-tree clearcutting effects on soil horizons and organic matter pools. *Soil Sci Soc Am J* 55:497-502
- Juice SM, Fahey TJ, Siccama TG, Driscoll CT, Denny EG, Eager C, Cleavitt NL, Minocha R, Richardson AD (2006) Response of sugar maple to calcium addition to northern hardwood forest. *Ecology* 87:1267-1280
- Kaal J, Baldock JA, Buurman P, Nierop KGJ, Pontevedra-Pombal X, Martinez-Cortizas A (2007) Evaluating pyrolysis–GC/MS and <sup>13</sup>C CPMAS NMR in conjunction with a molecular mixing model of the Penido Vello peat deposit, NW Spain. *Org Geochem* 38:1097-1111
- Kögel I, Hempfling R, Zech W, Hatcher PG, Schulten H-R (1988) Chemical composition of the organic matter in forest soils: 1. Forest litter. *Soil Sci* 146:124-136
- Kögel-Knabner I, Zech W, Hatcher PG (1988) Chemical composition of the organic matter in forest soils: the humus layer. *Zeitschrift für Pflanzenernährung und Bodenkunde* 151:331-340
- Kolattukudy PE (1980) Cutin, suberin, and waxes. *Biochem Plants* 4:571-645
- Landgraf D, Leinweber P, Makeschin F (2006) Cold and hot water-extractable organic matter as indicators of litter decomposition in forest soils. *J Plant Nutr Soil Sci* 169:76–82
- Leinweber P, Schulten H–R, Koerschens M (1995) Hot-water extracted organic matter: Chemical composition and temporal variations in a long-term field experiment. *Biol Fert Soils* 20:17-23
- Lemma B, Nilsson I, Kleja DB, Olsson M, Knicker H (2007) Decomposition and substrate quality of leaf litters and fine roots from three exotic plantations and a native forest on the southwestern highlands of Ethiopia. *Soil Biol Biochem* 39:2317-2328

- Malcolm RL, MacCarthy P (1986) Limitations in the use of commercial humic acids in soil and water research. *Environ Sci Technol* 20:904–911
- Mathers NJ, Jalota RK, Dalal RC, Boyd SE (2007)  $^{13}\text{C}$  NMR analysis of decomposing litter and fine roots in the semi-arid Mulga Lands of southern Queensland. *Soil Biol Biochem* 39:993-1006
- Mitchell MJ, Driscoll CT, Fuller RD, David MB, Likens GE (1989) Effect of whole-tree harvesting on the sulfur dynamics of a forest soil. *Soil Sci Soc Am J* 53:933–940
- Nelson PN, Baldock JA (2005) Estimating the molecular composition of a diverse range of natural organic materials from solid-state  $^{13}\text{C}$  NMR and elemental analyses. *Biogeochemistry* 72:1-34
- Nelson PN, Baldock JA, Clarke P, Oades JM, Churchman GJ (1999) Dispersed clay and organic matter in soil: their nature and association. *Aust J Soil Res* 37:289-315.
- Ogner, G (1985) A comparison of four different raw humus types in Norway using chemical degradations and CPMAS  $^{13}\text{C}$  NMR spectroscopy. *Geoderma* 35:343-353
- Preston CM, Trofymow JA, Sayer BG, Niu J (1997)  $^{13}\text{C}$  nuclear magnetic resonance spectroscopy with cross-polarization and magic-angle spinning investigation of the proximate analysis fractions used to assess litter quality in decomposition studies. *Can J Botany* 75:1601-1613
- Rass DP, Rumpel C, Dignac M (2005) Is soil carbon mostly root carbon? Mechanisms for a specific stabilization. *Plant Soil* 269:341–356
- Reeves AD (1995) The use of organic markers in the differentiation of organic inputs to aquatic systems. *Phys Chem Earth* 20:133-140
- Rice JA, MacCarthy P (1991) Statistical evaluation of the elemental composition of humic substances. *Org Geochem* 17:635–648
- Rosenberg W, Nierop KGJ, Knicker H, de Jager PA, Kreutzer K, Weib T (2003) Liming effects on the chemical composition of the organic surface layer of a mature Norway spruce stand (*Picea abies* [L.] Karst.). *Soil Biol Biochem* 35:155-165
- Schlesinger WH (1997) *Biogeochemistry: An analysis of global change*, 2nd Ed. Academic Press, San Diego, California, USA
- Soil Survey Staff (2010) *Keys to soil taxonomy*, 11th ed. USDA-Natural Resources Conservation Service, Washington, DC, 338 pp.

- Sparling GM, Vojvodic-Vukovic M, Schipper LA (1998) Hot-water-soluble C as a simple measure of labile soil organic matter: the relationship with microbial biomass C. *Soil Biol Biochem* 30:1469-1472
- Stevenson FJ (1994) *Humus chemistry: genesis, composition, reactions*, 2nd Ed. John Wiley & Sons, New York
- Ussiri DAN, Johnson CE (2003) Characterization of organic matter in a northern hardwood forest soil by  $^{13}\text{C}$  NMR spectroscopy and chemical methods. *Geoderma* 111:123-149
- Wilson MA (1987) *NMR techniques and applications in geochemistry and soil chemistry*. Pergamon Press, Oxford, England
- Wu Q-L, Schleuss U, Blume HP (1995) Investigation on soil lipid extraction with different organic solvents. *Zeitschrift für Pflanzenernährung und Bodenkunde* 158:347–350

**Table 1** Distribution of  $^{13}\text{C}$  NMR spectral intensity among model components for processed terrestrial organic matter (Nelson and Baldock 2005)

<b>Chemical Shift</b>	<b>Functional Group</b>	<b>Carbohydrate</b>	<b>Protein</b>	<b>Lignin</b>	<b>Lipid</b>	<b>Carbonyl</b>	<b>Char</b>
ppm		%	%	%	%	%	%
0 - 45	Alkyl C	0	35.4	10.5	75.6	0	0
45 - 60	N-Alkyl/Methoxyl C	0	22.6	13.8	4.5	0	1.7
60 - 95	O-Alkyl C	83.3	3.5	12.5	9	0	1.8
95 - 110	Di-O-Alkyl C	16.7	0	8.6	0	0	5.3
110 - 145	Aryl C	0	8.9	30.6	3.6	0	72.1
145 - 165	O-Aryl C	0	1.3	19.5	0.7	0	15.2
165 - 215	Amide/Carboxyl C	0	28.3	4.6	6.6	100	3.9

**Table 2** Analysis of variance results for differences in molecular composition between SOM and HWEOM, and for significance of depth-related changes for SOM and HWEOM

Molecular Component	SOM (N=28) vs. HWEOM (N=27)		Variation with increasing depth/horizon (N=4)			
			SOM		HWEOM	
	Result	p-value	Pattern	p-value	Pattern	p-value
Carbohydrate	HWEOM>SOM	<0.001	Decreasing	<0.001	Increasing	<0.001
Protein	HWEOM>SOM	<0.001	Increasing	0.046	Inconclusive	0.094
Lignin	SOM>HWEOM	0.007	Decreasing	0.013	Decreasing	<0.001
Lipid	SOM>HWEOM	<0.001	Increasing	<0.001	Decreasing	0.013
Carbonyl	HWEOM>SOM	<0.001	Inconclusive	0.894	Decreasing	0.007

**Table 3** Estimated hot-water extraction percentages of major molecular components in organic-horizon soils at Hubbard Brook.

Values in the table represent the percentage of each component in the soil that was extracted in the corresponding HWEOM sample.

The uncertainties ( $\pm$ ) shown are standard errors (N=4, except for Oa:2-3 cm, for which N=3)

<b>Molecular Component</b>	<b>Oi</b>	<b>Oe</b>	<b>Oa (0-1 cm)</b>	<b>Oa (1-2 cm)</b>	<b>Oa (2-3 cm)</b>	<b>Oa (3-4 cm)</b>	<b>Oa (4-5 cm)</b>
	%	%	%	%	%	%	%
Carbohydrate	3.0 $\pm$ 0.5	1.3 $\pm$ 0.08	1.5 $\pm$ 0.2	1.5 $\pm$ 0.1	1.6 $\pm$ 0.07	1.8 $\pm$ 0.3	1.8 $\pm$ 0.1
Protein	4.6 $\pm$ 1.0	1.5 $\pm$ 0.2	1.6 $\pm$ 0.2	1.4 $\pm$ 0.2	1.2 $\pm$ 0.1	1.3 $\pm$ 0.2	1.1 $\pm$ 0.1
Lignin	4.0 $\pm$ 1.0	0.6 $\pm$ 0.1	0.7 $\pm$ 0.3	0.6 $\pm$ 0.07	0.5 $\pm$ 0.07	0.9 $\pm$ 0.2	1.1 $\pm$ 0.5
Lipid	1.4 $\pm$ 0.2	0.4 $\pm$ 0.04	0.4 $\pm$ 0.08	0.2 $\pm$ 0.02	0.1 $\pm$ 0.01	0.1 $\pm$ 0.01	0.1 $\pm$ 0.01

**Table 4** Mean carbon, nitrogen and ash content in soils and HWEOM at the Hubbard Brook Experimental Forest, NH. The uncertainties ( $\pm$ ) shown are standard errors (N=4, except for Oa:2-3 cm, for which N=3)

Horizon/Layer	Hot-water Extracts	Soil			HWEOM		
	C Yield	Ash	C	N	Ash	C	N
	% of Soil C	g kg <sup>-1</sup>	g kg <sup>-1</sup>	g kg <sup>-1</sup>	g kg <sup>-1</sup>	g kg <sup>-1</sup>	g kg <sup>-1</sup>
Oi	3.4	61 $\pm$ 15	503 $\pm$ 12	22 $\pm$ 2	115 $\pm$ 11	427 $\pm$ 5	26 $\pm$ 1
Oe	1.1	90 $\pm$ 23	494 $\pm$ 12	26 $\pm$ 2	124 $\pm$ 29	408 $\pm$ 14	33 $\pm$ 2
Oa (0-1 cm)	1.1	139 $\pm$ 50	475 $\pm$ 25	25 $\pm$ 2	146 $\pm$ 29	408 $\pm$ 5	33 $\pm$ 2
Oa (1-2 cm)	0.97	175 $\pm$ 49	438 $\pm$ 30	25 $\pm$ 2	94 $\pm$ 10	415 $\pm$ 5	37 $\pm$ 4
Oa (2-3 cm)	0.95	267 $\pm$ 70	402 $\pm$ 38	23 $\pm$ 1	91 $\pm$ 8	416 $\pm$ 9	35 $\pm$ 1
Oa (3-4 cm)	0.98	292 $\pm$ 75	387 $\pm$ 46	22 $\pm$ 2	81 $\pm$ 5	413 $\pm$ 7	34 $\pm$ 0.8
Oa (4-5 cm)	0.93	394 $\pm$ 72	329 $\pm$ 40	18 $\pm$ 2	78 $\pm$ 4	423 $\pm$ 2	34 $\pm$ 0.9



## Figure Captions

**Fig. 1** Variation in percentage of carbon present as carbohydrates with soil depth for soil organic matter and hot-water extractable organic matter (HWEOM). The error bars represent standard errors.

**Fig. 2** Variation in the percentage of carbon present as proteins with soil depth for soil organic matter and hot-water extractable organic matter (HWEOM). The error bars represent standard errors.

**Fig. 3** Variation in the percentage of carbon present as lignin with soil depth for soil organic matter and hot-water extractable organic matter (HWEOM). The error bars represent standard errors.

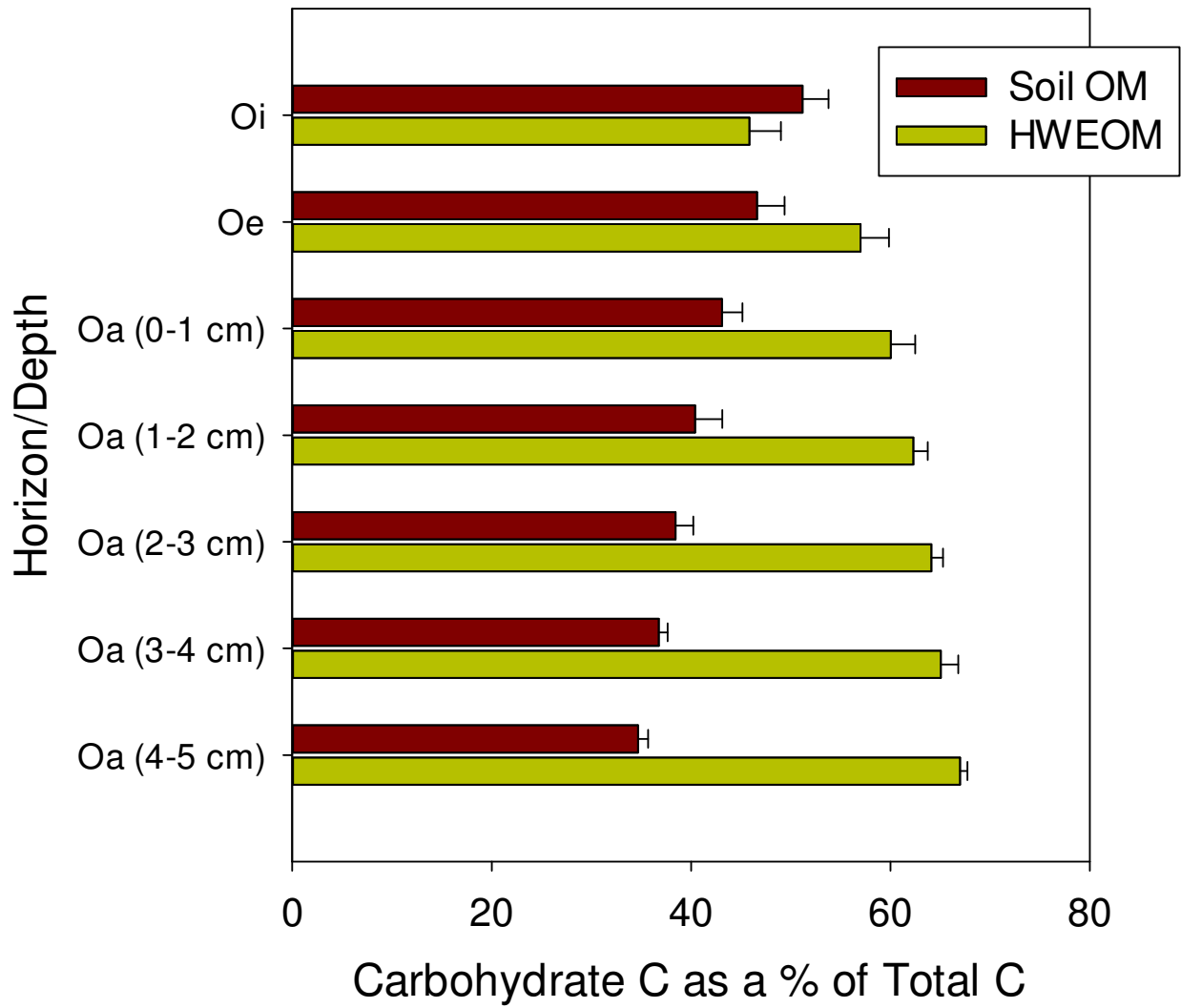
**Fig. 4** Variation in the percentage of carbon present as lipids with soil depth for soil organic matter and hot-water extractable organic matter (HWEOM). The error bars represent standard errors.

**Fig. 5** Variation in percentage of carbon present as carbonyl with soil depth for soil organic matter and hot-water extractable organic matter (HWEOM). The error bars represent standard errors.

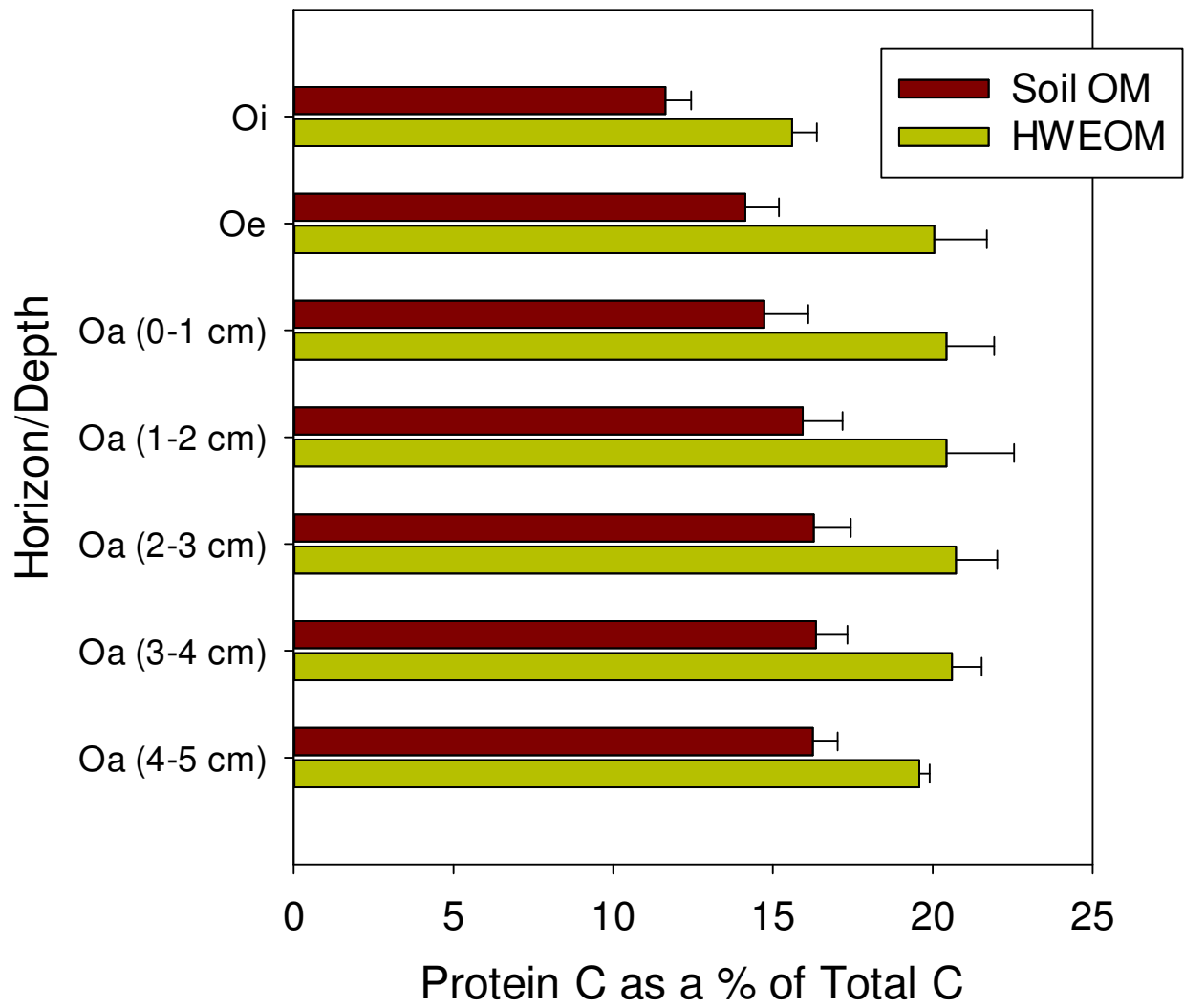
**Fig. 6** Measured vs. predicted spectral intensities for (a) soil organic matter and (b) hot-water extractable organic matter (HWEOM). The error bars represent standard errors (N=27 for HWEOM, 28 for whole soils).

**Fig. 7** Measured vs. predicted H:C ratios for soil organic matter and hot-water extractable organic matter (HWEOM). The data points are compared with a 1:1 line, representing equality between predicted and measured data.

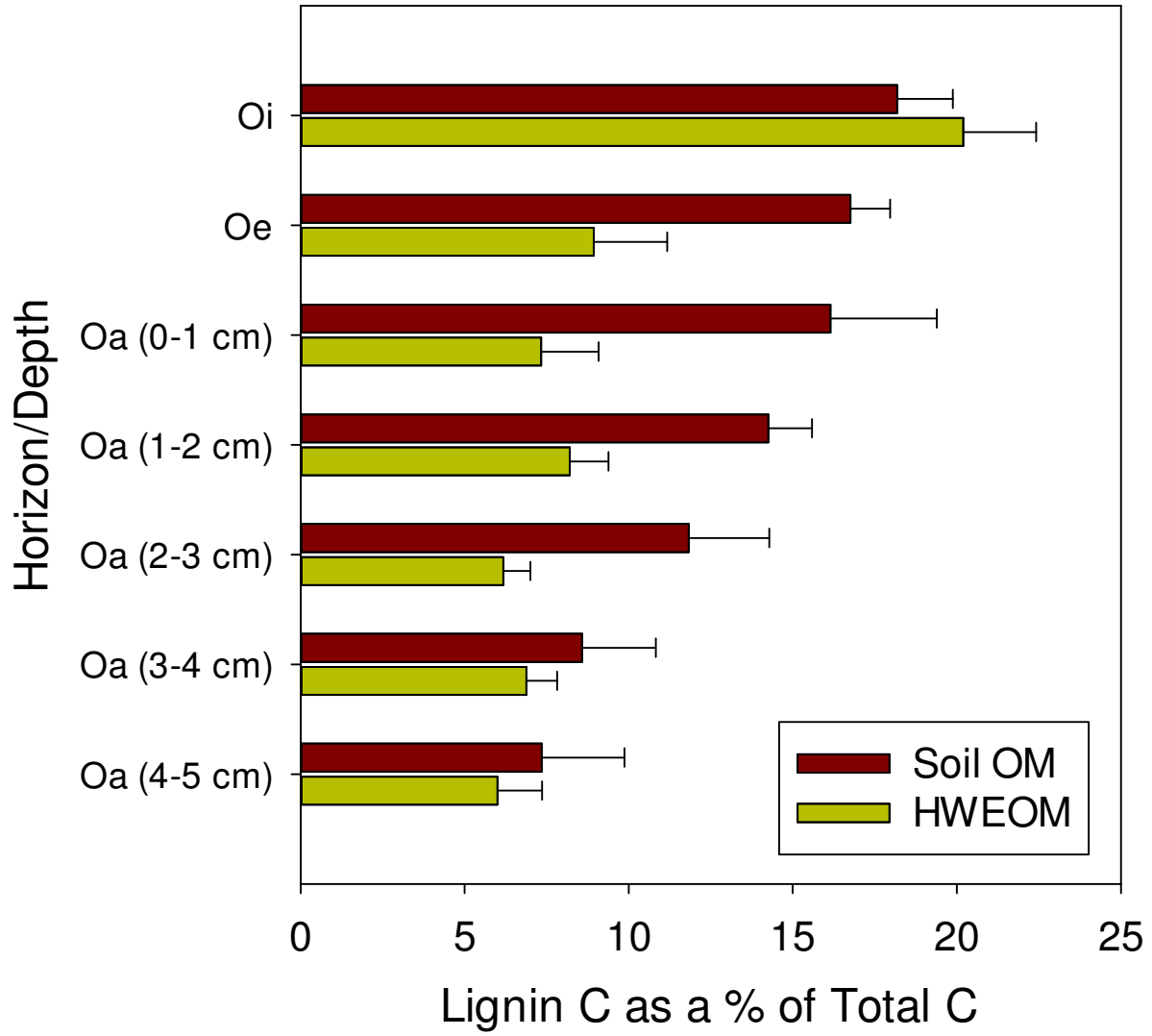
**Fig. 8** Measured vs. predicted O:C ratios for soil organic matter and hot-water extractable organic matter (HWEOM). The data points are compared with a 1:1 line, representing equality between predicted and measured data.



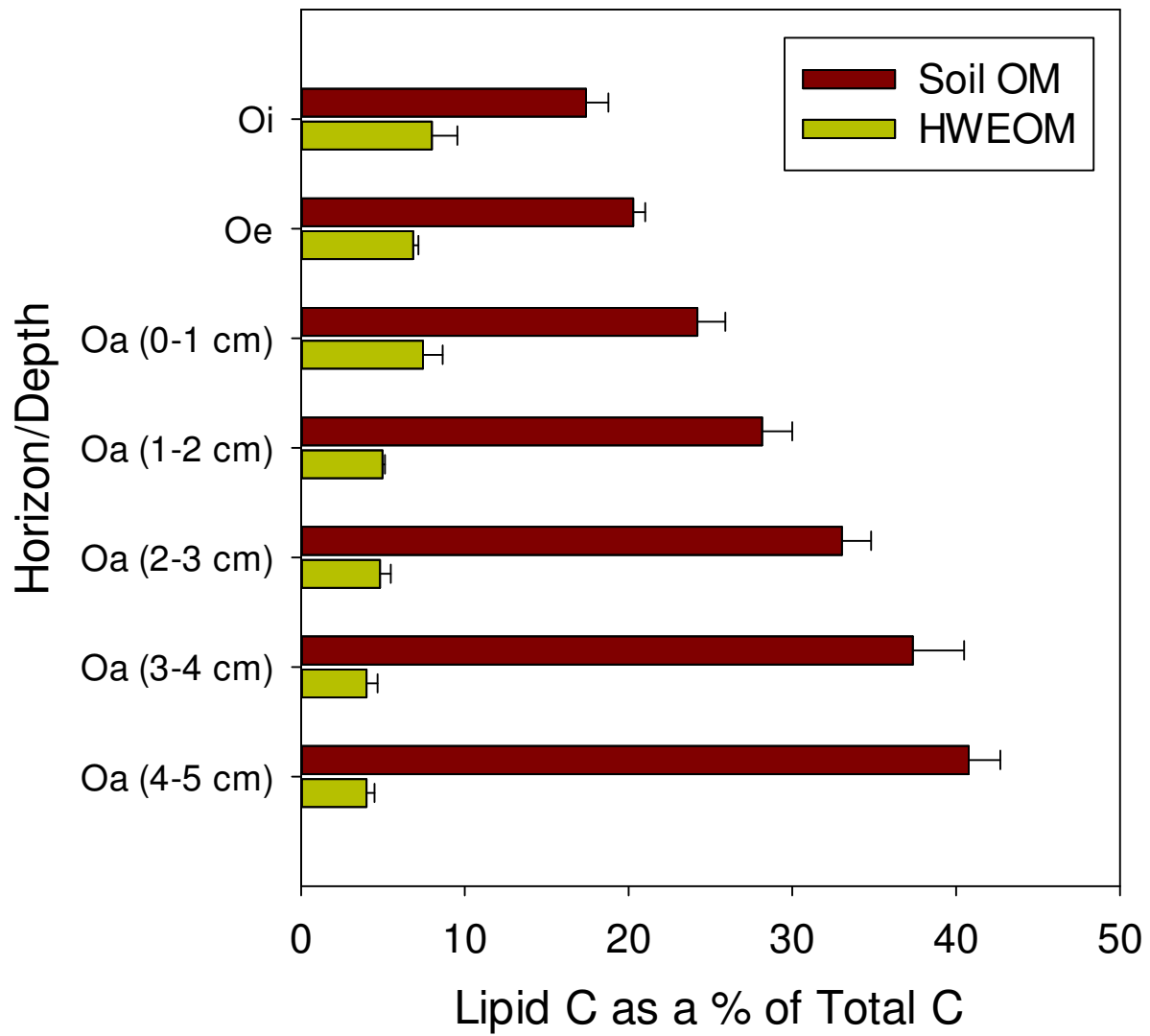
**Figure 1**



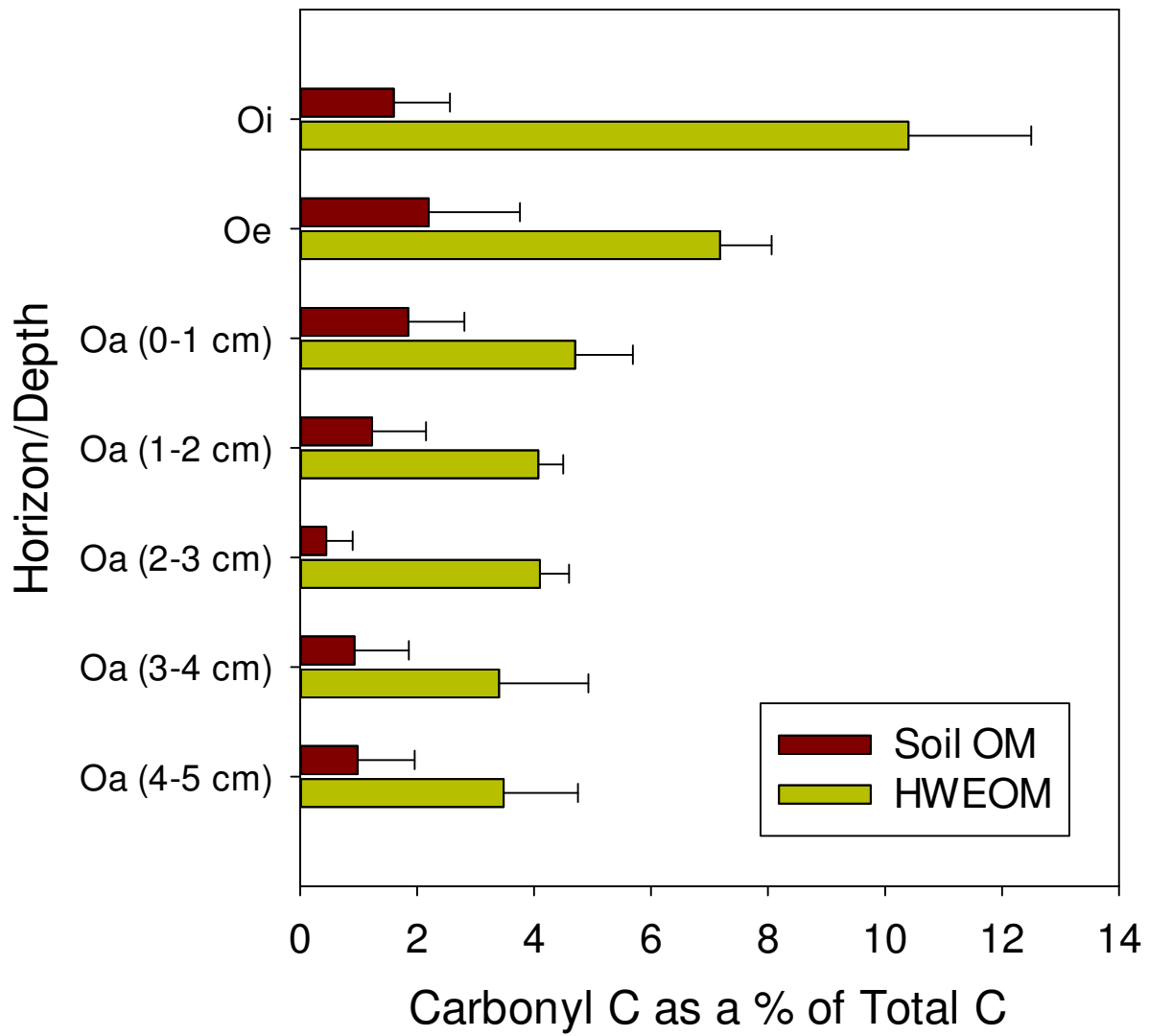
**Figure 2**



**Figure 3**



**Figure 4**



**Figure 5**

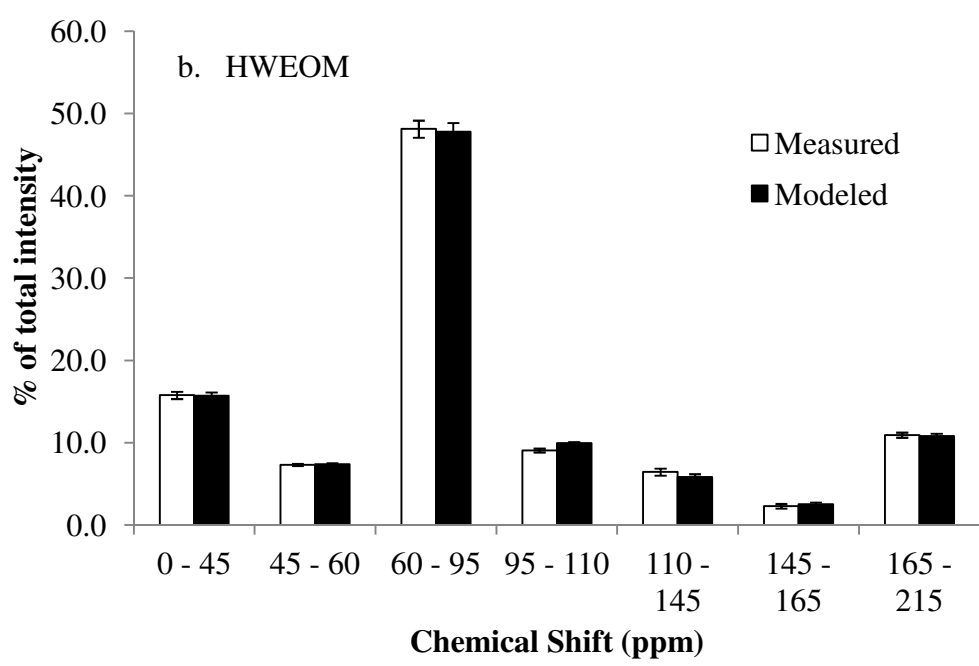
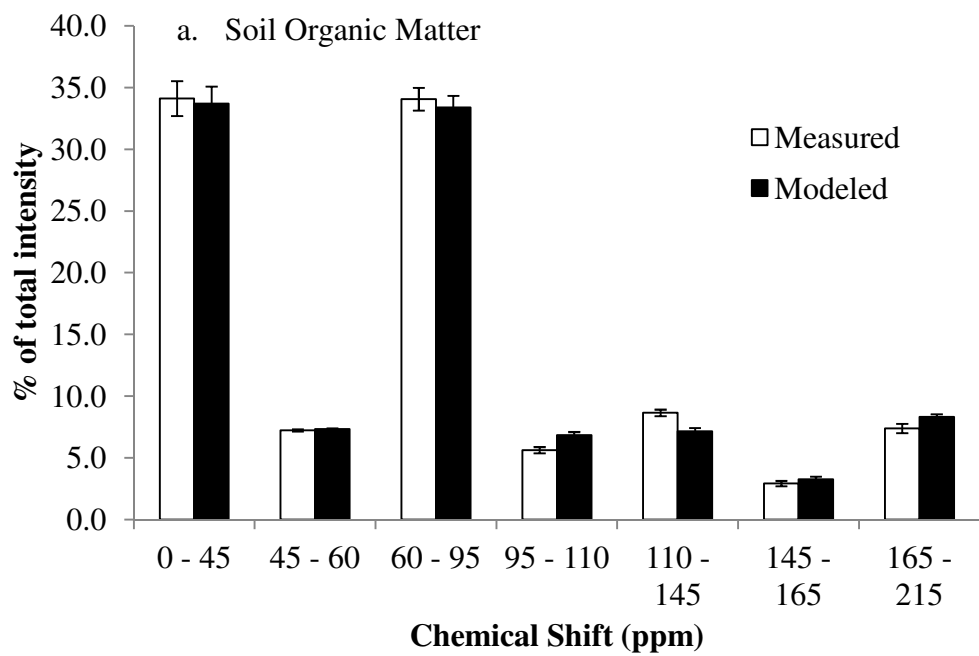
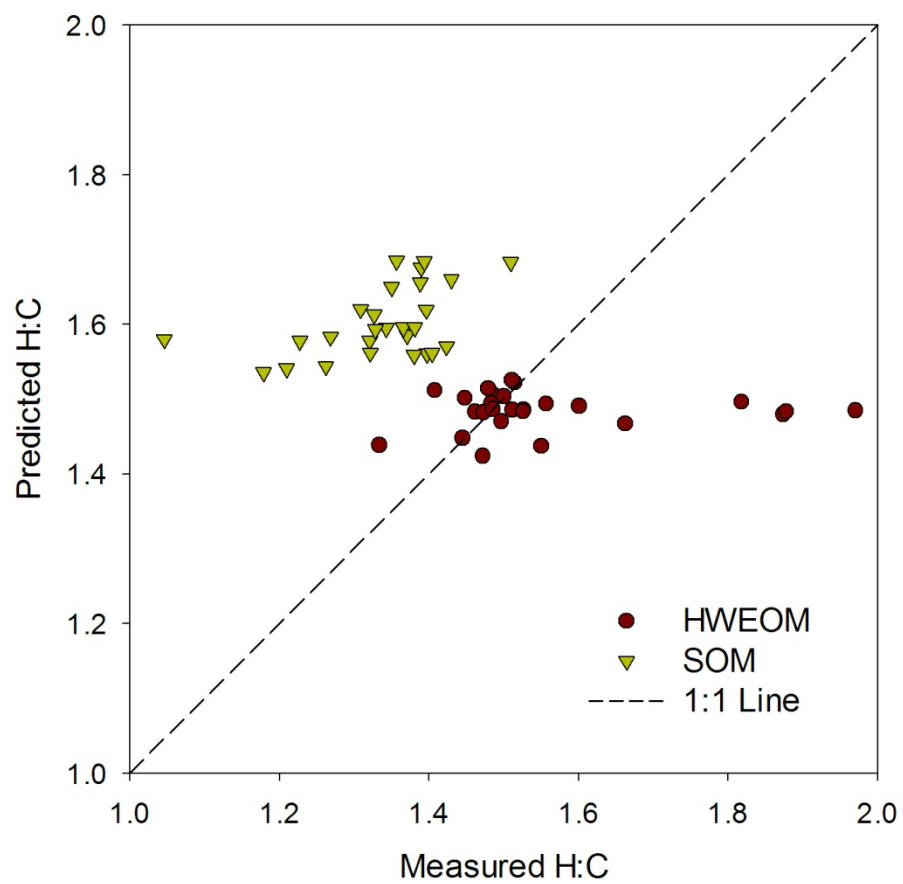
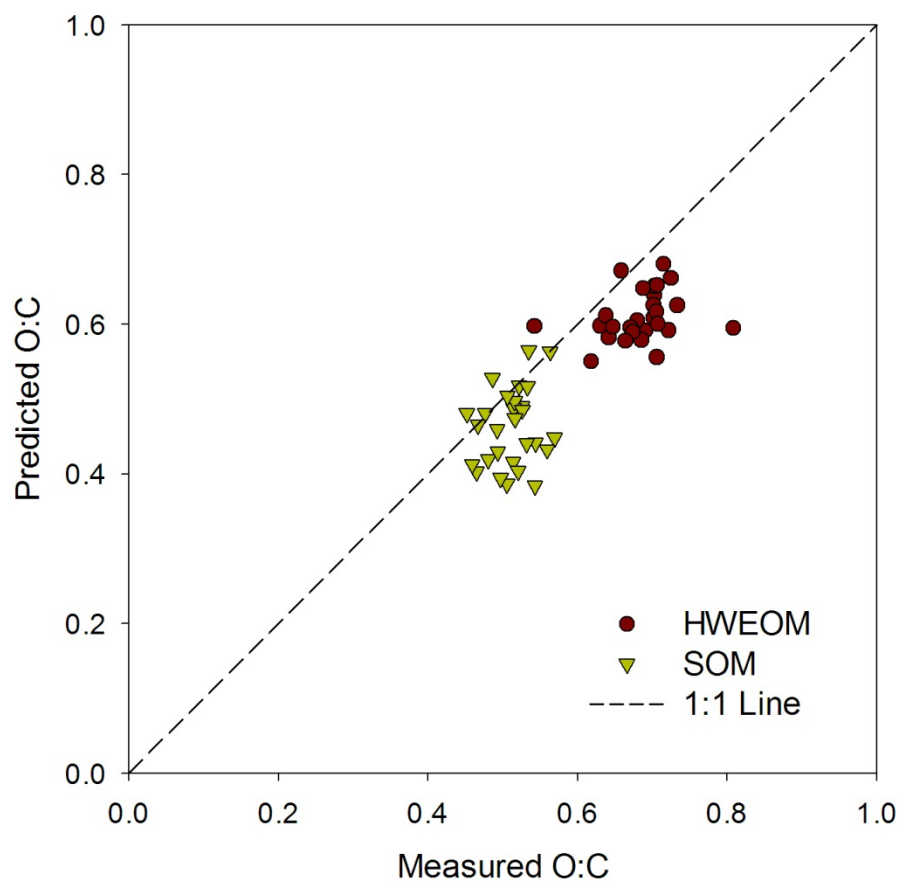


Figure 6





**Figure 7**



**Figure 8**