

Spatial molecular heterogeneity of POM during decomposition at different soil depths resolved by VNIR hyperspectral imaging

Julien Guigue¹, Christopher Just¹, Siwei Luo¹, Marta Fogt^{1,2}, Michael Schlöter^{1,2}, Ingrid Kögel-Knabner^{1,3}, and Eleanor Hobley^{1,4}

¹Chair of Soil Science, School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany

²Helmholtz Zentrum München, German Research Center for Environmental Health, Research Unit Environmental Genomics, Neuherberg, Germany

³Institute for Advanced Study, Technical University of Munich, Garching, Germany

⁴Zentrale Stelle für Informationstechnik im Sicherheitsbereich, Munich, Germany

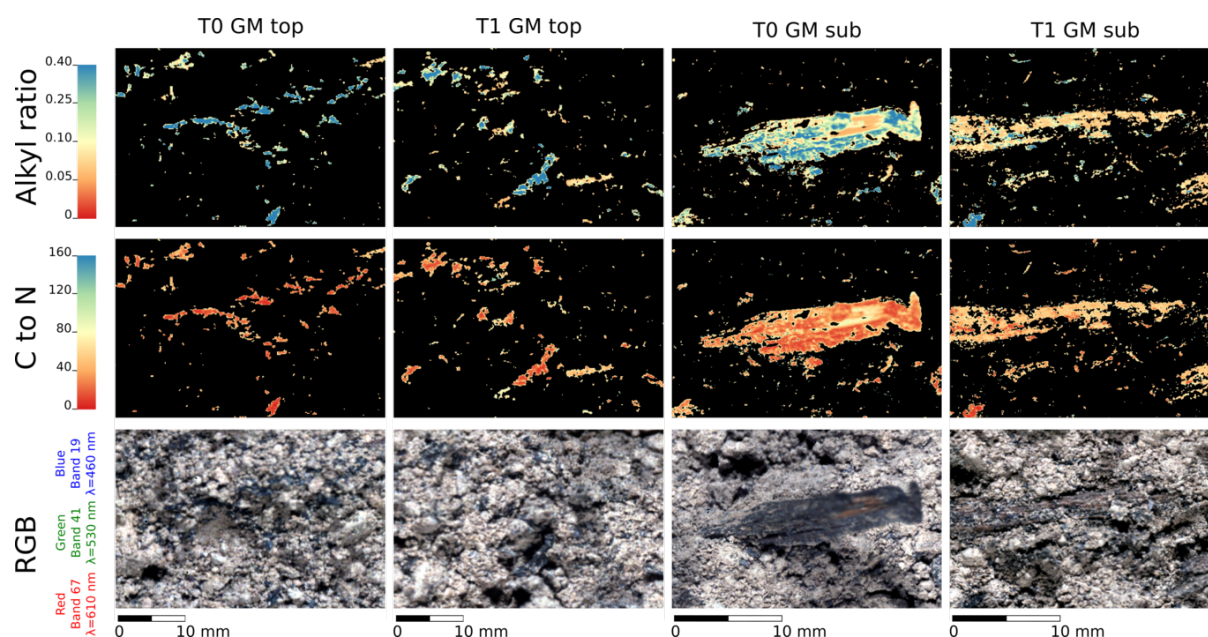
April 2, 2021

Abstract

Soil organic matter is composed of fractions with different functions and reactivity. Among these, particulate organic matter (POM) is the main educt of new inputs of organic matter in soils and its chemical fate corresponds to the first stages of the SOM decomposition cascade ultimately leading to the association of organic and mineral phases. We aimed at investigating the POM molecular changes during decomposition at a sub-millimetre scale by combining direct measurements of POM elemental and molecular composition with laboratory imaging VNIR spectroscopy. For this, we set up an incubation experiment to compare the molecular composition of straw and composted green manure, materials greatly differing in their C/N ratio, during their decomposition in reconstituted topsoil or subsoil of a Luvisol, and recorded hyperspectral images at high spatial and spectral resolutions of complete soil cores at the start and end of the incubation.

Hyperspectral imaging was successfully combined with machine learning ensembles to produce a precise mapping of POM alkyl/O-N alkyl ratio and C/N, revealing the spatial heterogeneity in the composition of both straw and green manure. We found that both types of organic amendment were more degraded in the reconstituted topsoil than in subsoil after the incubation. We also measured consistent trends in molecular changes undergone by straw, with the alkyl/O-N alkyl ratio slightly increasing from 0.06 to 0.07, and C/N dropping by about 40 units. The green manure material was very heterogeneous, with no clear molecular changes detected as a result of incubation.

The visualisation approach presented here enables high-resolution mapping of the spatial distribution of the molecular characteristics of organic particles in soil cores, and offers opportunities to disentangle the roles of POM chemistry and morphology during the first steps of the decomposition cascade of organic matter in soils.



KEYWORDS

alkyl/O-N alkyl ratio, artificial neural networks (ANN), C to N ratio, particulate organic matter (POM), soil organic matter (SOM), solid-state ^{13}C -CPMAS-NMR spectroscopy, subsoil, Visible-near infrared (VNIR) hyperspectral imaging

INTRODUCTION

Soils provide many services in terrestrial ecosystems and adequate management of agricultural soils support food security, water quality, biodiversity and climate regulation (Lal, 2016). Soil organic matter (SOM) quantity influences soil fertility and water holding capacity, and is associated with more diverse and resilient soil fauna and flora (Bossuyt *et al.*, 2005; Verbruggen *et al.*, 2016; Kravchenko *et al.*, 2019). Maintaining and increasing SOM quantity in agricultural soils is therefore often considered a keystone for their sustainable use (Chenu *et al.*, 2019).

The addition of organic residues into soil, either by direct return of crop residues or from exogenous sources, is a common practice in agricultural management aiming to enhance SOM quantity (Sisouvanh *et al.*, 2021; Maillard & Angers, 2014). SOM also controls other important soil parameters such as soil structural development or activity of (micro)organisms (Pascault *et al.*, 2013; Bai *et al.*, 2018). Organic fertilization through crop residue addition mainly results in the accumulation of coarse organic material at the annual to decadal timescale (van Wesemael *et al.*, 2019) while changes in the quantity of mineral-

associated organic matter (MAOM) occur generally on the scale of decades and more (Cambardella & Elliott, 1992; Cotrufo *et al.*, 2019). The coarse fraction of SOM, corresponding to particulate organic matter (POM), is composed of partially decomposed plant fragments, whose association with the soil mineral matrix is restricted to surface coatings of the organic particles by minerals or occlusion within aggregates (Golchin *et al.*, 1994; Wagai *et al.*, 2009). This fraction contrasts with the MAOM, which is composed of organic compounds with lower molecular weight and strongly attached to mineral phases (mostly to phyllosilicates and oxides) or protected within micro aggregates (von Lützow *et al.*, 2008). The accrual of SOM through POM accumulation has been proposed as a possible C sequestration strategy (Lavalée *et al.*, 2020) having the advantages, in contrast to MAOM, of not being constrained by saturation of mineral surfaces or C/N stoichiometry (Cotrufo *et al.*, 2019).

The composition of POM reflects both its source material and its degradation processes during decomposition. Initially, this fraction is primarily composed of two types of plant tissues, namely parenchymatic tissues (e.g. leaves and bark) and woody tissues, the former being mainly made up of (hemi-)cellulose walls and rich in proteins while the latter contains more lignin (Kögel-Knabner, 2002). During decomposition, POM serves as an energy source for microorganisms which take up C and other nutrients and respire a considerable fraction of the processed C. Under oxic conditions, the N used by microorganisms tends to be recycled, resulting in a decrease of the POM C to N ratio (C/N) during initial decomposition stages. Other indications of changes in POM properties during decomposition can be obtained from ¹³C nuclear magnetic resonance spectroscopy. Studies conducted on a diverse range of natural organic materials revealed that O/N alkyl structures are preferentially taken up by microorganisms, and that the relative proportion of alkyl C generally increases as decomposition progresses (Baldock *et al.*, 1997; Kögel-Knabner, 1997; Gao *et al.*, 2016). The ratio of alkyl C to O/N-alkyl C (hereafter noted alkyl ratio) is thus used as an indicator of degree of decomposition, with higher ratios indicating a greater degree of decomposition. These general dynamics of decomposition are modulated by the composition of the original material and soil conditions (Wilson *et al.*, 1983).

Water and oxygen availability are important for POM decomposition and can strongly vary with soil depth. In deeper horizons, decomposition might be restricted by lower temperature and oxygen availability in comparison to topsoils, even if there is little direct evidence of these constraints and that specific microbial taxa have temperature optima below 10 °C (Rumpel & Kögel-Knabner, 2011; Gao *et al.*, 2016). In addition, the spatial heterogeneity of soil components differs between topsoil and subsoil in agricultural soils, especially in fields where regular ploughing of the topsoil takes place (Chabbi *et al.*, 2009; Hobley *et al.*, 2018). SOM spatial distribution is more heterogeneous in the subsoil and this patchy distribution results in hotspots where SOM decomposition, and potentially stabilisation, as well as nu-

trient cycling is centralised (Heinze *et al.*, 2018; Heitkötter & Marschner, 2018; Hobley *et al.*, 2018). This highlights the importance of the location of organic substrates in niches, i.e. microdomains with various dynamics for their decomposition, and of techniques able to resolve such small scale heterogeneity.

Techniques to analyse the fine-scale spatial heterogeneity of SOM distribution and composition can help to elucidate processes of SOM decomposition. Laboratory VNIR imaging spectroscopy applied to Histosols has been shown to enable the identification of OM with different compositions (Steffens *et al.*, 2014; Granlund *et al.*, 2021). This method was also successfully used for the classification of diagnostic soil horizons and the mapping of several elements (C, N, Fe) at the pedon scale (Steffens & Buddenbaum, 2013; Hobley *et al.*, 2018; Sorenson *et al.*, 2020), as well as for the determination of soil structure arrangement and mapping of soil components (POM, Fe oxides, mineral matrix) at a sub-millimetre scale (Mueller *et al.*, 2021; Lucas *et al.*, 2020). Combined with modern prediction methods based upon machine learning and artificial intelligence (e.g. random forest or artificial neural networks), this high spatial-resolution analytical technique enables mapping of chemical composition in soil profiles at a sub-millimetre scale.

In this study we hypothesized that both soil depth and quality of organic amendment are strong determinants of the molecular composition of POM and its decomposition, and aimed at characterising in-situ the small scale molecular heterogeneity of organic amendments. Decomposition of POM in topsoil will be faster than in the subsoil, where nutrient and oxygen availability is limited, thereby retarding decomposition. Additionally, organic amendments high in N will be preferentially decomposed by soil microorganisms, whereas decomposition of organic amendments dominated by C compounds will be relatively slower due to N limitation during the decomposition process. To test these hypotheses we incubated soil samples from the surface (0-30 cm) and upper subsoil (30-60 cm) of a Luvisol and added either green manure with a low C/N or straw amendment with a high C/N for a period of 180 days. We used VNIR hyperspectral imaging combined with artificial intelligence modelling to map the composition of POM prior to and after incubation.

Materials and methods

Soil sample and incubation experiment

The soil used for the incubation experiment was sampled from an agricultural trial field at ‘Campus Klein-Altendorf’ experimental research station (50°37’16”N; 6°59’53”E), University of Bonn, Germany. The mean annual air temperature at Bonn-Rohleber is 10.3 °C and the mean annual precipitation is 816

mm for the period 1991-2010. The soil is a Haplic Luvisol (IUSS Working Group WRB, 2015) derived from quaternary Loess deposits, with a pH_{water} of 6.5 and a texture dominated by silt particles, with an enrichment in clay with depth (clay/silt/sand [%] 25-43/50-68/4-6 %). The organic C and total N contents are respectively 8.0 and 1.1 mg g^{-1} in the topsoil and 3.0 and 0.5 mg g^{-1} in the subsoil. Two soil depths (0-30 cm, and 30-60 cm), hereafter named topsoil and subsoil corresponding to the ploughed layer and the non-ploughed upper subsoil, were collected using a backhoe and directly sieved to 5 cm. Soil cores were prepared in 90-cm high PVC pipes ($\phi=7.5$ cm) by inserting, from bottom to top, 20 cm of quartz sand, 30 cm of subsoil, 5 cm of mixed topsoil and subsoil, and 25 cm of topsoil. Soils were amended with two types of organic materials, wheat straw and green manure, added either to the topsoil or to the subsoil. Straw material corresponds to harvest residues whereas the composted green manure was a mixture of green waste (trees, bushes and shrubs) from public green areas and parks. The organic material was incorporated by mixing it with the soil (1:4 volume mixing ratio for organic:soil) before packing the cores. Sixteen soil cores were prepared, corresponding to duplicates of the two types of organic material amendments (straw or green manure), admixed into one of the two soil depths (denoted top- or sub-) for initial sampling (T0) and after 180 days (T1) of incubation at 20 degC and 60 % maximum water-holding capacity determined gravimetrically.

VNIR hyperspectral imaging

Before recording hyperspectral images, soil cores were cut lengthwise, from bottom to top, into two equal halves and then dried at room temperature. For imaging, samples were illuminated with two 150-W halogen lamps. Hyperspectral images were recorded using a Hypspec VNIR-1800 camera (Norsk elektro optikk, Norway) after automatic dark background correction. The sensor was equipped with a 30-cm lens, giving a final field of view of approximately 9 cm for the 1800 detectors ($53 \times 53 \mu\text{m}^2$ per pixel). For each pixel, light reflectance intensity was measured for 186 bands in the region 400-990 nm (spectral resolution of 3.17 nm per band).

To account for potential unevenness in illumination and spectral response at different horizontal locations in the core, the spectral intensity (I) of the raw images of the soil cores were normalised to the defined reflectance (R) of calibration target for each wavelength (λ) and pixel (x):

$$R_{\Sigma\alpha\mu\pi\lambda\epsilon, \lambda, \xi} = \frac{I_{\Sigma\alpha\mu\pi\lambda\epsilon, \lambda, \xi}}{I_{T\alpha\rho\gamma\epsilon\tau, \lambda, \xi}} R_{T\alpha\rho\gamma\epsilon\tau, \lambda, \xi}$$

The background, plastic pipe, light overexposure and shadows were later masked using ENVI version 5.2 (Exelis Visual Information Solutions, Boulder, Colorado) by mapping the spectral intensity of the VNIR band at ~980 nm and the ratios of the spectral intensity of bands at 980 nm and 420 nm, and of bands at 630 nm and 420 nm. After comparing these with the original images, thresholds in the spectral response and ratios were identified and used to mask non-soil components of the images.

POM sampling in soil cores

We selected regions of interest (ROI) with different spectral signatures for sampling and analysis of C and N contents and for analysis of molecular structure information by solid-state ^{13}C -CPMAS-NMR spectroscopy. Regions of interest were sampled based upon score projections of the first three components of a principal component analysis (PCA) of the 186 bands in the VNIR spectra, which generally accounted for more than 95% of spectral variability in the cores. We sampled two ROI per core that received straw as a treatment (N=16). Because of the greater visual heterogeneity of the green manure, which could be roughly identified as wood-like or non-woody particles, we increased the number of ROI in these samples (N=31), sampling one to two ROI of visually identifiable wood-like material and two ROI of visually identifiable non-wood material per core (i.e. bark). Wood samples with a diameter larger than 5 mm were not sampled. In total, we sampled 47 ROIs, corresponding to straw or green manure at different stages of decomposition. All material was milled using an agate mortar and pestle prior to further analysis.

Solid-state ^{13}C CP-MAS-NMR spectroscopy

Solid-state ^{13}C cross-polarisation magic angle spinning nuclear magnetic resonance spectra of the ROI samples were obtained at a frequency of 50.3 MHz (Bruker DSX 200). The powdered samples were filled into a zirconium dioxide rotor and spun at a magic angle under a magnetic field of 6800 Hz with an acquisition time of 0.01024 s. A ramped ^1H pulse was applied during 1 ms contact time to prevent Hartmann-Hahn mismatches. An average of 5500 scans were accumulated with a delay time of 1 s according to the amount of sample and the carbon content.

Tetramethylsilane was equalised at 0 ppm as a reference for the chemical shifts. Phase and baseline correction of the acquired spectra were applied after Fourier transformation. The spectra were then integrated in following chemical shift regions according to (Beudert *et al.*, 1989) with slight adjustments according to (Mueller & Koegel-Knabner, 2009): -10 to 45 ppm (alkyl C), 45 to 110 ppm (O/N-alkyl C), 110 to 160 ppm (aromatic C) and 160 to 220 ppm (carboxyl C), spinning sidebands were included. We calculated the ratio of alkyl C to O/N-alkyl C (hereafter noted alkyl ratio) based on these integrated

shift regions.

C and N measurements

Total carbon and nitrogen contents of the ROI samples were measured by dry combustion (HEKAtech EuroEA 3000). Calibration was made against sulphanilamide ($C_6H_8N_2O_2S$, 41.8 % C and 16.3 % N) and BBOT ($C_{26}H_{26}N_2O_2S$, 72.5 % C and 6.5 % N). C/N ratio (hereafter noted C/N) is expressed as the mass ratio of the two elements. All measurements were performed with at least two analytical replicates. An additional replicate was analysed when the difference between analytical replicates exceeded 2 mg C g^{-1} .

Modelling

The alkyl ratio and C/N were modelled using the 186 band standard normal variate transformed VNIR reflectance spectra as predictor variables in random forest and artificial neural networks algorithms. The consistency of the spectra of a sample in the profile before and after powdering was checked before modelling. After removing five samples with an Euclidean distance of > 1.5 between the spectra from the core and the powder, four samples corresponding to features with overexposure or shadows leading to low quality spectral information, and four samples from a core replicate with inaccurate spectra, 34 samples were retained as the training dataset. The VNIR spectral data of each ROI were extracted and the mean spectra calculated, which were then used as predictors for POM composition.

Predictive modelling was performed using two algorithms implemented with R software version 4.0.3 (R Core Team, 2021). First, an unconstrained random forest (RF) algorithm was implemented after optimisation for the number of trees in the forest using the ‘party’ package (Hothorn *et al.*, 2006; Strobl *et al.*, 2007; Strobl *et al.*, 2008). Second, we implemented a random ensemble of fully-connected artificial neural networks (ensemble ANN) and optimized for the number of models in the ensemble. Individual ANN modelling was performed using the ‘neuralnet’ package (Fritsch *et al.*, 2019). Finally, we evaluated model averaging from the RF and ensemble ANN results, using both weighted and unweighted estimate averaging, with weighting applied proportional to the goodness-of-fit of the individual model estimates.

To implement the random ensemble ANN we independently trained an ensemble of neural network models, in which each model was initialised with random starting weights and an architecture of two hidden layers with a random number of nodes. The number of nodes was randomly selected between 3 and 185 for the first hidden layer and 2 to one less than the number of nodes in the first layer for the second hidden layer. The ensemble approach was applied in order to overcome the instability and high

sensitivity of ANN model performance to initialization which is to be expected for our small datasets (Kolen & Pollack, 1991). The random structure of the ANN models was selected to avoid optimisation and over-fitting issues given the small data-set used for model fit. The random ensemble was performed using n-out-of-n bootstrapping and was independently optimized for the number of models for each of the variables using the out-of-bag samples in each bootstrap. Optimization was performed using the MSE and percent variance explained by the model and the optimal number of models was selected according to the stability of model variance with increasing number of models. In the ANN models, the target variables were scaled to a range of 0-1 prior to model fitting and estimates rescaled to the original scale after prediction.

Model evaluation was performed using n out of n bootstrap estimation for both the RF and ANN models, with evaluation done using the out-of-bag model estimates using percentage of variance explained by the models as well as root mean-squared error (RMSE) and mean average error (MAE) to determine the best models. For the ANN, we evaluated both the mean and median of the out-of-bag estimates for predictive performance.

Before applying the predictive model to the images of whole soil cores, pixels were classified into two groups (POM or mineral soil) using a spectral angle mapper classification algorithm in ENVI version 5.2 (Exelis Visual Information Solutions, Boulder, Colorado) with a spectral angle threshold at 0.2 rad. The models were then used to predict the alkyl ratio as well as the C/N for each pixel classified as POM. After predicting the distribution of alkyl ratio or C/N in each core, values were checked for plausibility against the calibration range as well as against published literature values for these soil parameters.

Statistical testing

Statistical differences in the chemical characteristics of the organic materials were tested using 3- way ANOVA to evaluate the effects of incubation, type of added OM and depth on organic material chemical composition. Orthogonal contrasts were subsequently computed to estimate the effect of incubation on the different types of OM at the two depths before and after incubation. For the predictions of C/N and alkyl ratio based on hyperspectral imaging, the differences in the distributions between treatments and times were tested based upon the data from all pixels classified as POM using the Kolmogorov-Smirnov non-parametric test.

Results

Composition of POM

The alkyl ratio, alkyl, aryl and carboxyl functional group contributions respectively increased by 0.04-0.06, 1.8-2.6%, 3.3-6.0% and 2.2-2.6% as a result of incubation, while O/N alkyl and C/N decreased by 8.2-10.7% and 37-46 (Table 1). The changes in alkyl ratio and alkyl contributions resulting from the incubation were greater for green manure, while the changes in C/N, O/N alkyl, aryl and carboxyl contributions were greater for straw. For all molecular indicators, we measured a higher variability in the data for green manure compared to straw, as highlighted by the large standard deviations presented in Table 1.

	Organic Amendment	Incubation Time	n roi	Alkyl Ratio	Alkyl C	O/N- alkyl C	Aryl C	Car- boxyl C	C/N
					%	%	%	%	
Top- soil	GM	T0	4	0.07±0.03	5.0±2.2	76.8±4.2	15.6±2.8	2.2±1.7	98±50
		T1	4	0.1±0.06	6.2±3.1	68.1±8.2	21.4±2.3	4.1±4.6	69±38
	S	T0	4	0.06±0.01	4.7±0.4	81.4±0.8	10.9±0.4	3.1±0.5	92±7
		T1	4	0.09±0.01	6.4±0.6	68.9±4.3	18.0±1.9	6.5±1.8	50±7
Sub- soil	GM	T0	3	0.07±0.05	5.4±3.0	73.3±4.0	19.1±2.6	2.1±1.2	99±72
		T1	4	0.16±0.12	9.3±6.3	65.7±10.4	20.1±2.2	4.8±2.2	55±38
	S	T0	4	0.06±0.01	4.7±0.4	81.2±1.2	10.9±0.8	3.0±0.4	99±9
		T1	4	0.09±0.03	6.6±1.4	72.5±4.5	15.8±1.7	4.7±1.3	51±11

Table 1: Mean and standard deviation of ^{13}C -CPMAS-NMR functional groups distribution, and of C/N ratio of the ROI for different amendments (GM: green manure / S: straw), soil matrices (topsoil / subsoil) and incubation time (start: T0 / end: T1).

The ANOVA of the molecular indicators measured on ROI indicated significant differences ($p < 0.05$) between OM composition at T0 and T1, whereas differences between green manure and straw were significant only for O/N alkyl and aryl contributions, with higher O/N alkyl and lower aryl contributions

for straw compared to green manure (Table 1). No significant differences in POM composition between depths were found. No significant interactions between incubation time, depths and type of added OM were found so 3-way ANOVA models without interactions were applied for all variables presented in Table 1.

Prediction of POM properties from machine learning using hyperspectral images

For predicting C/N, the random forest modelling was optimal with 200 trees, whereas the random ensemble ANN was optimal with 100 ANNs in the ensemble. For predicting alkyl ratio, the random forest modelling was optimal with 400 trees, and the random ensemble ANN was optimal with 200 neural networks in the ensemble. Both modelling approaches yielded accurate predictions, with high to very high coefficients of determination ($R^2 > 0.7$). The predictions for C/N were better than for alkyl ratio, and in general, the random ensemble ANN outperformed the RF models, and the median estimates of the ANN were better than the mean estimates (Table S1).

For C/N, model averaging improved the predictions, with the best estimates obtained using a weighted average of the RF and median of the random ensemble ANN ($R^2=0.90$, RMSE 14.3, MAE 10.4). For alkyl ratio predictions, the random ensemble ANN outperformed the RF and model averaging yielded no improvement as assessed by the goodness-of-fit parameters. In predicting the measured values close to the maximum and minimum, the RF models tended to underestimate the higher and overestimate the lower, while the ensemble ANN models did not show this tendency (Figure 1 and Fig. S1).

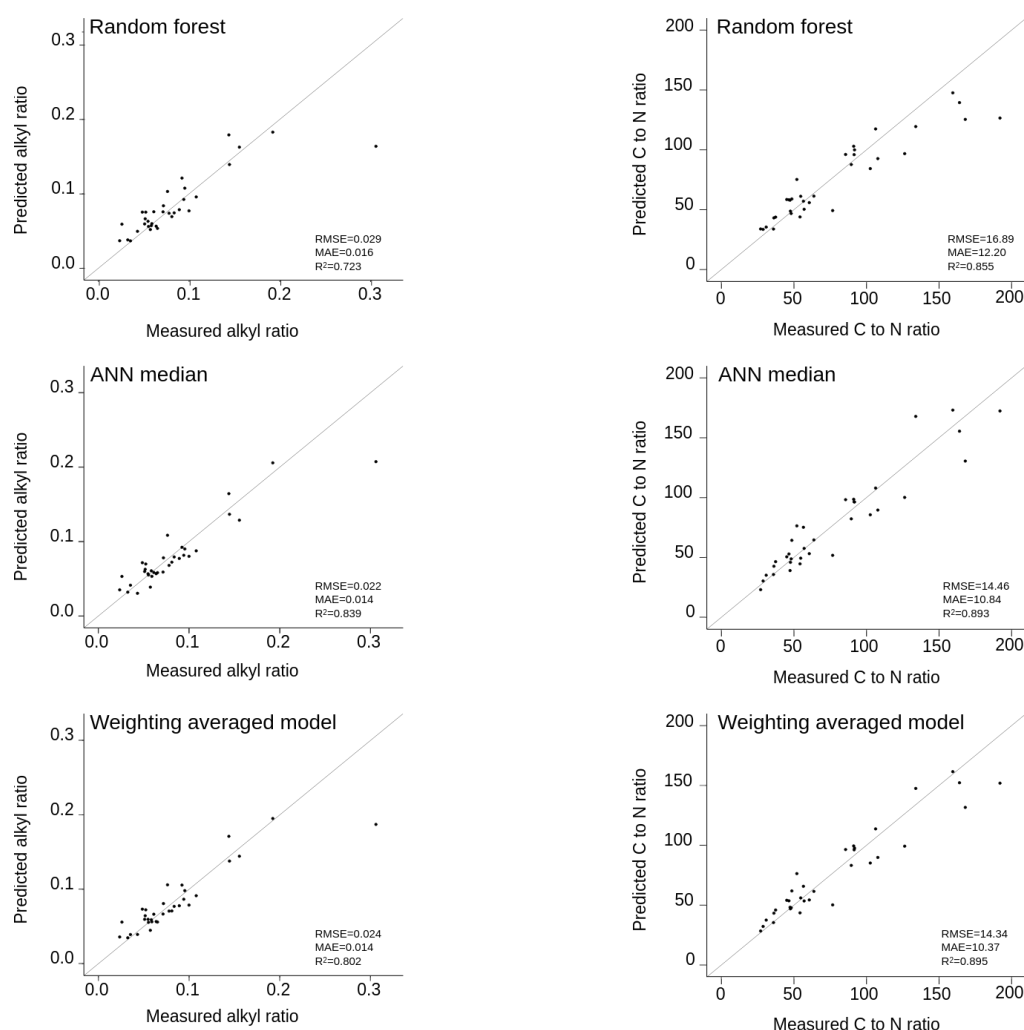


Figure 1: Alkyl ratio (left) and C/N (right) predicted from the random forest (top), ensemble ANN models (middle) and weighted averaged models as functions of measured values

The ANN models predicted impossible (< 0) and highly implausible C/N ratios (< 15) for a large number of pixels in the cores amended with green manure (Table 2 and S2), with up to 47 % of pixels containing predictions outside of the calibration range and deemed unreliable. For the alkyl ratio, the number of pixels predicted outside the calibration range using the ANN ensemble was very similar to that of C/N and corresponded to largely the same pixels. We thus considered the ANNs unreliable and used the weighted averaged models from the RF and the median predictions from the random ensemble ANNs for predictive purposes of the two types of OM amendment and for further interpretation. The model averaging resulted in the reduction in the proportion of pixels predicted out of the calibration range (20 % for green manure and < 0.2 % for straw).

			Alkyl				C/N								
			Ratio												
			npix-	mean	me-	SD	range	skew	Kurt.	mean	me-	SD	range	skew	Kurt.
			els		dian					dian					
Top-soil	GM	T0	233975	0.24	0.21	0.13	0.04	0.32	-	35.8	34.2	16.8	-2.4 –	0.65	0.64
							–	1.34					133.5		
							0.54								
	T1	288932	0.17	0.11	0.12	0.03	1.12	-	47.7	48.8	19.7	2.8 –	0.39	0.76	
							–	0.19					170.1		
							0.52								
S	T0	195318	0.07	0.06	0.01	0.03	0.89	0.93	101.3	101.2	19.1	8.8 –	0.24	0.83	
							–						201.6		
							0.19								
	T1	242590	0.07	0.07	0.02	0.02	5.41	51.02	68.2	59.6	26.5	-9.6 –	1.31	1.46	
							–						193.6		
							0.42								
Sub-soil	GM	T0	182448	0.17	0.11	0.12	0.03	1.04	-	51.5	49.2	26.9	-1.9 –	1.1	1.97
							–	0.35					191.4		
							0.52								
	T1	165390	0.15	0.09	0.12	0.03	1.43	0.62	54.1	53.3	27.1	-5.5 –	1.01	1.85	
							–						185.5		
							0.52								
S	T0	212277	0.06	0.06	0.01	0.03	4.22	86.56	107.1	107.9	15.3	3.2 –	-	2.83	
							–						198.5	0.1	
							0.39								
	T1	451221	0.07	0.07	0.01	0.02	0.87	24.36	78.8	73.3	27.2	3.6 –	0.74	0.05	
							–						200.2		
							0.45								

Table 2: Summary data for Alkyl ratio and C/N of POM in each soil column predicted from the weighted averaged models trained on the VNIR hyperspectra (skew: skewness of the distribution, Kurt: kurtosis of the distribution)

Predicting chemical properties of POM in soil cores before and after incubation

The number of pixels associated with POM varied for each hyperspectral image due to the differences in exposed POM area, ranging from 165390 to 451221 pixels per image (Table 2). For soil cores amended with straw, both the distribution of alkyl ratio and C/N were approximately unimodal, indicating a single peak centred around 0.05 - 0.08 for alkyl ratio and 60 - 110 for C/N, with a minor shoulder appearing in both after incubation (Figure 2). Concurring with the measured values of the ROI samples (Table 1), the predicted alkyl ratio of straw were significantly different as a result of the 180-day incubation (Table 3) in both the topsoil and subsoil, with medians slightly increasing from 0.06 to 0.07 (Table 2). Importantly, although the initial alkyl ratio of straw material was initially different between the topsoil and subsoil, the differences in the Kolmogorov-Smirnov D statistic decreased by over 50 % at the end of the incubation, indicating the greater similarity in alkyl ratio of straw added to both two depths (Table 3, Figure 2).

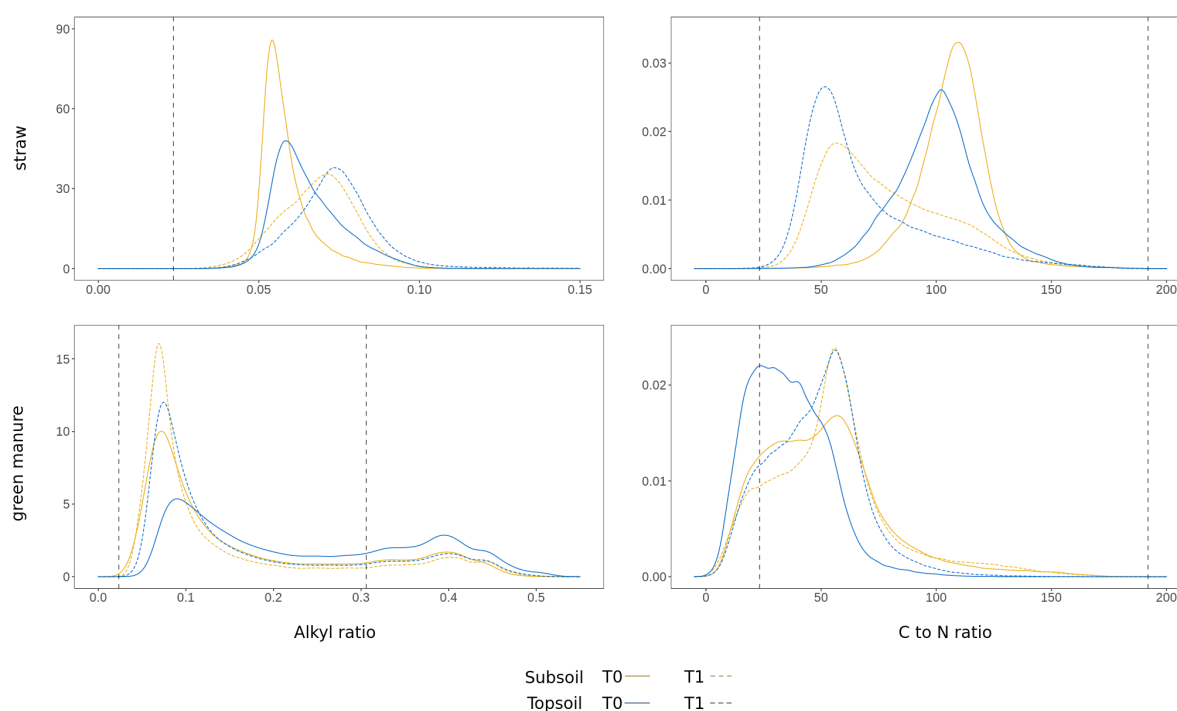


Figure 2: Density plots of predicted alkyl ratio and C/N of macro POM modelled for each amendment (green manure / straw), horizon (topsoil / subsoil) and incubation time (start: T0 / end: T1). The y-axis of each individual plot is the probability density function and is normalized so that the total area under the curve integrates to one

For the C/N of straw, the differences in the distributions remained similar when comparing the soil depths to which the material was added for the two incubation times and median C/N signatures were higher by 7 and 13 C/N units in subsoil (Table 2 and 3). The C/N was greatly modified by incubation (Table 3) with similar magnitudes at both depths corresponding to decreases by about 40 units, as highlighted by the high Kolmogorov-Smirnov D statistic (Figure 2).

Organic amendment	Comparison	Alkyl Ratio D	C/N D
GM	T0 to T1 (Topsoil)	0.29	0.28
	T0 to T1 (Subsoil)	0.13	0.08
	Topsoil to subsoil (T0)	0.28	0.29
	Topsoil to subsoil (T1)	0.15	0.1
Straw	T0 to T1 (Topsoil)	0.35	0.62
	T0 to T1 (Subsoil)	0.49	0.58
	Topsoil to subsoil (T0)	0.34	0.2
	Topsoil to subsoil (T1)	0.16	0.21

Table 3: Kolmogorov-Smirnov test statistic (D) for comparison of distributions of predicted C/N ratios and alkyl ratios from hyperspectral imaging. All differences were estimated to be statistically significant at $p < 0.001$ significance level

For the cores amended with green manure, there was no shift in the mean predicted alkyl ratio and C/N after incubation (Table 2), a result of the very high coefficients of variance, which range from 55 to 82 % for the predicted alkyl ratio and from 41 to 52 % for C/N . Nevertheless, the distributions of the predicted green manure material statistically differed before and after incubation as well as between the topsoil and subsoil. The alkyl ratio of green manure decreased during the incubation at both depths, and was larger in the topsoil at both sampling times (Table 2). There were greater differences in the composition of green manure material admixed to the topsoil and subsoil before incubation than after incubation, with D reduced by around 50 % for the alkyl ratio and 75 % for the C/N ratio after incubation (Figure 2 and Table 3).

Fine-scale mapping of POM composition

The high resolution of the VNIR imaging allowed us to identify and characterise organic particles ranging from a few millimetres down to sub-millimetre scale (Figures 3 and 4). For both green manure and straw, their chemical composition, i.e. alkyl ratio and C/N, could be mapped at a very fine scale, with the

maps revealing the spatial heterogeneity of both indicators for the two types of OM. This heterogeneity could be seen in both very small particles of a few hundreds of μm^2 , as well as within large particles of several mm^2 . Even in the case of straw, which was a lot more homogeneous than green manure, the chemical maps highlighted the spatial differences in the molecular composition.

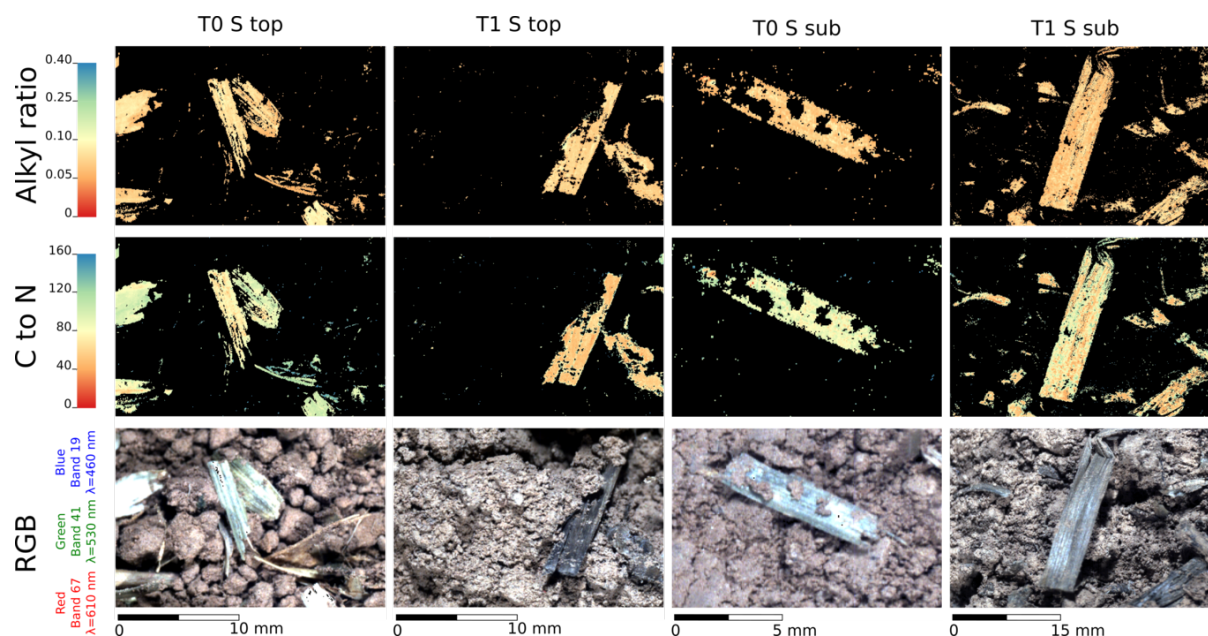


Figure 3: Chemical maps of wheat straw particles (S) before and after 180 days of incubation. The top row shows the distribution of the alkyl ratio of POM predicted from weighted averaged models. The distribution of C/N in POM predicted from weighted averaged models is represented on the second row. The bottom row contains RGB colour images (normalised against the reflectance target) of soil material with POM. The depicted areas correspond to a limited portion of soil cores and are presented as examples. T0: beginning of the incubation, T1: end of the 180-day incubation. S: straw, top: topsoil, sub: subsoil

For green manure, the chemical signatures of low alkyl ratio and large C/N, as well as large alkyl ratio and low C/N were co-localized (Figure 3). At the end of the incubation, we were able to map the decrease in the alkyl ratio of green manure particles and increase in their C/N signatures (Figures 2 and 3). Chemical maps for straw confirmed the co-localization of low alkyl ratio and large C/N, as well as the opposite. We generally detected a lower contribution of very small straw particles and, as observed for green manure, the variability in both alkyl ratio and C/N was similar between very small and within large particles (Figure 4).

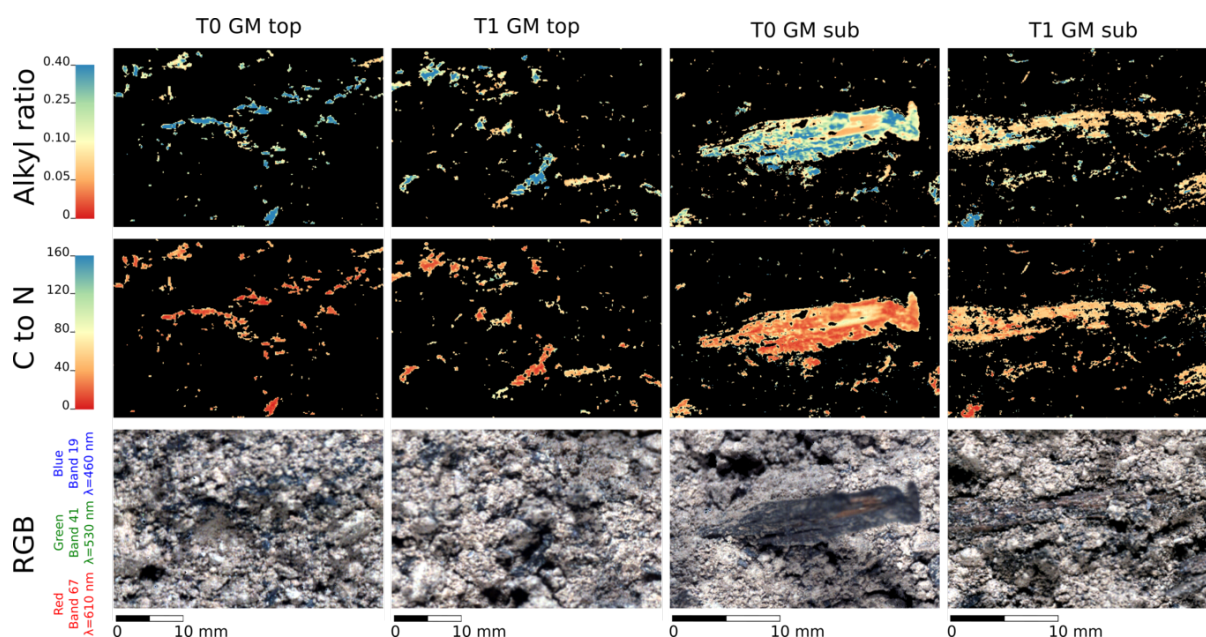


Figure 4: Chemical maps of green manure particles before and after 180 days of incubation. The top row shows the distribution of the alkyl ratio of POM predicted from weighted averaged models. The distribution of C/N in POM predicted from weighted averaged models is represented on the second row. The bottom row contains RGB colour images (normalised against the reflectance target) of soil material with POM. The depicted areas correspond to a limited portion of soil cores and are presented as examples. T0: beginning of the incubation, T1: end of the 180-day incubation. GM: green manure, top: topsoil, sub: subsoil

Discussion

VNIR hyperspectral imaging to map POM molecular composition

The successful modelling results for both the alkyl ratio and C/N of POM we achieved by linking VNIR with C/N measurements and ^{13}C -CPMAS-NMR spectroscopic data in machine learning ensemble models validates this method as appropriate for studying the decomposition of organic amendments and plant residues in soils at a sub-millimetre scale. (Steffens *et al.*, 2014) were already able to distinguish three different classes of OM at different stage of decomposition based on the composition determined by ^{13}C -CPMAS-NMR spectroscopy and on their VNIR reflectance spectra, but we went further by predicting molecular indicators, i.e. alkyl ratio and C/N, of POM at a very fine scale (Figure 3 and 4). (Xu *et al.*, 2020) achieved the mapping of so-called different soil OM fractions, namely soil organic carbon,

salt-extractable organic carbon and readily oxidizable organic C. These variables are strongly correlated with total organic carbon contents in the soil and do not reveal much about the composition of soil OM. To our knowledge, our work is the first to focus on POM molecular characterisation by VNIR hyperspectral imaging and to predict relevant proxies, C/N and alkyl/O-N alkyl ratio, for decomposition stage of organic particles at such a high resolution (Figure 2).

Classical VNIR spectroscopy has been used to estimate the contents of organic C, CaCO_3 , total N, Fe oxides and clays among other soil parameters (Ben-Dor & Banin, 1995; Chang *et al.*, 2001; Viscarra Rossel *et al.*, 2006), but also the type of soil OM. For example to differentiate between POM, mineral-associated organic matter (MAOM) and pyrogenic organic carbon (Hobley *et al.*, 2016; Sanderman *et al.*, 2021). Depending on the wavelength range of sensors, VNIR imaging methods have the potential to predict the contents of these soil parameters and the detailed mapping made possible by this technique provides a substantial advantage for the understanding of soil functioning by taking into account the spatial arrangement of soil components. Our results indicate that VNIR imaging combined with machine learning can be used to differentiate POM from the mineral soil and to predict molecular composition of the identified POM particles, revealing the heterogeneity in their molecular composition at a sub-millimetre scale (Table 2 and Figure 2).

However there are limitations that still require further consequent developments of this approach. Firstly, the predictions are dependent on the choice of model. This has been already shown in many studies before, and the choice of the most appropriate modelling algorithm, or combination of algorithms, varies between studies and modelled parameters (Hobley *et al.*, 2018; Xu *et al.*, 2020; Granlund *et al.*, 2021). By design, RF does not extrapolate the predictions beyond the calibration range. What initially appears to be an advantage, since extrapolation is not desirable in the use of most machine learning algorithms, leads to a compression of the number of predictions in the marginal areas of the calibration range (Figure S1). The ensemble ANN on the other hand has better goodness-of-fit parameters and tuning of the number of layers was not needed, only the number of models was. This reduces the chance of over-fitting a single ANN by constraining it to a certain architecture. We obtained excellent predictions inside the calibration range, but extrapolation was not constrained to plausible results, indicated by the very large proportion of nonsensical and impossible predictions outside the calibration range (Table S2).

Multilayer perceptrons like ANN algorithms are known to perform well in generalisation but perform poorly with extrapolation. ANN ensemble and model averaging overcomes these shortfalls to some extent (Hobley *et al.*, 2018) and delivered the best and most plausible results in our study, but the values beyond the calibration range must be interpreted with great care (Figure 2).

Another limitation of the modelling approach is that predictions are applied pixel wise ($53 \times 53 \mu\text{m}^2$ per pixel) whilst calibration is bound to ROI with larger areas (up to centimetre scale). The ROI areas must be large enough to sample the thinnest layer as possible to obtain the minimum quantities of material (usually a minimum of 100 mg) needed for C, N and ^{13}C -CPMAS-NMR spectroscopy measurements of the material measured by VNIR, which is only based on surface reflectance, resulting in the need for very shallow sampling. This is a classical limitation we noticed in other studies on SOM prediction based on hyperspectral imaging. Improvement of the calibration procedure requires methods producing spatially-resolved molecular information at a lower scale, such as the combination of scanning electron microscope with energy dispersive X-ray spectroscopy (SEM-EDX, e.g. (Hapca *et al.*, 2015)) or scanning transmission X-ray microscopy.

Molecular composition of straw and green manure at different stages of decomposition

The dominance of a single but shifting molecular signature in the wheat straw before and after incubation for both alkyl ratio and C/N is consistent with studies reporting decreasing proportions of straw carbon as carbohydrate and increasing proportions of aromatic compounds during straw decomposition (Cogle *et al.*, 1989; Gao *et al.*, 2016). The alkyl ratio increase indicates the decomposition of straw compounds, likely by a shift from substituted aliphatic alcohols and ethers to unsubstituted C in paraffinic structures (Wilson *et al.*, 1983; Kögel-Knabner, 2002), whereas the decrease in C/N is consistent with a loss of C caused by microbial respiration (Geissen & Brümmer, 1999) and the preservation of N by decomposers in their tissues and by-products. However, the greater heterogeneity of both alkyl ratio and C/N of straw at the end of the incubation (Table 2) indicates that decomposition was spatially heterogeneous and that some regions of the straw amendment decomposed faster than others. Although this could be explained by preferential decomposition of specific plant-derived structural compounds by microorganisms during the first stages of decomposition (Golchin *et al.*, 1997), this spatially heterogeneous degradation of the chemically homogeneous straw amendment is more likely attributable to abiotic factors, i.e. the microscale conditions of the microbial habitats surrounding POM that regulates soil moisture and microbial accessibility of the OM (Dungait *et al.*, 2012).

In contrast to the clear shift in POM composition during incubation of straw, changes in the composition of green manure were less clear, a result of the greater heterogeneity of green manure before and after incubation. Indeed, the general increase in C/N and decrease in alkyl ratio after incubation are contrary to typical changes induced by decomposition of organic materials (Golchin *et al.*, 1994; Baldock *et al.*, 1997; Stone *et al.*, 2001; Kögel-Knabner, 2002). We attribute this to a shift in the exposed surface of

the material before and after incubation. The green manure material contained highly different types of plant tissues, such as leaves, woody parts and bark material and was composted before we used it. While pieces of bark and partially decomposed green manure were observed at the beginning of the experiment, these particles may have significantly peeled off and dispersed during the incubation, exposing internal woody structures of green manure material. To better explain the changes in green manure chemistry, further research into the molecular composition of the various types of organic material present in the mixture and their decomposition is required.

Straw amendments are more decomposed in topsoils than in subsoils

The lower C/N and larger alkyl ratio of straw particles (Table 2), and the stronger shifts in distribution in the topsoil at the end of the incubation indicate a more advanced degradation than in the subsoil (Figure 2), which we attribute to the differences in the initial conditions of the two soil matrices that drive decomposition processes and rates. Generally, topsoils accommodate more diverse microbial communities and more microbial biomass than subsoils (Taylor *et al.*, 2002), and contain more SOC. The production of dissolved organic matter (DOM) is correlated with high SOC contents and abundance of microorganisms (Guigue *et al.*, 2015) and larger quantities of DOM are measured in topsoils than in subsoils (Kalbitz *et al.*, 2000; Kaiser & Kalbitz, 2012). DOM is composed of labile and energy-rich compounds easily converted either by microbial resynthesis or respiration (Strauss & Lamberti, 2002; Guggenberger & Kaiser, 2003; Kaiser & Kalbitz, 2012). This fuels the production of extracellular hydrolytic enzymes that contribute to POM decomposition (Berg & McClaugherty, 2014), and enables the faster decomposition of the added fresh organic matter in topsoils than subsoils.

C/N spatial distribution and coupling of C and N cycling

The C/N of above 40 after incubation (Figure 2) are well above typical lower bounds for SOM in central European agricultural soils, which rarely exceed C/N of 15 (Matschullat *et al.*, 2018). This indicates that the decomposition processes of the POM was not completed during the period of incubation and we assume that C/N of POM would continue to decrease if the incubation was prolonged. The initially lower C/N of green manure in the topsoil is likely the result of the prior composting of this material. It may also result from a greater proportion of N-rich compounds, such as bark, on the surface of the composted material when it was introduced to the topsoil for incubation.

The C/N of the POM fraction has been reported as being greater than 20 in soils under cropland, forest and grassland (Warren & Whitehead, 1988; Meijboom *et al.*, 1995; Hassink, 1995; Baldock *et al.*, 2003; Bimüller *et al.*, 2014). In addition, C/N ranging from 20 to 25 are generally accepted to be

the thresholds for the shift of microbial N immobilisation to N mineralisation (Nicolardot *et al.*, 2001; Robertson & Groffman, 2015) that stabilises the C/N by balancing C and N losses and further leading to a rapid mineralisation of POM. Straw decomposition is reported to start rapidly before slowing down during the process (de Willigen *et al.*, 2008), with decomposition mostly fuelled by hemicellulose while more recalcitrant ligneous materials are decomposed slower. Thus, the relative contribution of ligneous material increases during decomposition (Cogle *et al.*, 1989; Gao *et al.*, 2016), concurrent with decreases in the decomposition rate, corresponding to the preservation of remaining straw residues with a high C/N. Furthermore, lignin macromolecules decomposition is promoted when the macrostructures are firstly shredded by macro fauna (Scheu, 1992). The absence of such organisms in our experiment likely supported the preservation of large lignin-like moieties with high C/N.

Conclusion

The coupling of VNIR imaging with machine learning modelling was successful for the sub-millimetre scale mapping of molecular composition of various types of POM at distinct decomposition stages. Our novel approach based on model averaging of a random forest with an ensemble ANN overcomes issues relating to training of a single ANN and extrapolation of results beyond calibration ranges, presenting a new direction in machine learning and spectroscopy in soils. With this technique, we demonstrated the spatially heterogeneous changes in alkyl ratio and C/N of POM during decomposition with an overall increase in alkyl ratio and decrease in C/N of straw as a result of decomposition. The changes of the more heterogeneous green manure showed an opposite trend, likely associated with the preferential decomposition of N-rich bark tissues together with the preservation of less decomposable C-rich plant residues enriched in ligneous material. The decomposition of both straw and green manure was retarded in the subsoil compared to topsoil, as highlighted by smaller changes in the POM in the subsoil after the 6-month incubation in the subsoil. The visualisation approach presented has a great potential for applications aiming to investigate the spatial heterogeneity in molecular changes of organic particles during decomposition, and can help to disentangle the concurrent roles of accessibility and recalcitrance during the first steps of the decomposition cascade of organic matter in soils.

Acknowledgements

This work was financially supported by the German Federal Ministry for Education and Research, through the BonaRes initiative (BMBF, grant FKZ 031B0026B, Soil3 project and 031B0511C BonaRes centre). As a part of the BonaRes initiative, the authors would like to thank all contributors to the Soil3

project and the BonaRes centre.

Supplementary files

Hosted file

T0_S_sub_preds_Figure_S1.pdf available at <https://authorea.com/users/397797/articles/510586-spatial-molecular-heterogeneity-of-pom-during-decomposition-at-different-soil-depths-resolved-by-vnir-hyperspectral-imaging>

Hosted file

T0_GM_sub_preds_Figure_S1.pdf available at <https://authorea.com/users/397797/articles/510586-spatial-molecular-heterogeneity-of-pom-during-decomposition-at-different-soil-depths-resolved-by-vnir-hyperspectral-imaging>

Hosted file

Table_S1_Goodness_of_fit_all_models.xlsx available at <https://authorea.com/users/397797/articles/510586-spatial-molecular-heterogeneity-of-pom-during-decomposition-at-different-soil-depths-resolved-by-vnir-hyperspectral-imaging>

Hosted file

Table_S2_data_plausible.xlsx available at <https://authorea.com/users/397797/articles/510586-spatial-molecular-heterogeneity-of-pom-during-decomposition-at-different-soil-depths-resolved-by-vnir-hyperspectral-imaging>

References

- Bai, Z., Caspari, T., Gonzalez, M.R., Batjes, N.H., Mäder, P., Bünemann, E.K., Goede, R. de, Brussaard, L., Xu, M., Ferreira, C.S.S., Reintam, E., Fan, H., Mihelič, R., Glavan, M. & Tóth, Z. 2018. Effects of agricultural management practices on soil quality: A review of long-term experiments for Europe and China. *Agriculture, Ecosystems & Environment*, **265**, 1–7, (At: <https://linkinghub.elsevier.com/retrieve/pii/S016788091830224X>).
- Baldock, J.A., Currie, G.J. & Oades, J.M. 2003. Organic matter as seen by solid state ¹³C NMR and pyrolysis tandem mass spectrometry. In: *Advances in Soil Organic Matter Research*, pp. 45–60. W.S. Wilson. Woodhead Publishing.

- 460 Baldock, J.A., Oades, J.M., Nelson, P.N., Skene, T.M., Golchin, A. & Clarke, P. 1997. Assessing the
461 extent of decomposition of natural organic materials using solid-state ^{13}C NMR spectroscopy. *Soil*
462 *Research*, **35**, 1061–1084, (At: <http://www.publish.csiro.au/?paper=S97004>).).
- 463 Ben-Dor, E. & Banin, A. 1995. Near-Infrared analysis as a rapid method to simultaneously evaluate
464 several soil properties. *Soil Science Society of America Journal*, **59**, 364–372, (At: <http://doi.wiley.com/10.2136/sssaj1995.03615995005900020014x>).).
- 466 Berg, B. & McClaugherty. 2014. *Plant litter – decomposition, humus formation, carbon seques-*
467 *tration..* 3rd ed. Springer-Verlag Berlin Heidelberg. (At: [https://www.springer.com/gp/book/](https://www.springer.com/gp/book/9783642388200)
468 [9783642388200](https://www.springer.com/gp/book/9783642388200)).).
- 469 Beudert, G., Kögel-Knabner, I. & Zech, W. 1989. Micromorphological, wet-chemical and ^{13}C NMR
470 spectroscopic characterization of density fractionated forest soils. *The Science of The Total Environment*,
471 **81-82**, 401–408, (At: <https://linkinghub.elsevier.com/retrieve/pii/0048969789901484>).).
- 472 Bimüller, C., Mueller, C.W., Lützw, M. von, Kreyling, O., Kölbl, A., Haug, S., Schloter, M. & Kögel-
473 Knabner, I. 2014. Decoupled carbon and nitrogen mineralization in soil particle size fractions of a
474 forest topsoil. *Soil Biology and Biochemistry*, **78**, 263–273, (At: [http://linkinghub.elsevier.com/](http://linkinghub.elsevier.com/retrieve/pii/S0038071714002764)
475 [retrieve/pii/S0038071714002764](http://linkinghub.elsevier.com/retrieve/pii/S0038071714002764)).).
- 476 Bossuyt, H., Six, J. & Hendrix, P.F. 2005. Protection of soil carbon by microaggregates within earth-
477 worm casts. *Soil Biology and Biochemistry*, **37**, 251–258, (At: [https://linkinghub.elsevier.com/](https://linkinghub.elsevier.com/retrieve/pii/S0038071704002883)
478 [retrieve/pii/S0038071704002883](https://linkinghub.elsevier.com/retrieve/pii/S0038071704002883)).).
- 479 Cambardella, C.A. & Elliott, E.T. 1992. Particulate soil organic-matter changes across a grassland
480 cultivation sequence. *Soil Science Society of America Journal*, **56**, 777–783, (At: <http://doi.wiley.com/10.2136/sssaj1992.03615995005600030017x>).).
- 482 Chabbi, A., Kögel-Knabner, I. & Rumpel, C. 2009. Stabilised carbon in subsoil horizons is located in
483 spatially distinct parts of the soil profile. *Soil Biology and Biochemistry*, **41**, 256–261, (At: <https://linkinghub.elsevier.com/retrieve/pii/S0038071708003660>).).
- 485 Chang, C.-W., Laird, D.A., Mausbach, M.J. & Hurburgh, C.R., Jr. 2001. Near-infrared reflectance
486 spectroscopy–principal components regression analyses of soil properties. *Soil Science Society of America*
487 *Journal*, **65**, 480, (At: <https://www.soils.org/publications/sssaj/abstracts/65/2/480>).).
- 488 Chenu, C., Angers, D.A., Barré, P., Derrien, D., Arrouays, D. & Balesdent, J. 2019. Increasing organic
489 stocks in agricultural soils: Knowledge gaps and potential innovations. *Soil and Tillage Research*, **188**,
490 41–52, (At: <https://linkinghub.elsevier.com/retrieve/pii/S0167198718303738>).).

- 491 Cogle, A.L., Saffigna, P.G. & Barron, P.F. 1989. The use of ^{13}C -NMR for studies of wheat straw
492 decomposition. *Plant and Soil*, **113**, 125–128, (At: <http://link.springer.com/10.1007/BF02181930>.
493).
- 494 Cotrufo, M.F., Ranalli, M.G., Haddix, M.L., Six, J. & Lugato, E. 2019. Soil carbon storage informed
495 by particulate and mineral-associated organic matter. *Nature Geoscience*, **12**, 989–994, (At: <http://www.nature.com/articles/s41561-019-0484-6>.).
- 496
- 497 Dungait, J.A.J., Hopkins, D.W., Gregory, A.S. & Whitmore, A.P. 2012. Soil organic matter turnover is
498 governed by accessibility not recalcitrance. *Global Change Biology*, **18**, 1781–1796.
- 499 Fritsch, S., Guenther, F. & Wright, M.N. 2019. neuralnet: training of neural networks. (At: <https://github.com/bips-hb/neuralnet>.).
- 500
- 501 Gao, H., Chen, X., Wei, J., Zhang, Y., Zhang, L., Chang, J. & Thompson, M.L. 2016. Decomposition
502 dynamics and changes in chemical composition of wheat straw residue under anaerobic and aerobic
503 conditions. *PLoS ONE*, **11**, e0158172, (At: <https://dx.plos.org/10.1371/journal.pone.0158172>.
504).
- 505 Geissen, V. & Brümmer, G.W. 1999. Decomposition rates and feeding activities of soil fauna in deciduous
506 forest soils in relation to soil chemical parameters following liming and fertilization. *Biology and Fertility*
507 *of Soils*, **29**, 335–342, (At: <http://link.springer.com/10.1007/s003740050562>.).
- 508 Golchin, A., Baldock, J. & Oades, J.M. 1997. A model linking organic matter decomposition, chemistry
509 and aggregate dynamics.. In: *Soil processes and the carbon cycle*, pp. 245–266. Lal, R; Kimble, JM;
510 Follett, RF; Stewart, BA eds. Boca Raton, FL, USA: CRC Press.
- 511 Golchin, A., Oades, J.M., Skjemstad, J.O. & Clarke, P. 1994. Study of free and occluded partic-
512 ulate organic matter in soils by solid state ^{13}C CP/MAS NMR spectroscopy and scanning electron
513 microscopy. *Australian Journal of Soil Research*, **32**, 285–309, (At: [http://www.publish.csiro.au/
514 ?paper=SR9940285](http://www.publish.csiro.au/?paper=SR9940285).).
- 515 Granlund, L., Keinänen, M. & Tahvanainen, T. 2021. Identification of peat type and humification by
516 laboratory VNIR/SWIR hyperspectral imaging of peat profiles with focus on fen-bog transition in aapa
517 mires. *Plant and Soil*, **460**, 667–686, (At: [http://link.springer.com/10.1007/s11104-020-04775-
518 y](http://link.springer.com/10.1007/s11104-020-04775-y).).
- 519 Guggenberger, G. & Kaiser, K. 2003. Dissolved organic matter in soil : challenging the paradigm
520 of sorptive preservation. *Geoderma*, **113**, 293–310, (At: [https://doi.org/10.1016/S0016-7061\(02\)
521 00366-X](https://doi.org/10.1016/S0016-7061(02)00366-X).).

- Guigue, J., Lévêque, J., Mathieu, O., Schmitt-Kopplin, P., Lucio, M., Arrouays, D., Jolivet, C., Dequiedt, S., Chemidlin Prévost-Bouré, N. & Ranjard, L. 2015. Water-extractable organic matter linked to soil physico-chemistry and microbiology at the regional scale. *Soil Biology and Biochemistry*, **84**, 158–167, (At: <http://linkinghub.elsevier.com/retrieve/pii/S0038071715000619>).).
- Hapca, S., Baveye, P.C., Wilson, C., Lark, R.M. & Otten, W. 2015. Three-dimensional mapping of soil chemical characteristics at micrometric scale by combining 2D SEM-EDX data and 3D X-ray CT images. *PLoS ONE*, **10**, e0137205, (At: <http://dx.plos.org/10.1371/journal.pone.0137205>).).
- Hassink, J. 1995. Density fractions of soil macroorganic matter and microbial biomass as predictors of C and N mineralization. *Soil Biology and Biochemistry*, **27**, 1099–1108, (At: [https://doi.org/10.1016/0038-0717\(95\)00027-C](https://doi.org/10.1016/0038-0717(95)00027-C)).).
- Heinze, S., Ludwig, B., Piepho, H.-P., Mikutta, R., Don, A., Wordell-Dietrich, P., Helfrich, M., Hertel, D., Leuschner, C., Kirfel, K., Kandeler, E., Preusser, S., Guggenberger, G., Leinemann, T. & Marschner, B. 2018. Factors controlling the variability of organic matter in the top- and subsoil of a sandy Dystric Cambisol under beech forest. *Geoderma*, **311**, 37–44, (At: <https://linkinghub.elsevier.com/retrieve/pii/S001670611731279X>).).
- Heitkötter, J. & Marschner, B. 2018. Is There Anybody Out There? Substrate Availability Controls Microbial Activity outside of Hotspots in Subsoils. *Soil Systems*, **2**, 35, (At: <http://www.mdpi.com/2571-8789/2/2/35>).).
- Hobley, E.U., Brereton, A.J. & Wilson, B. 2016. Soil charcoal prediction using attenuated total reflectance mid-infrared spectroscopy. *Soil Research*, **55**, 86–92, (At: <http://www.publish.csiro.au/?paper=SR16068>).).
- Hobley, E.U., Steffens, M., Bauke, S.L. & Kögel-Knabner, I. 2018. Hotspots of soil organic carbon storage revealed by laboratory hyperspectral imaging. *Scientific Reports*, **8**, 13900, (At: <http://www.nature.com/articles/s41598-018-31776-w>).).
- Hothorn, T., Bühlmann, P., Dudoit, S., Molinaro, A. & Laan, M. van der. 2006. Survival ensembles. *Biostatistics*, **7**, 355–373, (At: <https://academic.oup.com/biostatistics/article-lookup/doi/10.1093/biostatistics/kxj011>).).
- Kaiser, K. & Kalbitz, K. 2012. Cycling downwards – dissolved organic matter in soils. *Soil Biology and Biochemistry*, **52**, 29–32, (At: <http://linkinghub.elsevier.com/retrieve/pii/S0038071712001381>).).
- Kalbitz, K., Solinger, S., Park, J.H., Michalzik, B. & Matzner, E. 2000. Controls on the dynamics of dis-

solved organic matter in soils: a review. *Soil Science*, **165**, 277–304, (At: https://journals.lww.com/soilsci/Abstract/2000/04000/Controls_on_the_Dynamics_of_Dissolved_Organic.1.aspx).

Kolen, J.F. & Pollack, J.B. 1991. Back propagation is sensitive to initial conditions. In: *Advances in neural information processing systems* (eds. Lippmann, R.P., Moody, J.E. & Touretzky, D.S.), pp. 860–867. 3rd ed. Morgan Kaufmann, San Mateo, California.

Kravchenko, A.N., Guber, A.K., Razavi, B.S., Koestel, J., Quigley, M.Y., Robertson, G.P. & Kuzyakov, Y. 2019. Microbial spatial footprint as a driver of soil carbon stabilization. *Nature Communications*, **10**, 3121, (At: <http://www.nature.com/articles/s41467-019-11057-4>).

Kögel-Knabner, I. 2002. The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter. *Soil Biology and Biochemistry*, **34**, 139–162.

Kögel-Knabner, I. 1997. ¹³C and ¹⁵N NMR spectroscopy as a tool in soil organic matter studies. *Geoderma*, **80**, 243–270.

Lal, R. 2016. Soil health and carbon management. *Food and Energy Security*, **5**, 212–222, (At: <http://doi.wiley.com/10.1002/fes3.96>).

Lavallee, J.M., Soong, J.L. & Cotrufo, M.F. 2020. Conceptualizing soil organic matter into particulate and mineral-associated forms to address global change in the 21st century. *Global Change Biology*, **26**, 261–273, (At: <https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.14859>).

Lucas, M., Pihlap, E., Steffens, M., Vetterlein, D. & Kögel-Knabner, I. 2020. Combination of imaging infrared spectroscopy and X-ray computed microtomography for the investigation of bio- and physicochemical processes in structured soils. *Frontiers in Environmental Science*, **8**, 42, (At: <https://www.frontiersin.org/article/10.3389/fenvs.2020.00042/full>).

Maillard, É. & Angers, D.A. 2014. Animal manure application and soil organic carbon stocks: a meta-analysis. *Global Change Biology*, **20**, 666–679, (At: <http://doi.wiley.com/10.1111/gcb.12438>).

Matschullat, J., Reimann, C., Birke, M., Santos Carvalho, D. dos, Albanese, S., Anderson, M., Baritz, R., Batista, M.J., Bel-Ian, A., Cicchella, D., Demetriades, A., De Vivo, B., De Vos, W., Dinelli, E., Ďuriš, M., Dusza-Dobek, A., Eggen, O.A., Eklund, M., Ersten, V., Fabian, K., Filzmoser, P., Flight, D.M.A., Forrester, S., Fügedi, U., Gilucis, A., Gosar, M., Gregorauskiene, V., De Groot, W., Gulan, A., Halamić, J., Haslinger, E., Hayoz, P., Hoogewerff, J., Hrvatovic, H., Husnjak, S., Jähne-Klingberg, F., Janik, L., Jordan, G., Kaminari, M., Kirby, J., Klos, V., Kwećko, P., Kutí, L., Ladenberger, A., Lima, A., Locutura, J., Lucivjansky, P., Mann, A., Mackovych, D., McLaughlin, M., Malýuk, B.I., Maquil, R., Meuli, R.G., Mol, G., Négrel, P., O'Connor, P., Oorts, K., Ottesen, R.T., Pasniecna,

A., Petersell, V., Pfeleiderer, S., Poňavič, M., Prazeres, C., Radusinović, S., Rauch, U., Sadeghi, M.,
 Salpeteur, I., Scanlon, R., Schedl, A., Scheib, A., Schoeters, I., Šefčík, P., Sellersjö, E., Slaninka, I.,
 Soriano-Disla, J.M., Šorša, A., Svrkota, R., Stafilov, T., Tarvainen, T., Tendavilov, V., Valera, P.,
 Verougstraete, V., Vidojević, D., Zissimos, A. & Zomeni, Z. 2018. GEMAS: CNS concentrations and
 C/N ratios in European agricultural soil. *Science of The Total Environment*, **627**, 975–984, (At: <https://linkinghub.elsevier.com/retrieve/pii/S0048969718302559>).

Meijboom, F.W., Hassink, J. & Noordwijk, M. van. 1995. Density fractionation of soil macroorganic
 matter using silica suspensions. *Soil Biology and Biochemistry*, **27**, 1109–1111.

Mueller, C.W. & Koegel-Knabner, I. 2009. Soil organic carbon stocks, distribution, and composition
 affected by historic land use changes on adjacent sites. *Biology and Fertility of Soils*, **45**, 347–359, (At: <http://link.springer.com/10.1007/s00374-008-0336-9>).

Mueller, C.W., Steffens, M. & Buddenbaum, H. 2021. Permafrost soil complexity evaluated by laboratory
 imaging Vis-NIR spectroscopy. *European Journal of Soil Science*, **72**, 114–119, (At: <http://doi.wiley.com/10.1111/ejss.12927>).

Nicolardot, B., Recous, S. & Mary, B. 2001. Simulation of C and N mineralisation during crop residue
 decomposition: A simple dynamic model based on the C:N ratio of the residues. *Plant and Soil*, **228**,
 83–103.

Pascual, N., Ranjard, L., Kaisermann, A., Bachar, D., Christen, R., Terrat, S., Mathieu, O., Lévêque,
 J., Mougél, C., Henault, C., Lemanceau, P., Péan, M., Boiry, S., Fontaine, S. & Maron, P.-A. 2013.
 Stimulation of different functional groups of bacteria by various plant residues as a driver of soil priming
 effect. *Ecosystems*, **16**, 810–822, (At: <http://link.springer.com/10.1007/s10021-013-9650-7>).

Robertson, G.P. & Groffman, P.M. 2015. Nitrogen transformations. In: *Soil microbiology, ecology and
 biochemistry* (ed. Paul, E.A.), pp. 421–446. 4th ed. Academic Press, Burlington, Massachusetts, USA.

Rumpel, C. & Kögel-Knabner, I. 2011. Deep soil organic matter—a key but poorly understood compo-
 nent of terrestrial C cycle. *Plant and Soil*, **338**, 143–158, (At: <http://link.springer.com/10.1007/s11104-010-0391-5>).

Sanderman, J., Baldock, J.A., Dangal, S.R.S., Ludwig, S., Potter, S., Rivard, C. & Savage, K. 2021.
 Soil organic carbon fractions in the Great Plains of the United States: an application of mid-infrared
 spectroscopy. *Biogeochemistry*, (At: <http://link.springer.com/10.1007/s10533-021-00755-1>).

Scheu, S. 1992. Decomposition of lignin in soil microcompartments: A methodical study with three
 different C-14-labelled lignin substrates. *Biology and Fertility of Soils*, **13**, 160–164, (At: <https://>

link.springer.com/article/10.1007/BF00336273.).

Sisouvanh, P., Trelo-ges, V., Isarangkool Na Ayutthaya, S., Pierret, A., Nunan, N., Silvera, N., Xayyathip, K. & Hartmann, C. 2021. Can organic amendments improve soil physical characteristics and increase maize performances in contrasting soil water regimes?. *Agriculture*, **11**, 132, (At: <https://www.mdpi.com/2077-0472/11/2/132>.).

Sorenson, P.T., Quideau, S.A., Rivard, B. & Dyck, M. 2020. Distribution mapping of soil profile carbon and nitrogen with laboratory imaging spectroscopy. *Geoderma*, **359**, 113982, (At: <https://linkinghub.elsevier.com/retrieve/pii/S0016706119302459>.).

Steffens, M. & Buddenbaum, H. 2013. Laboratory imaging spectroscopy of a stagnic Luvisol profile - High resolution soil characterisation, classification and mapping of elemental concentrations. *Geoderma*, **195-196**, 122-132, (At: <http://dx.doi.org/10.1016/j.geoderma.2012.11.011>.).

Steffens, M., Kohlpaintner, M. & Buddenbaum, H. 2014. Fine spatial resolution mapping of soil organic matter quality in a Histosol profile. *European Journal of Soil Science*, **65**, 827-839, (At: <http://doi.wiley.com/10.1111/ejss.12182>.).

Stone, A.G., Traina, S.J. & Hoitink, H.A.J. 2001. Particulate organic matter composition and pythium damping-off of cucumber. *Soil Science Society of America Journal*, **65**, 761-770, (At: <http://doi.wiley.com/10.2136/sssaj2001.653761x>.).

Strauss, E.A. & Lamberti, G.A. 2002. Effect of dissolved organic carbon quality on microbial decomposition and nitrification rates in stream sediments. *Freshwater Biology*, **47**, 65-74, (At: <http://doi.wiley.com/10.1046/j.1365-2427.2002.00776.x>.).

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T. & Zeileis, A. 2008. Conditional variable importance for random forests. *BMC Bioinformatics*, **9**, 307, (At: <http://www.biomedcentral.com/1471-2105/9/307>.).

Strobl, C., Boulesteix, A.-L., Zeileis, A. & Hothorn, T. 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, **8**, 25, (At: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-8-25>.).

Taylor, J.P., Wilson, B., Mills, M.S. & Burns, R.G. 2002. Comparison of microbial numbers and enzymatic activities in surface soils and subsoils using various techniques. *Soil Biology and Biochemistry*, **34**, 387-401, (At: <https://linkinghub.elsevier.com/retrieve/pii/S0038071701001997>.).

Verbruggen, E., Jansa, J., Hammer, E.C. & Rillig, M.C. 2016. Do arbuscular mycorrhizal fungi stabilize litter-derived carbon in soil?. *Journal of Ecology*, **104**, 261-269, (At: <http://doi.wiley.com/10.1111/>

1365–2745.12496.).

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J. & Skjemstad, J.O. 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, **131**, 59–75, (At: <https://doi.org/10.1016/j.geoderma.2005.03.007>).

Wagai, R., Mayer, L.M. & Kitayama, K. 2009. Nature of the “occluded” low-density fraction in soil organic matter studies: A critical review. *Soil Science and Plant Nutrition*, **55**, 13–25, (At: <http://dx.doi.org/10.1111/j.1747-0765.2008.00356.x>).

Warren, G.P. & Whitehead, D.C. 1988. Available soil nitrogen in relation to fractions of soil nitrogen and other soil properties. *Plant and Soil*, **112**, 155–165, (At: <http://link.springer.com/10.1007/BF02139991>).

Wilson, M.A., Heng, S., Goh, K.M., Pugmire, R.J. & Grant, D.M. 1983. Studies of litter and acid insoluble soil organic matter fractions using ¹³C-cross polarization nuclear magnetic resonance spectroscopy with magic angle spinning. *Journal of Soil Science*, **34**, 83–97, (At: <http://doi.wiley.com/10.1111/j.1365-2389.1983.tb00815.x>).

Xu, S., Wang, M. & Shi, X. 2020. Hyperspectral imaging for high-resolution mapping of soil carbon fractions in intact paddy soil profiles with multivariate techniques and variable selection. *Geoderma*, **370**, 114358, (At: <https://linkinghub.elsevier.com/retrieve/pii/S0016706118323838>).

Willingen, P. de, Janssen, B.H., Heesmans, H.I.M., Conjin, J.G., Velthof, G.J. & Chardon, W.J. 2008. *Decomposition and accumulation of organic matter in soil; comparison of some models*. Wageningen, The Netherlands. (At: <http://library.wur.nl/WebQuery/wurpubs/fulltext/15401>).

Wesemael, B. van, Chartin, C., Wiesmeier, M., Lützow, M. von, Hobley, E.U., Carnol, M., Krüger, I., Campion, M., Roisin, C., Hennart, S. & Kögel-Knabner, I. 2019. An indicator for organic matter dynamics in temperate agricultural soils. *Agriculture, Ecosystems & Environment*, **274**, 62–75, (At: <https://doi.org/10.1016/j.agee.2019.01.005>).

Lützow, M. von, Kögel-Knabner, I., Ludwig, B., Matzner, E., Flessa, H., Ekschmitt, K., Guggenberger, G., Marschner, B. & Kalbitz, K. 2008. Stabilization mechanisms of organic matter in four temperate soils: Development and application of a conceptual model. *Journal of Plant Nutrition and Soil Science*, **171**, 111–124, (At: <http://doi.wiley.com/10.1002/jpln.200700047>).

IUSS Working Group WRB. 2015. *World Reference Base for Soil Resources 2014, update 2015 International soil classification system for naming soils and creating legends for soil maps*. (AT Lebedev, Ed.).

677 World Soil Resources Report No 106, FAO, Rome. (At: [http://www.fao.org/publications/card/](http://www.fao.org/publications/card/en/c/942e424c-85a9-411d-a739-22d5f8b6cc41/)
678 [en/c/942e424c-85a9-411d-a739-22d5f8b6cc41/](http://www.fao.org/publications/card/en/c/942e424c-85a9-411d-a739-22d5f8b6cc41/)).

679 R Core Team. 2021. *R: A language and environment for statistical computing*. R Foundation for
680 Statistical Computing, Vienna, Austria. (At: <http://www.r-project.org/>).