



Origin of carbon in agricultural soil profiles deduced from depth gradients of C:N ratios, carbon fractions, $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values

Florian Schneider · Wulf Amelung · Axel Don

Received: 1 May 2020 / Revised: 6 November 2020 / Accepted: 12 November 2020 / Published online: 4 December 2020
© The Author(s) 2020

Abstract

Aims Agricultural soils in Germany store 2.54 Pg of organic carbon (C). However, information about how and when this C entered the soils is limited. This study illustrates how depth profiles of organic matter can shed light on different entry paths of organic C.

Methods Machine learning was used to explain total organic C (TOC), C:N, particulate organic C (POC), $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values down to 100 cm depth based on pedology, geology, climate and management-related variables from the German Agricultural Soil Inventory. We estimated TOC turnover rates based on the relationship between the proportion of maize (only C_4 plant) in crop rotations and soil $\delta^{13}\text{C}$ values.

Results In the upper 30 cm of cropland, fresh photosynthates added on average 0.2 to 0.8 Mg C $\text{ha}^{-1} \text{year}^{-1}$. Organic fertiliser was another source of topsoil C, especially in grassland. Sandy sites in north-west Germany contained historic C from past heathland and peatland. One third of German agricultural land was found to be

on colluvial and alluvial deposits, in which allochthonous C from upstream and upslope areas evidently increased the TOC content of subsoils. In and below hardpans, TOC content and C:N and POC:TOC ratios were low, indicating restricted root-derived C input.

Conclusions Our data indicate that ongoing management in German agricultural soils mainly affects topsoil C, while C storage in subsoils reveals significant legacies from allochthonous, buried or translocated C inputs. Specific attention should be focused on the sustainable loosening of hardpans that could result in a slow, but significant increase in subsoil C stocks.

Abbreviations

C:N	Total organic carbon to total nitrogen ratio.
MOC	Mineral-associated organic carbon.
POC	Particulate organic carbon.
R^2	Coefficient of determination.
RMSE	Root mean square error.
TOC	Total organic carbon.
TIC	Total inorganic carbon.

Responsible Editor: Lucas Silva.

F. Schneider (✉) · A. Don
Thünen Institute of Climate-Smart Agriculture, Bundesallee 65,
38116 Braunschweig, Germany
e-mail: florian.schneider@thuenen.de

W. Amelung
Institute of Crop Science and Resource Conservation – Soil
Science and Soil Ecology, University of Bonn, Nußallee 13,
53115 Bonn, Germany

Introduction

Globally, soils contain about 1,500 Pg of carbon bound within soil organic matter in the upper 100 cm (Batjes 1996). This is about twice the amount of carbon (C) that currently resides in the Earth's atmosphere as carbon dioxide (IPCC 2013). Nevertheless, soil organic C

content tends to decrease with greater depth, following the depth distribution of organic C input and its turnover. In agricultural soils, primary organic C inputs may originate from: (i) aboveground biomass that is not harvested, such as stubble, mulch and green manure, (ii) organic fertilisers including animal excreta, compost and biogas digestates, and (iii) root litter and rhizodeposits. The first two sources of organic C enter soil at the surface, with redistribution along the soil profile possibly only occurring via bioturbation, tillage and/or leaching of mobile organic C species. Root-derived C, however, enters soil directly belowground to a depth of one metre or more (Canadell et al. 1996). The relative contribution of aboveground biomass, organic fertiliser and roots to total C inputs varies considerably between different land uses and management regimes. Grassland, for instance, receives more C input via roots and rhizodeposits than cropland (Pausch and Kuzyakov 2018), while cropland receives a larger C input from aboveground in the form of harvest residues and stubble.

After C has entered the soil, it is prone to a range of transformation processes (Kögel-Knabner and Amelung 2014). Transformation results in significant contributions of microbial debris to total organic C (TOC) (Appuhn and Joergensen 2006; Liang et al. 2019; Miltner et al. 2012). Microbial transformations of TOC are usually accompanied by stable isotope discrimination processes. Undecomposed photosynthates from annual crops and grasses typically show $\delta^{13}\text{C}$ signals from -26 to -30‰ for C_3 plants and -10 to -14‰ for C_4 plants (Philp and Monaco 2012), as well as $\delta^{15}\text{N}$ signals from -8 to 9‰ depending on the nitrogen source (Craine et al. 2015). During microbial processing, the organic C fraction is enriched with heavy isotopes of C and N, which is reflected in increasing $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values (Boutton 1996; Natelhoffer and Fry 1988). If organic material is processed and decomposed, its C:N ratio also decreases from > 15 to the C:N ratio of microbes ranging from 5 to 8 (Amelung et al. 2018). This makes $\delta^{13}\text{C}$, $\delta^{15}\text{N}$ and C:N values proxies for characterising the degree of microbial processing of organic C and detecting the presence of relatively undecomposed litter. The latter can also be quantified directly by density fractionation, which separates soil organic C into particulate organic C (POC) and mineral-associated C (MOC).

To date, most work on C fractions, $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values, and C:N ratios has been conducted at

field or sub-field scale and focused on topsoil because: (i) organic matter in topsoil is more sensitive to land use or management treatments than subsoil, and (ii) measuring and detecting changes in POC, $\delta^{13}\text{C}$, $\delta^{15}\text{N}$ and C:N signatures of subsoil can be analytically challenging as subsoil typically contains much less organic matter than topsoil. Although globally more than half of the soil organic C in the upper metre is stored below 30 cm depth (Batjes 1996), much less is known about the origin of C in subsoil than in topsoil. Carbon input into subsoils may come from the surface soil and enter subsoil via leaching (Kindler et al. 2011), turbation (bio-, cryo-, or pelleturbation), or burial of topsoil C, depending on pedogenesis and climatic regions. However it may also come from the soil parent material, especially if the soil developed from relatively young and C-rich colluvial deposits (Doetterl et al. 2012), or directly from deep roots and their rhizodeposits.

Radiocarbon ages in soil profiles suggest that subsoil C-cycling is much slower than topsoil C-cycling. Once C has entered the subsoil, much of it seems to reside there for millennia (Balesdent et al. 2018). Crops deposit about one third of all root C below 30 cm depth (Jackson et al. 1996), but this number varies considerably in time and space. Given the long residence time of subsoil C, increasing root-derived C input in the subsoil on a large scale could remove significant amounts of C from the atmosphere. In this context, the breeding of plants with deep and bushy roots has been proposed as a negative emission technology for counteracting greenhouse gas-induced global warming (Kell 2011; 2012). However, there are various soil properties that restrict the growth of deep roots, such as compacted soil layers, which are commonly referred to as hardpans (Lynch and Wojciechowski 2015). Oldeman et al. (1991) projected the global extent of hardpans to be 68 million hectares and attributed hotspots of soil compaction in Europe to heavy machinery. For Germany, Schneider and Don (2019) recently estimated half of all cropland to be compacted due to either pedogenetic causes (37% of cropland) or anthropogenic causes (13% of cropland). However, the effect of hardpans on subsoil C has received limited scrutiny to date.

A better understanding of the origin of C in agricultural soils could improve the design of future climate-smart management practices for enhanced C sequestration. Here, we hypothesised that depth gradients of soil organic C reflect past organic C inputs and allow to

track the contribution of deep roots to subsoil C, unless subsoils contain large amounts of topsoil C or parent material-inherent C. We assumed that deep input of root C correlates positively with the depth gradients of TOC, C:N and POC:TOC, but negatively with $^{13}\text{C}:^{12}\text{C}$ and $^{15}\text{N}:^{14}\text{N}$ abundances, as elevated portions of undecomposed root material in the subsoil should also be reflected in lower $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values. We also assumed the depth distribution of root-derived C to be controlled by soil properties and management. We use Germany as a model country, aligning novel data with data from the first German Agricultural Soil Inventory (Poeplau et al. 2020) to evaluate the depth distribution of soil organic matter with respect to recent and historic C inputs into agricultural soils at regional scale. Specifically, our aims were to:

- describe the depth gradients of TOC, C:N ratio, POC, $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ in the upper metre and relate the observed patterns to physico-chemical characteristics of the soil matrix, parent material, climate, land use and/or management using a data-mining approach.
- investigate entry pathways and turnover of organic C in topsoils and subsoils.
- estimate the effect of hardpans in croplands on root-derived C input in subsoil.

Materials and methods

Study area

This study focused on agricultural soils in Germany, which cover an area of 166,451 km² (Destatis 2019). About 70% of this area is used for annual crops (117,309 km²), 28% is used as permanent grassland (47,134 km²), and the remaining agricultural land is used for perennial crops (Destatis 2019). Based on FAOSTAT (2019), the top three crops grown between 2009 and 2018 were wheat (*Triticum aestivum* L., mean grain yield: 7.6 t ha⁻¹), followed by barley (*Hordeum vulgare* L., mean grain yield 6.5 t ha⁻¹) and canola (*Brassica napus* L., mean grain yield 3.7 t ha⁻¹). German grassland is typically managed intensively and includes 1.4 times more pasture than meadow (Destatis 2019).

Sampling design & data acquisition

This study was based on the dataset from the first German Agricultural Soil Inventory, which comprises information on the soil and management at 3,104 sites covering all agricultural land in Germany in a regular 8 km x 8 km grid (Poeplau et al. 2020). At each site, soil profiles were dug to 100 cm depth. Soil morphology was characterised based on soil horizons following AD-HOC-AG Boden (2005), while composite soil samples for laboratory analysis were taken at fixed depth increments (0–10, 10–30, 30–50, 50–70 and 70–100 cm). If a horizon boundary was at least 5 cm above or below a sampling boundary, an additional soil sample was taken. This allowed laboratory and field data to be merged. Total organic C (TOC), total inorganic C (TIC) and total nitrogen were measured by dry combustion using an elemental analyser (LECO TRUMAC, St Joseph, MI, USA). Sand content was determined by wet-sieving, and silt and clay contents following the pipette method. Bulk density, rock fragment fraction and pH_{H2O} were measured as described by Jacobs et al. (2018).

As processes governing C stabilisation in organic soil differ from those in mineral soils, and as root-derived C inputs contribute only little to C stocks of organic soil, this study focused solely on mineral soils. Therefore, 165 sites with organic soil were removed from subsequent analyses, leaving 15,935 mineral soil samples from 2,939 sites for which TOC and C:N values were readily available. For the present study, $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ natural abundance as well as particulate organic C to total organic C (POC:TOC) ratios were also determined. The $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ measurements were restricted to 1,357 soil samples from 248 core sites. These core sites comprise a representative subset of mineral agricultural soils in Germany with respect to soil reference group, land use and C stock. Detailed criteria for the core site selection are described by Vos et al. (2018). Isotopes were analysed using an isotope-ratio mass spectrometer (IRMS, Thermo Fisher Scientific Delta plus). All $\delta^{13}\text{C}$ values refer to the organic C fraction only. Soil samples containing inorganic C were fumigated with hydrochloric acid, as described by Walthert et al. (2010), prior to $\delta^{13}\text{C}$ analyses. If total C content based on IRMS was higher than TOC based on dry combustion, inorganic C removal was repeated and $\delta^{13}\text{C}$ re-measured. For 183 samples, no reliable $\delta^{13}\text{C}$ measurements could be obtained for the organic C, mostly due to incomplete removal of inorganic C at TIC:TOC ratios over 5.

Values of $\delta^{15}\text{N}$ refer to total N. The $\delta^{13}\text{C}$ values are reported relative to Vienna Pee Dee Belemnite and $\delta^{15}\text{N}$ values are reported relative to atmospheric nitrogen following international standards.

POC:TOC ratios were determined using near-infrared spectroscopy (FT-NIRS; MPA, Bruker Optik GmbH, Ettlingen, Germany). The basic concept for this approach has recently been proven by Jaconi et al. (2019) who also used soil samples from the German Agricultural Soil Inventory, but focused on topsoil (0–10 cm) only. Here, we present POC:TOC ratios for all depth increments of the German Agricultural Soil Inventory. Our predictions of POC:TOC ratios are based on the same 105 topsoil samples used by Jaconi et al. (2019), plus additional subsoil samples from 27 of these 105 sites. The 27 subsoil samples were selected to cover the complete range of clay content and C:N ratios of subsoils in order to get a representative subset of the subsoil samples (Fig. A1.1). For each of these 27 subsoil samples, TOC was fractionated into POC and MOC using density fractionation. In brief, soil was dispersed in a sodium polytungstate solution with a density of 1.6 g cm^{-3} using ultrasonic power. POC was defined as the light C fraction with a density of $< 1.6\text{ g cm}^{-3}$ while MOC was defined as the remaining, heavier C fraction. We followed exactly the fractionation protocol, which was previously described by Jaconi et al. (2019) for topsoil. However, because the POC content of subsoil was about one order of magnitude lower than in the topsoil, up to 120 g soil instead of 10 g for topsoils were fractionated. This was time-consuming and cost-intensive (up to 620 g sodium polytungstate per sample), therefore not all the sites included in the study of Jaconi et al. (2019) could be considered for subsoil fractionation. TOC recovery for subsoil samples ranged from 80 to 120%. Spectral pre-treatments and models were implemented as proposed by Jaconi et al. (2019). Additional tree-based algorithms (ranger, cubist) and modelling approaches (log-ratio transformation of POC and MOC content vs. log-ratio transformation of POC:TOC and POC:MOC, separate prediction of POC and MOC contents vs. separate prediction of POC:TOC and MOC:TOC ratios) were also tested, as detailed in the Supplementary Material. Model performance was judged based on leave-one-site-out validation, i.e. the data was partitioned by sites and the values of each site were predicted based on models that were trained only based on the samples from the other sites (Jaconi et al., 2019). This validation procedure is more robust than partitioning the dataset just once into

training and testing (Hastie et al., 2009). An ensemble of three individual model combinations performed best (Fig. A1.2) and was used to predict the POC:TOC ratios of all unseen soil samples.

In order to enhance the comparability among sites, all soil data were aggregated to obtain one single value per depth increment (0–10 cm, 10–30 cm, 30–50 cm, 50–70 cm and 70–100 cm). This was done as follows: TOC and TIC concentrations, C:N and POC:TOC ratios, $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values, as well as contents of clay and sand, and $\text{pH}_{\text{H}_2\text{O}}$ values were aggregated by weighted means, with fine soil stock of the individual depth increments ($\text{Mg soil} < 2\text{ mm ha}^{-1}$) as the weighting factor. For continuous variables given in vol-% (rock fragment fraction) or in area-% (e.g. oximorphic features at the profile wall), the thickness of the respective depth increment served as the weighting factor. Categorical variables with only two levels (e.g. buried topsoil, yes or no) were dummy coded (yes = 1, no = 0) and also aggregated, with the thickness of the respective depth increment serving as the weighting factor. For categorical variables with more than two levels (soil horizon and stratigraphy), the factor level with the largest contribution to the depth increment was used. Information on crop rotations, tillage and management during the past decade came from farmer questionnaires. Site-specific C input data was estimated by Jacobs et al. (2020).

Analyses of organic C proxies along the soil profile

Each organic C proxy (TOC, C:N, POC:TOC, $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$) was predicted using machine learning on variables that provided information about current C input (e.g. root-derived C), geogenic C input (e.g. parent material), translocated topsoil C (e.g. horizon symbols), soil transport (e.g. slope) and C stabilisation (e.g. texture). Furthermore, we included climate-related explanatory variables because climate is a major soil-forming factor and can have various effects on C input and C stabilisation. Table 1 provides an overview of all 39 explanatory variables used. In the Supplementary Material, these explanatory variables are characterized in further detail (Table A2.1, Fig. A2.1, Fig. A2.2).

In a first step, the machine-learning algorithm was trained to predict the bulk soil value of each organic C proxy (TOC, C:N, POC:TOC, $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$) at a given depth solely as a function of these 39 explanatory variables. The resulting 25 models (five organic C proxies x five depth increments) were independent of the soil

Table 1 Overview of explanatory variables used to predict organic C proxies. For further details, German translations and maps of these explanatory variables, the reader is referred to Table A2.1, Fig. A2.1 and Fig. A2.2 in the Supplementary Material

Name	Type*	Explanation
Climate		
Pedo-climatic zone	F(15)	Pedo-climatic zones
Precipitation	C	Multi-annual mean of precipitation (mm year^{-1})
Summer days	C	Number of summer days (-)
Sunshine	C	Multi-annual mean of sunshine duration (hours year^{-1})
Temperature	C	Multi-annual monthly averaged daily minimum air temperature in 2 m height ($^{\circ}\text{C}$)
Geology		
Age of parent material	C	Age of the soil parent material (Ma)
Chronostratigraphy	F(3)	Chronostratigraphic units: Holocene, Pleistocene or older
Curvature1	F(3)	Vertical surface curvature: none, convex or concave
Curvature2	F(3)	Horizontal surface curvature: none, convex or concave
Elevation	C	Elevation above sea level (m)
Geomorphology	F(10)	Geomorphological terrain classes
Lithogeny	F(8)	Environmental conditions during lithogeny
Parent material	F(14)	Soil parent material
Petrology	F(9)	Petrology
Slope	C	Slope in degrees
Soil region	F(11)	Soil regions
Terrain	F(4)	Topographical relief
Management		
C input: fertiliser	C	Mean C input from organic fertilizer ($\text{t ha}^{-1} \text{year}^{-1}$)
C input: roots	C	Mean C input from roots and rhizodeposits below ground ($\text{t ha}^{-1} \text{year}^{-1}$)
C input: total	C	Mean total C input from above and below ground ($\text{t ha}^{-1} \text{year}^{-1}$)
Canola, beet and legumes	C	Share of canola (<i>Brassica napus</i>), beet (mostly <i>Beta vulgaris</i>) and legumes (mostly <i>Trifolium</i> sp.) in crop rotation (-)
Land use	F(3)	Annual crops, perennial crops or grassland
Livestock	C	Livestock units per hectare
Maize	C	Share of maize in the crop rotation of the past 10 years (%)
N input: fertiliser	C	Mean N input from fertilizers ($\text{t ha}^{-1} \text{year}^{-1}$)
Soil transport	F(3)	No soil transportation, soil accumulation (deposition) or abrasion (removal of earth)
Soil		
Biopores	C	Biopore density (vol-%)
Clay	C	Clay content (%)
Gr-horizon	C	Area percentage of profile wall with reducing conditions in > 300 days per year
Groundwater	F(5)	Mean depth to groundwater table
Oximorphic features	C	Area percentage of oximorphic features at profile wall
pH	C	pH in deionized water
Reductimorphic features	C	Area percentage of reductimorphic features at profile wall
Relic topsoil	C	Fossil or relic topsoil
Rock fragments	C	Rock fragment content (vol-%)
Sand	C	Sand content (%)
Soil group	F(11)	WRB Reference Soil Group

Table 1 (continued)

Name	Type*	Explanation
Soil horizon	F(11)	<p>Master soil horizon symbols based on AD-HOC-AG Boden (2005). Factor levels (English description based on Wittmann et al. 1997): A (mineral topsoil horizon; epipedon); B (mineral subsoil horizon characterized by a change of colour and mineral composition of the parent material resulting from the accumulation of topsoil constituents, which were removed from the overlying horizon or/and weathering in situ; contains less than 75 vol-% of residual parent rock and no lithogenic carbonate); C (unaltered or unweathered material similar to parent material); E (plaggen-horizon; humus content similar to Ah-horizon); G (horizon affected by groundwater, which causes the development of redoximorphic features); H (organic horizon originating from residue of peat-forming plants); M (horizon of fluvial, colluvial soils from translocated soil-material, humus content like Ah); P (subsoil horizon on claystone or clay (> 45% clay) with prismatic to subangular blocky structure and temporarily wide cracks); R (man-made horizon by deep cultivation (> 40 cm), humus content like Ah); S (subsoil horizon affected by a perched water table; surface water stagnation); T (subsoil horizon from the solution residue of carbonate rock, clay content > 65%, < 5% carbonate).</p>
TIC	C	Total inorganic carbon (g kg^{-1})

* C = continuous variable; F(L) = categorical variable with L levels

properties in the depth increments above or below. In a second step, tests were carried out to assess how much and to what depth model performances would improve if the algorithm was also informed about the measured values of the targeted organic C proxy in the uppermost sampling layer. Thus, TOC content from 10 to 100 cm depth was predicted, as was done in the first step, but here included site-specific TOC contents from 0 to 10 cm depth as an additional (40th) explanatory variable in the models. This was done in an analogous way to also predicting the other organic C proxies (C:N, POC:TOC, $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$). In the third and final step, the ratio between a given organic C proxy and its value in the uppermost sampling layer (0–10 cm) was predicted. Such ratios are referred to below as “normalised values” to distinguish them from “bulk soil values”. Normalised values were used to describe the depth gradients of organic C proxies. For example, if a site showed bulk soil TOC values of 1 g kg^{-1} in 50–70 cm and 5 g kg^{-1} in 0–10 cm, the normalised TOC value in 50–70 cm depth would be $1:5 = 0.2$ or 20%. Soil layers with normalised values above 300%, e.g. more than three times as much TOC in 50–70 cm than in 0–10 cm, were treated as outliers and excluded from further analyses. Furthermore, C:N ratios of carbonate-

rich soils with more than twice as much TIC as TOC were excluded because these measurements were potentially biased.

For machine learning, the party::cforest implementation of random forest and bagging ensemble algorithms was used (Hothorn et al. 2005; Strobl et al. 2008; Strobl et al. 2007) owing to its underlying strength in deriving variable importance from a mix of continuous and categorical predictors (Hapfelmeier et al. 2014). For each model, five folds with three repetitions were applied and the hyperparameter mtry was tuned using grid search on the following values: 3, 6, 13, 26, 39 or 59. The performance of each model was evaluated by root mean squared error (RMSE) and mean coefficients of determination (R^2) metrics. Permutation importance was calculated based on Hapfelmeier et al. (2014) using the mtry value that produced the highest R^2 . Variables of greater importance than would be expected from a theoretical model where all variables are equally important were considered influential (Hobley et al. 2015). The importance of influential variables was scaled to 100% for better comparability. Explanatory variables that correlated the most were elevation/temperature (Pearson's $r = -0.81$ in 0–10 cm depth) and sand/clay (Pearson's $r = -0.74$ in 0–10 cm depth). Because these Pearson's r

values were less extreme than the threshold values for collinearity applied in previous work (e.g. Hobley et al. 2015), in our study, collinearity was not regarded as a problematic issue.

Turnover and mean residence time of organic C

Maize (*Zea mays* L.) was the only C₄ plant in crop rotations, and as such introduced a natural ¹³C label into the soil. We estimated the fraction *F* of maize-derived C in TOC after *t* years of maize based on Balesdent et al. (1987) as:

$$F(t) = \frac{\delta^{13}C_{S1,t} - \delta^{13}C_{S0}}{\delta^{13}C_{P1} - \delta^{13}C_{P0}} \quad (1)$$

where S1 and S0 were the soils of sites with and without maize respectively and P1 and P0 were photosynthates from maize and from other crops respectively. $\delta^{13}C_{P1}$ was set to -12‰ (Balesdent et al. 1987). $\delta^{13}C_{S0}$ was defined as the intercept of a linear regression between the number of maize years (*t*) and the corresponding $\delta^{13}C$ values of soil ($\delta^{13}C_{S1,t}$). The slope of this regression times the number of maize years (*t*) plus $\delta^{13}C_{S0}$ yielded the $\delta^{13}C$ value of soil after *t* maize years ($\delta^{13}C_{S1,t}$). $\delta^{13}C_{P0}$ was assumed to be the same as $\delta^{13}C_{S0}$, which is a common simplification because the difference between C₃ and C₄-derived organic C is typically about one order of magnitude higher than turnover-induced C fractionation in soil (Balesdent et al. 1987; Boutton 1996). Carbon turnover was defined as the average increase in *F* per maize year.

Crop rotation was reported for a period spanning 1 to 17 years depending on the site. A pre-test revealed that using sites with crop rotations reported for ≤ 9 years resulted in overestimated organic C turnover (Fig. A3.1). Hence, all sites with only nine years of crop rotation data or less were omitted. Among the remaining sites (*n* = 124), only the past ten years were considered for calculating organic C turnover in order to increase the comparability among sites.

The mean residence time of organic C was estimated following Amelung et al. (2008) as $-t/\ln(1-F(t))$, where $1-F(t)$ represents the remaining proportion of C₃-C after *t* years of maize. This approach assumes: (i) TOC stocks to be in equilibrium and (ii) first-order kinetics for organic C decomposition (Balesdent and Mariotti 1996).

Soil compaction vs. root-derived deep C input

The level of soil compaction was classified according to Schneider and Don (2019) as not compacted (packing density $< 1.75 \text{ g cm}^{-3}$), moderately compacted ($1.75 \text{ g cm}^{-3} < \text{packing density} < 1.82 \text{ g cm}^{-3}$) or severely compacted (packing density $> 1.82 \text{ g cm}^{-3}$). A given site was classified as compacted at depth_{*i*} if a compacted soil layer occurred at or above depth_{*i*}. This was done in order to account for the fact that compacted soil layers (e.g. ploughpan at 30–50 cm) may act as a barrier to root-derived C inputs at greater depths (e.g. subsoil below ploughpan). Root-derived C inputs were inferred from normalised TOC, C:N, POC:TOC, $\delta^{13}C$ and $\delta^{15}N$ values, i.e. measured values of TOC, C:N, POC:TOC, $\delta^{13}C$ and $\delta^{15}N$ at a given depth_{*i*} divided by their respective value in the uppermost sampling layer (0–10 cm). Normalised values were preferred over bulk soil values because normalisation minimised site-specific differences in organic C. Other factors influencing the depth gradients of TOC, C:N, POC:TOC, $\delta^{13}C$ and $\delta^{15}N$, apart from root-derived C inputs, were assumed to be independent of the level of compaction. The relationship between root-derived C input and soil compaction was examined by depth (30–50 cm, 50–70 cm and 70–100 cm). Topsoil was not considered because the majority of hardpans occurred in subsoils. The analyses were restricted to cropland soil only (2,261 sites) because compaction in cropland was more severe than in grassland.

General statistics & software

All data analysis was performed using R (R Core Team 2019) in RStudio (RStudio Team 2019) and built on tidyverse packages (Wickham et al. 2019). Computationally intensive code was implemented in parallel using the foreach package (Microsoft and Weston 2020) and run on a high-performance server with 64 cores (128 threads). Spectral pre-treatments were done using prospectr (Stevens and Ramirez-Lopez 2013), while pls (Mevik et al. 2019), ranger (Wright and Ziegler 2015) and cubist (Kuhn and Quinlan 2018) were used for spectral modelling. Each key statistic in the text is accompanied by its 95% bias-corrected and accelerated (BCa) confidence interval based on ≥ 3000 bootstrapped resamples calculated using the boot package (Canty and Ripley 2019).

Results¹

TOC

In 0–10 cm depth, TOC content averaged 23.5 g kg^{-1} (95% CI, 23.0 to 24.1) and its interquartile (middle 50%) ranged from 12.9 to 29.6 g kg^{-1} . Variation in TOC at this uppermost depth increment was mostly caused by different land use (cropland vs. grassland) and clay content (Figs. 1 and 2). In 0–10 cm, the TOC content in grassland was 2.4 times (95% CI, 2.3 to 2.5) greater than that in cropland and increased significantly with clay content (Fig. A3.2). Land use and soil texture showed distinct regional patterns, which were reflected in the TOC map (Fig. 1): regions with dominant grassland use in south and north-west Germany were associated with elevated TOC values, while large areas with low clay content in north-east Germany showed low TOC values (Fig. A2.1, Fig. A2.2).

With increasing soil depth, the importance of variables to explain TOC content changed considerably (Fig. 2). Instead of land use, climate-related variables were the most important for predicting TOC in 10–30 cm depth. TOC content correlated positively with precipitation and negatively with the number of summer days (Fig. 2). The “Weser-Ems-Geest” climatic region in north-west Germany, which is characterised by a maritime climate (high precipitation, not many summer days) and historic heathland and peatland cover, showed the largest TOC content (Fig. 3d). In 30–70 cm depth, soil reference group and horizon symbols were most important for explaining TOC. For example, in 50–70 cm, Anthrosols contained 8.4 g kg^{-1} (95% CI, 7.3 to 9.8), Chernozems, Phaeozems and Vertisols contained $5.8 \text{ g TOC kg}^{-1}$ (95% CI, 5.5 to 6.2), while all other soil groups contained only $3.1 \text{ g TOC kg}^{-1}$ (95% CI, 2.9 to 3.2) (Figs. 3a and 4). At the same depth, M horizons designating colluvial deposits (mean 7.2 g kg^{-1} ; 95% CI, 6.5 to 8.0) and deep or buried A horizons (mean 5.8 g kg^{-1} ; 95% CI, 4.8 to 7.1) showed significantly larger TOC contents than other soil horizons (mean 3.5 g kg^{-1} ; 95% CI, 3.4 to 3.7) (Fig. 3b). Finally, below 70 cm depth, geological chronostratigraphy

and texture, both characterising the soil parent material, were the most important for predicting the contents of TOC in this subsoil depth. Holocenic sediments, which comprised mostly fluvial deposits along major river valleys (Elbe and Rhine), marine deposits along the north-west coastline and colluvial deposits at the foot of slopes, showed TOC contents that were three times (95% CI, 2.6 to 3.4) greater in 70–100 cm depth than soils from other, older parent materials. As with topsoil, the content of clay in subsoil also correlated positively with that of TOC (Fig. A3.2). Although TOC content in grassland was significantly higher than that in cropland at all depth increments (Fig. 1), land use was not relevant for explaining TOC below 10 cm depth (Fig. 2). Different TOC contents in the subsoils of cropland and grassland were related to preferential grassland use on Gleysols, presumably because of traffic restrictions for heavy machinery and elevated flood risks. Many Gleysols developed from C-rich holocenic sediments (Fig. 5). Furthermore, Gleysols may accumulate C due to oxygen limitation – both processes mask potential effects of land use on TOC in subsoils. If the different TOC contents in subsoils of cropland and grassland were due to land use, the relative difference between the TOC contents of cropland and grassland should have decreased with depth. However, the opposite was the case: the relative difference was lowest in 10–30 cm, where TOC content under grassland was 1.27 times (95% CI, 1.21 to 1.33) higher than that under cropland. Below 30 cm, the relative difference between grassland and cropland increased again to 1.62 times (95% CI, 1.41 to 1.90) higher TOC contents under grassland than that under cropland in 70–100 cm soil depth. Overall, predictions for bulk TOC without information from other depth increments were better in 0–10 cm (R^2 of 0.70) than below this depth (R^2 from 0.39 to 0.45). Including the TOC contents of 0–10 cm as an additional predictor of TOC in > 10 cm depth improved model performance considerably, confirming that TOC contents were site-specific and that TOC in topsoil and subsoil was related (Fig. A3.3). The importance of TOC in 0–10 cm for explaining TOC below 10 cm depth decreased with greater depth.

Depth gradients of TOC, which were expressed as normalised TOC values, depended less on clay content than was the case with TOC values of bulk soil (Fig. 2). Instead, normalised TOC values were

¹ Table 1 provides brief descriptions of the explanatory variables mentioned in this section. Further details are documented in the [Supplementary Material](#). Table A2.1 provides definitions and Figures A2.1 and A2.2 illustrate the spatial distribution of the parameters in maps

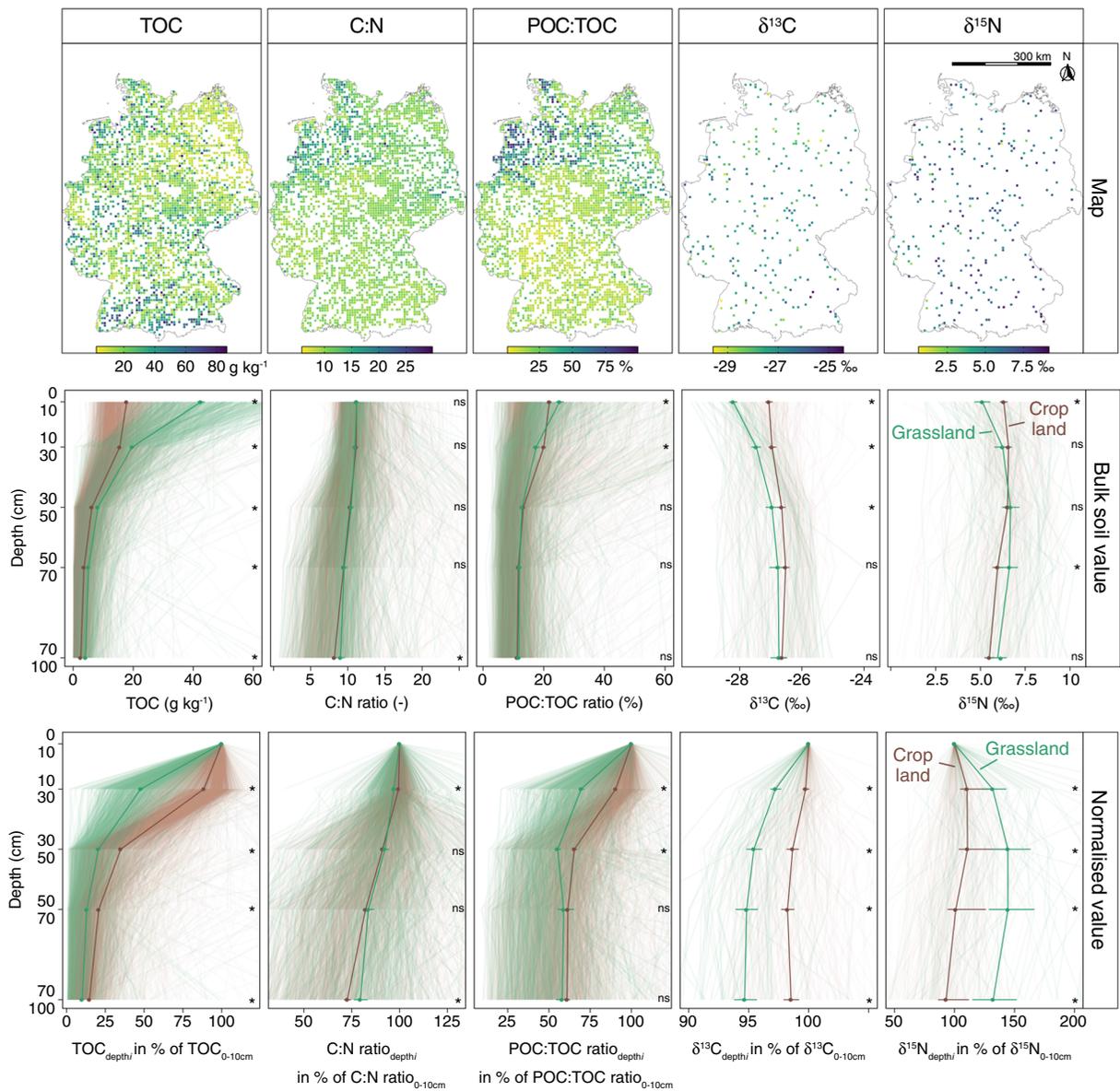
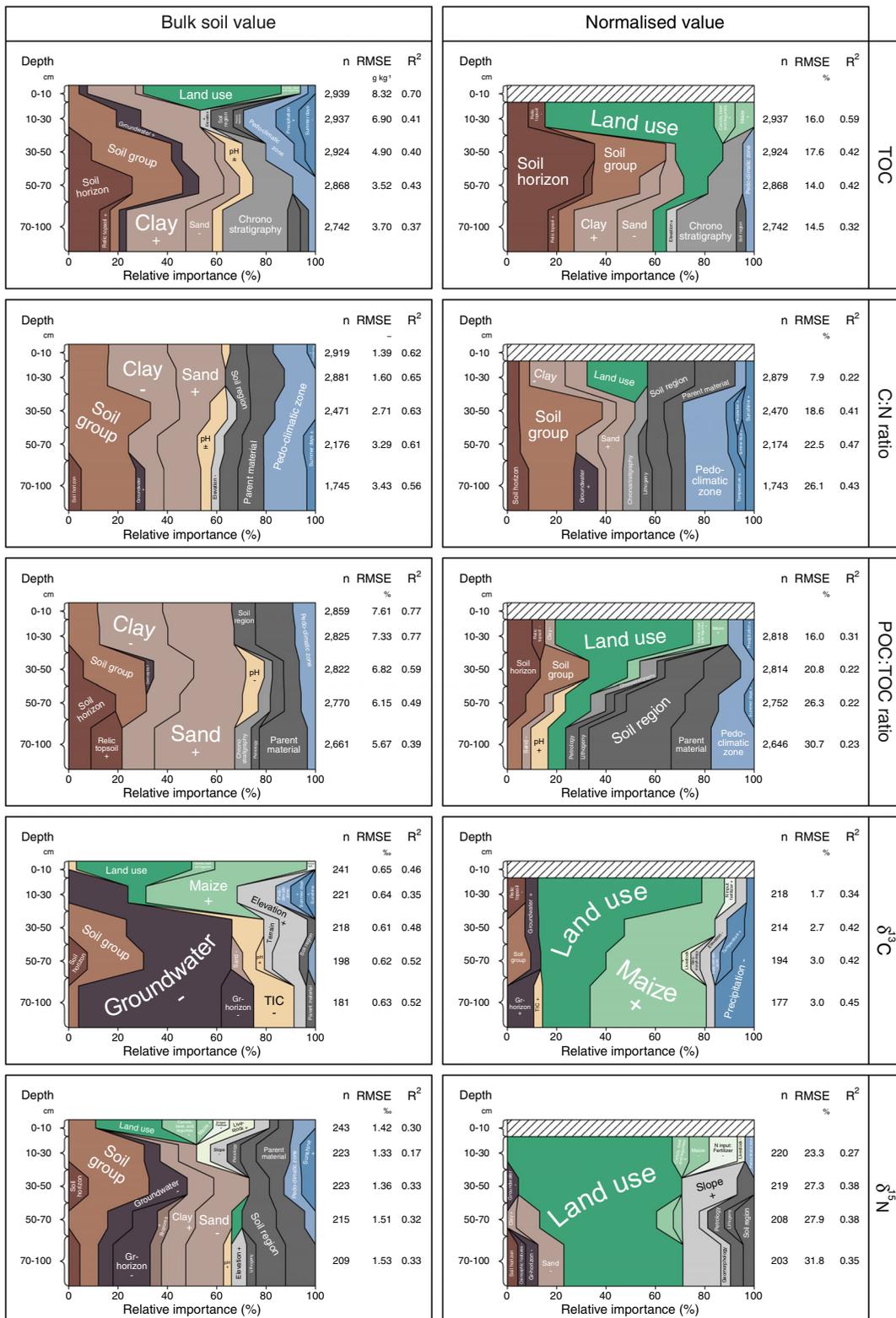


Fig. 1 Distribution of total organic carbon (TOC) content, C:N ratio, particulate organic carbon (POC) to TOC ratio, as well as $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values in German agricultural soils. Top row: map of bulk soil values in 0–10 cm. Middle row: depth distribution of individual and mean bulk soil values coloured by land use. Bottom

row: depth distribution of individual and mean normalised values coloured by land use. Normalised values were used to describe depth gradients. Error bars illustrate bootstrapped 95% confidence intervals around means and stars denote significant differences between means ($P < 0.05$)

related more to land use and soil horizons. In the upper 30 cm of grassland, TOC content decreased much more with depth than it did under cropland (Fig. 1), which is why the effect of land use on normalised TOC was most pronounced in 10–30 cm depth (Fig. 2). In 30–70 cm, differences between soil horizons were most important for

explaining normalised TOC values. For example, in 30–50 cm, M horizons and A horizons still contained 43% of TOC in 0–10 cm (95% CI, 42 to 45), while the other master horizons contained only 27% of TOC in 0–10 cm (95% CI, 26 to 27). Below 50 cm depth, geological chronostratigraphy became increasingly important for explaining normalised



◀ **Fig. 2** Important predictors of TOC content, C:N ratio, POC:TOC ratio, and $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values by depth. Left: results for predicting bulk soil values without information from other depth increments. Right: results for predicting normalised values that describe depth gradients. Earthy colours represent pedology, grey represents geology and geomorphology, green represents anthropogenic influence and blue represents climate-related variables. Areas are proportional to the relative importance of variables for predicting a given target at a given depth increment. Each model is characterised by the number of observations in the training data (“n”) and average errors from three times-repeated, five-fold cross validation (root mean square error “RMSE” and R-squared “R²”)

TOC contents, while parameters related to soil formation processes (soil horizon and soil reference group) became less important – a pattern similar to that observed for TOC values of bulk soil.

C:N ratio

In 0–10 cm, the mean C:N ratio of bulk soil was 11.2 (95% CI, 11.1 to 11.3) and its interquartile range was from 9.9 to 11.7. This variation was mostly related to texture. The C:N ratio increased with increasing sand content, especially above 70% sand, while it decreased exponentially with increasing clay content to around 10 for clay contents above 10% (Fig. A3.2). High C:N ratios clustered in particular in the “Old Drift” region (Fig. 3c, Fig. A2.2) and the “Weser-Ems-Geest” pedo-climatic zone (Fig. 3d, Fig. A2.2). These regions are located in north-west Germany and are characterised by historic heathland and peatland cover.

At greater depths, Random Forest identified similar drivers and performed equally well in explaining C:N ratios of bulk soil as for 0–10 cm soil depth (see above). In 30–70 cm, differences in C:N ratios by soil reference groups were most pronounced (Figs. 3a and 4) and consequently the variable importance of soil reference groups increased (Fig. 2). This was particularly because Podzol soils showed remarkably wide C:N ratios in subsoils. The C:N ratio of Podzols increased with greater depth, which was different to all the other soil reference groups (Figs. 3a and 4). Surprisingly, at all depth increments, the C:N ratio of cropland and grassland was very similar (Fig. 1) and consequently land use was not acknowledged as being important for predicting C:N ratios (Fig. 2). The site-specific nature of C:N was highlighted by a strong improvement of models for

subsoil C:N when informed about the respective C:N ratio in 0–10 cm (Fig. A3.3).

The C:N ratios changed much less with increasing soil depth than TOC. In 70–100 cm depth, bulk soil C:N was only 25% lower than in 0–10 cm depth (95% CI, 24 to 27). Even land use and tillage introduced only minor changes in the depth distribution of C:N ratios (Fig. 1). Normalisation of C:N ratios, i.e. expressing C:N ratios in terms of the percentage of their values in 0–10 cm, made land use important as a predictor of the ratios in 10–30 cm. It also increased the importance of groundwater and geological chronostratigraphy, whereas soil texture had less explanatory power (Fig. 2, Fig. A2.2).

POC:TOC ratio

The POC:TOC ratio showed the most distinct regional pattern of all the organic C indicators examined in this study (Fig. 1). In north-west German lowlands, bulk soil POC:TOC ratios above 30% were widespread in 0–10 cm of soil, while other regions had lower POC:TOC ratios. As for C:N ratios, soil texture and variables identifying historic heathland and peatland cover were most important for predicting POC:TOC ratios in 0–10 cm soil depth (see “Old Drift” in Fig. 3c; “Weser-Ems-Geest” in Fig. 3d). The POC:TOC ratio decreased exponentially with increasing clay content (Fig. A3.2).

With increasing depth, mean POC:TOC ratios decreased from 20% (95% CI, 19 to 20) in 0–10 cm to 11% (95% CI, 11 to 12) in 70–100 cm depth. Despite this decrease, the importance of explanatory variables hardly changed with depth, and texture remained the most important explanatory variable, which made POC:TOC in 0–10 cm a good predictor of POC:TOC at greater depths (Fig. A3.3). Major variations in subsoil POC:TOC were mostly associated with human soil profile modifications (Anthrosols, Fig. 4) and topsoil burial. Normalised POC:TOC ratios, which show the differences between sampling depths, were largely independent of soil texture, in contrast to the POC:TOC ratios of bulk soil. Instead, normalised POC:TOC ratios were driven by land use, soil reference group, soil region, parent material and pedo-climatic zone. In cropland, POC:TOC ratios decreased more gently with increasing depth than in grassland (Fig. 1). Podzols showed a much stronger decline of POC:TOC ratios with depth than other soil reference groups, even though the C:N ratio in Podzols increased with depth (Figs. 3a and 4). In the “Old Drift” soil region and

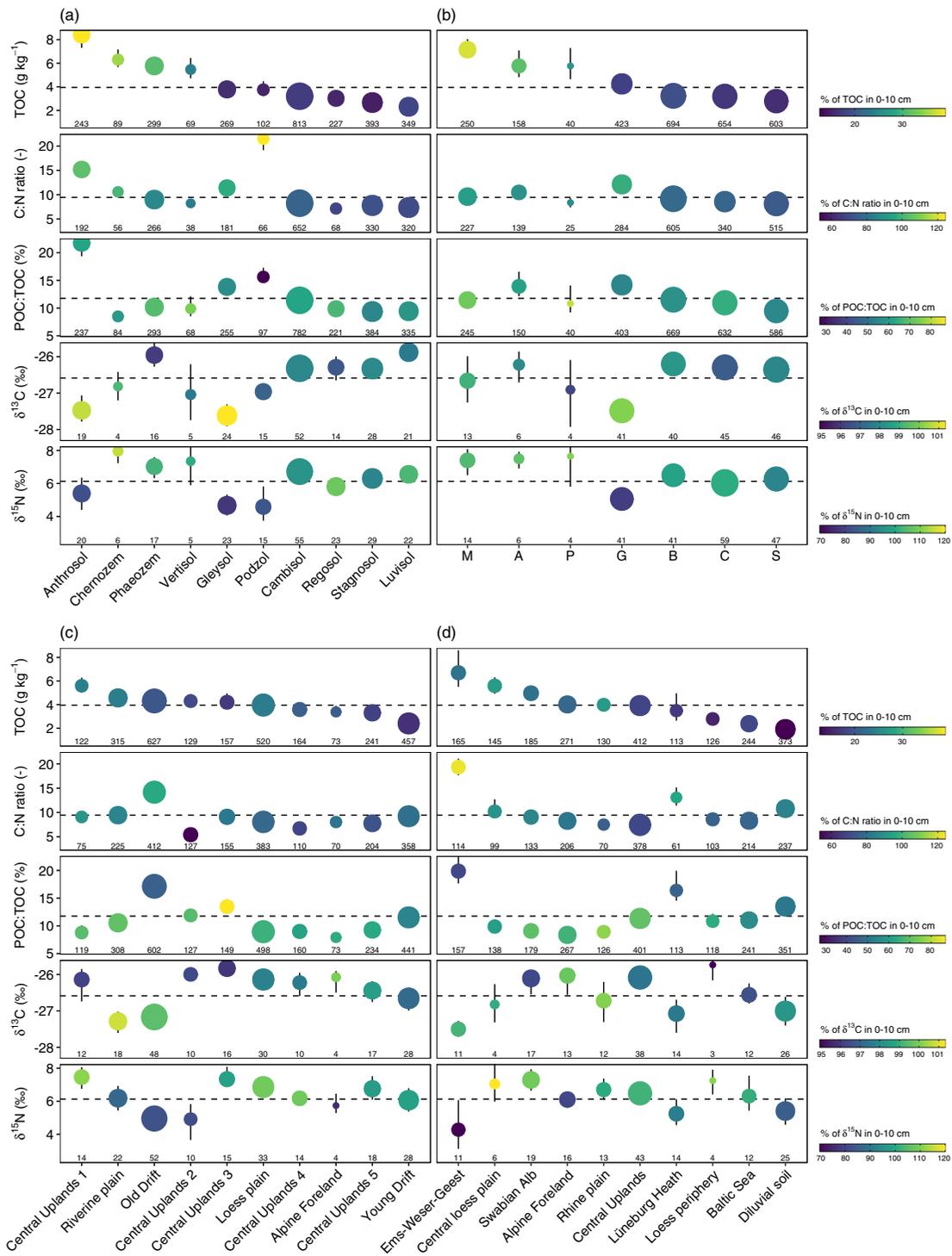


Fig. 3 TOC content, C:N and POC:TOC ratios, and δ¹³C and δ¹⁵N values in 50–70 cm depth by (a) WRB soil reference group, (b) soil horizon (“Horizonthaupsymbol”), (c) soil region (“Bodenregion”), and (d) pedo-climatic zone (“Bodenklimaraum”). Colours illustrate normalised values, which describe depth gradients. Sample numbers are characterized by dot

size and numbers below each dot. Dashed horizontal lines represent global mean values. For better clarity, only the most important factor levels are shown (for variables with more than ten factor levels, this is the ten most common levels; only factor levels with more than three observations in δ¹³C)

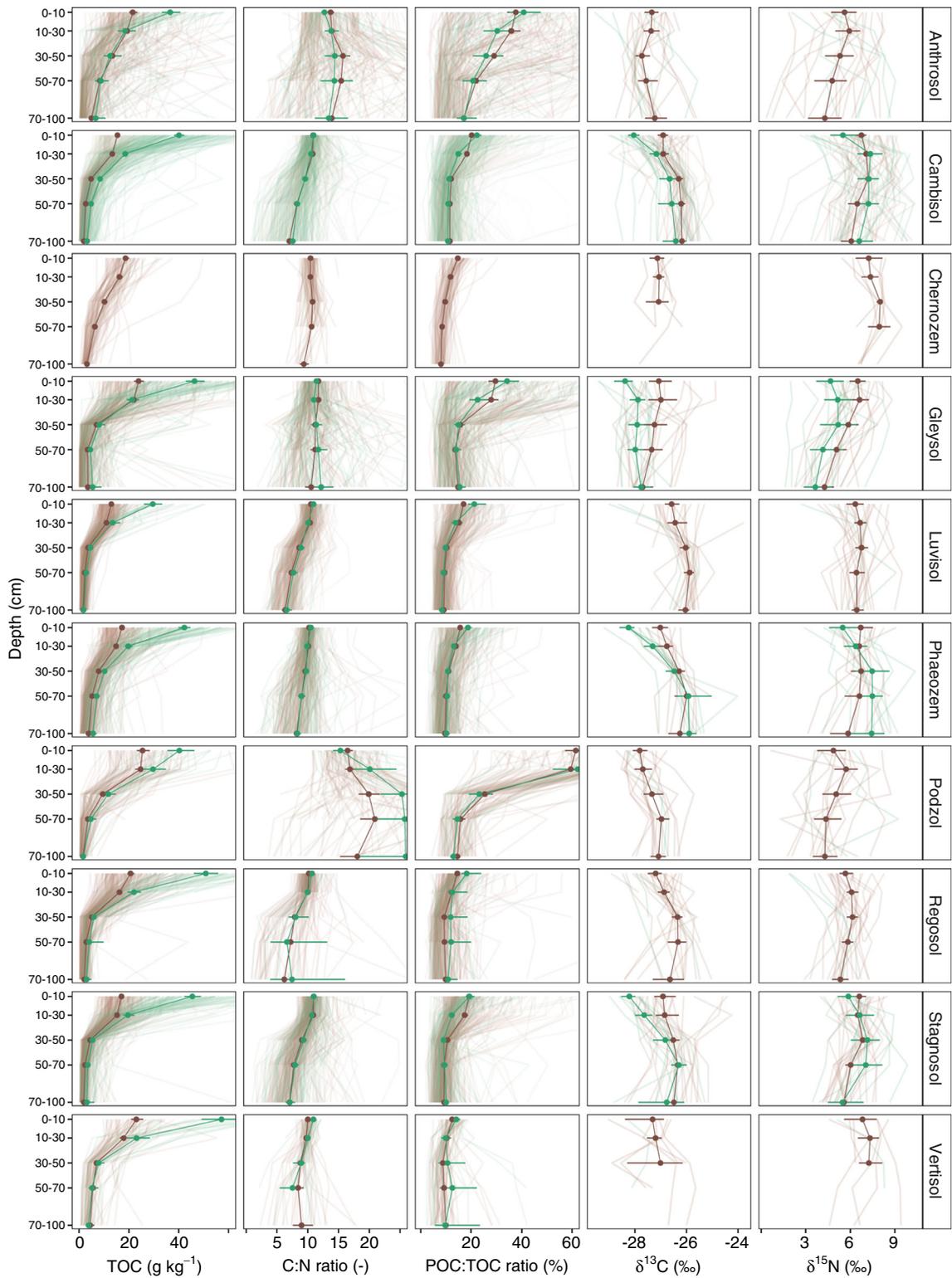
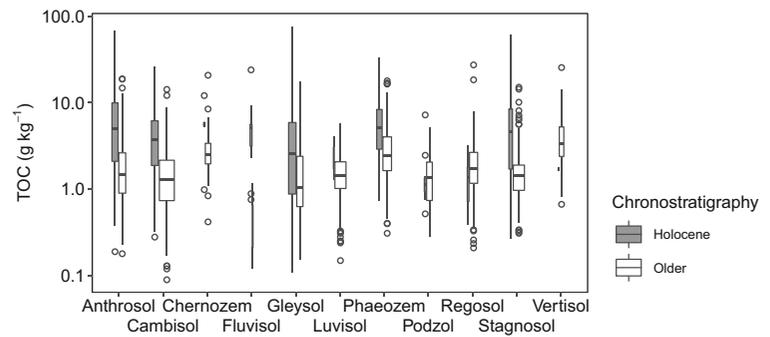


Fig. 4 Individual and mean depth distributions of TOC content, C:N and POC:TOC ratios, and $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values for cropland (brown) and grassland (green) by WRB soil reference group.

Means are only shown for $n > 5$. Error bars illustrate bootstrapped 95% confidence intervals of the means

Fig. 5 Total organic carbon content in 70–100 cm depth by soil group and chronostratigraphy. Box-plot width illustrates observation numbers



“Ems-Weser-Gest” pedo-climatic zone, which both had historic heathland and peatland cover, POC:TOC ratios declined much more strongly along the soil profile than in other regions (Fig. A2.2).

$\delta^{13}\text{C}$

In 0–10 cm, the mean $\delta^{13}\text{C}$ value was -27.4‰ (95% CI, -27.5 to -27.2), and differences in soil $\delta^{13}\text{C}$ were mostly explained by land use and the proportion of maize in the crop rotations (Figs. 2 and 6). The mean $\delta^{13}\text{C}$ value in the upper 10 cm of grassland was -28.2‰ (95% CI, -28.3 to -28.1) and that of cropland without maize was at -27.4‰ (95% CI, -27.5 to -27.2).

The mean $\delta^{13}\text{C}$ value increased along the soil profile, reaching -26.7‰ (95% CI, -26.8 to -26.5) in 70–100 cm (Fig. 1). Below 30 cm depth, groundwater explained most of the variation in $\delta^{13}\text{C}$ (Figs. 2 and 7). $\delta^{13}\text{C}$ correlated negatively with the groundwater table. In 70–100 cm, soil with mean groundwater tables shallower than 80 cm showed a mean $\delta^{13}\text{C}$ value of -27.7‰ (95% CI, -28.0 to -27.5), while soil with groundwater tables deeper than 200 cm showed a mean $\delta^{13}\text{C}$ of -26.3‰ (95% CI, -26.4 to -26.1). Similarly, in horizons with groundwater causing reducing conditions (Gr-horizons), an average $\delta^{13}\text{C}$ value of -27.7‰ (95% CI, -28.0 to -27.5) was detected, while in other horizons $\delta^{13}\text{C}$ was -26.5‰ (95% CI, -26.6 to -26.4) in 70–100 cm

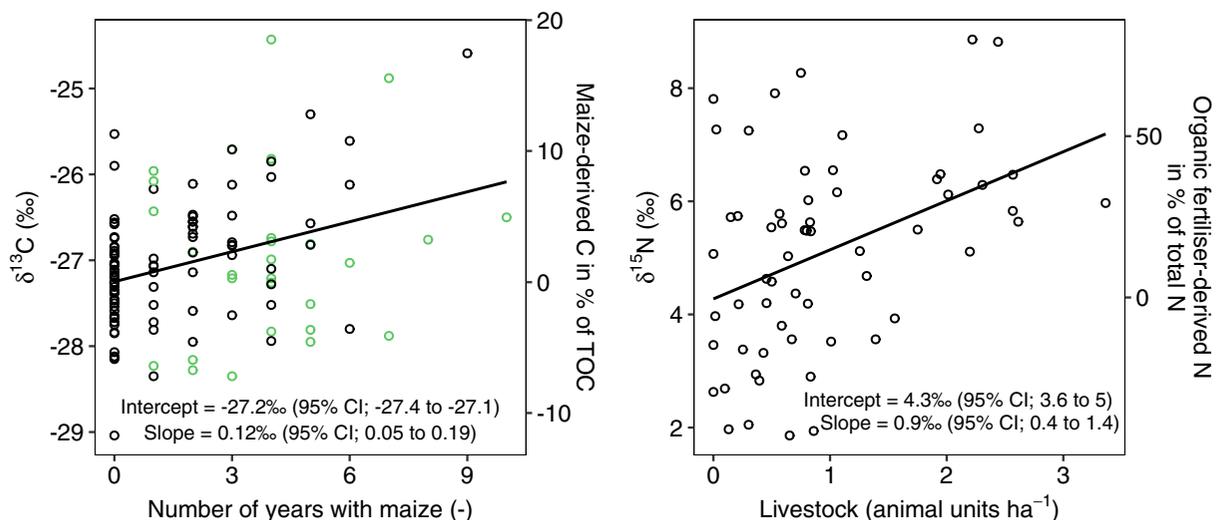
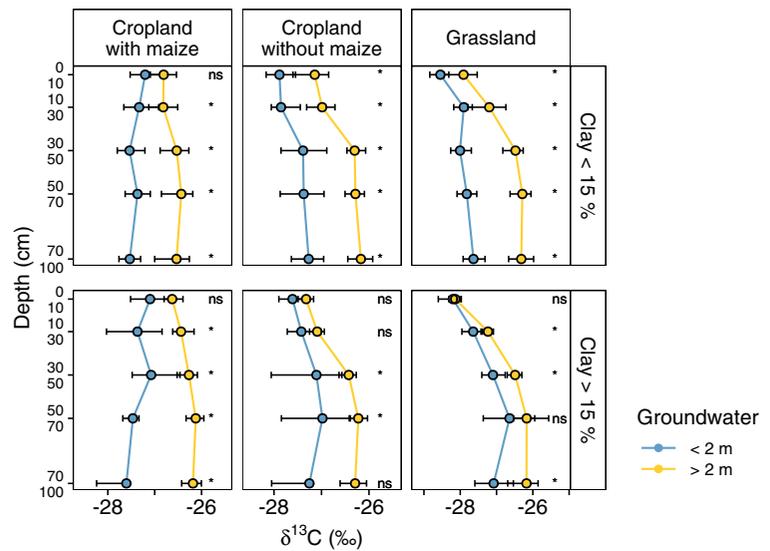


Fig. 6 Natural abundance of ^{13}C and ^{15}N in the upper 10 cm of bulk soil. Left: $\delta^{13}\text{C}$ in cropland as a function of the number of maize years in the ten years preceding sampling ($n=124$). Green colour shows sites where maize was the last crop before sampling. The secondary y-axis shows the proportion of maize-derived carbon (C) assuming a two-pool mixing model with the mean $\delta^{13}\text{C}$ without maize (-27.3‰) for the C_3 source and -12‰ for

the C_4 source. Maize was the only C_4 crop included in the crop rotations. Right: $\delta^{15}\text{N}$ in grassland as a function of livestock units per hectare ($n=61$). The secondary y-axis shows the share of organic fertiliser-derived nitrogen (N) assuming a two-pool mixing model with the mean $\delta^{15}\text{N}$ without livestock (4.3‰) for plant litter-derived N and 10‰ for organic fertiliser

Fig. 7 Mean $\delta^{13}\text{C}$ values of bulk soil by land use, soil texture and groundwater level. Sample numbers ranged from $n = 2$ to 27 for groundwater < 2 m, and from $n = 7$ to 48 for groundwater > 2 m, depending on depth increments and land use. Error bars represent bootstrapped 95% confidence intervals and stars denote significant differences ($P < 0.05$)



depth (Fig. 7). Furthermore, in 70–100 cm, $\delta^{13}\text{C}$ of TOC decreased with increasing TIC content. Subsoil $\delta^{13}\text{C}$ values below 50 cm depth were independent of the $\delta^{13}\text{C}$ in the surface soil (0–10 cm) due to the dominant influence of groundwater (Fig. A3.3). Cropland use and maize crops in particular smoothed the depth gradient of $\delta^{13}\text{C}$ since maize as a C_4 plant enriched ^{13}C in the topsoil (Fig. 7). Generally, changes in $\delta^{13}\text{C}$ along the depth profile were relatively small, with Gr-horizons in 70–100 cm even containing slightly more ^{13}C than in 0–10 cm depth (mean 101%; 95% CI: 99 to 103).

$\delta^{15}\text{N}$

In 0–10 cm, the mean $\delta^{15}\text{N}$ value was 6.0‰ (95% CI, 5.8 to 6.2). With increasing depth, $\delta^{15}\text{N}$ increased to 6.6‰ (95% CI, 6.3 to 6.8) in 30–50 cm before it decreased again and reached 5.6‰ (95% CI, 5.4 to 5.9) in 70–100 cm (Fig. 1). In 0–10 cm, differences in bulk soil $\delta^{15}\text{N}$ were mostly due to land use, management and variables identifying historic heathland and peatland in north-west lowlands (Fig. 2). In 0–10 cm depth, the $\delta^{15}\text{N}$ value under cropland was 1.2‰ points (95% CI, 0.7 to 1.7) higher than under grassland. Generally, $\delta^{15}\text{N}$ correlated positively with the proportion of canola (*Brassica napus*), beet (mostly *Beta vulgaris*) and legumes (mostly *Trifolium* sp.) in crop rotations (Fig. 2). Additionally, soil $\delta^{15}\text{N}$ values under both cropland and grassland were significantly lower in the “Ems-Weser-Geest” and “Old Drift” regions than in other regions (Fig. 3c, d). In grassland topsoil, soil $\delta^{15}\text{N}$ was

positively correlated with livestock density (Fig. 6). In the subsoil, land use and management lost explanatory power; instead, soil reference group and groundwater were important for predicting $\delta^{15}\text{N}$. In 70–100 cm depth, Gr-horizons were depleted in ^{15}N , showing 2.4‰ (95% CI, 1.9 to 2.8) lower $\delta^{15}\text{N}$ values than other horizons. At this depth, $\delta^{15}\text{N}$ in Gr-horizons averaged 78% of $\delta^{15}\text{N}$ in 0–10 cm (95% CI, 66 to 110), while in other horizons $\delta^{15}\text{N}$ averaged 104% of $\delta^{15}\text{N}$ in 0–10 cm (95% CI, 98 to 112). The depth gradient of $\delta^{15}\text{N}$ was thus best described by land use and groundwater influence. $\delta^{15}\text{N}$ values under grassland increased more with greater depth than they did under cropland because of lower values in the surface soil (Fig. 1).

Turnover of maize-derived soil carbon

Every growing season with maize increased the $\delta^{13}\text{C}$ value of cropland soil by on average 0.12‰ points (95% CI, 0.05 to 0.19) in 0–10 cm depth (Fig. 6). Based on Eq. 1, this translates to a mean turnover rate of 0.8% TOC year^{-1} (95% CI, 0.3 to 1.3). In 10–30 cm depth, the C turnover rate was roughly the same (95% CI, 0.3 to 1.2). These turnover rates indicate that C resides in the ploughing layer for around 100 to 300 years on average before being released back again into the atmosphere. Below 30 cm depth, the proportion of maize in crop rotations did not correlate with $\delta^{13}\text{C}$, which is why turnover and residence time of subsoil C could not be determined.

Soil compaction vs. root-derived deep C input

In cropland soil with hardpans, TOC content and C:N ratio declined more with increasing depth than in cropland soil without hardpans (Fig. 8). Their normalised values, i.e. values at depth_i relative to values in 0–10 cm, decreased significantly at increasing levels of soil compaction. The magnitude of this effect decreased with increasing soil depth. The above pattern was similar for POC:TOC ratios, but the significance here was not consistent between depth increments. In contrast to our initial hypothesis, the depth gradients of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values seemed not to be influenced by hardpans and showed no consistent trends with soil compaction.

Discussion

The current study is the first to provide a representative picture of the depth gradients of C:N, POC:TOC, $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ in Germany's mineral agricultural soils. Based on the large number of soil samples (2,931 sites * average of 5.4 depth increments) and high data quality (all soil analyses in one laboratory), our data-mining

approach revealed previously unseen patterns in TOC depth distribution and soil organic matter quality to 100 cm depth at regional scale. In the following, we discuss how these patterns can be used to trace different entry paths of organic C into German agricultural soils. We differentiate between:

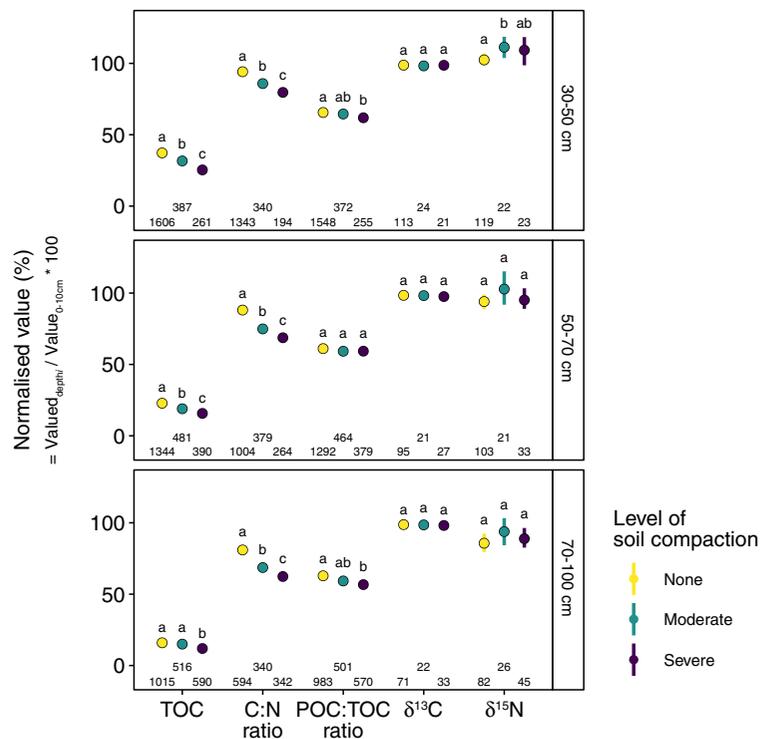
- autochthonous C, which covers aboveground and belowground C inputs from plants that grew on top of a given soil profile,
- C input from organic fertiliser (animal excreta), and
- parent material-inherent C, which comprises allochthonous C from plants that once grew somewhere else, e.g. upslope or upstream.

Also, we discuss the importance of vertical redistribution within soil profiles for explaining subsoil C, as well as texture and groundwater as major causes of variable C retention in German agricultural soils.

Origin of topsoil carbon

Most organic C entered the topsoil in the usual way, i.e. it originated from plants that grew on top of a given

Fig. 8 Normalised TOC, C:N, POC:TOC, $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values in cropland soil without compaction, moderate compaction and severe compaction. Mean values are shown with error bars representing bootstrapped 95% confidence intervals. Different letters indicate significant differences ($P < 0.05$). Sample numbers are shown below each dot



agricultural land, and not from somewhere else, and is therefore of autochthonous origin. The natural label of maize-derived C allows the fate of these autochthonous C inputs to be tracked. Our results suggest that each cropping year with maize replaced 0.3 to 1.3% of TOC in the upper 30 cm of cropland. Poeplau et al. (2020) estimate German cropland to store on average 61 Mg TOC ha⁻¹ in 0–30 cm (mineral soils only). Thus, autochthonous C contributed to the current soil organic C stock of German cropland at a mean rate of 0.3 to 1.3% of 61 Mg TOC ha⁻¹ year⁻¹ = 0.2 to 0.8 Mg C ha⁻¹ year⁻¹. This estimate is about one order of magnitude lower than the C input of 3.2 Mg ha⁻¹ year⁻¹, which was recently proposed by Jacobs et al. (2020). However, this number presented by Jacobs et al. (2020) refers to all plant residues and photosynthates that are not harvested, including stubble, mulch and other litter material > 2 mm. Soil organic C, however, includes only the organic C moieties < 2 mm. Plant residues > 2 mm have to be decomposed before entering the soil organic C pool. Decomposition is accompanied by respiration and organic C losses, which explains why less than 25% of the initial 3.2 Mg C input ha⁻¹ year⁻¹ that remains in the fields after harvest finally ends up in the fine soil. Once in the fine soil, organic C is likely to remain there for a mean period of 100 to 300 years before being mineralised and released back into the atmosphere. This mean residence time represents a rough approximation assuming steady state conditions and first-order decomposition rates (Gleixner et al. 2002) – we did not perform radiocarbon dating. Uncertainties also remain from the use of single-pool mixing models; due to uncertain data in the long term we did not perform estimations using two-pool C concepts (Derrien and Amelung 2011). Nevertheless, our estimate is in good agreement with commonly reported radiocarbon ages of TOC in the ploughing layer (e.g. Flessa et al. 2008).

Our method of calculating organic C turnover differed from most previous studies and contains additional uncertainties. Our calculations are based on: (i) sites with crop rotations instead of continuous maize cultivation, (ii) unknown maize-derived C inputs dating back more than 10 years, and (iii) about 120 core sites instead of just one or a few sites. We believe that (iii) compensates for potential noise introduced by (i) and (ii). Once organic C has entered the fine soil, the vast majority resides there for >> 10 years. We did not observe a systematic pattern between the time period since the last maize crop and the $\delta^{13}\text{C}$ value of soil (Fig. 6, green

dots). Hence, the exact timing of the last maize cropping in the past 10 years barely influenced organic matter turnover estimates in the range of 100 years and beyond. Unknown historic maize-derived C input was a potential source of bias in our turnover calculation. If the recorded share of maize in crop rotations within ten years before sampling correlated with historic maize use, we would have over-estimated TOC turnover. However, a pre-test suggested that this was not the case for the chosen reference period of ten years (cf. Fig. A3.1). Also, our estimate for C turnover (0.3 to 1.3% of TOC year⁻¹) is at the lower range of what has previously been observed in field trials and, therefore, an overestimation of organic C turnover is unlikely. Schiedung et al. (2017) reported much faster C turnover: here, after two years, maize-derived C had substituted 7.4 ± 3.2% of TOC in the topsoil (0–30 cm) and 2.9 ± 1.7% of TOC in the subsoil (30–50 cm), yet in their study shredded maize stubble was left as mulch on the soil surface, probably enhancing maize-derived C input into the topsoil. Flessa et al. (2000) reported that 15% of TOC in 0–30 cm was maize-derived after 37 years of continuous maize cropping, which translates into a turnover of 0.4% of TOC year⁻¹. Rasse et al. (2006) estimated topsoil organic C turnover to be 1.0% of TOC year⁻¹ for a Eutric Cambisol site in France. We therefore conclude that the simplified estimation of our turnover rate at 0.3 to 1.3% of TOC year⁻¹ is a reasonable magnitude for TOC turnover in the ploughing layer of German cropland.

In grassland, $\delta^{15}\text{N}$ values were indicative of additional C and N inputs in the form of organic fertilisers. Elevated $\delta^{15}\text{N}$ values in the topsoil of grassland with high livestock densities indicated a significant proportion of soil N to be derived from slurry or manure. Slurry-derived and manure-derived N tends to be enriched in ¹⁵N because of isotopic discrimination during digestion and ammonia volatilisation during storage, resulting in $\delta^{15}\text{N}$ values ranging from 6 to 13‰ (Kriszan et al. 2014). Each unit (500 kg) increase in livestock per hectare increased $\delta^{15}\text{N}$ of grassland in 0–10 cm by on average 0.9‰. This slope is in perfect agreement with the one documented by Kriszan et al. (2014), who studied the effect of stocking rates on $\delta^{15}\text{N}$ in the upper 5 cm of nine grassland sites in western Germany. Based on a positive correlation between $\delta^{15}\text{N}$ in topsoil and overall farm N balances, Kriszan et al. (2014) proposed using $\delta^{15}\text{N}$ as an indicator for N-use efficiencies at farm scale. Grassland of farms with positive N balances, i.e. greater N input than N export,

showed topsoil $\delta^{15}\text{N}$ values above 5.4‰ (Kriszan et al. 2014). In the present study, half of all grassland sites were above this threshold value, indicating that about half of Germany's grassland shares a decadal history of N surpluses due to excessive organic fertilisation. Assuming that $\delta^{15}\text{N}$ in grassland topsoil without slurry and manure fertilisation was 4.3‰ (intercept of linear regression line of $\delta^{15}\text{N}$ vs. livestock), average $\delta^{15}\text{N}$ of organic fertilizer was 10‰ and there was similar fractionation of soil N and amended slurry/manure N, then each livestock unit increased the share of organic fertiliser-derived N under grassland and 0–10 cm by 15 percentage points (95% CI, 8 to 22; Fig. 6). Considering that organic C and total N were highly correlated (grassland, 0–10 cm: $R^2 = 0.86$), the proportion of organic fertiliser-derived topsoil C can be assumed to be similar to the organic fertiliser-derived proportion of topsoil N. This estimate should be considered as a very approximate figure because divergence in both $\delta^{15}\text{N}$ of N input and fractionation factors for $\delta^{15}\text{N}$ in soil can be huge (Craine et al. 2015; Högberg 1997) – much higher than for $\delta^{13}\text{C}$. Nonetheless, the effect of livestock on $\delta^{15}\text{N}$ values in soil indicates that in addition to plant-derived C input, manure-derived and slurry-derived C has also contributed significantly to topsoil C stocks in German grassland.

The Random Forest algorithm identified systematic patterns in TOC contents, C:N and POC:TOC ratios between sampling sites and depths. These patterns were indicative of land use and management-driven differences in the quantity and quality of C inputs and subsequent redistribution of these inputs. In 0–10 cm, land use was the most important predictor of TOC and grassland showed a TOC content 2.4 times greater than cropland. This might be explained by grassland typically receiving a greater total root-derived C input than cropland (Don et al. 2009). Root-derived C resides in soil longer than other types of C input (Rasse et al. 2005). This supports the hypothesis that higher TOC content in grassland compared with that in cropland is caused by different quantities of root-derived C input. Differences in TOC contents by land use were further amplified by tillage-induced redistribution of TOC under cropland. Most cropland in Germany is ploughed annually to an average depth of about 30 cm, which decreases TOC content in 0–10 cm and increases it in 10–30 cm. Thus, differences in the depth gradient of TOC in the upper 30 cm between grassland and cropland arose from both different C input and tillage-

induced redistribution. Both processes together explain the large land use effects on the depth gradients of TOC contents, as well as of POC:TOC and $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values (Fig. 2, right) – tillage effects could not be separated from C input effects.

Interestingly, in 0–10 cm, Random Forest considered land use to be the most important variable for explaining TOC, but land use was irrelevant for explaining absolute C:N and POC:TOC ratios (Fig. 2, left). This was surprising considering the greater root-derived C input and much less frequent or absent tillage in grassland compared with cropland. A higher TOC content at equal C:N and POC:TOC ratios in grassland than in cropland can only be explained if both POC and MOC were equally elevated under grassland, indicating that the amount of MOC in the topsoil is governed not only by texture but by input as well: under steady state conditions, topsoil MOC content thus also increases with increasing C input (Rehbein et al. 2015).

In addition to recent land use and management effects on topsoil TOC, our study confirmed the legacy effects of historic heathland and peatland cover on organic C quality and quantity unique to the north-west part of Germany (Vos et al. 2018). Here, the maritime climate with relatively large amounts of rain in combination with forest clearances and subsequent spread of heathland favoured the podzolisation of old drift sands (Behre 2008; Schmidt and Roeschmann 2014). Ironpans that developed during podzolisation hindered water drainage and favoured the spread of peatland on a large scale. In the last few centuries, most of the peatland has been drained, the peat harvested, ironpans broken and former heathland and peatland converted to agricultural land (Schneider et al. 2017). Today, the affected soils still have an elevated TOC content (Springob and Kirchmann 2002), C:N ratios > 13 (Poeplau et al. 2020) and POC:TOC values > 30% (Vos et al. 2018). Detailed information about the extent of the historic heathland and peatland is limited and therefore could not be included as a covariate in this study. However, the region in which historic heathland and peatland have been found is characterised by a distinct climate (considerable rain, limited sunshine, mild winters), soil region (Old Drift), parent material (drift sand) and the pedo-climatic zone (Weser-Ems-Geest). Therefore, these variables informed the Random Forest models about historic heathland and peatland cover. Direct climate effects on recent C inputs or C turnover could not be inferred from the present study. Also, the

contribution of black carbon to TOC could not be elucidated in the present study because no proxy for black carbon has been analysed.

Origin of subsoil carbon

Below 30 cm depth, the ^{13}C label of maize and the ^{15}N label of organic fertiliser could no longer be detected, which confirms at national scale what many studies have previously shown at field scale: subsoil C and N are much less responsive to management-induced differences in C and N inputs than topsoil C and N (Poeplau and Don 2013). Subsoil C must therefore be older than topsoil C and exchanged with atmospheric C on a much longer timescale, millennial rather than centennial (Balesdent et al. 2018). In the Random Forest models, information on land use and management, which was recorded in the German Agricultural Soil Inventory for 10 years preceding sampling, was not relevant for explaining bulk soil TOC content below 30 cm depth. The only indication of significant exchange between subsoil C and the atmosphere was offered by the depth gradients of TOC, C:N and POC:TOC in soil profiles with hardpans (Fig. 8, Fig. A3.4). Below hardpans, TOC contents as well as C:N and POC:TOC ratios were relatively low, suggesting that hardpans significantly restricted root-derived litter input into subsoils. This confirms the existence of root-derived C inputs to the subsoil. However the question remains of how much deep roots and their rhizodeposits contribute to the TOC stocks of subsoils. If organic C were solely root-derived and the input of root-derived C were only a function of depth, it should be possible to predict the organic C quantity and quality along the soil profile from the organic C quantity and quality of the topsoils. However, the importance of topsoil C in 0–10 cm for explaining subsoil C decreased with increasing depth (Fig. A3.3) because:

- C input is not only a function of depth but also of root-restricting soil properties.
- the contribution of translocated C or parent material-inherited C increases with increasing depth, and/or.
- with increasing depth, texture-driven differences in the residence time of C increase – subsoil C is governed more by clay content than by root-derived C input.

This last point of texture-driven C retention has a huge effect on TOC storage, especially in the subsoil.

Deep in the soil profile, C input is typically low and C stabilisation governs how much C is being stored. Given two subsoils that only differ in C input, the subsoil receiving the higher C input will also show the larger TOC content. However, if a third subsoil not only differs in C input but also in soil texture, interpreting its TOC content is not straightforward because texture is key to C retention. This makes separating the effects of C input and C stabilisation challenging at sites with different textures. At a regional scale, however, as in German agricultural soils, texture tends to vary much more between sites than with soil depth², therefore texture-driven differences in C retention can be overcome by evaluating depth gradients of TOC instead of independent bulk soil values. Depth gradients allow inferences to be drawn about the depth distribution of organic C inputs even at sites with contrasting soil textures.

In the present study, we used a Random Forest algorithm to elucidate patterns behind the depth gradients of organic C in Germany's agricultural soils. To explain the depth gradients of TOC contents as well as C:N and POC:TOC ratios, the algorithm ranked variables identifying those variables describing soil parent materials (chronostratigraphy), lateral soil transport (soil horizon symbol "M"), topsoil burial (relic topsoil) and vertical translocation of C (soil groups³: Anthrosols, Chernozems and Podzols) as particularly important. About 35% of agricultural soils developed from relatively young, holocenic deposits, most of which were either of fluvio-marine or colluvic origin. Such alluvial and colluvial sediments typically comprise large amounts of allocthonous C from upslope or upstream areas (Doetterl et al. 2012, Mayer et al. 2019) and can bury significant

² In German agricultural soils, the average interquartile range of clay within soil profiles was roughly three times lower than the interquartile range of clay between sites.

³ In the German Agricultural Soil Inventory, soils were originally classified based on the national standard guideline (AD-HOC-AG Boden 2005), which builds on a genetic framework: soil horizons are classified according to their forming processes (genesis) and the resulting sequence of horizons gives rise to a specific soil group. This emphasis on soil formation distinguishes German soil taxonomy from many other soil classification systems, e.g. of WRB or US soil taxonomy, in which soil groups are primarily classified based on diagnostic soil properties. However, the resulting soil groups show parallels and groups of the German soil classification can roughly be translated to WRB reference soil groups according to AD-HOC-AG Boden (2005), which has been done in this study (see also Poeplau et al. 2020). Nonetheless, the soil groups used in this study remained indicative for soil forming processes, and therefore can be interpreted as "drivers" for soil organic C.

amounts of autochthonous C at their depositional sites (Chaopricha and Marín-Spiotta 2014). In view of elevated subsoil C contents and only minor decreases of TOC in soil profiles from holocenic deposits (Fig. 5), we hypothesise a significant proportion of allochthonous C and buried autochthonous C to be still preserved in subsoils. Floodplains and wide valley bottoms are widely acknowledged to act as major C sinks in the landscape (Sutfin et al. 2016). Hoffmann et al. (2009) estimated the non-alpine part of the Rhine basin to have accumulated about 1 Pg of organic C from upstream and upslope areas since the beginning of the floodplain deposition, which translates to a mean Holocene sequestration rate of roughly $50 \text{ kg C ha}^{-1} \text{ year}^{-1}$. Recent C sequestration rates in the Rhine basin are hard to quantify, but there is evidence that agricultural intensification has increased pre-human C sequestration rates of the Rhine basin by at least one order of magnitude due to stronger hillslope erosion (Hoffmann et al. 2013). This highlights the importance of lateral C fluxes in explaining soil organic C stocks, especially (i) when moving beyond pedon to landscape and regional scales (Van Oost et al. 2012), and (ii) for explaining subsoil C stocks due to their lower C turnover than topsoils. Organic C inherited from holocenic, i.e. relatively young, soil parent material appears to be a key contributor to subsoil C stocks in German agricultural soils.

Another major contributor to subsoil C was topsoil C translocated into the subsoil via anthropogenic soil profile modifications (e.g. deep ploughing). Numerous agricultural soil profiles have been modified by farmers far beyond annual tillage depths with the goal of site melioration. Such Anthrosols showed much higher TOC contents, and wider C:N and POC:TOC ratios in subsoils than most other soil groups. The depth gradients of these parameters were very noisy in Anthrosols (Fig. 4) but these depth gradients were also less steep on average than in other soil groups (Fig. 3a). The noise in the depth gradients of Anthrosols illustrates the diversity of their underlying formation. Anthrosols comprise a large variety of soils that have all been modified profoundly by human activity, including deep-ploughed soil, mining overburden, landfills and plaggen soil. Most deep soil profile modifications have accidentally translocated and buried great amounts of topsoil C in deeper soil layers. Previous studies have shown that such topsoil burial increases the residence time of C and thus soil organic C storage (Alcántara et al. 2016). Anthrosols mostly occurred either in the Old Drift landscape of north-

west Germany or in viticultural areas along the Rhine and Mosel rivers, but they covered a significant proportion (8%) of agricultural land in Germany, making anthropogenic soil profile modifications an important feature for explaining the variability in subsoil C stocks and organic C depth gradients at national scale.

Furthermore, soil biota seems to have transferred great amounts of topsoil C into the subsoil. In the German soil classification (AD-HOC-AG Boden 2005), epipedon with strong bioturbation is coded with the soil horizon symbol “Ax”. On average, Ax-horizons reached a depth of 76 cm (95% CI, 74 to 80). Most “Ax”-horizons occurred in Chernozems, which interestingly showed relatively high TOC but relatively low POC:TOC values in subsoils compared with other soil groups (Fig. 3a). This could be explained by the activity of anecic earthworms, which might not only bury topsoil C but also convert particulate organic matter to mineral-associated organic matter and thus decrease POC:TOC ratios (Vidal et al. 2019).

Finally, some topsoil C was translocated to subsoils during podzolisation. About 3.5% of all sampled sites were Podzols. Most of them developed from pleistocenic drift sands in north-west Germany. Podzols showed much higher C:N but lower POC:TOC ratios in subsoil compared with topsoil, indicating that instead of POC, dissolved organic C species with a high C:N ratio were illuviated in the subsoil of Podzols (Sauer et al. 2007). The great importance of the “Weser-Ems-Geest” pedoclimatic region and the “Old Drift” soil region in explaining the depth gradients of C:N and POC:TOC points to considerable amounts of refractory C from historic heathland and peatland cover in soil profiles in this region. However, many agricultural soils in the “Weser-Ems-Geest” area were also subject to deep soil profile modifications (Anthrosols) or podzolisation (Podzols), which makes it difficult to disentangle the individual processes to explain elevated subsoil C in this area.

Effects of texture and groundwater on organic C

The close correlation between organic C content and soil texture and with variables characterising oxygen availability (depth of groundwater table, Gr-horizons) underlines their influence on the retention and turnover of soil organic matter. Clay increases the C-storage capacity of soils because it offers a large, charged specific surface area for sorption of organic C (Hassink 1997; von

Luetzow et al. 2008). Mineral-associated organic C mostly consists of organic matter with relatively low C:N ratios (Jilling et al. 2018). The observed decrease in C:N and POC:TOC ratios with increasing clay content (Fig. 2) can therefore be attributed to the preferential retention and microbial conversion of organic C on mineral surfaces. Sorption of organic C on clay surfaces decreases C turnover (von Luetzow et al. 2008), whereas microbes allocated on mineral surfaces directly incorporate metabolised C into their biomass and necromass (Kögel-Knabner and Amelung 2014). This makes soil texture a key variable for explaining the C stocks of agricultural soils in Germany (Vos et al. 2019).

Groundwater can decrease C turnover because it restricts oxygen availability and therefore biological activity (Blazejewski et al. 2005; Marin-Spiotta et al. 2011). About 11% of all examined sites show Gr-horizons, i.e. reducing conditions in at least 9 out of 12 months (AD-HOC-AG Boden 2005). These horizons had much lower $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values than those of other soil horizons with higher oxygen availability at similar depths. Since microbial turnover constantly enriches ^{13}C and ^{15}N in the metabolites (Boutton 1996), low $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values in water-saturated Gr-horizons may indicate retarded biological decomposition (Natlhoffner and Fry 1988).

In soil with a similar C input, different $\delta^{13}\text{C}$ values of bulk soil tended to be associated with different POC:TOC ratios. In Gr-horizons, however, $\delta^{13}\text{C}$ and also $\delta^{15}\text{N}$ were low, while POC:TOC ratios were average. Instead of retarded conversion of POC to MOC and associated changes in POC:TOC, it could be retarded MOC turnover that keeps $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values in Gr-horizons relatively low – in other horizons, MOC turnover could lead to an enrichment of heavy isotopes, particularly on the time-scale of pedogenesis. Alternatively, the low $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values observed in hydric soil could be explained by compound-specific organic matter turnover (Gurwick et al. 2008). Under anaerobic conditions, ^{13}C -enriched carbohydrates might be decomposed at a higher rate, resulting in the selective preservation of ^{13}C -depleted lignin compounds (Benner et al. 1987; Spiker and Hatcher 1987). Finally, it should be borne in mind that some Gleysols developed from C-rich holocenic sediments. The degree to which allochthonous C from aquatic sources, which also has light isotope signatures (Finlay and Kendall 2007; Laskov et al. 2002), contributes to the TOC content and thus to $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values in the Gr-horizons warrants further clarification.

It is noteworthy that groundwater had a much stronger effect on bulk soil values of $\delta^{13}\text{C}$ than on its depth gradient. This was because groundwater-affected soil showed slightly more negative $\delta^{13}\text{C}$ values than sites without groundwater influence, even in topsoil. Generally, groundwater tables were too low to affect C decomposition even in the topsoil. Instead, different source signals probably contributed to the observed differences in the $\delta^{13}\text{C}$ values of topsoils. Plants growing at sites without groundwater influence might experience more frequent and severe drought stress than plants growing at sites with high groundwater levels. Under water stress, C_3 plants tend to get enriched in ^{13}C by up to 1‰ as a result of a decreased stomatal aperture (Boutton 1996). Thus, water stress-induced ^{13}C enrichment of C_3 photosynthates might explain the observed differences in $\delta^{13}\text{C}$ values of topsoils (Fig. 7). Together, the two processes of decreased decomposition of subsoil C at sites with high groundwater and ^{13}C -enriched-C input, especially in the topsoil at sites without groundwater influence, might explain why groundwater affected $\delta^{13}\text{C}$ values in both subsoil and topsoil in the same direction.

Limits of Random Forest algorithms in explaining soil organic C

The present study complements a series of previous studies that have used Random Forest to gain a better understanding about soil organic C dynamics (Padarian et al. 2020). Generally, such studies follow the same principle: First, a Random Forest model is trained to describe soil organic C with a set of explanatory variables (training data). Then, variable importance is evaluated by measuring the increase in model error after permuting (i.e. deleting the information of) each explanatory variable one at a time (Molnar 2019). Using this method, for example, Hobbey et al. (2015) found climate variables to be key for explaining soil organic C storage in Eastern Australia. Also, in southeast Germany, Wiesmeier et al. (2012) identified precipitation and temperature to be important for explaining soil organic C. However, for whole Germany, our study and Vos et al. (2019) found climate to be insignificant for explaining soil organic C. Such seemingly contrasting findings can be confusing - they illustrate three major shortcomings of Random Forests in explaining soil organic C:

- In contrast to mechanistic models, Random Forest models can only describe patterns within the predictor space they were trained on (Meyer & Pebesma 2020). Random Forest models that are trained to predict organic C in mineral soil (as in this study) will fail to predict organic C in organic soil because, per definition, the organic C content in organic soil is much higher than in mineral soil. This is merely because of the spread of soil organic C in the training data, not because of the contrasting processes governing organic C storage in mineral and organic soils. The latter is unknown to the model.
- The permutation importance of an explanatory variable depends on (i) its variance, and (ii) on the other explanatory variables included in the training data. If there is little variance in an explanatory variable, it will only be of minor importance for the model. For example, climate can only become important for explaining differences in soil organic C in the presence of significant longitudinal or altitudinal gradients in the target area (Adhikari et al. 2020). This might explain why in Australia (significant climate gradient), Random Forest identified climate as important for explaining soil organic C whereas in Germany (insignificant climate gradient) this was not the case. Furthermore, variable importance depends on the composition of the training dataset. For example, in the present study, if the Random Forest models had remained uniformed about soil groups or soil horizons, the models would have associated elevated soil organic C content in Chernozems and M-horizons with surrogate variables, e.g. soil texture. Selecting only meaningful explanatory variables, i.e. variables that are indicative for potential drivers of soil organic C, and having a wide coverage of these potential drivers in the training data facilitates the interpretability of Random Forest models.
- Random Forests can only reveal associations, not causality (Wadoux et al. 2020).

However, if these limitations are kept in mind, unravelling associations with Random Forest models can spark hypotheses for subsequent mechanistic studies (cf. Kell and Oliver 2004). In conclusion, Random Forest models provide a valuable tool to shed light on systems whose complexity is not yet fully understood, such as soil organic C dynamics at regional scales.

Loosening hardpans is climate smart

A significant bend in the depth gradients of TOC contents, C:N and POC:TOC ratios associated with hardpans suggests restricted deep root-derived C inputs in and below hardpans. In the case of TOC, the same trend was found even after mass correction to account for site-specific rock contents and bulk densities by evaluating depth gradients of TOC density, i.e. kg TOC ha⁻¹ at depth_i relative to 0–10 cm (Fig. A3.4). These results confirm that loosening compacted soil layers has the potential for additional C storage in subsoils because loosening of compacted soil can facilitate deep root-derived C inputs. If loosening of compacted soils could achieve the same depth gradients of TOC density as that observed in non-compacted soil profiles, this could increase C stocks in 30–100 cm by 2.3 t C ha⁻¹ on average. However, the site-specific C-sequestration potential of soil loosening depends on the magnitude and depth of compacted soil layers. At sites with severely compacted soil layers starting already at 30 cm depth, loosening of soils could potentially accumulate up to 9 t ha⁻¹ additional organic C in 30–100 cm, while at sites with medium compaction only at 70–100 cm depth, loosening could increase subsoil organic C stocks by about 1 t ha⁻¹ only. If all compacted soil layers in 30–100 cm depth were loosened so that the depth gradients of TOC density resembled those of non-compacted soil profiles, German cropland could store 0.03 Pg (4%) more organic C in total in 30–100 cm depth. This corresponds to 119,874 thousand tonnes of carbon dioxide equivalents or 1.8 times the annual greenhouse-gas emissions of Germany's entire agricultural sector (UBA 2019). Considering that clay content increased with increasing level of compactness and that clay correlated slightly positively with the depth gradient of TOC, the carbon dioxide sequestration potential of cropland could even be greater. However, from what is known about topsoils (Poeplau et al. 2011), it is likely that subsoil C stocks would need at least a few decades to reach their new C equilibrium after soil loosening. Furthermore, deep soil loosening is not a trivial measure and might have to be repeated after a few years (Schneider et al. 2017). Biological soil loosening with biopore forming, tap-rooted pioneer crops could provide a viable alternative to mechanical deep loosening methods (Kautz 2015). It is generally easier to loosen shallow hardpans than

those at greater depths. Loosening shallow hardpans also has greater potential for C sequestration and is thus most promising for climate mitigation.

Conclusions

Fresh photosynthates feed the long-term organic C stock of German cropland in 0–30 cm depth at a mean rate of 0.2 to 0.8 Mg C ha⁻¹ year⁻¹. Organic fertiliser is another important source of C input, especially in grassland. In parts of north-west Germany, sandy soils still contain elevated amounts of historic C from past heathland and peatland cover.

In subsoils, holocenic parent material harbours large amounts of allochthonous C from upstream or upslope areas. This parent material-inherited C is a major contributor to subsoil C, especially in soils that developed from alluvial or colluvial deposits. Furthermore, vertical displacement of topsoil C along the soil profile contributed significantly to subsoil C storage in German agricultural land, especially in Anthrosols (topsoil burial with deep ploughing), Chernozems (bioturbation) and Podzols (illuviation). Overall, the origin of parent material and related pedogenesis were key drivers of subsoil TOC storage. The only evidence of direct root-derived C input into subsoils was a significant bend in the depth profile of TOC contents as well as C:N and POC:TOC ratios in compacted soil layers, suggesting that hardpans restrict root-derived C inputs into subsoils. The sustained loosening of hardpans by means of mechanical or biological melioration could increase deep root-derived C inputs and thus result in a significant increase in subsoil C stocks, enhancing soil fertility and climate mitigation.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11104-020-04769-w>.

Acknowledgements We would like to thank Anette Giesemann and Ute Rieß for conducting IRMS measurements and Magdalena Gara, Lara Lindermann, Luisa Böning, Anna-Lisa Siemens, Fenja Steinberg and Stella Sonnenburg for sample preparation. Silke Weiß recorded all NIR spectra and Angelica Jaconi shared great R scripts for embarking on the spectral modelling. Cora Vos provided the laboratory data for topsoil C fractions. We thank Reinhard Well for his ideas on interpreting $\delta^{15}\text{N}$ data. Ali Sakhaee, Catharina Riggers, Christopher Poeplau, Cora Vos, Fabian Kalks, Julia Schröder, Mirjam Helfrich, Sofia Heukrodt, Sophie Drexler

and Tino Peplau commented on an earlier version of this manuscript. This study was part of the BonaRes research project Soil3 (grant number 031B0515E) funded by the German Federal Ministry of Education and Research (BMBF).

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adhikari K, Mishra U, Owens PR, Libohova Z, Wills SA, Riley WJ, Hoffman FM, Smith DR (2020) Importance and strength of environmental controllers of soil organic carbon changes with scale. *Geoderma* 375:114472
- Alcántara V, Don A, Well R, Nieder R (2016) Deep ploughing increases agricultural soil organic matter stocks. *Glob Change Biol* 22:2939–2956
- Amelung W, Brodowski S, Sandhage-Hofmann A, Bol R (2008) Chap. 6 combining biomarker with stable isotope analyses for assessing the transformation and turnover of soil organic matter. *Advances in agronomy*. Academic, Cambridge
- Amelung W, Blume H-P, Fleige H, Horn R, Kandeler E, Kögel-Knabner I, Kretzschmar R, Stahr K, Wilke B-M (2018) Scheffer/Schachtschabel Lehrbuch der Bodenkunde. Springer-Verlag, Berlin
- Appuhn A, Joergensen RG (2006) Microbial colonisation of roots as a function of plant species. *Soil Biol Biochem* 38:1040–1051. <https://doi.org/10.1016/j.soilbio.2005.09.002>
- Balesdent J, Mariotti A (1996) Measurement of soil organic matter turnover using ¹³C natural abundance. In: Boutton T, Yamasaki S (eds) *Mass spectrometry of soils*. Marcel Dekker, New York
- Balesdent J, Mariotti A, Guillet B (1987) Natural ¹³C abundance as a tracer for studies of soil organic matter dynamics. *Soil Biol Biochem* 19:25–30. [https://doi.org/10.1016/0038-0717\(87\)90120-9](https://doi.org/10.1016/0038-0717(87)90120-9)
- Balesdent J, Basile-Doelsch I, Chadoeuf J, Cornu S, Derrien D, Fekiacova Z, Hatté C (2018) Atmosphere–soil carbon transfer as a function of soil depth. *Nature* 559:599–602. <https://doi.org/10.1038/s41586-018-0328-3>
- Batjes NH (1996) Total carbon and nitrogen in the soils of the world. *Eur J Soil Sci* 47:151–163. DOI <https://doi.org/10.1111/j.1365-2389.1996.tb01386.x>

- Behre K-E (2008) Landschaftsgeschichte Norddeutschlands. Umwelt und Siedlung von der Steinzeit bis zur Gegenwart. Wachholtz Verlag, Neumünster
- Benner R, Fogel ML, Sprague EK, Hodson RE (1987) Depletion of ^{13}C in lignin and its implications for stable carbon isotope studies. *Nature* 329:708–710
- Blazewski GA, Stolt MH, Gold AJ, Groffman PM (2005) Macro- and micromorphology of subsurface carbon in riparian zone soils. *Soil Sci Soc Am J* 69:1320–1329
- Boden AD-HOC-AG (2005) *Bodenkundliche Kartieranleitung*, 5th edn. Schweizerbart, Stuttgart
- Boutton T (1996) Stable carbon isotope ratios of soil organic matter and their use as indicators of vegetation and climate change. In: Boutton T, Yamasaki S (eds) *Mass spectrometry of soils*. Marcel Dekker, New York
- Canadell J, Jackson RB, Ehleringer JR, Mooney HA, Sala OE, Schulze ED (1996) Maximum rooting depth of vegetation types at the global scale. *Oecologia* 108:583–595. <https://doi.org/10.1007/bf00329030>
- Canty A, Ripley B (2019) boot: Bootstrap R (S-Plus) Functions. R package version 1.3–24. <https://cran.rproject.org/web/packages/boot/index.html>. Accessed 12 Feb 2020
- Chaopricha NT, Marin-Spiotta E (2014) Soil burial contributes to deep soil organic carbon storage. *Soil Biol Biochem* 69:251–264. <https://doi.org/10.1016/j.soilbio.2013.11.011>
- Craine JM, Brookshire E, Cramer MD, Hasselquist NJ, Koba K, Marin-Spiotta E, Wang L (2015) Ecological interpretations of nitrogen isotope ratios of terrestrial plants and soils. *Plant Soil* 396:1–26
- Derrien D, Amelung W (2011) Computing the mean residence time of soil carbon fractions using stable isotopes: impacts of the model framework. *Eur J Soil Sci* 62:237–252
- Destatis (2019) *Land- und Forstwirtschaft. Statistisches Jahrbuch - Deutschland und Internationales*. Bundesministerium für Ernährung und Landwirtschaft, Bonn. www.destatis.de/jahrbuch. Accessed 12 Feb 2020
- Doetterl S, Six J, Van Wesemael B, Van Oost K (2012) Carbon cycling in eroding landscapes: geomorphic controls on soil organic C pool composition and C stabilization. *Glob Change Biol* 18:2218–2232. <https://doi.org/10.1111/j.1365-2486.2012.02680.x>
- Don A, Scholten T, Schulze ED (2009) Conversion of cropland into grassland: Implications for soil organic-carbon stocks in two soils with different texture. *J Plant Nutr Soil Sci* 172:53–62. <https://doi.org/10.1002/jpln.200700158>
- FAOSTAT (2019) *Crops*. Last update: February 6, 2020. Accessed 12 Feb 2020. <http://www.fao.org/faostat/en/#data/QC>
- Finlay JC, Kendall C (2007) Stable isotope tracing of temporal and spatial variability in organic matter sources to freshwater ecosystems. In: Michener R, Lajtha K (eds) *Stable isotopes in ecology and environmental science*. Blackwell Publishing, Malden, Oxford, Victoria
- Flessa H, Ludwig B, Heil B, Merbach W (2000) The origin of soil organic C, dissolved organic C and respiration in a long-term maize experiment in Halle, Germany, determined by C-13 natural abundance. *J Plant Nutr Soil Sci* 163:157–163. [https://doi.org/10.1002/\(sici\)1522-2624\(200004\)163:2<157::aid-jpln157>3.0.co;2-9](https://doi.org/10.1002/(sici)1522-2624(200004)163:2<157::aid-jpln157>3.0.co;2-9)
- Flessa H, Amelung W, Helfrich M, Wiesenberger GLB, Gleixner G, Brodowski S, Rethemeyer J, Kramer C, Grootes PM (2008) Storage and stability of organic matter and fossil carbon in a Luvisol and Phaeozem with continuous maize cropping: A synthesis. *J Plant Nutr Soil Sci* 171:36–51. <https://doi.org/10.1002/jpln.200700050>
- Gleixner G, Poirier N, Bol R, Balesdent J (2002) Molecular dynamics of organic matter in a cultivated soil. *Org Geochem* 33:357–366. [https://doi.org/10.1016/S0146-6380\(01\)00166-8](https://doi.org/10.1016/S0146-6380(01)00166-8)
- Gurwick NP, Groffman PM, Yavitt JB, Gold AJ, Blazewski G, Stolt M (2008) Microbially available carbon in buried riparian soils in a glaciated landscape. *Soil Biol Biochem* 40:85–96. <https://doi.org/10.1016/j.soilbio.2007.07.007>
- Hapfelmeier A, Hothorn T, Ulm K, Strobl C (2014) A new variable importance measure for random forests with missing data. *Stat Comput* 24:21–34. <https://doi.org/10.1007/s11222-012-9349-1>
- Hassink J (1997) The capacity of soils to preserve organic C and N by their association with clay and silt particles. *Plant Soil* 191:77–87. <https://doi.org/10.1023/a:1004213929699>
- Hastie T, Tibshirani R, Friedman J (2009) *The elements of statistical learning. Data Mining, Inference, and Prediction*. Springer, New York
- Hobley E, Wilson B, Wilkie A, Gray J, Koen T (2015) Drivers of soil organic carbon storage and vertical distribution in Eastern Australia. *Plant Soil* 390:111–127
- Hoffmann T, Glatzel S, Dikau R (2009) A carbon storage perspective on alluvial sediment storage in the Rhine catchment. *Geomorphology* 108:127–137
- Hoffmann T, Mudd S, Van Oost K, Verstraeten G, Erkens G, Lang A, Middelkoop H, Boyle J, Kaplan J, Willenbring J (2013) Humans and the missing C-sink: erosion and burial of soil carbon through time. *Earth Surf Dyn* 1:45–52
- Högberg P (1997) Tansley review no. 95 ^{15}N natural abundance in soil-plant systems. *New Phytol* 137:179–203
- Hothorn T, Bühlmann P, Dudoit S, Molinaro A, Van Der Laan MJ (2005) *Survival ensembles*. *Biostatistics* 7:355–373
- IPCC (2013) *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge
- Jackson RB, Canadell J, Ehleringer JR, Mooney HA, Sala OE, Schulze ED (1996) A global analysis of root distributions for terrestrial biomes. *Oecologia* 108:389–411. <https://doi.org/10.1007/BF00333714>
- Jacobs A, Flessa H, Don A, Heidkamp A, Prietz R, Dechow R, Gensior A, Poeplau C, Riggers C, Schneider F, Tiemeyer B, Vos C, Wittnebel M, Müller T, Säurich A, Fahrion-Nitschke A, Gebbert S, Hopstock R, Jaconi A, Kolata H, Lorbeer M, Schröder J, Laggner A, Weiser C, Freibauer A (2018) *Landwirtschaftlich genutzte Böden in Deutschland – Ergebnisse der Bodenzustandserhebung*. Johann Heinrich von Thünen-Institut, Braunschweig. Thünen Report 64. <https://doi.org/10.3220/rep1542818391000>
- Jacobs A, Poeplau C, Weiser C, Fahrion-Nitschke A, Don A (2020) Exports and inputs of organic carbon on agricultural soils in Germany. *Nutr Cycl Agroecosyst*. Early View. <https://doi.org/10.1007/s10705-020-10087-5>
- Jaconi A, Poeplau C, Ramirez-Lopez L, Van Wesemael B, Don A (2019) Log-ratio transformation is the key to determining soil organic carbon fractions with near-infrared spectroscopy. *Eur J Soil Sci* 70:127–139

- Jilling A, Keiluweit M, Contosta AR, Frey S, Schimel J, Schnecker J, Smith RG, Tiemann L, Grandy AS (2018) Minerals in the rhizosphere: overlooked mediators of soil nitrogen availability to plants and microbes. *Biogeochemistry* 139:103–122. <https://doi.org/10.1007/s10533-018-0459-5>
- Kautz T (2015) Research on subsoil biopores and their functions in organically managed soils: A review. *Renew Agric Food Syst* 30:318–327. <https://doi.org/10.1017/S1742170513000549>
- Kell DB (2011) Breeding crop plants with deep roots: their role in sustainable carbon, nutrient and water sequestration. *Ann Bot* 108:407–418. <https://doi.org/10.1093/aob/mcr175>
- Kell DB (2012) Large-scale sequestration of atmospheric carbon via plant roots in natural and agricultural ecosystems: why and how. *Philos Trans R Soc B Biol Sci* 367:1589–1597. <https://doi.org/10.1098/rstb.2011.0244>
- Kell DB, Oliver SG (2004) Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *BioEssays* 26:99–105
- Kindler R, Siemens J, Kaiser K, Walmsley DC, Bernhofer C, Buchmann N, Cellier P, Eugster W, Gleixner G, Grünwald T (2011) Dissolved carbon leaching from soil is a crucial component of the net ecosystem carbon balance. *Glob Change Biol* 17:1167–1185
- Kögel-Knabner I, Amelung W (2014) 12.7 - Dynamics, Chemistry, and Preservation of Organic Matter in Soils. In: Holland HD, Turekian KK (eds) *Treatise on Geochemistry* (Second Edition). Elsevier, Oxford
- Kriszan M, Schellberg J, Amelung W, Gebbing T, Pötsch EM, Kühbauch W (2014) Revealing N management intensity on grassland farms based on natural $\delta^{15}\text{N}$ abundance. *Agr Ecosyst Environ* 184:158–167. <https://doi.org/10.1016/j.agee.2013.11.028>
- Kuhn M, Quinlan R (2018) Cubist: Rule- and instance-based regression modeling. R package version 0.2.2. <https://CRAN.R-project.org/package=Cubist>. Accessed 12 Feb 2020
- Laskov C, Amelung W, Peiffer S (2002) Organic matter preservation in the sediment of an acidic mining lake. *Environ Sci Technol* 36:4218–4223
- Liang C, Amelung W, Lehmann J, Kästner M (2019) Quantitative assessment of microbial necromass contribution to soil organic matter. *Glob Change Biol* 25:3578–3590. <https://doi.org/10.1111/gcb.14781>
- Lynch JP, Wojciechowski T (2015) Opportunities and challenges in the subsoil: pathways to deeper rooted crops. *J Exp Bot* 66:2199–2210
- Marin-Spiotta E, Chadwick OA, Kramer M, Carbone MS (2011) Carbon delivery to deep mineral horizons in Hawaiian rain forest soils. *J Geophys Res*: 116(G03011). <https://doi.org/10.1029/2010JG001587>
- Mayer S, Kühnel A, Burmeister J, Kögel-Knabner I, Wiesmeier M (2019) Controlling factors of organic carbon stocks in agricultural topsoils and subsoils of Bavaria. *Soil Tillage Res* 192:22–32
- Mevik B-H, Wehrens R, Liland KH (2019) pls: Partial Least squares and principal component regression. R package version 2.7-1. <https://CRAN.R-project.org/package=pls>. Accessed 12 Feb 2020
- Meyer H, Pebesma E (2020) Predicting into unknown space? Estimating the area of applicability of spatial prediction models. arXiv preprint arXiv:200507939
- Microsoft, Weston S (2020) foreach: Provides foreach looping construct. R package version 1.4.8. <https://CRAN.R-project.org/package=foreach>. Accessed 12 Feb 2020
- Miltner A, Bombach P, Schmidt-Brücken B, Kästner M (2012) SOM genesis: microbial biomass as a significant source. *Biogeochemistry* 111:41–55. <https://doi.org/10.1007/s10533-011-9658-z>
- Molnar C (2019) Interpretable machine learning. A guide for making black box models explainable. <https://christophm.github.io/interpretable-ml-book/>. Accessed 12 Feb 2020
- Natelhoffer K, Fry B (1988) Controls on natural nitrogen-15 and carbon-13 abundances in forest soil organic matter. *Soil Sci Soc Am J* 52:1633–1640
- Oldeman LR, Hakkeling RTA, Sombroek WG (1991) World map of the status of human-induced soil degradation: an explanatory note. Global Assessment of Soil Degradation (GLASOD), International Soil Reference and Information Centre (ISRIC). United Nations Environment Programme (UNEP), Wageningen
- Padarian J, Minasny B, McBratney AB (2020) Machine learning and soil sciences: a review aided by machine learning tools. *Soil* 6:35–52
- Pausch J, Kuzyakov Y (2018) Carbon input by roots into the soil: quantification of rhizodeposition from root to ecosystem scale. *Global change biology* 24:1–12
- Philp RP, Monaco GL (2012) Applications of stable isotopes in hydrocarbon exploration and environmental forensics. In: Baskaran M (ed) *Handbook of environmental isotope geochemistry*. Springer, Berlin
- Poeplau C, Don A (2013) Sensitivity of soil organic carbon stocks and fractions to different land-use changes across Europe. *Geoderma* 192:189–201. <https://doi.org/10.1016/j.geoderma.2012.08.003>
- Poeplau C, Don A, Vesterdal L, Leifeld J, Van Wesemael B, Schumacher J, Gensior A (2011) Temporal dynamics of soil organic carbon after land-use change in the temperate zone – carbon response functions as a model approach. *Glob Change Biol* 17:2415–2427. <https://doi.org/10.1111/j.1365-2486.2011.02408.x>
- Poeplau C, Jacobs A, Don A, Vos C, Schneider F, Wittnebel M, Tiemeyer B, Heidkamp A, Prietz R, Flessa H (2020) Stocks of organic carbon in German agricultural soils—Key results of the first comprehensive inventory. *J Plant Nutr Soil Sci Early View*. <https://doi.org/10.1002/jpln.202000113>
- R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Rasse DP, Rumpel C, Dignac M-F (2005) Is soil carbon mostly root carbon? Mechanisms for a specific stabilisation. *Plant soil* 269:341–356
- Rasse DP, Mulder J, Moni C, Chenu C (2006) Carbon turnover kinetics with depth in a french loamy soil. *Soil Sci Soc Am J* 70:2097–2105. <https://doi.org/10.2136/sssaj2006.0056>
- Rehbein K, Sandhage-Hofmann A, Amelung W (2015) Soil carbon accrual in particle-size fractions under *Miscanthus x. giganteus* cultivation. *Biomass Bioenergy* 78:80–91. <https://doi.org/10.1016/j.biombioe.2015.04.006>
- RStudio T (2019) RStudio: Integrated development environment for R. RStudio, Inc., Boston

- Sauer D, Sponagel H, Sommer M, Giani L, Jahn R, Stahr K (2007) Podzol: Soil of the Year 2007. A review on its genesis, occurrence, and functions. *J Plant Nutr Soil Sci* 170:581–597. <https://doi.org/10.1002/jpln.200700135>
- Schiedung H, Tilly N, Hütt C, Welp G, Brüggemann N, Amelung W (2017) Spatial controls of topsoil and subsoil organic carbon turnover under C3–C4 vegetation change. *Geoderma* 303:44–51
- Schmidt R, Roeschmann G (2014) Norddeutsche Altmoränenlandschaften. In: Blume HP, Stahr K, Fischer W, Guggenberger G, Horn R, Frede HG, Felix-Henningsen P (eds) *Handbuch der Bodenkunde*. Wiley, Weinheim. <https://doi.org/10.1002/9783527678495.hbbk2005011>
- Schneider F, Don A (2019) Root-restricting layers in German agricultural soils. Part I: extent and cause. *Plant Soil* 442: 433–451. <https://doi.org/10.1007/s11104-019-04185-9>
- Schneider F, Don A, Hennings I, Schmittmann O, Seidel SJ (2017) The effect of deep tillage on crop yield – What do we really know? *Soil Tillage Res* 174:193–204
- Spiker EC, Hatcher PG (1987) The effects of early diagenesis on the chemical and stable carbon isotopic composition of wood. *Geochim Cosmochim Acta* 51:1385–1391. [https://doi.org/10.1016/0016-7037\(87\)90323-1](https://doi.org/10.1016/0016-7037(87)90323-1)
- Springob G, Kirchmann H (2002) C-rich sandy Ap horizons of specific historical land-use contain large fractions of refractory organic matter. *Soil Biol Biochem* 34:1571–1581
- Stevens A, Ramirez-Lopez L (2013) An introduction to the prospectr package. R package Vignette R package version 0.1.3. <https://cran.r-project.org/web/packages/prospectr/index.html>. Accessed 12 Feb 2020
- 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
- Strobl C, Boulesteix A-L, Kneib T, Augustin T, Zeileis A (2008) Conditional variable importance for random forests. *BMC Bioinformatics* 9:307
- Sutfin NA, Wohl EE, Dwire KA (2016) Banking carbon: a review of organic carbon storage and physical factors influencing retention in floodplains and riparian ecosystems. *Earth Surf Proc Land* 41:38–60
- UBA (2019) Berichterstattung unter der Klimarahmenkonvention der Vereinten Nationen und dem Kyoto-Protokoll 2019. In: *Nationaler Inventarbericht zum Deutschen Treibhausgasinventar 1990–2017*. UNFCCC-Submission. Umweltbundesamt, Dessau-Roßlau
- Van Oost K, Verstraeten G, Doetterl S, Notebaert B, Wiaux F, Broothaerts N, Six J (2012) Legacy of human-induced C erosion and burial on soil–atmosphere C exchange. *Proc Natl Acad Sci* 109:19492–19497
- Vidal A, Watteau F, Remusat L, Mueller CW, Nguyen Tu T-T, Buegger F, Derenne S, Quenea K (2019) Earthworm cast formation and development: a shift from plant litter to mineral associated organic matter. *Front Environ Sci* 7. <https://doi.org/10.3389/fenvs.2019.00055>
- von Luetzow M, Kogel-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. *J Plant Nutr Soil Sci* 171:111–124. <https://doi.org/10.1002/jpln.200700047>
- Vos C, Jaconi A, Jacobs A, Axel D (2018) Hot regions of labile and stable soil organic carbon in Germany–Spatial variability and driving factors. *Soil* 4:153–167
- Vos C, Don A, Hobley EU, Prietz R, Heidkamp A, Freibauer A (2019) Factors controlling the variation in organic carbon stocks in agricultural soils of Germany. *Eur J Soil Sci* 70: 550–564
- Wadoux AMJC, Samuel-Rosa A, Poggio L, Mulder VL (2020) A note on knowledge discovery and machine learning in digital soil mapping. *Eur J Soil Sci* 71:133–136
- Walther L, Graf U, Kammer A, Luster J, Pezzotta D, Zimmermann S, Hagedorn F (2010) Determination of organic and inorganic carbon, $\delta^{13}C$, and nitrogen in soils containing carbonates after acid fumigation with HCl. *J Plant Nutr Soil Sci* 173:207–216. <https://doi.org/10.1002/jpln.200900158>
- Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J (2019) Welcome to the Tidyverse. *J Open Source Softw* 4: 1686
- Wiesmeier M, Sporlein P, Geuss U, Hangen E, Haug S, Reischl A, Schilling B, von Lutzow M, Kogel-Knabner I (2012) Soil organic carbon stocks in southeast Germany (Bavaria) as affected by land use, soil type and sampling depth. *Global Change Biol* 18:2233–2245
- Wittmann O, Blume HP, Filipinski M (1997) Soil classification of the Federal Republic of Germany. *Mitt Deutsch Bodenkundl Ges* 84:253–275
- Wright MN, Ziegler A (2015) ranger: A fast implementation of random forests for high dimensional data in C++ and R. arXiv preprint arXiv:150804409

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