Towards a Quantitative Estimate of Anthropogenic Subsoil Compaction in European Croplands Based on National Soil Surveys

Laura Sofie Harbo, Marine Lacoste, Line Boulonne, Julien Wengler, Owen Fenton, Giulia Bondi, Patrick Tuohy, Amélie Marie Beucher, Mathieu Lamandé, Tommy D'hose, et al.

To cite this version:

Laura Sofie Harbo, Marine Lacoste, Line Boulonne, Julien Wengler, Owen Fenton, et al.. Towards a Quantitative Estimate of Anthropogenic Subsoil Compaction in European Croplands Based on National Soil Surveys. European Journal of Soil Science, 2025, 76 (4), pp.18. ⟨10.1111/ejss.70150⟩. ⟨hal-05233562⟩

HAL Id: hal-05233562
https://hal.inrae.fr/hal-05233562v1

Submitted on 1 Sep 2025

European Journal of Soil Science

WILEY

European Journal of Soil Science

RESEARCH ARTICLE OPEN ACCESS

Towards a Quantitative Estimate of Anthropogenic Subsoil Compaction in European Croplands Based on National Soil Surveys

Laura Sofie Harbo[1] | Marine Lacoste[2] | Line Boulonne[2] | Julien Wengler[2] | Owen Fenton[3] | Giulia Bondi[3] | Patrick Tuohy[4] | Amélie Marie Beucher[5] | Mathieu Lamandé[5] | Tommy D'Hose[6] | Florian Schneider[1]

[1]Thünen Institute of Climate-Smart Agriculture, Braunschweig, Germany | [2]INRAE, Info&Sols, Orléans, France | [3]Teagasc, Environmental Research Centre, Co. Wexford, Ireland | [4]Animal and Grassland Research and Innovation Centre, Teagasc, Co. Cork, Ireland | [5]Aarhus University, Department of Agroecology, Tjele, Denmark | [6]Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Melle, Belgium

Correspondence: Laura Sofie Harbo (laura.harbo@thuenen.de)

Received: 27 November 2024 | **Revised:** 13 June 2025 | **Accepted:** 18 June 2025

Funding: This work was supported by the EJP SOIL project SoilCompaC (862695), SOLGRAS project (GUDP 39009-21-1850) and Horizon Europe AI4SoilHealth (101086179).

Keywords: annual cropland | bulk density | data-driven reciprocal modelling | Denmark | Flanders | France | Germany | Ireland | packing density | random forest model

ABSTRACT

Subsoil compaction can lead to lower yields and reduced ecosystem functioning due to limited root growth of crops and can affect the cycling of nutrients and water within the soil. Subsoil compaction is often assessed using soil packing density, which accounts for textural differences in bulk density, and a reference threshold. Due to the high costs of sampling bulk density, subsoil compaction is rarely assessed at regional or national scales. While some soils are naturally compact, human activities, for example agriculture, are known to further increase subsoil compactness. The assessment of the anthropogenic component of the subsoil compactness is challenged by the lack of reference soil (i.e., unaffected by anthropogenic activities) that is otherwise comparable in terms of pedogenic and climatic parameters. In this study, a data-driven reciprocal modelling approach was used to model a reference subsoil bulk density for annual croplands based on observations from permanent grasslands, as grasslands are assumed to be free of anthropogenic subsoil compaction. The data originated from soil monitoring networks in five European countries (Belgium (Flanders), Denmark, France, Germany and Ireland). Depending on the country, the subsoil surpassed the compaction threshold of 1.71 $\mathrm{g\,cm^{-3}}$ packing density for 14%–52% of sites. The highest proportion of compacted sites was found in Flanders, while Denmark had the lowest proportion. Similarly, the highest estimated anthropogenic subsoil compaction was found in Flanders (mean 0.05 $\mathrm{g\,cm^{-3}}$) while the lowest was found in France (mean 0.00 $\mathrm{g\,cm^{-3}}$). Overall, the highest estimated anthropogenic subsoil compaction was found in loamy soils and soils with the lowest organic C content, such as eastern Germany and eastern Denmark. Based on our results, between 0% (France) and 47% (Flanders) of the annual cropland sites currently surpass the packing density threshold for compacted soil due to anthropogenic activities rather than due to pedogenic drivers.

1 | Introduction

In many croplands on mineral soils, densely packed soil layers (i.e., soil layers with high compactness) pose a severe barrier to vertical root elongation, limiting the plant uptake of water and nutrients from the subsoil, thus making crops more vulnerable to drought stress and periods of nutrient limitations, in turn affecting yields (Håkansson and Reeder 1994; Houskova and

Summary

- Subsoil compaction modelled in five regions (Denmark, Flanders, France, Germany, Ireland)
- 14%–52% of annual croplands in surveyed regions have compacted subsoil (30–50 cm).
- Subsoil compaction is primarily explained by soil texture and organic C (negative correlation).
- 20%–60% of existing subsoil compaction in annual croplands is human-induced.

Montanarella 2007; Lipiec et al. 2012; Tim Chamen et al. 2015; Schneider and Don 2019; Sonderegger et al. 2020). Compacted soil layers also affect water infiltration to deeper soil layers and impact the ecosystem functions (Jones et al. 2003; Hamza and Anderson 2005; Hu et al. 2021; Schroeder et al. 2022), which in turn can exacerbate flood events after heavy rain (Troldborg et al. 2013; Alaoui et al. 2018; Sonderegger et al. 2020). Compaction of agricultural soil, defined as the densification of soil by the decrease in air-filled porosity and characterised by an increase in bulk density (van den Akker 2008), can have both natural and anthropogenic causes. While clay illuviation can naturally increase soil bulk density (Buurman et al. 1998; Batey 2009), the use of heavy agricultural machines also increases bulk density in both topsoil and subsoil (Håkansson and Reeder 1994; Keller and Or 2022). The degree of compaction of soil layers can be quantified by the packing density, which is a rescaling of the bulk density accounting for the influence of soil texture on soil bulk density (Renger et al. 2014); the packing density is compared to a threshold to determine whether the soil layer is compacted, that is, surpasses the set compaction threshold (Section 2.5).

While soil bulk density is influenced by both pedogenic and geogenic site properties (Schroeder et al. 2022), land management is a primary contributor to the degree of compaction in both the top- and subsoil (Soane et al. 1980; Rabot et al. 2018; Schneider and Don 2019). In particular, traffic of heavy agricultural machinery on agricultural land in suboptimal conditions (e.g., wet soil) has been linked to subsoil compaction (Håkansson and Reeder 1994; Schjønning et al. 2015b; Schjønning and Lamandé 2018; Keller and Or 2022).

Natural recovery from subsoil compaction is a slow process, and subsoil compaction can persist over several decades (e.g., Berisso et al. 2012, 2013). Mechanical loosening of topsoils can be achieved by tillage (Xu and Mermoud 2001; Alvarez et al. 2009; Kadžienė et al. 2011; Priori et al. 2024), whereas subsoil loosening is more time- and energy-intensive, and the soil is prone to recompaction (Schäfer-Landefeld et al. 2004).

The dynamics and effects of compaction of agricultural soils have been studied widely using field experiments (Hamza and Anderson 2005; Zhang et al. 2024), but the extrapolation of the results to regional or national scales is complicated by the influence of site-specific conditions such as soil texture, geological origin, climate and local management history. Nonetheless, assessment of the severity and extent of soil compaction is an important element of current soil health debates at the European scale (EU 2023). Due to the substantial resources required to collect bulk density samples, detailed field investigations have rarely been conducted at regional, national or continental scales. The proportion of the European agricultural area that has been degraded as a result of compaction is thought to range from 3% (Panagos et al. 2024) to an estimated 20% (Oldeman et al. 1991). Nevertheless, the extent of subsoil compaction at the continental scale remains to be quantified, despite the estimation of Schjønning et al. (2015a) based on SPADE8 data that approximately one-quarter of European subsoils (0.25–0.7 m deep) exhibit a high normalised relative density, that is, the texture-corrected density. To the best of our knowledge, there is no quantitative estimate to date for the extent and severity of anthropogenic subsoil compaction at a European scale that is based on observational data.

While the threats of soil compaction have been recognised for a long time, the extent and severity of soil compaction beyond the field scale are challenging to quantify. Due to soil heterogeneity and spatial variability in pedogenic, climatic and anthropogenic influences, observational bulk density data from numerous sites are required to accurately assess soil compaction at a large regional or national scale. Additionally, the quantification of compaction due to anthropogenic activities requires reference sites for comparison, which can be challenging to find in Europe due to the millennia-long history of agriculture and extensive impact of humans on the environment (Venter et al. 2016a, 2016b). While long-term experiments are essential for studying soil dynamics, they cannot serve as reference sites for broader regional studies due to variations in soil and climatic conditions. In this study, permanent grasslands serve as the reference for the annual cropland sites. Previous studies, using more sensitive indicators for subsoil compaction than bulk density (i.e., isotropy), have shown that grassland subsoil is less affected by compaction as compared to arable subsoils (Horn et al. 2020; Mordhorst et al. 2019), however the correlation between the effect on isotropy and bulk density remains unclear (Horn et al. 2020). While some compaction has also been shown for grasslands with low bulk density (Jorajuria et al. 1997), grasslands have been found to be less compacted that annual croplands in the transitional zone between the topsoil and subsoil (25–35 cm) in northern Germany (Wiermann et al. 2025), indicating an overall lower compaction of grasslands. Additionally, grasslands are hardly tilled (often only for renewal) and typically have a higher content of soil organic matter as compared to annual croplands (Bondi et al. 2021; Guillaume et al. 2022) which can increase soil elasticity and thereby contribute to the soil's resistance to deformation (Soane 1990), suggesting that grasslands might be less anthropogenically compacted than annual croplands. Furthermore, grasslands can often have a grass sod that also might protect the soil against compaction (Emmet-Booth et al. 2018) and the effect of animal trampling is found primarily in the topsoil (Mulholland and Fullen 1991; Hiltebrand et al. 2025), so animal grazing is expected to have minor influence on the compactness of the subsoil. Altogether, there is little evidence of severe subsoil compaction in permanent grasslands.

Using machine learning to predict a reference bulk density is an alternative approach to the identification of paired plots or the matching of reference sites to suspected compacted sites across the available data sets. The data-driven reciprocal modelling approach uses data from a reference dataset (i.e., permanent grasslands) to train a statistical model that can estimate a reference value (i.e., bulk density) for the target dataset (i.e., annual croplands) (Schneider et al. 2021). Data-driven reciprocal modelling requires a lot of data, both in terms of input variables for the statistical model as well as in terms of observations, to function well. However, the data-driven modelling approach circumvents the need for paired plots as well as issues relating to a lack of comparability between the reference group and the target group of observations.

In this study, we use the data-driven reciprocal modelling approach alongside data from permanent grasslands and annual croplands across five European countries to (i) give an estimate of the extent and severity of subsoil (30–50 cm) compaction in annual croplands on mineral soil, (ii) quantify the anthropogenically driven subsoil compaction based on the data-driven reciprocal modelling approach and (iii) examine the extent to which management practices and site characteristics can explain the estimated changes in bulk density.

2 | Materials and Methods

2.1 | Soil Survey Data

Bulk density (BD) data from the upper subsoil (target depth: 30–50 cm) and associated soil-forming variables were compiled from nationally or regionally representative surveys of agricultural soils in Belgium (Flanders region only; Cmon), Denmark (DSMG), France (RMQS), Germany (BZE-LW) and Ireland (I-SIS) (Table 1). Specifically, for each soil survey site (i.e., sampling location, see Supplementary Section 1 for survey-specific information), the following data were extracted or calculated: dry bulk density ($g\,cm^{-3}$) of the fine soil (< 2 mm) (Equation 1), volumetric coarse fragment (> 2 mm) content (vol %), organic C content (weight %), total nitrogen content (weight %), clay and silt content (weight %), as well as pH and land use. All data were laboratory and field data, except the coarse fragment content of the Danish sites, which was not measured during the sampling campaigns. Instead, the coarse fragment data for Denmark were gap-filled based on soil texture and coarse fragment content relationships published in Harbo et al. (2022). As the sampling depth varies between the included surveys (Table 1; Supplementary Section 1), sites were included in the study only if the bulk density measurement was collected primarily from the 30–50 cm depth, that is, the majority of the sampling ring or excavated volume was within the target depth. The data from each survey are described in greater detail in Supplementary Section 1.

The bulk density of the fine soil (BD, $g\,cm^{-3}$) was calculated from the measured parameters as follows:

$$BD = \frac{Mass_{total} - Mass_{coarse}}{Volume_{total} - Volume_{coarse}} \quad (1)$$

where $Mass_{total}$ is the dry mass of the entire soil sample (g), $Mass_{coarse}$ is the dry mass of the coarse fragments (> 2 mm) (g), $Volume_{total}$ is the volume of the entire soil sample (cm^3) and

TABLE 1 | Overview of the soil surveys included in the study, including sampling years, number of sites before and after data filtering and reference sources. See details of site filtering criteria in Section 2.3.

Country (Region)	Survey	Sampling years	Sampling depth (cm)	Annual cropland sites		Permanent grassland sites		Reference(s)
				Total	Included	Total	Included	
Belgium (Flanders)	Cmon	2022–2024	42.5–47.5, 30–60	124	111	54	44	Oorts et al. (2024)
Denmark	DSMG	1987–1991	Horizon-based	491	123	55	19	Østergaard and Mamsen (1990)
France	RMQS (1st campaign)	2000–2009	30–50*	886	422	537	189	Arrouays et al. (2002) Jolivet et al. (2022) Munera-Echeverri et al. (2024)
Germany	BZE-LW	2011–2018	30–50	2234	1700	729	421	(Poeplau, Don, et al. 2020) (Poeplau, Jacobs, et al. 2020) (data)
Ireland	I-SIS	2007–2013	Horizon-based	2	0 (2 eligible)	75	40	Creamer et al. (2014)

*The French data is collected using both horizon-based and fixed depth sampling, see Supplementary Section 1 for more information. Abbreviations used in table: BZE-LW: Bodenzustandserhebung Landwirtschaft; Cmon: Carbon monitoring network; DSMG: Danish Soil Monitoring Grid; I-SIS: Irish Soil Information System; RMQS: Réseau de Mesure de la Qualité des Sols.

$Volume_{coarse}$ is the volume of the coarse fraction (cm^3) calculated based on the mass of the coarse fragments and either the standard rock density of 2.65 $g\,cm^{-3}$ or, where available, measured rock densities.

Sites with coarse fragment content > 5% have been excluded from the study due to differences in sampling methodologies between the data sources (see Section 2.3 Data filtering for further elaboration).

To harmonise the data from the five surveys, the texture data were converted from the local definitions (see Supplementary Section 1) to standardised clay [0–2 μm], silt [2–50 μm] and sand [50–2000 μm] particle size classes using log-linear interpolation.

2.2 | Derived data

Data from the soil surveys were augmented by the inclusion of the following climatic, paedogenic and geological covariates that were derived from maps based on the coordinates of the sites; mean annual temperature (MAT; bio1) and mean annual precipitation (MAP; bio12) were extracted from the climate database CHELSA (Climatologies at High Resolution for the Earth's Surface, Karger et al. 2017). The soil group of each site was extracted from the map of the World Reference Base (WRB) of soil resources (IUSS Working Group 2022; Poggio et al. 2021). Slope, elevation and wetness index were extracted from EcoDataCube (Hengl 2018) using exact site coordinates. Information regarding soil parent material was similarly collected from geological maps covering Europe and adjacent areas (Asch 2005). To preserve data confidentiality regarding exact soil survey site coordinates, a script was shared with data managers and run locally, where applicable. For more information regarding the map-based covariates, see Table S1.

Site-specific information regarding agricultural practices, crop rotations and farm size was not available from the soil surveys. However, field activities are known drivers of anthropogenic subsoil compaction. Therefore, generalised data on farming practices were collected from Eurostat at NUTS2 level (NUTS2 is a European standard for referencing administrative divisions of countries for statistical purposes and typically corresponds to regions, provinces or states). The mean farm size was calculated across all available years by dividing the total agricultural area (UAA) of each NUTS2 region by the total number of holdings (farms) for each NUTS2 region included in the study (Eurostat 2024b).

Similarly, as management practices likely impact the subsoil compaction, we included the mean proportion of cropland area with potato and sugar beet cropping, which was calculated by dividing the total area of both crops by the total cropping area (ARA) in the NUTS2 region, taking the mean value across all available years (Eurostat 2024c). Only this one category of crops was considered in order to avoid negative correlation of the crop-related covariates as well as closed-compositional data, which is problematic for statistical analyses (Aitchison 1982; Pawlowsky-Glahn and Egozcue 2006). The proportion of potato and sugar beet was chosen due to the overall weight of the crop when harvested, as well as the spatial variability of the covariate as compared to other crop categories.

Similar to the proportion of cropland with potato and sugar beet, the mean proportion of the area under conventional tillage practice (as opposed to conservational tillage and no tillage) for each NUTS2 was calculated by dividing the area under conventional tillage by the total cropland area (ARA) of the NUTS2 region, taking the mean across all available years (Eurostat 2024).

2.3 | Data Filtering

The sites were evaluated for inclusion in the study in three aspects: (i) data quality, (ii) site relevance and (iii) scope of the study.

Data quality was assessed using data quality flags, if available, as well as plausibility checks and whether all relevant covariates were available for the target depth (30–50 cm) at each site. Plausibility checks included verifying that the sum of the particle size classes was equal to 100% (only relevant for the Danish sites), as well as reasonable values for BD (0.7–2.2 $g\,cm^{-3}$). In total, 1200 sites were excluded due to missing or implausible data in the target depth across all five surveys (740 annual cropland sites and 460 permanent grassland sites).

Site relevance was determined by geographical location and by evaluating the land use history of the sites where such information was available; for the French dataset, grassland sites younger than 10 years as well as permanently fallow lands were excluded. Additionally, sites outside metropolitan France were not included in the study. For the German dataset, grassland sites were excluded from the study if they had an Ap horizon (plough layer), indicating anthropogenic influence from ploughing. German sites were also excluded if they were located on drained peatlands, as such sites are not considered mineral soils. For the Danish, Flemish and Irish data, detailed land use history was not available for evaluation of the sites, so all sites were kept. In total, 398 sites were excluded from the study based on land use history (131 annual cropland sites and 267 grassland sites).

To assure the sites were within the scope of the study, in this case mineral soils (< 10% organic C) with less than 5% coarse fragments, a total of 866 sites were excluded (557 annual cropland sites and 309 permanent grassland sites). Sites with > 5% coarse fragments were excluded to avoid sites where the coarse fragments might have affected the undisturbed soil samples for BD during the sample collection. Additionally, the soil surveys have different approaches to sampling soils with high amounts of coarse fragments, and such sites are therefore less comparable across the different soil surveys.

Sites can fall outside multiple cut-off values (e.g., both high organic C and low BD, or coarse fragments and land use history), and the total number of sites excluded is therefore not the sum of the sites excluded by category.

From 5187 sites available from the soil surveys (1450 grassland sites and 3737 annual cropland sites), 3069 sites remained after data filtering (713 grassland sites and 2356 annual cropland sites) (Figure 1), corresponding to 59% initial total sites (49% of initial grasslands and 63% of initial annual croplands). For the individual countries and region, the number of remaining sites varied; for Germany, 72% of sites were retained (76% annual cropland sites and 58% permanent grassland sites), while the Flemish dataset retained 87% of sites (90% annual cropland sites, 81% permanent grassland sites). In the Danish dataset, 26% of sites were retained (25% annual cropland sites, 35% permanent grassland sites), while in the French dataset, 43% of sites were retained (48% of annual cropland sites, 35% of permanent grassland sites).

The Irish survey (I-SIS) only included two sites on annual cropland. As two sites cannot be representative for all of Irish cropland, only permanent grassland data from Ireland were retained for inclusion as training data for the grassland reference model (see Section 2.4). 53% of the permanent grassland sites from Ireland passed the data filtering.

For most (97.3%) of the 3071 included sites, soil data were collected after the year 2000, with the exception of the Danish data (2.7% of sites), which were collected primarily in the 1980s (Table 1).

2.4 | Quantification of Anthropogenic Subsoil Compaction Under Cropland

2.4.1 | Estimating the Grassland Reference Subsoil BD

To determine the change in subsoil BD of annual croplands due to anthropogenic activities, a reference BD is necessary. Such reference BD needs to be derived from sites without anthropogenically induced compaction; here, the subsoil under permanent grassland sites was assumed to be unaffected by traffic- or trampling-induced compaction, and the permanent grassland subsoil served as a reference to the subsoil under annual cropland sites in this study.

Subsoil BD is influenced by many factors other than land management, which must be taken into account in order to have a more accurate reference BD for the annual cropland sites. In order to include the pedogenic and climatic influences on subsoil BD, we employed a data-driven reciprocal modelling approach as described by Schneider et al. (2021) to estimate the reference BD values for individual cropland sites. In data-driven reciprocal modelling, a statistical model is trained to predict the target variable (here, BD) of the reference data

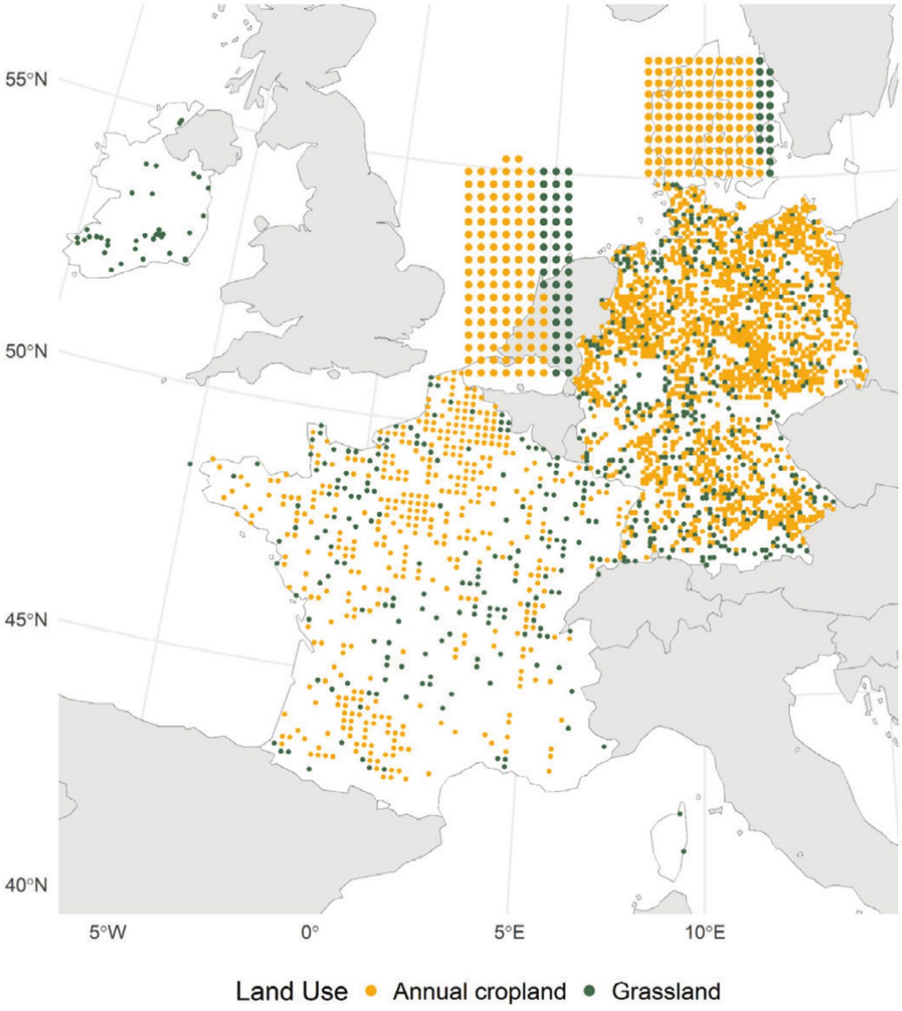

FIGURE 1 | Map of the sites included in the study; yellow circles represent annual cropland sites and green circles represent permanent grassland sites. As coordinates are not publicly available for Denmark and Flanders, the Danish and Flemish sites are arranged in grids and placed on the map for visualisation purposes. Countries and regions excluded from the study are shaded in grey. EuroGeographics for the administrative boundaries.

(i.e., data from an unaffected group, here, grasslands). This trained statistical model is then applied to the dataset of interest (i.e., affected group, here croplands) to estimate a site-specific reference value based on the identified predictors. In this study, the data-driven reciprocal model was implemented with Random Forest models.

The Random Forest model trained on the reference grassland data (henceforth the grassland reference model) included 54 covariates (Table S1). The model was trained on permanent grassland data from all five soil surveys. Prior to model training, Spearman's rank correlation coefficient was calculated for the continuous covariates; no correlation was above 0.75, and all covariates were retained.

The Random Forest model was trained with 500 trees. Its number of variables sampled at each split (*mtry*) and minimum node size were optimised using grid search. The optimal values were selected based on the highest concordance correlation coefficient between predicted and observed values in 10-fold cross-validation. The model optimisation process was validated using nested 10 × 10-fold cross-validation. Performance metrics were averaged over the 10 outer folds resulting in the following values (mean ± standard deviation): concordance correlation coefficient = 0.70 ± 0.04; coefficient of determination (R^2) = 0.54 ± 0.07; root mean square error (RMSE) = 0.11 ± 0.02 $\mathrm{g\,cm^{-3}}$; slope (observed vs. predicted) = 0.55 ± 0.06. The bias was negligible (0.00 ± 0.02).

For the optimised Random Forest model, permutation importance and partial dependence plots were calculated in order to evaluate the relative rank of the impact of the included covariates as well as the partial effect of the covariates on the target variable (i.e., bulk density). The grassland reference model included only covariates that were not influenced by land use or management (Table S1).

In the grassland reference model, organic C was by far the most important covariate for predicting BD (Figure 2A). There was a strong negative correlation between organic C content and the predicted BD (Figure 2B). Other than organic C, clay content, C-to-N (C/N) ratio as well as mean annual precipitation (MAP), also ranked at the top of covariates for predicting BD in grasslands. Similar to the organic C content, there was an overall negative correlation between the clay content and MAP covariates and the predicted subsoil BD, while the C/N ratio had a slight positive correlation. However, the effect of organic C on the predicted BD was considerably more pronounced than that of clay, C/N ratio and MAP (Figure 2B–E) as well as the other independent covariates included in the grassland reference model (Figure 2A). The remaining 46 covariates not included in Figure 2A had a minor influence on the model performance.

The results of an independent 10-fold cross-validation of the grassland reference model showed that the lowest observed BD values tended to be overestimated (i.e., appear above the 1:1 line), whereas the highest observed BD values tended to be underestimated (i.e., appear beneath the 1:1 line) (Figure 3A). However, there was no systematic bias towards either over- or underestimation, and this was likely an expression of regression to the mean. The information from the independent 10-fold cross-validation was used to compare the model performance on grassland sites to the model performance for annual cropland sites, rather than applying the grassland reference model to the permanent grassland sites, since applying the model to the data it is trained on is likely to result in overfitting.

The pattern of overestimation of the lowest values and underestimation of the highest values of the grassland reference model was generally also present when the grassland reference model was applied to the annual cropland sites (Figure 3B); however, the intention of the reciprocal modelling approach is to assess the difference between the modelled reference values and the observed values (i.e., the horizontal distance from observed BD to the 1:1 line), and an overall larger discrepancy between the two variables is therefore expected in Figure 3B. Indeed, a larger

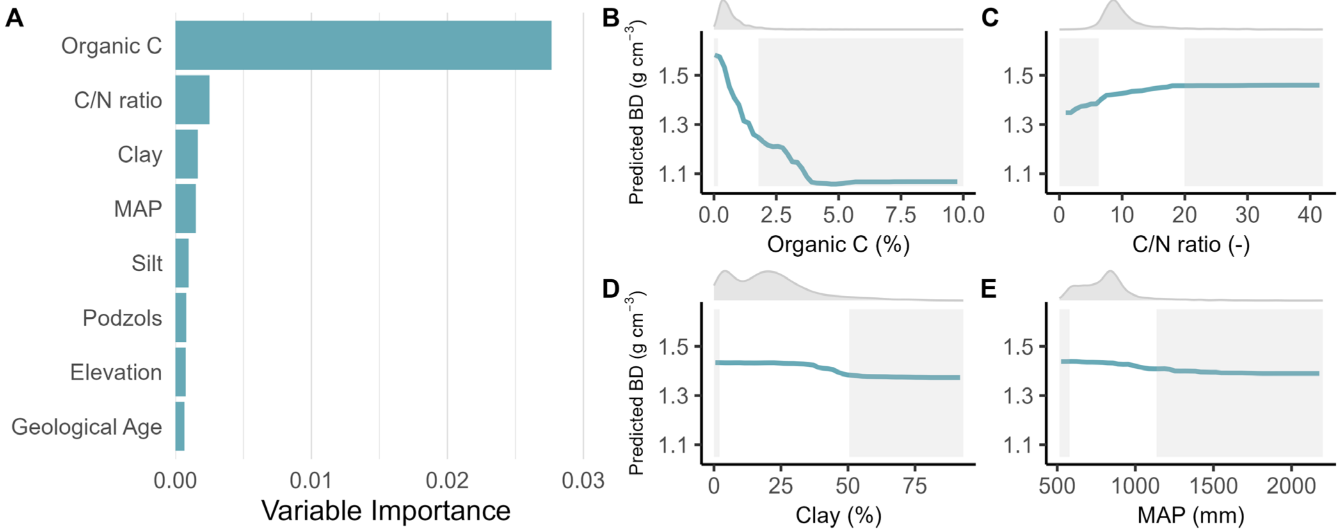

FIGURE 2 | (A) Top 8 covariates from the grassland reference model ranked according to variable importance. (B–E) Partial dependence plots illustrating the modelled partial effects of the top four most important predictors on bulk density from the grassland reference model. The grey boxes shade the lowest and highest 5% of observation ranges, respectively. Grey density plots show the distribution of the permanent grassland dataset.

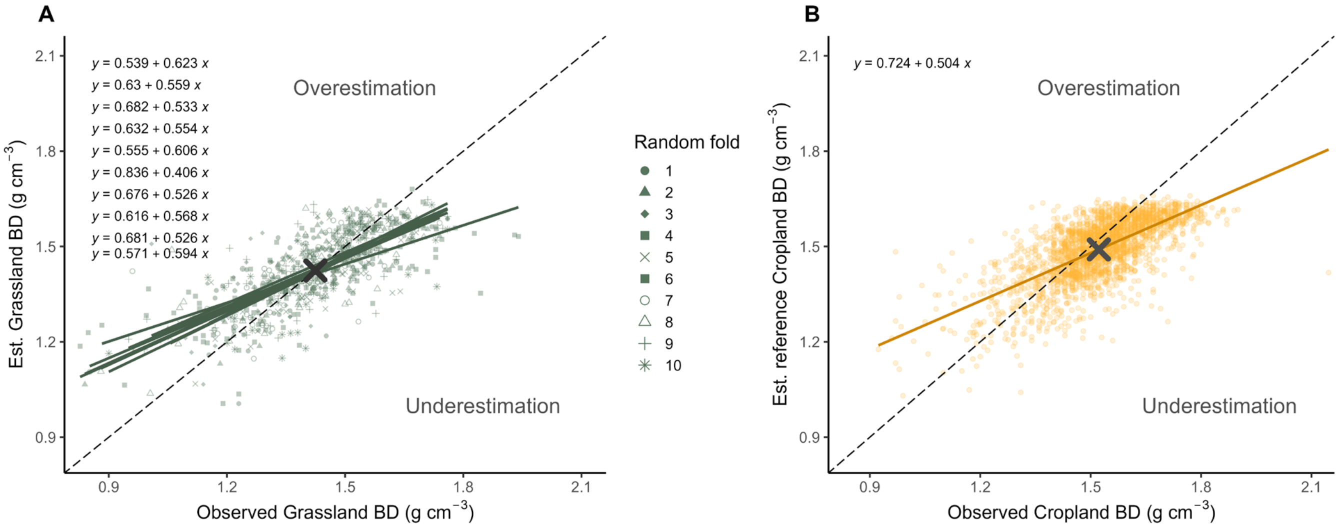

FIGURE 3 | (A) Estimated BD ($g\,cm^{-3}$) against the observed BD ($g\,cm^{-3}$) from the ten-fold random cross-validation of the grassland reference model for the permanent grassland sites ($n = 713$); the shape indicates the number of the random fold; a line is fit to each fold. (B) Estimated reference BD ($g\,cm^{-3}$) against the observed BD ($g\,cm^{-3}$) for the annual cropland sites that pass AOA ($n = 2358$) based on the grassland reference model. In both plots, the dashed lines indicate the 1:1 line, while the solid lines and equations show the best linear fit, and the X symbolises the mean value of both axes. Data from all surveys is included in A, while data from Ireland is excluded from B due to the low number of annual cropland sites.

proportion of the annual cropland sites were underestimated when the grassland reference model was applied to the annual cropland sites (Figure 3B). Additionally, the equation describing the best linear fit in Figure 3B was highly similar to the equations describing the linear fits for the 10 folds in Figure 3A, indicating that the model performance is stable across the range of BD values.

2.4.2 | Calculating the Difference in Subsoil BD

For each annual cropland site, a reference subsoil BD (BD_{ref}) was estimated using the grassland reference model. However, before applying the grassland reference model to the annual cropland sites, the site properties of the annual cropland sites were assessed to determine if they were sufficiently comparable to the grassland sites (training data). This was done using the Area of Applicability (AOA) framework established by Meyer and Pebesma (2021). All covariates included in the grassland reference model were included in the AOA assessment, weighted by variable importance. In total, 66 annual cropland sites (2.8% of the total number of annual cropland sites; 49 German sites, 13 Danish sites, 2 Flemish sites, 2 French sites) did not fall within the AOA and were discarded from the dataset before the grassland reference model was applied. The primary difference of the excluded sites was higher organic C content as well as C/N ratio, and they had a higher prevalence of Histosol and Podzol reference soil groups, as well as a lower prevalence of Luvisol reference soil groups in annual croplands as compared to the permanent grassland sites.

Subsequently, the difference in subsoil BD from the estimated reference value and the observed value (i.e., the estimated anthropogenic subsoil compaction; $\Delta BD_{anthropogenic}$) was calculated by subtracting the estimated reference bulk density (BD_{ref}) from the measured subsoil bulk density (BD_{obs}) for each annual cropland site (Equation (2)):

$$\Delta BD_{anthropogenic} = BD_{obs} - BD_{ref} \tag{2}$$

2.4.3 | Explaining the Difference in Subsoil BD

A second Random Forest model was trained on the calculated change in BD ($\Delta BD_{anthropogenic}$) of the annual cropland sites in order to analyse the effect of selected pedogenic, climatic and agricultural management covariates (Table S1). The Spearman's rank correlation coefficient of the covariates was calculated for the covariates; no correlation between covariates was above 0.75, and all covariates were retained. Similar to the grassland reference Random Forest model, this post hoc model was implemented with 1000 trees and mtry equal to the square root of the number of covariates. The target variable was $\Delta BD_{anthropogenic}$, and the independent covariates were organic C content, climatic parameters (MAT, MAP), soil texture (clay, silt), pH, and the NUTS2-derived management covariates (Table S1). The post hoc model was evaluated using five-fold random cross-validation.

2.4.4 | Correcting National Estimates for Regression to the Mean

Random Forest models are known to both underestimate high values and overestimate low values (Figure 3A), that is, perform well on average and for the middle range of covariates but poorly on individual observations (regression to the mean), especially for those with more extreme covariate values (Hengl et al. 2018). For individual sites, the model performance might therefore be inaccurate for BD values at the ends of the range (Figure 3A). Soil parameter estimates become more robust with spatial aggregation (Skalský et al. 2024), so the mean $\Delta BD_{anthropogenic}$ for the country or region ($\overline{\Delta BD_{country}}$) would be a more accurate estimate of the anthropogenically driven increase in subsoil BD. Thus, $\overline{\Delta BD_{country}}$ is subtracted from the site-specific observed BD (BD_{obs}) (Equation 3), which gives a corrected estimate of the reference BD for each annual cropland site:

$$BD_{ref,\ corrected} = BD_{obs} - \overline{\Delta BD_{country}} \tag{3}$$

The corrected estimate for the reference BD ($BD_{ref,\ corrected}$) is the reference value used in calculations to determine the proportion of annual cropland sites that surpass the packing density threshold (Section 2.5) without anthropogenic influence (i.e., naturally) for each country and region as well as across all observations. The site-specific BD_{obs} is used to calculate the proportion of annual cropland sites currently surpassing the packing density threshold. The difference between these two proportions gives the proportion of anthropogenically compacted annual cropland sites, indicative of the current area proportion of anthropogenically compacted annual cropland soil in the assessed areas of Europe.

2.5 | Packing Density

The subsoil compactness, that is, how densely packed the soil is, is evaluated using packing density (PD, $\mathrm{g\,cm^{-3}}$), which accounts for the variability in natural compactness across soil texture. Packing density is calculated after Renger et al. (2014) using Equation 4:

$$PD = BD + 0.005 * Clay + 0.001 * Silt \tag{4}$$

where BD is the fine soil bulk density ($\mathrm{g\,cm^{-3}}$), clay and silt are the mass proportion (%) of clay and silt particle size fractions. The multiplying factors of the clay and silt content have the unit $\mathrm{g\,cm^{-3}\,\%^{-1}}$. Going forward, the packing density takes on the subscript identifier of the bulk density used to calculate the packing density, for example, PD_{obs} is the packing density corresponding to the observed bulk density (BD_{obs}). To assess whether or not a site is compacted to the degree that soil functions are affected, a threshold packing density of $1.71\,\mathrm{g\,cm^{-3}}$ is used as suggested by Renger et al. (2014).

As Renger et al. (2014) argued that an equation that includes silt alongside clay and the corresponding threshold was more accurate than a previously used equation that only considers clay content (Equation 5), we chose to use the former equation which accounts for the silt fraction. In order to facilitate comparability to previous and potential future research, the results relating to packing density were also calculated using Equation 5, which has previously been used by, for example, Huber et al. (2008) and Panagos et al. (2024). When using Equation 5, a packing density threshold of $1.75\,\mathrm{g\,cm^{-3}}$ is typically used (Renger 1970). The results using Equation 5 and the threshold of $1.75\,\mathrm{g\,cm^{-3}}$ are given in Table S2 and Figure S4.

$$PD_{alternative} = BD + 0.009 * Clay \tag{5}$$

As soil texture is considered to be a static site property at a given site, a change in packing density must derive from changes in bulk density. Thus, for a given site, the change in bulk density is equal to the change in packing density ($\Delta BD = \Delta PD$).

In this study, a given site was considered compacted if the observed PD (PD_{obs}) was above the threshold of $1.71\,\mathrm{g\,cm^{-3}}$. A given site was considered naturally compact if the estimated grassland reference PD ($PD_{ref,\ corrected}$) was above the threshold. Due to the correction performed in Equation 3, all sites where $PD_{ref,\ corrected}$ surpasses the threshold also had PD_{obs} above the threshold, and all sites estimated to be naturally compact were also observed to be compacted.

2.6 | Data analysis and statistics

All analyses were performed in R version 4.3.1 (R Core Team 2023) and have relied on the' tidyverse' framework (Wickham et al. 2019). The Random Forest models were run using the 'ranger' package (Wright and Ziegler 2017). Maps were produced using the giscoR package (Hernangómez 2024) and using the 'ggparliament' package (Hickman et al. 2018). Eurostat information was accessed using the 'eurodata' package (Rutkowski 2023), and the texture triangle was generated using the 'ggtern' package (Hamilton and Ferry 2018).

R scripts containing the training and evaluation of both the grassland reference model and post hoc model, as well as statistical analyses, are available on Github (https://github.com/LauraHarbo/SoilCompaC_publication/).

3 | Results

3.1 | Estimated Anthropogenic Subsoil Compaction in Cropland

Based on the estimated difference between the observed BD and the predicted grassland reference BD, agricultural management activities have on average increased the BD of the upper subsoil of all the annual cropland sites by between 0.00 and $0.05\,\mathrm{g\,cm^{-3}}$ as compared to permanent grasslands, depending on the country or region (mean $0.03\,\mathrm{g\,cm^{-3}}$ across all sites) (Table 2).

At the site level, the estimated human-induced changes in BD were highly variable, including both estimated decreases and increases in subsoil BD of more than $\pm 0.3\,\mathrm{g\,cm^{-3}}$ (Figure S2). The estimated change in subsoil BD due to anthropogenic activities did not have a clear spatial pattern; estimated increases and decreases in subsoil BD were found across all regions (Figure 4A). The largest mean

TABLE 2 | Mean anthropogenically driven change in subsoil BD of annual croplands ($\mathrm{g\,cm^{-3}}$) and standard deviation ($\mathrm{g\,cm^{-3}}$) as estimated by reciprocal modelling for each country and region.

Country/ region	Number of annual cropland sites	Mean difference $\mathrm{g\,cm^{-3}}$	Standard deviation $\mathrm{g\,cm^{-3}}$
Denmark	110	0.04	0.118
Flanders	109	0.05	0.076
France	420	0.00	0.104
Germany	1651	0.04	0.093
All	2290	0.03	0.097

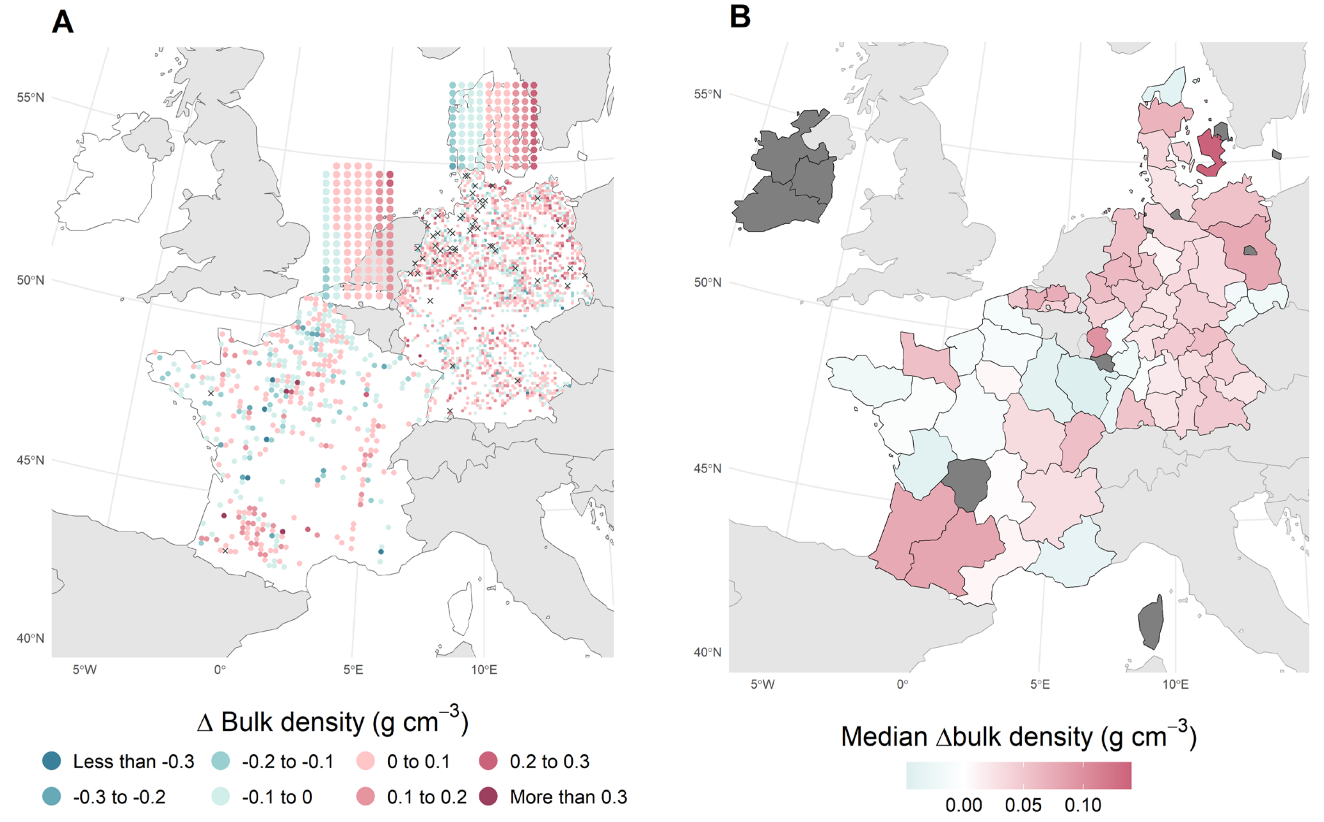

FIGURE 4 | (A) Estimated change in subsoil bulk density (g cm^{-3}) for each annual cropland site ($n = 2292$); the colour classes indicate estimated difference in subsoil bulk density. Black crosses indicate sites that did not fall within the area of applicability ($n = 66$, 13 not shown for Denmark). Point size varies by survey for readability purposes. The Danish and Flemish sites are arranged in a grid by estimated degree of change for visualisation purposes, as the coordinates are not publicly available. Ireland is excluded due to low number of sites. (B) Median estimated change in subsoil bulk density (g cm^{-3}) for each NUTS2 region; NUTS2 regions with less than 5 annual cropland sites that pass AOA are shaded dark grey. EuroGeographics for the administrative boundaries (NUTS2).

estimated change in subsoil BD was found in Denmark, Flanders and Germany (Table 2). The estimated mean change in subsoil BD varied across NUTS2 regions, and an average decrease in subsoil BD was found in six German regions, 11 regions in France and one in Denmark (Figure 4B). The greatest estimated increase in subsoil BD was observed in south-western France, eastern Germany and eastern Denmark (Figure 4B).

On average, the estimated increase in subsoil bulk density in France was 0. However, both extreme increases and decreases were found for individual sites in France (Figure 4A, Figure S2). For sites from France, decrease in subsoil BD was found more frequently (49% of annual cropland sites in France) as compared to Denmark (37%), Germany (33%) and Flanders (21%). Furthermore, the estimated negative changes were more pronounced in France as compared to the other countries and regions (Figure 2B). The negative changes in subsoil BD in France were primarily found in the regions of Lorraine, Bretagne, Champagne-Ardenne and Auvergne (Figure 4A).

3.2 | Anthropogenic Change in Subsoil BD Versus Pedo-Climatic and Management Variables (Post Hoc Model)

A second Random Forest model (post hoc model) was fitted to relate the estimated anthropogenic effect on the changes in subsoil BD to pedo-climatic variables (soil texture, organic C content, pH, MAT and MAP) and farm management (mean farm size, mean proportion of potatoes and sugar beets in crop rotation, and mean proportion of area under conventional tillage by NUTS2 region) (Table S1). The RMSE of the model was 0.009 (g cm^{-3}) with an R^2 value of 0.16 in fivefold random cross-validation. Restricting model fits to the individual surveys did not yield better overall model performance.

Silt and clay contents were the most important variables for explaining the estimated human-induced change in subsoil BD, followed by MAP and the organic C content (Figure 5). The area-aggregated management variables were generally less influential than pedo-climatic variables, with the exception of the mean farm size (Figure 5).

The estimated anthropogenically driven change of subsoil BD was highest at sites with the USDA soil texture categories sandy loam, sandy clay loam and loam, ranging from 10% to 20% clay and 20% to 30% silt (Figure 6A). Above about 30% silt, estimated changes in subsoil BD decreased with increasing silt contents (Figure 6C). For sites with low organic C content, there was a negative correlation between organic C content and estimated change in BD, while a positive correlation was found for sites with comparably higher organic C content (Figure 6D). The data indicated that, in the case of small farms (less than 80 ha), the smaller the farm size, the larger the change in BD. However, this negative correlation between farm size and estimated change in BD did not appear consistent for larger farms. There was a slight positive correlation between pH and BD change, with the exception of alkaline soils with pH values above 8. For the remaining covariates in the post hoc model, there were no strong correlations between the covariates in the range of the majority of observations and the estimated difference in subsoil BD (Figure 6).

3.3 | Extent of Anthropogenic Subsoil Compaction

Across the countries and region included in the study, the proportion of sites where the observed packing density (PD_{obs}) exceeded the packing density threshold of 1.71 g cm^{-3} varied (Figure 7C); the highest proportion of sites surpassing the threshold was found in Flanders (52%), followed by Germany (39%), France (28%) and Denmark (14%). Across all annual cropland sites in the study, 36% surpassed the packing density threshold of 1.71 g cm^{-3} (Table 3).

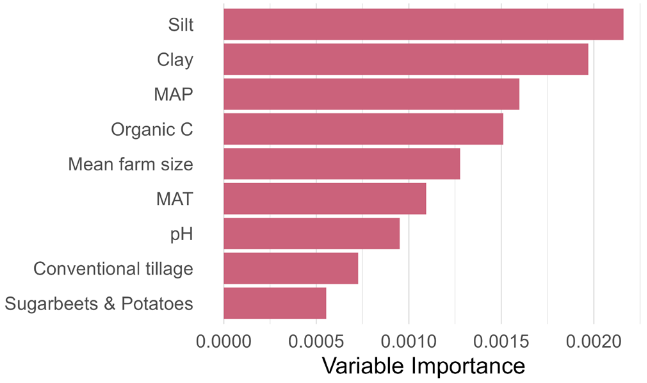

FIGURE 5 | Variable importance of the independent covariates of the post hoc model.

Overall, the proportion of sites where the estimated natural packing density surpassed the threshold was similar for all countries (approximately 25%), except Denmark, which was lower than the others (10%) (Table 3); this corresponds well with the data from the permanent grassland sites, where the proportion of sites that surpass the threshold ranges from 22% to 27%, except for Denmark, where no permanent grassland sites surpassed the threshold, and France, where 11% of the permanent grassland sites surpassed the threshold. In Denmark and France, the proportion of annual cropland sites that surpassed the packing density threshold due to anthropogenic influence was comparatively low (4% and 0%, respectively), while this proportion was higher for Flanders and Germany (25% and 13%, respectively) (Table 3). Thus, not only was the proportion of subsoil in Flanders and Germany that surpassed the packing density threshold of 1.71 g cm^{-3} larger than in Denmark and France, but a comparatively larger proportion of the sites surpassed the packing density threshold due to anthropogenic activities; 47% and 34% for Flanders and Germany, but only 0% and 27% for France and Denmark (Table 3).

While the sites with a packing density surpassing the threshold of 1.71 g cm^{-3} were grouped together in south-western and central north-eastern France (Figure 7A), the spatial distribution of the German sites with packing density above the threshold was more even across the country, with larger proportions found in eastern and southern Germany (Figure 7B). In Germany, there was no discernible difference in the distribution of naturally and anthropogenically compacted sites. Overall, anthropogenically compacted sites occurred in close proximity to naturally compact sites.

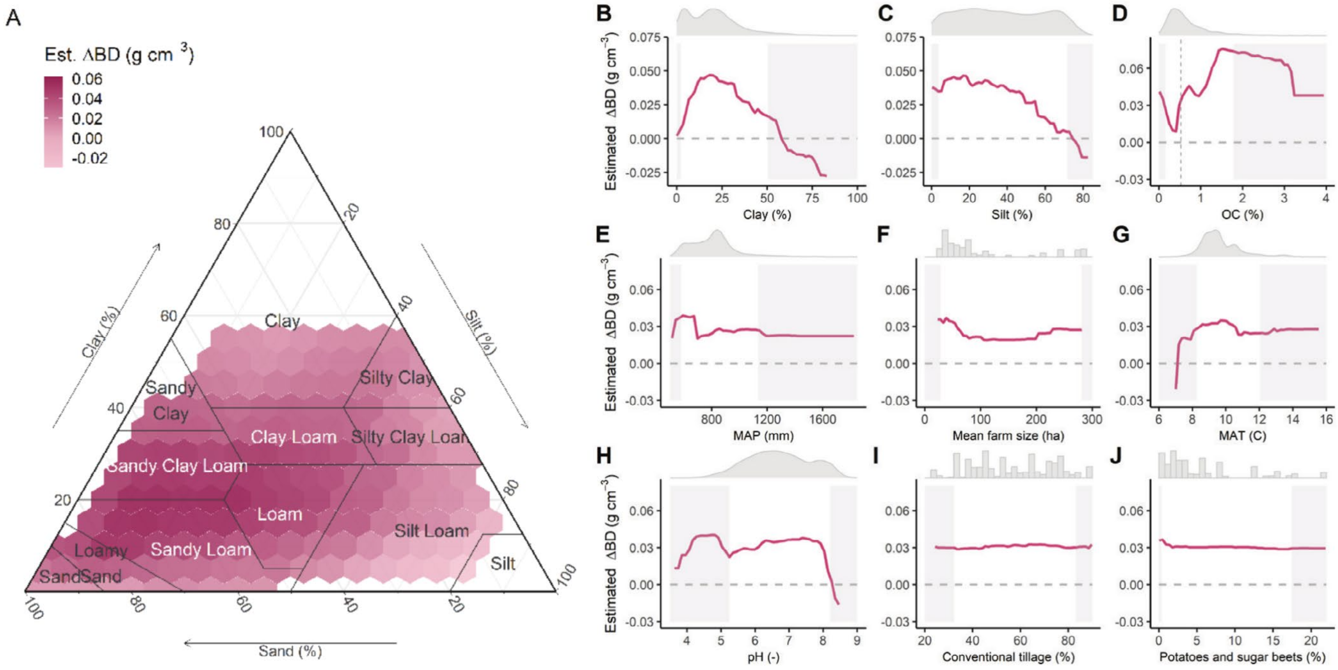

FIGURE 6 | (A) Partial dependence of the estimated changes in subsoil BD on soil texture (clay, silt and sand) and (B–J) univariate partial dependence plots for all variables in the post hoc model with variable distribution plots; histograms for data from Eurostat as data are grouped by NUTS2 regions. The grey boxes shade the lowest and highest 5% of observations, respectively. In (D), the vertical line indicates the mean organic C value.

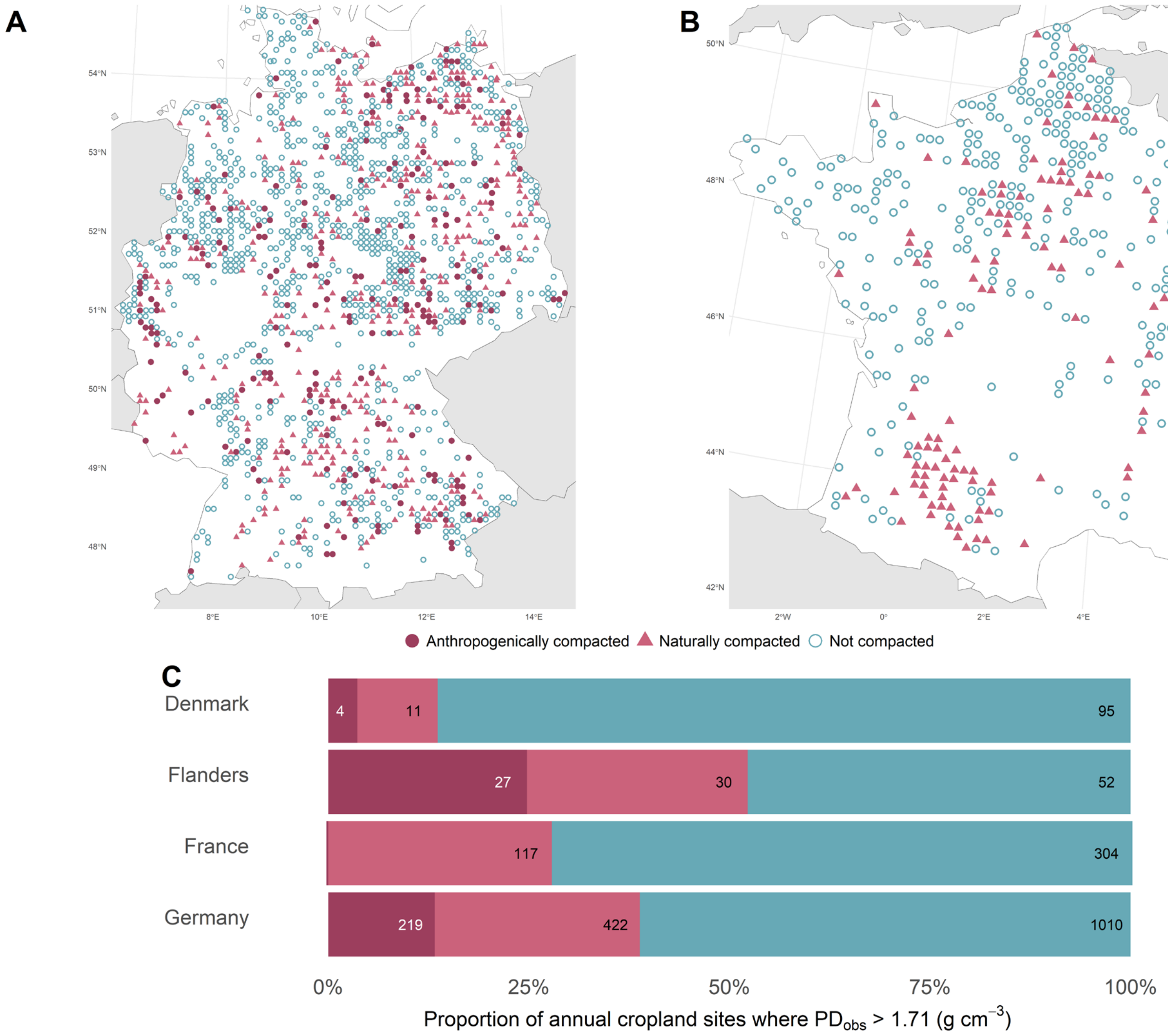

FIGURE 7 | Annual cropland sites where the observed packing density (PD_{obs}) surpassed the threshold of 1.71 g cm^{-3} in the subsoil for (A) Germany and (B) France. Coordinates from Denmark and Flanders were not publicly available. (C) The relative proportion of annual cropland sites where the observed packing density (PD_{obs}) surpassed the threshold of 1.71 g cm^{-3} in the subsoil. For all plots, the darker red shade ('Anthropogenically compacted') indicates that only PD_{obs} surpassed the threshold, while the lighter red shade ('Naturally compacted') indicates that both PD_{obs} and the reference packing density ($PD_{ref,\ corrected}$) surpassed the threshold. The blue shade indicates that PD_{obs} did not surpass the threshold. The number of observations (*n*) in each class is given on the bars. EuroGeographics for the administrative boundaries; for visualisation purposes, the maps do not share scale.

TABLE 3 | Number and relative proportion (%) of annual cropland sites that surpassed the threshold for compacted soil (1.71 g cm^{-3}) naturally ($PD_{ref,\ corrected}$), currently (PD_{obs}) and their estimated absolute and relative increase due to anthropogenic activities.

Country/Region	Sites where observed PD surpassed threshold; $PD_{obs} >= 1.71\ \mathrm{g\,cm^{-3}}$		Sites that naturally surpassed threshold; $PD_{ref,\ corrected} >= 0.71\ \mathrm{g\,cm^{-3}}$		Sites that surpassed threshold due to anthropogenic activities		Proportion of compacted sites due to anthropogenic activities
	n	**%**	***n***	**%**	***n***	**%**	**%**
Denmark	15	14	11	10	4	4	27
Flanders	57	52	30	28	27	25	47
France	116	28	117	28	0	0	0
Germany	641	39	422	26	219	13	34
All	829	36	580	25	249	11	30

4 | Discussion

4.1 | Extent and Severity of Subsoil Compaction

Depending on the country or region, 14%–52% of all assessed annual cropland sites on mineral soil had a subsoil PD_{obs} above the threshold of $1.71\,\mathrm{g\,cm^{-3}}$ (Figure 7 and Table 3), and on average up to 30% of the annual cropland sites that surpass the threshold do so due to anthropogenic influence. This is the first quantitative estimate based on observational data of both the extent and severity of subsoil compaction at a European scale, and the first time the anthropogenic component of subsoil compaction is quantified and mapped.

The range of estimates of the extent of subsoil compaction found in this study (Table S2) is similar to the general estimate of 20% across Europe by Oldeman et al. (1991) of undocumented origin, and also aligns well with the estimated 25% of European cropland with high texture-corrected subsoil density by Schjønning et al. (2015a).

In general, sites with anthropogenically compacted subsoil, that is, $PD > 1.71\,\mathrm{g\,cm^{-3}}$, were found in close proximity to sites where the natural subsoil packing density was estimated to be above the packing density threshold (Figure 7). Assuming that $PD_{reference}$ would be similar for sites that are closer together, this proximity between anthropogenically compacted sites and naturally more compact sites suggests that the anthropogenically compacted sites might have a relatively high $PD_{reference}$. Such sites would thus require a smaller increase in subsoil packing density to surpass the threshold as compared to sites with a lower $PD_{reference}$. Conversely, even sites with a larger predicted increase in subsoil BD did not always have a subsoil packing density that surpassed the packing density threshold (Figures 4, 7, S5). Consequently, merely comparing the value with the $1.71\,\mathrm{g\,cm^{-3}}$ threshold may not be sufficient for fully characterising the attribution of compaction to anthropogenic actions.

However, larger increases in subsoil BD were found more frequently for sites determined to be compacted, regardless of cause (i.e., only anthropogenic or both naturally and anthropogenically). In fact, the sites estimated to be naturally compact had a larger predicted increase in subsoil compactness (Figure S5), cementing that such sites may be further compacted by anthropogenic activities in addition to the naturally denser state. Overall, the packing density of the majority of the annual cropland sites is increased by agricultural activities, regardless of whether or not the observed packing density surpasses the threshold.

The reciprocal modelling predicted an increase in subsoil BD for the majority of annual cropland sites individually as well as for clusters of administrative NUTS2 regions as compared to grasslands; however, some regions had an overall negative estimated change in BD. While it is to be expected that some sites might have an estimated decrease in BD due to the Random Forest model's regression towards the mean, these negative values for entire regions were unexpected and should be investigated further. One contributing factor could be an unknown, regionally important covariate that was not included in the current version of the grassland reference model (e.g., soil depth, or specific climatic or pedogenic covariates). Another factor contributing to the unexpected negative results for NUTS2 regions may be uneven distribution of observations combined with differences in areas within the regions. The French sampling network is less dense than the other surveys (i.e., 16 × 16 km grid for France; 8 × 8 in Germany, 7 × 7 in Denmark), which may cause larger variability in the observed soil parameters. Given the parameters of the filtering criteria (Section 2.3) are likely unevenly distributed in space, sites are unlikely to have been filtered out evenly across space, thus leaving some regions with comparatively fewer points per area, reducing both the number of reference (grassland) sites and annual cropland sites. As such, the information available for the Random Forest models is scarcer, likely yielding poorer results in data-sparse regions. Indeed, the regions where the number of data points is low in France overlap with the regions where negative changes in BD were found (Figure S6). Furthermore, societal regional boundaries such as NUTS2 regions rarely align well with the boundaries of pedological, geological and climatic regions, which in themselves often are gradual rather than distinct, and visual representation of the data using NUTS2 regions is unlikely to reflect patterns aligned with pedoclimatic parameters, but are nonetheless convenient for visualisation and policy purposes.

The dataset from Denmark was substantially older than the datasets from the other surveys (Table 1), and the French dataset is the second oldest (Table 1). The observations from the Danish and French sites might thus be less affected by more recent increased machinery weight than sites from the other, more recent surveys. Schjønning et al. (2015b) reported that the soil stress applied at 0.5 m depth by combined harvesters produced in Denmark increased by a factor of 1.7 between 1981 and 2009. This age difference in the data may be a contributing factor to the lower estimate of anthropogenic subsoil compaction in Denmark and France as compared to the other countries and regions. However, as the subsoil compaction due to field traffic is accumulated over time, the Danish and French sites may also have reached a comparable level of compaction by the sampling time, given the century-long history of agriculture in Europe. Additionally, the grassland reference model predicts the BD of sites without intense agricultural management and should thus be agnostic to the sampling time of the annual cropland sites. Thus, it is more likely that the negative values derive from overestimation by the grassland reference model rather than the measured BD and corresponding PD values being lower due to sampling further in the past.

The clay and silt content of Danish soils was relatively low, and the packing density would therefore be lower in Denmark on average as compared to other regions with otherwise comparable BD_{obs}. Additionally, the eastern-most region in Denmark, which is not represented by any sites in the study, has the highest proportion of clay content and loam soils (Adhikari et al. 2013). Inclusion of data from this region might increase the overall subsoil compaction estimate for Denmark as well as the proportion of Danish sites that surpass the packing density threshold. Indeed, the estimated increase in subsoil BD increased eastwards in Denmark (Figure 4B).

In Bretagne, one of the regions with an unexpected average decrease in BD in the present study, statistical models that explain soil properties have previously produced unexpected findings (Pacini et al. 2023). Additionally, the BD values from the French soil survey are generally lower than the BD values from the other four soil surveys. As such, smaller increases in BD are to be expected (Equation 2) when subtracting the estimated reference BD (trained on all available data) from the observed BD, which is likely to be lower in France than that of the general dataset. Furthermore, the proportion of naturally compact sites was similar between France, Germany and Flanders, suggesting that the reference model overestimates the reference values for France, resulting in a smaller estimated proportion of anthropogenically compacted sites alongside the smaller average estimated increase in subsoil packing density in France. When the grassland reference model was trained only on the French data, the overall estimated change in BD for the cropland sites approximately doubled, which is to be expected if the range of BD values of the training data was lower than the range of BD values of the data the model is applied. However, the spatial pattern of larger and smaller changes in BD remained (Figure S7). As such, the estimated spatial pattern of anthropogenic subsoil compaction of the current modelling approach appears robust.

Overall, anthropogenic subsoil compaction levels in the studied European cropland regions were heavily affected by soil texture. Loamy-textured subsoils were the most affected by anthropogenic compaction, mirroring the findings in Zhang et al. (2024). Loamy-textured soils can contain many macropores, and the cohesive forces binding the soil particles together are relatively weak and can easily be overcome when stress is applied to the soil profile (USDA 1996). Sandy soils, in contrast, generally show poor aggregation and are naturally more densely packed. Additionally, heavy clay soil showed a relatively low increase in subsoil BD, likely due to very strong binding forces between soil particles (Raghavan et al. 1977; Murray and Quirk 1990). The post hoc model further revealed a potential dependency of change in subsoil BD on organic C content, likely due to the increased elasticity and resistance to deformation of soils with higher soil organic C contents (Soane 1990).

4.2 | Anthropogenic Component of Subsoil Compaction

The estimate of mean anthropogenic increase in the subsoil BD of cropland soil up to 0.05 g cm^{-3}, depending on the country or region, is in good agreement with observations from controlled field experiments on soil compaction. Keller et al. (2021) observed that trafficking wet, loamy soil with front wheel loads of 9 Mg increased the BD at 30 cm (i.e., the boundary between topsoil and subsoil in this study) by 0.09 g cm^{-3} directly after and by 0.06 g cm^{-3} about 1 year after the controlled compaction event. In contrast to the effect of a single compaction event as described by Keller et al. (2021), this study's compaction estimate reflects the cumulative historic sum of the various human-induced processes causing subsoil compaction. This cumulative historic sum might first and foremost be explained by the trafficking on cropland soil with machinery of steadily increasing weight (Keller and Or 2022). The upper subsoil, the target of this study, could additionally be compacted by smearing due to tillage, and possibly wheeling at the furrow bottom with potential wheel slip when cropland is ploughed (Pulido-Moncada et al. 2019; Seehusen et al. 2019; Ren et al. 2022). Lastly, the historic conversion of natural land to cropland has resulted in a significant decline in soil organic matter in Europe (Sanderman et al. 2017) which has increased the BD of the soil (Schneider and Don 2019), although the magnitude of the impact of the historic legacy effect on the subsoil BD and soil organic carbon below 30 cm in cropland is under debate (Poeplau et al. 2011; Emde et al. 2024).

The relationships between the estimated anthropogenic changes in subsoil BD and farm management practices (mean farm size, mean proportion of area under conventional tillage, mean proportion of area used for potatoes and sugar beets) in the post hoc model were more difficult to interpret as the farm management data were available for NUTS2 regions rather than for each annual cropland site. Additionally, the farm management practices were geographically clustered and thereby correlated with both pedogenic and climatic covariates (Figure S2). There is therefore a high likelihood that the effects of farm management-related parameters were masked by other covariates, despite filtering out covariates with a Spearman's rank coefficient larger than 0.75. In fact, the farm management covariates based on NUTS2 regions other than mean farm size did not improve the performance of the post hoc model. Soil properties together with regional-specific climatic conditions can, thus, be considered more important than generalised farming practices for explaining differences in human-induced changes in subsoil BD in European annual croplands on mineral soils; however, site-specific management practices (e.g., frequency of heavy machinery usage, timing of field activities, etc.) are still likely highly influential factors on subsoil compaction.

Given the lack of evidence in this modelling exercise of the otherwise well-established correlation between field trafficking intensity and subsoil compaction (Seehusen et al. 2019; Pulido-Moncadaet al., 2019; Keller et al. 2021; Keller and Or 2022), targeted investigations and experiments will likely be able to explore management-driven effects on soil compaction more efficiently than a modelling approach. However, inclusion of site-specific management information may improve the performance of the post hoc model, given that such information is available.

To improve the use of the packing density threshold, the impacts of subsoil compaction on yields and soil function ought to be studied in greater detail to determine if the threshold of packing density ≥ 1.71 g cm^{-3} (1.75 g cm^{-3} when using Equation 5) is also a meaningful boundary value with regards to soil functions other than root growth (e.g., water transport). Furthermore, it may be necessary to investigate whether a single threshold value for all soil layers is the preferable approach or if a depth-specific limit would give more accurate estimates of the effect of subsoil compaction on soil functions and fertility, as deeper soil layers in general have higher packing density values due to greater BD with depth and since the root density decreases with depth. Additionally, assessment of the applicability of the given threshold should also be studied further across a larger range of pedoclimatic conditions in order to assess geographic relevance of the threshold. It is suggested that depth-, pedoclimatic- and soil function-dependent thresholds may offer a more accurate reflection of the overall impact of soil compaction on soil functions than a single threshold for all functions, pedoclimates and soil layers. Furthermore, other measures for quantifying soil compactness, such as anisotropy, may give additional insight into the dynamics, effects and causes of soil compaction (Mordhorst et al. 2019), which is not possible if only packing density is used to evaluate soil compactness.

4.3 | Study Limitations

Even though we in this study evaluated the most comprehensive dataset available on subsoil bulk density in Europe, it is difficult to quantify the annual cropland area within the assessed European regions that follows the area-relevant filtering criteria of the data (mineral soil and $<5\%$ coarse fragments). It was assumed that the included sites were representative of the area in question. Therefore, the proportion of sites that exceeded the packing density threshold can be interpreted as the proportion of study-relevant annual cropland area that exceeded the threshold for a given country or region.

Given the extensive filtering of the sites from the surveys due to the reasons outlined in Section 2.3, the data included in the study were likely less spatially representative than in the original surveys. In particular, many sites were excluded due to the filtering criterion of no more than 5% coarse fragments. Coarse fragments in soil samples can impact the measurement of BD in multiple ways; firstly, coarse fragments can be too large to fit within a ring sample or too numerous to allow for placement of the ring. Secondly, if the density of the coarse fragments is not measured, the estimate of the fine soil BD will be less accurate as a standard density of quartz will be used; this effect scales proportionally with the coarse fragment content. While other methods for measuring BD than ring samples are available, these are likely also affected by the presence of high volumes of coarse fragments (e.g., excavated volumes may also be complicated by coarse fragments). Given that the effect of disregarding (or misrepresenting) coarse fragments on the fine soil BD estimate is proportional to the coarse fragment volume (Harbo et al. 2022; Fenton et al. 2024), BD estimates for soils with high coarse fragment volume are more likely to be affected to a larger extent than sites with low amounts of coarse fragments. Furthermore, the surveys have different methods to sample sites with high coarse fragment content, making such measurements less comparable across surveys.

The grassland data were filtered based on the available land use history to exclude sites with recent influence of anthropogenic activities to achieve a reference model with minimal anthropogenic influence. However, detailed information about land use history and pedogenic horizons was not available from the Irish, Flemish or Danish datasets, and no information was available for any soil survey regarding the long-term land use history of the sites. If a proportion of the permanent grassland sites from these surveys has recent anthropogenic influence on the subsoil BD, the estimated human-induced change in subsoil BD would be underestimated. This underestimation was indeed found when the grassland reference model was trained on all available grasslands from Germany as compared to being trained only on German grasslands without identified Ap horizons (data not shown), proving the lingering effect of previous ploughing on BD despite land use change. Extending this concept, the anthropogenic subsoil compaction might have been further underestimated in this study, as permanent grasslands are managed, though less frequently than croplands, using heavy machines, including for harvest and fertilisation, which may have caused an increase in subsoil bulk density as compared to a theoretically unaffected site.

The grassland reference model both underestimates large BD values and overestimates low BD values (Figure 3), which is attributable to regression to the mean, as discussed in section 2.4. However, as soil bulk density generally correlates negatively with organic C content (De Vos et al. 2005; Schrumpf et al. 2011), there may be an indirect correlation between the organic C content and predicted change in BD. At sites where the organic C content is high, the bulk density can generally be expected to be low, and our predicted reference BD would therefore be more likely to be overestimated (Figure 3). A greater likelihood of overestimation of the reference means a higher risk of predicting a negative change in BD, and would also result in misassigning a decrease in BD to anthropogenic activities. The dynamic related to overestimation of low BD values may also be an explanatory component as to why the mean increase in subsoil BD in France is 0, given that the French BD values were, on average, lower compared to the other soil surveys included in the study.

On the other end of the spectrum, sites with low organic C content are more likely to have high BD and are, as such, at risk of underprediction of the reference BD by our approach. Thus, in our study, the estimated anthropogenic change in BD is more likely to be overestimated for sites with low organic C content, resulting in attributing changes to anthropogenic activities that are, in part, model artefacts.

Given that there are substantially more cropland sites with a relatively low organic C content as compared to the grassland training data, and that the attribution of negative changes in subsoil bulk density does not affect the designation of sites as either 'naturally compact' or 'anthropogenically compacted', the calculated proportion of anthropogenically compacted sites represents the upper limit of the proportion of sites that surpass the packing density threshold due to anthropogenic activities.

It may be feasible to enhance the grassland reference model by incorporating additional covariates to diminish the number of sites where the estimated anthropogenic subsoil compaction was negative. The population of annual cropland sites included in the study was overall comparable to the permanent grassland sites; the grassland reference model was trained as per the AOA assessment. However, the annual cropland sites where a decrease in subsoil BD was predicted might have been at the opposing end of the predictor space, or there might have been few or no grassland sites in the local area (e.g., NUTS2 region) of the annual cropland sites, which would then lack local references.

The underlying reasons for the occasional predicted decreases in subsoil BD thus remain unresolved; however, there was a clear correlation between sites with estimated loosening and sites with lower-than-average BD_{obs} (Figure 3). On one hand, natural processes and some management practices may be able to loosen the subsoil (e.g., deep rooting crops (Reintam et al. 2006; Flávio Neto et al. 2015) and sward lifting (De Boer et al. 2018)), but it is currently unknown whether such practices were implemented on the study sites. On the other hand, the estimated negative values might also be due to the fact that Random Forest models are not able to predict the highest and lowest values in the dataset (Kuhn and Johnson 2013), and thus the lowest BD values were generally overestimated while the highest values were likely underestimated by the grassland reference model. Overall, modelling limitations appear to be more reasonable explanations for the unexpected results, given that subsoil loosening is not typically observed in connection with human activities, and both a soil wetness indicator and MAP were included in the model and showed only minor importance for the model performance (Figure 2). That said, further investigation of regions of both predicted decrease in subsoil BD and/or overall low subsoil BD_{obs} might shed more light on general dynamics of subsoil BD and compaction processes of such regions.

Altogether, the limitations of the modelling approach result in regression to the mean, while the effect of potential recent anthropogenic influence on the permanent grassland sites would result in a systematic underestimation of the anthropogenic component of the current compactness of annual cropland sites in the investigated countries and regions. Additionally, the inherent negative correlation between organic C content and soil BD means that there is a risk of overestimating the anthropogenic component of the subsoil compaction. While the difference in the input data from the five regions and countries in the study may have caused some estimated changes in subsoil BD to be negative, the spatial pattern of subsoil BD increase is robust.

5 | Conclusion

The quantification of anthropogenically driven subsoil compaction conducted in this study represents a significant contribution to the assessment of the area affected by subsoil compaction in the studied regions of Europe. The methodology developed in the present study proved useful in elucidating a spatial pattern of subsoil BD change and in identifying the most important factors for explaining subsoil compaction in cropland. It may be applied to more datasets where the relevant covariates are available for both annual cropland and grassland sites.

Further research is required to clarify the effects of agricultural practices as well as local soil and climatic conditions on subsoil compaction, as the site-specific data on agricultural management practices were not available for this study. It is recommended that national and pan-national monitoring networks should include samples from the subsoil layer in addition to the topsoil, in order to facilitate further studies at larger scales. Furthermore, data regarding agricultural management practices should be accessible for sites in the monitoring networks, as this may facilitate additional research on the impact of human activities on soil at the European scale. Lastly, the development of pan-European standards for soil sampling or transfer functions to ensure higher degrees of comparability between datasets might further improve studies of soil parameters at a continental scale.

Author Contributions

Laura Sofie Harbo: investigation, data curation, formal analysis, visualization, writing – original draft, writing – review and editing, methodology. **Marine Lacoste:** data curation, resources, writing – review and editing. **Line Boulonne:** data curation, resources, writing – review and editing. **Julien Wengler:** data curation, resources, writing – review and editing. **Owen Fenton:** data curation, resources, writing – review and editing. **Giulia Bondi:** data curation, resources, writing – review and editing. **Patrick Tuohy:** data curation, resources, writing – review and editing. **Amélie Marie Beucher:** data curation, resources, writing – review and editing. **Mathieu Lamandé:** resources, data curation, writing – review and editing, project administration, funding acquisition. **Tommy D'Hose:** data curation, project administration, resources, funding acquisition, writing – review and editing. **Florian Schneider:** conceptualization, methodology, data curation, supervision, resources, project administration, writing – review and editing, writing – original draft, funding acquisition.

Acknowledgements

A special thank you to Elron Wiedermann for his guidance regarding the potential effects of various cropping systems and selection of agricultural management covariates from Eurostat, and to David Emde for running scripts on the HPC. Additional thanks to the soil organic matter research group at the Thünen Institute of Climate-Smart Agriculture for commenting and discussing previous versions of the manuscript. This study was funded by the European Union's Horizon 2020 research and innovation programme via the EJP SOIL project SoilCompaC (Grant No. 862695) and Horizon Europe AI4SoilHealth (Grant No. 101086179). Mathieu Lamandé's contribution to this study was also partly financially supported by the Ministry of Environment and Food of Denmark via the SOLGRAS project (GUDP Grant no. 39009-21-1850). Open Access funding enabled and organized by Projekt DEAL.

Data Availability Statement

Scripts for the reciprocal modelling approach, statistics and figures are available on GitHub (https://github.com/LauraHarbo/SoilCompaC-prelim/tree/main; temporary repository, a permanent repository will be created after the peer review process, when the analyses are finalized). Data included in the model is available from the direct sources, if data protect regulations and agreements with data managers allow.

References

Adhikari, K., R. B. Kheir, M. B. Greve, et al. 2013. "High-Resolution 3-D Mapping of Soil Texture in Denmark." *Soil Science Society of America Journal* 77, no. 3: 860–876. https://doi.org/10.2136/sssaj2012.0275.

Aitchison, J. 1982. "The Statistical Analysis of Compositional Data." *Journal of the Royal Statistical Society. Series B (Methodological)* 44, no. 2: 139–177.

Akker, J. J. H. van den. 2008. Environmental Assessment of Soil for Monitoring Volume I: Indicators & Criteria. 10.2788/93515.

Alaoui, A., M. Rogger, S. Peth, and G. Blöschl. 2018. "Does Soil Compaction Increase Floods? A Review." *Journal of Hydrology* 557: 631–642. https://doi.org/10.1016/j.jhydrol.2017.12.052.

Alvarez, C. R., M. A. Taboada, F. H. Gutierrez Boem, A. Bono, P. L. Fernandez, and P. Prystupa. 2009. "Topsoil Properties as Affected by Tillage Systems in the Rolling Pampa Region of Argentina." *Soil Science Society of America Journal* 73, no. 4: 1242–1250. https://doi.org/10.2136/sssaj2008.0246.

Arrouays, D., C. Jolivet, L. Boulonne, G. Bodineau, N. Saby, and E. Grolleau. 2002. "A New Initiative in France: A Multi-Institutional Soil Quality Monitoring Network." *Comptes rendus de l'Académie d'Agriculture de France* 88, no. 5: 93–103.

Asch, K. 2005. "Internationale Geologische Karte von Europa Und Den Angrenzenden Regionen 1:5.000.000 (IGME5000)." In *PDF, JPEG, TIFF, gedruckt, Shapefile*. Federal Institute for Geosciences and Natural Resources. https://doi.org/10.25928/IGME-5000.

Baritz, R., M. Fuchs, R. Hartwich, D. Krug, and S. Richter. 2005. "Soil Regions of the European Union and Adjecent Countries 1:5000000 Version 2.0 (EUR5000)." https://metadata.europe-geology.eu/record/full/AE71FFEE-1AE9-4624-AE3F-F49513FE9DCB?dlang=ger.

Batey, T. 2009. "Soil Compaction and Soil Management – a Review." *Soil Use and Management* 25, no. 4: 335–345. https://doi.org/10.1111/j.1475-2743.2009.00236.x.

Berisso, F. E., P. Schjønning, T. Keller, et al. 2012. "Persistent Effects of Subsoil Compaction on Pore Size Distribution and Gas Transport in a Loamy Soil." *Soil and Tillage Research* 122, no. June: 42–51. https://doi.org/10.1016/j.still.2012.02.005.

Berisso, F. E., P. Schjønning, T. Keller, et al. 2013. "Gas Transport and Subsoil Pore Characteristics: Anisotropy and Long-Term Effects of Compaction." *Geoderma* 195–196, no. March: 184–191. https://doi.org/10.1016/j.geoderma.2012.12.002.

Bondi, G., L. O'Sullivan, O. Fenton, R. Creamer, I. Marongiu, and D. P. Wall. 2021. "Trafficking Intensity Index for Soil Compaction Management in Grasslands." *Soil Use and Management* 37, no. 3: 504–518. https://doi.org/10.1111/sum.12586.

Buurman, P., A. G. Jongmans, and M. D. PiPujol. 1998. "Clay Illuviation and Mechanical Clay Infiltration—Is There a Difference?" *Quaternary International Revisitation of Concepts in Paleopedology* 51–52: 66–69. https://doi.org/10.1016/S1040-6182(98)90225-7.

Chamen, T., W. C. Moxey, A. P. Moxey, W. Towers, B. Balana, and P. D. Hallett. 2015. "Mitigating Arable Soil Compaction: A Review and Analysis of Available Cost and Benefit Data." *Soil and Tillage Research* Soil Structure and its Functions in Ecosystems: Phase matter & Scale matter 146: 10–25. https://doi.org/10.1016/j.still.2014.09.011.

Creamer, R., I. Simo, B. Reidy, et al. 2014. Irish Soil Information System - Synthesis Report.

De Boer, H. C., J. G. C. Deru, and N. Van Eekeren. 2018. "Sward Lifting in Compacted Grassland: Effects on Soil Structure, Grass Rooting and Productivity." *Soil and Tillage Research* 184, no. December: 317–325. https://doi.org/10.1016/j.still.2018.07.013.

De Vos, B., M. Van Meirvenne, P. Quataert, J. Deckers, and B. Muys. 2005. "Predictive Quality of Pedotransfer Functions for Estimating Bulk Density of Forest Soils." *Soil Science Society of America Journal* 69, no. 2: 500–510. https://doi.org/10.2136/sssaj2005.0500.

Emde, D., C. Poeplau, A. Don, S. Heilek, and F. Schneider. 2024. "The Centennial Legacy of Land-Use Change on Organic Carbon Stocks of German Agricultural Soils." *Global Change Biology* 30, no. 8: e17444. https://doi.org/10.1111/gcb.17444.

Emmet-Booth, J. P., G. Bondi, O. Fenton, et al. 2018. "GrassVESS: A Modification of the Visual Evaluation of Soil Structure Method for Grasslands." *Soil Use and Management* 34, no. 1: 37–47. https://doi.org/10.1111/sum.12396.

EU. 2023. "Proposal for a Directive on Soil Monitoring and Resilience." https://environment.ec.europa.eu/publications/proposal-directive-soil-monitoring-and-resilience_en.

Eurostat. 2024. "Agricultural Practices." https://doi.org/10.2908/EF_MP_PRAC.

Fenton, O., G. Bondi, C. J. Bracken, et al. 2024. "Relative and Absolute Difference in Soil Organic Carbon Stocks in Grassland Soils in Ireland: Impact of Rock Fragments, Bulk Density and Calculation Methods." *Geoderma Regional* 36, no. March: e00769. https://doi.org/10.1016/j.geodrs.2024.e00769.

Guillaume, T., D. Makowski, Z. Libohova, L. Bragazza, F. Sallaku, and S. Sinaj. 2022. "Soil Organic Carbon Saturation in Cropland-Grassland Systems: Storage Potential and Soil Quality." *Geoderma* 406: 115529. https://doi.org/10.1016/j.geoderma.2021.115529.

Håkansson, I., and R. C. Reeder. 1994. "Subsoil Compaction by Vehicles With High Axle Load—Extent, Persistence and Crop Response." *Soil and Tillage Research* 29, no. 2–3: 277–304. https://doi.org/10.1016/0167-1987(94)90065-5.

Hamilton, N. E., and M. Ferry. 2018. "{ggtern}: Ternary Diagrams Using {ggplot2}." *Journal of Statistical Software* Code Snippets 87, no. 3: 1–17. https://doi.org/10.18637/jss.v087.c03.

Hamza, M. A., and W. K. Anderson. 2005. "Soil Compaction in Cropping Systems." *Soil and Tillage Research* 82, no. 2: 121–145. https://doi.org/10.1016/j.still.2004.08.009.

Harbo, L. S., J. E. Olesen, Z. Liang, B. T. Christensen, and L. Elsgaard. 2022. "Estimating Organic Carbon Stocks of Mineral Soils in Denmark: Impact of Bulk Density and Content of Rock Fragments." *Geoderma Regional* 30: e00560. https://doi.org/10.1016/j.geodrs.2022.e00560.

Hengl, T. 2018. EcoDataCube. https://stac.ecodatacube.eu/dtm_elevation/collection.json.

Hengl, T., M. Nussbaum, M. N. Wright, G. B. M. Heuvelink, and B. Gräler. 2018. "Random Forest as a Generic Framework for Predictive Modeling of Spatial and Spatio-Temporal Variables." *PeerJ* 6, no. August: e5518. https://doi.org/10.7717/peerj.5518.

Hernangómez, D. 2024. "giscoR: Download Map Data From GISCO API – Eurostat Zenodo." https://doi.org/10.5281/ZENODO.4317946.

Hickman, R., Z. Meers, and T. J. Leeper. 2018. "Ggparliament: Parliament Plots." https://CRAN.R-project.org/package=ggparliament.

Hiltebrand, C., T. Keller, I. Bachmann Rieder, and S. Doetterl. 2025. "Changes in Soil Quality on Horse Paddock Trails and the Influence of Paddock Grids." *Soil Use and Management* 41, no. 1: e70028. https://doi.org/10.1111/sum.7002.

Horn, R., A. Mordhorst, H. Fleige, et al. 2020. "Soil Type and Land Use Effects on Tensorial Properties of Saturated Hydraulic Conductivity in Northern Germany." *European Journal of Soil Science* 71, no. 2: 179–189. https://doi.org/10.1111/ejss.12864.

Houskova, B., and L. Montanarella. 2007. Towards an European Soil Data Center in Support of the EU Thematics Strategy for Soil Protection.

Hu, W., J. Drewry, M. Beare, A. Eger, and K. Müller. 2021. "Compaction Induced Soil Structural Degradation Affects Productivity and Environmental Outcomes: A Review and New Zealand Case Study." *Geoderma* 395, no. August: 115035. https://doi.org/10.1016/j.geoderma.2021.115035.

Huber, S., G. Prokop, D. Arrouays, et al. 2008. *Environmental Assessment of Soil for Monitoring*. Vol. I, Indicators & Criteria. Publications Office. https://doi.org/10.2788/93515.

IUSS Working group. 2022. *World Reference Base for Soil Resources. International Soil Classification System for Naming Soil and Creating Legends for Soil Maps*. 4th ed. International Union of Soil Sciences (IUSS).

Jolivet, C., J.-L. Almeida-Falcon, P. Berché, et al. 2022. "French Soil Quality Monitoring Network Manual. RMQS2: Second Metropolitan Campaign, 2016–2027." https://doi.org/10.17180/KC64-NY88.

Jones, R. J. A., G. Spoor, and A. J. Thomasson. 2003. "Vulnerability of Subsoils in Europe to Compaction: A Preliminary Analysis." *Soil and Tillage Research* Experiences With the Impact and Prevention of Subsoil Compaction in the European Union, 73, no. 1: 131–143. https://doi.org/10.1016/S0167-1987(03)00106-5.

Jorajuria, D., L. Draghi, and A. Aragon. 1997. “The Effect of Vehicle Weight on the Distribution of Compaction With Depth and the Yield of Lolium/Trifolium Grassland.” *Soil and Tillage Research* 41: 1–12.

Kadžienė, G., L. J. Munkholm, and J. K. Mutegi. 2011. “Root Growth Conditions in the Topsoil as Affected by Tillage Intensity.” *Geoderma* 166, no. 1: 66–73. https://doi.org/10.1016/j.geoderma.2011.07.013.

Karger, D. N., O. Conrad, J. Böhner, et al. 2017. “Climatologies at High Resolution for the Earth's Land Surface Areas.” *Scientific Data* 4, no. 1: 170122. https://doi.org/10.1038/sdata.2017.122.

Keller, T., and D. Or. 2022. “Farm Vehicles Approaching Weights of Sauropods Exceed Safe Mechanical Limits for Soil Functioning.” *Proceedings of the National Academy of Sciences* 119, no. 21: e2117699119. https://doi.org/10.1073/pnas.2117699119.

Keller, T., T. Colombi, S. Ruiz, et al. 2021. “Soil Structure Recovery Following Compaction: Short-Term Evolution of Soil Physical Properties in a Loamy Soil.” *Soil Science Society of America Journal* 85, no. 4: 1002–1020. https://doi.org/10.1002/saj2.20240.

Kome, G., R. Kogge, B. Kogge Enang, B. Palmer Kfuban Yerima, and M. G. R. Lontsi. 2018. “Models Relating Soil pH Measurements in H2O, KCl and CaCl2 for Volcanic Ash Soils of Cameroon.” *Geoderma Regional* 14, no. September: e00185. https://doi.org/10.1016/j.geodrs.2018.e00185.

Kuhn, M., and K. Johnson. 2013. “Measuring Performance in Regression Models.” In *Applied Predictive Modeling*. Springer. https://doi.org/10.1007/978-1-4614-6849-3_5.

Lipiec, J., R. Horn, J. Pietrusiewicz, and A. Siczek. 2012. “Effects of Soil Compaction on Root Elongation and Anatomy of Different Cereal Plant Species.” *Soil and Tillage Research* 121, no. May: 74–81. https://doi.org/10.1016/j.still.2012.01.013.

Massey, P., C. O'Connor, P. Sills, et al. 2014. “Irish Soil Information System: Laboratory Standard Operating Procedures (2007-S-CD-1-S1), EPA STRIVE Programme 2007–2013, Report 7, EPA, Dublin,”.

Meyer, H., and E. Pebesma. 2021. “Predicting into Unknown Space? Estimating the Area of Applicability of Spatial Prediction Models.” *Methods in Ecology and Evolution* 12, no. 9: 1620–1633. https://doi.org/10.1111/2041-210X.13650.

Miller, R. O., and D. E. Kissel. 2010. “Comparison of Soil pH Methods on Soils of North America.” *Soil Science Society of America Journal* 74, no. 1: 310–316. https://doi.org/10.2136/sssaj2008.0047.

Moeys, J., W. Shangguan, R. Petzold, et al. 2024. Soiltexture: Functions for Soil Texture Plot, Classification and Transformation. https://cran.r-project.org/web/packages/soiltexture/index.html.

Mordhorst, A., H. Fleige, I. Zimmermann, et al. 2019. “Anisotropie der gesättigten Wasserleitfähigkeit in Böden der Hauptnaturräume Schleswig-Holsteins (Norddeutschland) unter Acker- und Grünlandnutzung.” *Die Bodenkultur: Journal of Land Management, Food and Environment* 70, no. 1: 33–45. https://doi.org/10.2478/boku-2019-0003.

Mulholland, B., and M. A. Fullen. 1991. “Cattle Trampling and Soil Compaction on Loamy Sands.” *Soil Use and Management* 7, no. 4: 189–193. https://doi.org/10.1111/j.1475-2743.1991.tb00873.x.

Munera-Echeverri, J.-L., L. Boulonne, N. P. A. Saby, et al. 2024. *Datasets on Bulk Density and Coarse Fragment Content From the French Soil Quality Monitoring Network*, 110767. Data in Brief. https://doi.org/10.1016/j.dib.2024.110767.

Murray, R. S., and J. P. Quirk. 1990. “Interparticle Forces in Relation to the Stability of Soil Aggregates.” In *Soil Colloids and Their Associations in Aggregates*, edited by M. F. De Boodt, M. H. B. Hayes, A. Herbillon, E. B. A. De Strooper, and J. J. Tuck, 439–461. Springer US. https://doi.org/10.1007/978-1-4899-2611-1_16.

Neto, F., J.,. E. da Costa Severiano, K. A. de Pinho Costa, W. S. G. Junnyor, W. G. Gonçalves, and R. Andrade. 2015. “Biological Soil Loosening by Grasses From Genus Brachiaria in Crop-Livestock Integration.” *Acta Scientiarum Agronomy* 37, no. September: 375–383. https://doi.org/10.4025/actasciagron.v37i3.19392.

Oldeman, L. R., R. T. A. Hakkeling, and W. G. Sombroek. 1991. *World Map of the Status of Human-Induced Soil Degradation*. International Soil Reference and Information Centre. https://www.isric.org/sites/default/files/isric_report_1990_07.pdf.

Oorts, K., F. Amery, S. Lettens, et al. 2024. Dataset Bodemkoolstofmonitoringnetwerk Cmon - 2022-2024. Departement Omgeving, Vlaams Planbureau Voor Omgeving, Brussel.

Østergaard, H. S., and P. Mamsen. 1990. *Kvadratnet for Nitratundersøgelser i Danmark: Oversigt 1986–1989*. Landbrugets Rådgivningscenter Landskontoret for Planteavl.

Pacini, L., P. Arbelet, S. Chen, et al. 2023. “A New Approach to Estimate Soil Organic Carbon Content Targets in European Croplands Topsoils.” *Science of the Total Environment* 900: 165811. https://doi.org/10.1016/j.scitotenv.2023.165811.

Panagos, P., D. De Rosa, L. Liakos, M. Labouyrie, P. Borrelli, and C. Ballabio. 2024. “Soil Bulk Density Assessment in Europe.” *Agriculture, Ecosystems & Environment* 364, no. April: 108907. https://doi.org/10.1016/j.agee.2024.108907.

Pawlowsky-Glahn, V., and J. J. Egozcue. 2006. “Compositional Data and Their Analysis: An Introduction.” *Geological Society, London, Special Publications* 264, no. 1: 1–10. https://doi.org/10.1144/GSL.SP.2006.264.01.01.

Poeplau, C., A. Don, H. Flessa, A. Heidkamp, A. Jacobs, and R. Prietz. 2020. “Erste Bodenzustandserhebung Landwirtschaft – Kerndatensatz.“ https://doi.org/10.3220/DATA20200203151139.

Poeplau, C., A. Don, L. Vesterdal, et al. 2011. “Temporal Dynamics of Soil Organic Carbon after Land-Use Change in the Temperate Zone – Carbon Response Functions as a Model Approach.” *Global Change Biology* 17, no. 7: 2415–2427. https://doi.org/10.1111/j.1365-2486.2011.02408.x.

Poeplau, C., A. Jacobs, A. Don, et al. 2020. “Stocks of Organic Carbon in German Agricultural Soils—Key Results of the First Comprehensive Inventory.” *Journal of Plant Nutrition and Soil Science* 183, no. 6: 665–681. https://doi.org/10.1002/jpln.202000113.

Poggio, L., L. M. de Sousa, N. H. Batjes, et al. 2021. “SoilGrids 2.0: Producing Soil Information for the Globe With Quantified Spatial Uncertainty.” *Soil* 7, no. 1: 217–240. https://doi.org/10.5194/soil-7-217-2021.

Priori, S., M. Zanini, V. Falcioni, and R. Casa. 2024. “Topsoil Vertical Gradient in Different Tillage Systems: An Analytical Review.” *Soil and Tillage Research* 236: 105947. https://doi.org/10.1016/j.still.2023.105947.

Pulido-Moncada, M., L. J. Munkholm, and P. Schjønning. 2019. “Wheel Load, Repeated Wheeling, and Traction Effects on Subsoil Compaction in Northern Europe.” *Soil and Tillage Research* 186: 300–309. https://doi.org/10.1016/j.still.2018.11.005.

R Core Team. 2023. “R: A Language and Environment for Statistical Computing.” Vienna, Austria. https://www.r-project.org/.

Rabot, E., M. Wiesmeier, S. Schlüter, and H.-J. Vogel. 2018. “Soil Structure as an Indicator of Soil Functions: A Review.” *Geoderma* 314: 122–137. https://doi.org/10.1016/j.geoderma.2017.11.009.

Raghavan, G. S. V., E. McKyes, and B. Beaulieu. 1977. “Prediction of Clay Soil Compaction.” *Journal of Terramechanics* 14, no. 1: 31–38. https://doi.org/10.1016/0022-4898(77)90025-8.

Reidy, B., I. Simo, P. Sills, and R. E. Creamer. 2016. “Pedotransfer functions for Irish soils – estimation of bulk density (BD) per horizon type.” *Soil* 2: 25–39.

Reintam, E., J. Kuht, K. Trükmann, and W. Leal Filho. 2006. “Effect of Annual Lupines Growing on Compacted Sandy Loam Stagnic Luvisol: (Wirkung Des Anbaus von Einjähriger Lupine Auf Einen Verdichteten

Sandig-Lehmigen Stagnic Luvisol).” *Archives of Agronomy and Soil Science* 52, no. 2: 171–181. https://doi.org/10.1080/03650340600581885.

Ren, L., W. M. Cornelis, G. Ruysschaert, J. De Pue, P. Lootens, and T. D'Hose. 2022. “Quantifying the Impact of Induced Topsoil and Historical Subsoil Compaction as Well as the Persistence of Subsoiling.” *Geoderma* 424, no. October: 116024. https://doi.org/10.1016/j.geoderma.2022.116024.

Renger, M. 1970. “Über Den Einfluß Der Dränung Auf Des Gefüge Und Die Wasserfurchlässigkeit Bindiger Böden.” *Mitteilungen Der Deutschen Bodenkundlichen Gesellschaft* 11: 23–28.

Renger, M., K. Bohne, and G. Wessolek. 2014. Bodenphysikalische Kennwerte und Berechnungsverfahren für die Praxis, Teil II.

Rutkowski, A. 2023. Eurodata: Fast and Easy Eurostat Data Import and Search. https://CRAN.R-project.org/package=eurodata.

Sanderman, J., T. Hengl, and G. J. Fiske. 2017. “Soil Carbon Debt of 12,000 Years of Human Land Use.” *Proceedings of the National Academy of Sciences* 114, no. 36: 9575–9580. https://doi.org/10.1073/pnas.1706103114.

Schäfer-Landefeld, L., R. Brandhuber, S. Fenner, H.-J. Koch, and N. Stockfisch. 2004. “Effects of Agricultural Machinery With High Axle Load on Soil Properties of Normally Managed Fields.” *Soil and Tillage Research* 75, no. 1: 75–86. https://doi.org/10.1016/S0167-1987(03)00154-5.

Schjønning, P., and M. Lamandé. 2018. “Models for Prediction of Soil Precompression Stress From Readily Available Soil Properties.” *Geoderma* 320, no. June: 115–125. https://doi.org/10.1016/j.geoderma.2018.01.028.

Schjønning, P., J. J. H. van den Akker, T. Keller, et al. 2015a. “Soil compaction.” In *Chapter 6 of Stolte, et al. (Eds.), Report on Soil Threats to Soil Quality*. Advances of Agronomy. EU-Project RECARE.

Schjønning, P., J. J. H. van den Akker, T. Keller, et al. 2015b. “Chapter Five - Driver-Pressure-State-Impact-Response (DPSIR) Analysis and Risk Assessment for Soil Compaction—A European Perspective.” In *Advances in Agronomy*, edited by D. L. Sparks, vol. 133, 183–237. Academic Press. https://doi.org/10.1016/bs.agron.2015.06.001.

Schneider, F., and A. Don. 2019. “Root-Restricting Layers in German Agricultural Soils. Part I: Extent and Cause.” *Plant and Soil* 442, no. 1: 433–451. https://doi.org/10.1007/s11104-019-04185-9.

Schneider, F., C. Poeplau, and A. Don. 2021. “Predicting Ecosystem Responses by Data-driven Reciprocal Modelling.” *Global Change Biology* 27, no. 21: 5670–5679. https://doi.org/10.1111/gcb.15817.

Schroeder, R., H. Fleige, C. Hoffmann, H. J. Vogel, and R. Horn. 2022. “Mechanical Soil Database—Part I: Impact of Bulk Density and Organic Matter on Precompression Stress and Consequences for Saturated Hydraulic Conductivity.” *Frontiers in Environmental Science* 10, no. February. https://doi.org/10.3389/fenvs.2022.793625.

Schrumpf, M., E. D. Schulze, K. Kaiser, and J. Schumacher. 2011. “How Accurately Can Soil Organic Carbon Stocks and Stock Changes Be Quantified by Soil Inventories?” *Biogeosciences* 8, no. 5: 1193–1212. https://doi.org/10.5194/bg-8-1193-2011.

Seehusen, T., R. Riggert, H. Fleige, R. Horn, and H. Riley. 2019. “Soil Compaction and Stress Propagation after Different Wheeling Intensities on a Silt Soil in South-East Norway.” *Acta Agriculturae Scandinavica Section B Soil and Plant Science* 69, no. 4: 343–355. https://doi.org/10.1080/09064710.2019.1576762.

Skalský, R., G. Barančíková, J. Makovníková, Š. Koco, J. Halas, and J. Kobza. 2024. “Regional Topsoil Organic Carbon Content in the Agricultural Soils of Slovakia and Its Drivers, as Revealed by the Most Recent National Soil Monitoring Data.” *Environmental Challenges* 14: 100816. https://doi.org/10.1016/j.envc.2023.100816.

Sleutel, S., T. D'Hose, G. Ruysschaert, S. Lettens, and B. De Vos. 2021. Actualisatie En Verfijning Van De Onderbouwing Van Een Methodiek Voor De Systematische Monitoring Van Koolstofvoorraden in De Bodem.

Soane, B. D. 1990. “The Role of Organic Matter in Soil Compactibility: A Review of Some Practical Aspects.” *Soil and Tillage Research* 16, no. 1: 179–201. https://doi.org/10.1016/0167-1987(90)90029-D. A Tribute to Prof. IR. H. Kuipers.

Soane, B. D., P. S. Blackwell, J. W. Dickson, and D. J. Painter. 1980. “Compaction by Agricultural Vehicles: A Review I. Soil and Wheel Characteristics.” *Soil and Tillage Research* 1: 207–237. https://doi.org/10.1016/0167-1987(80)90026-4.

Sonderegger, T., S. Pfister, and S. Hellweg. 2020. “Assessing Impacts on the Natural Resource Soil in Life Cycle Assessment: Methods for Compaction and Water Erosion.” *Environmental Science & Technology* 54, no. 11: 6496–6507. https://doi.org/10.1021/acs.est.0c01553.

Troldborg, M., I. Aalders, W. Towers, et al. 2013. “Application of Bayesian Belief Networks to Quantify and Map Areas at Risk to Soil Threats: Using Soil Compaction as an Example.” *Soil and Tillage Research* 132: 56–68. https://doi.org/10.1016/j.still.2013.05.005.

USDA. 1996. *Soil Quality Resource Concerns: Compaction. Soil Quality Information Sheet*. United States Department of Agriculture - Natural Resources Conservation Service. https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb5294092.pdf.

Venter, O., E. W. Sanderson, A. Magrach, et al. 2016a. “Global Terrestrial Human Footprint Maps for 1993 and 2009.” *Scientific Data* 3, no. 1: 160067. https://doi.org/10.1038/sdata.2016.67.

Venter, O., E. W. Sanderson, A. Magrach, et al. 2016b. “Sixteen Years of Change in the Global Terrestrial Human Footprint and Implications for Biodiversity Conservation.” *Nature Communications* 7, no. 1: 12558. https://doi.org/10.1038/ncomms12558.

Wickham, H., M. Averick, J. Bryan, et al. 2019. “Welcome to the Tidyverse.” *Journal of Open Source Software* 4, no. 43: 1686. https://doi.org/10.21105/joss.01686.

Wiermann, C., M. H. Bockwoldt, D. Uteau, S. Peth, A. Mordhorst, and H. Fleige. 2025. “Present State of Soil Structure: A Comparison of Cropland and Grassland Soils in North Germany.” *Soil and Tillage Research* 253: 106681. https://doi.org/10.1016/j.still.2025.106681.

Wright, M. N., and A. Ziegler. 2017. “Ranger : A Fast Implementation of Random Forests for High Dimensional Data in C++ and R.” *Journal of Statistical Software* 77, no. 1: 1–17. https://doi.org/10.18637/jss.v077.i01.

Xu, D., and A. Mermoud. 2001. “Topsoil Properties as Affected by Tillage Practices in North China.” *Soil and Tillage Research* 60, no. 1: 11–19. https://doi.org/10.1016/S0167-1987(01)00167-2.

Zhang, B., Y. Jia, H. Fan, et al. 2024. “Soil Compaction Due to Agricultural Machinery Impact: A Systematic Review.” *Land Degradation & Development* 35, no. 10: 3256–3273. https://doi.org/10.1002/ldr.5144.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.