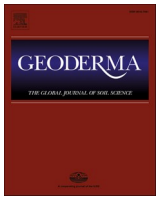
Geoderma 442 (2024) 116772

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma



Soil, climate, time and site factors as drivers of soil structure evolution in agricultural soils from a temperate-boreal region

Tobias Klöffel [a,*], Jennie Barron [a], Attila Nemes [b,c], Daniel Giménez [d], Nicholas Jarvis [a]

[a] Department of Soil and Environment, Swedish University of Agricultural Sciences, Uppsala, Sweden
[b] Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, Norway
[c] Division of Environment and Natural Resources, Norwegian Institute of Bioeconomy Research, Norway
[d] Department of Environmental Sciences, Rutgers University, New Brunswick, NJ, USA

ARTICLE INFO

Handling Editor: Y. Capowiez.

Keywords:
Agricultural intensification
Parent material
Pore-size distribution
Soil compaction
Soil organic carbon

ABSTRACT

The evolution of soil structure in agricultural soils is driven by natural and anthropogenic factors including inherent soil properties, climate and soil management interventions, all acting at different spatial and temporal scales. Although the causal relationships between soil structure and these individual factors are increasingly understood, their relative importance and complex interactive effects on soil structure have so far not been investigated across a geo-climatic region.

Here we present the first attempt to identify the relative importance of factors that drive the evolution of soil structure in agricultural soils as well as their direction of effect with a focus on the temperate-boreal zone. This was done using a random forest (RF) approach including soil, climate, time, and site factors as covariates. Relative entropy, as quantified by the Kullback-Leibler (KL) divergence, was used as a quantitative index of soil structure, which is derived from the particle-size distribution and soil water retention data, and integrates the effects of soil structure on pores from the micrometre-scale to large macropores. Our dataset includes 431 intact topsoil and subsoil samples from 89 agricultural sites across Sweden and Norway, which were sampled between 1953 and 2017. The relative importance of covariates for the evolution of soil structure was identified and their non-linear and non-monotonic effects on the KL divergence were investigated through partial dependence analysis. To reveal any differences between topsoils (0–30 cm; $n = 174$) and subsoils (30–100 cm; $n = 257$), the same analysis was repeated separately on these two subsets.

The covariates were able to explain on average more than 50% of the variation in KL divergence for all soil samples and when only subsoil samples were included. However, the predictions were poorer for topsoil samples (≈ 35%), underlining the complex dynamics of soil structure in agricultural topsoils. Parent material was the most important predictor for the KL divergence, followed by clay content for all soil samples and sampling year for only subsoil samples. Mean annual air temperature ranked third and annual precipitation ranked fourth for subsoil samples. However, it remains unclear whether the effects of climate factors are direct (e.g., freezing and thawing, wetting and drying, rainfall impact) or indirectly expressed through interactions with soil management. The partial dependence analysis revealed a soil organic carbon threshold of around 3% below which soil structure starts to deteriorate. Besides this, our results suggest that subsoil structure in the agricultural land of Sweden deteriorated steadily during the 1950′s to 1970′s, which we attribute to traffic compaction as a consequence of agricultural intensification. We discuss our findings in the light of data bias, laboratory methods and multicollinearity and conclude that the approach followed here gave valuable insights into the drivers of soil structure evolution in agricultural soils of the temperate-boreal zone. Theses insights will be of use to inform soil management interventions that address soil structure or soil properties and functions related to it.

* Corresponding author.
E-mail address: tobias.kloffel@slu.se (T. Klöffel).

https://doi.org/10.1016/j.geoderma.2024.116772
Received 17 July 2023; Received in revised form 18 December 2023; Accepted 3 January 2024
Available online 12 January 2024

1. Introduction

Soil structure is a crucial factor for sustaining the production capacity of agricultural systems by controlling the flow of water, air and nutrients to roots, determining their ability to explore the soil medium and by regulating the activity of soil biota (Bengough et al., 2011; Erktan et al., 2020; Or et al., 2021). Furthermore, soil structure governs the accessibility of organic matter to soil biota and therefore affects the flows of carbon, nutrients and energy (Janzen, 2015; Meurer et al., 2020a). It is thus of vital interest to identify the main drivers of structure formation and evolution.

Soils are structured at all scales, ranging from the largest pores formed by soil tillage and the actions of soil fauna and plant roots (Meurer et al., 2020a; Or et al., 2021) to the turnover of organic matter and the aggregation of clay minerals by so-called 'cementing agents' affecting mainly smaller pores (Fukumasu et al., 2022; Meurer et al., 2020b; Totsche et al., 2018). The state of soil structure in agricultural soils is ever-changing, driven by natural and anthropogenic factors, which act at a wide range of spatial and temporal scales (Bodner et al., 2008; Mohammed et al., 2020; Or et al., 2021).

Various metrics have been used to describe the current state of soil structure that refer either to the soil solid phase or pore space (Rabot et al., 2018). Commonly, these metrics only reveal information about soil structure at a specific scale, focusing mostly on the larger structural pores (Rabot et al., 2018). The soil water retention curve (SWRC), describing the relationship between the two major state variables of soil water – water content (θ) and pressure head (h) – has the potential to characterize soil structure across different spatial scales since it carries information about the pore-size distribution over a wide range of pore sizes (Rabot et al., 2018; Vogel et al., 2010). However, since the pore-size distribution is the result of both properties of the soil mineral phase (e.g., particle-size distribution, particle shape) and soil structure, a major challenge has been to separate the effects of soil texture and structure. While previous indices of soil structure based on the SWRC did not account the effects of soil texture (e.g., Dexter, 2004; Yoon and Giménez, 2012), this issue has been addressed in a recently proposed index of soil structure by Klöffel et al. (2022). This index quantifies the difference between two pore/void-size distributions (VSD) using the concept of relative entropy, also known as the Kullback-Leibler (KL) divergence (Kullback & Leibler, 1951), where one VSD is for the structured soil and the second is derived from soil texture for a hypothetical soil without structure. The KL divergence can therefore serve as an integrated measure of soil structure across spatial scales.

The KL divergence as an index of soil structure has so far only been tested on a small dataset (Klöffel et al., 2022). However, an important advantage of this index is that only data on particle-size distribution (PSD), total porosity and soil water retention are required. In this respect, an increasing number of databases are now available that include these properties from a multitude of soil types and locations that cover broad geo-climatic areas (e.g., European Commission, 2013; Gupta et al., 2022a; Lilburne et al., 2012; Nemes et al., 2001). In each soil sample in these databases, the pore space structure has developed under unique boundary conditions, that is, they were exposed to different soil management regimes and climates, formed on different parent materials, were located at different depths, and were sampled at different times. There is ample evidence that these and other soil forming factors have strong implications for the evolution and current state of soil structure and its effects on, for example, soil hydraulic properties (Bodner et al., 2013a; Gupta et al., 2022b; Hirmas et al., 2018; Lin et al., 2006; Or et al., 2021; Schlüter et al., 2011; Wu et al., 2023). Thus, it should be possible to identify the main drivers of soil structure formation in agricultural soils as well as their relative importance over a wide range of spatial and temporal scales by extracting quantitative information on soil structure from these databases.

The twin aims of this study were (i) to further test the utility of the KL divergence (Klöffel et al., 2022) as an integrated measure of soil structure on a larger dataset and (ii) to identify the main drivers of soil structure evolution and their relative importance in agricultural soils of a temperate-boreal region. To these ends, we made use of existing soil survey databases containing measurements of soil water retention and particle-size distribution for agricultural soils in Sweden and Norway and applied a random forest analysis to explore relationships between the KL divergence and a range of covariates characterizing soil, climate, time and site factors that are known or expected to affect soil structure.

2. Materials and methods

2.1. The KL divergence as an index of soil structure

The KL divergence is a quantitative measure of how one probability distribution differs from a second one, which is considered to be the reference distribution. Applied as an index of soil structure, the KL divergence is determined from two VSDs, where one VSD is that of the structured soil and the other that of a hypothetical so-called reference soil (Klöffel et al., 2022). The latter is defined as the same soil 'without structure'. This means that the VSD of the reference soil is in principle solely determined by the isotropic packing of soil particles, which we assume depends only on the PSD and packing density (Fiès and Bruand, 1998; Gupta and Larson, 1979). The VSD of the structured soil is derived from the measured SWRC, while the VSD of the hypothetical reference soil is derived from the measured PSD. The original approach followed by Klöffel et al. (2022) to derive the KL divergence of a soil was slightly modified in this study and can be broadly subdivided into three steps, which are described in the following subsections.

2.1.1. Modelling the pore-size distribution of the structured soil

The first step is to estimate the VSD of the structured soil from the SWRC. We translated the soil water pressure head (h) [cm] to an equivalent pore radius (r_v) [cm] using the Young-Laplace relationship to obtain (r_v, θ) pairs. Assuming physical properties of water at 20 °C and full contact between liquid and solid phases, this relationship can be written as (e.g., Brutsaert, 1966)

$$r_v = -\frac{0.149}{h} \tag{1}$$

Describing the SWRC of a structured soil requires a model that is able to account for both the 'textural' and 'structural' pore domain (Alaoui et al., 2011; Nimmo, 1997). The presence of these two domains is reflected in multi-, often bimodal SWRCs, for example due to the presence of soil macropores (Durner, 1994; Jensen et al., 2019; Ross and Smettem, 1993; Zhang et al., 2022). However, multimodality in SWRCs does not necessarily imply a well-developed soil structure, but can also result from the grading of soil particles (Fredlund et al., 2000; Klöffel et al., 2022). Conversely, soil structure formation at much smaller scales (e.g., micro-aggregation arising from the association of soil organic matter with mineral particles) does not likely contribute to multimodality in SWRCs (Hwang and Choi, 2006).

Here, we used the bimodal Kosugi (1996) model to describe the SWRC of the structured soil, where θ [$cm^3\ cm^{-3}$] is expressed as a function of r_v (e.g., Pollacco et al., 2017). The model is given as

$$\theta(r_v) = \theta_r + \frac{1}{2}(\theta_{s,1} - \theta_r)\,erfc\left[\frac{ln\left(\frac{r_v}{r_{v,m,1}}\right)}{\sqrt{2}\,\sigma_1}\right] + \frac{1}{2}\theta_{s,2}\,erfc\left[\frac{ln\left(\frac{r_v}{r_{v,m,2}}\right)}{\sqrt{2}\,\sigma_2}\right] \tag{2}$$

where θ_s [$cm^3\ cm^{-3}$] is the saturated water content, θ_r [$cm^3\ cm^{-3}$] is the residual water content, r_m [cm] is the median pore radius, σ [-] is the standard deviation of $\ln(r_v)$, *erfc(.)* denotes the complementary error function, and the subscripts '1′ and '2′ refer to the smaller and larger pore domain respectively. The VSD of the structured soil is obtained by differentiating Eq. (2) with respect to r_v.

2.1.2. Modelling the pore-size distribution of the reference soil without structure

The VSD of the hypothetical reference soil without structure is derived from the measured PSD of the fine-earth fraction (i.e., particles <2 mm in diameter). Many models have been developed to derive a SWRC from the cumulative PSD (e.g., Arya et al., 1999; Arya and Heitman, 2015; Chang et al., 2019; Haverkamp and Parlange, 1986; Mohammadi and Vanclooster, 2011; Pollacco et al., 2020; You et al., 2022), and one major difference among these models is how the relationship between particle and pore size is defined. Chang et al. (2019) demonstrated that a simple linear relationship yields satisfactory results for soils with mixed particle sizes. We therefore used their suggested factor of 0.3 to translate particle radius (r_p) into the pore radius of the reference soil. Thus, the i^{th} equivalent pore radius is given as

$$r_{v,i} = 0.3 r_{p,i} \tag{3}$$

However, we made two exceptions to the scaling factor of 0.3 in Eq. (3). First, the scaling factor for clay-sized particles ($r_p = 1$ μm) was reduced to 0.1, such that $r_v = 0.1 r_p$. The fraction of clay-sized particles is thereby directly related to the amount of water retained at permanent wilting point (θ_{pwp}), since the pore radius of 0.1 μm resulting from this equation is equivalent to $h = -15{,}000$ cm according to the Young-Laplace relationship. This modification was motivated by the strong correlation between clay content and water content at permanent wilting point that has been demonstrated in many studies (e.g., Bagnall et al., 2022; Chang et al., 2019; Kätterer et al., 2006; Pollacco et al., 2020), as well as the one presented here (see Subsection 2.2.2), suggesting that clay particles strongly contribute to the creation of pores of this equivalent radius. The second exception results from the notion that a scaling factor of 0.3 may not be justifiable for very sandy soils. Sand-sized particles often have a rounded shape similar to spheres. Consequently, for very sandy soils, the packing properties of the soil particles should be comparable with those of sphere packing. Furthermore, the presence of only sand-sized particles means that there is a lack of smaller particles that can fit within the voids created by larger grains (Fiès and Bruand, 1998). Indeed, preliminary calculations of the KL divergence showed unrealistic values for soils with sand contents larger than 92 %. We therefore used a different scaling factor for these soils, which we derived from Gupta & Larson (1979) who presented scaling factors for mono-sized sphere packing at different densities. We chose a scaling factor of 0.5, which is in line with the theoretical packing density we defined for coarse-textured soils without structure (see next paragraph).

To calculate the water contents of the (r_v, θ) pairs of the reference soil, we first defined the 'dry end' and the 'wet end' of the SWRC, that is, the water content at the permanent wilting point (θ_{pwp}) and at saturation (θ_s). For the dry end, we assumed that soil structure had no effect on θ_{pwp} and thus this value was set equal to that of the structured soil. For the wet end, θ_s was set equal to the total porosity of a soil without structure, which, in turn, was derived from our dataset in the following way: we split the dataset according to soil texture following the FAO classification scheme (FAO-UNESCO, 1974) into 'fine' (clay content >35 %), 'coarse' (clay content <18 %, sand content >65 %) and 'medium' (all others). The first percentile of the total porosities for each class was then used as total porosities of the reference soil amounting to 0.37, 0.31 and 0.34 $cm^3\ cm^{-3}$ for the classes fine, medium and coarse respectively. The remaining water contents for the (r_v, θ) pairs were calculated by assigning a fraction of the pore space between θ_s and θ_{pwp} to each pore radius proportional to the weight fraction of the corresponding particle radius (Arya and Heitman, 2015; Arya and Paris, 1981). In summary, θ corresponding to the i^{th} pore radius is given by

$$\theta_i = \begin{cases} \theta_{pwp}, i = 1 \\ \sum_{i=2}^{n-1} \left[(\theta_s - \theta_{pwp}) \dfrac{w_i - w_{i-1}}{1 - w_{clay}} \right] + \theta_{pwp}, i \geq 2 < n \\ \theta_s, i = n \end{cases} \tag{4}$$

where n is the total number of particle-size fractions, w_i [g g^{-1}] is the weight of the i^{th} particle-size fraction and w_{clay} [g g^{-1}] is the weight fraction of clay-sized particles ($r_p = 1$ μm).

As noted before, bimodality of SWRCs can also be induced by particle grading, for example, if the PSD peaks in both coarse and fine particle-size fractions (Fredlund et al., 2000; Pieri et al., 2006). To account for this possibility and to achieve a good description of the SWRC, we also fitted the bimodal Kosugi (1996) model in Eq. (2) to the (r_v, θ) pairs of the reference soil. As for the structured soil, the SWRC was differentiated with respect to r_v to obtain the VSD.

2.1.3. Calculating the KL divergence

Klöffel et al. (2022) presented an analytical solution to calculate the KL divergence, where the VSDs of both structured and reference soil are unimodal. However, as described in Subsections 2.1.1 and 2.1.2, we applied a bimodal model for both soils to achieve the best fits possible, for which no easy (if any) analytical solution exists. We therefore approximated the KL divergence numerically using the trapezoidal rule by subdividing the pore radius domain into N segments, each representing a trapezoid with a width of $(r_{v,upper} - r_{v,lower})/N$, where $r_{v,upper}$ and $r_{v,lower}$ are upper and lower pore radii:

$$D_{KL}(P||Q) = \left[\sum_{i=1}^{N} p(r_{v,i}) \log \frac{p(r_{v,i})}{q(r_{v,i})} \right] \times \frac{(r_{v,upper} - r_{v,lower})}{N} \tag{5}$$

In Eq. (5), D_{KL} denotes the KL divergence [-], $p(r_{v,i})$ and $q(r_{v,i})$ represent the value of the VSD of the structured soil and of the reference soil for the i^{th} segment boundary respectively. Here, we defined the upper pore radius limit as $r_{v,upper} = 1{,}500$ μm and the lower limit as $r_{v,lower} = 0.1$ μm. These choices were based on the assumptions that water retention measurements are unreliable for $h > -1$ cm (corresponding to $r_{v,upper}$ >1500 μm according to the Young-Laplace relationship) and, as noted before, that pores that remain water-filled at permanent wilting point are not affected by soil structure development.

2.1.4. Fitting the soil water retention curves

We fitted Eq. (2) to the measured and derived (r_v, θ) pairs of the structured and reference soil with the least squares non-linear fitting algorithm (Levenberg-Marquardt) implemented in the Python module *lmfit* (Newville et al., 2022). In fitting the SWRC, we set θ_r to zero for both reference and structured soil. This resulted in the optimization of six model parameters: $\theta_{s,1}$, $r_{m,1}$, σ_1, $\theta_{s,2}$, $r_{m,2}$ and σ_2.

The least squares fitting algorithm is sensitive to initial parameter values and thus, if not chosen properly, the algorithm may not converge to an acceptable solution. Here, the initial parameter values for the structured soil were derived from the PSD following a similar approach as described in Klöffel et al. (2022). For the reference soil, we tested different combinations of initial parameter values and found a set within physically reasonable limits (Fernández-Gálvez et al., 2021) that achieved good convergence in most cases. Table S1 shows the default initial parameter values for fitting the SWRC of the structured and reference soils. Each fit was evaluated visually and, if an acceptable convergence was not achieved with the default initial values, they were adjusted manually to see whether this improved the convergence. If this was not the case, the sample was excluded from further analysis.

2.2. Soil physical data

2.2.1. Description of datasets

As described in Section 2.1, the KL divergence as an index of soil structure only requires data on basic soil physical and hydraulic properties: soil texture, total porosity and soil water retention. The datasets from Sweden and Norway included in this study are described in the following.

2.2.1.1. Sweden. The majority of the Swedish soil physical data was obtained from the European HYdropedological Data Inventory (EU-HYDI) (European Commission, 2013). This data, which includes 1,596 samples from 187 soil profiles in arable land (including grassland for ley production), was collected by the Department of Soil Sciences at the Swedish University of Agricultural Sciences (SLU) between 1955 and 1973. The sampling distribution shows a decreasing coverage from south to north, generally reflecting the distribution of arable land in Sweden. However, it should be noted that, as a result of past research focus, the soils included in this dataset are not fully representative of the agricultural soils of Sweden with respect to their basic properties such as texture (Piikki and Söderström, 2019). In particular, the data shows a bias towards soils within the silty and clayey soil texture classes (see also Fig. 1 in Kätterer et al., 2006).

Each soil profile in this dataset was sampled at 10 cm intervals to a depth of at least 100 cm if the parent material allowed. The particle-size distribution of the fine earth fraction ($\leq$2 mm) was measured from disturbed soil samples using wet sieving for particle diameters >60 μm and using the pipette method for the smaller size fractions (Kätterer et al., 2006). The data were reported in six particle-size classes according to Swedish standards: clay (<2 μm), fine silt (2–6 μm), medium silt (6–20 μm), coarse silt (20–60 μm), fine sand (60–200 μm), and medium and coarse sand (200–2,000 μm). Four replicates of undisturbed soil cores (100 cm^3 volume) and disturbed soil samples (ca. 15 cm^3 volume) were used for water retention measurements (Wilkert, 1983). On the undisturbed cores, (h, θ) pairs were obtained using sand boxes at $h \geq -100$ cm and pressure plate apparatus at -100 cm $> h > -5{,}000$ cm, while disturbed samples were used to measure (h, θ) pairs at $h < -5{,}000$ cm on the pressure plate apparatus. Total porosity was derived from particle density and dry bulk density, where the former was determined from volume displacement of a fine earth sub-sample in ethyl alcohol and the latter from sample drying at 105 °C for 48 h.

We included additional data collected by the Department of Soil and Environment at SLU between 1974 and 1997, which is not available in the EU-HYDI database. This dataset included measurements on 591 soil samples from 64 sites across Sweden, sampled at varying depths, with the measurement procedures and data reporting carried out as described above. Finally, data from an additional 45 samples were obtained from a series of soil descriptions of long-term soil fertility experiments (Kirchmann, 1991; Kirchmann et al., 2005; Kirchmann et al., 1999; Kirchmann et al., 1996; Kirchmann and Eriksson, 1993). This series includes ten soil profiles across southern and central Sweden, where sampling was performed between 1991 and 2004 outside the treatments tested in these experiments and to a depth of 100 cm. While soil texture and total porosity were measured according to soil horizons, water retention data was given in 10 cm depth intervals. Thus, to make the data consistent, we generated weighted averages of θ for each horizon from the water retention measurements. The methods applied to determine soil water retention, soil texture and total porosity were identical to the other Swedish data.

2.2.1.2. Norway. The Norwegian soil physical dataset considered for this study was collected during the "SOILSPACE" project (Norges Forskningsrådet pr. no. #240663) between 2015 and 2019 (e.g., Koestel et al., 2018). It includes 261 soil samples taken at various depths from 50 arable sites including grassland. The particle-size distribution of the fine

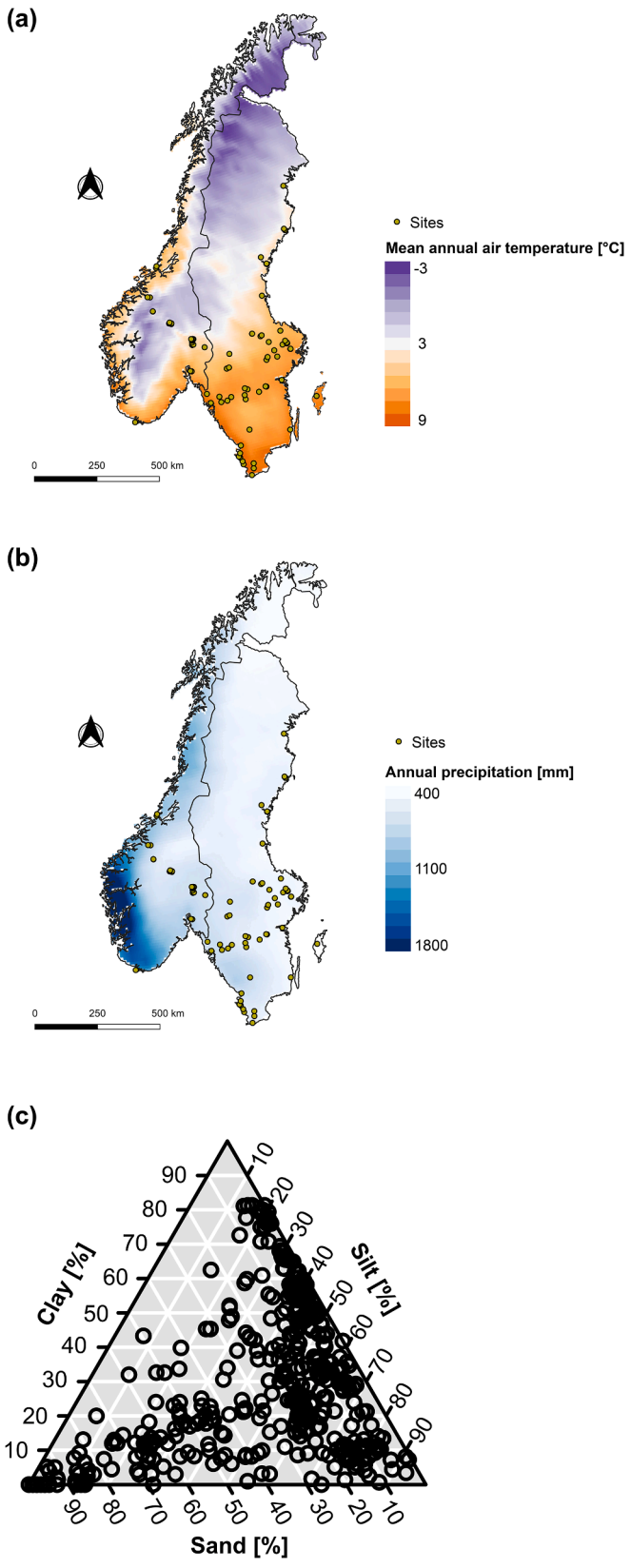


Fig. 1. Geographical (a, b) and textural distribution (c) of the final sites (n = 89) and samples (n = 431). The geographical distribution is shown in terms of mean annual air temperatures (a) and annual precipitation (b). The climate data was retrieved from the MARS Meteorological Database (Toreti, 2014) and shows mean values between 1979 and 2022.

earth fraction (≤2 mm) was measured from disturbed soil samples using wet sieving for particle diameters >63 μm and using the integral suspension pressure method (PARIO, Durner et al., 2017) to derive the <63 μm portion of the distribution. Sample preparation took place according to Norwegian standards (Børresen and Krogstad, 2015). After curve-fitting to the quasi-continuous particle-size distribution data within the 0.01–2000 μm particle diameter range using the PARIO software, the same six particle-size classes as for Sweden were derived. Soil water retention was determined on undisturbed soil cores of either 200 or 250 cm^3 in volume (height = 6 cm): (h, θ) pairs were measured by sand boxes for $h \geq -100$ cm, by the simplified evaporation method using the ku-pF device (UGT GmbH, Germany) for ca. -50 cm $\geq h \geq -900$ cm, by pressure plate at $h = -1000$ and -3000 cm, and by the dew point method (Campbell et al., 2007) using the WP4C device (METER Group Inc. USA) for ca. $h \leq -3500$ cm. Undisturbed sub-samples 10 cm^3 in volume (height = 1 cm) were used for pressure plate measurements, while dew point measurements were made on disturbed sub-samples taken from the original cores. The full range of measurements was not available for all samples in the database. Total porosity was equated with the measured θ_s.

2.2.2. Data manipulation, inclusion criteria and final dataset

The datasets on soil physical and hydraulic properties described above differ with respect to laboratory procedures and the number and range of measured data points both for water retention and soil texture. This required some manipulation to ensure the data was consistent and comparable between the different sources. We set the measured total porosity equal to θ_s if not measured specifically. Since θ_{pwp} was not measured for some of the soil samples in the datasets, we estimated θ_{pwp} from clay content (w_{clay}) for these samples using a pedotransfer function developed from our own data. Specifically, a second degree polynomial function was fitted to (w_{clay}, θ_{pwp}) pairs including all soil samples for which w_{clay} and θ_{pwp} were reported ($n = 493$; $R^2 = 0.89$):

$$\theta_{pwp} = 0.014 + 0.572 w_{clay} - 0.202 w_{clay}^2 \quad (6)$$

Subsequently, we defined several inclusion criteria with respect to the water retention data with the aims to (i) avoid systematic errors between the different measurement techniques, (ii) set the range of measured points, (iii) ensure the points are balanced across this range, and (iv) make the different datasets comparable. We only used data in the $0 > h \geq -15{,}000$ cm range measured by either sand box, pressure plate or the dew point technique; we required each sample to have at least one (h, θ) pair at $h < -3000$ cm; we required (h, θ) pairs to have $\theta \leq \theta_s$ and have at least six (h, θ) pairs remaining to facilitate the fitting of the six-parameter model. In addition, samples were excluded if they were taken from >1 m soil depth or if the water retention data were clearly tri-modal or showed obvious inconsistencies such as non-monotonic behaviour. After applying these criteria, we obtained a final dataset that comprises 431 soil samples from 89 sites across Sweden and Norway. Fig. 1a and b show the locations of these sites superimposed on climate maps, while the range of soil textures covered is shown in Fig. 1c.

2.3. Covariates

The covariates used to analyse the variation in KL divergence across our study area were selected based on their availability for all sites and their known or expected relevance for soil structure evolution in agricultural soils. The selection of covariates was also done with the aim to minimize multi-collinearity as far as possible, which is important when investigating partial dependencies (see Subsection 2.4.2). Nevertheless, some degree of multi-collinearity cannot be prevented. For example, naturally enough, there is a strong negative relationship between soil organic carbon content and soil depth in our dataset (Spearman's $r = -0.74$, $p < 0.001$). We included in total nine covariates, both continuous and categorical, in our analysis of the variation in KL divergence across the study area. These covariates can be subdivided into four categories:

i. Soil properties: clay content, silt content, soil organic carbon content (SOC), soil depth.
ii. Climate: mean annual air temperature (T_m), annual precipitation (P_a).
iii. Time: sampling year, sampling season.
iv. Geology: Parent material.

Details on the nine covariates are provided in Table 1 and their frequency distributions are shown in Fig. S1. Note that we included silt instead of sand content because sand content showed a higher correlation with clay content in our final dataset. Although information about land cover was available in the databases, we did not include it as a covariate in the analysis. This is because only the land cover at the time of sampling was recorded, while information on long-term crop rotations and tillage practices was lacking. Preliminary model testing confirmed the low explanatory power of the available information on land cover at the time of sampling.

Clay and silt contents were obtained from the soil texture data, where we used the Swedish and Norwegian thresholds for the silt fraction (>2 μm and <60 μm particle diameter). Furthermore, data on sampling year, sampling season and soil depth were reported during sample collection for all datasets and were thus readily available. The same was true for parent material, except for some more recent Swedish data. These gaps were filled using a quaternary geology map from the Geological Survey of Sweden. Soil organic carbon content was measured directly by combustion for all Swedish and Norwegian datasets except for the EU-HYDI data. Here, SOC contents had to be derived from loss on ignition (LOI) measurements. We used an empirical function specifically developed for the Swedish data to convert LOI measurements to SOM contents, and a factor of 0.5 to convert SOM to SOC contents (Ljung, 1987; Pribyl, 2010). This gave the following equation:

$$SOC = \frac{LOI - K}{2} \quad (8)$$

where LOI [g/g] refers to the weight loss on ignition and K is a factor depending on clay content (w_{clay}):

$$\begin{aligned} K &= 0.1\ w_{clay}, & w_{clay} &\leqslant 20\% \\ K &= 1.06 + 0.047\ w_{clay}, & w_{clay} &> 20\% \end{aligned}$$

Finally, data on mean annual air temperature (T_m) and annual precipitation (P_a) for Sweden and Norway were obtained from the European Joint Research Centre (JRC) MARS Meteorological Database (Toreti, 2014). This database includes daily observations from European weather stations between 1979 and 2022, which are interpolated on a 25x25 km grid. The grid points closest to each of the 89 sites were determined with the Haversine distance formula.

2.4. Random forest (RF) analysis

A random forest (RF) comprises a large number of decision trees, where each individual tree considers a bootstrapped subset of the total number of observations. Decision trees are usually grown by sequentially searching the variable space of each covariate for the best split of the subset, that is, the split that minimizes the mean squared error. Random forests are special in this context as they only take a random subset of these covariates into account while growing the decision tree. The rationale behind only considering subsets of observations and covariates is that the variance of the random forest model is considerably reduced and its prediction accuracy is increased (James et al., 2013).

We performed the RF analyses for three scenarios: (i) all samples, (ii) only topsoil samples, and (iii) only subsoil samples. Samples with soil

Table 1
Covariates included in this study.

Category	Variable name	Variable type	Range/categories	Unit
Soil properties	Clay content	Continuous	[0, 81.3]	weight-%
	Silt content	Continuous	[0, 93.3]	weight-%
	Soil depth [a]	Continuous	[0, 93]	cm
	Soil organic carbon content	Continuous	[0, 8.7]	weight-%
Climate	Mean annual air temperature (T_m)	Continuous	[2.3, 8.8]	°C
	Annual precipitation (P_a)	Continuous	[430, 1215]	mm
Time	Sampling year	Continuous	[1954, 2017]	year
	Sampling season [b]	Categorical	*spring*, *summer*, *autumn*, *winter*	–
Geology	Parent material	Categorical	*aeolian sediments*, *fluvial sediments*, *glacial sediments*, *gyttja*, *lacustrine sediments*, *marine sediments*, *organic material*, *postglacial sediments*, *shore*, *till*	–

[a] Soil depth was not considered for topsoil and subsoil analyses.
[b] Spring: April-May, Summer: June-August, Autumn: September-November, Winter: December-March.

depth ≤30 cm were regarded as topsoil (n = 174) and the others as subsoil (n = 257). The same approach was applied to each of the three scenarios, except that the variable 'soil depth' was removed as a covariate for the scenarios where only topsoil or only subsoil samples were considered. For each scenario, the RF analysis was repeated for 100 random subsets, each comprising 90 % of the observations. From these repetitions, we could test the robustness of the RF analysis by identifying whether its performance depended on the inclusion of a subgroup of samples. The RF was built from 500 decision trees and three covariates per tree were considered, which is the standard of one third of the covariates commonly applied for regression problems (James et al., 2013). We used so-called 'out-of-bag' (OOB) error estimates as a measure of the performance of the RF models. This entails using decision trees derived from bootstrapped subsets of observations to make predictions for the remaining observations. Predictions are made for each of the observations from each tree in the RF model. These predictions are then averaged and compared with the true observations resulting in the OOB error estimates (James et al., 2013). Subsequently, the OOB error estimates were used to calculate the coefficient-of-determination (R^2). We estimated the relative importance of each covariate for the KL divergence by calculating the increase in total mean squared error on removing them from the model (e.g., Gupta et al., 2022b). We also used partial dependence plots to illustrate the non-linear and potentially non-monotonic effect of each covariate on the KL divergence (e.g., Jorda et al., 2015). This is done by fixing a value for the variable of interest and calculating the average model output over the whole range of the other covariates (James et al., 2013). All operations related to the RF analysis were performed using the 'randomForest' package in R (Liaw and Wiener, 2002).

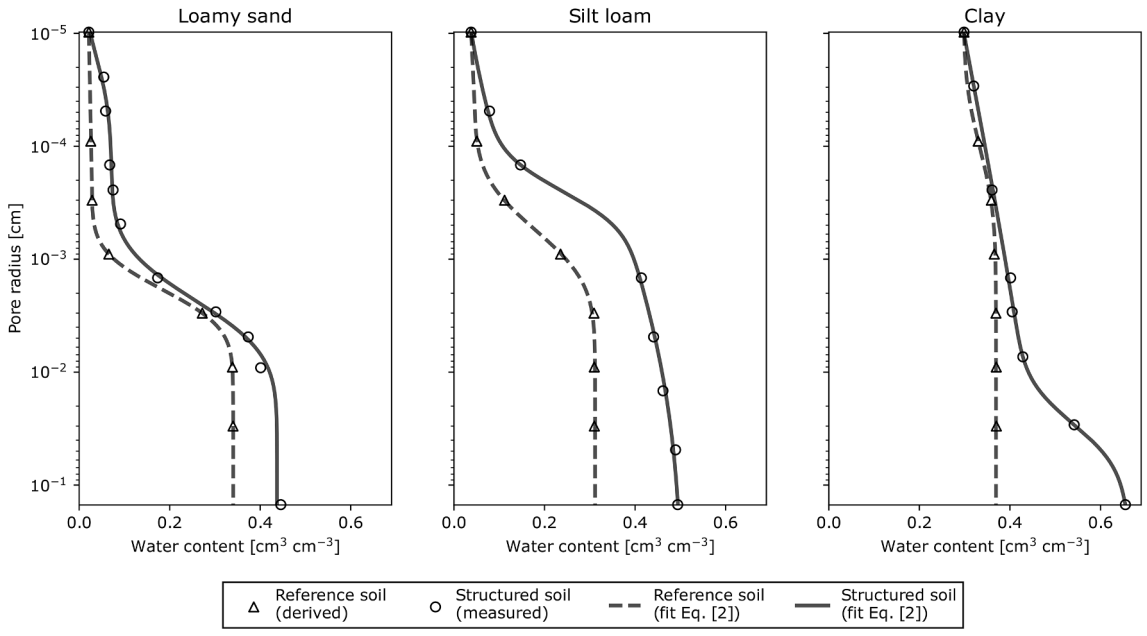


Fig. 2. Example fits of Eq. [2] to derived (reference soil) and measured (structured soil) water retention data of a coarse- (loamy sand), medium- (silt loam) and fine-textured (clay) soil. Soil texture classes according to Soil Science Division Staff (2017).

3. Results and discussion

3.1. Model fitting

Fig. 2 shows three example SWRCs fitted to (r_v, θ) pairs of the structured and reference soils. The bimodal Kosugi (1996) model (Eq. (2)) gave excellent fits for the soil samples in the different soil texture classes. This is evident from the mean root-mean-square errors (RMSE), which were 0.001±0.005 cm^3 cm^{-3} and 0.005±0.005 cm^3 cm^{-3} for the reference and structured soils respectively.

3.2. Performance of random forests

The performance of the RF models, that is, their ability to explain the variation in KL divergence using the nine covariates as expressed by the coefficient of determination (R^2), is shown in Fig. 3. Each boxplot representing one of the three scenarios shows the R^2 values for the 100 model runs. Most of the RF models developed on all samples and subsoil samples explained more than 50 % of the variance in KL divergence. This is encouraging as there are several additional factors that are not accounted for in this study that affect soil structure either directly or indirectly, for example soil management, clay mineralogy, reactive (hydro-)oxides, soil biota and the presence and type of exchangeable cations (Arthur et al., 2013; Regelink et al., 2015; Wu et al., 2023).

In contrast, none of the models developed on topsoil samples could explain more than 50 % of the variance (Fig. 3). Fig. 3 also shows that there is a larger variation in R^2 values for topsoil samples compared with the other two scenarios. This indicates that the performance of each of the 100 model runs with topsoil samples depended strongly on the samples included when the RF model was built. In other words, for the topsoil samples there existed several subsets for which predictions were particularly poor. While this could be partly related to the relatively small sample size ($n = 174$), we also see this as a sign of the complexity inherent in the evolution of soil structure in agricultural topsoils. Furthermore, the poorer performance of models when only topsoil samples were included implies that the KL divergence in topsoils is to a significant extent controlled by covariates that were not included in the RF analysis. In particular, soil tillage and subsequent consolidation as well as shrink-swell and freeze-thaw processes in response to short-term weather conditions can have large effects on the structure of agricultural topsoil (Assouline, 2006; Bodner et al., 2013b; Ghezzehei and Or, 2003; Or et al., 2021; Sandin et al., 2017; Unger, 1991). Furthermore, topsoils show the highest density of roots and the largest activity of soil biota, especially where root-restricting layers are present below plough depth (Bengough et al., 2011). Effects of root growth and soil biota activity on the VSD are species dependent and the result of several feedback mechanisms (Bodner et al., 2014; Leuther et al., 2023; Lucas et al., 2022; Lucas et al., 2019; Meurer et al., 2020a), which most likely contributes to the variance which cannot be explained by our RF models. In contrast, subsoils are less exposed to these processes and it seems that their structure can be better predicted with the covariates included in the analysis.

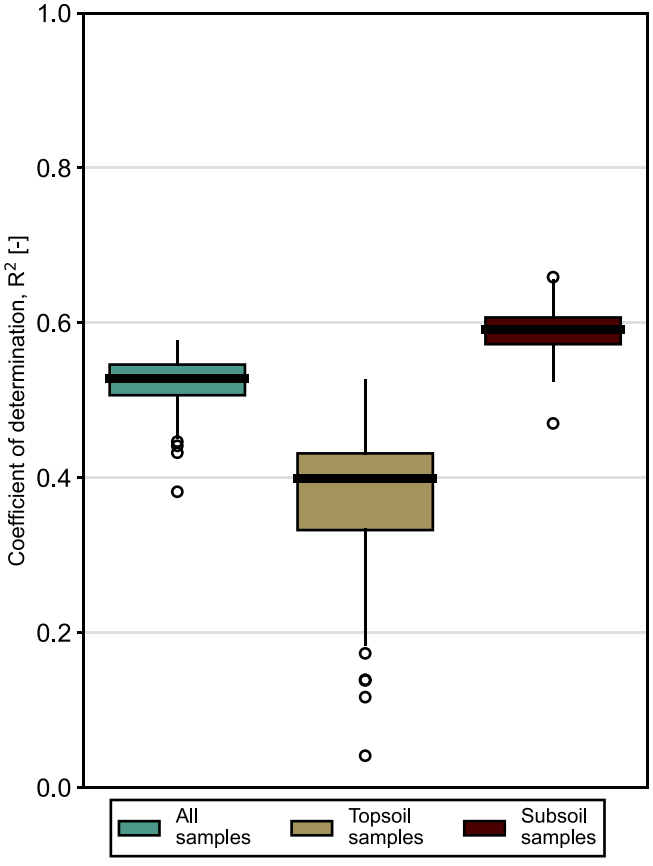


Fig. 3. Performance of random forest models for the three scenarios (all samples, topsoil samples, subsoil samples) based on out-of-bag error estimates.

As a result of the poorer performance of models when only topsoil samples were included, we focus the following discussion on the scenarios where all samples and only subsoil samples were included in the analysis.

3.3. Relative importance of covariates and effects on the KL divergence

For the scenarios where all samples and only subsoil samples were included, Fig. 4 shows the relative importance of each variable for explaining the variance in KL divergence, quantified as the increase in mean squared error (MSE) when this variable was removed from the 100 RF models. Note that the standard errors in Fig. 4 are hardly visible due to the small variability between the individual RF models, highlighting the robustness of the models. Figs. 5 and 6 show the partial dependence plots for all samples and subsoil samples respectively. For completeness, the relative importance and partial dependence plots for the topsoil samples are presented in the supplementary material (Fig. S2 and Fig. S3 respectively). However, care should be taken with the interpretation of these results for the reasons mentioned above.

3.3.1. Soil properties and parent material

Parent material was the most important factor both for all samples and only subsoil samples (Fig. 4). The partial dependence plots suggest much larger KL divergences for soils developed on lake-derived sediments and, in particular, “gyttja” as compared with the other parent materials (Figs. 5 and 6). “Gyttja” is a Swedish term that refers to quaternary lacustrine sediments with high clay and SOC contents. Soils that develop on these sediments are characterised by a large porosity, usually more than 60 %, and well-developed fractures, especially in the subsoil arising from irreversible shrinkage following drainage (Berglund, 1996; Berglund & Berglund, 2010). These characteristics result in large KL divergences throughout the soil profile. The differences in KL divergence among the other parent materials are comparatively small, which suggests that the strong structural development in gyttja soils was the main reason why parent material was ranked as the most important variable. To test this, we re-ran the analysis after removing the gyttja soils from our dataset. We found that the order of relative importance was unaffected, while the importance of parent material for explaining the variation of KL divergence in the subsoil even seemed to increase relative to the other covariates (Fig. 4). The reasons for this rather surprising result are not clear to us.

Clay content was the most important soil property, ranking overall second if all samples were included in the RF analysis and fifth when only subsoil samples were included (Fig. 4). The trends in the partial dependence plots were similar in both scenarios, showing a decrease in

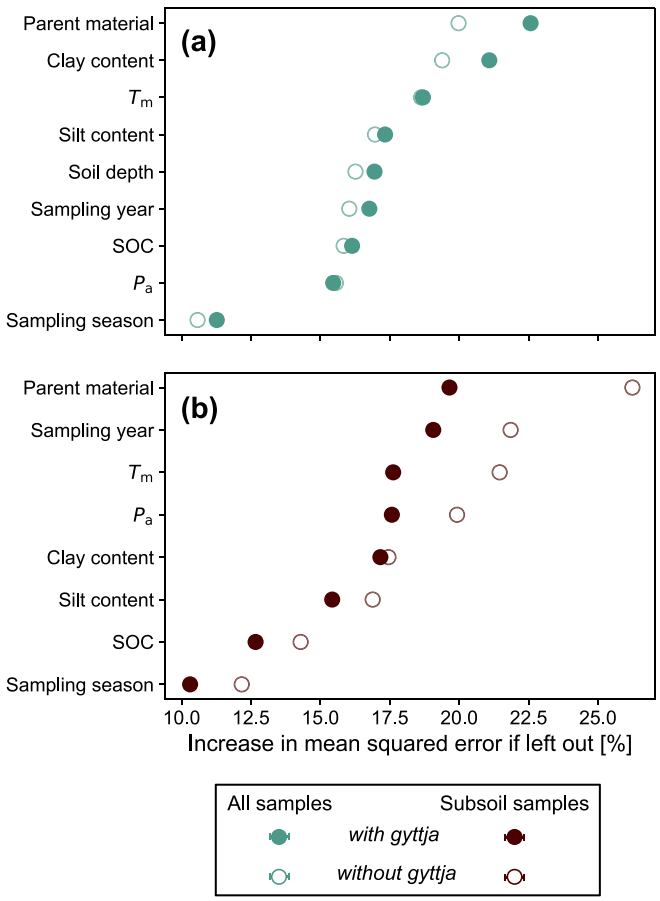


Fig. 4. Relative importance of covariates for explaining the variation in KL divergence for the scenarios including all samples (a) and only subsoil samples (b). The relative importance is expressed as the increase in mean squared error if a specific covariate is left out of the models. Filled circles represent the case where all samples were included in the random forest analysis, while empty circles represent the case without "gyttja" soils. Error bars (hardly visible) indicate standard errors.

KL divergence with increasing clay contents until a value of around 60 %. This was followed by a sharp increase in KL divergence and relatively constant values at larger clay contents. The sharp increase in KL divergence at around 60 % clay content can be explained by the gyttja soils, which have the highest clay contents in our dataset. This is illustrated by the dashed lines in Figs. 5 and 6 for the analysis carried out without the gyttja samples, which lack the sharp increase in KL divergence. The gradual decrease in KL divergence with increasing clay contents is a little surprising at first sight, given that swelling and shrinking upon wetting and drying in soils with relatively high clay contents is known to enhance soil structure development (Bodner et al., 2013a; Diel et al., 2019). However, with the exception of the soils in the south-west of Sweden, the mineralogy of clay soils in Scandinavia is dominated by minerals such as illite, which are less prone to swelling and shrinking. We suggest that the decrease in KL divergence with increasing clay content may be the result of an increased susceptibility to soil compaction, a well-known consequence of intensive agriculture in the temperate-boreal zones with short growing seasons and predominantly wet soil conditions (Batey, 2009). For example, Gebhardt et al. (2009) noted a much higher susceptibility to harmful compaction of soils falling into the clay loam and heavy clay texture class as compared with more coarse-textured soils. Similarly, Smith et al. (1997) found a bell-shaped relationship between susceptibility towards uni-axial soil compaction and clay-plus-silt contents with a peak at around 70 %. In comparison to clay contents, silt contents were less important for explaining the variance in KL divergence (Fig. 4). Similar to clay content, a slight decreasing trend in the KL divergence with increasing silt content until a value of around 60 % can be noted (Figs. 5 and 6).

Soil depth was the fifth most important covariate for the scenario when all soil samples were included in the RF analysis (Fig. 4). The partial dependence plot showed a sharp decrease in KL divergence with increasing soil depth, reaching a minimum at around 30 cm (Fig. 5). We interpret this minimum as a sign of soil compaction at and just below plough depth (Batey, 2009; Håkansson and Medvedev, 1995). In contrast, the uppermost few centimetres in arable soils are commonly subject to regular and intensive processes that loosen the soil such as ploughing, seedbed preparation, root growth and the activity of soil biota (Meurer et al., 2020a; Or et al., 2021). The partial dependence plot suggests that the KL divergence remains relatively constant with increasing soil depth below plough depth, with the KL divergence even showing a slight tendency to increase (Fig. 5). The reason for this is not clear, but it may be a consequence of mechanical stresses and compaction from farm vehicle traffic, which are greater in the upper subsoil (Hadas, 1994). Fig. 5 also shows that the subsoils of gyttja have a strongly developed structure that buffers the decrease in KL divergence with increasing soil depth.

Soil organic carbon content had the smallest effect of all soil properties on the KL divergence, ranking sixth and seventh for all samples and subsoil samples respectively (Fig. 4). This finding is in apparent contrast to other studies that reported significant effects of SOC on various soil structural properties, including the KL divergence (e.g., Anderson et al., 1990; Klöffel et al., 2022; Naveed et al., 2014). We propose several reasons for this. First, SOC contents for 71 % of the samples included here were estimated from loss on ignition (LOI) data using an empirical function, thereby introducing errors that may have distorted the relationship between SOC and KL divergence (Chatterjee et al., 2009; Hoogsteen et al., 2015). This is supported by the fact that LOI samples showed a weaker correlation (Spearman's $r = 0.31$, $p < 0.001$) between SOC and KL divergence compared with samples where SOC contents were measured directly (Spearman's $r = 0.46$, $p < 0.001$). Second, Klöffel et al. (2022) showed that the KL divergence is most sensitive to structure formation in the larger pore region, because this is where the VSD of the structured and reference soil have the potential to diverge most. In agricultural soils, SOC is often dominated by the mineral-associated organic matter fraction (Fukumasu et al., 2022; Poeplau et al., 2018) and thus mostly drives soil structure formation at smaller scales (e.g., through aggregation) with little direct effect on larger structural pores (Jarvis et al., 2017). Third, subsoil samples of low SOC content represent a considerable proportion of the samples in our dataset (see Figure S1), for which the positive effects of SOC on soil structure may not be apparent. Finally, regular mechanical disturbance, for example through soil tillage, may override the effect of SOC by periodically or permanently destroying larger structural pores. This explanation is supported by the partial dependence plots, which revealed an interesting effect of SOC on the KL divergence, showing a steep increase in KL divergence with SOC contents less than ca. 2 % and reaching a plateau at larger SOC contents with a maximum at around 3 % (Figs. 5 and 6). While the increasing trend for lower SOC contents is in line with the expected development of soil structure, expressed as wider VSDs and higher porosities with increasing SOC contents (Fukumasu et al., 2022; Jarvis et al., 2017; Klöffel et al., 2022; Meurer et al., 2020b; Zhang et al., 2021), the plateau at 3 % contradicts the increase in total porosity for SOC contents beyond this value, which is observed for other land uses (Robinson et al., 2022). Thus, the partial dependence plots derived from our dataset suggest that a critical SOC content (Loveland and Webb, 2003) may exist in tilled and trafficked agricultural soils below which soil structure starts to deteriorate and that this threshold is around 3 %. We note that our dataset contains only few samples with SOC contents above this threshold ($n = 23$; Fig. S1) so that any increases in KL divergence with SOC contents larger than 3 % may not have be revealed. Furthermore, we previously noted the significant negative

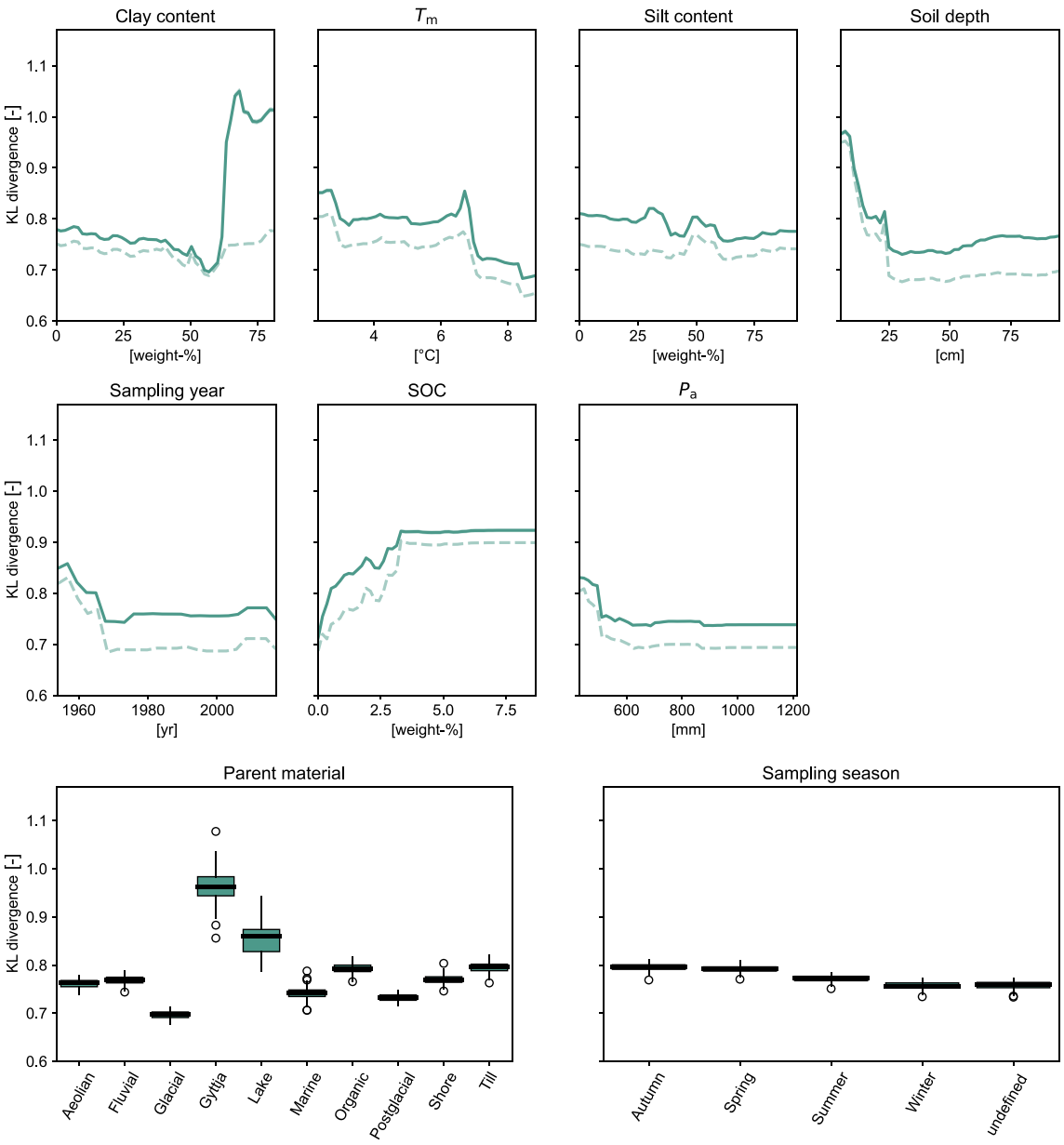


Fig. 5. Partial dependence plots for the scenario including all soil samples. Both continuous and categorical covariates, which were used to explain the variation in KL divergence, are shown. For the continuous covariates, the solid lines represent the case including all soil samples and the dashed lines represent the case without "gyttja" soils.

correlation between SOC content and soil depth in our dataset. However, the shape of the partial dependence plots is the same for subsoil and topsoil (Fig. 6 and Fig. S3), which gives us some confidence that the relationship between SOC content and KL divergence revealed here is largely the effect of SOC.

3.3.2. Climate

The annual average air temperature, T_m, was the third most important variable both when all samples and only subsoil samples were included in the RF analysis (Fig. 4). The partial dependence plots show that the KL divergence tends to be larger in colder climates, which is more prominent in the case where all samples were included (Figs. 5 and 6). This difference between scenarios is due to the inclusion of topsoil samples (Fig. S3). The smaller KL divergences in the warmer regions of Sweden and Norway may be a consequence of more intensive agriculture, which is associated with the use of heavier machinery and thus an increased risk of soil structure deterioration by compaction (Björklund et al., 1999; Keller et al., 2019). Another likely reason for the larger KL divergences in colder regions could be the effects of freezing and thawing on soil structure. We roughly approximated the number of freeze-thaw cycles in the soil from our air temperature data, where one cycle was defined by a drop below −0.5 °C and subsequent increase above 0.5 °C. We found a strong negative correlation between the estimated number of freeze-thaw cycles and T_m (Spearman's $r = -0.80$; $p < 0.001$). Freeze-thaw cycles can have a 'structure-recovering' effect in soils whose structure has previously been degraded, even in the subsoil (Bryk et al., 2017; Gregory et al., 2009; Ma et al., 2019). In southern Sweden, where the largest T_m are found, intense freeze-thaw cycles are rare and freezing does not penetrate far into the soil rendering structure-recovering effects ineffective. This supports the results of Hirmas et al. (2018), who found a positive correlation between the frequency of freezing and effective porosity (i.e., the difference between total porosity and water content at field capacity) across different climate regions of the US. Finally, it can also be noticed that the peak in KL

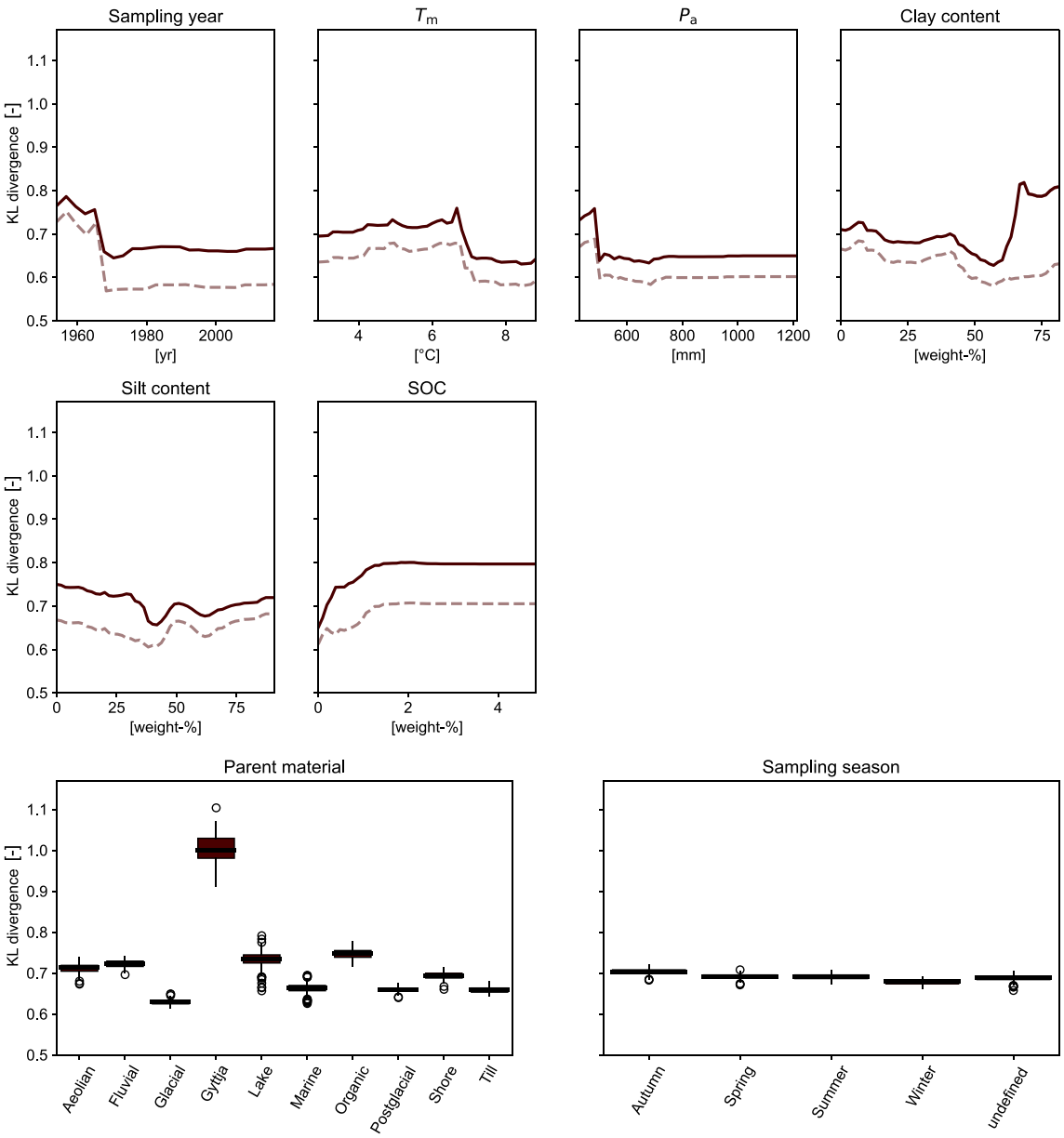


Fig. 6. Partial dependence plots for the scenario including only subsoil samples. Both continuous and categorical covariates, which were used to explain the variation in KL divergence, are shown. For the continuous covariates, the solid lines represent the case including all soil samples and the dashed lines represent the case without "gyttja" soils.

divergence at T_m values of around 6.5 °C in Figs. 5 and 6 is caused by the gyttja soils.

The second climate variable, the annual average precipitation, P_a, was less important than T_m in both scenarios, ranking fourth for the subsoil samples, while it was the third least important variable when all samples were included (Fig. 4). The partial dependence plots in Figs. 5 and 6 show similar trends with the largest KL divergences in dry climates, decreasing to minimum values at a P_a of ca. 500 to 600 mm, with relatively constant KL divergences with further increases in annual precipitation. This finding is in line with previous studies, which showed directly or indirectly that soils in drier climates tend to have a stronger developed structure and more heterogeneous VSDs (Caplan et al., 2019; Hirmas et al., 2018; Jarvis et al., 2013; Jorda et al., 2015; Wu et al., 2023). The reasons for this, however, are still unclear. It is possible that the effect is indirect and due to an increased risk of soil compaction in wetter regions (Jarvis et al., 2013). Shrinking and swelling may also be a relevant process in this context, as shrinkage would be more pronounced in drier climates (Bodner et al., 2013a; Caplan et al., 2019).

3.3.3. *Time*

Sampling year was the second most important covariate if only subsoil samples were included in the RF analysis, while it was less important for the scenario with all samples (Fig. 4). The partial dependence plots showed a decline in KL divergence between the 1950′s and 1970′s (Figs. 5 and 6) and was relatively constant thereafter. These results suggest a gradual deterioration of soil structure in the recent past, which seems to have been especially prevalent in the subsoil. This deterioration most probably resulted from the considerable increase in production intensity between the 1950′s and 1990′s, in particular in southern Sweden (Björklund et al., 1999). Key aspects of this intensification were an overall increase in the weight of agricultural machinery (Keller et al., 2019; Parvin et al., 2022) and more repeated trafficking, both of which would tend to enhance vertical stresses transmitted to the soil and aggravate the severity of soil compaction (Hadas, 1994;

Håkansson and Medvedev, 1995; Keller and Or, 2022). It can be noted that the larger KL divergences found for samples taken between the 1950′s and 1970′s do not appear to be an artifact due to collinearity. For example, they do not coincide with the sampling of soils that are naturally more structured such as the gyttja soils. In fact, the opposite is the case, as most of the samples of gyttja soils were taken after the 1970′s, which therefore buffered the decline in KL divergence with time (Figs. 5 and 6). Similarly, the warmer sites, which showed lower KL divergences, were not preferentially sampled in more recent years, as shown by a lack of correlation between sampling year and T_m (Spearman's $r = -0.03$, $p = 0.50$).

In contrast to sampling year, sampling *season* was the least important variable in all scenarios (Fig. 4), and the partial dependence plots showed little variation among the different seasons (Figs. 5 and 6). This is in contrast to many studies that stress the importance of within-season variability of soil pore structure (e.g., Alletto et al., 2015; Messing and Jarvis, 1993; Schwen et al., 2011). However, this within-season variability has only been investigated in topsoil, where it mainly results from the consolidation of soil following mechanical loosening by tillage (Ghezzehei and Or, 2003). The dominance of subsoil samples in our dataset might therefore have played a role. However, sampling season ranked lowest even for the scenario where only topsoil samples were included (Fig. S2). It seems more likely that the low relevance of sampling season may be because many of the topsoil samples in our database were sampled in summer, presumably after the soil had already undergone some degree of consolidation, while relatively few were taken soon after tillage in spring (Fig. S1).

4. Summary and conclusions

We conclude that an index of soil structure based on relative entropy, the Kullback-Leibler (KL) divergence, gave useful insights into the drivers of soil structure evolution in agricultural soils of a region in the temperate-boreal zone. The random forest (RF) analysis explained on average more than 50 % of the variation in KL divergence for a large dataset of samples from Sweden and Norway. However, the performance was worse in the topsoil, highlighting the complex dynamics of soil structural evolution in agricultural topsoils and the need to include other covariates than those chosen here, such as information on soil tillage and recent weather conditions. Our analysis revealed the leading role of properties of the soil mineral phase for driving the evolution of soil structure, in particular properties resulting from the nature of the parent material and clay content. This raises the question to what degree soil structure is "pre-determined" by inherent soil properties, which are difficult to manage. Furthermore, we found relatively large effects of mean annual air temperature and average annual precipitation on the KL divergence. Although we were unable to answer the question whether these effects are direct, for example through swelling and shrinking and/or freezing and thawing processes, or indirectly expressed through interactions with soil management (e.g., increased risks of traffic compaction in wetter regions), they might be relevant in the context of climate change. In particular, as climate change in the temperate-boreal zone is expected to result in wetter soil conditions and the northward expansion of intensive agriculture due to warmer temperatures, the results suggest an increased risk of soil structure deterioration in the future. Soil organic carbon content turned out to be the least relevant soil property for explaining the variation in KL divergence in our study. However, our results suggest a threshold SOC content at around 3 % below which soil structure deteriorates. Further investigations are needed to test whether the de-coupling between SOC increase and soil structure formation may be the result of regular soil tillage and thus a typical feature of agricultural soils. Lastly, our analysis confirmed signs of soil structure deterioration, in particular in the upper subsoil. Through the inclusion of sampling year as a covariate, we could reveal that this deterioration worsened between the 1950's and 1970's, which is most probably related to the increase in machinery weight and agricultural intensification in the recent past. This finding implies the need for soil structure monitoring to unveil such trends at an early stage that allows for timely mitigation measures. We note that the dataset used here suffers to some degree from sampling bias as a result of past research interests in Sweden. This calls for a uniform spatial representation of soil physical data in the study area, which may be achieved through targeted sampling campaigns. Furthermore, we note that our results are likely to be strictly only valid for agricultural soils of the temperate and boreal zone. A similar study in other regions would therefore be a logical next step.

CRediT authorship contribution statement

Tobias Klöffel: Formal analysis, Methodology, Software, Writing – original draft, Writing – review & editing, Data curation. **Jennie Barron:** Funding acquisition, Supervision, Writing – review & editing. **Attila Nemes:** Data curation, Writing – review & editing. **Daniel Giménez:** Data curation, Methodology, Writing – review & editing. **Nicholas Jarvis:** Methodology, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We are grateful to Örjan and Kerstin Berglund (Department of Soil and Environment, SLU) for providing additional data for the Swedish soil physical dataset. Thanks also to Claudia von Brömssen (Department of Energy and Technology, SLU) for statistical support in the random forest analysis. Finally, we thank the Faculty of Natural Resources and Agricultural Sciences (SLU) for financially supporting this work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.geoderma.2024.116772.

References

Alaoui, A., Lipiec, J., Gerke, H.H., 2011. A review of the changes in the soil pore system due to soil deformation: A hydrodynamic perspective. Soil and Tillage Research 115–116, 1–15. https://doi.org/10.1016/j.still.2011.06.002.

Alletto, L., Pot, V., Giuliano, S., Costes, M., Perdrieux, F., Justes, E., 2015. Temporal variation in soil physical properties improves the water dynamics modeling in a conventionally-tilled soil. Geoderma 243–244, 18–28. https://doi.org/10.1016/j.geoderma.2014.12.006.

Anderson, S.H., Gantzer, C.J., Brown, J.R., 1990. Soil physical properties after 100 years of continuous cultivation. Journal of Soil and Water Conservation 45, 117–121.

Arthur, E., Moldrup, P., Schjønning, P., De Jonge, L.W., 2013. Water Retention, Gas Transport, and Pore Network Complexity during Short-Term Regeneration of Soil Structure. Soil Science Society of America Journal 77, 1965–1976. https://doi.org/10.2136/sssaj2013.07.0270.

Arya, L.M., Heitman, J.L., 2015. A Non-Empirical Method for Computing Pore Radii and Soil Water Characteristics from Particle-Size Distribution. Soil Science Society of America Journal 79, 1537–1544. https://doi.org/10.2136/sssaj2015.04.0145.

Arya, L.M., Leij, F.J., van Genuchten, M.T., Shouse, P.J., 1999. Scaling Parameter to Predict the Soil Water Characteristic from Particle-Size Distribution Data. Soil Science Society of America Journal 63, 510–519. https://doi.org/10.2136/sssaj1999.03615995006300030013x.

Arya, L.M., Paris, J.F., 1981. A Physicoempirical Model to Predict the Soil Moisture Characteristic from Particle-Size Distribution and Bulk Density Data. Soil Science Society of America Journal 45, 1023–1030. https://doi.org/10.2136/sssaj1981.03615995004500060004x.

Assouline, S., 2006. Modeling the Relationship between Soil Bulk Density and the Water Retention Curve. Vadose Zone Journal 5, 554–563. https://doi.org/10.2136/vzj2005.0083.

Bagnall, D.K., Morgan, C.L.S., Cope, M., Bean, G.M., Cappellazzi, S., Greub, K., Liptzin, D., Norris, C.L., Rieke, E., Tracy, P., Aberle, E., Ashworth, A., Bañuelos Tavarez, O., Bary, A., Baumhardt, R.L., Borbón Gracia, A., Brainard, D., Brennan, J., Briones Reyes, D., Bruhjell, D., Carlyle, C., Crawford, J., Creech, C., Culman, S., Deen, W., Dell, C., Derner, J., Ducey, T., Duiker, S.W., Dyck, M., Ellert, B., Entz, M., Espinosa Solorio, A., Fonte, S.J., Fonteyne, S., Fortuna, A., Foster, J., Fultz, L., Gamble, A.V., Geddes, C., Griffin-LaHue, D., Grove, J., Hamilton, S.K., Hao, X., Hayden, Z.D., Howe, J., Ippolito, J., Johnson, G., Kautz, M., Kitchen, N., Kumar, S., Kurtz, K., Larney, F., Lewis, K., Liebman, M., Lopez Ramirez, A., Machado, S., Maharjan, B., Martinez Gamiño, M.A., May, W., McClaran, M., McDaniel, M., Millar, N., Mitchell, J.P., Moore, P.A., Moore, A., Mora Gutiérrez, M., Nelson, K.A., Omondi, E., Osborne, S., Alcalá, L.O., Owens, P., Pena-Yewtukhiw, E.M., Poffenbarger, H., Ponce Lira, B., Reeve, J., Reinbott, T., Reiter, M., Ritchey, E., Roozeboom, K.L., Rui, I., Sadeghpour, A., Sainju, U.M., Sanford, G., Schillinger, W., Schindelbeck, R.R., Schipanski, M., Schlegel, A., Scow, K., Sherrod, L., Sidhu, S., Solís Moya, E., St. Luce, M., Strock, J., Suyker, A., Sykes, V., Tao, H., Trujillo Campos, A., Van Eerd, L.L., Verhulst, N., Vyn, T.J., Wang, Y., Watts, D., Wright, D., Zhang, T., Honeycutt, C.W., 2022. Carbon-sensitive pedotransfer functions for plant available water. Soil Science Society of America Journal 86, 612–629. https://doi.org/10.1002/saj2.20395.

Batey, T., 2009. Soil compaction and soil management – a review. Soil Use and Management 25, 335–345. https://doi.org/10.1111/j.1475-2743.2009.00236.x.

Bengough, A.G., McKenzie, B.M., Hallett, P.D., Valentine, T.A., 2011. Root elongation, water stress, and mechanical impedance: a review of limiting stresses and beneficial root tip traits. Journal of Experimental Botany 62, 59–68. https://doi.org/10.1093/jxb/erq350.

Berglund, K., 1996. Properties of cultivated gyttja soils. International Peat Journal 6, 5–23.

Berglund, Ö., Berglund, K., 2010. Distribution and cultivation intensity of agricultural peat and gyttja soils in Sweden and estimation of greenhouse gas emissions from cultivated peat soils. Geoderma 154, 173–180. https://doi.org/10.1016/j.geoderma.2008.11.035.

Björklund, J., Limburg, K.E., Rydberg, T., 1999. Impact of production intensity on the ability of the agricultural landscape to generate ecosystem services: an example from Sweden. Ecological Economics 29, 269–291. https://doi.org/10.1016/S0921-8009(99)00014-2.

Bodner, G., Loiskandl, W., Buchan, G., Kaul, H.-P., 2008. Natural and management-induced dynamics of hydraulic conductivity along a cover-cropped field slope. Geoderma 146, 317–325. https://doi.org/10.1016/j.geoderma.2008.06.012.

Bodner, G., Scholl, P., Kaul, H.-P., 2013a. Field quantification of wetting–drying cycles to predict temporal changes of soil pore size distribution. Soil and Tillage Research 133, 1–9. https://doi.org/10.1016/j.still.2013.05.006.

Bodner, G., Scholl, P., Loiskandl, W., Kaul, H.-P., 2013b. Environmental and management influences on temporal variability of near saturated soil hydraulic properties. Geoderma 204–205, 120–129. https://doi.org/10.1016/j.geoderma.2013.04.015.

Bodner, G., Leitner, D., Kaul, H.-P., 2014. Coarse and fine root plants affect pore size distributions differently. Plant Soil 380, 133–151. https://doi.org/10.1007/s11104-014-2079-8.

Børresen, T., Krogstad, T., 2015. Field and laboratory exercises (JORD200). Norwegian University of Life Sciences (NMBU), Norway.

Brutsaert, W., 1966. Probability laws for pore-size distributions. Soil Science 101, 85–92.

Bryk, M., Kołodziej, B., Słowińska-Jurkiewicz, A., Jaroszuk-Sierocińska, M., 2017. Evaluation of soil structure and physical properties influenced by weather conditions during autumn-winter-spring season. Soil and Tillage Research 170, 66–76. https://doi.org/10.1016/j.still.2017.03.004.

Caplan, J.S., Giménez, D., Hirmas, D.R., Brunsell, N.A., Blair, J.M., Knapp, A.K., 2019. Decadal-scale shifts in soil hydraulic properties as induced by altered precipitation. Science. Advances 5, eaau6635. https://doi.org/10.1126/sciadv.aau6635.

Chang, C., Cheng, D., Qiao, X., 2019. Improving estimation of pore size distribution to predict the soil water retention curve from its particle size distribution. Geoderma 340, 206–212. https://doi.org/10.1016/j.geoderma.2019.01.011.

Chatterjee, A., Lal, R., Wielopolski, L., Martin, M.Z., Ebinger, M.H., 2009. Evaluation of Different Soil Carbon Determination Methods. Critical Reviews in Plant Sciences 28, 164–178. https://doi.org/10.1080/07352680902776556.

Dexter, A.R., 2004. Soil physical quality. Geoderma 120, 201–214. https://doi.org/10.1016/j.geoderma.2003.09.004.

Diel, J., Vogel, H.-J., Schlüter, S., 2019. Impact of wetting and drying cycles on soil structure dynamics. Geoderma 345, 63–71. https://doi.org/10.1016/j.geoderma.2019.03.018.

Durner, W., 1994. Hydraulic conductivity estimation for soils with heterogeneous pore structure. Water Resources Research 30, 211–223. https://doi.org/10.1029/93WR02676.

Durner, W., Iden, S.C., Von Unold, G., 2017. The integral suspension pressure method (ISP) for precise particle-size analysis by gravitational sedimentation: ISP METHOD FOR PARTICLE-SIZE ANALYSIS. Water Resources Research 53, 33–48. https://doi.org/10.1002/2016WR019830.

Erktan, A., Or, D., Scheu, S., 2020. The physical structure of soil: Determinant and consequence of trophic interactions. Soil Biology and Biochemistry 148, 107876. https://doi.org/10.1016/j.soilbio.2020.107876.

European Commission, 2013. European HYdropedological Data Inventory (EU-HYDI). (JRC Technical Reports). Institute for Environment and Sustainability, LU.

FAO-UNESCO, 1974. Soil map of the world, 1 : 5 000 000, Volume I, Legend. UNESCO, Paris.

Fernández-Gálvez, J., Pollacco, J.A.P., Lilburne, L., McNeill, S., Carrick, S., Lassabatere, L., Angulo-Jaramillo, R., 2021. Deriving physical and unique bimodal soil Kosugi hydraulic parameters from inverse modelling. Advances in Water Resources 153, 103933. https://doi.org/10.1016/j.advwatres.2021.103933.

Fiès, J.C., Bruand, A., 1998. Particle packing and organization of the textural porosity in clay-silt-sand mixtures. European Journal of Soil Science 49, 557–567. https://doi.org/10.1046/j.1365-2389.1998.4940557.x.

Fredlund, M.D., Fredlund, D.G., Wilson, G.W., 2000. An equation to represent grain-size distribution. Canadian Geotechnical Journal. https://doi.org/10.1139/t00-015.

Fukumasu, J., Jarvis, N., Koestel, J., Kätterer, T., Larsbo, M., 2022. Relations between soil organic carbon content and the pore size distribution for an arable topsoil with large variations in soil properties. European Journal of Soil Science 73. https://doi.org/10.1111/ejss.13212.

Gebhardt, S., Fleige, H., Horn, R., 2009. Effect of compaction on pore functions of soils in a Saalean moraine landscape in North Germany. Journal of Plant Nutrition and Soil Science 172, 688–695. https://doi.org/10.1002/jpln.200800073.

Ghezzehei, T.A., Or, D., 2003. Pore-Space Dynamics in a Soil Aggregate Bed under a Static External Load. Soil Science Society of America Journal 67, 12–19. https://doi.org/10.2136/sssaj2003.1200.

Gregory, A.S., Watts, C.W., Griffiths, B.S., Hallett, P.D., Kuan, H.L., Whitmore, A.P., 2009. The effect of long-term soil management on the physical and biological resilience of a range of arable and grassland soils in England. Geoderma 153, 172–185. https://doi.org/10.1016/j.geoderma.2009.08.002.

Gupta, S., Larson, W., 1979. A Model for Predicting Packing Density of Soils Using Particle-Size Distribution1. Soil Science Society of America Journal - SSSAJ 43. https://doi.org/10.2136/sssaj1979.03615995004300040028x.

Gupta, S., Papritz, A., Lehmann, P., Hengl, T., Bonetti, S., Or, D., 2022a. Global Soil Hydraulic Properties dataset based on legacy site observations and robust parameterization. Scientific Data 9, 444. https://doi.org/10.1038/s41597-022-01481-5.

Gupta, S., Papritz, A., Lehmann, P., Hengl, T., Bonetti, S., Or, D., 2022b. Global Mapping of Soil Water Characteristics Parameters— Fusing Curated Data with Machine Learning and Environmental Covariates. Remote Sensing 14, 1947. https://doi.org/10.3390/rs14081947.

Hadas, A., 1994. Soil compaction caused by high axle loads— review of concepts and experimental data. Soil and Tillage Research 29, 253–276. https://doi.org/10.1016/0167-1987(94)90064-7.

Håkansson, I., Medvedev, V.W., 1995. Protection of soils from mechanical overloading by establishing limits for stresses caused by heavy vehicles. Soil and Tillage Research 35, 85–97. https://doi.org/10.1016/0167-1987(95)00476-9.

Haverkamp, R., Parlange, J.-Y., 1986. Predicting the water-retention curve from particle-size distribution: 1. Sandy soils without organic matter. Soil Science 142, 325–339.

Hirmas, D.R., Giménez, D., Nemes, A., Kerry, R., Brunsell, N.A., Wilson, C.J., 2018. Climate-induced changes in continental-scale soil macroporosity may intensify water cycle. Nature 561, 100–103. https://doi.org/10.1038/s41586-018-0463-x.

Hoogsteen, M.J.J., Lantinga, E.A., Bakker, E.J., Groot, J.C.J., Tittonell, P.A., 2015. Estimating soil organic carbon through loss on ignition: effects of ignition conditions and structural water loss: Refining the loss on ignition method. Eur J Soil Sci 66, 320–328. https://doi.org/10.1111/ejss.12224.

Hwang, S.I., Choi, S.I., 2006. Use of a lognormal distribution model for estimating soil water retention curves from particle-size distribution data. Journal of Hydrology 323, 325–334. https://doi.org/10.1016/j.jhydrol.2005.09.005.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning, Springer Texts in Statistics. Springer New York, New York, NY. https://doi.org/10.1007/978-1-4614-7138-7.

Janzen, H.H., 2015. Beyond carbon sequestration: soil as conduit of solar energy: Soil carbon and energy. European Journal of Soil Science 66, 19–32. https://doi.org/10.1111/ejss.12194.

Jarvis, N., Koestel, J., Messing, I., Moeys, J., Lindahl, A., 2013. Influence of soil, land use and climatic factors on the hydraulic conductivity of soil. Hydrology and Earth System Sciences 17, 5185–5195. https://doi.org/10.5194/hess-17-5185-2013.

Jarvis, N., Forkman, J., Koestel, J., Kätterer, T., Larsbo, M., Taylor, A., 2017. Long-term effects of grass-clover leys on the structure of a silt loam soil in a cold climate. Agriculture, Ecosystems & Environment 247, 319–328. https://doi.org/10.1016/j.agee.2017.06.042.

Jensen, J.L., Schjønning, P., Watts, C.W., Christensen, B.T., Munkholm, L.J., 2019. Soil Water Retention: Uni-Modal Models of Pore-Size Distribution Neglect Impacts of Soil Management. Soil Science Society of America Journal 83, 18–26. https://doi.org/10.2136/sssaj2018.06.0238.

Jorda, H., Bechtold, M., Jarvis, N., Koestel, J., 2015. Using boosted regression trees to explore key factors controlling saturated and near-saturated hydraulic conductivity: Effects of soil properties on hydraulic conductivity. European Journal of Soil Science 66, 744–756. https://doi.org/10.1111/ejss.12249.

Kätterer, T., Andrén, O., Jansson, P.-E., 2006. Pedotransfer functions for estimating plant available water and bulk density in Swedish agricultural soils. Acta Agriculturae Scandinavica, Section B - Soil & Plant Science 56, 263–276. https://doi.org/10.1080/09064710500310170.

Keller, T., Or, D., 2022. Farm vehicles approaching weights of sauropods exceed safe mechanical limits for soil functioning. Proceedings of the National Academy of Sciences 119. https://doi.org/10.1073/pnas.2117699119 e2117699119.

Keller, T., Sandin, M., Colombi, T., Horn, R., Or, D., 2019. Historical increase in agricultural machinery weights enhanced soil stress levels and adversely affected soil functioning. Soil and Tillage Research 194, 104293. https://doi.org/10.1016/j.still.2019.104293.

Kirchmann, H., 1991. Properties and Classification of Soils of the Swedish Long-term Fertility Experiments: I. *Sites at Fors and Kungsängen*. Acta Agriculturae Scandinavica 41, 227–242. https://doi.org/10.1080/00015129109439905.
Kirchmann, H., Eriksson, J., 1993. Properties and Classification of Soils of the Swedish Long-Term Fertility Experiments: II. Sites at Örja and Orup. Acta Agriculturae Scandinavica, Section B - Soil & Plant Science 43, 193–205. https://doi.org/10.1080/09064719309411242.
Kirchmann, H., Snäil, S., Eriksson, J., 1996. Properties and Classification of Soils of the Swedish Long-Term Fertility Experiments: III. Sites at Västraby and S. Ugglarp. Acta Agriculturae Scandinavica, Section B - Soil & Plant Science 46, 86–97. https://doi.org/10.1080/09064719609413120.
Kirchmann, H., Eriksson, J., Snäll, S., 1999. Properties and Classification of Soils of the Swedish Long-Term Fertility Experiments: IV. Sites at Ekebo and Fjärdingslöv. Acta Agriculturae Scandinavica, Section B - Soil & Plant Science 49, 25–38. https://doi.org/10.1080/09064719950135678.
Kirchmann, H., Snäll, S., Eriksson, J., Mattsson, L., 2005. Properties and classification of soils of the Swedish long-term fertility experiments: V. Sites at Vreta Kloster and Högåsa. Acta Agriculturae Scandinavica, Section B - Soil & Plant Science 55, 98–110. https://doi.org/10.1080/09064710510008711.
Klöffel, T., Jarvis, N., Yoon, S.W., Barron, J., Giménez, D., 2022. Relative entropy as an index of soil structure. European Journal of Soil Science 73. https://doi.org/10.1111/ejss.13254.
Koestel, J., Dathe, A., Skaggs, T.H., Klakegg, O., Ahmad, M.A., Babko, M., Giménez, D., Farkas, C., Nemes, A., Jarvis, N., 2018. Estimating the Permeability of Naturally Structured Soil From Percolation Theory and Pore Space Characteristics Imaged by X-Ray. Water Resources Research 54, 9255–9263. https://doi.org/10.1029/2018WR023609.
Kosugi, K., 1996. Lognormal Distribution Model for Unsaturated Soil Hydraulic Properties. Water Resources Research 32, 2697–2703. https://doi.org/10.1029/96WR01776.
Kullback, S., Leibler, R.A., 1951. On Information and Sufficiency. The Annals of Mathematical Statistics 22, 79–86.
Leuther, F., Mikutta, R., Wolff, M., Kaiser, K., Schlüter, S., 2023. Structure turnover times of grassland soils under different moisture regimes. Geoderma 433, 116464. https://doi.org/10.1016/j.geoderma.2023.116464.
Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest}. R News 2, 18–22.
Lilburne, L.R., Hewitt, A.E., Webb, T.W., 2012. Soil and informatics science combine to develop S-map: A new generation soil information system for New Zealand. Geoderma 170, 232–238. https://doi.org/10.1016/j.geoderma.2011.11.012.
Lin, H., Bouma, J., Pachepsky, Y., Western, A., Thompson, J., Van Genuchten, R., Vogel, H.-J., Lilly, A., 2006. Hydropedology: Synergistic integration of pedology and hydrology: OPINION. Water Resources Research 42. https://doi.org/10.1029/2005WR004085.
Ljung, G., 1987. Mekanisk analys – Beskrivning av en rationell metod för jordartsbestämning (Avdelningsmeddelande 87:2). Institutionen för markvetenskap, Avdelningen för lantbrukets hydroteknik, Sveriges lantbruksuniversitet, Uppsala, Sverige.
Loveland, P., Webb, J., 2003. Is there a critical level of organic matter in the agricultural soils of temperate regions: a review. Soil and Tillage Research 70, 1–18. https://doi.org/10.1016/S0167-1987(02)00139-3.
Lucas, M., Schlüter, S., Vogel, H.-J., Vetterlein, D., 2019. Roots compact the surrounding soil depending on the structures they encounter. Scientific Reports 9, 16236. https://doi.org/10.1038/s41598-019-52665-w.
Lucas, M., Nguyen, L.T.T., Guber, A., Kravchenko, A.N., 2022. Cover crop influence on pore size distribution and biopore dynamics: Enumerating root and soil faunal effects. Frontiers in Plant Science 13, 928569. https://doi.org/10.3389/fpls.2022.928569.
Ma, Q., Zhang, K., Jabro, J.D., Ren, L., Liu, H., 2019. Freeze–thaw cycles effects on soil physical properties under different degraded conditions in Northeast China. Environmental Earth Sciences 78, 321. https://doi.org/10.1007/s12665-019-8323-z.
Messing, I., Jarvis, N.J., 1993. Temporal variation in the hydraulic conductivity of a tilled clay soil as measured by tension infiltrometers. Journal of Soil Science 44, 11–24. https://doi.org/10.1111/j.1365-2389.1993.tb00430.x.
Meurer, K., Barron, J., Chenu, C., Coucheney, E., Fielding, M., Hallett, P., Herrmann, A. M., Keller, T., Koestel, J., Larsbo, M., Lewan, E., Or, D., Parsons, D., Parvin, N., Taylor, A., Vereecken, H., Jarvis, N., 2020a. A framework for modelling soil structure dynamics induced by biological activity. Global Change Biology 26, 5382–5403. https://doi.org/10.1111/gcb.15289.
Meurer, K., Chenu, C., Coucheney, E., Herrmann, A.M., Keller, T., Kätterer, T., Nimblad Svensson, D., Jarvis, N., 2020b. Modelling dynamic interactions between soil structure and the storage and turnover of soil organic matter. Biogeosciences 17, 5025–5042. https://doi.org/10.5194/bg-17-5025-2020.
Mohammadi, M.H., Vanclooster, M., 2011. Predicting the Soil Moisture Characteristic Curve from Particle Size Distribution with a Simple Conceptual Model. Vadose Zone Journal 10, 594–602. https://doi.org/10.2136/vzj2010.0080.
Mohammed, A.K., Hirmas, D.R., Nemes, A., Giménez, D., 2020. Exogenous and endogenous controls on the development of soil structure. Geoderma 357, 113945. https://doi.org/10.1016/j.geoderma.2019.113945.
Naveed, M., Moldrup, P., Vogel, H.-Jö., Lamandé, M., Wildenschild, D., Tuller, M., de Jonge, L.W., 2014. Impact of long-term fertilization practice on soil structure evolution. Geoderma 217–218, 181–189. https://doi.org/10.1016/j.geoderma.2013.12.001.
Nemes, A., Schaap, M.G., Leij, F.J., Wösten, J.H.M., 2001. Description of the unsaturated soil hydraulic database UNSODA version 2.0. Journal of Hydrology 251, 151–162. https://doi.org/10.1016/S0022-1694(01)00465-6.
Newville, M., Otten, R., Nelson, A., Stensitzki, T., Ingargiola, A., Allan, D., Fox, A., Carter, F., Michał, Osborn, R., Pustakhod, D., Lneuhaus, Weigand, S., Aristov, A., Glenn, Deil, C., Mark, Hansen, A.L.R., Pasquevich, G., Foks, L., Zobrist, N., Frost, O., Stuermer, Azelcer, Polloreno, A., Persaud, A., Nielsen, J.H., Pompili, M., Caldwell, S., Hahn, A., 2022. lmfit/lmfit-py: 1.1.0. https://doi.org/10.5281/ZENODO.598352.
Nimmo, J.R., 1997. Modeling Structural Influences on Soil Water Retention. Soil Science Society of America Journal 61, 712–719. https://doi.org/10.2136/sssaj1997.03615995006100030002x.
Or, D., Keller, T., Schlesinger, W.H., 2021. Natural and managed soil structure: On the fragile scaffolding for soil functioning. Soil and Tillage Research 208, 104912. https://doi.org/10.1016/j.still.2020.104912.
Parvin, N., Coucheney, E., Gren, I.-M., Andersson, H., Elofsson, K., Jarvis, N., Keller, T., 2022. On the relationships between the size of agricultural machinery, soil quality and net revenues for farmers and society. Soil Security 6, 100044. https://doi.org/10.1016/j.soisec.2022.100044.
Pieri, L., Bittelli, M., Pisa, P.R., 2006. Laser diffraction, transmission electron microscopy and image analysis to evaluate a bimodal Gaussian model for particle size distribution in soils. Geoderma 135, 118–132. https://doi.org/10.1016/j.geoderma.2005.11.009.
Piikki, K., Söderström, M., 2019. Digital soil mapping of arable land in Sweden – Validation of performance at multiple scales. Geoderma 352, 342–350. https://doi.org/10.1016/j.geoderma.2017.10.049.
Poeplau, C., Don, A., Six, J., Kaiser, M., Benbi, D., Chenu, C., Cotrufo, M.F., Derrien, D., Gioacchini, P., Grand, S., Gregorich, E., Griepentrog, M., Gunina, A., Haddix, M., Kuzyakov, Y., Kühnel, A., Macdonald, L.M., Soong, J., Trigalet, S., Vermeire, M.-L., Rovira, P., van Wesemael, B., Wiesmeier, M., Yeasmin, S., Yevdokimov, I., Nieder, R., 2018. Isolating organic carbon fractions with varying turnover rates in temperate agricultural soils – A comprehensive method comparison. Soil Biology and Biochemistry 125, 10–26. https://doi.org/10.1016/j.soilbio.2018.06.025.
Pollacco, J.A.P., Webb, T., McNeill, S., Hu, W., Carrick, S., Hewitt, A., Lilburne, L., 2017. Saturated hydraulic conductivity model computed from bimodal water retention curves for a range of New Zealand soils. Hydrology and Earth System Sciences 21, 2725–2737. https://doi.org/10.5194/hess-21-2725-2017.
Pollacco, J.A.P., Fernández-Gálvez, J., Carrick, S., 2020. Improved prediction of water retention curves for fine texture soils using an intergranular mixing particle size distribution model. Journal of Hydrology 584, 124597. https://doi.org/10.1016/j.jhydrol.2020.124597.
Pribyl, D.W., 2010. A critical review of the conventional SOC to SOM conversion factor. Geoderma 156, 75–83. https://doi.org/10.1016/j.geoderma.2010.02.003.
Rabot, E., Wiesmeier, M., Schlüter, S., Vogel, H.-J., 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.
Regelink, I.C., Stoof, C.R., Rousseva, S., Weng, L., Lair, G.J., Kram, P., Nikolaidis, N.P., Kercheva, M., Banwart, S., Comans, R.N.J., 2015. Linkages between aggregate formation, porosity and soil chemical properties. Geoderma 247–248, 24–37. https://doi.org/10.1016/j.geoderma.2015.01.022.
Robinson, D.A., Thomas, A., Reinsch, S., Lebron, I., Feeney, C.J., Maskell, L.C., Wood, C. M., Seaton, F.M., Emmett, B.A., Cosby, B.J., 2022. Analytical modelling of soil porosity and bulk density across the soil organic matter and land-use continuum. Scientific Reports 12, 7085. https://doi.org/10.1038/s41598-022-11099-7.
Ross, P.J., Smettem, K.R.J., 1993. Describing Soil Hydraulic Properties with Sums of Simple Functions. Soil Science Society of America Journal 57, 26–29. https://doi.org/10.2136/sssaj1993.03615995005700010006x.
Sandin, M., Koestel, J., Jarvis, N., Larsbo, M., 2017. Post-tillage evolution of structural pore space and saturated and near-saturated hydraulic conductivity in a clay loam soil. Soil and Tillage Research 165, 161–168. https://doi.org/10.1016/j.still.2016.08.004.
Schlüter, S., Weller, U., Vogel, H., 2011. Soil-structure development including seasonal dynamics in a long-term fertilization experiment. Journal of Plant Nutrition and Soil Science 174, 395–403. https://doi.org/10.1002/jpln.201000103.
Schwen, A., Bodner, G., Scholl, P., Buchan, G.D., Loiskandl, W., 2011. Temporal dynamics of soil hydraulic properties and the water-conducting porosity under different tillage. Soil and Tillage Research 113, 89–98. https://doi.org/10.1016/j.still.2011.02.005.
Smith, C.W., Johnston, M.A., Lorentz, S., 1997. Assessing the compaction susceptibility of South African forestry soils. I. The effect of soil type, water content and applied pressure on uni-axial compaction. Soil and Tillage Research 41, 53–73. https://doi.org/10.1016/S0167-1987(96)01084-7.
Staff, S.S.D., 2017. Soil survey manual, USDA Handbook No, 18. ed. Government Printing Office, Washington, D.C.
Toreti, A., 2014. Gridded Agro-Meteorological Data in Europe. European Commission, Joint Research Centre (JRC) [Dataset] PID: http://data.europa.eu/89h/jrc-marsop4-7-weather_obs_grid_2019.
Totsche, K.U., Amelung, W., Gerzabek, M.H., Guggenberger, G., Klumpp, E., Knief, C., Lehndorff, E., Mikutta, R., Peth, S., Prechtel, A., Ray, N., Kögel-Knabner, I., 2018. Microaggregates in soils. Journal of Plant Nutrition and Soil Science 181, 104–136. https://doi.org/10.1002/jpln.201600451.
Unger, P.W., 1991. Overwinter Changes in Physical Properties of No-Tillage Soil. Soil Science Society of America Journal 55, 778–782. https://doi.org/10.2136/sssaj1991.03615995005500030024x.
Vogel, H.-J., Weller, U., Schlüter, S., 2010. Quantification of soil structure based on Minkowski functions. Computers & Geosciences 36, 1236–1245. https://doi.org/10.1016/j.cageo.2010.03.007.

Wilkert, P., 1983. Studier av markprofiler i svenska åkerjordar - En faktasammanställning (No. 130). Department of Soil Science, Division of Agricultural Hydrotechnics, Uppsala, Sweden.

Wu, X., Yuan, Z., Li, D., Zhou, J., Liu, T., 2023. Geographic variations of pore structure of clayey soils along a climatic gradient. CATENA 222, 106861. https://doi.org/10.1016/j.catena.2022.106861.

Yoon, S.W., Giménez, D., 2012. Entropy Characterization of Soil Pore Systems Derived From Soil-Water Retention Curves. Soil Science 177, 361–368. https://doi.org/10.1097/SS.0b013e318256ba1c.

You, T., Li, S., Guo, Y., Wang, C., Liu, X., Zhao, J., Wang, D., 2022. A superior soil–water characteristic curve for correcting the Arya-Paris model based on particle size distribution. Journal of Hydrology 613, 128393. https://doi.org/10.1016/j.jhydrol.2022.128393.

Zhang, X., Neal, A.L., Crawford, J.W., Bacq-Labreuil, A., Akkari, E., Rickard, W., 2021. The effects of long-term fertilizations on soil hydraulic properties vary with scales. Journal of Hydrology 593, 125890. https://doi.org/10.1016/j.jhydrol.2020.125890.

Zhang, Y., Weihermüller, L., Toth, B., Noman, M., Vereecken, H., 2022. Analyzing dual porosity in soil hydraulic properties using soil databases for pedotransfer function development. Vadose Zone Journal 21. https://doi.org/10.1002/vzj2.20227.