Received: 15 July 2021 | Revised: 26 March 2022 | Accepted: 13 May 2022

DOI: 10.1111/ejss.13254

ORIGINAL ARTICLE

European Journal of Soil Science WILEY

Relative entropy as an index of soil structure

Tobias Klöffel[1] | **Nicholas Jarvis**[1] | **Sung Won Yoon**[2] | **Jennie Barron**[1] | **Daniel Giménez**[3]

[1]Department of Soil and Environment, Swedish University of Agricultural Sciences, Uppsala, Sweden

[2]Department of Nano, Chemical & Biological Engineering, Seokyeong University, Seoul, South Korea

[3]Department of Environmental Sciences, Rutgers University, New Brunswick, New Jersey, USA

Correspondence
Tobias Klöffel, Department of Soil and Environment, Swedish University of Agricultural Sciences, Uppsala, Sweden.
Email: tobias.kloffel@slu.se

Funding information
Swedish University of Agricultural Sciences

Abstract

Soil structure controls key soil functions in both natural and agro-ecosystems. Thus, numerous attempts have been made to develop methods aiming at its characterization. Here we propose an index of soil structure that uses relative entropy to quantify differences in the porosity and pore(void)-size distribution (VSD) between a structured soil derived from soil water retention data and the same soil without structure (a so-called reference soil) estimated from its particle-size distribution (PSD). The difference between these VSDs, which is the result of soil structure, is quantified using the Kullback–Leibler Divergence (KL divergence). We applied the method to soil data from two Swedish field experiments that investigate the long-term effects of soil management (fallow vs. inorganic fertilizer vs. manure) and land use (afforested land vs. agricultural land dominated by grass/clover ley) on soil properties. The KL divergence was larger for the soil receiving regular addition of manure compared with the soils receiving no organic amendments. Furthermore, soils under afforested land showed significantly larger KL divergences compared to agricultural soils near the soil surface, but smaller KL divergences in deeper soil layers, which closely mirrored the distribution of organic matter in the soil profile. Indeed, a significant positive correlation ($r = 0.374$, $p < 0.001$) was found between soil organic carbon concentrations and KL divergences across all sites and treatments. Despite challenges related to modelling the VSD of the reference soil without structure, the proposed index proved useful for evaluating differences in soil structure in response to soil management and land-use change and reflected the expected effects of soil organic matter on soil structure. We conclude that relative entropy shows great potential to serve as an easy-to-use index of soil structure, as it only requires widely available data on soil physical and hydraulic properties.

Highlights

- A new index of soil structure is proposed based on relative entropy
- A method is developed that separates the effects of soil texture and structure on the pore space

- The index identified soil structural differences in response to land use and soil organic carbon concentrations (SOC)
- The index shows the potential to serve as an easy-to-use metric of soil structure

KEYWORDS

Arya and Paris model, calcium nitrate, dual porosity, fertilization, Kosugi model, manure, pore-size distribution, soil physical quality

1 | INTRODUCTION

Soil structure, defined as the spatial arrangement of soil solid constituents and the pore space (Rabot et al., 2018), is a key determining factor for a multitude of soil processes supporting the functioning of natural and agricultural ecosystems (Dominati et al., 2010; Or et al., 2021; Robinson et al., 2009). Examples are the regulation of biochemical cycling by controlling trophic interactions of soil biota (Erktan et al., 2020), the support of crop vigour (Or et al., 2021) and the control of fundamental hydrological processes such as surface runoff, infiltration and drainage (Fatichi et al., 2020; Mueller et al., 2013). Soil structure can also serve as an indicator of soil development (Bucka et al., 2021). The quantification and assessment of soil structure as well as the development of related indices are therefore of critical importance and have been the subject of extensive research in the past (e.g., Crawford et al., 1995; Dexter, 2004; Reynolds et al., 2009; Vogel et al., 2010; Yoon & Giménez, 2012). Although the use of advanced imaging technologies (e.g., X-ray computed tomography) to investigate soil structure is gaining increasing popularity (e.g., Jarvis, Forkman, et al., 2017; Jarvis, Larsbo, & Koestel, 2017; Lucas et al., 2020; Luo et al., 2008), their availability remains limited and their application expensive, while the processing of large datasets is time-consuming (Young et al., 2001). To avoid these limitations, it is desirable to develop a quantitative index of soil structure that only requires data from routine soil physical analyses.

The Minkowski functions represent a concise way to describe the geometry and topology of a multi-scale binary medium like soil (Vogel et al., 2010). These are defined as the volume and connectivity of the pore phase and the surface area and curvature of its interface with the solids, all expressed as a function of pore diameter. Many endeavours to quantify soil structure have focused directly or indirectly on the use of the soil water retention curve (SWRC) as a proxy for the pore(void)-size distribution (VSD), that is, one of the Minkowski functions, and its integral, the total soil porosity (ϕ). The SWRC describes the functional relationship between the water content (volumetric or gravimetric, θ) and matric potential (ψ_{m}) and allows the estimation of ϕ and the VSD, both of which are strongly affected by the physical, chemical and biological processes underlying the dynamic evolution of soil structure (Meurer, Barron, et al., 2020; Regelink et al., 2015). However, the direct quantification of soil structure from the SWRC is difficult since it is also influenced by the pore space created by the random packing of soil particles, also referred to as textural pore space (Nimmo, 1997). Thus, as noted by Yoon and Giménez (2012), robust quantification of soil structure from the SWRC and VSD requires a method that is insensitive to the particle-size distribution (PSD).

Entropy, being a measure of complexity, information and "order", has been recognized as a potential indicator for soil change such as the formation or degradation of soil structure (Dexter, 1977; Lin, 2011; Meurer, Barron, et al., 2020; Yoon, 2009; Yoon & Giménez, 2012). Structure-forming processes including the activity of soil biota, the influence of roots, wet-dry and freeze–thaw cycles, and soil tillage can be regarded as actions that "add information" to a given soil volume, thereby increasing the entropy of properties and functions related to soil structure, such as the VSD. For example, the multi-scale nature of structure-forming processes can result in a bimodal or multimodal VSD (e.g., Dexter et al., 2008; Durner, 1994; Reynolds, 2017) and, in this way, increase its heterogeneity. This implies that the entropy of a VSD is minimized for a hypothetical soil devoid of any structural features (Meurer, Barron, et al., 2020) and that the difference in entropy between such a hypothetical soil and a natural soil will increase with the degree of structural development in the latter. Therefore, relative entropy, being a measure of entropy *difference* between two distributions, may be a suitable indicator for the degree of soil structure. Indeed, the concept of relative entropy is not new to soil science and has been applied in various contexts (e.g., Hou & Rubin, 2005; Kim et al., 2016; Tarquis et al., 2008).

In this study, we propose the use of relative entropy as an index of soil structure. With this approach, we aim to eliminate the effects of soil texture on the VSD by exploiting the relationship between the PSD and the

VSD. Furthermore, we only use data from routine soil physical analyses. In the following, we first explain relative entropy and show how it can be used as an index of soil structure. We then apply it to soil data from two Swedish field experiments investigating the long-term effects of (i) soil management practices (bare fallow vs. mineral fertilizer addition vs. farmyard manure addition) and (ii) land use (tree plantations vs. crop rotations dominated by grass/clover ley) on soil properties. Finally, we discuss the result of these preliminary tests and the suitability of relative entropy as an index of soil structure.

2 | MATERIALS AND METHODS

2.1 | Relative entropy as an index of soil structure

Relative entropy, also known as the Kullback–Leibler Divergence (KL divergence), originates from information theory and is a dimensionless measure of the difference between two probability distributions of the same random variable (Goodfellow et al., 2016; Kullback & Leibler, 1951). One of the probability distributions acts as a reference distribution to which the second distribution is compared. The choice of the reference probability distribution is important since the KL divergence is asymmetric, that is, it differs with respect to this choice. Another property of the KL divergence is that it is always non-negative, such that a value of zero is obtained for two identical distributions (Bishop, 2006; Goodfellow et al., 2016).

In the case of two discrete probability distributions of the same random variable X, the KL divergence is defined as follows (MacKay, 2002):

$$D_{KL}\left(P\middle\|Q\right) = \sum_{x\in X} P(x)\log\left(\frac{P(x)}{Q(x)}\right) \tag{1}$$

where $Q(x)$ is the value of the reference probability distribution Q at x and $P(x)$ is the value of the probability distribution P at x. If X is continuous, the KL divergence reads as follows (Bishop, 2006):

$$D_{KL}\left(P\middle\|Q\right) = \int_{-\infty}^{\infty} p(x)\log\left(\frac{p(x)}{q(x)}\right) \tag{2}$$

where q is the probability density function (PDF) of the reference distribution and p is the PDF of the second distribution. Figure 1 illustrates how the KL divergence is determined and how it increases as the difference between two probability density curves increases.

Soil structure is manifested through its direct impact on the soil pore space characteristics by enhancing ϕ and changing the median pore radius (e.g., Bodner et al., 2013; Kreiselmeier et al., 2019) as well as the heterogeneity (i.e., variance) of the VSD (e.g., Crawford et al., 1995; Hwang & Choi, 2006). The latter has been suggested to be the result of processes such as aggregation of primary particles and the influence of soil biota, roots and microcracking (Hwang & Choi, 2006). Broader VSDs show larger standard deviations and, thus, lead to a larger entropy (Yoon & Giménez, 2012). This motivates the use of the KL divergence as a measure to quantify changes in VSD in response to the formation and degradation of soil structure. For this, a suitable reference distribution is required.

The most suitable reference to quantify changes in VSD due to soil structure is a hypothetical soil devoid of any structural features, which we adopt and define here as the "reference soil". The porosity and VSD of a soil without structure are a function of the size distribution, shape and packing of its particles (e.g., Arya & Heitman, 2015; Arya & Paris, 1981; Crisp & Williams, 1971; Fiès & Bruand, 1998; Fiès & Stengel, 1981; Haverkamp & Parlange, 1986; Nimmo et al., 2007). Strictly speaking, this implies that the reference soil is unique for any natural soil given a specific PSD. This definition makes the reference soil conceptually equivalent to the textural component of a natural soil (Childs, 1969; Meurer, Barron, et al., 2020).

The KL divergence can be applied as an index of soil structure by following the individual steps 1–4 described below. We focus on the lognormal distribution to describe the PSD and VSD of soil (e.g., Brutsaert, 1966; Hwang & Choi, 2006; Kosugi, 1996), since it is uniquely defined by parameters with a clear physical meaning. These are the median (or geometric mean) particle or pore radius and the standard deviation (σ), which characterizes the broadness of a PSD or VSD (Kosugi, 1996).

2.1.1 | Step 1: Modelling the PSD of the reference soil

The Kosugi (1996) model is adopted to model the PSD in the following way (Hwang & Powers, 2003):

$$g(r_P) = \frac{1}{r_P\sigma_P\sqrt{2\pi}}\exp\left\{-\frac{(\ln r_P - \ln r_{m,P})}{2\sigma_P^2}\right\} \tag{3}$$

where r_P is the particle radius [cm], $r_{m,P}$ is the median particle radius [cm] and σ_P is the standard deviation of the PSD.

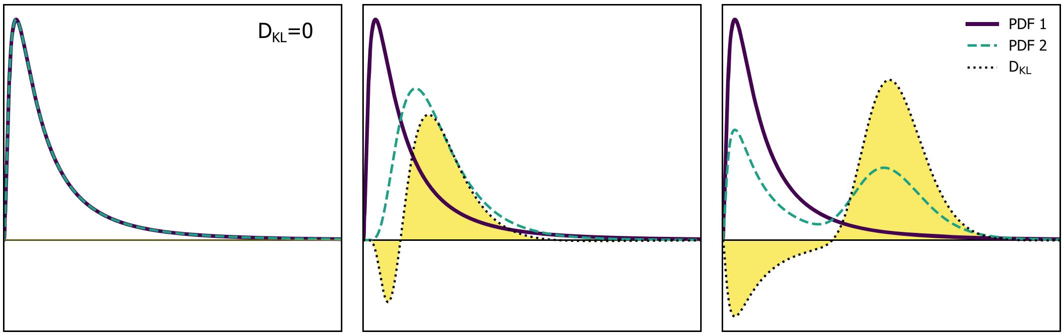

FIGURE 1 Illustration of how the KL divergence (D_{KL}) is determined. Each subplot shows two probability density curves that differ from one another increasingly from left to right. The purple solid curve represents the reference distribution to which the green dashed curve is compared. The black dotted curve is obtained by solving the expression in Equation (1) for each point along the horizontal axis. Integrating the yellow area above/below the dotted curve yields the KL divergence, which equals zero ($D_{KL} = 0$) when the two curves are identical (left subplot)

2.1.2 | Step 2: Modelling the VSD of the structured soil

The VSD of the structured soil is estimated from soil water retention measurements, that is, θ-ψ_m measurement pairs. First, ψ_m is mapped to an equivalent pore radius (r) through the Young–Laplace relationship:

$$r = -\frac{2\gamma\cos\delta}{\rho_w g \psi_m} \tag{4}$$

where ψ_m and r are in [cm], γ is the surface tension at the air-water interface [g s^{-2}], δ is the contact angle between the water and solid phase [°], ρ_w is the density of water [g cm^{-3}], and g is the acceleration due to gravity [cm s^{-2}]. In this study, we assume full contact between water and solid phase and the physical properties of water at 20°C (Brutsaert, 1966). This simplifies Equation (4) to $r = -0.149 \div \psi_m$.

Many mathematical expressions have been proposed to model the relationship between ψ_m/r and θ (Assouline & Or, 2013; Sillers et al., 2001), some of which are derived from the lognormal distribution (Brutsaert, 1966; Kosugi, 1994, 1996). We adopt the model by Kosugi (1996), which, for the unimodal case, is given as follows:

$$f_S(r) = \frac{(\theta_s - \theta_r)}{r\sigma_S\sqrt{2\pi}} \exp\left\{-\frac{(\ln r - \ln r_{m,S})}{2\sigma_S^2}\right\} \tag{5}$$

where θ_s and θ_r are the saturated and residual water contents [m^3 m^{-3}], respectively, σ_S is the standard deviation of the lognormal VSD [−] and $r_{m,S}$ is the median pore-radius [cm]. The subscript "S" indicates that the parameters refer to the VSD of the structured soil.

Water retention measurements suggesting a bimodal or multimodal VSD can be modelled by superimposing two or more unimodal distributions (Dexter et al., 2008; Durner, 1994; Pollacco et al., 2017; e.g., Ross & Smettem, 1993). Generally, this procedure leads to an improved description of the SWRC in structured soils (Reynolds, 2017). Using the Kosugi (1996) model, this gives:

$$f_S(r) = \sum_{i=1}^{n} \frac{(\theta_{s,i} - \theta_{r,i})}{r\sigma_{S,i}\sqrt{2\pi}} \exp\left\{-\frac{(\ln r - \ln r_{m,S,i})}{2\sigma_{S,i}^2}\right\} \tag{6}$$

where n indicates the number of superimposed pore classes/domains (i.e., lognormal distributions) and the subscript i defines the affinity of a parameter to the respective lognormal distribution. Note that for $i \geq 2$, θ_r no longer represents the residual water content, but the saturated water content of the $i-1$th pore class/domain (Pollacco et al., 2017).

2.1.3 | Step 3: Modelling the VSD of the reference soil

The PSD is translated into the VSD of the reference soil without structure using the model by Arya and Paris (1981). This model has been noted to perform particularly well for soils with little structural development (Nimmo et al., 2007). It assumes that the VSD and PSD are linearly related, which implies shape similarity between both distributions. The few attempts that have been undertaken to experimentally investigate the textural pore space showed that particles of different sizes result in pores of specific sizes and characteristics (Fiès et al., 1981; Fiès & Bruand, 1990, 1998; Fiès & Stengel, 1981). This suggests that the VSD of the textural pore space closely follows the PSD. In particular, we assume that the PSD defines the main properties of the textural VSD such as σ and the median pore radius. Moreover, the assumption of shape similarity should be valid for a soil without structure as has been shown for poly-disperse sphere packs with dense random packing (Assouline & Rouault, 1997).

Arya and Paris (1981) assume that a natural soil can be represented by several uniform sphere packs, which are packed in "discrete domains" and subsequently assembled to give the same bulk density (ρ_b) as the natural soil. First, the PSD curve is subdivided into a number of size fractions (usually 20 or more) and the relative abundance of each fraction is multiplied with the sample weight yielding weight fractions. Given the soil particle density (ρ_s) and ρ_b, these weight fractions provide information on the number of uniform spheres in each domain, which are then assembled into a hypothetical closed-packed cube with the void ratio of the bulk soil (e). Finally, r is calculated for the i^{th} size fraction with the following relationship:

$$r_i = 0.816 r_{P,i} \sqrt{e n_i^{(1-\alpha_i)}} \tag{7}$$

where $r_{P,i}$ is the mean particle radius of the i^{th} size fraction [cm] and α_i is a scaling parameter that links the ideal sphere pack to the natural soil (Arya et al., 1999). In a later study, Arya et al. (1999) present and discuss different ways to determine this parameter. Note that α_i becomes 1 when the ideal sphere pack and the natural soil are equivalent.

To determine the VSD of the reference soil using Equation (7), e is calculated by

$$e = \frac{\phi_{tex}}{1-\phi_{tex}} \tag{8}$$

where ϕ_{tex} denotes the porosity of the textural pore space. As noted before, detailed empirical studies on the textural pore space remain scarce. This might be due to challenges related to this task such as controlling the packing of soil particles, which exacerbates comparability between individual experiments. Furthermore, theoretical studies based on multicomponent sphere packs (e.g., Farr & Groot, 2009; Gupta & Larson, 1979; Shen et al., 2019) are not especially applicable to estimate ϕ_{tex} of real soils because soil particles in these models are represented as spheres that are packed in the densest way possible. This can yield unrealistically small values of porosity. In fact, ϕ in these models can approach zero when the PSD becomes increasingly right-skewed, that is, when a large number of small spheres can be fitted into the gaps between larger ones (Farr & Groot, 2009; Gupta & Larson, 1979). Here we assume a random closed packing of soil particles in the reference soil and follow Nimmo (2013), who suggested that a value of 0.30 should be an appropriate estimate for ϕ_{tex}. We acknowledge that this value may not be a reasonable estimate for all soils. However, as we demonstrate below, the KL divergence is relatively insensitive to ϕ_{tex}, so that its precise estimation is not of critical importance.

The shape similarity between PSD and VSD allows $r_{P,i}$ in Equation (7) to be replaced with $r_{m,P}$ to obtain the median pore radius of the reference soil ($r_{m,R}$). This we do setting α to 1, which gives:

$$r_{m,R} = 0.816\, r_{m,P}\sqrt{e} \tag{9}$$

Furthermore, the assumption that the VSD of a soil without structure closely follows its PSD justifies that:

$$\sigma_R = \sigma_P \tag{10}$$

where σ_R is the standard deviation of the VSD of the reference soil. Finally, it is reasonable to assume that θ_r is the same for both structured and reference soil and that θ_s can be approximated by ϕ_{tex} so that the VSD of the reference soil can be described by the following:

$$f_R(r) = \frac{(\phi_{tex}-\theta_r)}{r\sigma_R\sqrt{2\pi}} \exp\left\{-\frac{(\ln r - \ln r_{m,R})}{2\sigma_R^2}\right\} \tag{11}$$

2.1.4 | Step 4: Calculating the KL-divergence between reference and structured soil

Substituting the VSDs for the structured (Equation (5)) and reference soils (Equation (11)) into Equation (2), where $q(x)$ represents the VSD of the reference soil (f_R) and $p(x)$ the VSD of the structured soil (f_S), gives the analytical expression for the KL divergence as (see Appendix A for derivation):

$$D_{KL}\left(f_S \middle\| f_R\right) = (\theta_s-\theta_r)\left(log\frac{(\theta_s-\theta_r)\sigma_R}{(\phi_{tex}-\theta_r)\sigma_s} - \frac{1}{2} + \left[\frac{\sigma_S^2+(\log r_{m,S}-\log r_{m,R})^2}{2\sigma_R^2}\right]\right) \tag{12}$$

It is clear from Equation (12) that the KL divergence increases as θ_s, σ_S and the difference between $r_{m,S}$ and $r_{m,R}$ increase, whereas it decreases as ϕ_{tex} and σ_R

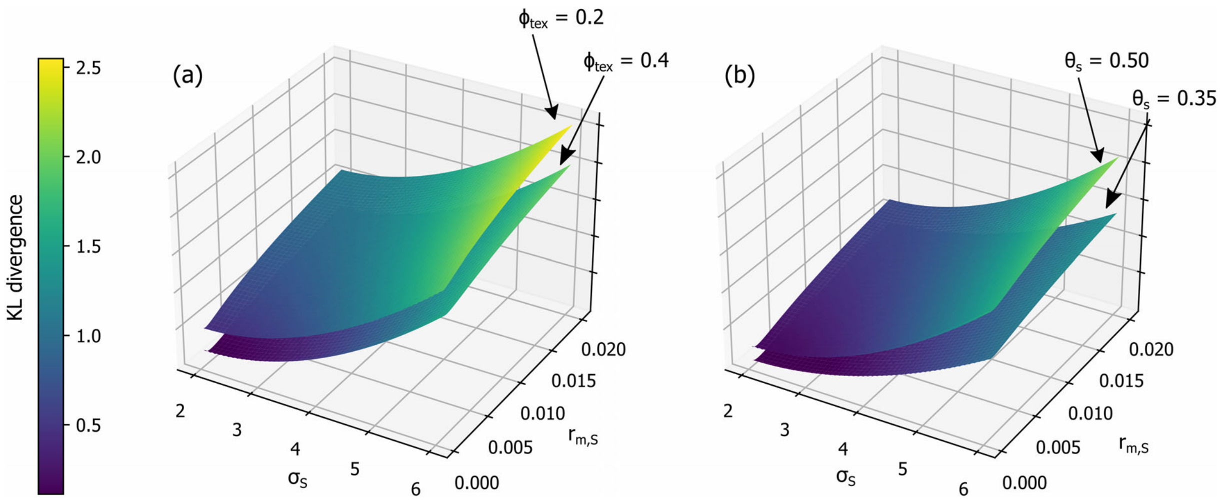

FIGURE 2 3D plots illustrating the sensitivity of parameters for calculating the KL divergence related to the structured soil (i.e., with subscript "S" in Equation (12)) with varying textural porosity (ϕ_{tex}, (a)) and saturated water content (θ_s, (b)). The parameters related to the reference soil were fixed at values that fall into the range of natural soils (see Tables S1 and S2), where the median pore radius ($r_{m,R}$) and the standard deviation (σ_R) were set to 0.002 cm and 2, respectively. The residual water content was set to zero. In subplot (a) θ_s was fixed at 0.50 $m^3\ m^{-3}$ and in subplot (b) ϕ_{tex} was fixed at 0.35 $m^3\ m^{-3}$

increase. However, the KL divergence is not equally sensitive to each of these parameters. The standard deviations σ_S and σ_R appear as squared terms, which means that they should have a stronger impact on the KL divergence than the other parameters. The 3D plots shown in Figure 2 illustrate the sensitivity of the parameters in Equation (12) that are related to the structured soil, whilst the parameters of the reference (non-structured) soil remain fixed. This is shown for different values of ϕ_{tex} (Figure 2a) and θ_s (Figure 2b). It can be seen that $r_{m,S}$ and ϕ_{tex} show relatively minor effects on the KL divergence, while θ_s becomes more relevant with increasing KL divergence by acting as a scalar (Equation (12)). The largest effect on the KL divergence, however, is exerted by σ_S.

An analytical expression for the KL divergence cannot be obtained when the soil water retention curve is best described by a bimodal VSD, which is often the case for structured soils (Dexter et al., 2008; Jensen et al., 2019; Reynolds, 2017). This proved to be the case for nearly all of the samples in the applications of the method to the two field experiments described below. We therefore determined the KL divergence numerically by inserting Equation (6) (instead of Equation (5)) for $p(x)$ in Equation (2) and applying discrete integration similar to Riemann's integral (e.g., Axler, 2020). For this, a pore-size range with lower and upper limits at 0 and 10 cm was defined and subdivided into four sub-intervals (0–10^{-5} cm, 10^{-5}–10^{-3} cm, 10^{-3}–0.1 cm, 0.1–10 cm). Each sub-interval was partitioned into 5×10^8 rectangles of equal width and the height of each rectangle was determined by solving the expression in Equation (1) at the mid-point of each rectangle. The number of rectangles was considered a good compromise between computation time and accuracy. Finally, the area of each rectangle was calculated and all areas were summed to obtain the KL divergence. The procedure can be summarized with the following equation:

$$D_{KL}\left(P\middle\|Q\right) = \sum_{j=1}^{4}\sum_{i=1}^{5\times10^8}\left(x_{j,i+1}-x_{j,i}\right) \left[P\left(\frac{x_{j,i}+x_{j,i+1}}{2}\right)\log\left(\frac{P\left(\frac{x_{j,i}+x_{j,i+1}}{2}\right)}{Q\left(\frac{x_{j,i}+x_{j,i+1}}{2}\right)}\right)\right] \tag{13}$$

where j refers to the subinterval and i $(i+1)$ to the left (right) position of the border of each rectangle. The other variables are the same as in Equation (1). The procedure was successfully validated for a unimodal case by comparing results obtained with the analytical solution (Equation (12)) with those obtained by the discrete integration method.

2.2 | Site descriptions and measurements

In the following, we demonstrate the use of the KL divergence as an index of soil structure for two case studies from Swedish field experiments. The first dataset is from a field experiment initiated in 1956 to monitor the long-term effects of mineral nitrogen fertilizers and organic matter amendments on soil organic matter contents, crop yields and physical soil properties (Kirchmann et al., 1994). The experimental site is located near the Swedish University of Agricultural Sciences at Ultuna, close to Uppsala (59.92°N, 17.65°E). The topsoil has a clay loam texture and the soil was classified as a *Eutric Cambisol* (FAO, 1989). The soil PSD at Ultuna was measured for seven classes using sedimentation and wet sieving (Kirchmann et al., 1994). Note that the soil texture was assumed to be the same for all treatments due to the small size (4 m^2) and close proximity of the plots. All plots in this experiment are managed by hand with digging in autumn and spring and have been planted with fodder maize since the year 2000. Further details about the experiment and site conditions are described in Kirchmann et al. (1994). Details on sampling for soil water retention are given in Svensson (2020) and shortly summarized here. Undisturbed soil cores (65.5 mm inner diameter and 74.8 mm height) were collected during early autumn in 2019 before harvest from plots with three different treatments: a bare fallow treatment with no additions (hereafter "fallow" treatment) and two cropped treatments. One of these is fertilized with calcium nitrate ($Ca[NO_3]_2$) at a rate of 80 kg N ha^{-1} (hereafter "$Ca[NO_3]_2$" treatment), while the other receives biennial additions of solid cow manure at a rate of 9.5 t ha^{-1} (hereafter "manure" treatment). Two replicate cores per treatment were sampled from four blocks (eight replicates per treatment in total) in between rows of maize just below the soil surface. Of these, one replicate core from the fallow treatment had to be discarded. Water retention was measured for each core on a suction plate at ψ_m of −10, −30, −100, −300 and − 600 hPa. Furthermore, dry ρ_b and ρ_s were determined for each replicate and ϕ was calculated. Particle density was calculated from the volume displacement of a sample of fine earth (<2 mm) with ethanol. Tables 1 and 2 show selected soil properties for the three treatments at Ultuna.

The second dataset was taken from Messing et al. (1997), who measured soil hydraulic properties at several sites in southern Sweden on adjacent fields with similar site conditions but under different land uses. One field represented agricultural land that had been afforested with aspen (*Populus deltoides*) or silver birch (*Betula pendula*) 30 years before the study was conducted (hereafter termed the "FOR" treatment), while the other field represented current agricultural land use dominated by grass/clover leys in rotation with cereals (hereafter termed the "AGR" treatment). For this application, three of the five sites studied by Messing et al. (1997) were selected (Almnäs, Siggebohyttan, Vik) due to their relatively coarse-textured soils (Table 1), as fine-textured soils are covered by the Ultuna case study. Undisturbed soil samples were collected with cylindrical soil cores (inner diameter 72 mm, height 50 mm) from 0–35 cm depth at four (Almnäs and Siggebohyttan) and from 0–30 cm depth at six (Vik) depth intervals. Three to four replicates per depth interval were sampled at each site and for each treatment. Water retention was measured at six values of ψ_m, namely −5, −30, −50, −100, −300 and − 600 hPa using porous sand blocks for −5 hPa and ceramic plates for the other pressure heads. The PSD in seven classes was measured on disturbed soil samples taken at 10 to 15 cm depth with wet sieving and sedimentation using the pipette method. We assume that the soil texture at this depth is representative of the full depth ranges due to past (FOR treatment) and ongoing (AGR treatment) tillage. Bulk density and ρ_s, which were used to calculate ϕ, were determined from the undisturbed and disturbed samples, respectively. Selected properties of the soils from the three sites are summarised in Tables 1 and 2.

2.3 | Fitting distributions and statistical analysis

All fitting was done with the least-squares method (Levenberg–Marquardt algorithm) available in the Python module *SciPy* (Virtanen et al., 2020). The parameters $r_{m,P}$ and σ_P in Equation (3) were obtained by integration and subsequent fitting to the cumulative PSD. Note that only unimodal PSDs can be modelled with Equation (3) although multimodal PSDs do exist (Fredlund et al., 2000). This issue is addressed in the discussion below.

The modelling of the VSD on the water retention measurements of the structured soils was mostly done assuming a bimodal VSD (Equation (6) with $i = 2$). This improved the goodness-of-fit as compared with using Equation (5) assuming a unimodal VSD. The bimodality of the pore system from the Ultuna data clearly resulted from a well-developed macropore system in this fine-textured soil. In contrast, many of the samples from the dataset collected by Messing et al. (1997) show bimodality in the size range of matrix pores. In addition, some of the samples in this dataset suggest a third pore domain reflecting the presence of macropores. Clearly, a bimodal model cannot be made to fit satisfactorily to data that indicates three pore regions. Preliminary testing showed that the estimated KL divergence is very sensitive to the quality of the fit to the water retention measurements across a wide range of soil water tensions. This third (macropore) region was therefore effectively neglected in the fitting of the bimodal model to the data for the coarse-textured soils at Almnäs, Siggebohyttan and Vik, by excluding the measured ϕ in the fitting. The measured ϕ was only included in the curve fitting for the samples at Ultuna, with its value fixed at a pore radius of 3 mm ($\psi_m \approx$

TABLE 1 Soil properties of sites selected to demonstrate the KL divergence

Site	Treatment	Sand[a] $[\mathrm{g\,g^{-1}}]$	Silt[b] $[\mathrm{g\,g^{-1}}]$	Clay[c] $[\mathrm{g\,g^{-1}}]$
Ultuna[d]	-	0.22	0.41	0.37
Almnäs[e]	AGR	0.69	0.20	0.11
	FOR	0.66	0.23	0.11
Siggebohyttan[e]	AGR	0.44	0.48	0.08
	FOR	0.61	0.31	0.08
Vik[e]	AGR	0.53	0.38	0.09
	FOR	0.59	0.33	0.08

Abbreviations: AGR, agricultural land; FOR, afforested land.

[a]2–0.06 mm particle diameter.

[b]0.06–0.002 mm particle diameter.

[c]<0.002 mm particle diameter.

[d]Data from Kirchmann et al. (1994) (assumed to be representative for all treatments at Ultuna).

[e]Data from Messing et al. (1997) (assumed to be representative for the entire investigated depth range).

−0.5 hPa), which is equivalent to assuming that there were no pores larger than this.

Adopting the Kosugi (1996) model to describe a bimodal VSD requires the optimization of seven parameters (see Equation (6)). Determining seven fitting parameters by inverse modelling against datasets comprising six data points clearly raises the issue of non-uniqueness, that is, the likelihood that different parameter sets will fit the measurements equally well as judged by some goodness-of-fit measure (Beven, 1993; Fernández-Gálvez et al., 2021). We therefore investigated ways to constrain the fitting to ensure unique solutions. First, we set θ_r to zero, which reduced the number of parameters to six. This is also justifiable in principle because θ was not measured at ψ_m-values less than −600 hPa and, consequently, the residual water content would anyway not be identifiable. We then tested constraining θ_s (the sum of $\theta_{s,1}$ and $\theta_{s,2}$) to equal the measured ϕ to further reduce the number of fitting parameters. However, this notably reduced the quality of the model fit to the data across a wide range of ψ_m-values, especially in cases where a third macropore domain was present. Thus, both $\theta_{s,1}$ and $\theta_{s,2}$ were included as fitting parameters. Besides this, we tested the hypothesis that the small pore region only comprises textural pores by setting $\sigma_{S,1}$ equal to σ_R ($=\sigma_P$) and $r_{m,S,1}$ equal to a fraction of $r_{m,P}$, which reduced the number of fitting parameters to four. However, this procedure also reduced the quality of the model fit to the water retention data, demonstrating that the small pore region also comprises structural pores. Hence, both $\sigma_{S,1}$ and $r_{m,S,1}$ were retained as fitting parameters. The final model for describing the VSD of the structured soil, therefore, requires six parameters to be optimized as follows:

$$f_S(r) = \frac{\theta_{s,1}}{r\sigma_{S,1}\sqrt{2\pi}}\exp\left\{-\frac{(\ln r - \ln r_{m,S,1})}{2\sigma_{S,1}^2}\right\} + \frac{\theta_{s,2}}{r\sigma_{S,2}\sqrt{2\pi}}\exp\left\{-\frac{(\ln r - \ln r_{m,S,2})}{2\sigma_{S,2}^2}\right\} \quad (14)$$

We adopted a procedure for fitting Equation (14) to our data which acknowledges the likelihood of non-unique solutions (Fernández-Gálvez et al., 2021; Pollacco et al., 2017) and therefore uncertainty in the derived KL divergences. Firstly, we found that choosing appropriate initial parameter value guesses for the fitting algorithm was crucial to improve convergence towards physically realistic parameter values. Hence, we derived the set of initial parameter value guesses following physically-based considerations: $\theta_{s,1}$ was assumed to be close to but larger than ϕ_{tex}, so that its initial estimate was set to 0.35 $\mathrm{m^3\,m^{-3}}$. Initial testing revealed a physically plausible correlation between σ_P and $\sigma_{S,1}$ (the standard deviation of the smaller pore domain), such that $2\sigma_P$ was found to be a good initial guess for $\sigma_{S,1}$. Similarly, a correlation was detected between $r_{m,S,1}$ (the median pore radius of the smaller pore domain) and $r_{m,P}$, such that $0.04r_{m,P}$ was considered an appropriate initial guess for $r_{m,S,1}$. Finally, an initial guess for $\theta_{s,2}$ was derived from the difference between the measured porosity ϕ and the initial guess for $\theta_{s,1}$ ($=0.35\,\mathrm{m^3\,m^{-3}}$). Selecting appropriate initial parameter guesses for $\sigma_{S,2}$ and $r_{m,S,2}$ was more difficult since the larger pore domain could either reflect macropore (Ultuna) or matrix pore regions (Almnäs, Siggebohyttan, Vik). To address this issue, we produced 100 initial parameter sets for each VSD to be fitted, where $\sigma_{S,2}$ ranged from 0.2 to 2 and $r_{m,S,2}$ from 0.001 to 0.008 cm, which were considered to be physically realistic ranges for these parameters (Fernández-Gálvez et al., 2021).

TABLE 2 Soil organic carbon concentrations (SOC) and total porosities (ϕ) for all sites and treatments

		Treatment									
		AGR		FOR		Fallow		$Ca(NO_3)_2$		Manure	
Site	**Depth [cm]**	**SOC [%]**	**ϕ [cm^3 cm^{-3}]**	**SOC [%]**	**ϕ [cm^3 cm^{-3}]**	**SOC[a] [%]**	**ϕ [cm^3 cm^{-3}]**	**SOC[a] [%]**	**ϕ [cm^3 cm^{-3}]**	**SOC[a] [%]**	**ϕ [cm^3 cm^{-3}]**
Ultuna	0–5	-	-	-	-	0.9	0.47	1.2	0.53	2.1	0.58
Almnäs	0–5	4.5	0.49	3.8	0.62	-	-	-	-	-	-
	5–15	4.3	0.47	2.2	0.53	-	-	-	-	-	-
	15–25	4.0	0.49	1.5	0.50	-	-	-	-	-	-
	25–35	2.1	0.49	0.4	0.47	-	-	-	-	-	-
Siggebohyttan	0–5	2.5	0.49	5.3	0.59	-	-	-	-	-	-
	5–15	1.9	0.46	3.0	0.52	-	-	-	-	-	-
	15–25	1.8	0.47	2.2	0.51	-	-	-	-	-	-
	25–35	0.5	0.44	0.6	0.43	-	-	-	-	-	-
Vik	0–5	3.1	0.52	5.3	0.63	-	-	-	-	-	-
	5–10	2.9	0.54	3.6	0.58	-	-	-	-	-	-
	10–15	2.9	0.51	2.8	0.56	-	-	-	-	-	-
	15–20	2.9	0.49	2.8	0.52	-	-	-	-	-	-
	20–25	2.5	0.51	2.2	0.54	-	-	-	-	-	-
	25–30	2.0	0.50	1.8	0.54	-	-	-	-	-	-

Note: The values represent averages over replicates.
Abbreviations: AGR, agricultural land; FOR, afforested land.
[a]Values extracted from Figure 6 in Svensson (2020).

Finally, similar to fitting the cumulative PSD, Equation (14) was integrated and optimized against the water retention data for these 100 combinations of initial parameter guesses. Any physically implausible optimized parameter sets were then discarded. This was considered to be the case if (i) any of the six final parameter values were negative, (ii) $\sigma_{S,1}$ was smaller than σ_R ($= \sigma_P$), or (iii) the sum of $\theta_{s,1}$ and $\theta_{s,2}$ was more than 10% larger than the measured ϕ. The root mean squared error (RMSE) was used to evaluate the goodness-of-fit of these physically plausible parameter sets, in order to define a number of "acceptable" parameter sets with which to calculate the KL divergence. This was achieved by retaining all parameter sets with RMSE values less than 10% larger than the best-fit parameter set (i.e., smallest RMSE). The KL divergences of these acceptable parameter sets were then calculated numerically (Equation (13)) and the median KL divergence was used as an index of soil structure for each sample/replicate for subsequent statistical analyses.

The unimodal model (Equation (5)) proved sufficient for modelling the VSD of only a small number of samples from the soils at Vik. In these cases, the initial parameter guesses for σ_S and $r_{m,S}$ were derived in the same way as for $\sigma_{S,1}$ and $r_{m,S,1}$ in the bimodal case. The initial guess for θ_s was set equal to the measured ϕ. The same physical constraints as for the bimodal model were applied in the optimization of the unimodal model, that is, σ_S was only allowed to be larger than σ_R ($=\sigma_P$) and θ_s was not allowed to be 10% larger than the measured ϕ. For the curve fitting, Equation (5) was integrated and the three parameters were optimized against the water retention data. The analytical solution of Equation (12) was used to calculate the KL divergences.

The KL divergences determined for the three treatments at Ultuna and the two land uses at the three sites in Messing et al. (1997) were tested for significant differences using R (R Core Team, 2019). Each group of replicates was tested for normality using the Shapiro–Wilk test. The Tukey method available from the *emmeans* package (Lenth, 2020) was used for pairwise comparison between treatments for the Ultuna data and, within each location, between treatments and depth intervals for data from Almnäs and Siggebohyttan. The KL divergences of one replicate from Vik did not pass the normality test. Hence, pairwise comparison between treatments and depth intervals for this site were done with Dunn's method using the *PMCMR* package (Pohlert, 2014) with automatic Holm *p*-value correction. Differences were considered significant if $p < 0.05$. Correlations were investigated with the Pearson correlation coefficient (Pearson's r).

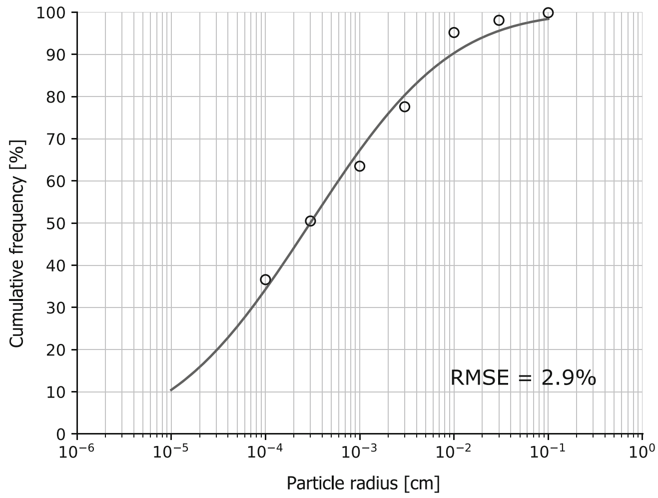

FIGURE 3 Cumulative particle-size distribution for the soil at Ultuna (circles indicate measured data and the line was fitted using Equation (3)). The root-mean-square error (RMSE) is indicated on the plot

3 | RESULTS

The PSD of the fine-textured soil at Ultuna was modelled well with Equation (3) (Figure 3; RMSE = 2.9%). Furthermore, the double lognormal model of Equation (14) gave excellent fits to the water retention data for all treatments at Ultuna (Figures 4, S1 and S2) with a largest RMSE of 0.0076 m^3 m^{-3}. Figure 4 shows fits and KL divergences for the example of the $Ca(NO_3)_2$ treatment at Ultuna. Most of the time, the best fits according to RMSE (indicated by the red triangles) are close to the median KL divergences (indicated by the orange horizontal lines), which were used for statistical analysis. Some variation is evident in the larger pore region, which also affects the KL divergence as indicated by the boxplots. The number of acceptable fits for the replicates ranged between one and 29. Figure 5 shows that the means of the KL divergences increase in the order fallow < $Ca(NO_3)_2$ < manure treatment. Although the mean values are larger for the manure treatment compared to the fallow and $Ca(NO_3)_2$ treatment, these differences were not significant ($p = 0.076$ for fallow vs. manure, and $p = 0.091$ for $Ca(NO_3)_2$ vs. manure). The difference in KL divergence between the $Ca(NO_3)_2$ and fallow treatment was also not significant ($p = 0.99$). The pattern shown in Figure 5 is supported by the values of the fitted parameters, which indicate no large differences in VSD between the $Ca(NO_3)_2$ and fallow treatments, except that the modelled total pore space ($\theta_{s,1} + \theta_{s,2}$) is on average larger in the former compared to the latter (Table S1).

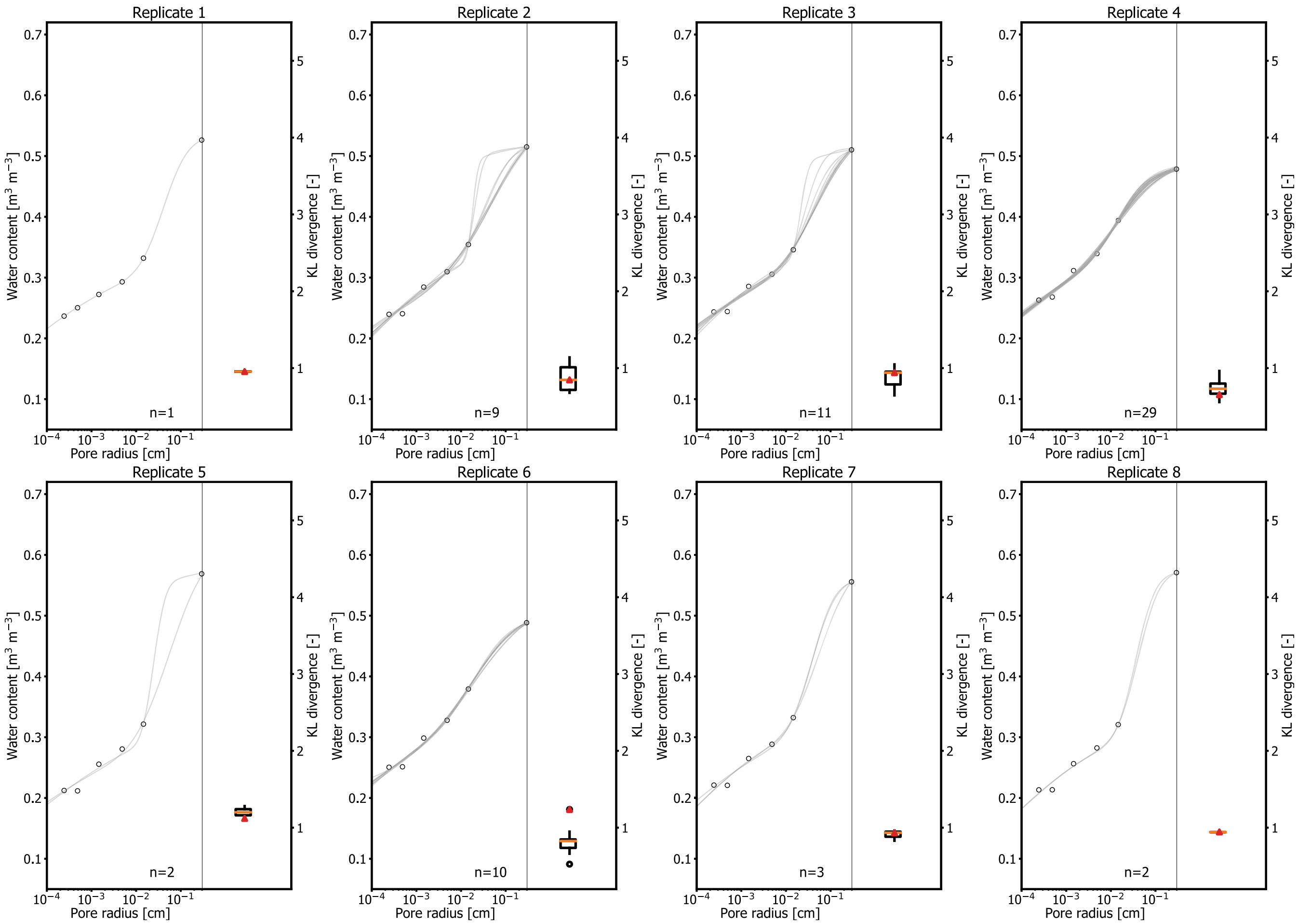

FIGURE 4 Acceptable model fits of cumulative pore-size distributions for each sample replicate (left side of each subplot) with corresponding KL divergences shown as boxplots (right side of each subplot) for the $ca(NO_3)_2$ treatment at Ultuna. On the left side of each subplot, open circles indicate measured values, where the right-most value represents the total measured porosity, which was fixed at a pore radius of 3 mm. The grey lines indicate acceptable model fits and *n* the number of acceptable model fits for the individual sample. Bars in boxplots on the right side of each subplot indicate 1.5 times the interquartile range and open circles indicate values outside this range. The orange horizontal line shows the median KL divergence. This value was used for statistical analysis. The red triangles show the KL divergence of the best fit according to the root mean square error

The fits of Equation (3) to the measured PSDs of the coarse-textured soils at Almnäs, Siggebohyttan and Vik showed larger RMSE's (2.9%–7.3%; Figures S3–S5) as compared to the fine-textured soil at Ultuna. This is because the relative abundance of the fine silt and clay fractions was underestimated, especially at Almnäs and Siggebohyttan (Figures S3–S5). The double lognormal model of Equation (14) and the single lognormal model of Equation (4) yielded excellent fits to the water retention data of these soils with RMSE's below 0.012 $m^3\ m^{-3}$ across all sites, treatments and depths (Figures 6, S6–S10). The example of the FOR treatment at Siggebohyttan is shown in Figure 6. Similar to the Ultuna site, the best fits according to the RMSE are close to the median KL divergences (Figures 6, S6–S10). The fitting parameters obtained for each site and treatment are given in Table S2. Figure 7 shows the KL divergence for the AGR and FOR treatments at Almnäs, Siggebohyttan and Vik for each depth interval. At all three sites, the KL divergence of the FOR treatments decreases significantly with soil depth. At Vik, the KL divergence showed a slight increase for the AGR treatment in the deepest soil layer (Figure 7c), while at Almnäs and Siggebohyttan the KL divergence increased in the soil layer at 15–25 cm depth followed by a decrease in the deepest soil layer (Figure 7a,b). However, there were no significant differences in soil depth in the AGR treatment. Furthermore, the FOR treatments showed larger KL divergences in the upper soil layers and smaller KL divergences in the deeper soil layers compared to the AGR treatments at all three sites (Figure 7). These differences were significant for the uppermost soil layer at Almnäs.

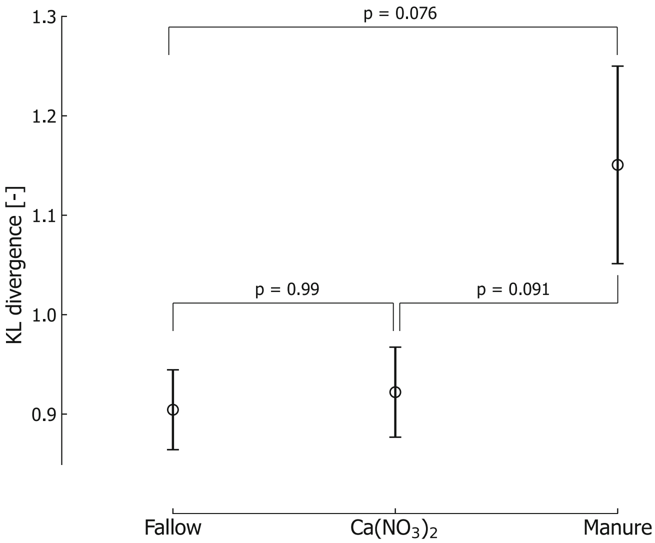

FIGURE 5 KL divergences for the different treatments at Ultuna. Error bars indicate standard errors

It is evident from the optimized model parameters that, at all four sites, larger KL divergences are associated with wider pore-size distributions (i.e., larger $\sigma_{S,1}$ and/or $\sigma_{S,2}$) and/or a larger modelled total pore space (i.e., large $\theta_{s,1}+\theta_{s,2}$) (Tables S1 and S2). A significant relationship was detected between the optimized standard deviations of the PSD (σ_P) and the small pore region ($\sigma_{S,1}$) (Pearson's $r = 0.783$, $p < 0.001$; Figure 8a). Furthermore, although the trend is less strong, the median pore radius of the small pore region ($r_{m,S,1}$) is also significantly correlated with the median particle radius across all sites and treatments ($r_{m,P}$) (Pearson's $r = 0.250$, $p = 0.004$; Figure 8b).

4 | DISCUSSION

Soil PSDs are sometimes more complex than can be described with simple unimodal distributions due to complicated breakdown processes or because soil particles originate from different parent materials (Gardner, 1956). The soils at Almnäs and Siggebohyttan are rather poorly graded, being dominated by the sand fraction, but with a small "hump" in the clay fraction. The lognormal distribution of Equation (3) is not suited to describing bimodal or gap-graded PSDs, and therefore does not match this data very well (Figures S3–S5). As a result, the modelled PSD appears to be narrower than is indicated by the measured values at these sites. This suggests that the KL divergence would have been smaller if the PSD had been modelled more accurately. Nevertheless, the relative difference in KL divergence between the FOR and AGR treatments would not have been affected in this case. More flexible PSD models have been proposed such as the ones by Fredlund et al. (2000) or Assouline et al. (1998), which have been shown to perform better than the lognormal distribution for a broad range of soil textures (Hwang, 2004). These models could be used to determine the VSD of the reference soil given that a linear relationship with the PSD is assumed. The lognormal distribution was chosen here mainly because it allows an explicit analytical solution (Equation (12)) and because its parameters have inherent physical meaning.

One debatable aspect of the method presented here is the linearity that we assumed between the PSD and VSD for the reference soil. It is clear that such an assumption would most likely not be valid for structured soils (Crisp & Williams, 1971; Haverkamp & Parlange, 1986; Hwang & Choi, 2006; Hwang & Powers, 2003). It does not even hold for simulations of tetrahedral (i.e., closest possible) packing of multicomponent sphere packs (Assouline et al., 1998; Assouline & Rouault, 1997). Nevertheless, previous studies of textural porosity do suggest a strong link between the PSD and the VSD (Fiès & Bruand, 1990, 1998). Whether this link is strictly linear remains to be investigated. We chose the model by Arya and Paris (1981) for the linear transformation from PSD to VSD of the reference soil, setting the scaling parameter α to 1, which implies no difference between an ideal sphere pack and the reference soil (Arya et al., 1999). The nature of the scaling parameter α has been strongly debated in the literature and its value seems to vary from soil to soil (Arya et al., 1999, 2008; Basile & D'Urso, 1997; Vaz et al., 2005). It is not clear, however, whether this variation is the result of soil structure, soil texture or both because the Arya and Paris (1981) model has mostly been tested on structured soils. Nevertheless, α commonly shows values close to 1 for a variety of soil textures (Vaz et al., 2005), which is why we adopted it here. Several other models for translating particle radii into pore radii have been proposed (Arya & Heitman, 2015; Chan & Govindaraju, 2004; e.g., Haverkamp & Parlange, 1986; Mohammadi & Vanclooster, 2011; Pollacco et al., 2020), which would lead to different results, since the KL divergence automatically depends on the model selected for this purpose. We tested the more recent model by Arya and Heitman (2015) on the same dataset and obtained smaller equivalent pore radii for the reference soil than for the Arya and Paris (1981) model. This increased the KL divergences but had only negligible effects on the relative differences between treatments and sites.

All treatments at Ultuna including the bare fallow treatment showed a bimodal VSD, which is probably the result of the high clay content at this site (37%). Nevertheless, differences in measured ϕ are clearly visible between the treatments with the manure treatment

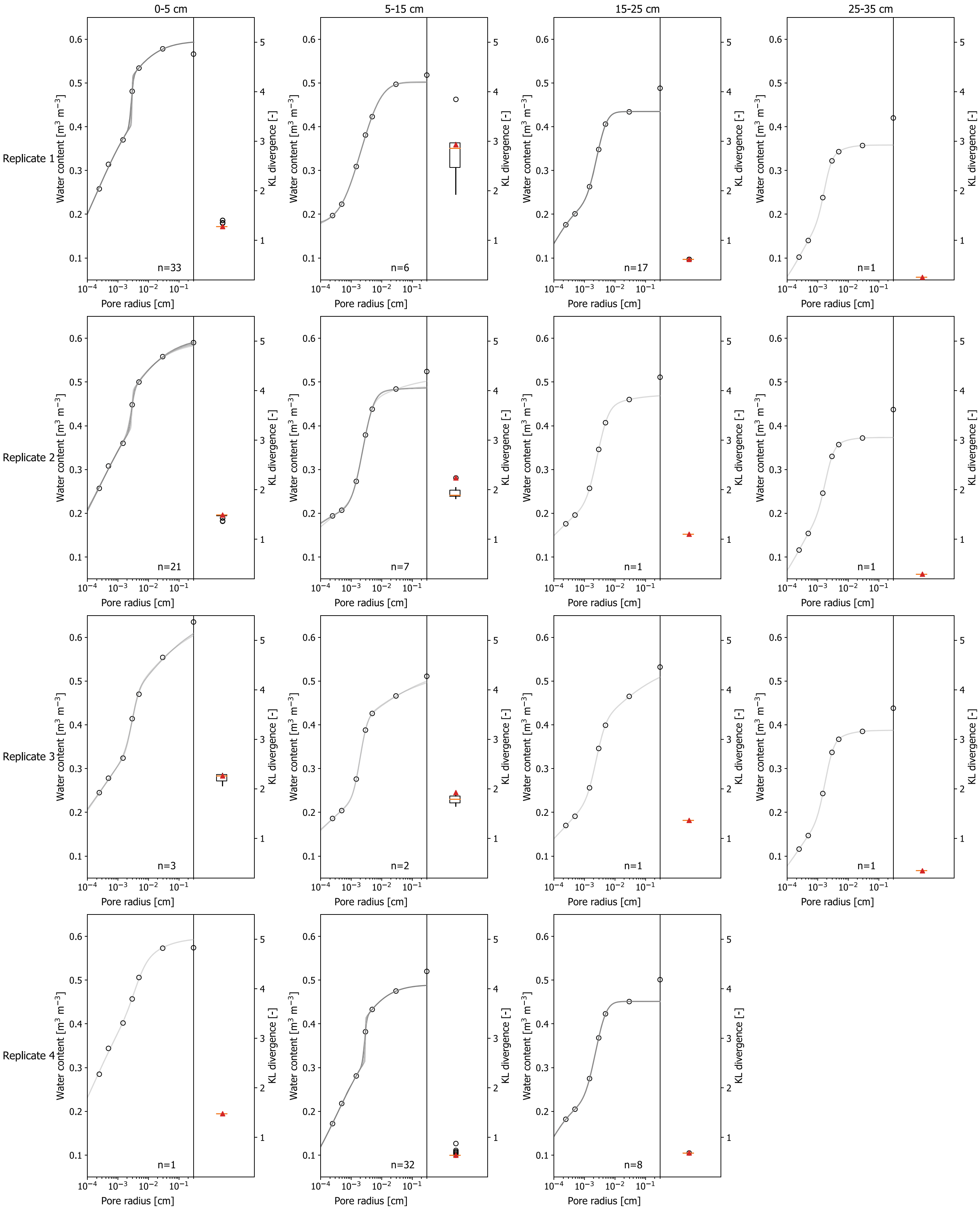

FIGURE 6 Acceptable model fits for each sample (left side of each subplot) with corresponding KL divergences shown as boxplots (right side of each subplot) for the afforested land at Siggebohyttan. For explanation of the individual figure features, see the caption for Figure 4

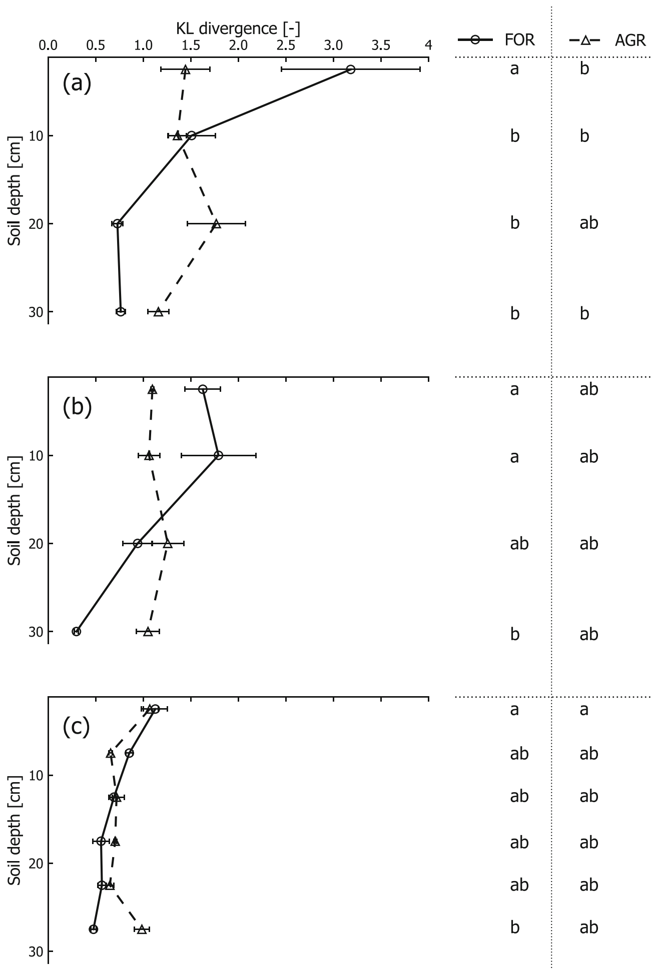

FIGURE 7 The variation of the median KL divergence of the acceptable parameter sets with soil depth in the agricultural (AGR) and afforested land (FOR) at (a) Almnäs, (b) Siggebohyttan, and (c) Vik. Symbols indicate arithmetic means and error bars indicate standard errors for each treatment and depth. Letters on the right side of the figure indicate significant differences ($p < 0.05$) in KL divergence between treatments and depth intervals for each site. Statistical comparisons between sites were not conducted

showing the largest and the bare fallow treatment the smallest value (Table 2). Water retention measurements and X-ray computed tomography analyses of other studies corroborate that the long-term addition of animal manure increases ϕ and changes the relative abundance of pore-size classes with main effects on macroporosity (e.g., Anderson et al., 1990; Naveed et al., 2014; Pagliai & Vignozzi, 1998; Zhang et al., 2021). This increases the heterogeneity as well as the broadness of the VSD (see values of $\sigma_{S,1}$ in Table S2), which explains the larger KL divergence in the manure treatment compared to the other two treatments. Apart from a larger ϕ, it seems that the regular addition of $Ca(NO_3)_2$ fertilizer and the presence of crops did not lead to noticeable differences in the VSD of this treatment compared to the bare fallow as indicated by the similar KL divergences in these two treatments. We found no studies that directly investigated the influence of $Ca(NO_3)_2$ addition on soil structure. While calcium is considered an important driver for micro-aggregation (Pihlap et al., 2021; Totsche et al., 2018), its overall effect on the VSD has been found to be limited even with the addition of far larger amounts than practiced in the long-term field site at Ultuna (Frank et al., 2020; Mamedov et al., 2021). Frank et al. (2020) noted that regular tillage can undermine the effects of liming on the pore space and that fine-textured soils require considerable amounts of lime for effects to be visible. We assume that the amount of calcium added as $Ca(NO_3)_2$ was not sufficient to induce detectable changes in the VSD of the fine-textured soil at Ultuna. Instead, the observed increase in macroporosity compared with the fallow treatment (Table 2) should be the result of crop growth, which creates root biopores and enhances soil faunal activity due to the input of carbon (Meurer, Barron, et al., 2020).

For the dataset reported by Messing et al. (1997), our results show a clear positive correlation (Pearson's $r = 0.374$, $p < 0.001$) between soil organic carbon concentrations and KL divergences across all three sites and both FOR and AGR treatments (Figure 9). The different trends in KL divergence with soil depth between the FOR and AGR treatments can therefore be largely explained by the depth distribution of soil organic matter, which is much more homogeneous in the AGR treatment (Table 2). Regular soil tillage probably contributed to the homogenization of both soil organic matter and KL divergence in the AGR treatments. Soil organic matter is known to be an important driver for soil structural development in the form of aggregation at the micro-scale (Chenu & Cosentino, 2011; Dignac et al., 2017; Vidal et al., 2021; Witzgall et al., 2021). Many studies have shown that soil organic matter has a significant positive effect on total porosity (Jarvis, Forkman, et al., 2017; Johannes et al., 2017; Meurer et al., 2020,b), whereas only a few studies have investigated the effects of SOM on the VSD. However, experiments have found impacts on a wide range of pore diameters, including both smaller matrix pores and larger mesopores (Fukumasu et al., 2022; Kirchmann & Gerzabek, 1999; Meurer, Chenu, et al., 2020; Sekucia et al., 2020; Zhang et al., 2021), which implies an increase in the heterogeneity of the VSD and, therefore, the KL divergence. The strong relationships observed between the model parameters of the PSD and the small pore region of the bimodal model (Figure 8a,b) do suggest that the latter is dominated by textural pore space. However, it seems clear that the small pore region does not consist exclusively of textural pores, since a simpler four-parameter model that we tested based on this assumption did not give good fits to

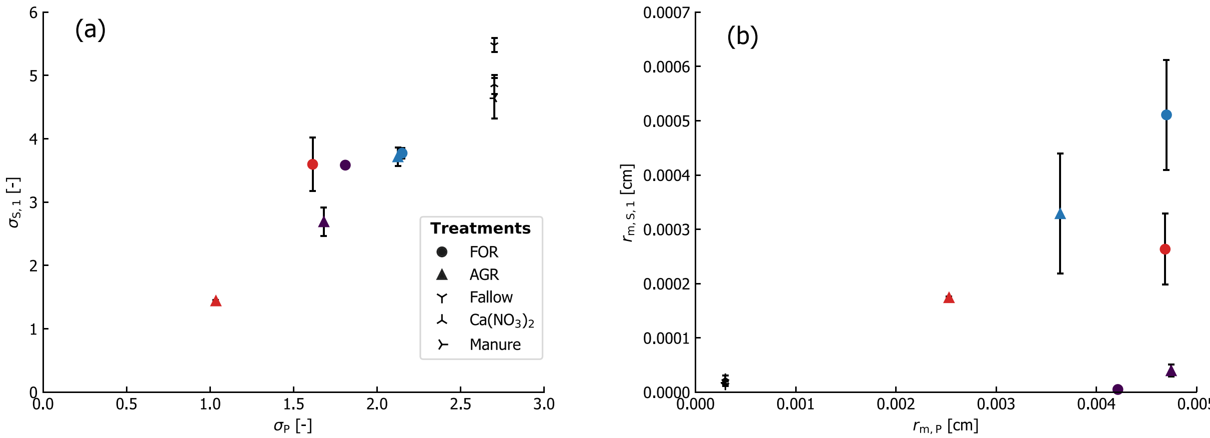

FIGURE 8 Illustration of the physical link between particle-size distribution and pore-size distribution of the small pore region (subscript 1 in Equation (14)) across sites and treatments (FOR = afforested land, AGR = agricultural land). (a) Shows the standard deviation parameter of the PSD (σ_P) against the standard deviation parameter of the small pore region ($\sigma_{S,1}$), and (b) shows the median particle radius ($r_{m,P}$) against the median pore radius of the small pore region ($r_{m,S,1}$). Error bars indicate standard errors and sites are represented by different colours (Almnäs: Dark purple, Siggebohyttan: Red, Vik: Blue, Ultuna: Black symbols)

the data. Thus, our results demonstrate that the small pore region also includes structural pore space, presumably related to aggregation by soil organic matter. Finally, Figure 9 shows that the KL divergences at Ultuna follow a similar trend with soil organic carbon concentrations (both increase in the order fallow < $Ca(NO_3)_2$ < manure), although in this case soil organic carbon is probably mostly acting as a proxy for the impacts of biological activity on macroporosity.

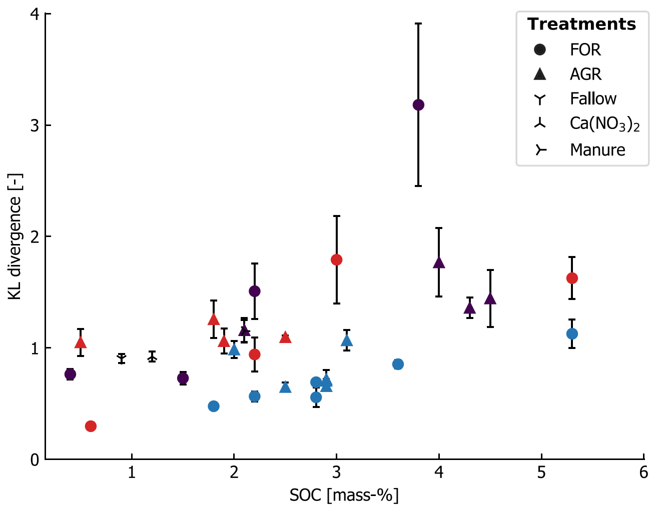

FIGURE 9 Illustration of the relationship between KL divergences and soil organic carbon concentrations (SOC) across sites and treatments (FOR = afforested land, AGR = agricultural land). Error bars indicate standard errors and sites are represented by different colours (Almnäs: Dark purple, Siggebohyttan: Red, Vik: Blue, and Ultuna: Black symbols)

5 | CONCLUSIONS

In this study, we described and demonstrated the applicability of relative entropy, the Kullback–Leibler (KL) divergence, as an index of soil structure. A large KL divergence, which may arise from combinations of a large structural porosity, large median pore size and a wide distribution of pore sizes (i.e., a large standard deviation) is indicative of a well-developed soil structure. We showed that the KL divergence follows expected trends in soil structural development between different treatments and management systems (tree plantations vs. agricultural land; bare fallow vs. $Ca(NO_3)_2$ addition vs. manure addition). The significant correlation found between soil organic carbon concentrations and KL divergences across the tested range of soil textures and management systems underlines this finding. We conclude therefore that the KL divergence may also have the potential to serve as an indicator of soil physical quality in agricultural soils under different management systems. Finally, because only routine soil data are required for this method, we expect it to be particularly useful for assessing the degree of soil structure for existing larger datasets of soil physical and hydraulic properties.

Some uncertainties of the presented method remain regarding the derivation of the VSD of the reference soil, which requires additional experimental efforts that focus on the study of textural porosity. Careful application of the curve fitting procedures is also necessary in order to ensure that the results are not unduly affected by problems related to non-uniqueness and parameter identification. In this respect, the calculated KL divergence values are also very sensitive to the quality of the fit. Further testing is necessary to confirm the general applicability of the method in contrasting soil types. However, from the results presented here, we conclude that relative entropy shows potential as an index of soil structure.

AUTHOR CONTRIBUTIONS

Tobias Klöffel: Formal analysis (equal); methodology (equal); writing – original draft (equal). **Nicholas Jarvis:** Methodology (equal); supervision (equal); writing – review and editing (equal). **Sung Won Yoon:** Methodology (equal). **Jennie Barron:** Project administration (equal); supervision (equal); writing – review and editing (equal). **Daniel Giménez:** Methodology (equal); writing – review and editing (equal).

ACKNOWLEDGEMENTS

We are grateful to David Nimblad Svensson, Johannes Koestel and Ingmar Messing (Department of Soil and Environment, SLU) for supplying the data used in this study and for helpful discussions. We thank Johannes Forkman (Department of Crop Production Ecology, SLU) for discussions on statistical analyses and John Nimmo (U.S.G.S) for helpful discussions on the issue of textural porosity. Finally, we are grateful for the helpful comments by the reviewers.

FUNDING INFORMATION

This work was financially supported by faculty PhD funding from the Swedish University of Agricultural Sciences (SLU).

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work of this paper.

DATA AVAILABILITY STATEMENT

Data available on request due to privacy/ethical restrictions. The Python scripts to calculate the KL divergence from two lognormal distributions or using discrete integration is publicly available in the GitHub repository: https://github.com/tobikloeff.

ORCID

Tobias Klöffel https://orcid.org/0000-0002-5518-376X
Jennie Barron https://orcid.org/0000-0002-3292-3438

REFERENCES

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*(1), 117–121.

Arya, L. M., Bowman, D. C., Thapa, B. B., & Cassel, D. K. (2008). Scaling soil water characteristics of golf course and athletic Field Sands from particle-size distribution. *Soil Science Society of America Journal*, *72*(1), 25–32. https://doi.org/10.2136/sssaj2006.0232

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*(6), 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*(3), 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*(6), 1023–1030. https://doi.org/10.2136/sssaj1981.03615995004500060004x

Assouline, S., & Or, D. (2013). Conceptual and parametric representation of soil hydraulic properties: A review. *Vadose Zone Journal*, *12*(4), 1–20. https://doi.org/10.2136/vzj2013.07.0121

Assouline, S., & Rouault, Y. (1997). Modeling the relationships between particle and pore size distributions in multicomponent sphere packs: Application to the water retention curve. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, *127*(1–3), 201–210. https://doi.org/10.1016/S0927-7757(97)00144-1

Assouline, S., Tessier, D., & Bruand, A. (1998). A conceptual model of the soil water retention curve. *Water Resources Research*, *34*(2), 223–231. https://doi.org/10.1029/97WR03039

Axler, S. (2020). Riemann Integration. In S. Axler (Ed.), *Measure, integration & real analysis* (Vol. 282, pp. 1–12). Springer International Publishing. https://doi.org/10.1007/978-3-030-33143-6_1

Basile, A., & D'Urso, G. (1997). Experimental corrections of simplified methods for predicting water retention curves in clay-loamy soils from particle-size determination. *Soil Technology*, *10*(3), 261–272. https://doi.org/10.1016/S0933-3630(96)00020-7

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, *16*(1), 41–51. https://doi.org/10.1016/0309-1708(93)90028-E

Bishop, C. (2006). *Pattern recognition and machine learning*. Springer-Verlag. https://www.springer.com/gp/book/9780387310732

Bodner, G., Scholl, P., & Kaul, H.-P. (2013). 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

Brutsaert, W. (1966). Probability laws for pore-size distributions. *Soil Science*, *101*(2), 85–92.

Bucka, F. B., Pihlap, E., Kaiser, J., Baumgartl, T., & Kögel-Knabner, I. (2021). A small-scale test for rapid assessment of the soil development potential in post-mining soils. *Soil and Tillage Research*, *211*, 105016. https://doi.org/10.1016/j.still.2021.105016

Chan, T. P., & Govindaraju, R. S. (2004). Estimating soil water retention curve from particle-size distribution data based on Polydisperse sphere systems. *Vadose Zone Journal*, *3*(4), 1443–1454. https://doi.org/10.2113/3.4.1443

Chenu, C., & Cosentino, D. (2011). Microbial regulation of soil structural dynamics. In K. Ritz & I. Young (Eds.), *The*

architecture and biology of soils: Life in inner space (pp. 37–70). CABI. https://doi.org/10.1079/9781845935320.0037
Childs, E. C. (1969). Chapter 6: The soil pore space. In *Introduction to the Physical Basis of Soil Water Phenomena* (pp. 92-96). *John Wiley & Sons.*
Crawford, J. W., Matsui, N., & Young, I. M. (1995). The relation between the moisture-release curve and the structure of soil. *European Journal of Soil Science*, *46*(3), 369–375. https://doi.org/10.1111/j.1365-2389.1995.tb01333.x
Crisp, D. J., & Williams, R. (1971). Direct measurement of pore-size distribution on artificial and natural deposits and prediction of pore space accessible to interstitial organisms. *Marine Biology*, *10*(3), 214–226. https://doi.org/10.1007/BF00352810
Dexter, A. R. (1977). A statistical measure of the structure of tilled soil. *Journal of Agricultural Engineering Research*, *22*(1), 101–104. https://doi.org/10.1016/0021-8634(77)90099-3
Dexter, A. R. (2004). Soil physical quality. *Geoderma*, *120*(3–4), 201–214. https://doi.org/10.1016/j.geoderma.2003.09.004
Dexter, A. R., Czyż, E. A., Richard, G., & Reszkowska, A. (2008). A user-friendly water retention function that takes account of the textural and structural pore spaces in soil. *Geoderma*, *143*(3–4), 243–253. https://doi.org/10.1016/j.geoderma.2007.11.010
Dignac, M.-F., Derrien, D., Barré, P., Barot, S., Cécillon, L., Chenu, C., Chevallier, T., Freschet, G. T., Garnier, P., Guenet, B., Hedde, M., Klumpp, K., Lashermes, G., Maron, P.-A., Nunan, N., Roumet, C., & Basile-Doelsch, I. (2017). Increasing soil carbon storage: Mechanisms, effects of agricultural practices and proxies. A review. *Agronomy for Sustainable Development*, *37*(2), 14. https://doi.org/10.1007/s13593-017-0421-2
Dominati, E., Patterson, M., & Mackay, A. (2010). A framework for classifying and quantifying the natural capital and ecosystem services of soils. *Ecological Economics*, *69*(9), 1858–1868. https://doi.org/10.1016/j.ecolecon.2010.05.002
Durner, W. (1994). Hydraulic conductivity estimation for soils with heterogeneous pore structure. *Water Resources Research*, *30*(2), 211–223. https://doi.org/10.1029/93WR02676
El-Baz, A., Nayak, T.K. (2004). Efficiency of composite sampling for estimating a lognormal distribution. *Environmental and Ecological Statistics*, *11*(3), 283–294. https://doi.org/10.1023/B:EEST.0000038016.20656.21
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
FAO. (1989). *Soil map of the world* (Technical Paper No. 20). ISRIC.
Farr, R. S., & Groot, R. D. (2009). Close packing density of polydisperse hard spheres. *The Journal of Chemical Physics*, *131*(24), 244104. https://doi.org/10.1063/1.3276799
Fatichi, S., Or, D., Walko, R., Vereecken, H., Young, M. H., Ghezzehei, T. A., Hengl, T., Kollet, S., Agam, N., & Avissar, R. (2020). Soil structure is an important omission in earth system models. *Nature Communications*, *11*(1), 522. https://doi.org/10.1038/s41467-020-14411-z
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. (1990). Textural porosity analysis of a silty clay soil using pore volume balance estimation, mercury porosimetry and quantified backscattered electron scanning image (BESI). *Geoderma*, *47*(3–4), 209–219. https://doi.org/10.1016/0016-7061(90)90029-9
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*(4), 557–567. https://doi.org/10.1046/j.1365-2389.1998.4940557.x
Fiès, J. C., & Stengel, P. (1981). Densite texturale de sols naturels. I. - Methode de mesure. *Agronomie*, *1*(8), 651–658.
Fiès, J. C., Stengel, P., Bourlet, M., Horoyan, J., & Jeandet, C. (1981). Densité texturale de sols naturels II. - Eléments d'interprétation. *Agronomie*, *1*(8), 659–666.
Frank, T., Zimmermann, I., & Horn, R. (2020). Lime application in marshlands of northern Germany—Influence of liming on the physicochemical and hydraulic properties of clayey soils. *Soil and Tillage Research*, *204*, 104730. https://doi.org/10.1016/j.still.2020.104730
Fredlund, M. D., Fredlund, D. G., & Wilson, G. W. (2000). An equation to represent grain-size distribution. *Canadian Geotechnical Journal*, *37*, 817–827. 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*(1), e13212. https://doi.org/10.1111/ejss.13212
Gardner, W. R. (1956). Representation of soil aggregate-size distribution by a logarithmic-Normal distribution. *Soil Science Society of America Journal*, *20*(2), 151–153. https://doi.org/10.2136/sssaj1956.03615995002000020003x
Gil, M. (2011). *On Renyi divergence measures for continuous alphabet sources* [thesis]. https://qspace.library.queensu.ca/handle/1974/6680
Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning* (illustrated edition). The MIT Press.
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*, 758–764. https://doi.org/10.2136/sssaj1979.03615995004300040028x
Haverkamp, R., & Parlange, J.-Y. (1986). Predicting the water-retention curve from particle-size distribution: 1. Sandy soils without organic matter. *Soil Science*, *142*(6), 325–339.
Hou, Z., & Rubin, Y. (2005). On minimum relative entropy concepts and prior compatibility issues in vadose zone inverse and forward modeling. *Water Resources Research*, *41*, W1245. https://doi.org/10.1029/2005WR004082
Hwang, S. I. (2004). Effect of texture on the performance of soil particle-size distribution models. *Geoderma*, *123*(3), 363–371. https://doi.org/10.1016/j.geoderma.2004.03.003
Hwang, S. I., & Powers, S. E. (2003). Lognormal distribution model for estimating soil water retention curves for sandy soils. *Soil Science*, *168*(3), 156–166. https://doi.org/10.1097/01.ss.0000058888.60072.e3
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*(1–4), 325–334. https://doi.org/10.1016/j.jhydrol.2005.09.005
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

Jarvis, N., Larsbo, M., & Koestel, J. (2017). Connectivity and percolation of structural pore networks in a cultivated silt loam soil quantified by X-ray tomography. *Geoderma*, *287*, 71–79. https://doi.org/10.1016/j.geoderma.2016.06.026

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

Johannes, A., Matter, A., Schulin, R., Weisskopf, P., Baveye, P. C., & Boivin, P. (2017). Optimal organic carbon values for soil structure quality of arable soils. Does clay content matter? *Geoderma*, *302*, 14–21. https://doi.org/10.1016/j.geoderma.2017.04.021

Kim, J., Ivanov, V. Y., & Fatichi, S. (2016). Environmental stochasticity controls soil erosion variability. *Scientific Reports*, *6*, 22065. https://doi.org/10.1038/srep22065

Kirchmann, H., & Gerzabek, M. H. (1999). Relationship between soil organic matter and micropores in a long-term experiment at Ultuna, Sweden. *Journal of Plant Nutrition and Soil Science*, *162*, 493–498.

Kirchmann, H., Persson, J., & Carlgren, K. (1994). *The Ultuna long-term soil organic matter experiment, 1956-1991* (Department of Soil Sciences, reports and dissertations no. 17). Swedish University of Agricultural Sciences.

Kosugi, K. (1994). Three-parameter lognormal distribution model for soil water retention. *Water Resources Research*, *30*(4), 891–901. https://doi.org/10.1029/93WR02931

Kosugi, K. (1996). Lognormal distribution model for unsaturated soil hydraulic properties. *Water Resources Research*, *32*(9), 2697–2703. https://doi.org/10.1029/96WR01776

Kreiselmeier, J., Chandrasekhar, P., Weninger, T., Schwen, A., Julich, S., Feger, K.-H., & Schwärzel, K. (2019). Quantification of soil pore dynamics during a winter wheat cropping cycle under different tillage regimes. *Soil and Tillage Research*, *192*, 222–232. https://doi.org/10.1016/j.still.2019.05.014

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, *22*(1), 79–86.

Lenth, R. (2020). *Emmeans: Estimated marginal means, aka least-squares means* [manual]. https://CRAN.R-project.org/package=emmeans

Lin, H. (2011). Three principles of soil change and Pedogenesis in time and space. *Soil Science Society of America Journal*, *75*(6), 2049–2070. https://doi.org/10.2136/sssaj2011.0130

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

Luo, L., Lin, H., & Halleck, P. (2008). Quantifying soil structure and preferential flow in intact soil using X-ray computed tomography. *Soil Science Society of America Journal*, *72*(4), 1058–1069. https://doi.org/10.2136/sssaj2007.0179

MacKay, D. J. C. (2002). *Information theory*. Cambridge University Press.

Mamedov, A. I., Fujimaki, H., Tsunekawa, A., Tsubo, M., & Levy, G. J. (2021). Structure stability of acidic Luvisols: Effects of tillage type and exogenous additives. *Soil and Tillage Research*, *206*, 104832. https://doi.org/10.1016/j.still.2020.104832

Messing, I., Alriksson, A., & Johansson, W. (1997). Soil physical properties of afforested and arable land. *Soil Use and Management*, *13*(4), 209–217. https://doi.org/10.1111/j.1475-2743.1997.tb00588.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. (2020). A framework for modelling soil structure dynamics induced by biological activity. *Global Change Biology*, *26*(10), 5382–5403. https://doi.org/10.1111/gcb.15289

Meurer, K. H. E., Chenu, C., Coucheney, E., Herrmann, A. M., Keller, T., Kätterer, T., Nimblad Svensson, D., & Jarvis, N. (2020). Modelling dynamic interactions between soil structure and the storage and turnover of soil organic matter. *Biogeosciences*, *17*(20), 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*(2), 594–602. https://doi.org/10.2136/vzj2010.0080

Mueller, L., Shepherd, G., Schindler, U., Ball, B. C., Munkholm, L. J., Hennings, V., Smolentseva, E., Rukhovic, O., Lukin, S., & Hu, C. (2013). Evaluation of soil structure in the framework of an overall soil quality rating. *Soil and Tillage Research*, *127*, 74–84. https://doi.org/10.1016/j.still.2012.03.002

Naveed, M., Moldrup, P., Vogel, H.-J., 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

Nimmo, J. R. (1997). Modeling structural influences on soil water retention. *Soil Science Society of America Journal*, *61*(3), 712–719. https://doi.org/10.2136/sssaj1997.03615995006100030002x

Nimmo, J. R. (2013). Porosity and pore size distribution. In *Reference Module in Earth Systems and Environmental Sciences*. Elsevier. https://doi.org/10.1016/B978-0-12-409548-9.05265-9

Nimmo, J. R., Herkelrath, W. N., & Laguna Luna, A. M. (2007). Physically based estimation of soil water retention from textural data: General framework, new models, and streamlined existing models. *Vadose Zone Journal*, *6*(4), 766–773. https://doi.org/10.2136/vzj2007.0019

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

Pagliai, M., & Vignozzi, N. (1998). Use of manures for soil improvement. In *Handbook of soil conditioners: Substances that enhance the physical properties of soil: Substances that enhance the physical properties of soil* (pp. 119–139). CRC Press.

Pihlap, E., Steffens, M., & Kögel-Knabner, I. (2021). Initial soil aggregate formation and stabilisation in soils developed from calcareous loess. *Geoderma*, *385*, 114854. https://doi.org/10.1016/j.geoderma.2020.114854

Pohlert, T. (2014). *The pairwise multiple comparison of mean ranks package (PMCMR)* [manual]. https://CRAN.R-project.org/package=PMCMR

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

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*(6), 2725–2737. https://doi.org/10.5194/hess-21-2725-2017

R Core Team. (2019). *R: A language and environment for statistical computing* [manual]. https://www.R-project.org/

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

Reynolds, W. D. (2017). Use of bimodal hydraulic property relationships to characterize soil physical quality. *Geoderma*, *294*, 38–49. https://doi.org/10.1016/j.geoderma.2017.01.035

Reynolds, W. D., Drury, C. F., Tan, C. S., Fox, C. A., & Yang, X. M. (2009). Use of indicators and pore volume-function characteristics to quantify soil physical quality. *Geoderma*, *152*(3–4), 252–263. https://doi.org/10.1016/j.geoderma.2009.06.009

Robinson, D. A., Lebron, I., & Vereecken, H. (2009). On the definition of the natural Capital of Soils: A framework for description, evaluation, and monitoring. *Soil Science Society of America Journal*, *73*(6), 1904–1911. https://doi.org/10.2136/sssaj2008.0332

Ross, P. J., & Smettem, K. R. J. (1993). Describing soil hydraulic properties with sums of simple functions. *Soil Science Society of America Journal*, *57*(1), 26–29. https://doi.org/10.2136/sssaj1993.03615995005700010006x

Sekucia, F., Dlapa, P., Kollár, J., Cerdá, A., Hrabovský, A., & Svobodová, L. (2020). Land-use impact on porosity and water retention of soils rich in rock fragments. *Catena*, *195*, 104807. https://doi.org/10.1016/j.catena.2020.104807

Shen, C., Liu, S., Xu, S., & Wang, L. (2019). Rapid estimation of maximum and minimum void ratios of granular soils. *Acta Geotechnica*, *14*(4), 991–1001. https://doi.org/10.1007/s11440-018-0714-x

Sillers, W. S., Fredlund, D. G., & Zakerzaheh, N. (2001). Mathematical attributes of some soil–water characteristic curve models. *Geotechnical and Geological Engineering*, *19*(3/4), 243–283. https://doi.org/10.1023/A:1013109728218

Svensson, D. N. (2020). *Influence of cropping and fertilization on soil pore characteristics in a long-term field study* [Master's thesis, Swedish University of Agricultural Sciences]. https://stud.epsilon.slu.se/15593/

Tarquis, A. M., Bird, N. R. A., Whitmore, A. P., Cartagena, M. C., & Pachepsky, Y. (2008). Multiscale entropy-based analysis of soil transect data. *Vadose Zone Journal*, *7*(2), 563–569. https://doi.org/10.2136/vzj2007.0039

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*(1), 104–136. https://doi.org/10.1002/jpln.201600451

Vaz, C. M. P., de Freitas Iossi, M., de Mendonça Naime, J., Macedo, Á., Reichert, J. M., Reinert, D. J., & Cooper, M. (2005). Validation of the Arya and Paris water retention model for Brazilian soils. *Soil Science Society of America Journal*, *69*(3), 577–583. https://doi.org/10.2136/sssaj2004.0104

Vidal, A., Klöffel, T., Guigue, J., Angst, G., Steffens, M., Hoeschen, C., & Mueller, C. W. (2021). Visualizing the transfer of organic matter from decaying plant residues to soil mineral surfaces controlled by microorganisms. *Soil Biology and Biochemistry*, *160*, 108347. https://doi.org/10.1016/j.soilbio.2021.108347

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … van Mulbregt, P. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Vogel, H.-J., Weller, U., & Schlüter, S. (2010). Quantification of soil structure based on Minkowski functions. *Computers & Geosciences*, *36*(10), 1236–1245. https://doi.org/10.1016/j.cageo.2010.03.007

Witzgall, K., Vidal, A., Schubert, D. I., Höschen, C., Schweizer, S. A., Buegger, F., Pouteau, V., Chenu, C., & Mueller, C. W. (2021). Particulate organic matter as a functional soil component for persistent soil organic carbon. *Nature Communications*, *12*(1), 4115. https://doi.org/10.1038/s41467-021-24192-8

Yoon, S. W. (2009). *A measure of soil structure derived from water retention properties: A Kullback-Leibler distance approach* [dissertation]. Rutgers University.

Yoon, S. W., & Giménez, D. (2012). Entropy characterization of soil pore systems derived from soil-water retention curves. *Soil Science*, *177*(6), 361–368. https://doi.org/10.1097/SS.0b013e318256ba1c

Young, I. M., Crawford, J. W., & Rappoldt, C. (2001). New methods and models for characterising structural heterogeneity of soil. *Soil and Tillage Research*, *61*(1–2), 33–45. https://doi.org/10.1016/S0167-1987(01)00188-X

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

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: 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*(3), e13254. https://doi.org/10.1111/ejss.13254

APPENDIX A: Derivation of the KL divergence for two lognormal distributions

The KL divergence has been derived for two lognormal distributions before (e.g., El-Baz et al., 2004; Gil, 2011). However, for the specific case of two VSDs, Equation (12) may not be trivial from these derivations (Yoon, 2009). Hence, we show step-by-step how Equation (12) is developed from Equations (5) and (11) in the following.

We start with Equation (2) and substitute $p(x)$ with Equation (5) and $q(x)$ with Equation (11). Since both are lognormal distributions, the lower integral limit is adapted to 0, giving:

$$D_{KL}\left(f_S\middle\|f_R\right) = \int_0^\infty f_S(r)\log\left(\frac{\frac{(\theta_s-\theta_r)_S}{r\sigma_S\sqrt{2\pi}}\exp\left\{-\frac{(\log r-\log r_{m,S})^2}{2\sigma_S^2}\right\}}{\frac{(\theta_s-\theta_r)_R}{r\sigma_R\sqrt{2\pi}}\exp\left\{-\frac{(\log r-\log r_{m,R})^2}{2\sigma_R^2}\right\}}\right)dr \tag{A1}$$

which can be written as follows:

$$D_{KL}\left(f_S\middle\|f_R\right) = \log\frac{(\theta_s-\theta_r)_S\,\sigma_R}{(\theta_s-\theta_r)_R\,\sigma_S}\int_0^\infty f_S(r)dr + \int_0^\infty f_S(r)\left[-\frac{(\log r-\log r_{m,S})^2}{2\sigma_S^2} + \frac{(\log r-\log r_{m,R})^2}{2\sigma_R^2}\right]dr \tag{A2}$$

The first integral in Equation (A2) can be substituted with $(\theta_s - \theta_r)_S$ and the equation transformed to

$$D_{KL}\left(f_S\middle\|f_R\right) = (\theta_s-\theta_r)_S\log\frac{(\theta_s-\theta_r)_S\,\sigma_R}{(\theta_s-\theta_r)_R\,\sigma_S} + \frac{1}{2\sigma_S^2}\left(-\int_0^\infty f_S(r)(\log r-\log r_{m,S})^2dr\right) + \frac{1}{2\sigma_R^2}\left(-\int_0^\infty f_S(r)(\log r-\log r_{m,R})^2dr\right) \tag{A3}$$

The two integrals in Equation (A3) can be solved using the following relationship:

$$\mathrm{E}\left[X^2\right] = \sigma^2 - \mathrm{E}[X]^2 \tag{A4}$$

where E denotes the expected value and X is a random variable and where $E[X]$ can be expressed as follows:

$$\mathrm{E}[X] = \int_{-\infty}^\infty x f(x)\,dx \tag{A5}$$

and $E[X^2]$ as:

$$\mathrm{E}\left[X^2\right] = \int_{-\infty}^\infty x^2 f(x)\,dx \tag{A6}$$

where $f(x)$ denotes a probability density function of X.

Applying the relationships Equation (A4) through (A6) to Equation (A3) and after further simplification we obtain the following:

$$D_{KL}\left(f_S\middle\|f_R\right) = (\theta_s-\theta_r)_S\log\frac{(\theta_s-\theta_r)_S\,\sigma_R}{(\theta_s-\theta_r)_R\,\sigma_S} + \frac{(\theta_s-\theta_r)_S}{2} + (\theta_s-\theta_r)_S\left[\frac{\sigma_S^2+(\log r_{m,S}-\log r_{m,R})^2}{2\sigma_R^2}\right] \tag{A7}$$

Finally, the term $(\theta_s - \theta_r)_S$ can be factorized to yield Equation (12):

$$D_{KL}\left(f_S\middle\|f_R\right) = (\theta_s-\theta_r)_S\left(\log\frac{(\theta_s-\theta_r)_S\,\sigma_R}{(\theta_s-\theta_r)_R\,\sigma_s} - \frac{1}{2} + \left[\frac{\sigma_s^2+(\log r_{m,S}-\log r_{m,R})^2}{2\sigma_R^2}\right]\right) \tag{A8}$$