Quantification of the pore size distribution of soils: assessment of existing software using tomographic and synthetic 3D images

A.N. Houston, W. Otten, R. Falconer, O. Monga,
P.C. Baveye and S.M. Hapca

The published article is available from doi: https://doi.org/10.1016/j.geoderma.2017.03.025

Quantification of the pore size distribution of soils: Assessment of existing software using tomographic and synthetic 3D images.

A. N. Houston[a], W. Otten[a,b], R. Falconer[a,c], O. Monga[d], P.C.Baveye[e], S.M. Hapca[a,f*]

[a] School of Science, Engineering and Technology, Abertay University, 40 Bell Street, Dundee DD1 1HG, UK.

[b] current address: School of Water, Energy and Environment, Cranfield University, Cranfield MK43 0AL, UK.

[c] current address: School of Arts, Media and Computer Games , Abertay University, 40 Bell Street, Dundee DD1 1HG, UK.

[d] UMMISCO-Cameroun, UMI 209 UMMISCO, University of Yaoundé, IRD, University of Paris 6, F-93143 Bondy Cedex, France.

[e] UMR ECOSYS, AgroParisTech, Université Paris-Saclay, Avenue Lucien Brétignières, Thiverval-Grignon 78850, France.

[f] current address: Dundee Epidemiology and Biostatistics Unit, School of Medicine, Dundee University, Kirsty Semple Way, Dundee DD2 4BF, UK.

[*] Corresponding author. E-mail: s.z.hapca@dundee.ac.uk

Abstract:

The pore size distribution (PSD) of the void space is widely used to predict a range of processes in soils. Recent advances in X-ray computed tomography (CT) now afford novel ways to obtain exact data on pore geometry, which has stimulated the development of algorithms to estimate the pore size distribution from 3D data sets. To date there is however no clear consensus on how PSDs should be estimated, and in what form PSDs are best presented. In this article, we first review the theoretical principles shared by the various methods for PSD estimation. Then we select methods that are widely adopted in soil science and geoscience, and we use a robust statistical method to compare their application to synthetic image samples, for which analytical solutions of PSDs are available, and X-ray CT images of soil samples selected from different treatments to obtain wide ranging PSDs. Results indicate that, when applied to the synthetic images, all methods presenting PSDs as pore volume per class size (i.e., Avizo, CTAnalyser, BoneJ, Quantim4, and DTM), perform well. Among them, the methods based on Maximum Inscribed Balls (Bone J, CTAnalyser, Quantim4) also produce similar PSDs for the soil samples, whereas the Delaunay Triangulation Method (DTM) produces larger estimates of the pore volume occupied by small pores, and Avizo yields larger estimates of the pore volume occupied by large pores. By contrast, the methods that calculate PSDs as object population fraction per volume class (Avizo, 3DMA, DFS-FIJI) perform inconsistently on the synthetic images and do not appear well suited to handle the more complex geometries of soils. It is anticipated that the extensive evaluation of method performance carried out in this study, together with the recommendations reached, will be useful to the porous media community to make more informed choices relative to suitable PSD estimation methods, and will help improve current practice, which is often ad hoc and heuristic.

Keywords: porous media, soil, pore size distribution, computed tomography, X-ray.

1. Introduction

In the 1930s and 40s, soil physicists like Haines (1930) and Childs (1940) came to acknowledge that the size distribution of soil particles, routinely measured since the 18th century (Baveye, 2013), provided very little useful information concerning the retention of water and its transport in soils. This realization led to a shift of emphasis from soil particles to the "water-occupied void space [...,] which largely determines the gross physical properties of soils" (Childs and Collis-George, 1948). These authors suggested that soil voids, or "pores", could be linked to straight capillaries of varying diameters, and that their size distribution would provide the type of direct quantitative information needed to describe the functioning of soils.

This perspective has since become one of the hallmarks of soil physics, and it is adopted in most soil physics textbooks to explain the principles that govern the retention of water in soils and its movement. Thus a significant body of research has been devoted to the use of the pore size distribution (PSD) to predict a wide range of processes of interest, such as gas diffusion, water retention and flow, mechanical resistence, carbon dynamics, microbial colonization, and root penetration (Monga *et al.,* 2009; Pajor *et al*., 2010; Kravchenko *et al.*, 2011a; Falconer *et al.,* 2012; Schmidt *et al.*, 2012; Cazelles *et al.,* 2013; Juarez *et al*., 2013; Zaffar and Lu 2015), as well as to assess the effect of different management practices and degradation processes on soil productivity (Kravchenko *et al.* 2011b, Dal Ferro *et al.*, 2012; Muñoz-Ortega *et al.*, 2014; Naveed *et al.*, 2014a; Rab *et al.*, 2014).

In parallel with the application of the PSD to predict the impact on soil processes, methods to measure the PSD have been evolving, an endeavor that is greatly complicated by the extreme heterogeneity of soils, and in particular by the presence of a wide range of pore sizes and morphologies. Over the years, various techniques have been proposed to evaluate the PSD, based alternatively on the analysis of moisture retention curves or nitrogen adsorption isotherms, or on mercury intrusion porosimetry (Echeverría *et al.*, 1999; Filimonova, Hajnos *et al.*, 2006; Dexter *et al.*, 2008; Dal Ferro *et al.*, 2012). However, each of these methods still suffers from a number of limitations. A common one relates to the fact that the resulting pore size distribution is unavoidably influenced by the connectivity of pores, between the inner portion of samples and their periphery. Furthermore, none of the available techniques can detect isolated pores, which, as a result of the dynamic nature of soil structure, may become reconnected over time. The analysis of N_2 adsorption method is suitable only for small pores less than 0.1 µm in diameter, whereas the determination of the PSD based on the moisture retention curve runs into difficulties in swelling soils, because of pore drainage and shrinkage (Zong *et al.*, 2014).

The major technological advances in non-destructive imaging techniques that occured in the last decade, in particular the commercialization of affordable bench-top X-ray Computed Tomography (X-ray CT) systems, have changed dramatically the way we look at the internal geometry of soil voids (Ketcham and Carlson, 2001; Wildenschild *et al.*, 2002; Kaestner *et al*., 2008; Taina *et al.*, 2008; Wildenschild and Sheppard, 2013). Especially since the development of efficient, non-operator-dependent algorithms to segment the grayscale images provided by CT scanners (Sheppard *et al.*, 2004; Iassonov *et al.*, 2009; Baveye *et al.*, 2010; Schlüter *et al.*, 2010; Hapca *et al.*, 2013; Houston *et al.*, 2013a; Schlüter *et al.*, 2014), it is now possible to get a reliable perspective on how intricate and convoluted the geometry of soil voids is, down to submicron scales.

These past few years, various algorithms have been proposed to extract PSDs from 3D CT images. Unlike with other soil characteristics, e.g., porosity and specific surface area, for which there is a clear consensus over the estimation approach, there is no general agreement, nor a clear sense of direction regarding an appropriate method for estimating PSDs. Several algorithms have been proposed (Table 1), each of which has limitations in terms of pore space model representation. In many cases, authors developed software to address specific situations. It is unclear if these developments were driven by a lack of familiarity with existing methods, by specific computational or programming language constraints, or by authors seeking further improvement of existing methods. Many of these available methods do however share common algorithms (Table 1), which raises the question of whether generalisations can be made. A few methods have reached the stage of user-friendly software that is either commercially (e.g., Avizo) or freely available (e.g., ImageJ, Quantim4). Even though these various methods often share the same theoretical basis, specific requirements associated with their application in various disciplines, like hydrology or ecology, have led to PSDs being reported in different ways, either as pore volume and surface distribution per class size, or as body and throat population distribution per class size. Conceptually, this poses no real problem, as indeed distinct formulations may be more appropriate in particular cases than in others, but it has made it difficult to compare the performance of the different algorithms and to determine their limitations.

In this general context, the objective of this study is to review existing PSD estimation methods from both a theoretical and practical perspective, and to compare their performance on a selection of synthetic 3D images as well as X-ray CT images of soils of different types. The computer packages selected for this comparison have all been used in the past to determine the PSD of soils, represent distinct types of algoritms (see details below), and are all readily available. They include, respectively, the commercially licenced programs ***Avizo*** (FEI Visualization Sciences Group) and ***CTAnalyser*** (Skyscan-Bruker), freely available ImageJ plugins BoneJ and Skeletonize3D (Arganda-Carreras et al. 2008), used in conjunction with the “Exact Signed Euclidean Distance Transform” (Borgefors, 1986), and hereafter referred to as ***DFS-FIJI***, the 2005 open-source release of ***3DMA*** (Lindquist *et al.*, 2000), the open-source library ***Quantim4*** (Vogel, 1997; Vogel and Roth, 2001), and the program ***DTM*** developed by Monga *et al.* (2007, 2009) based on Delaunay triangulation.

2 Theoretical approaches to PSD estimation and applications

Despite the plethora of methods that have been developed, some common steps and methods can be identified. The common steps consist of first identifying objects within the image, then estimating a size measure per object, and finally forming a distribution from these measures. In the case of a natural porous medium such as soil, the first of these steps can be made difficult by the occurrence of tortuous interconnected pore clusters. Such clusters are considered to be composed of pore bodies that connect with each other and each such connection may be described as a pore throat (Lindquist and Venkatarangan, 1999). Much effort has been invested during recent decades into automatic methods for identifying pore bodies and throats within digital images. All methods make use of a dichotomous image consisting only of pore object versus solid background. The distance transform (Borgefors, 1986) is embodied in many approaches, since the resulting Distance Map (DM) image has numerous uses. It transforms a classified image (e.g., pore versus solid) into a DM image whose elements are assigned a value representing their distance from the nearest pore-solid interface. Local maxima of the distance transform define points that can be used to extract the medial axis of objects, and also offers a means of accelerating search procedures on the object space. The tools of mathematical morphology (Serra, 1982) also appear within several approaches, as a means of extracting the discrete skeleton (a homologue of the medial axis) as well as other transformations of pore objects. The main techniques for identifying throats and bodies within segmented images include **medial axis extraction**, **maximum inscribed balls, morphological opening, and object separation by watersheds** (see Table 1). A brief description of these techniques is presented in the following sections.

2.1 Medial axis

The medial axis, first proposed by Blum (1973) as an image analysis tool for object shape recognition, has been intensively used for the purpose of pore space modelling. It is defined as the topological skeleton running through the middle of pore channels. Several approaches for medial axis extraction have been proposed, including skeletonization by morphological thinning or burning algorithms, methods based on distance transform and Voronoi tessellation methods.

The morphological thinning approach operates directly on the binary image, resulting in a discrete image description of the pore space skeleton (Baldwin *et al.*, 1996). The process is based on iterative application of morphological erosion operations, which must be constrained and ordered according to a local topological structure within the image (Lee *et al.*,1994). The iterative application leads to the pore space skeleton, then a skeleton distance function is defined as the Euclidean distance from each skeleton pixel to the nearest solid pixel. However, despite the use of constraints, there is no guarantee of a uniquely determined result, especially for pore objects that are asymmetric with respect to the skeletal axis. Analogous to the thining method is **the *pore space burning algorithm*** (Linquist et al., 1996) which can be described as a fire that starts at the pore boundary and spreads with uniform speed burning everything in its path until the different wavefronts eventually meet in the middle. The set of all points where the fire directionally extinguishes itself provides the skeleton of the medial axis. A size measure is given by the time at which the fire reaches any unburned point, known as the burn number.

Another approach to medial axis extraction relies on the use of the distance transform to detect ***ridges (local maxima) in the distance map*** image via analysis of zero-crossing points in its spatial gradient (Siddiqi and Pizer, 2008). Once the location of the medial axis points has been determined, accurate geometric description (i.e. including surface orientation) of the medial axis can be obtained using the structure tensor (Heyden and Kahl, 2011). This is a covariance matrix formed from weighted combinations of gradient vectors in the local neighbourhood of a point. The eigensystem of this covariance matrix reveals local anisotropy in object structure and hence can be used to infer dimensionality (point, line or plane). A disadvantage of this approach is the computational cost: A large number of covariance matrices must be constructed and their eigensystems determined.

Voronoi tessellation has also been proposed for medial axis extraction. It consists of partitionning the pore space into 3D Voronoi regions based on seed points placed on the boundary of the pore objects (Delerue *et al.*, 1999; Delerue and Perrier, 2002). The medial axis can then be extracted from the the subset of the Voronoi facets located inside the pore surface and further filtering according to some angle criteria. A size measure for these features can be determined from a distance transform image (computed separately) or by explicit search. If material exhibits a resolved granular structure with well defined pore objects, this approach provides a good approximation of the medial axis. In general, however, the method is highly unstable with respect to small details of pore shapes.

Therefore, for pore objects that are irregular and complex in shape, an alternative is to use Delaunay triangulation to decompose the boundary of pore objects into 3D surface elements (Monga et al, 2007, 2009). Voronoi regions are then produced from these surface elements and filtering is applied as before to approximate the medial axis. For a precise description of network structure, the decomposition into surface elements may have to be very detailed which leads to extreme computational cost. In practice a balance between accuracy and smoothness is achieved by locally adapting the surface tessellation.

2.2 Maximum inscribed balls (MIB)

This technique finds the largest inscribed spheres centred on each voxel of the pore space that just touches the pore surface. Those that are fully overlapped by larger spheres (engulfed) are removed; the remaining spheres are called maximal balls and cover fully the pore space. Within the pore-ball description, balls that touch or overlap are considered linked to one another by pore channels, hence a graph description consisting of nodes (balls representing pore space) and edges (the channels linking pore space) may be extracted (Silin and Patzek, 2006). Finding the minimal set of maximum-sized balls that accurately describe pore space, requires a search procedure to locate all engulfed balls and then eliminate them from the pore-ball description. This is straightforward in the case of a “simply engulfed” ball but challenging in the case of “compound engulfment”. The combinatoric nature of this search problem means that the algorithm employed must be considered carefully in relation to problem size and computational capacity. As a result some implementations of the MIB procedure make use of medial axis function as a support to fit the inscribed spheres, reducing in this way considerably the search space.

2.3 Morphological opening

This algorithm iterates over increasing level thresholds on the distance map of the pore space, constructing both an "opening map" image and also a mask image that guides subsequent iterative construction (Vogel, 1997). Within each iteration, a morphological structuring element (a ball of radius indicated by the current distance threshold) is applied at locations on the boundary of the solid background as dictated by the mask image. The opening map records the distance threshold at which each image element has been so "opened", while the mask image helps eliminate redundant operations. Although the distance map may use the Euclidean distance metric, the reliance on morphological operations means that it is impractical to generate an opening map of continuous Euclidean distance measure. Only integer-valued distances are recorded, hence the opening map contains a subset of the Euclidean measure, considered in the present work to be a "morphological distance measure".

2.4 Object separation method

This technique makes use of a distance transform of the binary image to create a distance map to which a watershed transformation is applied to separate the pore space into pore objects (Rabbani *et al.*, 2014). This is achieved by identifying watershed basins around each local maximum of the distance transform, resulting in one pore object associated with every local maximum. When pores have a rough surface, application of this technique can break the pore space into many small objects due to additional local maxima near the surface. A main limitation of this partitioning method is the use of spherical structuring elements when identifying watershed basins, which might not cope very well when subject to tortuous interconnected pore clusters.

3 Materials and Methods

Performance of existing PSD software was evaluated on a selection of X-ray CT soil images as well as 3D synthetic images that were constructed based on a simple 3D ball pore geometry at different porosity levels. Comparison was possible among the methods providing the same type of PSD output, either in the form of pore volume fraction per size interval (for BoneJ, CTAnalyser, Quantim4 and Avizo), or object population fraction per size interval (for 3DMA, DFS-FIJI and Avizo). Avizo was the only software in the study that would provide both types of outputs. Additionally, for the synthetic images it was possible to compare the methods against the exact analytical solution.

3.1 Image data

Soil images

Undisturbed soil samples were obtained from the top 5 cm of Mid Pilmore at the James Hutton Institute (JHI, Dundee, UK) under three tillage regimes (no tillage, minimum tillage and ploughed), as previously described by (Sun *et al.*, 2010; Pérez-Reche *et al.*, 2012; Hapca *et al.*, 2013; Houston, *et al.*, 2013b; Juarez *et al.*, 2013). This gave four soil treatments with three replicates per treatment, yielding a total of 12 samples.

Soil images were obtained using a HMX225 X-ray micro-tomography system (NIKON Metrology, UK). The undisturbed samples were scanned at 150 kV and 50 μA using a 2 mm aluminium filter to obtain 1200 angular projections with 4 exposures per frame. A molybdenum target was used. The repacked samples were scanned at 125 kV and 131 μA using a 0.5 mm aluminium filter and 3010 projections. All radiographs were reconstructed into a 3D volume using CT-Pro v.2.0 (NIKON Metrology, UK). For each sample a 512^3 voxels region of interest at the centre of the sample volume was selected and reconstructed at 50 μm resolution. Reconstructed images were mapped from 32-bit floating point to 8-bit unsigned using the outlier rejection method (Houston, *et al.*, 2013b). Segmentation was achieved using Adaptive Window Indicator Kriging (Houston, *et al.*, 2013a) incorporating hysteresis threshold determination as described in Schlüter *et al.* (2010). Standard morphological measures of the pore space including porosity, pore surface area, and connectivity were calculated for each of the soil samples and used as an intial soil treatment comparison. The porosity was calclulated as the image volume fraction occupied by the pore space, the pore surface area was computed for each segmented image according to the prescription in Ohser and Mucklich (2000). The surface area is a dimensionless parameter being calculated relative to the outer surface area of the cube in order to enable for comparisons of different volumes. The connectivity was estimated as the volume fraction of pore space that connects with the surface of the image volume (Houston et al., 2013b).

Synthetic images

Synthetic images were constructed algorithmically by applying a constrained boolean model of 3D balls to create the pore space. The objective of this approach is to obtain images containing clusters of pore bodies, the surface of each cluster being a set of truncated spheres that inter-connect by circular pore throats. In addition to these clusters, a number of non-intersecting spherical pore bodies are also typically present within such images. The choice of spherical pore bodies can be motivated by the fact that it leads to relatively simple design and implementation of the synthesis algorithm, which allows one to formulate appropriate constraints and to determine analytic measures. Specifically, by using only spheres, it allows every pore body to be easily identified and clearly discriminated from all others, at the same time it ensures that every throat aperture is significantly smaller than the bodies it connects.

Once the network structure of pore bodies associated with the synthetic image is available (Figure 1), the procedure is to inspect individual clusters and to delete the smaller ball whenever the overlap between a pair of balls does not meet prescribed criteria. The criteria are selected so as to ensure first that a circular “throat” of intersection is unambiguously defined in every case, and second that each pore body is enclosed by a spherical surface that can be clearly discriminated from that of all other bodies. As well as limiting the degree of overlap (in terms of volume) between any pair of balls, it is important to detect any overlap of throat intersection between three or more balls. This latter condition implies combinatorial processing of the ball descriptions, i.e., each ball needs to be checked against all others in all possible combinations. In practice however, it suffices to detect and rectify triple intersections because quadruple or larger intersections may be decomposed into conjunctions of triple intersections. Given that for every case of multiple intersections, all balls except the two largest ones are deleted, there is no dependency upon the order in which triple intersections are detected. As a result a searching algorithm of $O(n^3)$ order was applied as a simple “brute force” approach to deal with the multiple intersections problem. Additionally, individual spherical bodies are permitted to intersect, subject to a number of constraints, so as to produce more complex pore networks. The purpose of the constraints is to ensure a well-defined circular throat aperture between each intersecting body pair. This means ensuring that the distance between the centre points of overlapping spheres is neither too large nor too small, and also that each circular aperture is distinct from all others. The criteria presented above were used to generate three synthetic images with parameters chosen so that to produce different porosity levels (0.17, 0.24 and 0.29, respectively). For the first two images, additional constraints were used to ensure that individual bodies were fully contained within the image, without intersecting the image boundaries. An exception was made in the case of the third sample (of 0.29 porosity), for which a number of spherical bodies were permitted to touch (without being truncated) the upper and lower surfaces of the image, creating a vertically percolating pore network. For each synthetic image, an exact analytical measure of the pore size distribution was defined by labelling each ball with the corresponding diameter and calculating the relative frequency of balls per size diameter to derive a measure of object population per size interval or, alternatively, by calculating the pore volume occupied by balls of same diameter to derive a measure of pore volume fraction per size interval. The exact analytical solution was further compared with those obtained by the various algorithms.

3.2 Image preparation and application of specific PSD analysis methods

All the PSD methods make use of a binary image consisting only of pore objects versus solid background, converted as necessary to compatible file formats such as TIFF, BMP or RAW format images. It was in some cases necessary to designate object versus background image elements. The exceptions to this include Quantim4 and 3DMA, both of which implicitly identify zero-valued elements as being pore (displayed black). Another exception is ImageJ plugins, which typically identify objects as consisting of the 8-bit element value 255 (usually displayed white) and so require black-pore-object images to be inverted. The specific sequence of operations required for each analysis method together with a brief description of the underlying theoretical approach is given in the supplementary material.

3.3 Statistical evaluation of the PSD analysis results provided by the different methods

Performance evaluation and comparison of the PSD methods presented above was conducted on the synthetic images and the soil image data.

Statistical analysis of the PSD results was conducted by fitting a two-parameter gamma distribution model to the PSDs provided by the different methods for each of the fifteen image samples. The gamma distribution was chosen on the basis that it is a positive distribution, which, depending on the values of the two parameters (shape and scale parameters), can be very flexible in covering a variety of shapes ranging from positively skewed to symmetric. As a result the gamma distribution was a good model candidate to fit the different shapes of the PSDs produced by the different methods and the different soil types or synthetic images. The Non-Linear Mixed-Effect procedure in R (*nmle* package in R v.3.1.1) was used to fit the gamma distribution to the data and to investigate significant difference in the PSD model parameters (for both shape and scale simultaneously) estimated for the different methods. Method comparison was conducted first on all the soil images (twelve samples). For methods performance comparison on the soil data, methods and soil treatments (with four levels no tillage, minimum tillage, ploughed and sieved). were introduced in the model as fixed factors and the soil samples as random factors. To asses the consistency of the methods throughout the soil treatments an interaction terms between methods and treatments were also investigated. A second analysis was also conducted on each of the soil treatment samples (three replicates) separately and on the synthetic images (three replicates). In this analysis, methods were introduced into the model as fixed factors and samples as both fixed and random factor. To assess the consistency of the methods throughout the different samples, interaction effects between methods and samples were also investigated. The methods comparison analysis was perfomermed separately, first for the pore volume fraction based PSD methods and then for the object population fraction based methods.

4 Results

4.1 PSD of the synthetic image data

The distributions of pore volume fraction per size interval estimated by Bone J, CTAnalyser, Quantim4, Avizo and DTM are in good agreement with the analytical solutions. The parameters of the gamma distribution fitted to the PSDs estimated by these five methods are not significantly different from the analytical solution (p-values>0.81). As illustrated in Figure 2, these results are consistent for all three synthetic samples. Avizo and DTM seem to generate some fictitious small diameter results, however. The distributions produced by Avizo and DTM also exhibit some slight irregularities compared to the other three methods. In the case of Avizo, this may be linked to problems evident within the separated object maps, where the separating surfaces in some cases seem excessive in number, giving rise to fragmentary objects.

Estimation of PSD by 3DMA, DFS and Avizo in terms of object population fraction per size interval, shows significant interaction effects between the different methods and the three synthetic image samples (p-values<0.001), indicating that these methods are not stable in their estimation when subjected to a range of pore space morphologies. Compared to the analytical solution (Figure 3), Avizo has a tendency to overestimate the pore population fraction of small class size in the sample with small porosity (A1 -17% porosity). The DFS method in general, overestimates the population fraction of small pores, whereas the 3DMA method overestimated the population fraction of large class size pores.

4.2 PSD of the soil image data

Standard morphological measures of the four treatments, including porosity, pore surface area and connectivity are presented in Table 2. Soil porosity and connectivity were not significantly different among no-tillage, medium-tillage and ploughed treatments (p-values>0.05) whereas the sieved soils had a porosity and connectivity significantly lower than the other three treatments (p-values<0.001). In terms of pore surface area only the ploughed soils appear to be significantly different from sieved soils (p-value<0.001) and medium-tillage soils (p-value<0.05), all the other pairwise comparisons being not significant (p-value>0.05), possibly reflecting a relatively large within-treatment variability (Table 2).

The PSD analysis of soil data, expressed as *pore volume per class size*, revealed that BoneJ, CTAnalyser and Quantim4 are in close agreement with each other but differ significantly from both AVIZO and DTM (in the scale parameter, p-value<0.001 and p-value=0.002 respectively). Examination of the map images for MIB methods (BoneJ, CTAnalyser, Quantim4, and DTM) reveals that many partially filling ball objects are created where pores have a complex shape, a feature that is widespread in the case of soil pores. This feature is very clear in 3D images, but is unfortunately hard to convey adequately in 2D images. Discrepancies in our perception of the connectivity and geometry of the pore space based on 2D and 3D images are well known and unavoidable (Hapca et al., 2011, 2015). Detailed analyses should therefore be based on 3D images. Nevertheless, the cross-section in Figure 4 illustrates well the fact that in some of the wide, complex-shaped pores, instead of having large balls of the relevant diameter, one often finds several smaller balls, occupying less volume. As a result, the MIB based methods produced larger estimates of the pore volume occupied by small pores with less volume being occupied by large pores, compared to Avizo (Figure 5). This tendency gets even more noticeable in the case of PSDs calculated by DTM, which is much skewed at the lower end indicating that most of the large pores get fragmented into very small pores. Comparison of the PSDs for the different soil treatments based on DTM indicates that no-tillage and minimum-tillage treatments were not significantly different in terms of PSD shape and scale parameters (p-values>0.11), whereas all other pairwise treatment differences were significant (p-value<0.05). In turn, based on Avizo only, the sieved and no-tillage treatment appears to be significantly different in terms of PSD scale parameter (p-value=0.046), all other pairwise treatment differences being not significant (p-values>0.08). BoneJ, CTAnalyser and Quantim4 methods were consistent with each other showing that the no-tillage and minimum-tillage treatments were not significantly different in terms of PSD shape and scale parameters (p-values>0.28). The same thing happens with the ploughed and sieved treatments (p-values>0.22). As illustrated in Figure 5, the PSD of the no-tillage and minimum-tillage treatments share similar profiles with more pores of larger size as compared to the ploughed and sieved treatments. The above analysis shows that different methods obtain different estimates for PSD, and assessments of treatment effects are affected by the method chosen.

Comparison of soil PSDs provided by 3DMA, DFS and Avizo in terms of *object population fraction per class size* showed significant differences among methods for all four soil treatments (p-values<0.05). In addition, PSD estimation by the three methods was inconsistently different for the different soil treatments, the fitted gamma model indicating significant interaction effects between methods and treatments (p-values<0.05). As illustrated in Figure 6, for the sieved soil and the no-tillage treatment, there is a relatively good visual agreement in the PSD estimation in particular for classes of larger size, however for the ploughed treatment there is an obvious discrepancy between the methods, with Avizo providing larger frequency estimates of large class size pores as compared to the other two methods (p-values<0.05). As for the minimum-tillage treatment, all three methods appear to disagree in their PSD estimation (p-values<0.05). In particular, the DFS method this time appears to overestimate the frequency of large class size pores as compared to Avizo and 3DMA methods. Comparison of the different soil treatments based on Avizo indicated that the sieved soils were significantly different from all the other treatments in terms of PSD, with all the other pairwise treatment comparison not being significant (p-values>0.10). In turn, the DFS method identified significant differences in terms of PSD between the ploughed soil and the other treatments, and between the sieved treatment and the no-tillage treatment (p-values<0.05), all other pairwise treatment comparison not being significant (p-values>0.08). Finally, the 3DMA method identified all soil treatments as being significantly different in terms of PSD. The lack of agreement between these methods suggest that object population fraction based PSD methods such as 3DMA, DFS or Avizo are not necessarily suited for soil data, in particular for the purpose of soil treatment comparison.

5 Discussion and conclusions

The Pore Size Distribution (PSD) has been widely used as a means of characterising the physical structure of geomaterials including soils, since at least the mid-20th century, with links to both fluid transport properties and the availability of ecological habitat. However, for soils, which are very heterogeneous in their physical structure due to a wide range of pore sizes and morphologies, estimation of the PSD is particularly challenging. Despite significant work on the development of both traditional invasive techniques and non-destructive 3D image analysis methods, there is still no consensus on what method should be used. In this context, the purpose of our work was to present a theoretical review of underlying methodologies and to compare available methods for application in soil science through a statistical framework.

The statistical framework developed in this study for PSD method comparison is based on a gamma distribution model fitted to the PSDs estimated by the different methods for the different soil types. Then, a nonlinear mixed-effect procedure was considered in order to statistically compare the estimated parameters of the gamma distribution model for the different cases. To our knowledge, this is the first time a robust statistical method is developed and used for the purpose of PSD comparison. In the last few years, a number of authors have instead proceeded to a visual comparison of PSDs (e.g., Al-Raoush *et al.,* 2003, Al-Raoush and Wilson 2005; Dong *et al.,* 2008; Ngom *et al.,* 2011). In principle, these two approaches could be viewed as complementary. Our perspective, nevertheless, is that , as with the methods used to threshold CT images (Baveye et al., 2010), an approach that is objective, *i.e.*, does not rely on operator judgment, is likely to lead to more reliable conclusions.

A total of seven methods were considered for assessment in this study, which include the commercially licenced Avizo and CTAnalyser, freely available plugins BoneJ and Skeletonize3D (called here DFS-FIJI) for ImageJ, the 2005 open-source release of 3DMA, as well as the open source libraries Quantim4 and DTM. It was found that all methods presenting the PSD as pore volume per class size (this includes Avizo, CTAnalyser, BoneJ, Quantim4 and DTM) were in good agreement with the analytical solution when tested on the synthetic images. Avizo makes use of spherical structuring elements when identifying watershed basins, while the other four methods share an MIB-based approach to PSD calculation, which explains the good agreement with the analytical solution on the synthetic samples. In turn, a great discrepancy was found between the analytical solution and the methods for which PSD is calculated as object population fraction per class size, in particular for 3DMA and DFS-FIJI. Differences in method estimation appeared to get even wider in the case of soil images, with only CTAnalyser, BoneJ and Quantim4 providing consistently similar distributions for the different soil types, the rest of the methods being all different. These findings are in agreement with some previous studies (Al-Raoush *et al*., 2003, Al-Raoush and Wilson, 2005; Dong *et al.,* 2008; Ngom *et al.,* 2011), which have also reported differences among the PSD estimation methods tested. In the study by Al-Raoush *et al.* (2003), 3D images of synthetic structures of spheres regularly and randomly packed were used to compare the performance of a medial axis approach for pore network extraction against a method based on modified Delaunay tessellation. The two methods provided similar PSD results when tested on synthetic regular packing, but great discrepancies were found when the methods were applied to randomly packed spheres. In a different study, Dong *et al.* (2008) compared four methods, medial axis (Lindquist *et al.,* 1996), maximal ball (Sillin and Patzek, 2006), velocity based (Øren *et al.,* 2006) and grain recognition based algorithm (Øren and Bakke, 2003), on 3D rock microstructure images of both sandstone and carbonate obtained from process based reconstructions and X-ray micro-tomography. Again it was reported that depending on the type of structure and type of images, there is a difference in the level of agreement among PSD estimates provided by the four methods. In particular very little agreement was found for those images presenting pores of low sphericity angular shapes. In a more recent study by Ngom *et al*. (2011), a Delaunay Triangulation Method (DTM) for PSD estimation (Monga *et al.,* 2007, 2008) was compared against the 3DMA method (Linquist *et al.,* 2000) based on two soil samples from two different treatments, a ploughed soil and a grassland soil. It was reported that for both sample images the DTM method tended to fragment the pore space into small pores resulting in PSD with a higher pick at small class sizes as compared to the 3DMA method, which presented lower pick at small class sizes but longer tails for large class sizes.

In the current study, the DTM method was also compared against several other methods for PSD estimations. It was found that while on synthetic images the DTM was in good agreement with the other methods and with the analytical solution, when tested on the soil images the PSD distribution generated by the DTM method was very skewed to the lower end due to many fragmentary pore objects being created for soil pores with complex shape. In turn the watershed-based Aviso method appears to separate the pore space into larger objects as compared to the MIB-based method resulting in low peak, longer-tailed PSDs. However, despite these clear differences in overall performance on soil images (Figure 5), a characteristics of the volume per class size PSD estimation methods was the consistency in the PSD profiles produced by these methods independent of the soil type, suggesting only mild interaction effects between the methods and the soil type on the PSD estimation. This further indicates that each of these PSD estimation methods can be reliably used for the purpose of soil type assessment and comparison.

The second type of methods, based on object population fraction per class size, was less consistent in terms of PSD estimation when applied to both synthetic images and soil image data. In the case of the synthetic images, the 3DMA and DFS estimations for PSD were different from the analytical solution for all three synthetic images. The DFS method appeared to fragment the pore space into smaller pores resulting in PSDs being more skewed towards the lower end as compared to the analytical solution, and this performance was consistent for all three synthetic images. In turn the 3DMA method produced completely different distribution profiles for the three synthetic images with the low porosity image (sample 1), having a longer tailed distribution compared to the analytical solution while for the high porosity sample (sample 3) the distribution had a higher pick at the lower end as compared to the analytical solution (Figure 3). The Avizo method was in agreement with the analytical solution for the high porosity samples, but failed to perform well on the low porosity sample (sample 1), identifying more pore objects of larger size than in reality. An explanation to this can be that for the low porosity sample there is less degree of overlap between the 3D ball objects, and therefore for the overlapping balls the shape of the objects are not too complex to be separated by the watersheds (in particular when a small ball overlaps with a large ball), and so this is kept as one big pore object during the PSD estimation. When applied to soil images, the level of agreement between the three methods depended on the soil type; for the no-tillage and sieved soil all three methods provided very similar PSDs, whereas for the medium tillage and ploughed soils the estimation in the PSD by the different methods was very different. This inconsistency in method performace when applied to different soil treatments was statistically confirmed by the significant interaction effects between methods and the soil types when the gamma model was fitted to the data (Figure 6). This further indicates that the PSD methods based on object population fraction per class size are less reliable to be used for the purpose of soil treatment assessment and comparison.

In general, the lack of agreement among the PSD estimation methods can be attributed to the way each of these methods handle tortuous interconnected pore clusters or rough surface pores, which can lead in some cases to many fragmentary small objects being created along the pore surface. The volume contribution of these small objects is still negligible and therefore volume-based PSD methods are less affected by these artefacts, whereas if a large amount of small pore objects is created, this can have a high impact on the shape of PSDs reporting the relative frequency of objects per class size.

Whereas this study presents an up-to-date theoretical and practical assessment of existing methods for PSD estimation from 3D porous media images, of main interest is the performance of PSD methods on 3D soil images, which pose additional challenges due to the heterogeneous nature of the inner pore structure. Based upon the current analysis, we recomend that PSD be presented as a pore volume per size class, which for the methods tested gave the greatest consistency and confidence that the methods can be used for relative comparisons of samples. Of the methods tested, Bone J, CTAnalyser, Quantim4, Avizo and DTM were in good agreement with the analytical solutions for pore volume per size class. For soil however, only the methods based on MIB (Bone J, CTAnalyser, Quantim4) produced consistent results. We also found that methods based on object population fraction per class size produced unstable results for both the synthetic samples but in particular for the more complex soil samples. We therefore recommend that these methods be avoided till improved further.

Acknowledgements

The authors gratefully acknowledge the assistance provided by Hans-Jörg Vogel (Helmholtz UFZ, Germany) in the installation and use of Quantim4, N Corps (Bruker-Skyscan) who contributed a copy of CTAnalyser for evaluation, Alexandra Kravchenko (Michigan State University, Michigan, USA) for general assistance with 3DMA, and Masha Prodanovic (University of Texas at Austin, USA) for guidance on pore analysis with 3DMA. RF, PB, OM and WO received support from the French ANR (projects ANR-09-SYSCOM MEPSOM and ANR-15-CE01-0006 Soilμ-3D).

References

Al-Kharusi, A.S., Blunt, M.J., 2007. Network extraction from sandstone and carbonate pore space images. J. Pet. Sci. Eng. 56(4), 219-231.

Al-Raoush, R.I., Thompson, K., Willson, C.S., 2003. Comparison of network generation techniques for unconsolidated porous media. Soil Sci Soc Am J. 67(6), 1687-1700.

Al-Raoush, R.I., Willson, C.S., 2005. Extraction of physically realistic pore network properties from three-dimensional synchrotron X-ray microtomography images of unconsolidated porous media systems. J. Hydrol. 300(1-4), 44-64.

Arganda-Carreras, I., Sorzano, C.O.S., Marabini, R., Carazo, J.M., Ortiz-de-Solorzano, C., Kybic, J., 2006. Consistent and elastic registration of histological sections using vector-spline regularization. In: Computer Vision Approaches to Medical Image Analysis, in: Beichel R.R & Sonka M. (Eds.), Lecture Notes in Computer Science, Springer Berlin Heidelberg 4241, pp. 85-95.

Armstrong, R.T., Wildenschild, D., Bay, B.K., 2015. The effect of pore morphology on microbial enhanced oil recovery. J. Pet. Sci. Eng. 130, 16-25.

Arns, C.H., 2004. A comparison of pore size distributions derived by NMR and X-ray-CT techniques. Phys. A 339(1-2), 159-165.

Baldwin, C.A., Sederman, A.J., Mantle, M.D., Alexander, P., Gladden, L.F., 1996. Determination and characterization of the structure of a pore space from 3D volume images. J. Colloid Interface Sci. 181(1), 79-92.

Baveye, P.C., Laba, M., Otten, W., Bouckaert, L., Dello Sterpaio, P., Goswami, R.R., Grinev, D., Houston, A., Hu, Y., Liu, J., Mooney, S., Pajor, R., Sleutel, S., Tarquis, A., Wang, W., Wei, Q., Sezgin, M., 2010. Observer-dependent variability of the thresholding step in the quantitative analysis of soil images and X-ray microtomography data. Geoderma 157(1-2), 51-63.

Baveye, P.C., 2013. Jean-Baptiste De Beunie (1717–1793): Unsung pioneer of the study of soils. Soil Sci. 178(2), 55-59.

Beckingham, L.E., Peters, C.A., Um, W., Jones, K.W., Lindquist, W.B., 2013. 2D and 3D imaging resolution trade-offs in quantifying pore throats for prediction of permeability. Adv. Water Resour. 62, 1-12.

Blum, H., 1973. Biological shape and visual science (part I). J.Theoret. Biol. 38(2), 205-287.

Borgefors, G., 1986. Distance transformations in digital images. Comput. Vision Graph. 34(3), 344-371.

Cazelles, K., Otten, W., Baveye, P.C., Falconer, R.E., 2012. Soil fungal dynamics: Parametrization and sensitivity analysis of modelled physiological processes, soil architecture and carbon distribution. Ecol. Model. (248), 165-173.

Childs, E.C., 1940. The use of soil moisture characteristics in soil studies. Soil Sci. 50(1), 239-252.

Childs, E.C., Collis-George, N.C., 1948. Soil geometry and soil-water equilibria. Trans. Faraday Soc. 44(3), 78-85.

Dal Ferro, N., Delmas, P., Duwig, C., Simonetti, G, Morari, F., 2012. Coupling X-ray microtomography and mercury intrusion porosimetry to quantify aggregate structures of a cambisol under different fertilisation treatments. Soil Tillage Res. 119, 13–21.

Delerue, J.F., Perrier, E., Yu, Z.Y., Velde, B., 1999. New algorithms in 3D image analysis and their application to the measurement of a spatialized pore size distribution in soils', Phys. Chem. Earth Pt. A. 24(7), 639-644.

Delerue, J.-F., Perrier, E., 2002. DXSoil, a library for 3D image analysis in soil science, Comput. Geosci. 28(9), 1041–1050.

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.

Dong, H., Fjeldstad, S., Alberts, L., Roth, S., Bakke S., Øren, P.-E., 2008. Pore network modelling on carbonate: a comparative study of different Micro-CT network extraction methods. In: International symposium of the society of core analysis, Society of Core Analysis, 2008.

Dong, H., Blunt, M., 2009. Pore-network extraction from micro-computerized-tomography images. Phys. Rev. E 80(3), p.036307.

Doube, M., Kłosowski, M.M., Arganda-Carreras, I., Cordelières, F.P., Dougherty, R.P., Jackson, J.S., Schmid, B., Hutchinson, J.R, Shefelbine, S.J., 2010. BoneJ: Free and extensible bone image analysis in ImageJ. Bone 47(6), 1076–1079.

Dougherty, R., Kunzelmann, K.-H., 2007. Computing local thickness of 3D structures with ImageJ. Microsc. Microanal. 13(S02).

Echeverría, J.C., Morera, M.T., Mazkiarán, C., Garrido, J.J., 1999. Characterization of the porous structure of soils: adsorption of nitrogen (77K) and carbon dioxide (273K), and mercury porosimetry. Eur. J. Soil Sci. 50(3), 497–503.

Falconer R., Houston A., Otten, W., Baveye, P., 2012. Emergent behaviour of fungal dynamics: influence of soil architecture and moisture distribution. Soil Sci. 177, 111-119.

Filimonova, S.V., Knicker, H., Kögel-Knabner, I., 2006. Soil micro- and mesopores studied by N_2 adsorption and 129Xe NMR of adsorbed xenon. Geoderma 130(3-4), 218-228.

Haines, W.B., 1930. Studies in the physical properties of soils. The hysteresis effect in caoillary properties, and the modes of moisture distribution associated herewith. J. Agric. Sci. 20, 97-116.

Hajnos, M., Lipiec, J., Świeboda, R., Sokołowska, Z., Witkowska-Walczak, B., 2006. Complete characterization of pore size distribution of tilled and orchard soil using water retention curve, mercury porosimetry, nitrogen adsorption, and water desorption methods. Geoderma135, 307–314.

Hapca S.M, Wang, Z.X., Otten, W., Wilson, C, Baveye, P.C., 2011. Automated statistical method to align 2D chemical maps with 3D X-ray computed micro-tomographic images of soils. Geoderma, 164(3-4), 146-154.

Hapca, S.M., Houston, A.N., Otten, W., Baveye, P.C., 2013. New local thresholding method for soil images by minimizing grayscale intra-class variance. Vadose Zone J. 12(3).

Hapca S., Baveye, P.C., Wilson, C., Lark, R.M., Otten, W., 2015. Three-dimensional mapping of soil chemical characteristics at micrometric scale by combining 2D SEM-EDX data and 3D X-ray CT images. PLoS ONE 10(9): e0137205.

Heyden, A., Kahl, F. (Eds), 2011. Image Analysis. Lecture Notes in Computer Science, Springer Berlin Heidelberg.

Houston, A.N., Otten, W., Baveye, P.C., Hapca, S., 2013a. Adaptive-window indicator kriging: A thresholding method for computed tomography images of porous media, Comput. Geosci. 54, 239-248.

Houston, A.N., Schmidt, S., Tarquis, A.M., Otten, W., Baveye, P.C., Hapca, S.M., 2013b. Effect of scanning and image reconstruction settings in X-ray computed microtomography on quality and segmentation of 3D soil images, Geoderma, 207-208, 154-165.

Iassonov, P., Gebrenegus, T., Tuller, M., 2009. Segmentation of X-ray computed tomography images of porous materials: A crucial step for characterization and quantitative analysis of pore structures, Water Resour. Res. 45(9).

Jiang, Z. Wu, K., Couples, G., Van Dijke, M.I.J., Sorbie, K.S., Ma, J., 2007. Efficient extraction of networks from three-dimensional porous media. Water Resour. Res. 43(12).

Juarez, S., Nunan, N., Duday, A.-C., Pouteau, V., Schmidt, S., Hapca, S., Falconer, R., Otten, W., Chenu, C., 2013. Effects of different soil structures on the decomposition of native and added organic carbon. Eur. J. Soil Biol. 58, 81-90.

Kaestner, A., Lehmann, E., Stampanoni, M., 2008. Imaging and image processing in porous media research. Adv. Water Resour. 31(9), 1174–1187.

Ketcham, R.A, Carlson, W.D., 2001. Acquisition, optimization and interpretation of X-ray computed tomographic imagery: applications to the geosciences. Comput. Geosci. 27(4), 381–400.

Kravchenko, A., Falconer, R.E., Grinev, D., Otten, W., 2011a. Fungal colonization in soils with different management histories: modeling growth in three-dimensional pore volumes. Ecol. Appl. 21(4), 1202-1210.

Kravchenko, A.N. Wang, W., Smucker, A.J.M, Rivers, M., 2011. Long-term differences in tillage and land use affect intra-aggregate pore heterogeneity. Soil Sci. Soc. Am. J. 75(5), 1658-1666.

Lee, T.C., Kashyap, R.L., Chu, C.N., 1994. Building skeleton models via 3-D medial surface axis thinning algorithms. Graph. Models Im. Proc. 56(6), 462–478.

Lindquist, W.B., Lee, S.-M., Coker, D.A., Jones, K.W., Spanne, P., 1996. Medial axis analysis of void structure in three-dimensional tomographic images of porous media. J. Geophys. Res. 101(B4), 8297.

Lindquist, W.B., Venkatarangan, A., 1999. Investigating 3D geometry of porous media from high resolution images. Phys. Chem. Earth Pt. A. 24(7), 593–599.

Lindquist, W.B., Venkatarangan, A., Dunsmuir, J, Wong, T., 2000. Pore and throat size distributions measured from synchrotron X-ray tomographic images of Fontainebleau sandstones. J. Geophys. Res. 105(B9), 21509.

Lindquist, W.B., 2002. Quantitative analysis of three-dimensional X-ray tomographic images. International Symposium on Optical Science and Technology. International Society for Optics and Photonics, 103–115.

Luo, L., Lin, H., Li, S., 2010. Quantification of 3-D soil macropore networks in different soil types and land uses using computed tomography. J. Hydrol. 393 (1-2), 53–64.

Merkus, H.G., 2009. Particle size measurements: fundamentals, practice, quality. Springer. p.534.

Monga, O., Ngom, N.F., Delerue, F.J., 2007. Representing geometric structures in 3D tomography soil images: Application to pore-space modeling. Comput. Geosci. 33(9), 1140-1161.

Monga, O., Bousso, M., Garnier, P., Pot, V., 2009. Using pore space 3D geometrical modelling to simulate biological activity: Impact of soil structure, Comput. Geosci. 35(9), 1789-1801.

Muñoz-Ortega, F.J., San José Martínez, F., Caniego Monreal, F.J., 2014. Volume, surface, connectivity and size distribution of soil pore space in CT images: Comparison of samples at different depths from nearby natural and tillage areas. Pure Appl. Geophys. 172(1), 167-179.

Naveed, M., Moldrup, P., Vogel, H.-J., Lamandé, M., Wildenschild, D., Tuller, M., Wollesen, L., 2014a. Impact of long-term fertilization practice on soil structure evolution. Geoderma 217-218, 181-189.

Naveed, M., Arthur, E., de Jonge, L.W., Tuller, M., Moldrup, P., 2014b. Pore structure of natural and regenerated soil aggregates: An X-ray computed tomography analysis. Soil Sci. Soc. Am. J. 78(2), 377-386.

Ngom, N.F. Garnier, P., Monga, O., Peth, S., 2011. Extraction of three-dimensional soil pore space from microtomography images using a geometrical approach. Geoderma 163(1-2), 127–134.

Øren, P.-E., Bakke, S., 2003. Reconstruction of berea sandstone and pore-scale modelling of wettability effects. J. Pet. Sci. Eng. 39, 177-199.

Øren, P.-E., Bakke, S., Rueslatten, H.G., 2006. Digital core laboratory:Rock and flow properties derived from computer generated rocks. Proceedings of International Symposium of the Society of Core Analysis, trondheim, Norway.

Pajor, R., Falconer, R., Hapca, S., Otten, W., 2010. Modelling and quantifying the effect of heterogeneity in soil physical conditions on fungal growth. Biogeosciences. 7(11), 3731-3740.

Pérez-Reche, F.J., Taraskin, S.N., Otten, W., Viana, M.P., Costa, L. da F, Gilligan, C.A., 2012. Prominent effect of soil network heterogeneity on microbial invasion. Phys. Rev. Lett. 109(9), 98102.

Peth, S.,Horn, R., Beckmann, F., Donath, T., Fischer, J., Smucker, A.J.M., 2008. Three-dimensional quantification of intra-aggregate pore-space features using synchrotron-radiation-based microtomography. Soil Sci. Soc. Am. J. 72(4), 897-907.

Pierret, A., Capowiez, Y., Belzunces, L., Moran, C.J., 2002. 3D reconstruction and quantification of macropores using X-ray computed tomography and image analysis. Geoderma 106(3-4), 247-271.

Prodanović, M., Lindquist, W.B., Seright, R.S., 2006. Porous structure and fluid partitioning in polyethylene cores from 3D X-ray microtomographic imaging. J. Colloid Interface Sci. 298(1), 282-297.

Rab, M.A., Haling, R.E., Aarons, S.R., Hannah, M., Young, I.M., Gibson, D., 2014. Evaluation of X-ray computed tomography for quantifying macroporosity of loamy pasture soils. Geoderma 213, 460-470.

Rabbani, A., Jamshidi, S., Salehi, S., 2014. An automated simple algorithm for realistic pore network extraction from micro-tomography images. J. Pet. Sci. Eng. 123, 164-171.

Schlüter, S., Sheppard, A., Brown, K., Wildenschild, D., 2014. Image processing of multiphase images obtained via X-ray microtomography: A review. Water Resour. Res. 50(4), 3615–3639.

Schlüter, S., Weller, U., Vogel, H.-J., 2010. Segmentation of X-ray microtomography images of soil using gradient masks. Comput. Geosci. 36(10), 1246–1251.

Schmidt, S, Bengough, A.G., Gregory P.J., Grivev, D.V., Otten, W., 2012. Estimating root-soil contact from 3D X-ray microtomographs. Eur. J. Soil Sci. 63(6), 776-786.

Serra, J.P., 1982. Image analysis and mathematical morphology, Volume 1. Academic Press.

Sheppard, A.P., Sok, R.M., Averdunk, H., 2004. Techniques for image enhancement and segmentation of tomographic images of porous materials. Phys. A 339(1-2), 145-151.

Siddiqi, K., Pizer, S.M. (Eds), 2008. Medial representations. In: Computational Imaging and Vision, Dordrecht, Springer, Netherlands.

Silin, D., Patzek, T., 2006. Pore space morphology analysis using maximal inscribed spheres. Phys. A 371(2), 336-360.

Sun, B., Hallett, P.D., Caul, S., Daniell, T.J., Hopkins, D.W., 2010. Distribution of soil carbon and microbial biomass in arable soils under different tillage regimes. Plant Soil 338(1-2), 17-25.

Taina, I.A., Heck, R.J., Elliot, T.R., 2008. Application of X-ray computed tomography to soil science: A literature review. Can. J. Soil Sci. 88(1), 1-19.

Talabi, O., AlSayari, S., Iglauer, S., Blunt, M.J., 2009. Pore-scale simulation of NMR response. J. Pet. Sci. Eng. 67(3-4), 168-178.

Thovert, J.F.,Yousefian, F., Spanne, P., Jacquin, C.G., Adler, P.M., 2001. Grain reconstruction of porous media: application to a low-porosity Fontainebleau sandstone. Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys. 63, 061307.

Vaz, C.M.P., Manieri, J.M., de Maria, I.C., Tuller, M., 2011. Modeling and correction of soil penetration resistance for varying soil water content. Geoderma 166(1), 92-101.

Vogel, H.-J., 1997. Morphological determination of pore connectivity as a function of pore size using serial sections. Eur. J. Soil Sci. 48(3), 365-377.

Vogel, H.-J., Roth, K., 2001. Quantitative morphology and network representation of soil pore structure. Adv. Water Resour. 24(3-4), 233-242.

Vogel, H.-J., Weller, U., Schlüter, S., 2010. Quantification of soil structure based on Minkowski functions. Comput. Geosci. 36(10), 1236–1245.

Wang, W., Kravchenko, A.N., Johnson, T., Srinivasan, S., Ananyeva, K.A., Smucker, A.J.M, Rose, J.B., Rivers, M.L., 2013. Intra-aggregate pore structures and distribution by water flow within and movement out of soil macroaggregates. Vadose Zone J. 12(4).

Wildenschild, D, Sheppard, A.P., 2013. X-ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems. Adv. Water Resour. 51, 217-246.

Wildenschild, D., Vaz, C.M., Rivers, M.L., Rikard, D., Christensen, B.S.B., 2002. Using X-ray computed tomography in hydrology: systems, resolutions, and limitations, J. Hydrol. 267(3-4), 285-297.

Zaffar, M., Lu., S.G., 2015. Pore size distribution of clayey soils and its correlation with soil organic matter. Pedosphere 25(2), 240-249.

Zong, Y., Yu, X., Zhu, M., Lu, S., 2014. Characterizing soil pore structure using nitrogen adsorption, mercury intrusion porosimetry, and synchrotron-radiation-based X-ray computed microtomography techniques. J. Soils Sediments 15(2), 302-312.

Tables

Table 1. Examples of methods for pore size distribution estimation, the main algorithm they are based upon and the main measures they produce as output to present the pore sizes in a porous medium. The reported studies had a **main** element of method development or application and are all based on information from 2D thin sections or 3D data sets obtained with X-ray CT. See text for description of main algoritms and methods.

Publication	Software	Method
Baldwin et al. 1996	Autors - development	Medial axis by morphological thinning
Lindquist et al., 1996, 2000	3DMA - development	Medial axis by morphological thinning/burning of pore space
Vogel 1997	Quantim4 - development	Morphological opening by errosion and dilation using an incremental spherical structuring elements
Vogel and Roth, 1998, 2001	Quantim4 - application	Morphological opening (Vogel 1997)
Lindquist & Venkatarangan, 1999	3DMA - application	Medial axis by morphological thinning (Lindquist *et al*. 1996)
Delerue et al., 1999	Authors - development	Medial axis by Voronoi tessellation, maximum inscribed balls
Thovert et al., 2001	Authors - development	Maximum inscribed balls
Lindquist, 2002	3DMA - development	Medial axis by morpholocal thinning
Delerue and Perrier, 2002	DXView - development	Medial axis by Voronoi tessellation, maximum inscribed balls
Pierret et al., 2002	Authors - development	Morphological opening (using a 32 face “sphere” structuring element)
Arns, 2004	Author - application	Maximum inscribed balls (Thovert *et al*. 2001)
Al-Raoush and Wilson, 2005	3DMA - application	Medial axis by morphologal thinning (Lindquist *et al*. 1996)
Silin and Patzek, 2006	Authors - development	Maximum inscribed ball
Prodanovic et al., 2006	3DMA - development	Medial axis buy morphological thinning
Al-Kharusi and Blunt, 2007	Authors - development	Maximum inscribed balls
Jiang et al., 2007	Authors - development	Medial axis by morphological thinning prioritized by Euclidean distance
Monga et al., 2007, 2009	DTM - development	Delaunay triangulation and maximum inscribed balls
Peth et al., 2008	3DMA - application	Medial axis by morphological thinning (Lindquist *et al.,* 2000)
Dong and Blunt, 2009	Authors - development	Maximum inscribed ball
Talabi et al., 2009	Authors - application	Maximum inscribed balls (Al-Kharusi and Blunt, 2007).
Doube et al., 2010	BoneJ (ImageJ) - development	Maximum inscribed ball, medial axis by finding ridges on an Euclidean distance map
Luo et al., 2010	Avizo5 - application	Object separation, 3D skeletonization.
Kravchenko et al. 2011b	3DMA - application	Medial axis by morphological thinning (Lindquist *et al.,* 2000)
Ngom et al., 2011	DTM and 3DMA - application	Delaunay triangulation (Monga *et al.,*2007, 2009), medial axis (Lindquist *et al.* 1996)
Vaz et al., 2011	Authors - application	Morphological opening (Vogel and Roth, 1998; Pierret *et al.,* 2002)
Beckingham et al, 2013	3DMA - application	Medial axis by morphological thinning (Lindquist *et al.,*1996, 2000)
Wang et al., 2013	3DMA- application	Medial axis by morphological thinning (Lindquist *et al.,* 2000)
Rabbani et al., 2014	Authors - development	Object separation by distance and watershed transform
Naveed et al., 2014b	BoneJ (ImageJ) - application	Maximum inscribed balls (Doube *et al.,* 2010)
Munoz-Ortega et al. 2015	Quantim4 - application	Morphological opening (Vogel 1997; Vogel *et al.* 2010)
Armstrong et al., 2015	3DMA - application	Medial axis by morphological thinning (Lindquist 2002; Prodanović *et al.* 2006).

Table 2. Summary of soil morphological measures of the four soil treatment showing porosity, pore surface area, and pore connectivity sample mean±SE (n=3).

Morphological soil properties	No tillage	Medium tillage	Ploughed	Sieved
Porosity	0.108±0.012	0.118±0.018	0.144±0.021	0.044±0.002
Surface area	17.039±1.081	14.619±2.464	25.183±5.224	9.675±0.485
Connectivity	0.764±0.063	0.909±0.020	0.880±0.059	0.299±0.043

Legend to figures

Figure 1. Graphical renderings of a synthetic sample (corresponding to 17% porosity) illustrating from left to right (a) network formed by intersecting balls, (b) isolated (dark) and intersecting (light) balls, and (c) colour labelling of cluster image elements based on the 6-connected neighbourhood. In the latter image, only the largest 250 clusters are assigned a distinct colour, the remainder are shown in transparent grey.

Figure 2. Distribution of pore volume fraction per size interval for the synthetic images and the corresponding Gamma distribution fit; comparison of PSD methods (BoneJ, CTAnaliser, Quantim 4, DTM and Avizo) against the analytical solution.

Figure 3. Distribution of pore object population fraction per size interval for the synthetic images and the corresponding Gamma distribution fit; comparison of PSD methods (Avizo, DFS-FIJI, 3DMA) against analytical solution.

Figure 4. (a) Illustrative cross-sectional thresholded image through one of the soil samples (sample M1-1), with black pixels representing the solid phase and white pixels the pore space. (b) Image of the same cross-section with the pore space approximated with balls, using DTM. At the top left and at the bottom of this image, there is evidence of partial filling of pores due to edge effects, which can be eliminated by selecting a smaller image after approximation by balls. Throughout the image, pores with complex geometries tend to be partially filled by a combination of small and slightly large balls.

Figure 5. Distribution of pore volume fraction per size interval on a selection of soil images, as calculated by BoneJ, CTAnalyser and Quantim 4, DTM and Avizo, and the corresponding Gamma distribution fit. A single Gamma distribution was fitted to BoneJ, CTAnalyser and Quantim4 as these three methods were found nor significantly different.

Figure 6. Distribution of pore object population fraction per size interval on a selection of soil images, as calculated by Avizo, DFS-FIJI and 3DMA, and the corresponding Gamma distribution fit.

Figure 1

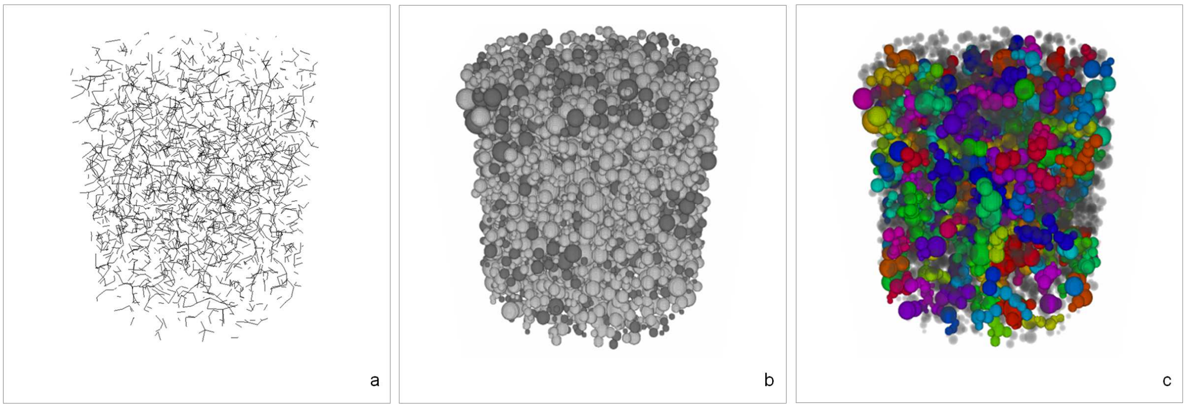


Figure 2

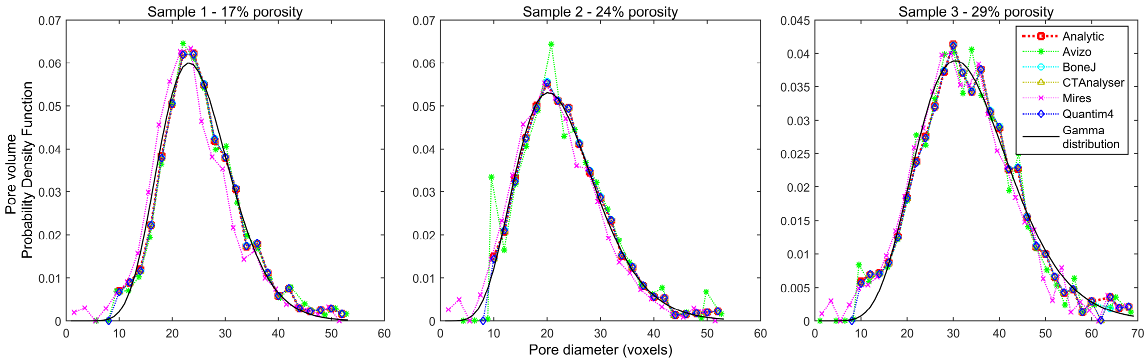

Sample 1 - 17% porosity
Sample 2 - 24% porosity
Sample 3 - 29% porosity
Pore volume
Probability Density Function
Pore diameter (voxels)
Analytic
Avizo
BoneJ
CTAnalyser
Mires
Quantim4
Gamma distribution

Figure 3

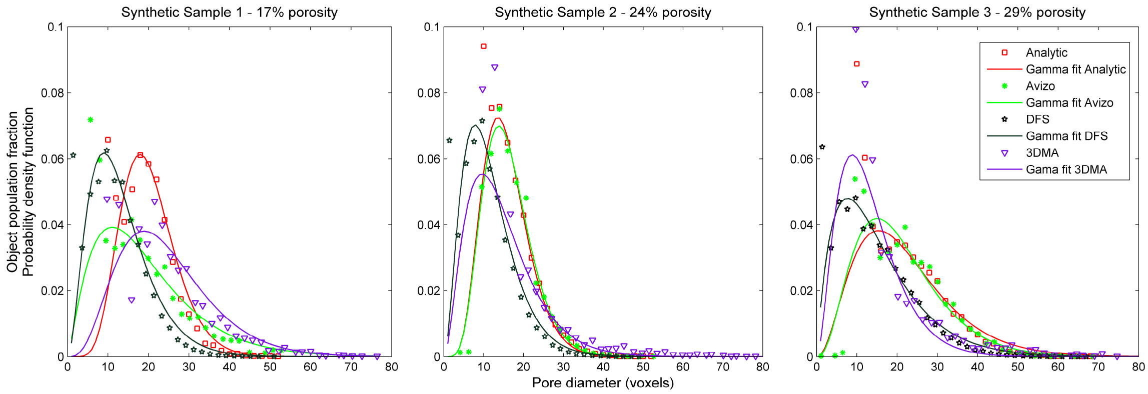

Synthetic Sample 1 - 17% porosity
Synthetic Sample 2 - 24% porosity
Synthetic Sample 3 - 29% porosity
Object population fraction
Probability density function
Pore diameter (voxels)
Analytic
Gamma fit Analytic
Avizo
Gamma fit Avizo
DFS
Gamma fit DFS
3DMA
Gama fit 3DMA

Figure 4

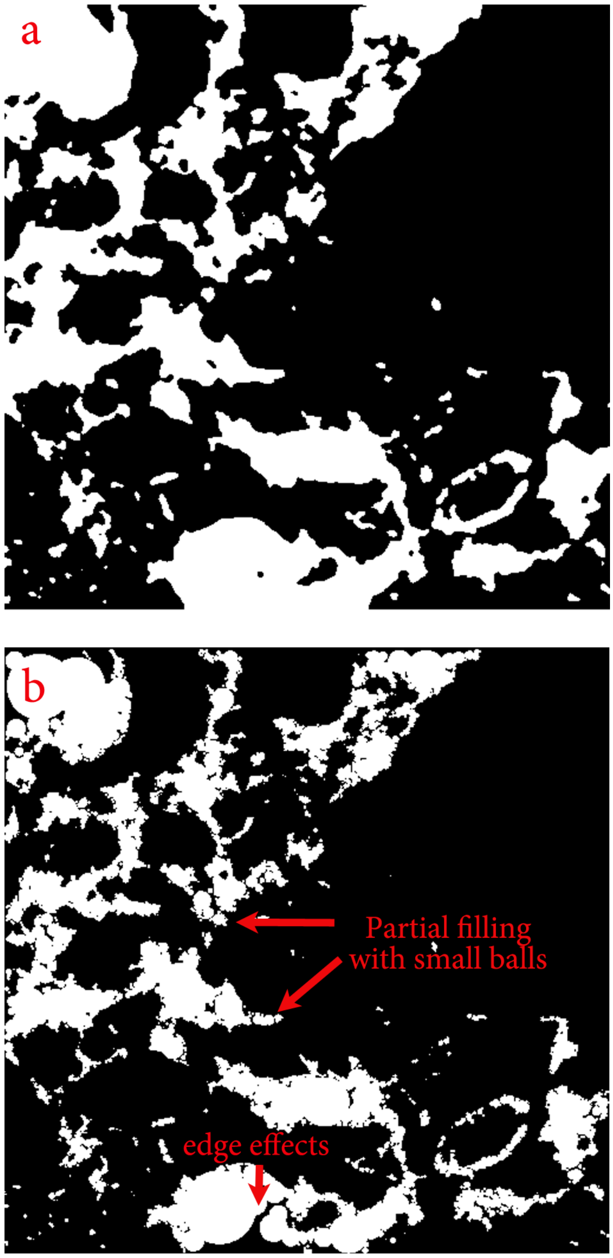
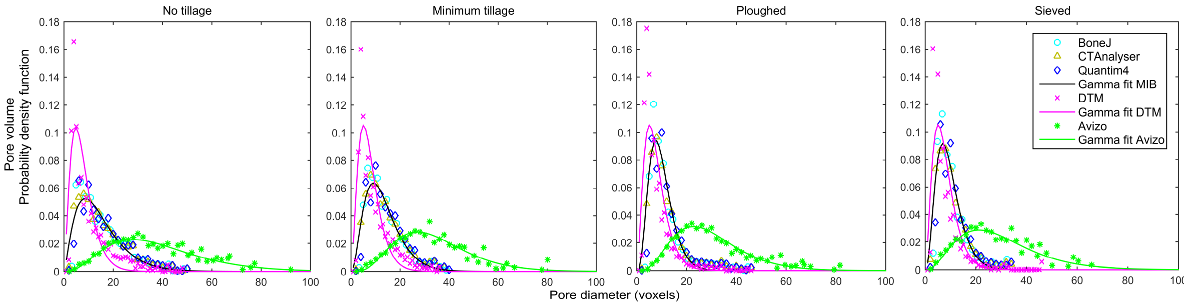
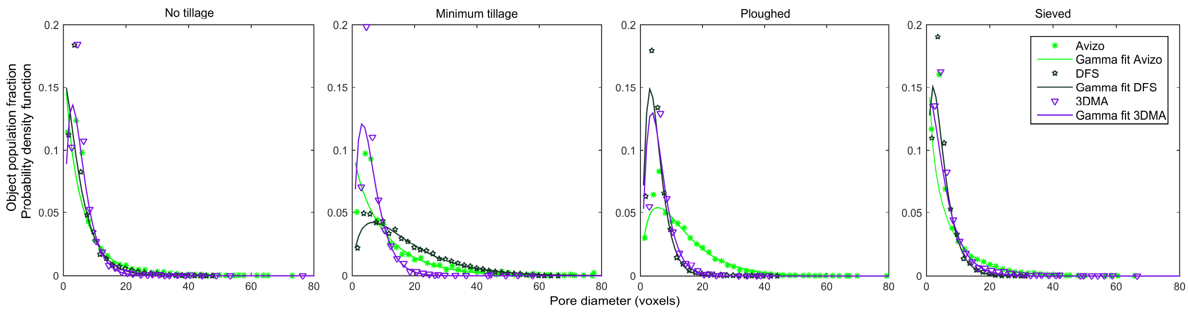


Figure 5



No tillage
Minimum tillage
Ploughed
Sieved
Pore volume
Probability density function
Pore diameter (voxels)
BoneJ
CTAnalyser
Quantim4
Gamma fit MIB
DTM
Gamma fit DTM
Avizo
Gamma fit Avizo

Figure 6


No tillage
Minimum tillage
Ploughed
Sieved
Object population fraction
Probability density function
Pore diameter (voxels)
Avizo
Gamma fit Avizo
DFS
Gamma fit DFS
3DMA
Gamma fit 3DMA
0.2
0.15
0.1
0.05
0
20
40
60
80

Supplementary material

Description of the computer packages used for PSD estimation

Avizo Fire

Avizo (version 7.1) is commercial software consisting of a base application (providing the ability to process and visualise data in many formats) plus a range of optional software modules that extend the basic functionality (FEI Visualization Sciences Group). The “Fire” extension package provides a selection of such extensions appropriate for materials science and these were used within the present study.

The ‘separate objects’ feature of Avizo Fire, which operates by constructing watersheds within the background phase of the image, was used to divide pore objects into size classes (Avizo documentation). These watersheds are allowed to project through objects in the form of surfaces explicitly represented using image elements, which may result in some slight distortion of measures (such as volume) on objects separated in this manner.

The effect of watershed separation applied to soil pores is that larger and more tortuous pore clusters are partitioned into a number of smaller objects. This tends to reduce the incidence of object concavity within the image, producing a population of more convex objects. Each object can then be individually labelled (using region-growing on a specified local neighbourhood) and finally measured using specific functions built into Avizo. For the present study, labelling was carried out using the 6-connected neighbourhood and the built-in measure functions “Width3d”, “EqDiameter”, “Unweighted” and “Volume3d” were collected per object. The “Width3d” measure is an estimate of the Feret diameter (Merkus, 2009) and “EqDiameter” is the diameter of an equivalent sphere (i.e. one whose volume equals that of the object). Feret diameter was estimated using 30 samples which is the default value suggested by the software. The “Unweighted” function provided the number of objects per object diameter while the “Volume3d” measure provides the volume estimated by point counting.

CTAnalyser

CTAnalyser (version 1.13) is provided commercially along with X-ray CT equipment as part of an overall imaging solution (Skyscan-Bruker microCT). This software implements a size measure entitled “structure thickness” (commonly known as “trabecular thickness” when bone is analysed) by a method based on fitting maximal balls within the object. This is achieved via analysis of a distance map but the metric used to form the distance map is not specified in the documentation, nor is the resulting map exported. The software does however permit the final size map image to be saved to disk. The elements of the size map image are 8bit indices denoting the size category of the covering ball, the index zero denotes background phase elements while the smallest objects are denoted by an index of one: objects belonging to larger size categories are assigned indices in ascending order. An accompanying report text file allows each index value to be related to a size measure and also gives the volume (estimated by point counting) for each size class. After designating the objects of interest using a thresholding operation, the remaining processing is fully automatic, i.e. no user specified parameters are involved.

As the information reported by CTAnalyser is rather limited, the categorical size map image was used to calculate corresponding pore volume per size category. This functional measure calculation on the size map image was achieved using own software.

BoneJ (FIJI/ImageJ)

BoneJ (Doube *et al.*, 2010) is a freely available plugin module within the FIJI image analysis platform which is a software distribution of ImageJ (version 1.47). BoneJ implements a “structure thickness” measure based on MIB fitting along the medial axis, which is derived by finding ridges on an Euclidean distance map (Dougherty and Kunzelmann, 2007). It results in a map image of MIB diameters, from which a volume weighted distribution can be obtained. The image histogram feature of FIJI can be used for this purpose, bearing in mind that volume estimates are obtained by point-counting. As part of the investigative work of the current study, it was determined that BoneJ version 1.3.12 (released 29th April 2014) and earlier versions, produce MIB map images that do not conserve image structure. Specifically, each MIB object generated within the map image overlaps the image background, leading to inflated estimates of pore volume. As a result, each MIB diameter is enlarged by approximately two image elements, which although small at diameter level, it is sufficient to noticeably bias the estimated size distribution. This problem was corrected in this study (and the subsequent BoneJ releases) by masking of the size measure image, i.e., setting to zero any measure that lies outside of the original object, as defined by the original dichotomous image of pore versus solid.

Quantim4

Quantim4 (version 4.8) is an open source C/C++ function library (http://www.ufz.de/index.php?en=39198) applicable mainly to Linux systems (the use of features specific to the GNU g++ compiler mean that the code is not easily portable to other systems). The analysis of images with Quantim4 requires some programming ability, but owing to the convenient high level functions provided by the Quantim4 library a useful analysis program can be both small and simple in structure.

Quantim4 uses a ball shaped structuring element in a sequence of morphological operations guided by the Euclidean distance map. Morphological openings are applied to individual image elements and these operations are both parameterised and ordered according to a distance measure on those image elements (Vogel, 1997; Vogel and Roth, 2001). This approach is equivalent to the direct fitting of maximal balls, as described in (Coeurjolly, 2012) and achieves results that are quite similar in practice. The underlying algorithm can be briefly summarized as follows: first a distance map image is computed using the squared Euclidean distance metric, then this distance map is used to construct the "Open Map" (terminology provided by Quantim4 documentation) by mathematical morphology. The final processing stage computes Minkowski functionals (Vogel *et al.*, 2010) for thresholds of the Open Map, providing cumulative measures (including the volume fraction of objects) per size interval. The size measure reported by Quantim4 is determined as the diameter of a sphere containing a volume equivalent to that of the morphological structuring element for each size class.

DTM

The method developed by Monga and co-workers (Monga *et al.,* 2007, 2009, Ngom *et al.,* 2011) involves a number of successive steps. The first consists of selecting pore boundary points, defined as points in the interior of pores, which have at least one neighbor voxel that does not belong to the pore space. A 3D Delaunay triangulation of boundary points is then computed using the very fast code developed by George (2004). All tetrahedrons that are not included entirely in the pore space are removed, and Delaunay spheres, i.e., spheres passing through the four vertices of a given tetrahedron, are computed for the tetrahedra that remain. These Delaunay spheres are maximal in the sense that they are fully contained within the pore space and that no other sphere (within the pore space) contains it. The centers of all the Delaunay spheres are then assumed to constitute the "skeleton" of the pore space, referred to either as 'medial axes" (Ngom *et al.,* 2011) or "Lambda-skeleton". This approximation is reasonable because it can be shown that when the sampling of a surface defining a volume shape tends to 0, then the set of the centers of Delaunay spheres converges uniformly to the shape skeleton. The last step of the method then involves the use of heuristc algorithms to compute a minimal set of maximal balls covering the Lambda skeleton, with "minimal" interpreted in a cardinal sense. The basic idea of the heuristic is to place iteratively the biggest ball, by maintaining a minimal covering with the already selected balls. Once the minimal set of maximal spheres is obtained, the distribution of spheres can be used easily to compute a pore size distribution.

3DMA

The 2005 release of 3DMA is an open source package consisting of many image analysis algorithms invoked via a hierarchy of text menus (the 2011 or later release of 3DMA is commercial software and provides a graphical interface; the commercial version was not assessed within the present study). Analysis of pore size in 3DMA is based on detecting the location of pore throats, these being the narrow apertures that separate a pore cluster into distinct bodies. This is achieved by analysis of a “burn map” of the pore space obtained using the pore space burning algorithm, which is equivalent to a form of distance map obtained using either the Manhattan or chessboard distance metric (Lindquist *et al.*, 1996, 2000; Lindquist, 2002). The latter metric was used in this study. The discrete skeleton of the burn map is determined using the LKC algorithm (Lee *et al.*, 1994) and then local minima of the burn number (distance measure) on the skeleton are used to guide the search for minimum area planar throats. The result of this analysis is a pair of binary encoded data files, one listing throat locations and the other body locations. In the former case the estimated area for each throat is also given, while the element (voxel) count for each body is given in the latter case. The distributions of these data can be plotted from within 3DMA, but we elected also to process the data files using our own software in order to control histogram binning.

DFS-FIJI

This method combines two software tools available within FIJI (ImageJ, version 1.47) that are both applied to the image of pore structure. The first tool, “Skeletonize3D” (Arganda-Carreras *et al*. 2008) generates a discrete skeleton map image of pore space using the LKC algorithm (Lee *et al.,*1994). The second tool generates the Euclidean distance map of pore space, using an unspecified algorithm (nor is any author credited). The conjunction of these two map images (preserving the Euclidean measure only where the skeleton is defined) is referred to, within the present work, as a Discrete Skeleton Function (henceforth DSF), an approximate representation of the medial axis function (Blum, 1973). Discarding spatial information, the DSF can be interpreted as a population of local radius measures and hence may be used directly to form a population diameter distribution. Alternatively a means of approximating a volume measure per skeleton element is to treat each as being the centre of a disk, calculating the area of the disk and then extruding this by one voxel to obtain the volume of a circular cylinder. The total volume estimate obtained in this way does not resemble the true pore volume, but might be scaled so as to plot a crude estimate of volume fraction versus diameter. Irrespective of the manner of presentation of the DSF, imprecision is introduced by using only discrete information (map images) without reconstructing the underlying continuous medial axis function. Where the local pore diameter is even valued, the discrete skeleton map rounds the axial location to the nearest element, hence the selected Euclidean distance measure will be in error by ±0.5 elements.