Computers and Geosciences 123 (2019) 20–37

Contents lists available at ScienceDirect

Computers and Geosciences

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

Research paper

From spheres to ellipsoids: Speeding up considerably the morphological modeling of pore space and water retention in soils

Alain Trésor Kemgue[a], Olivier Monga[b,*], Serge Moto[a], Valérie Pot[c], Patricia Garnier[c], Philippe C. Baveye[d], Abdelaziz Bouras[e]

[a] UMMISCO-Cameroun, UMI 209 UMMISCO, University of Yaoundé, IRD, University of Paris 6, Bondy Cedex F, 93143, France
[b] Sorbonne Université, IRD, UMMISCO, Bondy, F-93143, France
[c] UMR 1402 ECOSYS, INRA, Avenue Lucien Brétignières, Thiverval-Grignon, 78850, France
[d] UMR 1402 ECOSYS, AgroParisTech, Avenue Lucien Brétignières, Thiverval-Grignon, 78850, France
[e] Department of Computer Science, Qatar University, P. O. Box 2713, Doha, Qatar

ARTICLE INFO

Keywords:
Pore scale
Synchroton X-ray micro computed tomography
Soil air-water interfaces
Computational geometry
3D volume segmentation
Morphological modeling

ABSTRACT

In recent years, technological advances have stimulated researchers to try to unravel the extremely complex microscale processes that control the activity of microorganisms in soils. In particular, significant work has been carried out on the development of models able to accurately predict the microscale distribution of water, and the location of air–water interfaces in pores. A comparison, by Pot et al. (2015), of two different modeling approaches with actual synchrotron-based tomography data, shows that a two-phase lattice Boltzmann model (LBM) is able to predict remarkably well the location of air–water interfaces but is extremely slow, whereas a morphological model (MOSAIC), representing the pore space as a collection of spherical balls, provides a reasonable approximation of the observed air–water interfaces when each ball is allowed to drain independently, but does so blazingly fast. Interfaces predicted by MOSAIC, however, tend to have nonphysical shapes. In that general context, the key objective of the research described in the present article, based on the same tomography data as Pot et al. (2015), was to find out to what extent the use of ellipsoids instead of spherical balls in MOSAIC could not appreciably speed up computations, or at least, at equal computational time, provide a quantitatively better approximation of water-air interfaces. As far as we know, this is the first time an ellipsoids-based approximation of the soil pore space is proposed. A secondary objective was to assess whether ellipsoids might yield smoother, more physical, interfaces. Simulation results indicate that the use of ellipsoids provides a sizeable increase in accuracy in the prediction of air-water interfaces, an approximately 6-fold drop in computation time, and much more realistic-looking interfaces, compared to what is obtained with spherical balls. These observations are encouraging for the use of models based on geometric primitives to describe a range of microscale processes, and to address the still daunting issue of upscaling to the macroscopic scale.

1. Introduction

Over the last two decades, significant progress has been achieved in the microscopic characterization of the geometry of interstices in porous media, and in particular in soils. Since the mid-90s, researchers have had access to X-ray micro-computed tomography (X-ray μCT) equipment at the various synchrotrons facilities around the world, providing 3-D images of soils at micrometric resolution. More recently, in the past 15 years, soil scientists have also had dedicated table-top scanners at their disposal, whose resolution is still steadily improving and is now as fine as 0.3 μm for small samples. At the same time, significant efforts have been devoted to the development of efficient, operator-independent segmentation algorithms that allow the 3D grayscale images produced by X-ray μCT scanners to be transformed into binary, i.e., black and white, images (Houston et al., 2013; Baveye et al., 2018). These images afford detailed information about the geometry and topology of pores in soils, at least of the portion of the pores that are visible at the finite resolution of μCT images (Baveye et al., 2017, 2018).

The availability of high-resolution information about the geometry of soils has stimulated a significant research effort by modelers in recent years (Baveye, 2015; Falconer et al., 2012; Houston et al., 2013; Masse et al., 2007; Monga et al., 2008, 2009; Nguyen-Ngoc et al., 2013; Nguyen-Ngoc et al., 2015; Vogel et al., 2015). In earlier days, attempts

* Corresponding author.
E-mail address: Olivier.Monga@ird.fr (O. Monga).

https://doi.org/10.1016/j.cageo.2018.11.006
Received 19 February 2018; Received in revised form 6 November 2018; Accepted 9 November 2018
Available online 12 November 2018
0098-3004/

to model microscale processes in soils relied on the assimilation of pores to idealized geometrical structures, typically geometrically-regular networks of cylinders (e.g., Held and Celia, 2001) or angular wedges (e.g. Ebrahimi and Or, 2015a,b) of various sizes. Network features (e.g., pore size distribution) were determined by comparison of model outputs with macroscopic measurements, like the soil moisture retention curve. By contrast, research over the last 15 years has focused on the actual geometry of pores, as revealed in μCT images. Among the different approaches that have been explored, a new generation of network models has emerged, where the geometrical configuration of the networks is obtained directly from inspection of binary 3D CT images of soils. Using this technique and including the cross-section areas and arc-length of the pores in the description of the different branches of the network, Pérez-Reche et al. (2012) were able calculate the probability for a microbe to enter individual pores and, broadly, to propagate through soils.

Several other approaches have been explored as well, to describe various retention or transport processes in soils, and when coupled with suitable descriptions of microbial activity, to model microbially-mediated processes as well. In this context, the lattice-Boltzmann method has been, and continues to be, used extensively (e.g., Vogel et al., 2015; Portell et al., 2018; Baveye et al., 2018). For example, Pot et al. (2015) show that a two-phase, two relaxation-times (TRT) lattice-Boltzmann model produces remarkably accurate predictions of air-water interfaces in individual pores of a silty loam soil. In a few cases, LBM predicts the presence of air in pores where none is observed experimentally, whereas occasionally the opposite occurs. Nevertheless, these discrepancies are very minor. In general, the results of the comparison are remarkably good, with a high level of agreement between measured and predicted locations of air–water interfaces. These observations suggest that LBM-based models can be used reliably to determine where water is located in the pore space of soils, information that is crucial to describe properly the activity of soil bacteria, archaea, and fungi.

One sizeable drawback of the LBM is the very long time that computations take to process the hundreds of millions of voxels involved in typical situations. For example, it took of the order of 20 h for the longest simulations carried out by Pot et al. (2015). With techniques like domain decomposition, the use of computer clusters with multiple processors, and computations using graphic processors (GPUs) instead of CPUs, it will probably be possible to speed the computations up drastically in years to come (e.g., Vidal et al., 2010), but the size and resolution of 3D images that one will want to use will also steadily increase for a variety of reasons (e.g., representativeness of simulations), so that no matter what computational breakthroughs occur, LBM modeling is likely to remain a computer-intensive endeavor (Baveye et al., 2018).

An alternative to the LBM approach, which may help alleviate some of its hefty computational demands and be very useful in applications, consists of approximating the pore space in soils in a piecewise fashion by a network of so-called “volume primitives.” These can be simple geometric shapes (balls, cylinders, cones, torus, generalized cylinders …) that may be transformed and combined at will to represent more complex geometries, and to which capillary theory can be applied to predict their filling or emptying. Pot et al. (2015) proposed a model, named MOSAIC, which uses balls as volume primitives. Pot et al. (2015) show that, when pores are allowed to drain regardless of the status of their neighbors, and when the critical diameter at which pores empty is adjusted to fit the experimentally measured degree of water saturation, MOSAIC produces a reasonable agreement between measured and predicted locations of air–water interfaces and that it does so extremely fast compared to the LBM approach. Pot et al. (2015) found that, on a similar type of computer (typical desktop computer with two 2.6 GHz processors), running the LBM program on the cases they considered takes around 19 h, whereas exactly the same cases take about 5 min when run with MOSAIC (Ngom et al., 2011, 2012; Monga et al., 2014).

Balls are among the simplest geometrical primitives that could be used to approximate the shape of interstices in soils. At least in principle, there is a possibility that other types of primitives, for example conical wedges, ellipsoids, tori, or generalized cylinders, would require a smaller number of individual objects in any given soil, and therefore would lead to a significantly shorter computational time to obtain a fit of the same quality to water retention data. Indeed, the speed of the algorithm determining water-air interfaces is roughly proportional to the number of primitives. Alternatively, with a number of objects similar to that of the spherical balls used in Pot et al. (2015), one could perhaps expect a much higher prediction accuracy than what these authors found. At this juncture, it is not clear which, if any, of these two perspectives is valid.

In this general context, the key objective of the research described in the present article was to assess the extent to which ellipsoids could advantageously replace the spherical primitives considered in the MOSAIC model by Pot et al. (2015). As far as we are aware, this is the first attempt to model pore space in soils using ellipsoids. In order to illustrate the relevance of ellipsoids-based pore space geometrical modeling, we show that this representation can be used for the determination of the three dimensional distribution of water and air in soil pores. Exactly the same cases are analyzed as in this earlier article, and the modelled results are compared with those obtained experimentally, with the LBM approach, or with the spherical primitives, respectively. A statistic, the mean absolute error, is used to compare the modelled predictions and measured data relative to the location of

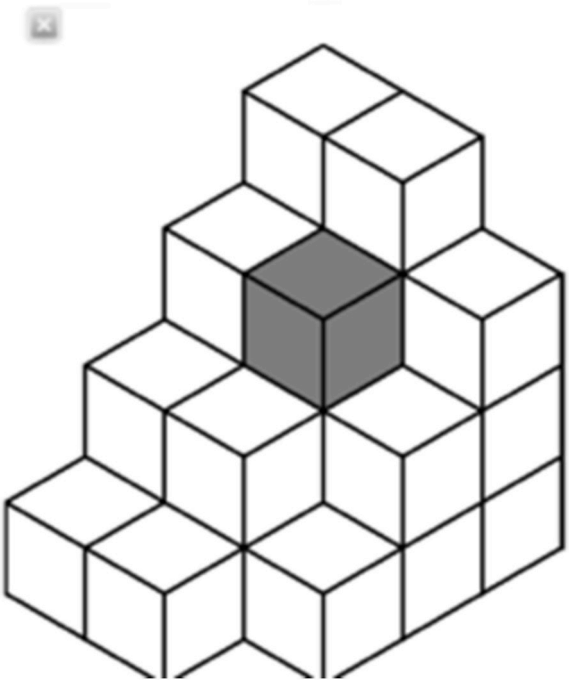

Fig. 1. Illustrative representation of voxels in 3D images of soils.

water-air menisci in the different cases.

2. Computational modeling

2.1. Geometrical pore space modeling: from voxels to ellipsoids

2.1.1. Context

The key advantage of primitive-based representations over voxel-based descriptions, like those considered in the Lattice-Boltzmann approach, is that the former explicitly associate shape descriptors to the pore space, and therefore do not require one to handle the prohibitively high numbers of voxels, typically in the tens of millions, that constitute this pore space. The situation is not unique to soils, but occurs in general whenever one needs to extract an intrinsic shape representation from raw sensor data (Alsallakh, 2014, Monga, 2007, Ayache, 2003, Banégas et al., 2001, Blinn, 1982; Caumon et al., 2004, 2005; Doube, 2015; Druoton et al., 2013; Malladi et al., 1995; Schlüter et al., 2010; Terzopoulos, 1991). The basic principle of most shape modeling methods consists of looking for piecewise approximations by analytic surface- or volume primitives. The representation by piecewise primitives is computed through the explicit, or sometimes implicit, optimization of a non-linear functional integrating an approximation error term and an antagonist scale term (Monga, 2007; Mumford, 1989).

In the case of natural shapes, 3D image segmentation remains a difficult and open problem due to the difficulty to find a suitable functional, as well as tractable minimization schemes. The reason is that natural shapes have not been designed as manufactured items using analytic surfaces and volumes. Then, in a certain sense, the algorithm has to invent a meaningful analytic representation of the shape initially described, typically, by dozen of millions of voxels. By meaningful, we mean that the shape representation should faithfully approximate the shape, be robust to small shape changes, have a kind of "continuity" with respect to initial description, be compact, and also be adapted to its ultimate use. Indeed, the current state of the art indicates that very little research has been devoted so far to the 3D shape modeling for complex natural volume shapes. This is due in part to the related subsequent mathematical and algorithmic difficulties that have to be faced, and also to the lack of motivation linked to specific application contexts. It is likely that the first point will remain true maybe for a few decades but will be more and more overcome by the increasing performance of computers. One could argue that the second hurdle is not going to last much longer, in view of the intensifying research effort in the area of natural complex systems modeling and especially for soil science applications (Shomar et al., 2014; Darwish et al., 2014; Vogel and Roth, 2001; Wen and Gomez Hernandez, 1996; Whiffin et al., 2007; Eickhorst and Tippkötter, 2008; Freije, 2015; Kuiper et al., 2015; Vogel, 2015; Baveye et al., 2018). For instance (Delerue and Perrier, 2002; Ngom et al., 2011, 2012), proposed to use balls, tori, and generalized cylinders to represent pore space from computed tomography images in order to simulate microbial decomposition in soil.

2.1.2. Problem formal statement

Let S be a volume shape defined by a set of voxels as follows:

Let I (i,j,k) be a 3D discrete binary image.

Let S be the volume defined by all voxels M(i,j,k) of I set to one:

$$M(i,j,k) \in S \Leftrightarrow I(i,j,k) = 1$$

In the continuous space, each voxel could be considered as a cube (see Fig. 1):

The goal consists of finding out piece wise approximations of shape S using typical primitives having compactness, robustness and invariance properties. The invariance properties are: invariance by rigid transformation of the initial data, invariance with respect to the course of data, etc. The robustness properties are linked to the stability of the representation to small changes. Of course, as in every approximation scheme, the scale notion should be explicitly expressed within this representation. Given that this representation will be used to simulate complex spatialized phenomena taking place inside the shape, deep-learning based descriptors cannot, at least directly, be applied to solve this problem.

Let V be the set of voxels defining the volume shape:

$$V = \{v_1,\ v_2,\ ...v_n\}$$

Let E be a set of ellipsoids such that each pair is either disjoint or tangent:

$$E = \{e_1, e_2, ...e_m\};\ \forall\ (i,j),\ (e_i \cap e_j = \varnothing)\ \vee\ (e_i \cap e_j\ \in S(E))$$

where S(E) is the set of regular surfaces of affine space E.

Following the basic principle described in Monga (2007), we define the following functional:

$$C(V,E) = (\Phi(VE) + \Phi(E\Theta V)) + \lambda \sum_{j=1}^{j=m} \rho(e_j)$$

where Φ is the function that associates to a volume shape the numerical value of its volume, is the function that associates to a surface shape the numerical value of its area, and λ defines the scale of the representation $C(V,E)$ (Monga, 2007, Mumford, 1989).

We look for sets of ellipsoids E minimizing $C(V,E)$. In order to address this optimization problem, we propose to use a "split and merge" strategy involving clustering and region-growing algorithms.

2.1.3. Optimal set of spherical balls

The first step in the morphological modeling of soil pores by ellipsoids consists of identifying the minimal set of spherical ρballs including the λ-skeleton of the shape. The shape is determined within the voxels associated with the pores in the thresholded)i.e., binary) CT images. As explained in detail in several previous publications (Monga et al., 2007; Ngom et al., 2012), this minimal set of balls is obtained after subjecting the image to several successive operations. First, voxels at the pore boundary are identified. They are defined as solid phase voxels that have at least one pore voxel in their immediate neighborhood. Second, once the boundary voxels are all known, a so-called "Delaunay triangulation" is computed, which by definition connects boundary voxels via tetrahedra with faces that are as equilateral as possible. Third, for each tetrahedron of the Delaunay triangulation, the center and radius of the circumscribed sphere are calculated. Fourth, one proceeds to the "pruning" of Delaunay spheres, which means that one evaluates the extent of overlap between a Delaunay sphere and the set of voxels forming the pore space, and one discards the spheres for which this overlap is less than an arbitrary number. Indeed, the numerical discretization implies that some maximal balls are not completely included within the shape. In order to select also these balls, we relaxed the constraint using a threshold set to 90% after trial and error. We found that the results were not sensitive to 3–5% threshold variations. Fifth, in principle the next step consists of extracting from all the spheres the so-called "λ-skeleton", defined as the subset of balls whose radius is at least λ (in units of voxel widths). But in the present

research, it is assumed for convenience that $\lambda = 1$, so that, practically, all balls are retained. Sixth, the spherical balls are ranked in decreasing order according to their radius. Seventh, starting with the largest ball and going down the list, the next ball in the list is considered, and it is discarded if this ball overlaps previous balls that were kept by more than a percentage of the value of the ball volume. In previous articles (Monga et al., 2007; Ngom et al., 2012), we chose this threshold in order to ensure that the shape skeleton is included within the union of the balls. In a preliminary study, we found that a global recovery percentage of 20% was fitting regarding the shape skeleton inclusion constraint adoted by Monga et al. (2007) and Ngom et al. (2012). This process is repeated until there are no more balls to consider, at which point one has obtained the smallest (or near smallest) set of balls, covering the skeleton of the pore space.

2.1.4. Clustering of balls

In applications in other contexts, several authors have developed statistical techniques to find connected and compact set of voxels (e.g., Mokhtari et al., 2014; Banégas et al., 2001; Baillard et al., 2000; Alsallakh, 2014). One such technique is a variant of Diday's dynamic clustering method, also known as the "k-means" algorithm. Basically, one searches for the best partition of a given cloud of points into k classes. The best partition maximizes the inter-class variance (the variance between distinct classes) and minimizes the sum of intra-class variances (the variance inside each class). In order to apply this technique to the minimal set of balls described earlier, each ball has to be assigned a weight, equal to its volume. Then, using for k a value equal to 30% of the total number of balls to cluster, the balls are divided in k groups, so as to minimize the sum, for all the balls, of the products (weight x distance^2), where "distance" is measured between the center of each ball and the barycenter of the group to which the ball belongs. The principle of setting k to a percentage of the total number of maximal balls is that the value of k should be linked to the complexity of the shape, which is related to the initial number of maximal balls. According to preliminary trials, we found that a value of 30% for k was reasonable. Also, the value of this percentage can influence the computing time, but less so the final result due to the region-growing step that follows.

Once the k groups are obtained, points on the surfaces of the balls in each group are selected randomly, and a Matlab algorithm is applied to these points to obtain the convex envelope of all the points in the group. The algorithm produces a set of triangular facets approximating the convex envelope. Then, trough minimization of algebraic distances (Bricault and Monga, 1997) one tries to determine the parameters of the ellipsoid that best fits the convex envelope in each group.

For each ellipsoid, the percentage of the volume of the ellipsoid that overlaps the set of pore voxels is calculated. If this percentage is less than a threshold of 90%, the ellipsoid is considered "invalid." Only "valid" ellipsoids are kept. The balls associated with invalid ellipsoids are once again pooled all together, and the k-means algorithm is applied to them, with the same k value (which now leads to a smaller number of groups since there are fewer balls to divide). This process of generating ellipsoids and pruning those that are invalid, is continued until it is no longer possible to come up with valid ellipsoids, at which stage the remaining balls are kept as they are. Indeed, this clustering stage is equivalent to segmenting the shape skeleton using the set of maximal balls computed through 3D Delaunay triangulation (George and Borouchaki, 1998). Afterwards, to each part of the skeleton partition, we attach a corresponding set of maximal balls. A key asset of this method is that it provides ellipsoids whose axes fit by construction with the skeleton of the shape and also that are sticking to the shape boundaries. These strong geometrical properties explain why the drainage simulation can perform well. Indeed, the goal is to obtain a pore space representation that provides a good approximation of the volume shape defined by the segmented image.

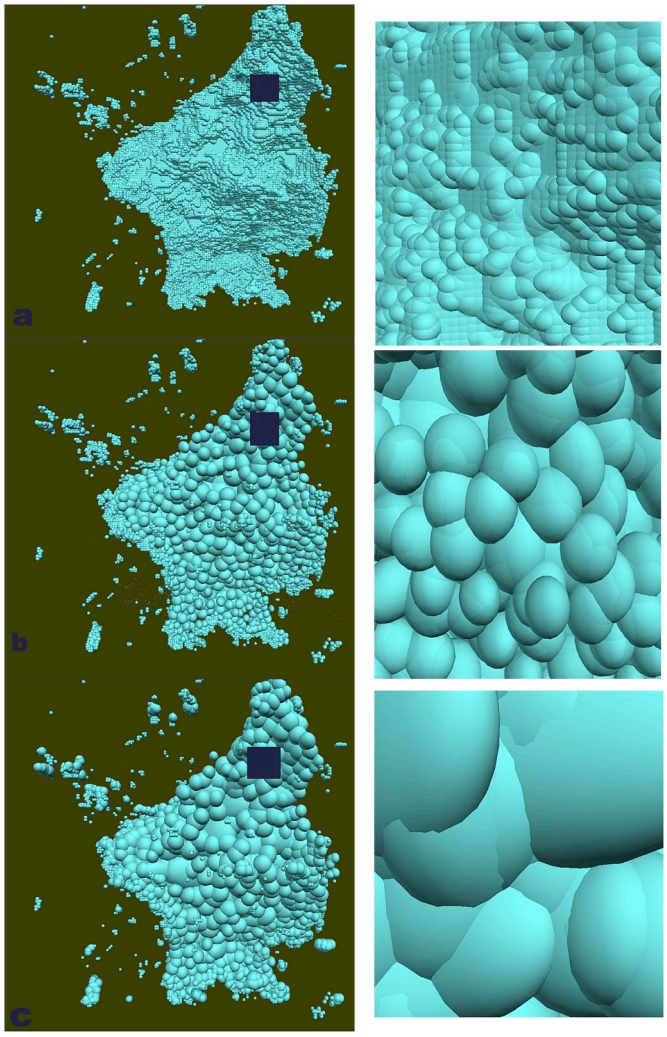

Fig. 2. Comparison of ellipsoids-versus ball-based representations of the pore space of ROI p1bkko4a. The first row shows the ball-based representation of the pore space. The second row shows primary ellipsoids resulting from the application of the split-merge algorithm, and the third row shows final ellipsoids produced by the region-growing algorithm. The images on the right are magnified versions of the regions under the squares in the images on the left.

2.1.5. Merging of ellipsoids

The above procedure yields a sometimes large number of ellipsoids, which can overlap slightly with neighboring ones, although the design of the method makes this overlap relatively small, as a rule. Nevertheless, since there are still many ellipsoids, it is useful to find out if their number cannot be reduced, to eventually speed up calculations made with them. Merging of ellipsoids is done by systematically looking at pairs of neighboring ellipsoids, and determining if a new ellipsoid obtained on the basis of all the balls associated with the two

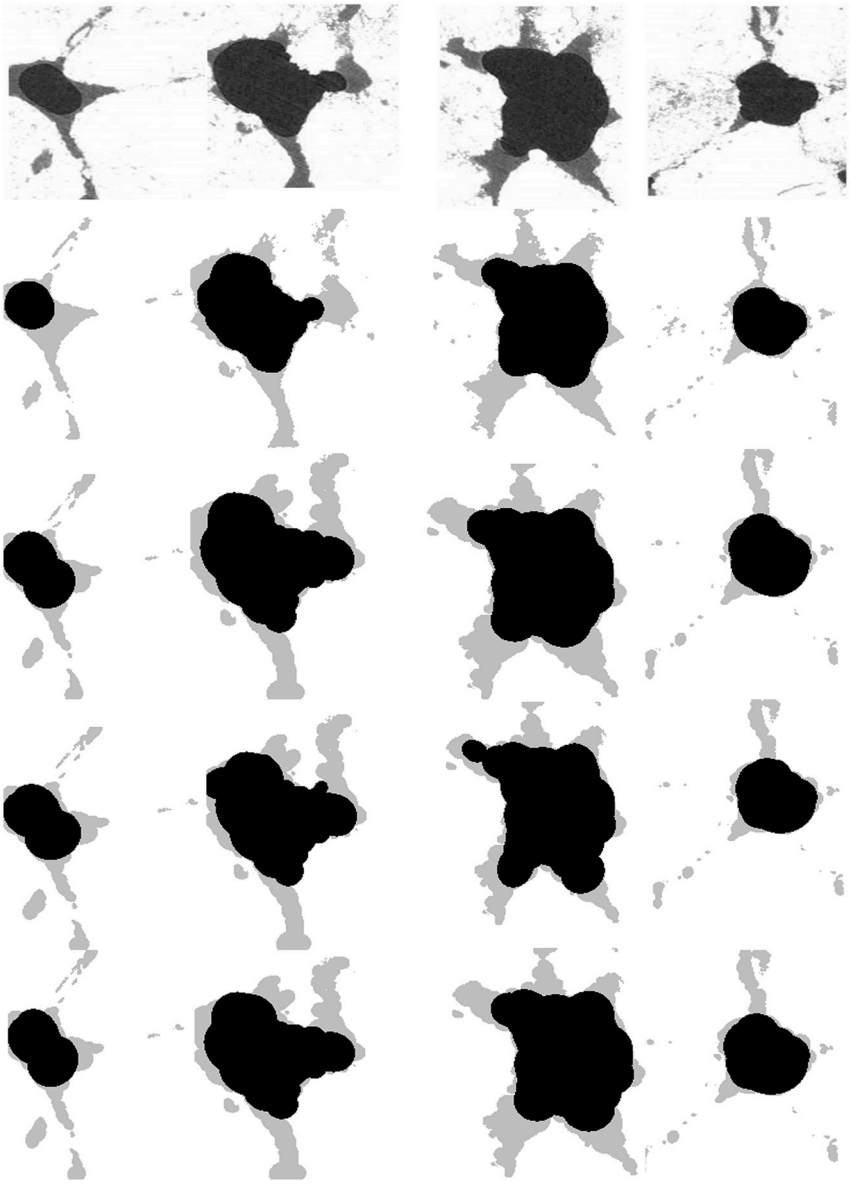

Fig. 3. Comparison between segmented SR-μCT data (first row) and model predictions of the air-water distribution in the connected pore space of ROI p1bkk04a at different vertical sections z = 22,64, 106, and 148 (from left to right). The water saturation index of the pore is 0.45. Solid phase is in white, water is in light gray, and air is in black. The second row corresponds to predictions using spherical balls, whereas the third, fourth, and fifth result from modeling with ellipsoids defined on the basis, respectively, of the minimal radius (3rd row), maximal radius (4th row), and mean radius (5th row). To drain ellipsoids, the threshold diameter is taken as 300 μm according to Young-Laplace equation.

individual ellipsoids is still a "valid" ellipsoid, in the same sense as above. Each valid ellipsoid that is identified in this manner is kept, and the systematic process of reviewing ellipsoids and merging them is continued until it is no longer possible to do further merging. This merging stage allows to provide a final result in the sense of the segmentation completeness.

In order to optimize this merging stage, we use a data structure described by Faugeras and Hebert (1984) and Monga (1987), which allows the iterative merging of the best pair of ellipsoids with respect to the fitting error. In a first step, we compute the adjacency graph of the ellipsoids obtained at the previous step. We determine the connectivity by means of the corresponding set of balls associated with each ellipsoid. Then, for each pair of connected ellipsoids, we sample the points of the union of the set of balls. We compute the convex hull and then the best fitting ellipsoid. We calculate the ratio between the ellipsoid volume and the summation of the volumes of the two corresponding ellipsoids. In order to increase the precision of the error evaluation, we can also come back to the initial set of voxels as previously. We take the minimum between the ratio of the volume of the ellipsoid divided by summation of the two initial ellipsoids, and the inverse. The merging cost is set to this ratio which is also used to calculate the criterion defining the merging predicate (see Monga, 1987). In annex is included a detailed description of the implemented algorithms.

2.2. *Draining from pore space morphological representation: from ellipsoids to air-water interfaces*

This section deals with the computation of air-water interfaces on the basis of the ellipsoids-based representation of the pore space.

Initially, all ellipsoids (see sub-section 2.1) are filled by water. The

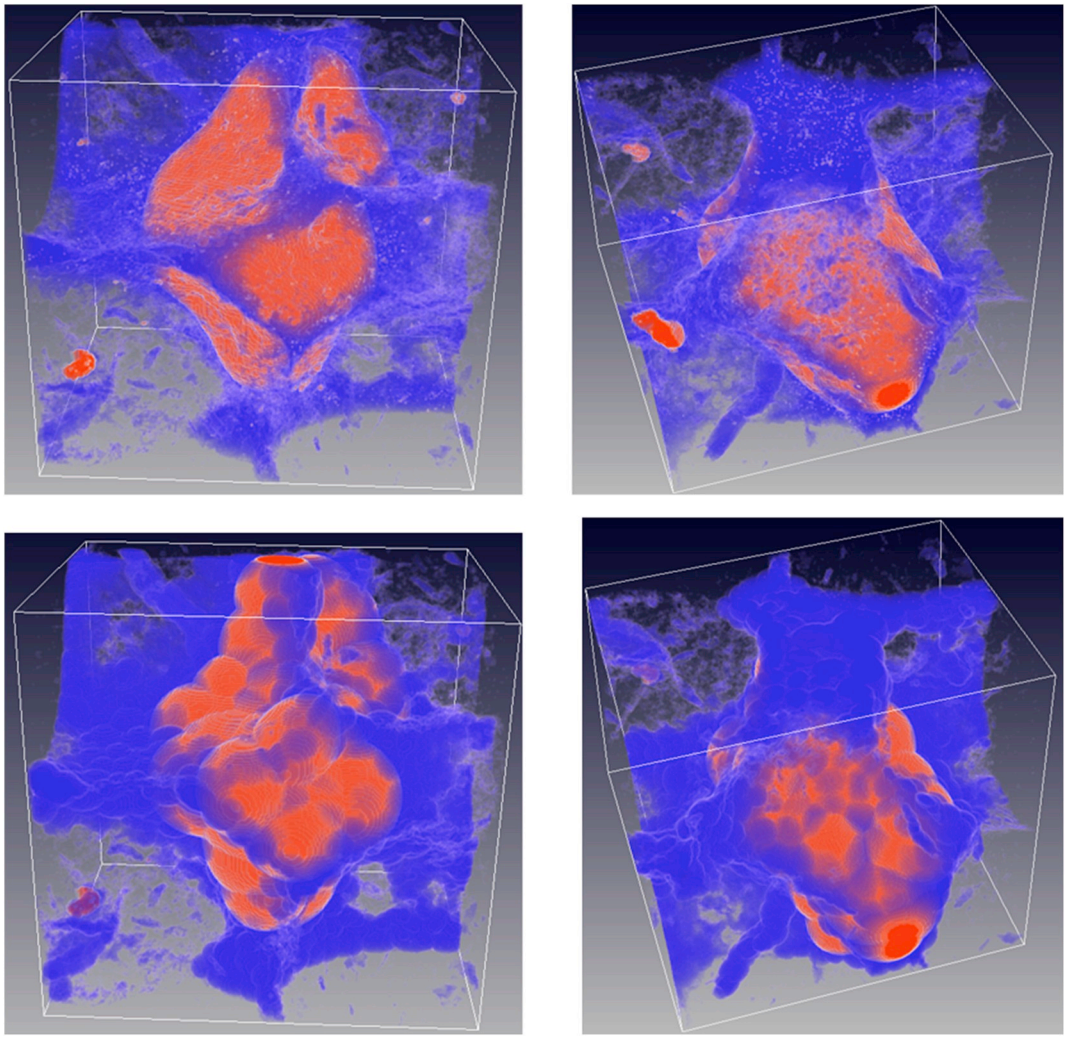

Fig. 4. Region of Interest p1bkk04a viewed from two different perspectives, presented side-by-side. The solid phase is not represented, whereas water is in blue and air is in red. The images on top correspond to the segmented SR-μCT images, and the bottom images show ellipsoids-based predictions, using the minimal axis and a threshold of 300 μm to evaluate their saturation level. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

final water content distribution is simulated by an iterative drainage procedure, following Pot et al. (2015). In this scheme, the Young–Laplace equation gives the equivalent maximum radius, r_{eq} $[m]$, of the pores that are filled with water at a fixed matric potential, ψ $[kPa]$:

$$\psi = 2\sigma cos(\theta) \,/\, r_{\mathrm{eq}} \tag{1}$$

Where σ is the surface tension of water in contact with air $[\mathrm{Nm}-1]$, θ is the contact angle of water on the solid surface, assumed in this study to be zero for a fully wetting fluid.

The conditions under which a given ellipsoid is allowed to empty are similar to those adopted in Pot et al. (2015), in the sense that an ellipsoid is allowed to empty even if none of the neighboring ellipsoids are empty. Like with the description of drainage using the ball-based representation (Monga et al., 2014), we can either fill with air or water the primitives (balls, ellipsoids, cylinders etc.) only according to their geometrical features, or take into account the adjacency relationships. In the first case, a ball will empty only if its radius is larger than a given value given by the Young-Laplace equation. In the second case, a ball will empty if its radius is larger than a given threshold and also if it is not surrounded by balls whose radius is smaller than the threshold. However, since unlike balls, an ellipsoid has 3 axes of symmetry, one has to decide which “radius” along which axis is used as a criterion to determine if the ellipsoid empties. Options are the smallest radius of the three, the largest of the three, and an average of all three. In the following, all 3 options are investigated. Eventually, when all the ellipsoids have been considered using a uniform rule for all of them, the model accounts for the spatial distribution of water in the soil, which can be directly compared to the air-water distribution observed in the segmented 3D SR-μCT images, and also to the model predictions obtained by Pot et al. (2015) with a lattice-Boltzmann approach. Of course the above statement requires that the same physical equations be followed.

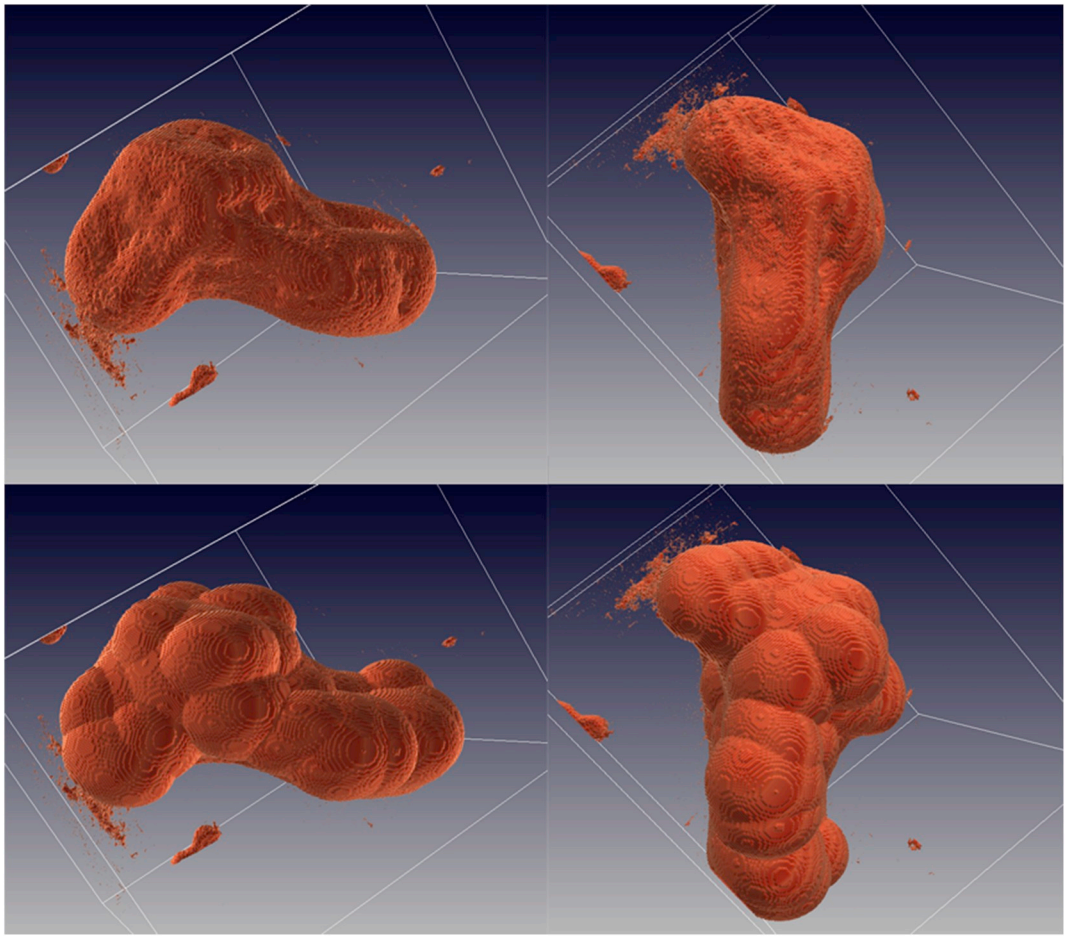

Fig. 5. Air phase in Region of Interest p1bkk04a viewed from two different perspectives, presented side-by-side. The images on top correspond to the segmented SR-μCT images, and the bottom images show the predictions using ellipsoids, using the minimal axis and a threshold of 300 μm to evaluate their saturation level.

3. Soil data set

The different steps leading to the 3D images used in the research recounted in this paper (Sections 3.1 to 3.4) have been described in detail by Pot et al. (2015). A shortened account is given here, in order for the text to be reasonably self-contained.

3.1. Soil preparation

Surface soil samples obtained in Versailles (France), with 17% clay, 56% silt and 27% sand, were passed through a 2 mm sieve. Soil aggregates between 2 and 3.15 mm in size were collected from the coarse fraction and their gravimetric moisture content was adjusted at 0.205 kg kg^{-1}. They were then packed uniformly at a bulk density of 1200 kg m^{-3} into cylinders, 50 mm in diameter and 40 mm high. The resulting soil columns were subsequently sawed to about 8 mm thick slices. Nine $6 \times 6 \times 8$ mm^3 soil cubes were sampled out of each slice with the help of a razor blade in order to minimize disturbance during cutting.

3.2. Equilibration at set matric potential

With an especially-developed, miniaturized experimental suction device, the small soil cubes were equilibrated at fixed matric potentials close to water saturation. The soil cubes were first equilibrated at a matric potential of −2 kPa for two days using de-aired water. Then, all soil cubes were slowly saturated and were kept equilibrated at 0 kPa for 1 h in order to re-move residual entrapped air. Then the syringes, in series of three, were directly equilibrated at specific matric potentials of −0.5 kPa, −1 kPa, and −2 kPa, respectively, for 24 h.

3.3. Synchrotron X-ray μCT (SR-μCT) scans

Microtomography measurements were carried out at the SR-μCT facility operated by the GKSS Research Center at HASYLAB (Hamburger Synchrotron Strahlungslabor) in Hamburg, Germany. An X-ray energy of 30 keV and rotation of the CCD camera in 0.2° steps from 0 to 180° yielded a voxel resolution (after reconstruction) of 4.6 μm.

3.4. Image pre-processing and analysis

Reconstructed image datasets were preprocessed in several ways.

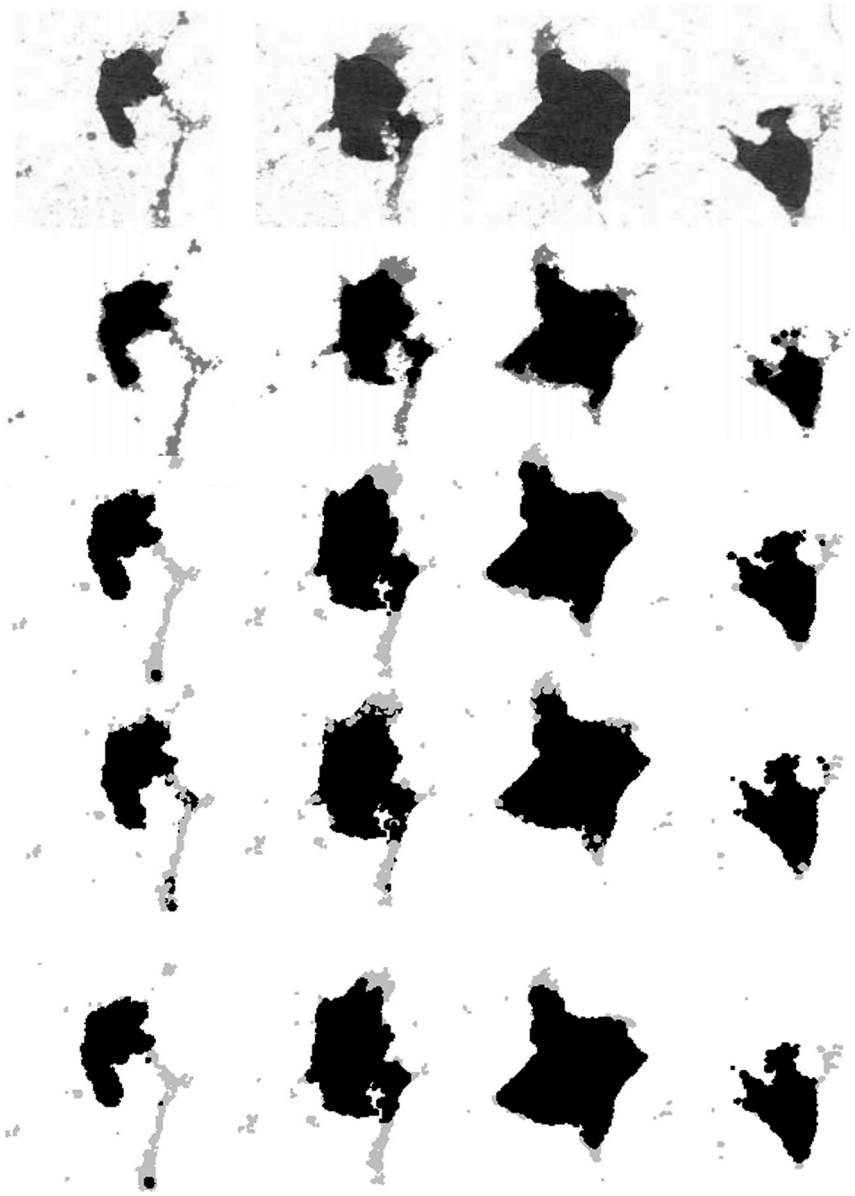

Fig. 6. Comparison between segmented SR-μCT data (first row) and model predictions of the air-water distribution in the connected pore space of ROI p1bkk03a at different vertical sections z = 18, 24, 34, and 65 (from left to right). The water saturation index of the pore is 0.20. Solid phase is in white, water is in light gray, and air is in black. The second row corresponds to predictions using spherical balls, whereas the third, fourth, and fifth result from modeling with ellipsoids defined on the basis, respectively, of the minimal radius (3rd row), maximal radius (4th row), and mean radius (5th row). To drain ellipsoids, the threshold diameter is taken as 73.6 μm according to Pot et al., (2015).

First, inscribed cubic subvolumes, of size 900 × 900 × 500 voxels, were extracted to avoid boundary effects at the edges of the sample. To facilitate computations mainly for LBM method, the resolution of the images was coarsened by a factor of two, to 9.2 μm. A rank order median filter was applied to reduce image noise. To obtain a ternary image, a global threshold value was automatically computed according to Otsu (1979) from the gray value histogram by maximizing the intra-class variance of the matrix and pore classes while minimizing the gray value variances within the two classes. The threshold to segment the water- and gas-filled pore regions was selected at the minimum between the two distinct peaks.

The relative fractions of the three phases (gas, water, matrix) were subsequently computed regarding the number of voxels attached to each phase. Porosity was computed as the sum of the relative fractions of gas and water and the water saturation index, Sw, was equal to the ratio of the relative fraction of water over porosity. Air–water triangulated isosurfaces were calculated by the Isocontour tool of ParaView (hrens et al., 2005) from the extracted 3D data including the segmented air phase of the images. The surface area was then simply computed by integration and translated into physical units of mm^2 via multiplication with the squared voxel resolution.

Houston et al. (2013) showed that different pre-treatments on the gray-scale image and different segmentation methods can provide different segmented images. Therefore, the outcomes of modeling are still dependent on the chosen method. Although recommendations are sometimes provided (Vaz et al., 2011; Houston et al., 2013), there is still a lack of standardization in the acquisition and processing of soil images.

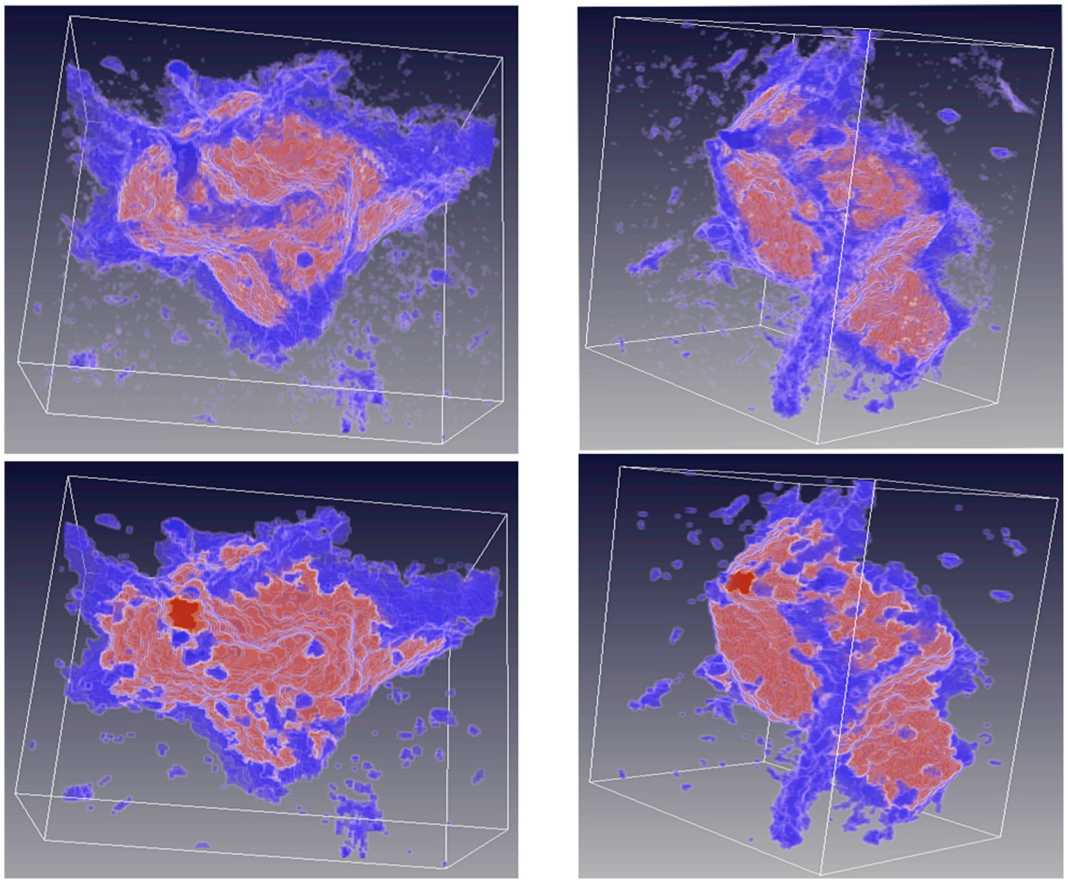

Fig. 7. Region of Interest p1bkk03a viewed from two different perspectives, presented side-by-side. The solid phase is not represented, whereas water is in blue and air is in red. The images on top correspond to the segmented SR-μCT images, and the bottom images show ellipsoids-based predictions, using the minimal axis and a threshold of 73.6 μm to evaluate their saturation level. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

4. Assessing model predictions of air-water interfaces

One goal of the present work was to assess whether a morphological model using ellipsoids predicted more accurately and/or faster than a model based on spherical balls the distribution of water in the various single pores analyzed by Pot et al. (2015). Therefore, exactly the same cases, associated with one cube each out of the series at −2 kPa and −1 kPa (cubes bkk03a and bkk04a, respectively) were considered in the research described here. Practically, we used the first data set to calibrate optimally the parameters and the second one to validate.

To evaluate the spatial agreement of the position of the menisci between simulations and measurements, the mean absolute error, MAE, was computed as follows:

$$MAE = (\Sigma_{i\text{-}1}{}^{n} |s_i\text{-}m_i|)/n \quad (2)$$

Where the summation is taken over all pore voxels n, and s_i and m_i are the thresholded densities of the voxels for the simulated and measured images respectively. Indeed, we attach to each ellipsoid the corresponding set of voxels in order to calculate MAE.

5. Results and discussion

The description of the results and their discussion in the following follows closely section 3.2. in Pot et al. (2015), dealing with the pore-scale modeling of air-water interfaces. Figs. 2–11 and Tables 1 and 2 illustrate the estimation of air-water interfaces using ellipsoids. The experimental data are the same as in Pot et al. (2015).

The computation of MAE associated with drainage using the Young-Laplace equation for sample p1bkk04a at matric potential of −1 kpa, and for sample p1bkk03a and p4bkk03a ata matric potential of −2 kpa (Table 1) has been carried out with different functions of the three ellipsoids radii. Indeed, at least for our data sets, the results (MAE values) are not very sensitive to the choice of function (minimal radius, mean radii, …). A plausible reason for this observation is that the radii of the ellipsoids, for these specific data, are relatively homogeneous. However, we recommend to choose the minimal radius because the Young-Laplace equation has been designed for cylinders, and when an ellipsoid has an elongated shape, its minimal radius is a good approximation to the radius of an equivalent cylinder.

Compared to the ball-based approach described by Pot et al. (2015), there is a clear gain of the ellipsoids-based approach in terms of computing time. Insofar as the gain in quality is concerned, one way to quantify it is via the ratio of MAEs, specifically the ratio of the MAE

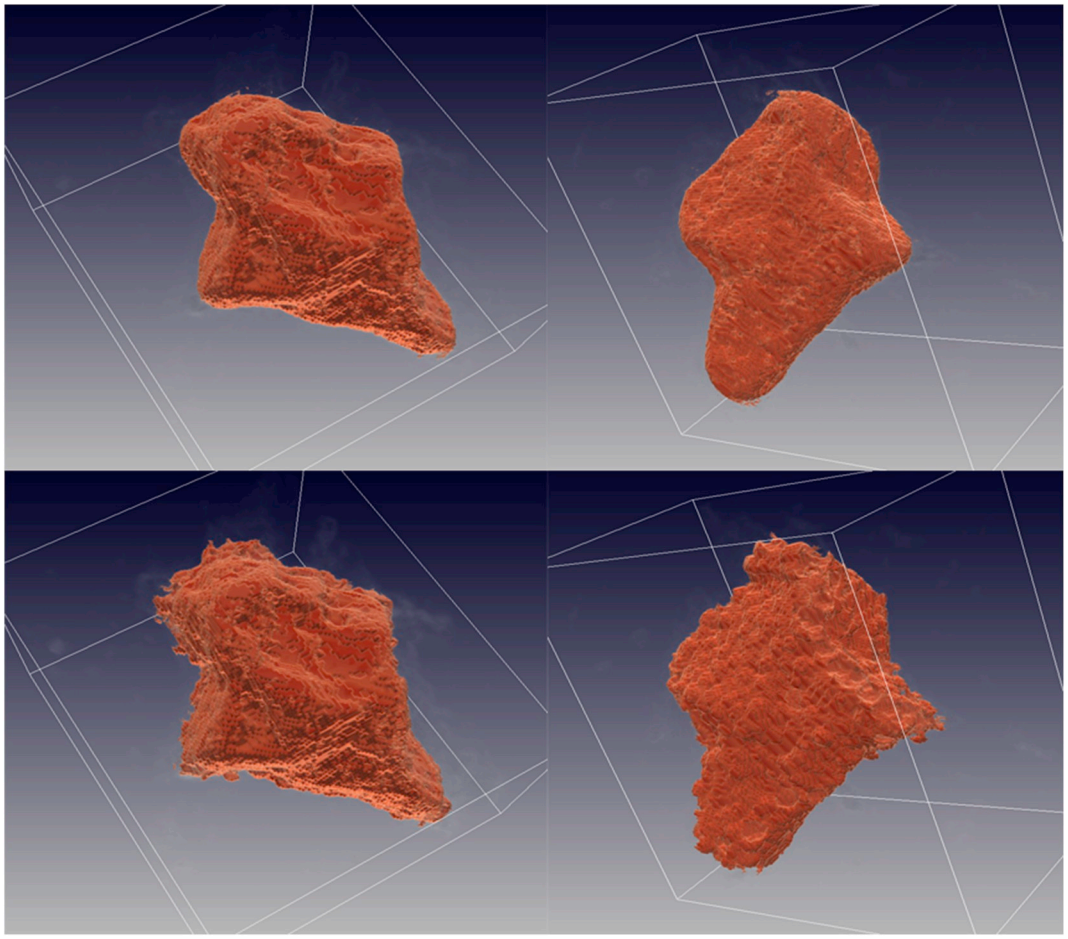

Fig. 8. Air phase in Region of Interest p1bkk03a viewed from two different perspectives, presented side-by-side. The images on top correspond to the segmented SR-μCT images, and the bottom images show the predictions using ellipsoids, using the minimal axis and a threshold of 73.6 μm to evaluate their saturation level.

obtained with ball-based representation to the MAE obtained with an ellipsoid-based case. For the ROI p1bkk04a, there is a two-fold decrease of the MAE obtained with the maximal balls, and a larger 5.75 ratio with the minimal set of maximal balls. For ROI p1bkk03a and ROI p4bkk03a, the values of MAE are comparable (in the case of R_{min}) if one considers the ball-based representation with maximal balls, but the ratio of MAEs is around 2 relative to the minimal set of maximal balls.

It is useful to try to estimate fitting errors when approximating the pore space by means of the set of ellipsoids and by means of balls (Table 2). There are two different error measurements: ρ_1 and ρ_2, respectively, quantify the is the fitting error linked to the percentage of the union of the ellipsoids resp. balls not included within pore space. The ratio ρ_1 is smaller for ellipsoids than for balls and ρ_2 is smaller for balls than for ellipsoids. This is due to the fact that we computed ellipsoids using the convex hull of the set of corresponding balls and therefore somewhat overestimated the pore space. Indeed, as noticed in Pot et al. (2015), the ball-based coverage underestimates the porosity. The use of ellipsoids corrects this problem but increases a little bit the porosity. Also, in Table 2, to parallel what was done in Pot et al. (2015), we have also considered all maximal balls, which have strong recovery, and not only the minimal set of balls recovering the skeleton as in Monga et al. (2014). The reason for this is that for applications to drainage, the results are much more precise by taking into account the set of all maximal balls.

For ROI p1bkk04a, the balls- and ellipsoids representations (Fig. 2) lead to model predictions of the air water distribution depicted in Figs. 3–5. It should be stressed that the balls-based representation used here, like in Pot et al. (2015), includes all maximal balls of the pore space shape. Indeed, to simulate drainage, we consider the union of all maximal balls whose radius is less than the threshold computed thanks to the Young-Laplace equation. Given that the maximal balls intersect strongly, they do not define at all a piecewise approximation of pore space. On the other hand, the ellipsoids, by construction, have very weak intersections, and can be construed as providing a piecewise approximation of the pore space. Therefore, the ellipsoids-based representation could be used for many other applications such as flow simulation (e.g., diffusion), biological dynamics simulation, pore space classification etc. Indeed, within the context of water retention, a more valid comparison of the balls-based representation of pore space versus the ellipsoids one, would have been to use not all maximal balls, but only the minimal set of balls recovering the skeleton. This set of balls defines a piece wise approximation of the pore space (Monga et al., 2007; Ngom et al., 2012), and has been used successfully to simulate microbial decomposition of organic matter (Monga et al., 2014). Then,

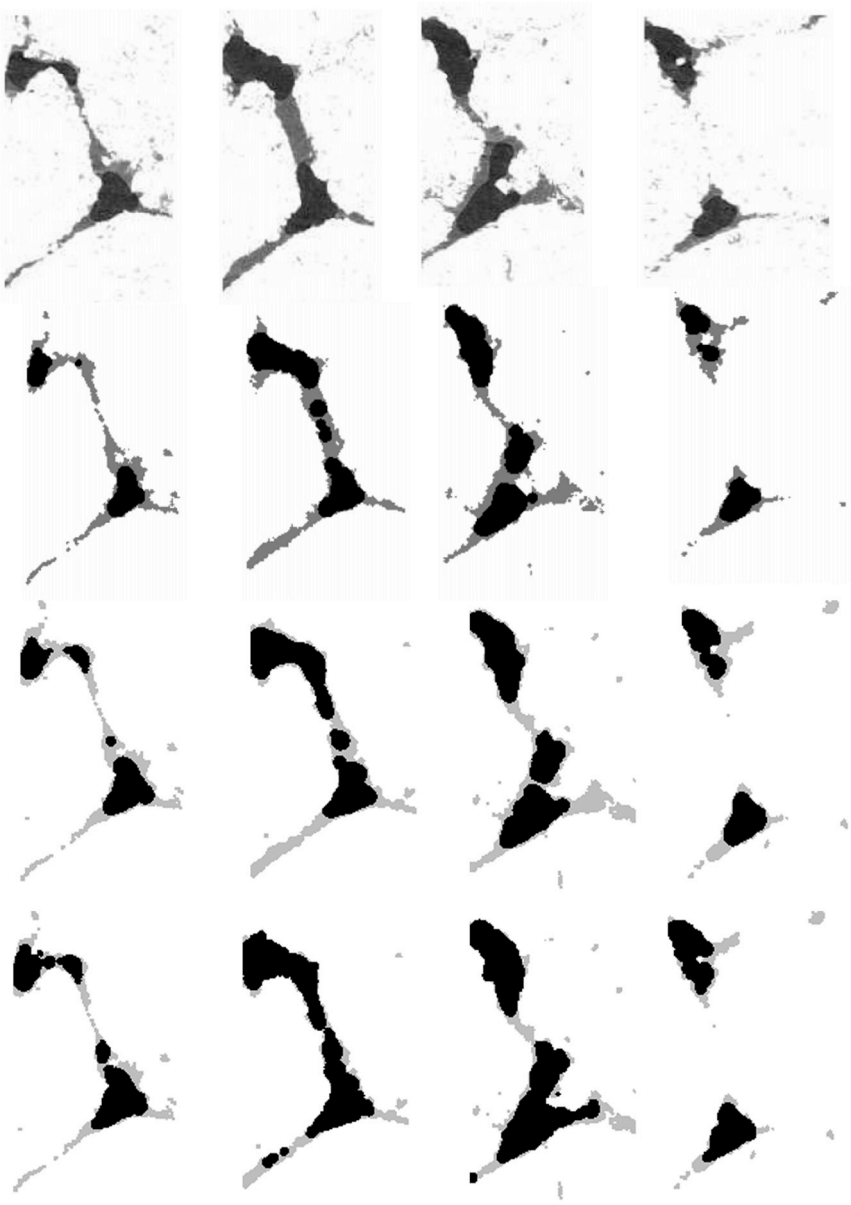

Fig. 9. Comparison between segmented SR-μCT data (first row) and model predictions of the air-water distribution in the connected pore space of ROI p4bkk03a at different vertical sections z = 30,39,54, and 63 (from left to right). The water saturation index of the pore is 0.20. Solid phase is in white, water is in light gray, and air is in black. The second row corresponds to predictions using spherical balls, whereas the third and fourth result from modeling with ellipsoids defined on the basis, respectively, of the minimal radius (3rd row), and maximal radius (4th row). To drain ellipsoids, the threshold diameter is taken as 112 μm according to Pot et al., (2015).

initially we had also computed the air-water interfaces using this minimal set of balls. Indeed, due to the fact that this schematic representation does not include a significant part of pore space, the water retention results are approximate. That is the reason why in the work described in Pot et al. (2015), we use the set of all maximal balls to simulate draining phase. Another way for comparing ellipsoids and balls-based approach for drainage simulation would be to compute all maximal ellipsoids of the shape and then to implement the same strategy than the one used with the set of all maximal balls. Unfortunately, a method has yet to be developed that allows one, with reasonable computing time, to determine all maximal ellipsoids of a 3D volume shape.

The visualizations of the water retention using the ellipsoids-based method for ROI p1bkk03a (Figs. 6–8) and for ROI p4bkk03a (Figs. 9–11), respectively, are in line with the quantitative performance evaluation given by the calculation of the MAE values (in Ttable 1). Indeed, the ellipsoids-based method provides air-water interfaces fitting fairly well with the experimental ones except in some cases where ellipsoids do not perform better than balls (Fig. 9). The predicted air-water interfaces kept the same globulous shape than with the balls (see Figs. 3, 5, 6 and 10).

Of course, globally, one would hope that the ellispiods-based representations could approximate experimental data better than they do, but nevertheless one could argue that they afford a reasonable estimation of the air-water interfaces in a much shorter computing time (15–30 min using a regular PC) than with the lattice-Boltzmann approach (a few days). The key explanation for this is that the ellipsoids-based method uses a more sophisticated representation of pore space geometry than is provided by the original set of pore voxels. The initial computation of the ellipsoids-based pore space representation takes

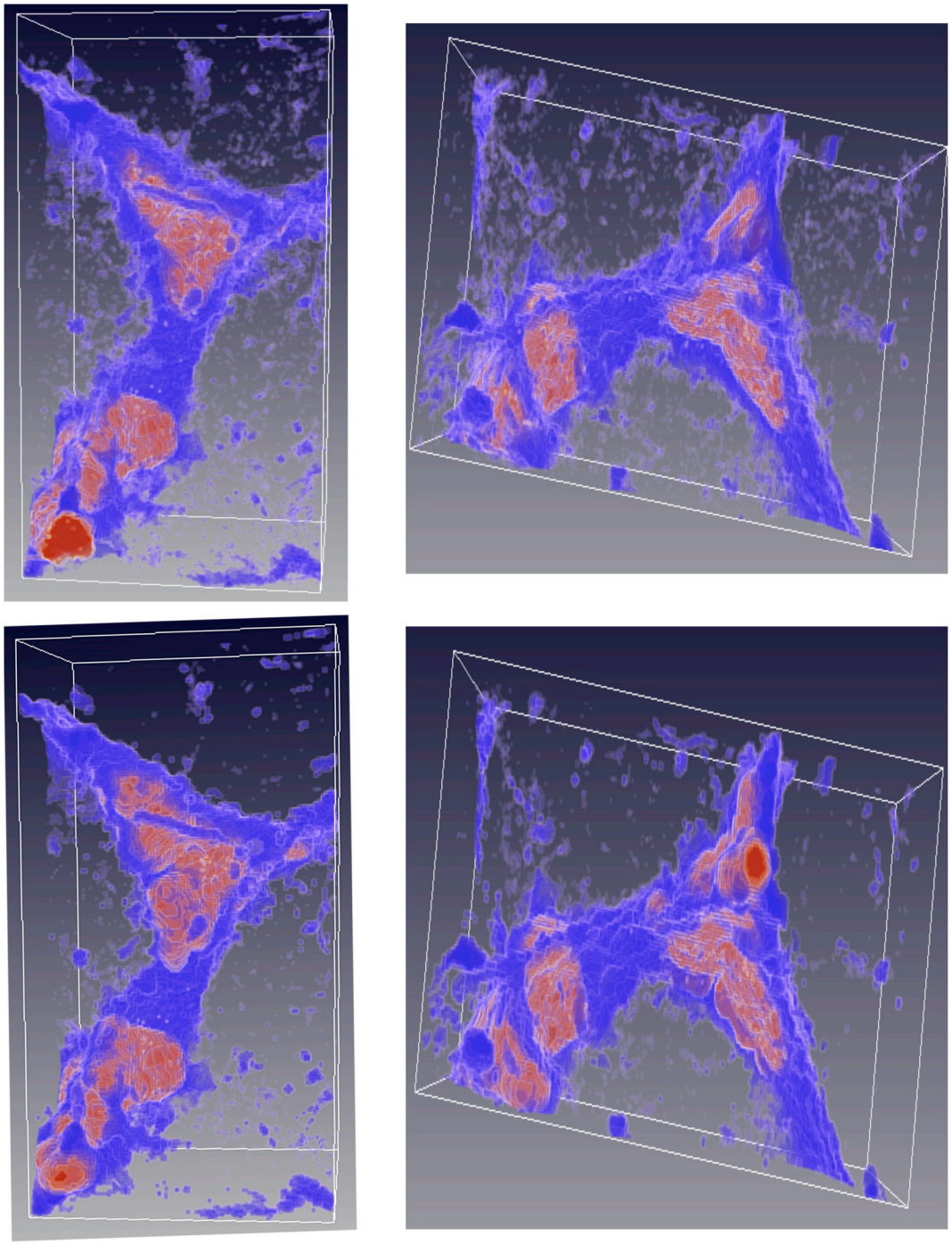

Fig. 10. Region of Interest p4bkk03a viewed from two different perspectives, presented side-by-side. The solid phase is not represented, whereas water is in blue and air is in red. The images on top correspond to the segmented SR-μCT images, and the bottom images show ellipsoids-based predictions, using the minimal axis and a threshold of 112 μm to evaluate their saturation level. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

between 15 and 25 min CPU time using Matlab coding (see Table 3). Especially for this specific task, requiring sophisticated data structures, an implementation in C or C++ would run much faster.

6. Conclusions

The key results of the research described in this article show that ellipsoids provide a better approximation of the pore space and of its level of saturation with water than does a representation with spherical geometric primitives (balls). In one case (ROI p1bkk04a), the value of the mean absolute error associated with predictions of air-water interfaces obtained with ellipsoids even approaches that obtained with the lattice-Boltzmann model. This is accompanied in all cases considered by a slightly better than 6-fold decrease in the time required for the computations, relative to the original MOSAIC model, using spheres, which increases still the computational advantage of models based on

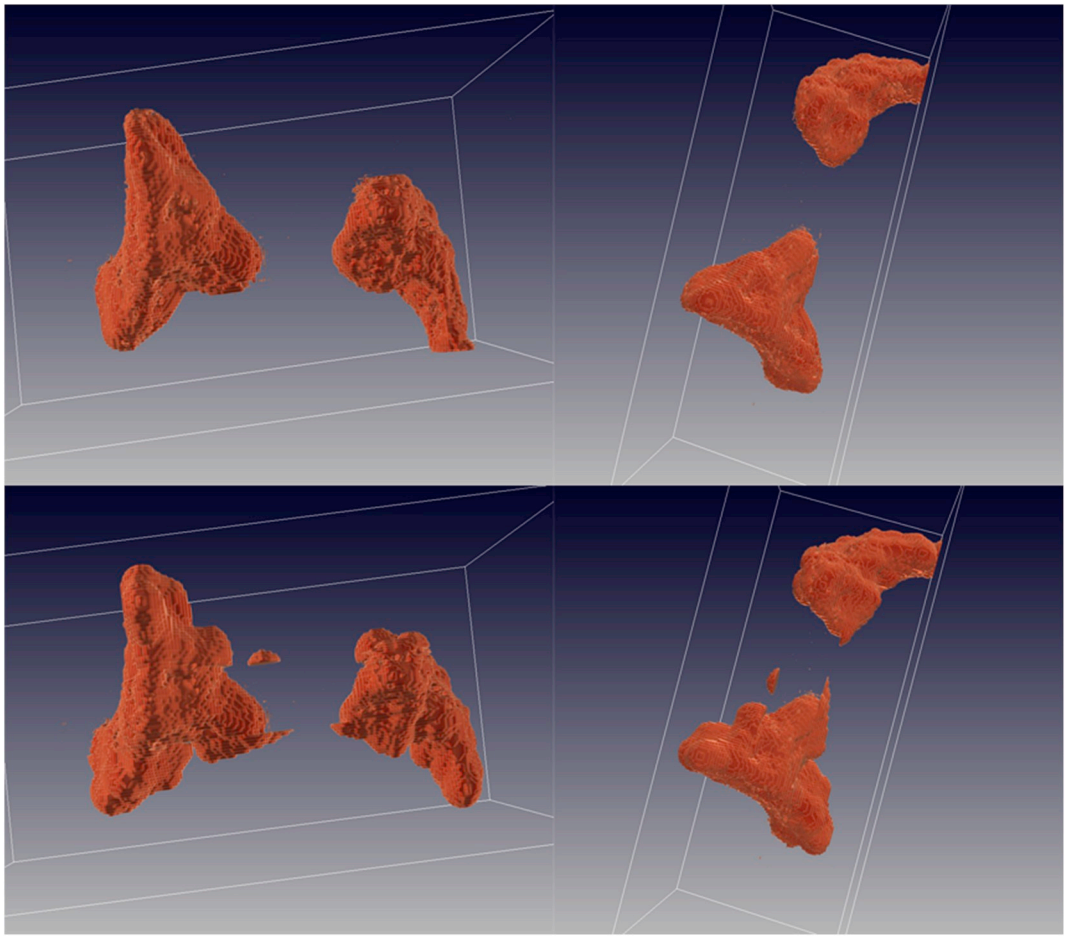

Fig. 11. Air phase in Region of Interest p4bkk03a viewed from two different perspectives, presented side-by-side. The images on top correspond to the segmented SR-μCT images, and the bottom images show the predictions using ellipsoids, using the minimal axis and a threshold of 112 μm to evaluate their saturation level.

Table 1
Values of the mean absolute error (MAE) between the segmented SR-μCT image and the geometrical representations (set of maximal balls and ellipsoids) and of the computation time (in seconds) for the different simulation cases. The diameter threshold was set according to Young-Laplace Equation for ROI p1bkk04a and according to Pot et al. for the two other data sets.

	Balls: minimal set of maximal balls recovering the skeleton	Balls: set of all maximal balls	Ellipsoids (Rmin)	Ellipsoids (Rmean)	Ellipsoids (Rmax)
ROI p1bkk04a					
MAE	0.46	0.16	0.08	0.08	0.08
Quality gain (MAE ratio) versus set of all maximal balls			2.00	2.00	2.00
Quality gain (MAE ratio) versus minimal set of maximal balls recovering the skeleton			5.75	5.75	5.75
Computation time (s)		12409	1933	1933	1934
Time gain (Computing time ratio)			6.41	6.42	6.42
Diameter threshold according to the Young-Laplace equation		300 μm	300 μm	300 μm	300 μm
ROI p1bkk03a					
MAE	0.41	0.18	0.17	0.21	0.20
Quality gain (MAE ratio) versus set all maximal balls			1.06	0.86	0.9
Quality gain (MAE ratio) versus minimal set of maximal balls recovering the skeleton			2.41	1.95	2.05
Computation time (s)		5476	879	816	820
Time gain (Computing time ratio)			6.20	6.70	6.68
Diameter threshold according to Pot et al., (2015)	73.6 μm	73.6 μm	73.6 μm	73.6 μm	73.6 μm
ROI p4bkk03a					
MAE	0.56	0.30	0.31	0.31	0.38
Quality gain (MAE ratio) versus set of all maximal balls			0.97	0.97	0.8
Quality gain (MAE ratio) versus minimal set of maximal balls recovering the skeleton			1.80	1.80	1.47
Computation time (s)		6765	1166	1115	1112
Time gain (Computing time ratio)			5.80	6.06	6.08
Diameter threshold according to Pot et al., (2015)	112 μm	112 μm	112 μm	112 μm	112 μm

Table 2
Pore space approximation errors between the segmented SR-μCT image and the geometrical models: $\rho 1 = \frac{V_{esp}}{V_{sol}}$ and $\rho 2 = \frac{V_{prim}}{V_{sol}}$, where V_{esp} is the number of pore space voxels of the thresholded SR-μCT images not included within the primitives, V_{prim} is the number of primitive voxels that are not included within the pore space of the thresholded SR-μCT images, and V_{sol} is the total number of pore space voxels of the thresholded SR-μCT images.

	Voxels	Balls: minimal set of maximal balls recovering skeleton	Balls: set of all maximal balls	First ellipsoids	Final ellipsoids
p1bkk04a					
	0	0.2442	0.0607	0.0582	0.0670
	0	0.0226	0.0773	0.1305	0.1467
Number of primitives	2079943	26493	151797		4072
p1bkk03a					
	0	0.2592	0.0865	0.0318	0.0239
	0	0.0215	0.0531	0.1140	0.1258
Number of primitives	280338	4941	125984		2011
p4bkk03a					
	0	0.4033	0.1190	0.0875	0.0968
	0	0.0479	0.0628	0.1502	0.1664
Number of primitives	219353	5257	134417		2505

Table 3
Computing times for providing ellipsoids; C language implementation would certainly decrease a lot the computing times compared to Matlab code.

	ROI p1bkk04a	ROI p1bkk03a	ROI p4bkk03a
Total CPU time for providing ellipsoids (Matlab code)	1370 s.	985 s.	1017 s.

geometric primitives relative to the high time-consuming lattice-Boltzmann approach. In addition, the shape of the air-water interfaces obtained with ellipsoids is less bulbous than that obtained with balls, and is much closer to what is found in the segmented synchrotron data. These different observations suggest that ellipsoid-based representations of the pore space of soils could serve as very valuable substitutes to ball-based representations. Although there is still a significant gap in terms of accuracy of predictions with the lattice-Bolztmann approach, ellipsoid-based models should prove useful in applications where computational time is an important constraint. This will be the case for any attempt to increase the size of soil samples being simulated, or when upscaling the description of microscale soil processes.

CRediT authorship contribution statement

Alain Trésor Kemgue: Methodology. **Olivier Monga:** Supervision. **Serge Moto:** Supervision. **Valérie Pot:** Data curation. **Patricia Garnier:** Writing-original draft. **Philippe C. Baveye:** Writing-original draft. **Abdelaziz Bouras:** Writing – review & editing.

Acknowledgments

The research described in this article was made possible through NPRP grant #9-390-1-088 from the Qatar National Research Fund, and in part also through a grant from the Agence Nationale de la Recherche (ANR, France) to project Soilμ3D ANR-15-CE01-0006-01.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cageo.2018.11.006.

Annex detailed algorithms

The corresponding code can be downloaded at:
https://github.com/kemgue/toellipsoids

Summary

We describe here in detail the algorithms implemented. The fitting process of the porous media by a minimal set of ellipsoids is done in two steps. The first step called initial segmentation uses either Algorithm 1 or Algorithm 2 described below and the second step called region growth or merging uses Algorithm 3 also mentioned. The ellipsoids obtained at the end of process are also called regions.

Initial segmentation:

Algorithm 1: Division-merge using clusters (k-means based)
Algorithm 2: Division-merge using octree data structure (octree based)

Final segmentation:

Algorithm 3: region growth or merging

Algorithm 1. Division-merge using clusters (k-means based)

BEGIN ALGORITHM
1) Data structure
AllBalls : will contain all the balls read in a file
AllEllipsoids : will contain the set of ellipsoids computed from the initial balls
X : percentage used to calculate K, the expected cluster number of K-means algorithm
Threshold : the balls approximation by set of ellipsoids threshold ;
BallsVolumes : will contain at a given time the volume of all balls of **AllBalls**
2) read the balls file and initialize data structures.
FOR_1 1<=l<=N , retrieve x_l,y_l,z_l coordinates of center of ball l and radius r_l
AllBalls [l,1]=x_l; **AllBalls** [l,2]=y_l; **AllBalls** [l,3]=z_l;
BallsVolumes[l]= (4/3)*pi*r_l^3;
END FOR_1
3) Calculation of cluster number or subset balls expected from K-means algorithm
K=(**X***N)/100
WHILE_1 noEmpty(**AllBalls**) and **K** > 2 and NotConverged()
SubSets =K-means(**AllBalls**,Vol,K);//SubSets contains K clusters of balls of **AllBalls**
FOR_2 1<=i<=K, we calculate the connected components of SubSets_i
Compons = relatedComponents(SubSets[i]); // **Compons** contain p connected components of SubSets_i
FOR_3 1<=j<=p, compute ellipsoid that fitting balls of Compons_j with fitting error errorApp_j
[errorApp_j,ellipsoid_j] = **ellipsoideFitting**(Compons[j]);
IF_1 (errorApp_j check Threshold) then we keep ellipsoid and remove its balls from **AllBalls**
addEllipsoid(**AllEllipsoids**,ellipsoid_j);
deleteBalls(**AllBalls**,Comp[j]);
END IF_1
END FOR_3
END FOR_2
N= Cardinal(**AllBalls**); K=(**X***N)/100; // calculation of new cluster number value of K and repeat the iteration
END WHILE_1

4) Either **AllBalls** is empty or we have **K**<=2, or we converge

IF_2 NoEmpty(**AllBalls**)= **FALSE** then, goto the end of algorithm. We obtains **M** ellipsoïds in **AllEllipsoids** that fitting all starting balls
GOTO END
// We calculate all the related components of the remaining AllBalls and we get p components
ELSE we try to apply another method to recover another ellipsoids from **AllBalls**
//We calculate all the connected components of remaining balls in **AllBalls** and we get p components.
Compons = **relatedComponents**(**AllBalls**);
FOR_4 1<=j<=p, we compute the cluster number of Compons_j
Nballs= Cardinal(Comp[j]); K=(X*Nballs)/100; Vol=BallsVolumes(Comp[j]);
SubSets = **K-means**(Compons[j],**Vol**,K); //SubSet contains K clusters of connected balls of Compons_j
FOR_5 1<=i<=K, calculate ellipsoid that fitting the balls of Subsets_i with fitting error errorApp_i
[errorApp_i,ellipsoid_i] = **ellipsoidFitting**(SubSet[i]);
IF_3 (errorApp_i check Threshold then we keep ellipsoid and remove its balls from **AllBalls**
addEllipsoid(**AllEllipsoids**,ellipsoid_i);
deleteBalls(**AllBalls**,SousEns[i]);
END IF_3
END FOR_5
END FOR_4
END IF_2
5) All remaining balls in **AllBalls are isolated** balls and will be considered as ellipsoids. So we they are added to **AllEllipsoids**.
Nballs= **Cardinal**(**AllBalls**);
FOR_6 1<=j<=p
addBalls(**AllEllipsoids**,**AllBalls**[i]);
deleteBalls(**AllBalls**,**AllBalls**[i]);
END FOR_6
// We have at the end M ellipsoïds in total in **AllEllipsoids** that fitting starting set of balls or porous media.
END ALGORITHM

Algorithm 2. Initial segmentation: division-merge based on octree structure

BEGIN ALGORITHM

1) Data structures
Threshold : division threshold (fitting threshold)
ImagePoint : data structure with two fields :
=> number : balls number in the point image
=> labels : labels of all balls in the point image
M : matrix of **ImagePoint** or 3D image with DIM x DIM x DIM dimension
AllBalls : array of dimension N balls and four columns. This array is also called balls features array
AllBalls [l,1]= x coordinate of ball l
AllBalls [l,2]= y coordinate of ball l
AllBalls [l,3]= z coordinate of ball l
AllBalls [l,4]= radius r of ball l
AllEllipsoids : contains set of all ellipsoids expected that fitting the set of balls.
2) Initialization : read the balls file and fill the structure **AllBalls** and 3D image M
FOR_1 1<=l<=N ball labeled l with i,j,k coodinates and radius r in the **AllBalls**,
M[i,j,k].number=1; M[i,j,k].labels[1]=l; **AllBalls**[l,1]=i; **AllBalls**[l,2]=j;

AllBalls[l,3]=k; **AllBalls**[l,4]=r;
END FOR_1
// if M[i,j,k].number == 1 then the point image is a ball else M[i,j,k].number == 0, 1<=i<=DIM, 1<=j<=DIM, 1<=k<=DIM

2) Top-Down octree construction from 3D image of dimension DIMxDIMxDIM of **ImagePoint**. Octree depth is log2(DIM).
At the beginning, the octree root contains the whole 3D image **M** of image point of all balls.
N = cardinal{M[i,j,k].nbre=1 ,1<=i<=DIM, 1<=j<=DIM, 1<=k<=DIM}

FUNCTION RecursiveCalls(M,DIM)
we fit the set of N balls of image M by an ellipsoid with fitting error **ErrorApp**.
setBall=**getBallsImages**(**AllBalls**,**M**) ;
[errorApp,ellipsoid] = **ellipsoidFitting**(setBall);
IF_1 (errorApp check Threshold), then keep ellipsoid stop octree division
addEllipsoid(**AllEllipsoid,**ellipsoid);
GOTO END
ELSE we split 3D image **M** in eight **DIM/2 x DIM/2 x DIM/2** sub-images **M1,...M8**
FOR_2 1<=t<=8 sub-image Mt,
recusiveCalls(Mt, **DIM/2**);
END FOR_2
END IF_1
END FUNCTION
The program stops when there is no longer any division. The set of ellipsoids is currently **AllEllipsoids** structure
END ALGORITHM

Algorithm 3. Region growth or merge

Algorithm 3 : region growth or merge
BEGIN ALGORITHM
1) Data structures
RegionFeatures : contains all initial ellipsoids features obtained from Algorithm 1 or 2
BallsFeatures : contains all subsets balls of every initial ellipsoids
Graphs : graph to store the adjacency relations between the initial ellipsoids or regions. The nodes are the ellipsoids labels(identification number of ellipsoid) and the edges are the relation between ellipsoids. The edge weight between two ellipsoids is fitting error obtained by merging two ellipsoids by one.
Heap : minimal heap where the merge errors between every two adjacency ellipsoids will be stored. A heap node is a structure with three fields. Two ellipsoids labels to merge and merging error that will be considered as the heap key.

Threshold : threshold of the merging error of two adjacency ellipsoids
2) read initial ellipsoids file and initialize the data structures
FOR_1 1<=l<=N,
//get the subset of balls and the ellipsoid features l
RegionFeatures[l]= Ellipsoid_l;
BallsFeatures[l]=SubSetBalls_l;
END FOR_1
3) Initialize graph and heap.
//Calculation of graph edges and inserting its in the minimal heap according their weight.
countRegion=cardinal(**RegionFeatures**);
FOR_2 1<=i<=(countRegion -1)

FOR_3 (i+1)<=j<=countRegion
IF_1 EdgeExist (i,j)= **TRUE**
[error_i_j,ellipsoid_i_j] = mergeEllipsoids(**BallsFeatures** [i], **BallsFeatures** [j]);

IF_2(error_i_j check the cricterion **Threshold**)
graphIndex++; heapIndex++;
Heap(heapIndex)=[error_i_j,i,j];
IF_3 EdgeExistInGraphs(i,j)=**FALSE**
//edge (i,j)not exist yet in the graph, then insert it.
Graphs(graphIndex)=[i,j,0,0,error_i_j];
ELSE
Update pointers for all predecessor edges that contain either node i or j with graphIndex
//update of all predecessors edges pointers that contain either node i or j with graphIndex
graphUpdate(**Graphs,** graphIndex) ;
Graphs (graphIndex)=[i,j,0,0, error_i_j];
END IF_3
END IF_2
END IF_1
END FOR_3
END FOR_2
4) merge phase. Pop from the heap edge that contains the smallest error ; Check whether this edge exists in the graph structure. If it is the case, merge the ellipsoids of that node to make a new one ; Assign a new label to the new ellipsoid and update the graph.
WHILE_1 noEmpty (**Heap**)
edgeMin=**POP**(Heap);
IF_4 edgeExistInHeap(**Heap**, areteMin)=TRUE
//minimal edge exists in the graph and and its error is that which is in the graph
IF_5 (getErrorOfEdge(edgeMin) check Threshold)
// if error edge named Min don't check the merging threshold, goto the end of fusion algorithm and break the loop WHILE.
GOTO END ALGORITHM
END IF_5
//merge of the two ellipsoids whose labels are contained in edgeMin
i=edgeMin[1]; j=edgeMin[2];
k=ComputeLabelNewEllipsoid(i,j);
[error_k,ellipsoid_k] = mergeEllipsoids(**BallsFeatures** [i], **BallsFeatures** [j]);
IF_6(error_k check Threshold)
//Update the new ellipsoid features and it subset balls.
updateEllipsoidFeatures(**RegionFeatures**,ellipsoid_k);
updateBallsFeatures(**BallsFeatures**,ellipsoid_k);
updateHeap(**Heap**,k,error_k);
updateGraph(**Graphs**,k,error_k);
END IF_6
END IF_5
END WHILE_1

//All ellipsoids are in **RegionFeatures** structure

END ALGORITHM

References

Alsallakh, B., 2014. Visual methods for analyzing probabilistic classification data. IEEE Trans. Visual. Comput. Graph. 20 (12), 1703–1712.

Ayache, N., 2003. Epidaure: a research project in medical image analysis, simulation and robotics at INRIA. IEEE Trans. Med. Imag. 22 (10), 1185–1201.

Baillard, C., Hellier, P., Barillot, C., 2000. Cooperation between level set techniques and dense 3D registration for the segmentation of brain structures. Pattern Recogn. 1, 991–994.

Banégas, F., Jaeger, M., Michelucci, D., Roelens, M., 2001. The ellipsoidal skeleton in medical applications. In: Proceedings of the Sixth ACM Symposium on Solid Modeling and Applications, pp. 30–38.

Baveye, P.C., 2015. Grand challenges in the research on soil processes. Front. Environ. Sci. 3, 10. https://doi.org/10.3389/fenvs.2015.00010.

Baveye, P.C., Pot, V., Garnier, P., 2017. Accounting for sub-resolution pores in models of water and solute transport in soils based on computed tomography images: are we there yet? J. Hydrol. 555, 253–256. https://doi.org/10.1016/j.jhydrol.2017.10.021.

Baveye, P.C., Otten, W., Kravchenko, A., Balseiro Romero, M., Beckers, É., Chalhoub, M., et al., 2018. Emergent properties of microbial activity in heterogeneous soil microenvironments: different research approaches are slowly converging, yet major challenges remain. Front. Microbiol. 8, 1364. https://doi.org/10.3389/fmicb.2017.01364.

Blinn, J., 1982. A generalization of algebraic surface drawing. ACM Trans. Graph. 1 (3), 235–256.

Bricault, I., Monga, O., 1997. From volume medical images to quadratic surface patches. CVGIP Image Underst. 67 (1), 24–38.

Caumon, G., Lepage, F., Sword, C.H., Mallet, J.-L., 2004. Building and editing a sealed geological model. Math. Geol. 36 (4), 405–424.

Caumon, G., Levy, B., Castani, L., Paul, J.C., 2005. Visualization of grids conforming to geological structures: a topological approach. Comput. Geosci. 31 (6), 671–680.

Darwish, M.A., Abdulrahim, H.K., Mohieldeen, Y., 2014. Qatar and GCC water security. Desalination Water Treat. 55, 2302–2325. https://doi.org/10.1080/19443994.2014.947782.

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

Doube, M., 2015. The ellipsoid factor for quantification of rods, plates, and intermediate forms in 3D geometries. Front. Endocrinol. 6, 15. https://doi.org/10.3389/fendo.2015.00015.

Druoton, L., Garnier, L., Langevin, R., 2013. Iterative construction of Dupin cyclides characteristic circles using non-stationary Iterated Function Systems (IFS). Comput. Aided Des. 45 (2), 568–573.

Ebrahimi, A., Or, D., 2015a. Hydration and diffusion processes shape microbial community organization and function in model soil aggregates. Water Resour. Res. 51 (12).

Ebrahimi, A., Or, D., 2015b. Hydration and diffusion processes shape microbial community organization and function in model soil aggregates. Water Resour. Res. 51, 9804–9827. https://doi.org/10.1002/2015WR017565.

Eickhorst, T., Tippkötter, R., 2008. Detection of microorganisms in undisturbed soil by combining fluorescence in situ hybridization (FISH) and micropedological methods. Soil Biol. Biochem. 40, 1284–1293.

Falconer, R.E., Houston, A.N., Otten, W., Baveye, P.C., 2012. Emergent behavior of soil fungal dynamics: influence of soil architecture and water distribution. Soil Sci. 177, 111–119.

Faugeras, O.D., Hebert, M., 1984. The representation, recognition, and locating of 3-D objects. Int. J. Robot Res. 5 (3), 27–52.

Freije, A.M., 2015. Heavy metal, trace element and petroleum hydrocarbon pollution in the Arabian Gulf: Review. J. Assoc. Arab Univ. Basic Appl. Sci. 17, 90–100.

George, P.-L., Borouchaki, H., 1998. Delaunay Triangulation and Meshing. Butterworth-Heinemann, London, U.K.

Held, R.J., Celia, M.A., 2001. Pore-scale modeling extension of constitutive relationships in the range of residual saturations. Water Resour. Res. 37 (1), 165–170.

Houston, A.N., Schmidt, S., Tarquis, A.M., Otten, W., Baveye, P.C., Hapca, S.M., 2013. Effect of scanning and image reconstruction settings in X-ray computed microtomography on quality and segmentation of 3D soil images. Geoderma 207–208 (1), 154–165. https://doi.org/10.1016/j.geoderma.2013.05.017.

Kuiper, N., Rowell, C., Shomar, B., 2015. High levels of molybdenum in Qatar's groundwater and potential impacts. J. Geochem. Explor. 150, 16–24.

Malladi, R., Sethian, J., Vemuri, B., 1995. Shape modeling with front propagation: a level set approach. IEEE Trans. Pattern Anal. Mach. Intell. 17 (2), 158–175.

Masse, D., Cambier, C., Brauman, A., et al., 2007. MIOR: an individual-based model for simulating the spatial patterns of soil organic matter microbial decomposition European. J. Soil Sci. 58 (5), 1127–1135.

Mokhtari, B., Melkemi, K.E., Michelucci, D., Foufou, S., 2014. Shape recognition and retrieval method based on dynamic clustering, shape recognition and retrieval method based on dynamic clustering. In: Horvath, I., Rusak, Z. (Eds.), Proceedings of TMCE 2014, May 19–23, 2014, Budapest, Hungary.

Monga, O., 1987. An optimal region growing algorithm for image segmentation. Int. J. Pattern Recogn. Artif. Intell. 1 (3), 351–376.

Monga, O., 2007. Defining and computing stable representations of volume shapes from discrete trace using volume primitives : application to 3D image analysis in soil science. Image Vis Comput. 25, 1134–1153.

Monga, O., Ngom, F., Delerue, J.F., 2007. Representing geometric structures in 3D tomography soil images: application to pore space modeling. Comput. Geosci. 33 (9), 1140–1161. https://doi.org/10.1016/j.cageo.2006.11.015.

Monga, O., Bousso, M., Pot, V., Garnier, P., 2008. 3D geometrical structures and biological activity: application to microbial soil organic matter decomposition in pore space. Ecol. Model. 216 (3–4), 291–302. https://doi.org/10.1016/j.ecolmodel.2008.04.015.

Monga, O., Bousso, M., Pot, V., Garnier, P., 2009. Using pore space 3D geometrical modelling to simulate biological activity: impact of soil structure. Comput. Geosci. 35 (9), 1789–1801. https://doi.org/10.1016/j.cageo.2009.02.007.

Monga, O., Garnier, P., Pot, V., Coucheney, E., Nunan, N., Otten, W., Chenu, C., 2014. Simulating microbial degradation of organic matter in a simple porous system using the 3D diffusion based model MOSAIC. Biogeosciences 11, 2201–2209.

Mumford, D., Shah, J., 1989. Optimal approximation by piecewise smooth functions and associated variational problems, Comm. Pure Appl. Math. 42, 577–685.

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.

Ngom, N.F., Monga, O., Ould Mohamed, M., 2012. 3D shape extraction segmentation and representation of soil microstructures using generalized cylinders. Comput. Geosci. 39, 50–63.

Nguyen-Ngoc, D., Leye, B., Monga, O., Garnier, P., Nunan, N., 2013. Modeling microbial decomposition in real 3D soil structures using partial differential equations. Int. J. Geosci. 4, 15–26. https://doi.org/10.4236/ijg.2013.410A003.

Nguyen-Ngoc, D., Leye, B., Monga, O., Garnier, P., Nunan, N., 2015. Simulating biological dynamics using partial differential equations: application to decomposition of organic matter in 3D soil structure. Vietnam Journal of Mathematics. Springer Verlag.

Otsu, N.A., 1979. Threshold selection method from gray level histograms. IEEE Trans. Syst. Man Cybern. 9, 6266.

Pérez-Reche, F.J., Taraskin, S.N., Otten, W., Viana, M.P., da, F., Costa, L., Gilligan, C.A., 2012. Prominent effect of soil network heterogenity on microbial invasion. Phys. Rev. Lett. 109 (9), 098102. https://doi.org/10.1103/PhysRevLett.109.098102.

Portell, X., Pot, V., Garnier, P., Otten, W., Baveye, P.C., 2018. Microscale heterogeneity of the spatial distribution of organic matter can promote bacterial biodiversity n soils" Insights from computer simulations. Front. Microbiol. 9, 1583. https://doi.org/10.3389/fmicb.2018.01583.

Pot, V., Peth, S., Monga, O., Vogel, L.E., Genty, A., Garnier, P., Vieublé-Gonod, L., Ogurreck, M., Beckmann, F., Baveye, P.C., 2015. Three-dimensional distribution of water and air in soil pores: comparison of two-phase two-relaxation-times lattice-Boltzmann and morphological model outputs with synchrotron X-ray computed tomography data. Advances in Water Research 84, 87–102.

Shomar, B., Darwish, M., Rowell, C., 2014. What does integrated water resources management from local to global perspective mean? Qatar as a case study, the very rich country with no water. Water Resour. Manag. 28, 2781–2791.

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

Terzopoulos, D., 1991. Dynamic 3D models with local and global deformations: deformable superquadrics. IEEE PAMI 13 (7), 703–714.

Vaz, C.M.P., de Maria, I.C., Lasso, P.O., Tuller, M., 2011. Evaluation of an advanced benchmark micro-computed tomography system for quantifying porosities and pore-size distributions of two Brazilian Oxisols. Soil Sci. Soc. Am. J. 75, 832–841.

Vidal, D., Roy, R., Bertrand, F., 2010. On improving the performance of large parallel lattice Boltzmann flow simulations in heterogeneous porous media. Comput. Fluids 39, 324–337.

Vogel, H.J., Roth, K.I., 2001. Quantitative morphology and network representation of soil pore structure. Adv. Water Resour. 24, 233–242.

Vogel, L., Makowski, D., Garnier, P., Vieublé-Gonod, L., Raynaud, X., Nunan, N., Chenu, C., Falconer, R., Pot, V., 2015. Modelling the effect of soil meso- and macropores topology on the biodegration of a soluble carbon substrate. Adv. Water Resour. 83, 123–136.

Wen, X.H., Gomez-Hernandez, J.J., 1996. Upscaling hydraulic conductivities in heterogeneous media: an overview. J. Hydrol. 183https://doi.org/10.1016/S0022–1694(96)80030–8. ix–xxxii.

Whiffin, V.S., van Paassen, L.A., Harkes, M.P., 2007. Microbial carbonate precipitation as a soil improvement technique. Geomicrobiol. J. 24 (5), 417–423.