Geoderma 157 (2010) 51–63

Contents lists available at ScienceDirect

Geoderma

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

Observer-dependent variability of the thresholding step in the quantitative analysis of soil images and X-ray microtomography data

Philippe C. Baveye [a,*], Magdeline Laba [b], Wilfred Otten [a], Liesbeth Bouckaert [c,1], Patricia Dello Sterpaio [a], Rohit R. Goswami [d], Dmitri Grinev [a], Alasdair Houston [a], Yaoping Hu [e], Jianli Liu [f], Sacha Mooney [g], Radoslaw Pajor [a], Steven Sleutel [c], Ana Tarquis [h], Wei Wang [i], Qiao Wei [e], Mehmet Sezgin [j]

[a] SIMBIOS Centre, Abertay University, Bell Street, Dundee, Scotland DD1 1HG, UK
[b] School of Civil and Environmental Engineering, Hollister Hall, Cornell University, Ithaca, New York 14853, USA
[c] Department of Soil Management and Soil Care, Ghent University, 653 Coupure Links, 9000 Gent, Belgium
[d] Geosyntec consultants, 5901 Broken Sound Parkway N.W., Suite 300, Boca Raton, Florida 33487, USA
[e] Department of Electrical and Computer Engineering, Schulich School of Engineering, University of Calgary, 2500 University Drive, NW, Calgary, Alberta, Canada, T2N 1N4
[f] Department of Soil Physics and Salinity, Institute of Soil Science, Chinese Academy of Sciences, 71 East Beijing Road, Nanjing 210008, China
[g] Environmental Sciences, School of Biosciences, Biology Building, University of Nottingham, University Park, Nottingham NG7 2RD, UK
[h] Dept. of Applied Mathematics to Agriculture Engineering, Polytechnic University of Madrid (U.P.M.), Ciudad Universitaria s.n., 28040 Madrid, Spain
[i] Department of Crop & Soil Sciences, Michigan State University, East Lansing, Michigan 48824, USA
[j] Tübitak Marmara Research Center, Information Technologies Research Institute, Gebze, Kocaeli, Turkey

ARTICLE INFO

Article history:
Received 19 June 2009
Received in revised form 11 March 2010
Accepted 15 March 2010

Keywords:
Image analysis
Computed tomography
Lattice–Boltzmann modelling
Pore geometry
Soil architecture

ABSTRACT

For the investigation of many geometrical features of soils, computer-assisted image analysis has become a method of choice over the last few decades. This analysis involves numerous steps, regarding which subjective decisions have to be made by the individuals conducting the research. This is particularly the case with the thresholding step, required to transform the original (color or greyscale) images into the type of binary representation (e.g., pores in white, solids in black) needed for fractal analysis or simulation with Lattice–Boltzmann models. Limited information exists at present on whether different observers, analyzing the same soil, would be likely to obtain similar results. In this general context, the first objective of the research reported in this article was to determine, through a so-called "round-robin" test, how much variation exists among the outcomes of various image thresholding strategies (including any image pre-treatment deemed appropriate), routinely adopted by soil scientists. Three test images – of a field soil, a soil thin section, and a virtual section through a 3-dimensional CT data set – were thresholded by 13 experts, worldwide. At the same time, variability of the outcomes of a set of automatic thresholding algorithms, applied to portions of the test images, was also investigated. The experimental results obtained illustrate the fact that experts rely on very different approaches to threshold images of soils, and that there is considerable observer influence associated with this thresholding. This observer dependence is not likely to be alleviated by adoption of one of the many existing automatic thresholding algorithms, many of which produce thresholded images that are equally, or even more, variable than those of the experts. These observations suggest that, at this point, analysis of the same image of a soil, be it a simple photograph or 3-dimensional X-ray CT data, by different individuals can lead to very different results, without any assurance that any of them would be even approximately "correct" or best suited to the objective at hand. Different strategies are proposed to cope with this situation, including the use of physical "standards", adoption of procedures to assess the accuracy of thresholding, benchmarking with physical measurements, or the development of computational methods that do not require binary images.

1. Introduction

In the early 1970s, the commercialization of image analysis instruments, like the Quantimet 720 (Fisher, 1971; Nawrath and Serra, 1979a,b), with dedicated minicomputers and integrated scanners, contributed significantly to promote the quantitative use of photographs in the study of soils and natural porous media. For the first time, it became possible to make fast, relatively accurate measurements on scanned images produced by cameras, light and electron microscopes, or X-ray spectrometers. Soil scientists rapidly seized that opportunity (e.g., Jongerius et al., 1972; Murphy et al., 1977a,b). They continued to do so in the 1980s (e.g., Ringrose-Voase

* Corresponding author.
E-mail address: P.Baveye@Abertay.ac.uk (P.C. Baveye).
[1] The order of authors 4–16 is alphabetical.

0016-7061/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.geoderma.2010.03.015

and Bullock, 1984), after the development of personal computers, and particularly in the 1990s, as inexpensive scanners, digital cameras, and versatile image manipulation software made it increasingly straightforward to acquire images and to analyze them (often with the help of fractal geometry), for a variety of purposes (e.g., Hallaire and Cointepas, 1993; Bielders et al., 1996; Deleporte et al., 1997; DeLeo et al., 1997; Beaudet-Vidal et al., 1998; Baveye, 2002; Ohrstrom et al., 2002; Dathe and Baveye, 2003; Ohrstrom et al., 2004; Morris and Mooney, 2004; Pendleton et al., 2005; Persson et al., 2005; Wantanaphong et al., 2006; Lipsius and Mooney, 2006; Vogel et al., 2006; Otten and Gilligan, 2006; Jacobson et al., 2007; Marcelino et al., 2007; Mooney and Morris, 2008; Papadopoulos et al., 2008; Persson and Olsson, 2008; Tarquis et al., 2008). In recent years, synchrotron-based X-ray computed tomography (CT) and table-top X-ray micro-CT scanners have allowed researchers to visualize in 3 dimensions the structure and composition of soils at micrometric resolutions, and have enabled significant advances to be made in our understanding of the functioning of soils at previously unexplored spatial scales (e.g., Garnier et al., 1998; Baveye et al., 2002; Elliot and Heck, 2007a,b; Sleutel et al., 2008).

In the midst of this technological evolution, Thompson et al. (1992) advised caution in the interpretation of data generated by analysis of images of soils, arguing that this analysis involves many successive steps, all of which can be affected by artefacts or subjectivity. Before measurements can be performed, unless one can simply take pictures in the field, soils need to be sampled, dried in some way, solidified by impregnation with one of a number of available resins, and cured for a period of time. At the stage when pictures of the soil are taken (in the field or in the laboratory), different lighting arrangements, cameras, lenses, resolutions, aperture and settings for exposition have to be selected among many possible choices. If computed tomography (CT) scanners are used to generate three-dimensional data, the resolution of the scanning, and a number of settings related to attenuation and contrast, can differ, depending on who does the scanning. Finally, the resulting 2-D or 3-D images of the soil have to be thresholded or segmented to produce a binary image, to which are then applied a wide range of statistical or mathematical methods. All these steps involve operational decisions that can vary from one observer to another.

To alleviate some of the resulting variability, Thompson et al. (1992) recommended that efforts should be made to standardize practices, in particular in terms of the thresholding step, to "assist operators in identifying discrepancies before measurement." To that end, these authors suggested that standard images of soil pore space could be used for calibration of thresholding protocols. At present, it is unclear to what extent this advice has been heeded and what type of standardization, if any, has occurred. Certainly, the concept of standard image to which Thompson et al. (1992) were referring has not made much headway since 1992, as evinced by Marcelino et al.'s (2007) renewed call for standardisation, which is almost word-for-word identical to Thompson et al.'s (1992). None of the various soil science laboratories operating CT scanners in the world appears to have an agreed-upon standard that is run systematically to validate subsequent analyses of the data. Nevertheless, it is possible that, since many of the groups carrying out image analysis on photographs of soils or CT data end up using the same software, some standardisation has occurred anyway.

In similar situations, when questions arise about the variability associated with the way different institutions or individuals approach a particular process, it is customary to carry out a so-called "round robin" or "ring" test. A number of knowledgeable participants are provided independently with the same materials/data and are asked to apply the given process to them as they would routinely. Such round-robin tests have been carried out in soil science in the past, with interesting results, in particular to assess inter- and intra-laboratory variability in the chemical analysis of soils (e.g., Sager, 1999; Cools et al., 2004; Creamer et al., 2009). Tests involving individual researchers have also been carried out. Murphy et al. (1985), for example, asked seven experienced micromorphologists to characterize independently, using the same reference handbook, a number of thin sections obtained from soil horizons representing soil materials formed by a variety of pedogenic processes. In spite of differences among the seven descriptions, the overall degree of uniformity among them was considered encouraging. More recently, Brown et al. (1996) gave five modellers a full description of a field experiment carried out to determine the leaching potential of a novel pesticide, and asked them to use the same three mathematical models to predict concentrations of pesticide in soil water at a depth of one meter, and in the soil itself over a 1 m profile, 220 days after application. The simulation outcome revealed appreciable differences in the way modellers approached their analysis, and in the quantitative results they obtained. Further round-robin tests of pesticide fate or exposure modelling by Francaviglia et al. (2000) and Beulke et al. (2006) led to essentially similar observations.

A full-fledged round-robin test of the quantitative analysis of soil images would require getting different observers separately in the same pits in the field, so they could photograph soil profiles as each of them sees fit, or circulating the same soil columns among different laboratories around the world, where CT scanners are routinely used, so that each observer can adjust the settings of their instrument as deemed appropriate. Aside from logistical (and therefore financial) hurdles in such a test that would need to be overcome, a major impediment is that some of the more intrusive steps involved in the full image analysis process (such as resin impregnation, cutting of soil blocks, or acid etching of thin sections to enhance contrast) cannot be duplicated by different observers on exactly the same samples, giving rise to uncertainty in treatment comparison. Therefore, until some of these challenges can be resolved, it is reasonable to envisage a round-robin test only with respect to specific steps of the whole image analysis process. There have been some attempts in that area, especially in relation with the thresholding/segmentation of images. Baveye et al. (1998), Ogawa et al. (1999) and Boast and Baveye (2006) applied two different thresholding algorithms to images of a soil profile, as a preliminary step in the fractal characterization of a preferential pathway. Marcelino et al.'s (2007) carried out a mini round-robin test with two individuals, one of whom replicated his analysis at two different times, to assess the influence of the method of microscopic visualization and thresholding (manual, semi-automatic, or automatic algorithms) on a number of soil parameters. Tarquis et al. (2008) applied four different threshold criteria to transform computed tomography grey-scale imagery of four Brazilian soils into binary imagery to estimate their mass fractal and entropy dimensions. They found that the threshold criteria used had a direct influence on the porosity obtained, varying from 8 to 24% in one of the samples, and on the fractal dimensions. Nevertheless, since they involve either few individual observers or a limited number of different image manipulation techniques, these various comparisons remain very limited in scope, and a larger-scale investigation appears necessary.

In this general context, this article is meant to pursue and improve on the preliminary efforts of Baveye et al. (1998), Ogawa et al. (1999), Boast and Baveye (2006), Marcelino et al. (2007), and Tarquis et al. (2008). The key objective is to assess, through a round-robin test, how much variation exists among the outcomes of image thresholding strategies (including any image pre-treatment deemed appropriate), routinely adopted by soil scientists. To that end, three test images – of a field soil, a soil thin section, and a virtual section through a 3-dimensional CT data set – were thresholded by 13 experts, worldwide. The majority of these experts were soil scientists, having published articles dealing with image analysis. Other experts had a recognized track record of image analyses in other fields (e.g., material science, remote sensing, or biofilm research). The outcome of their manipulation of the images was compiled and analyzed. At the same time, because application of automatic thresholding algorithms may offer at least a partial solution to the observer-dependence identified in the

round-robin test, variability among the outcomes of a set of algorithms was also investigated. Subsets of the three test images were subjected to forty different thresholding algorithms, described in Sezgin and Sankur (2004), and the variability of their results was compared to that of the experts consulted.

2. Materials and methods

2.1. Test images — round robin

Three test images – of a field soil, a soil thin section, and a virtual section through a 3-dimensional CT data set – were selected for this research. The key criterion in selecting these images was the fact that they either had been used for a piece of research leading to a publication in the soils literature, or they were about to be used in such a context. Therefore, these images were deemed representative of images or CT data analyzed by soils researchers. That does not mean that these images or CT data were of optimal quality for thresholding, i.e., that the field soil was uniformly illuminated, that no part of the images was under- or overexposed, or that the contrast throughout the images was optimized. Of course, as mentioned by one of the reviewers of this article, the quality of the images to some extent determines the similarity, or lack thereof, of the experts' outputs. Had a picture of a dyed preferential pathway been obtained in a uniformly illuminated, uniformly packed sand or sandy soil, the starting image submitted to experts would have been almost binary already, and of course, the thresholded images produced by the experts would have been very similar. However, that is a somewhat ideal situation, and it was thought better to envisage a more routine case, with images typically used by soils researchers.

The first image, labelled "Pathway" (Fig. 1), has been described in detail and analyzed earlier by Baveye et al. (1998), Ogawa et al. (1999) and Boast and Baveye (2006). This image results from a preferential flow experiment performed on July 14, 1995, in the old Cornell University orchard, in Ithaca, New York. A metal ring was pushed into the surface layer of the soil. Twenty litres of a 1% solution (10 g/L) of blue food colouring (F&DC #1) were poured inside the cylinder and rapidly infiltrated into the soil. Fifteen minutes later, a 1.8 m deep trench was dug, tangential to the outer surface of the metal ring. Initial digging was done with a backhoe, followed by carefully removing soil with shovels in order to obtain as vertical as possible a soil profile. Using a hand-held camera, under natural light, several colour slides were taken of soil faces exposed at different stages during the carving out of the soil. The slides were scanned and the resulting digitized images were stored in red-green-blue (RGB) colour-coding format on a CD-ROM. A 1691 by 2696 pixels portion of one of the images was rectified (rotated so a yardstick on the image would be vertical) and cropped to produce the "Pathway" image. A white label is located in the upper right portion of the image, and spade marks are clearly visible in the lower part of the profile, which also is less lit than the upper part. Plant roots were not removed before the picture was taken and are noticeable near the surface. The histograms of the red and green channels (RGB color decomposition), as well as of the cyan channel (CMYK color decomposition) all exhibit a major peak and a pronounced shoulder. The major peak is especially narrow, and the shoulder distinct, in the histogram of the cyan channel.

The light transmission photograph labelled "Thinsect" (Fig. 1) was obtained by placing directly on a piece of photographic paper, and illuminating for a set time, a resin-impregnated thin section (100 micrometer-thick) of a sample of soil obtained in 1992 in the research station of the Institut National des Sols at Glidji (Aneho), in the "terres de Barre" region of southern Togo. Further details about this soil are available in Bielders et al. (1996). A portion of the

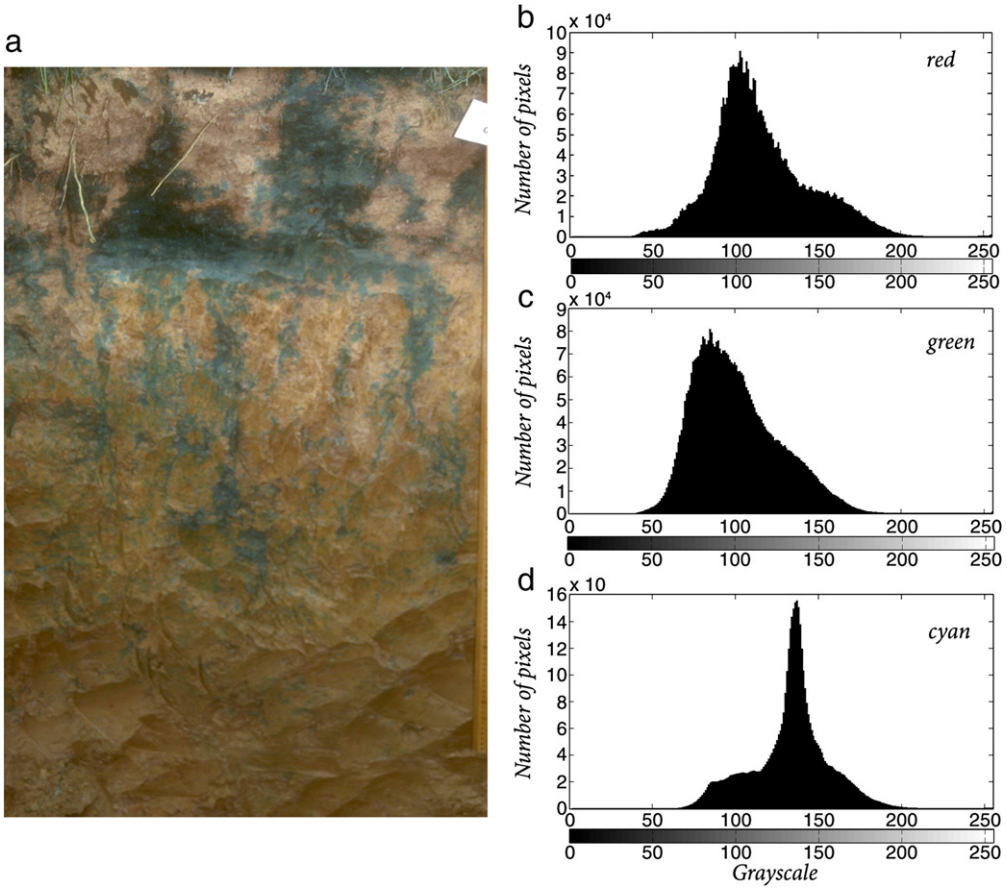

Fig. 1. (a) Pathway test image. (b)–(d) Greyscale histograms of the red, green, and cyan channels. In these histograms, zero is black and 255 is white, by convention.

photograph, below the surface crust, with dimensions of 4098 by 4098 pixels was cropped. It corresponds to a physical area of 7×7 mm on the thin section. The histogram (Fig. 2b) is characterized by a high number of points at a greyscale value of zero, associated with black particles present throughout the image. For greyscale values up to an apparent inflexion point at 224, the number of pixels increases monotonically. At a greyscale value of 237, the histogram exhibits a peak that is clearly associated with the resin-filled pores. This resin has a gravely appearance and here and there includes gas bubbles.

The image labelled "Nanotom" is associated with an arable sandy loam from Bullion Field, an experimental site in the Scottish Crop Research Institute (Invergowrie, U.K.). The soil was air-dried, sieved and repacked at a bulk-density of 1.2 Mgm^{-3} and the porosity was calculated from the bulk-density and the particle density of the soil as 52%. The soil was impregnated with Crystic xB52n9 resin and solidified (Harris et al., 2003), and a cube (1 cm^3) was cut out. The cube was then scanned with a Nanotom X-ray Computed Tomography instrument (Phoenix, Wunstorf, Germany) at a resolution of 2.5 µm. A 1900 by 1900 pixel virtual slice was selected within the resulting 3-dimensional image for further treatment and thresholding (Fig. 3a). The histogram associated with this image is remarkably unimodal, with a single, slightly skewed peak at a grayscale value of 208 (Fig. 3b). The lighter (greater grayscale values), somewhat granular material is the resin, in which at least at one location (center left of the image) a desiccation crack is clearly visible. Even though the original image was greyscale, in the process of preparing the final image for the round robin test, the software used – Adobe Photoshop CS3 Extended version 10.0.1 (Adobe Systems, Inc., San Jose, California) – stored the file in a TIFF format that some programs interpret as being RGB instead of greyscale. This explains why some experts refer to the image as RGB and select one of its channels for treatment.

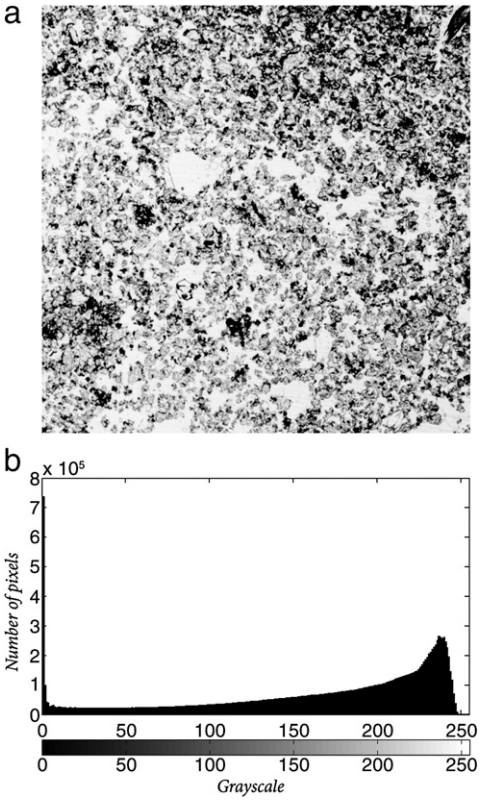

Fig. 2. (a) Thinsect test image. (b) Corresponding greyscale histogram.

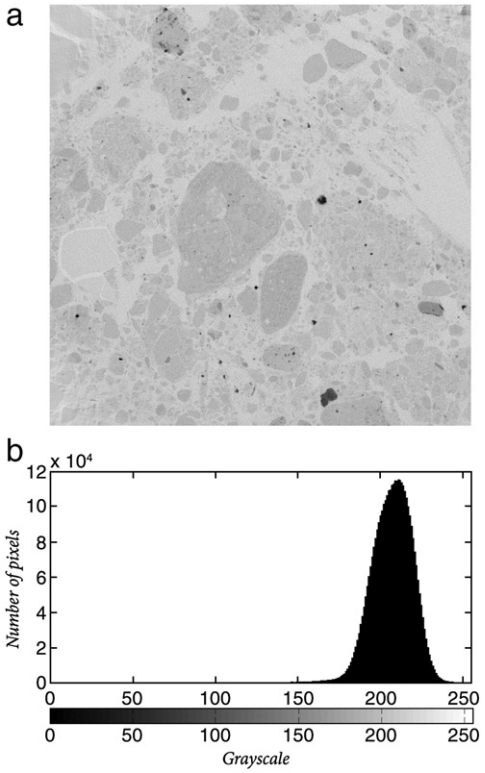

Fig. 3. (a) Nanotom test image. (b) Corresponding greyscale histogram.

2.2. Thresholding by experts

The three test images, Pathway, Thinsect, and Nanotom, were uploaded in TIFF format, without any compression, on a public web site. Various individuals with recognized expertise in image analysis and, in many cases, having published articles on its application to soils, were invited to download these images and threshold them. They were given succinct information on what each image was, and on how the resulting binary outputs were expected to look (e.g., pathway in black for Pathway image; pores in white, solids in black for the other two images). No other information was provided, in particular on any intended use of the image at a later stage (such as for Lattice–Boltzmann modelling or fractal analysis). Each expert was encouraged to apply to the images, as much as possible, the standard procedure he/she uses in his/her research at this point, including any pre-treatment of the images they might deem suitable, e.g., to enhance their contrast, de-noise them or highlight specific features. Some of the experts who responded declined our invitation to threshold the Pathway image because they had no experience dealing with color images, but they nevertheless analyzed the other two images. In the interest of brevity, the procedure followed by each expert is described in detail in the Supplemental material (see below).

2.3. Test sub-images

The compare automatic thresholding algorithms, decision was made to use the software OTIMEC (see next section). The current version of this program operates on greyscale images that are at most 256 by 256 pixels in size, and would need to be substantially altered to deal with the much bigger test images Pathway, Thinsect, and Nanotom. Therefore, 256 by 256 pixel sub-images of these original test images were selected for analysis. For the Pathway image, the cyan channel (CMYK color decomposition) was retained. In each case, different subimages were cropped, and their greyscale histograms

were examined, until a sub-image was found that had a histogram quantitatively similar to that of the original image. For the Pathway sub-image (Fig. 4a), it was not possible to find a situation in which correspondence was optimal, probably because of the geometry of the preferential pathway. The histogram of the sub-image does not reproduce the shoulders exhibited by the histogram of the whole image (Fig. 1d). But for the other two sub-images, and especially for the Thinsect image (Fig. 4b), the similarity of the histograms with those of the original images is satisfactory.

2.4. Comparison of automatic thresholding algorithms

Comparison of a large number of automatic thresholding algorithms is greatly facilitated by the fact that Sezgin and Sankur (2004) provided a very comprehensive classification of these algorithms and, especially, because they coded over thirty eight of them in a computer program, OTIMEC, which is freely available for download (at URL http://mehmetsezgin.net/; last accessed March 10, 2010).

The classification adopted by Sezgin and Sankur (2004) involves 6 categories, according to the type of information the thresholding algorithms are relying on. These categories (and, in parenthesis, the different associated components of the OTIMEC code) are:

1. Histogram shape-based methods. This group of methods achieves thresholding based on selected shape properties of the histogram, like concavities of the histogram relative to its hull (Shape_Rosenfeld), peaks or valleys (Shape_Sezan), overlapping peaks searched via curvature analysis (Shape_Olivio), or lobes of the histogram approximated via a rectangular two-step function (Shape_Tamesh).
2. Clustering-based methods. Greyscale values of pixels are clustered in two parts as background and foreground (object), or alternately are modelled as a mixture of two Gaussians. Since the two clusters correspond to the two lobes of a histogram (assumed distinct), some authors search for the midpoint of the peaks (Cluster_Ridler and Cluster_Yanni). Otsu (Cluster–Otsu) suggested minimizing the weighted sum of within-class variances of the foreground and background pixels to establish an optimum threshold. Minimum-error algorithms (Cluster_Lloyd and Cluster_Kittler) assume that the image can be characterized by a mixture distribution of foreground and background pixels. Another class of algorithms (Cluster_Jawahar a–b) assign fuzzy clustering memberships to pixels depending on their difference from the two main class means.
3. Entropy-based methods. This class of algorithms exploits the entropy of the distribution of the greyscale values in a scene. The maximization of the entropy of the thresholded image is interpreted as indicative of maximum information transfer (Entropy_Pun a–b, Entropy_Kapur, Entropy_Yen, Entropy_Sahoo) Other authors try to minimize the cross-entropy between the input grey-level image and the output binary image as indicative of preservation of information (Entropy_Li, Entropy_Brink) or a measure of fuzzy entropy (Entropy_Shanbag).
4. Object attribute-based methods. These algorithms select the threshold value based on some attribute quality or similarity measure between the original image and the binarized version of the image. These attributes can take the form of edge matching (Attribute_Hertz), grey-level moments (Attribute_Tsai,) connectivity (Attribute_O'Gormann), or stability of segmented objects (Attribute_Pikaz). Some other algorithms evaluate directly the resemblance of the original grey-level image to the binary image using fuzzy measures (Attribute_Huang).
5. Spatial methods. This class of algorithms utilizes not only the grey values distribution, but also dependency of pixels within a neighborhood, for example, in the form of co-occurrence probabilities (Spatial_Pal a–b), second-order entropy (Spatial_Abutaleb), or by capturing the spatial dependence of pixels as binary block patterns (Spatial_Beghdadi).
6. Local methods. In this class of algorithms, a threshold is calculated at each individual pixel, which depends on some local statistics like the local contrast (Local_Yasuda, Local_White, Local_Bernsen), local variance (Local_Niblack, Local_Sauvola), or surface-fitting parameters of the pixel neighborhood. Other algorithms (Local_-Palumbo, Local_Kamel) rely on a center-surround scheme for determining the threshold. A surface fitted to the grey-level landscape can also be used as a local threshold (Local_Yanowitz).

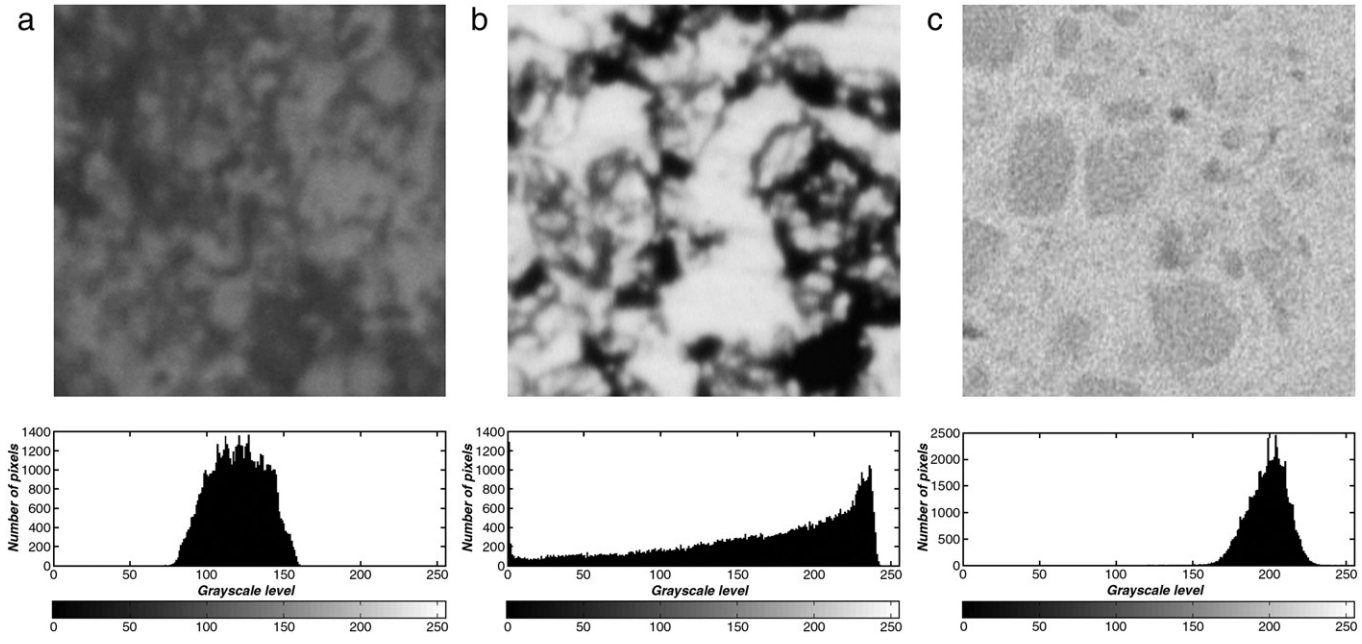

Fig. 4. (a) Pathway sub-image and its corresponding greyscale histogram. (b) Thinsect sub-image and greyscale histogram. (a) Nanotom sub-image and greyscale histogram.

3. Results

3.1. Round-robin test

Faced with the same test images, the 13 experts who participated in the Round Robin test manifested great variability in the software used. It appears clearly that up to now the image analysis software business has not caused any standardization of methods used in soil science research, as hypothesized in the introduction. Even the availability of freeware, like ImageJ, does not seem to prevent some experts from diversifying their software choices, including commercial packages. Similarly, the experts' responses show great diversity in whether or not they carried out pre-treatment steps prior to the thresholding proper, and in the method used both for pre-treatment and for thresholding. Even individuals within the same research groups ended up in some cases using very different methodologies.

This lack of standardization translates directly into thresholded images that differ substantially even with respect to the simplest quantitative measures (Table 1). For the thresholded Pathway images, the percentage of the image occupied by the preferential pathway ranges from a low of 16.85% (expert 4) to a high of 28.58% (expert 2). The mean percentage in all images is 22.85%, and the standard deviation is not enormous, at 3.63. Yet, qualitatively, the images that are associated with the extremes of the observed range (depicted in Fig. 5a and b) appear different. In the image (Fig. 5b) with the lower relative pathway area, the pathway is restricted to an anaemic-looking backbone of connected or nearly-connected macropores. At the other extreme (Fig. 5a), besides a fatter backbone, many isolated "islands" are present in the image.

The "porosity" of the thresholded Thinsect image, i.e., the relative area of each image that corresponds to pores, ranges from 24.02% (expert 8) to 75.50% (expert 4). The mean is 53.22% and the standard deviation is 20.47. This amounts to a far more substantial, three-fold difference in porosity, leading to pore spaces (Fig. 5c and d) that appear markedly different. In the high-porosity case (Fig. 5c), the solid particles seem almost punctual and are spatially dispersed, with a slight tendency to be more densely packed in the upper right corner of the image. By contrast, in the low-porosity image (Fig. 5d), soil particles in general appear bulkier and contact each other, thereby changing dramatically the connectivity of pores (at least in 2-D and at the resolution considered). The fact that information on porosity is available for the soil of which the Thinsect image is a depiction, makes it possible to assess at least qualitatively the appropriateness of some of the porosity values obtained after image thresholding. In the area of the thin section in which the image was cropped, the soil has a band-averaged porosity of 45% (Bielders et al., 1996), slightly below the mean of the experts' outputs.

Table 1
Comparison of quantitative measures obtained on the images thresholded by the different experts. For the Pathway image, the numbers correspond to the relative area (in %) of the image covered by the stained pathway. For the other images (Nanotom and Thinsect), the numbers correspond to the porosity (open pore space) in the image, as a percentage of the area of the whole image.

Expert	Pathway image	Thinsect image	Nanotom image
	Preferential pathway area (%)	Porosity (%)	Porosity (%)
1	–	25.76	1: 30.11 2: 51.14
2	28.58	72.24	57.10
3	1: 26.43 2: 26.45	1: 72.39 2: 72.33	1: 58.03 2: 58.12 3: 58.35 4: 57.92
4	16.85	75.50	36.57
5	Intermeans: 21.05 Min-error: 23.34	–	–
6	19.56	1: 32.50 2: 51.78	60.11
7	21.26	33.48	72.71
8	20.66	24.02	51.32
9	–	39.80	67.79
10	–	68.82	1: 54.58 2: 12.92
11	–	27.47	51.30
12	–	Iterative: 73.20 Otsu: 73.48 Entropy: 66.05 Ind. Kriging: 64.83	Iter: 65.58 Otsu: 65.58
13	24.35	31.09	54.32

For the Nanotom image, estimates of the porosity range from 12.92% (expert 10, case 2) to a high of 72.71% (expert 7). The mean porosity is 53.53% and the standard deviation is 14.36%. The thresholded image with the lowest porosity clearly stands out of the pack, and stems from the perspective adopted by the expert in this case to lose as little as possible of the solid mass. This expert considered that other perspectives were possible (including one with which the same expert found a much higher porosity of 54.58%, see Table 1). Among the remaining thresholded images, the one with the lowest porosity was obtained by expert 1 (case 1), and has a porosity of 30.4% (Fig. 5f). Between large solid particles, aggregates of what appears to be numerous small particles are obstructing much of the space, isolating many pores laterally from adjacent ones. By contrast, the geometry of the pore space in the image with the highest porosity (expert 7; Fig. 5e) is very open, with only a tiny portion of the pore space laterally occluded from the bulk.

Again, since a measured porosity value is available for the soil of which the Nanotom image is a depiction, the appropriateness of some of the porosity values obtained after image thresholding can be ascertained. This assessment has to be done with caution, because the actual porosity applies to a 3-dimensional sample, whose porosity may not be the same as that, evaluated in 2-dimension, in any cross-section through the sample. Nevertheless, one would expect statistically that, in general, the porosity of any cross-section would be in the vicinity of that of the 3-D sample. Unknowingly, the majority (10 of 18) of the estimated porosities after thresholding were within approximately 14% of the physically-based value, which in itself could be viewed as a positive observation. Another way to look at the data is that 6 estimated porosities ended up lower, sometimes considerably so, than 52%, while 12 estimates overshot the porosity, by up to 20.71%. Therefore, there appears to be a tendency, at least by some of the experts, to overestimate the porosity. This tendency does not appear to be at all correlated with the level of familiarity of the experts with soils.

3.2. Comparison of automatic algorithms outputs

The results just described indicate that experts, relying on various combinations of subjective procedures (e.g., manual contrast adjustment, visual thresholding) and automated steps, produced thresholded images that in some cases vary significantly with respect to simple quantitative measures, like porosity. In this context, it seems worthwhile to enquire whether the situation would have improved had the experts relied more systematically on automated methods. Of course, individual experts could choose different algorithms, but it might be the case that, overall, automated algorithms yield consistently less variable outcomes. Therefore, comparison of common automated algorithms would be instructive.

Since the Pathway sub-image corresponds only to a very small part of the original Pathway image, one should not compare directly the values found in the two tests for the relative pathway areas. In this case, as in the other images, only comparisons of the variability in each test are meaningful. In this context, a first observation is that, compared to the narrow range of experts' estimates for the relative pathway area, (Table 1 and Fig. 6a), the outputs of automatic algorithms are far more spread out, as they span the full spectrum

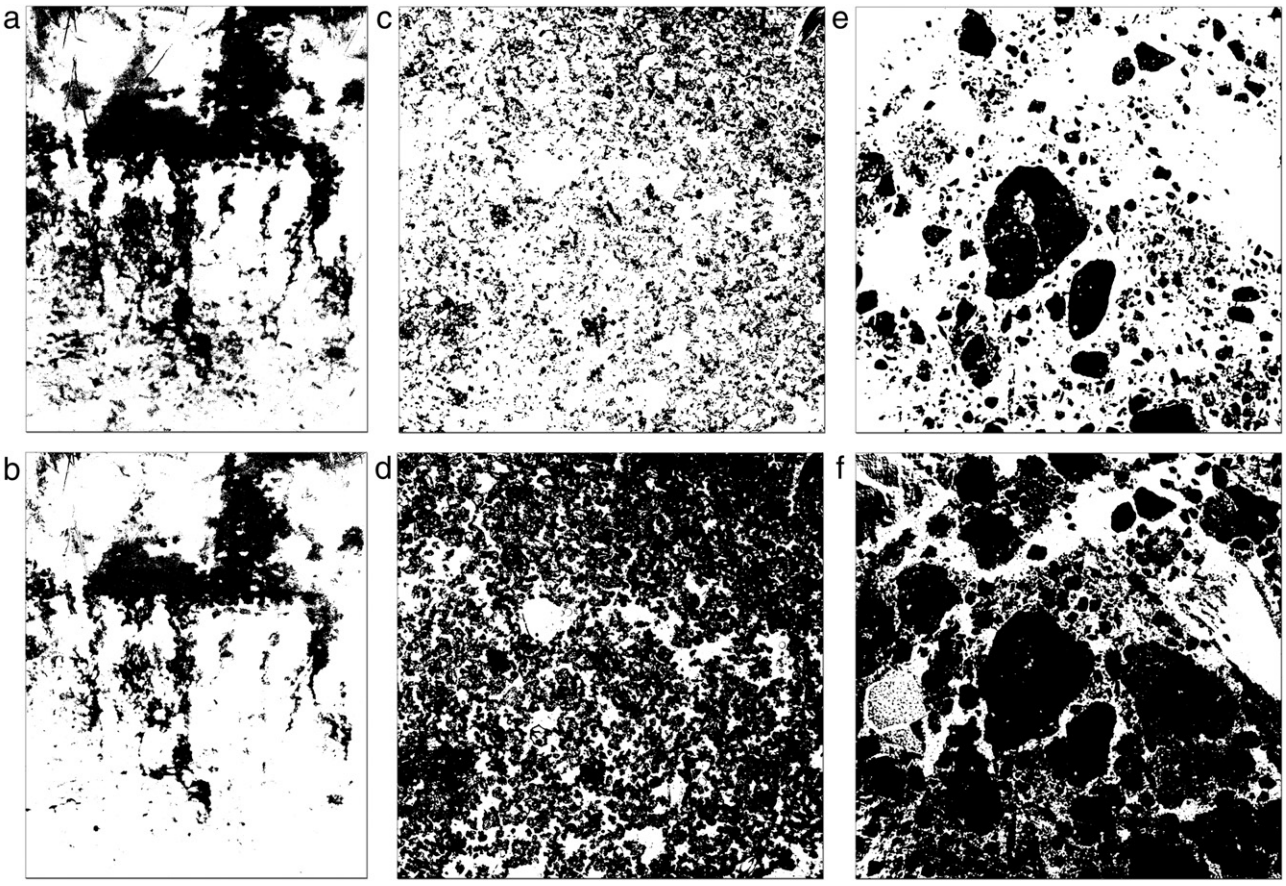

Fig. 5. (a) Thresholded test images with the highest (top) and lowest (bottom) relative pathway area (in %) or porosity (%). These images were thresholded by (a) expert 2, (b), expert 4, (c) expert 4, (d), expert 8, (e) expert 7 (b), expert 1, case 1.

from zero to a hundred percent (Fig. 6b). The extreme values are clearly due to the failure of some of the algorithms to converge to values that make physical sense. Some algorithms failed to converge altogether and their "outcomes" are not displayed in Fig. 6b. Six algorithms (Shape_Olivio, Cluster_Kittler, Attribute_Pikaz, Local_Palumbo, Local_Sauvola, and Local_White) converged but produced entirely white images, with no preferential pathway present. Conversely, five algorithms (Entropy_Brink, Attribute_O'Gorman, Spatial_Pal 1, Spatial_Pal b, Spatial Abutaleb) converged to entirely black images. Between these two extremes, pathway areas are spread from 23% to 98%, much more widely than in Fig. 6a. However, there is a distinct peak at 50 (with 9 images), and a concentration of half of all the images between 47 and 52.

Comparison of the outputs of automatic algorithms for the Thinsect sub-image (Fig. 7) yields some of the same results as for the Pathway subimage. The fact that there are almost no miss-convergence issues (just one, with Attribute_O'Gorman, yielding a zero porosity), suggests that the features of this sub-image, e.g., the shape of its histogram or its texture, correspond better to the various hypotheses underlying the different algorithms. Besides a peak with 5 images at a porosity of 69%, and two smaller peaks with 3 images each at 66% and 78%, there is considerable scatter in the porosity values obtained, definitely between 50 and 96%.

With respect to the lack of convergence, the situation appears to be exactly opposite with the Nanotom subimage. As with the Pathway subimage, this may be due to the fact that the histogram of this sub-image exhibits only one peak and shows no evidence of bimodality (mixing of two distributions), as is explicitly assumed by many thresholding algorithms, in one form or another. Out of 36 non-convergent algorithms, 18 produced aberrant results, with either zero, 99% or 100% porosity. The five algorithms resulting in zero porosity are Shape_sesan, Cluster_Jawahar a, Entropy_Pun b, Entropy_Brink, and Attribute_O'Gorman. The 9 algorithms resulting in 100% porosity are Shape_Olivio, Shape_Ramesh, Entropy_Yen, Entropy_Sahoo, Attribute_Sezgin, Local_White, Local_Palumbo, Local_Kamel, and Local_Sauvola. Besides these extremes, there is again a peak at an intermediate porosity value (57%, 5 images), which is slightly above, but within a few percent of, the physically-based porosity estimate for the soil sample. There is also a slight concentration around the peak, (but less marked in this case), and a considerable spread of estimated porosities, between 42 and 88%.

Further analysis of the results depicted in Figs. 6–8 shows that Attribute_O'Gorman fails to converge in all 3 subimages, and that additionally, Shape_Olivio, Local_White, Local_Palumbo, Local_Sauvola, and Entropy_Brink fail to converge for the Pathway and Nanotom sub-images. Besides that, however, there seems to be little noticeable pattern in the porosities obtained with the different algorithms for the 3 sub-images. For example, the algorithm Spatial_Abutaleb yields a high porosity of 73% for sub-image Thinsect, but one of the lowest non-trivial porosities (11%) for the Nanotom sub-image.

A similar variation in thresholding outputs was observed recently by Iassonov et al. (2009), who applied 14 algorithms to industrial and synchrotron computed tomography images of macroporous soils, sand-bentonite mixtures, and precision glass beads. Even in the case of a sample of (presumably millimetric-size) glass beads, with a measured porosity of 0.508 and scanned with a synchrotron-based X-ray CT system at a resolution of 5.9 μm, the image-derived porosities still ranged from 0.499 to 0.576, i.e., varied by more than 15%! Iassonov et al. (2009) concluded that even the two thresholding methods that

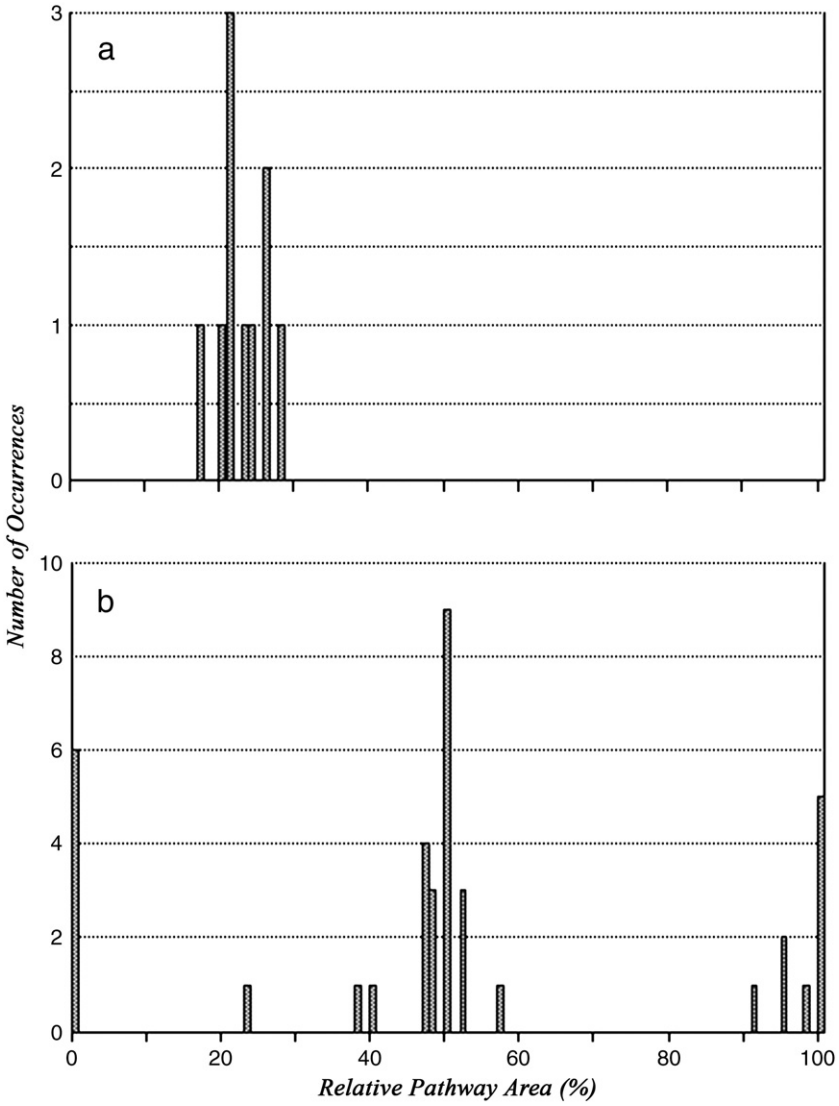

Fig. 6. Comparison of the frequency of occurrence of specific values of the relative pathway area for the Pathway image (a; expert output) and for the Pathway sub-image (b; automatic algorithms).

performed consistently better than the others still would require significant supervision by a skilled operator for routine use.

4. Conclusion and perspectives

The results presented in the previous section illustrate the fact that experts rely on very different approaches to threshold images of soils and that there is observer bias associated with this thresholding. This observer dependence is not likely to be alleviated by adoption of one of the existing automatic thresholding algorithms, some of which produce thresholded images that are clearly pathological (e.g., zero or 100% porosity) and can be discarded, while the remaining ones yield outcomes that are as variable as those of the experts. The upshot is that, at this point, analysis of the same image of a soil, be it a photograph or 3-dimensional X-ray CT data, can lead to very different results, without any assurance that any of them would be "correct" or, more pragmatically, be suited to the objective at To cope with this situation, a preliminary and relatively straightforward step would be to require authors who use image analysis to specify in detail what method they adopt to threshold images, and how sensitive the final outcome of their analysis is to the threshold value selected. Availability of pre- and post-thresholding images, perhaps as supplemental information linked to publications, would also be useful. However, neither of these measures alleviates the observer-dependence issue.

To remediate this situation, several avenues could be explored and deserve discussion. The first, suggested by Thompson et al. (1992) and reiterated recently by Sleutel et al. (2008), would be to agree on a set of "standards", which researchers carrying out image analysis in soil science would systematically use to benchmark the methodology they use. As simple as that concept might be, the selection of standards that would be representative of the range of situations in which image analysis is used in soil science is not straightforward. In addition, criteria would have to be developed to determine when a given methodology, benchmarked against one or more standards, can be considered "acceptable". Given the amount of consultation this approach would require among different research groups, it seems realistic to consider that progress along these lines is likely to be slow, even if standardisation efforts were launched immediately.

An alternative to standardizing methods of analysis, or to relying on agreed-upon physical standards, would be to complement each thresholding procedure with an accuracy assessment, as is done routinely in other fields that rely on image analysis. In the classification of vegetative patterns based on satellite or airborne imagery, researchers randomly select a sufficient number of areas in each mapped class, then go to the field and determine the true

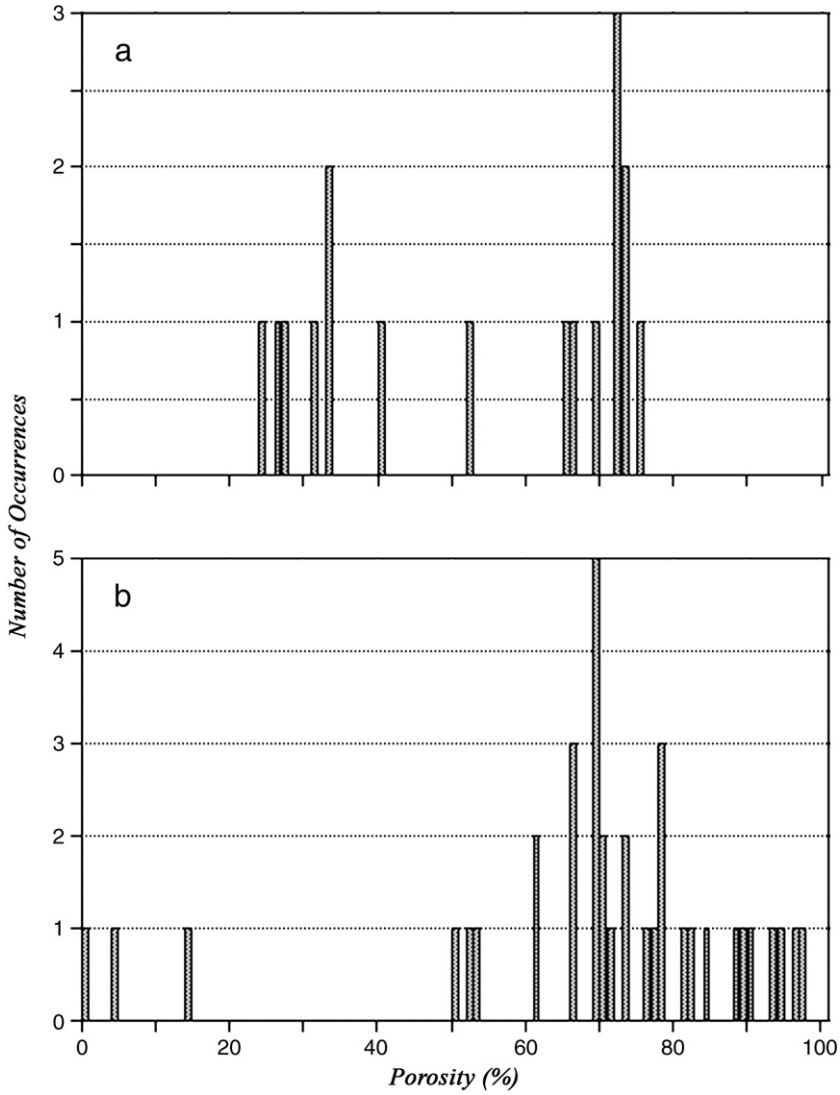

Fig. 7. Comparison of the frequency of occurrence of specific values of the porosity for the Thinsect image (a; expert output) and for the Thinsect sub-image (b; automatic algorithms).

vegetation at each location (e.g., Congalton, 1991; Congalton and Green, 1999; Laba et al., 2008, 2010). Comparison of predicted and actual classes leads to different quantitative estimates of the accuracy of a given classification.

To find out if this is feasible with images like those used in this research, a classical accuracy assessment procedure was carried out. The first step in this procedure is to convert the binary output file into a vector file. This was done using ArcInfo (a Geographical Information System by ESRI, Redlands, CA). The resulting polygons are stratified by size to eliminate small single-pixel sliver polygons. From these remaining polygons a "blind" stratified random sample of polygons from the two classes is chosen. The selected unlabeled polygons are overlayed on the original (un-thresholded) image and labeled.

This exercise was carried out for the Pathway image, thresholded by experts 2, 4 and 8. The results are summarized in a "contingency" table (Table 2), which, for each expert, provides the "expert-predicted" and the "directly-observed" classes associated with each polygon. The greyed diagonal boxes correspond to matches, whereas the off-diagonal boxes are mismatches. For each row, the ratio (in %) of the matches to the total number of pixels represents the "user's accuracy", and the ratio (in %) of the mismatches to the total number of pixels is typically referred to in the remote sensing literature as the "error of commission". The two similar parameters for the columns are known as the "producer's accuracy" and the "error of omission". The ratio of the number of matches to the total number of polygons is known as the "overall accuracy". It appears from the illustrative calculations in Table 2 that expert 4 obtained a far better error of commission (6%) for the dyed portion of the soil, and a reasonably good (11%) error of omission for the undyed soil. This leads to an overall accuracy of 70.6%, far better than that obtained by expert 2 and especially by expert 8.

Clearly, this method to estimate the accuracy of the thresholding of an image rests heavily on independent observations. In the case of the Pathway image, any visual hint of blue hue in the color image is a good indication of the presence of dye, and by zooming in on the selected polygons, it is virtually always possible to ascertain their likely class. For the other two images (Nanotom and Thinsect), simple reanalysis of the original pictures yields no reliable information, even after zooming in. In most cases, the polygons appear blurred, or their greyscale value is indeterminate, as if the polygons contained a mixture of the two classes. In these cases, one would have had to view the soils at a different scale and preferably with a different technique, for example by using a confocal laser microscope (DeLeo et al., 1997) to visualize soil thin sections and determine whether given polygons are associated with solids or pores.

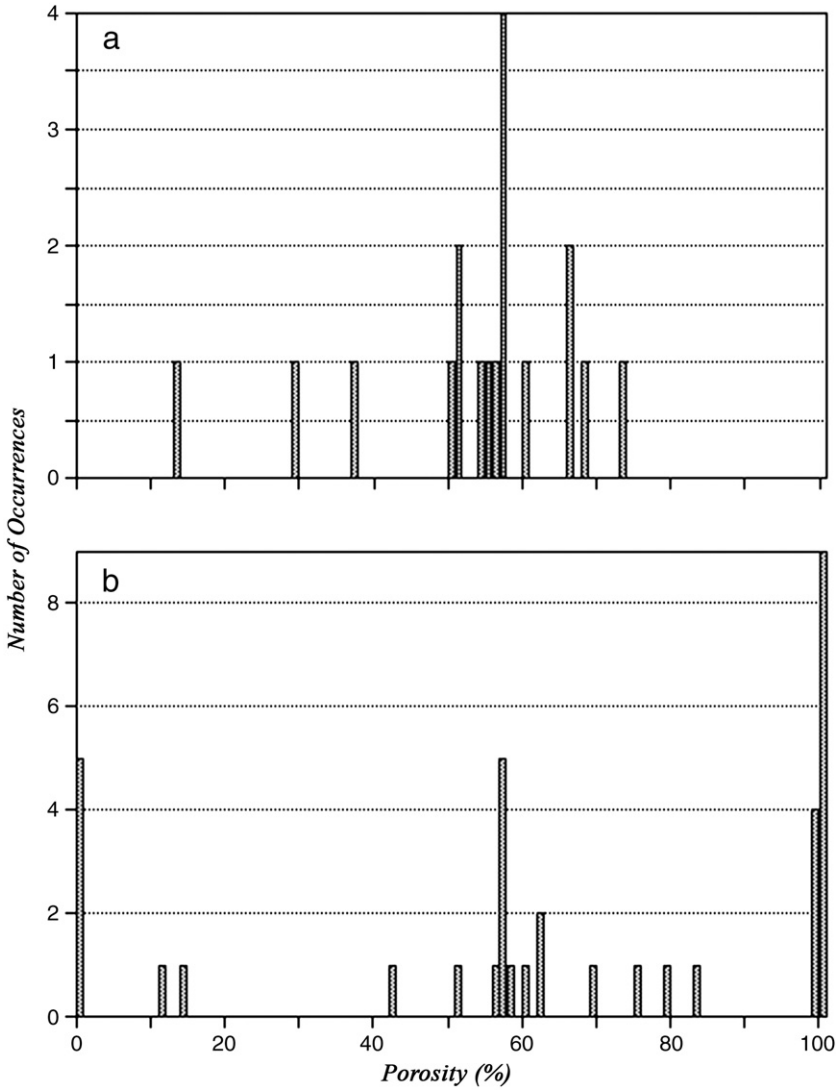

Fig. 8. Comparison of the frequency of occurrence of specific values of the porosity for the Nanotom image (a; expert output) and for the Nanotom sub-image (b; automatic algorithms).

Another promising avenue for progress is in the physical benchmarking of image analysis results. Once a soil sample is dry, and before any further treatment (e.g., resin impregnation) or measurement, it is easy to determine its porosity, for example via pycnometry. Alternatively, the porosity can be readily calculated from easily measurable variables such as bulk density and particle density. The value of the porosity thus obtained is then available to evaluate any representation of the pore geometry produced for example after thresholding CT data. Similarly, measurements of the moisture characteristic curve or the saturated hydraulic conductivity of soil samples can be compared to Lattice–Boltzmann simulations to benchmark the pore connectivity resulting from the use of specific thresholding protocols. A clear limitation of this type of physical benchmarking is that, strictly speaking, it can only apply to operations performed on 3-D data. This is because no physical measurements can be performed on 2-dimensional soil samples, and also because nothing guarantees that a given two-dimensional cross-section through a soil sample exhibits the same porosity or pore connectivity as the soil sample as a whole. Furthermore, unlike bulk physical measurements, image analysis is affected by the resolution of measurements. For example, pycnometry provides a volume estimate of the full continuum of pore sizes, whereas image analysis is limited by a lower size class related to the image resolution. For this reason, among others, one should always expect a discrepancy between the physically measured and estimated total porosity. Therefore, one would have to develop criteria to determine how big a numerical deviation could be tolerated before benchmarking results are no longer considered acceptable. Given all the information available with this approach, it would be conceivable to somehow constrain the thresholding process using the macroscopic, physical measurements. Again, in this case, research is needed to determine how this constraining should be carried out. For specific uses of the thresholded images, it might be appropriate to force the porosity of the resulting binary images to be exactly equal to the measured porosity, but for other applications, such a strict constraint might not make sense. In case two physical measurements are performed on the same soil sample, prior to image acquisition, the problem becomes one of multicriteria optimization, with an outcome probably conditional on the ultimate usage of the thresholded images.

Since there are so many unanswered questions related to the thresholding of soil images, perhaps a solution would be to explore ways to use these images without having to threshold them, in other words find ways to preserve the information that is contained in the greyscale values of the pixels or voxels of the image. This information is complex, and in some cases unreliable. The anti-aliasing filter of ordinary cameras, and Bayer interpolation, blur the images to an

Table 2
Contingency tables associated with the assessment of the accuracy of the thresholded Pathway images produced by experts 2, 4 and 8.

	Expert		Directly observed		Row total	User's accuracy (%)	Error of commission (%)	Overall accuracy (%)
			Undyed	Dyed				
Expert - predicted	2	Undyed	17	6	23	74	26	57.1
		Dyed	12	7	19	37	63	
		Column total	29	13	42			
		Producer's accuracy (%)	59	54				
		Error of omission (%)	41	46				
	4	Undyed	8	9	17	47	53	70.6
		Dyed	1	16	17	94	6	
		Column total	9	25	34			
		Producer's accuracy (%)	89	64				
		Error of omission (%)	11	36				
	8	Undyed	9	6	15	60	40	35.5
		Dyed	14	2	16	13	87	
		Column total	23	8	31			
		Producer's accuracy (%)	39	25				
		Error of omission (%)	61	75				

extent that may make individual pixel analysis meaningless. The use of common graphics applications and file formats means that information may be discarded or discounted. Images are typically displayed using 8 bits (1 part in 256) of precision per colour channel, i.e., the maximum widely supported by display devices. Contemporary scientific imaging apparatus may offer 16 bits (1 part in 65536) of precision per sample yet this is often linearly mapped to lower precision, either at the point of file export (conversion to some more portable format) or implicitly within many analyses (e.g. typical application software displays intensity histograms using at most a few hundred intervals). Such mapping may be satisfactory when directed by a human expert (e.g., the rejection of outliers is particularly important before precision is reduced) yet many users, unfamiliar with the implementation details of file formats and algorithms, remain oblivious to this process or its consequences. This reduction of image information amounts to a limited form of segmentation that can remove useful discriminating features from the intensity distribution. However, if such pitfalls are avoided, it does seem possible that physical properties can be estimated below the scale of individual image elements (pixels or voxels) by using image greyscale information. Such a "mean-field sub-voxel" approach is particularly challenging with regard to randomly heterogeneous materials such as soil; much investigation and verification remains to be done in this respect. Yet the appeal is clear: information is largely lost when one assumes a pixel or voxel to be entirely filled with either solid or pore space. Therefore, a richer description of a soil may result if the greyscale values are retained.

This is possible in a number of contexts. For example, instead of applying fractal geometry to binary images of soils, one could forego fractals completely and characterize soils quantitatively using multifractal measures, which under various conditions, can be meaningfully applied directly to greyscale images. Similarly, whereas traditional Lattice–Boltzmann (LB) models of water and solute transport in soils require a binary representation of the pore space, a number of alternative, mesoscopic formalisms, referred to as "grey" or "partial bounce-back" LB models (e.g., Chen and Zhu, 2008; Walsh et al., 2009), have been developed, in which the original information concerning the physical composition of individual pixels or voxels (given by their greyscale value) is retained, unchanged. Recently, Gommes et al. (2009) proposed a new method to measure the tortuosity of a porous or permeable material that is based on a geodesic reconstruction of the pore or permeation space from tomography data. Unlike earlier methods, this geodesic reconstruction can be applied directly to grayscale tomograms, without requiring any thresholding.

Another potentially fruitful avenue for research consists of trying to produce a binary representation of soils without having to carry out a 'standard' thresholding. Inspiration along those lines is provided by the fast-expanding literature on sub-pixel classification of remotely-sensed images and on super-resolution mapping, in which researchers have devised methods to handle what is referred to as the "mixed pixel" problem (e.g., Foody and Doan, 2007). By making some simplifying assumptions (e.g., on the uniform density of the solid phase), one could consider that the greyscale value of a pixel or voxel of a dry soil can be translated directly into a fraction of the space occupied by solid phase. For example, a voxel greyscale value of 75 in 8-bit representation (with 0 = black, and 255 = white) amounts to 71% (by volume) of solids and 29% of air. Instead of arbitrarily deciding that this voxel is all solid (as is done in standard thresholding), one could subdivide this voxel into n^3 sub-voxels of equal sizes, randomly fill as close to 71% of them with solids as possible, and consider that the remaining sub-voxels are air-filled. If the same procedure is applied to each voxel in the 3-D image, one of many possible sub-voxel resolution images, compatible with the measured image, is obtained. It can be used, e.g., for fractal characterization or LB modelling. By repeating the whole process multiple times, following a traditional Monte Carlo approach, probability distributions are eventually obtained for fractal dimensions or LB modelling outputs. For calculation-intensive uses of soil images, for example LB modelling or fungal growth (Falconer et al., 2008), this type of Monte Carlo approach will require significantly more computer processing time than is currently the case with thresholded images. However, the fact that Monte Carlo codes are particularly easy to parallelize opens up many possibilities for large-scale computations involving multiple processors, including "grid" computing (Renard et al., 2009) or the fast-developing "cloud" computing (Buyya et al., 2009).

It appears therefore that there are quite a few different, mutually compatible avenues to tackle the thresholding issue. Further research is undoubtedly needed to determine which one of these avenues is the most promising (and to find a way to deal with the other Achille's heel of image analysis: image resolution) but there is every hope that, within the not too distant future, ways will be found to transform image analysis into a very robust, observer-independent tool with which to study the properties of soils.

Acknowledgements

Thanks are due to Drs Andre Egbert and Hanna Jacobs from GE Sensing and Inspection Technologies Phoenix X-ray for their hospitality during D.V. Grinev's visit and for their assistance in obtaining the Nanotom image.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.geoderma.2010.03.015. These supplementary data, as well as downloadable copies of the three test images, are also accessible at http://www.simbios.ac.uk/muSIMCT/publicationCT.php.

References

Baveye, P., 2002. Comment on "evaluation of biofilm image thresholding methods". Water Research 36, 805–806.
Baveye, P., Boast, C.W., Ogawa, S., Parlange, J.Y., Steenhuis, T., 1998. Influence of image resolution and thresholding on the apparent mass fractal characteristics of preferential flow patterns in field soils. Water Resources Research 34, 2783–2796.
Baveye, P., Rogasik, H., Wendroth, O., Onasch, I., Crawford, J.W., 2002. Effect of sampling volume on the measurement of soil physical properties: simulation with X-ray tomography data. Measurement Science & Technology 13, 775–784.
Beaudet-Vidal, L., Fradin, V., Rossignol, J.P., 1998. Study of the macroporosity of reconstituted anthropic soils by image analysis. Soil & Tillage Research 47, 173–179.
Beulke, S., Brown, C.D., Dubus, I.G., Galicia, H., Jarvis, N., Schaefer, D., Trevisan, M., 2006. User subjectivity in Monte Carlo modeling of pesticide exposure. Environmental Toxicology and Chemistry 25, 2227–2236.
Bielders, C.L., Baveye, P., Wilding, L.P., Drees, L.R., Valentin, C., 1996. Tillage-induced spatial distribution of surface crusts on a sandy paleustult from Togo. Soil Science Society of America Journal 60, 843–855.
Boast, C., Baveye, P., 2006. Alleviation of an indeterminacy problem affecting two classical iterative image thresholding algorithms. International Journal of Pattern Recognition and Artificial Intelligence 20, 1–14.
Brown, C.D., Baer, U., Guther, P., Trevisan, M., Walker, A., 1996. Ring test with the models LEACHP, PRZM-2 and VARLEACH: variability between model users in prediction of pesticide leaching using a standard data set. Pesticide Science 47, 249–258.
Buyya, R., Yeo, C.S., Venugopal, S., Broberg, J., Brandic, I., 2009. Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility. Future Generation Computer Systems 25 (6), 599–616.
Chen, Y., Zhu, K., 2008. A study of the upper limit of solid scatters density for grey lattice Boltzmann method. Acta Mechanica Sinica 24, 515–522.
Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of the Environment 37 (1), 35–46.
Congalton, R.G., Green, K., 1999. Assessing the accuracy of remotely sensed data: principles and practices. Lewis Publisher, New York.
Cools, N., Delanote, V., Scheldeman, X., Quataert, P., De Vos, B., Roskams, P., 2004. Quality assurance and quality control in forest soil analyses: a comparison between European soil laboratories. Accreditation and Quality Assurance 9, 688–694.
Creamer, R.E., Bellamy, P., Black, H.I.J., Cameron, C.M., Campbell, C.D., Chamberlain, P., Harris, J., Parekh, N., Pawlett, M., Poskitt, J., Stone, D., Ritz, K., 2009. An inter-

laboratory comparison of multi-enzyme and multiple substrate-induced respiration assays to assess method consistency in soil monitoring. Biology and Fertility of Soils 45, 623–633.
Dathe, A., Baveye, P., 2003. Dependence of the surface fractal dimension of soil pores on image resolution and magnification. European Journal of Soil Science 54, 453–466.
DeLeo, P.C., Baveye, P., Ghiorse, W.C., 1997. Use of confocal laser scanning microscopy on soil thin-sections for improved characterization of microbial growth in unconsolidated soils and aquifer materials. Journal of Microbiological Methods 30, 193–203.
Deleporte, S., Hallaire, V., Tillier, P., 1997. Application of image analysis to a quantitative micromorphological study of forest humus. European Journal of Soil Biology 33, 83–88.
Elliot, T.R., Heck, R.J., 2007a. A comparison of optical and X-ray CT technique for void analysis in soil thin section. Geoderma 141, 60–70.
Elliot, T.R., Heck, R.J., 2007b. A comparison of 2D vs. 3D thresholding of X-ray CT imagery. Canadian Journal of Soil Science 87, 405–412.
Falconer, R.E., Bown, J.L., White, N.A., Crawford, J.W., 2008. Modelling interactions in fungi. Journal of the Royal Society Interface 5 (23), 603–615.
Fisher, C., 1971. The new quantimet 720. The Microscope 19, 1–20.
Foody, G.M., Doan, H.T.X., 2007. Variability in soft classification prediction and its implications for sub-pixel scale change detection and super resolution mapping. Photogrammetric Engineering and Remote Sensing 73 (8), 923–933.
Francaviglia, R., Capri, E., Klein, M., Hosang, J., Aden, K., Trevisan, M., Errera, G., 2000. Comparing and evaluating pesticide leaching models: results for the Tor Mancina data set (Italy). Agricultural Water Management 44, 135–151.
Garnier, P., Angulo-Jaramillo, R., DiCarlo, D.A., Bauters, T.W.J., Darnault, C.J.G., Steenhuis, T.S., Parlange, J.-Y., Baveye, P., 1998. Dual-energy synchrotron X-ray measurements of rapid soil density and water content changes in swelling soils during infiltration. Water Resources Research 34 (11), 2837–2842.
Gommes, C.J., Bons, A.J., Blacher, S., Dunsmuir, J.H., Tsou, A.H., 2009. Practical methods for measuring the tortuosity of porous materials from binary or Gray–Tone tomographic reconstructions. Aiche Journal 55, 2000–2012.
Hallaire, V., Cointepas, J.P., 1993. Image-analysis of the macroporosity of an orchard soil. Agronomie 13, 155–164.
Harris, K., Young, I.M., Gilligan, C.A., Otten, W., Ritz, K., 2003. Effect of bulk density on the spatial organisation of the fungus Rhizoctonia solani in soil. FEMS Microbiology Ecology 44 (1), 45–56.
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 Resources Research 45, W09415.
Jacobson, A.R., Dousset, S., Andreux, F., Baveye, P.C., 2007. Electron microprobe and synchrotron X-ray fluorescence mapping of the heterogeneous distribution of copper in high-copper vineyard soils. Environmental Science & Technology 41, 6343–6349.
Jongerius, A., Schoonderbeek, D., Jager, A., Kowalinski, S., 1972. Electro-optical soil porosity investigation by means of Quantimet-B equipment. Geoderma 7 (3–4), 177–198.
Laba, M., Downs, R., Smith, S., Welsh, S., Neider, C., White, S., Richmond, M., Sullivan, P., Philpot, W., Baveye, P., 2008. Mapping invasive wetland plants in the Hudson River National Estuarine Research Reserve using QuickBird satellite imagery. Remote Sensing of Environment 112 (1), 286–300.
Laba, M., Blair, B., Downs, R., Monger, B., Philpot, W., Smith, S., Sullivan, P., Baveye, P.C., 2010. Use of textural measurements to map invasive wetland plants in the Hudson River National Estuarine Research Reserve with IKONOS satellite imagery. Remote Sensing of Environment 114 (4), 876–886.
Lipsius, K., Mooney, S.J., 2006. Using image analysis of tracer staining to examine the infiltration patterns in a water repellent contaminated sandy soil. Geoderma 136, 865–875.
Marcelino, V., Cnudde, V., Vansteelandt, S., Caro, F., 2007. An evaluation of 2D-image analysis techniques for measuring soil microporosity. European Journal of Soil Science 58, 133–140.
Mooney, S.J., Morris, C., 2008. Morphological approach to understanding preferential flow using image analysis with dye tracers and X-ray computed tomography. Catena 73, 204–211.
Morris, C., Mooney, S.J., 2004. A high-resolution system for the quantification of preferential flow in undisturbed soil using observations of tracers. Geoderma 118, 133–143.
Murphy, C.P., Bullock, P., Turner, R.H., 1977a. Measurement and characterization of voids in soil thin-sections by image analysis .1. Principles and techniques. Journal of Soil Science 28, 498–508.
Murphy, C.P., Bullock, P., Biswell, K.J., 1977b. Measurement and characterization of voids in soil thin-sections by image analysis .2. Applications. Journal of Soil Science 28, 509–518.
Murphy, C.P., McKeague, J.A., Bresson, L.M., Bullock, P., Kooistra, M.J., Miedema, R., Stoops, G., 1985. Description of soil thin sections: an international comparison. Geoderma 35, 15–37.
Nawrath, R., Serra, J., 1979a. Quantitative image analysis — applications using sequential transformations. Microscopica Acta 82, 113–128.
Nawrath, R., Serra, J., 1979b. Quantitative image analysis—theory and instrumentation. Microscopica Acta 82, 101–111.
Ogawa, S., Baveye, P., Boast, C.W., Parlange, J.Y., Steenhuis, T., 1999. Surface fractal characteristics of preferential flow patterns in field soils: evaluation and effect of image processing. Geoderma 88, 109–136.
Ohrstrom, P., Persson, M., Albergel, J., Zante, P., Nasri, S., Berndtsson, R., Olsson, J., 2002. Field-scale variation of preferential flow as indicated from dye coverage. Journal of Hydrology 257, 164–173.
Ohrstrom, P., Hamed, Y., Persson, M., Berndtsson, R., 2004. Characterizing unsaturated solute transport by simultaneous use of dye and bromide. Journal of Hydrology 289, 23–35.
Otten, W., Gilligan, C.A., 2006. Soil structure and soil-borne diseases: using epidemiological concepts to scale from fungal spread to plant epidemics. European Journal of Soil Science 57, 26–37.
Papadopoulos, A., Bird, N.R.A., Mooney, S.J., Whitmore, A.P., 2008. Fractal analysis of pore roughness in images of soil using the slit island method. Vadose Zone Journal 7, 456–460.
Pendleton, D.E., Dathe, A., Baveye, P., 2005. Influence of image resolution and evaluation algorithm on estimates of the lacunarity of porous media. Physical Review E 72 (4) Art. No. 0413062005.
Persson, M., Olsson, J., 2008. Scaling analyses of high-resolution dye tracer experiments. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques 53, 1286–1299.
Persson, M., Haridy, S., Olsson, J., Wendt, J., 2005. Solute transport dynamics by high-resolution dye tracer experiments—image analysis and time moments. Vadose Zone Journal 4, 856–865.
Renard, P., Badoux, V., Petitdidier, M., Cossu, R., 2009. Grid computing for earth science. EOS, Transactions, American Geophysical Union 90 (14), 117–119.
Ringrose-Voase, A.J., Bullock, P., 1984. The automatic recognition and measurement of soil pore types by image analysis and computer programs. Journal of Soil Science 35, 673–684.
Sager, M., 1999. Current interlaboratory precision of exchangeable soil fraction measurements. Accreditation and Quality Assurance 4, 299–306.
Sezgin, M., Sankur, B., 2004. Survey over image thresholding techniques and quantitative performance evaluation. Journal of Electronic Imaging 13, 146–168.
Sleutel, S., Cnudde, V., Masschaele, B., Vlassenbroek, J., Dierick, M., Van Hoorebeke, L., Jacobs, P., De Neve, S., 2008. Comparison of different nano- and micro-focus X-ray computed tomography set-ups for the visualization of the soil microstructure and soil organic matter. Computers & Geosciences 34, 931–938.
Tarquis, A.M., Heck, R.J., Grau, J.B., Fabregat, J., Sanchez, M.E., Anton, J.M., 2008. Influence of thresholding in mass and entropy dimension of 3-D soil images. Nonlinear Processes in Geophysics 15, 881–891.
Thompson, M.L., Singh, P., Corak, S., Straszheim, W.E., 1992. Cautionary notes for the automated analysis of soil pore-space images. Geoderma 53, 399–415.
Vogel, H.J., Cousin, I., Ippisch, O., Bastian, P., 2006. The dominant role of structure for solute transport in soil: experimental evidence and modelling of structure and transport in a field experiment. Hydrology and Earth System Sciences 10, 495–506.
Walsh, S.D.C., Burwinkle, H., Saar, M.O., 2009. A new partial-bounceback Lattice–Boltzmann method for fluid flow through heterogeneous media. Computers and Geosciences 35 (6), 1186–1193.
Wantanaphong, J., Mooney, S.J., Bailey, E.H., 2006. Quantification of pore clogging characteristics in potential permeable reactive barrier (PRB) substrates using image analysis. Journal of Contaminant Hydrology 86, 299–320.