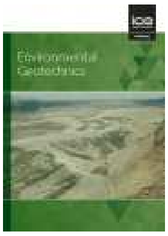
Article:
Xu, L, Xu, R, Shashank, BS et al. (3 more authors) (2024) An improved quantification method for characterisation of clay microstructure using SEM. Environmental Geotechnics, 11 (4). pp. 319-338. ISSN 2051-803X

https://doi.org/10.1680/jenge.21.00036



Environmental Geotechnics

An Improved Quantification Method for Characterization of ClayMicrostructure Using SEM

ENGE-2021-036-R3 | Paper

Submitted on: 16-11-21

Submitted by: Riqing Xu, Xiaohui CHEN, Liyang Xu, B. S. Shashank, Jianlin Yu, Chuangzhou Wu

Keywords: FABRIC/STRUCTURE OF SOILS, GEOENVIRONMENT, POROUS-MEDIA CHARACTERISATION

Information of the authors

Corresponding author: Riqing Xu*

Professor, Research Center of Coastal and Urban Geotechnical Engineering, Zhejiang University, China; Engineering research center of urban underground space development of Zhejiang Province, China.

Email: xurq@zju.edu.cn

Corresponding author: Xiaohui Chen*

Professor, School of Civil Engineering, University of Leeds, Leeds, UK

Email: x.chen@leeds.ac.uk

First author: Liyang Xu

PhD student, Research Center of Coastal and Urban Geotechnical Engineering, Zhejiang University, China; Engineering research center of urban underground space development of Zhejiang Province, China.

Email: 11712037@zju.edu.cn

Co-author: B. S. Shashank

Assistant Professor, Dept. of Civil Engineering, Birla Institute of Technology and Science (BITS) Pilani, Pilani-333031, India

Email: shashank.bs@pilani.bits-pilani.ac.in

Co-author: Jianlin Yu

Associate Professor, Research Center of Coastal and Urban Geotechnical Engineering, Zhejiang University, China; Engineering research center of urban underground space development of Zhejiang Province, China

Email: yujianlin72@126.com

Co-author: Chuangzhou Wu (Ph.D, P.E)

Professor, Institute of Port, Coastal, and Offshore Engineering, Ocean College, Zhejiang University, Zhoushan 316021, China

Email: ark_wu@zju.edu.cn

An Improved Quantification Method for Characterization of Clay Microstructure Using SEM

Liyang Xu[1,2], Riqing Xu*[1,2], B. S. Shashank[3], Xiaohui Chen*[4], Jianlin Yu[1,2], Chuangzhou Wu[5]

[1] Research Center of Coastal and Urban Geotechnical Engineering, Zhejiang University, China

[2] Engineering research center of urban underground space development of Zhejiang Province, China.

[3] Dept. of Civil Engineering, Birla Institute of Technology and Science (BITS) Pilani, India

[4] School of Civil Engineering, University of Leeds, Leeds, UK

[5] Institute of Port, Coastal, and Offshore Engineering, Ocean College, Zhejiang University, Zhoushan, China

Abstract: Clay is a material widely used in environmental geotechnics. Quantification of clay microstructure properties using SEM images is intuitively the simplest way, but the challenges remain in extracting more credible quantified micro-parameters that can represent the clay holistic macroscopic properties from SEM images and with higher efficiency. In this paper, an improved quantification method was proposed to deal with this issue. The influence of different milling and drying methods on quantification results was analyzed to improve reliability and efficiency. The *A*-*K* threshold determination method (*A*-*K* method) was proposed to calculate the optimal threshold in image processing. Also, the interval estimation method was introduced to obtain the optimal magnification and number of images. Then, this quantification method was applied to Hangzhou clay. Results show that anisotropy in clay microstructure should be considered when choosing the observed surface. And there is no significant difference in micro-parameters between liquid nitrogen frozen-vacuum drying and critical point drying. A comparison among the *A-K* method, artificial method, and theoretical methods that can determine threshold shows that the proposed *A-K* method combines the advantages of all methods. The optimal magnification and number of images can be determined by mathematical theory, which also improves reliability and efficiency.

Keywords:

Fabric of soils; Geoenvironment; Porous-media characterisation; SEM

List of notations

A_p real area of the particle

A' circumcircle area of clay particle

A_1 pore area

A_0 area of total SEM image

D_v fractal dimension

F mean shape factor

F_i shape factor for each clay particle

G_s specific gravity

H_m orientation probability entropy

k number of equally divided areas in the whole particle direction range

L actual perimeter of the clay particle

m total number of clay particles in the SEM image

n_i image porosity

n the porosity of clay

$N(r)$ number of square boxes

$N(\mu, \sigma^2)$ normal distribution

P perimeter of a circle that has the same area as the clay particle

$P_i(\alpha)$ percentage of soil particles whose directions α belong to a specific range

r side length of a square box

R_0 roundness

S^2 variance of the sample $X = (X_1, \cdots X_n)$

$t_{n-1}\frac{\alpha}{2}$ upper quantile

n_{3-D} 3-D surface porosity

V_{pore} volume of pore

$\bar{X}$ mean value of the sample $X = (X_1, \cdots X_n)$

α confidence coefficient

α_p orientation of soil particle

γ unit weight

μ unknown mean value of $N(\mu, \sigma^2)$

σ^2 unknown standard deviation of $N(\mu, \sigma^2)$

w water content

w_L liquid limit

w_P plastic limit

1. Introduction

Clay is a common material widely used in geoenvironmental engineering (Alba et al., 2009; Dohrmann et al., 2013; Tournassat et al., 2015), ranging from landfill clay liner to the barrier in the nuclear waste disposal (Chen et al., 2012; Estabragh et al., 2018; Toprak et al., 2018) (Fig.1). In engineering practice, it's crucial to understand the physical and mechanical properties of these clay barriers, such as the strength, mechanical stability, and hydraulic conductivity. The macro-behaviour of clays depends chiefly on their microstructure. For example, the shear strength of clay can be improved by increasing inter-cluster cementation bonding and reducing the pore space (Horpibulsuk et al. 2010). And microstructural changes in the clays may cause settlement and bearing capacity problems (Oztoprak and Pisirici, 2011).

Therefore, it's significant to study the micro-mechanisms of clay behaviour by effectively quantifying clay microstructure parameters such as image porosity, fractal dimension, mean shape factor, et al. (Pusch and Weston, 2003; Bennett and Hulbert M, 2012).

The efforts to explore the clay microstructure began since Terzaghi (1925) first proposed the concept of soil microstructure. And Scanning Electron Microscope (SEM) is one of the most popular equipment for micro research (Bohor and Hughes, 1971; Horpibulsuk et al., 2010; Liu et al., 2011; Liu et al., 2017). However, there are still many problems in the quantification of clay microstructure, and the challenges remain in obtaining ‘reliable’ quantified information of clay microstructure effectively and easily (Liu et al., 2017; Di Remigio et al., 2018). Reliable quantification means that the micro-parameters obtained from a limited number of images with a selected magnification should represent the clay microstructure more credibly.

The main difficulties in the quantification process of clay microstructure by SEM are shown in Fig.1 (Mazumder et al., 2018; Yang and Liu, 2019). The first issue is the clay sample preparation, and the sample milling and drying methods are discussed in this paper. There are many ways to mill and dry the samples for SEM experiments (Kaczyński and Trzciński 1997; Kjellsen et al., 2003; Janecek and Robert, 2016; Bangaru et al., 2019). The optimal method for clay samples is still uncertain. Another issue is the reliable representativeness of the results obtained from SEM images. The reliable representativeness is mainly related to three aspects as follows: (1) the magnification of SEM images; (2) the number of SEM images; (3) image processing.

In terms of the magnification, too low magnification will make it hard to distinguish between pores and particles in the SEM image, while too high magnification makes it difficult to recognise the micro-structural features since the whole image is occupied only by a few particles or pores (Lin et al. 2018; Bangaru et al. 2019; Wu and Chu, 2020). As for the number, since SEM images can only show the features of clay in the order of microns, a large number of SEM images are required to improve the representativeness of the quantification results, but an increased number of SEM images also costs longer time and higher fee. There are many tools and approaches for image processing, such as MATLAB®, Image-Pro Plus software, and Pores / Particles and Cracks Analysis System (PCAS), to extract micro-parameters from the SEM images (Cox and Budhu, 2008; Prakongkep et al., 2010; Liu et al., 2011; Taillon et al., 2018). The key for these image processing methods is the thresholding used to identify the particles and the pores in the binary image. However, it's still a big challenge to determine the optimal threshold value for SEM images of clay, and there is significant scope for advancements and refinements of the used techniques (Taillon et al., 2018).

The aim of this paper is to develop an improved quantification method for the characterization of clay microstructure using SEM. The method should be more reliable and more efficient. The influences of different milling methods and drying methods on the clay microstructure are quantitatively characterized by micro-parameters to achieve these goals. Besides, the following three aspects are discussed: 1. the magnification of

SEM image. 2. the number of SEM images. 3. image processing. Subsequently, the clay from Hangzhou, China, was studied as an example.

2. Factors influencing quantification of clay microscopic characterization

Influencing factors include SEM sample preparation and SEM-based quantitative analysis. SEM sample preparation mainly involves milling and drying. The SEM-based quantitative analysis mainly focuses on image processing, magnification, and number of images.

2.1 Milling directions

The milling direction may affect the clay microstructure characterization since clay is an anisotropic material at the macro scale. The clay sample considered in this study for SEM analysis was a small cuboid whose length, width, and depth were about 5 mm, 5 mm, and 2 mm, respectively. This SEM sample is milled from a remoulded sample, which is a cylindrical specimen with 39.1 mm diam and 80 mm height. The method to mill is a big issue.

2.2 Drying techniques

The clay sample should be dry when testing in the vacuum environment of SEM. But if the clay is dried, the pores between clay particles would become smaller (Burton et al., 2015; Sun and Cui, 2018), then the distance between particles would become smaller, which could change the interparticle forces such as van der Waals and capillary interactions, and hence the arrangement of the particles. Therefore, the clay microstructure would be modified, resulting in parameters corresponding to an entirely different microstructure. The method to dry is also a big issue.

2.3 Magnification and quantity of SEM image

Several decisions are needed to be made before imaging (Trzciński, 2004). Among them, the optimal selection of magnification and the number of images were most vital. The key to the problem was to improve efficiency and reliability so that with a limited number of images, all representative features have been captured and at the selected appropriate magnification. In this paper, the interval estimation method is introduced to make an optimal selection.

2.4 Image processing

The SEM image, for example, Fig.2a, is a grey photo of clay, whose relatively dark region represented pore space, and the relatively lighter region referred to as clay particle. The thresholding of the image is carried out to separate the two parts. The total range of threshold values is from 0 to 255. If the threshold value is set 0 in the PCAS program, the whole image is white, which means that all elements in the image are clay particles (Fig. 2b). On the contrary, if the threshold value is set to be 255, the whole image is black, i.e. all elements in the image are regarded as pores (Fig. 2c).

In the image processing for general images, there are two approaches to calculate the threshold value, one is the artificial approach, the other is the theoretical approach. The artificial approach compares the binarized images with the original SEM image by naked eyes. This artificial method can obtain the threshold directly, but it is time-consuming and error-prone. As for the theoretical approach, there are many methods, such as the method of iterative global thresholding (Shaikh et al. 2011), Otsu's method (Otsu, 1979), and the method of local properties based thresholding (Cheremkhin and Kurbatova, 2019). The theoretical approach can compute the threshold using algorithms such as iteration and variation of iteration (Bansal and Maini, 2013). However, many existing theoretical methods are not suitable for SEM images of clay since the particles and pores in the images are very small.

To combine the advantages of both the artificial approach and the theoretical approach, and avoid their disadvantages, the *A-K* threshold determination method is proposed in this paper. This new method also takes the natural porosity into account, and this method can improve the efficiency and reliability of image processing. It should be noted that the natural porosity is the fraction of the volume of pores over the total clay volume, and the natural porosity can be measured in a conventional geotechnical experiment.

3. Methodology for quantitative analysis of microstructures in SEM

The four main factors mentioned above should be considered to improve the clay quantification method.

3.1 The method of milling

The key issue in milling operation is to obtain an undisturbed surface. To study the influence of milling directions on the quantification result, samples from cross-section and vertical-section of Hangzhou clay were compared. The cross-section is parallel to the bottom base of the cylindrical remoulded clay sample, while the vertical-section is perpendicular to the bottom base. The methods to obtain the cross-section surface and vertical-section surface are shown in Fig. 3a and Fig. 3b, respectively. At first, the remoulded clay sample was broken into a strip about 20 mm (length) × 5 mm (width) × 5 mm (height) from the center. Then two parallel grooves of 0.5mm depth were carved in the middle of this strip. The distance between the grooves was 2 mm. After that, the sample was carefully broken by hand along the groove to get the undisturbed surface for SEM observation (SEM sample).

3.2 The method of drying

To study the influence of drying techniques on the quantification result, samples dried by four popular methods, including liquid nitrogen frozen-vacuum drying, critical point drying, air drying, and oven drying (Delage and Lefebvre G,1984; Dey et al., 1989; Janecek and Robert, 2016; Lindroth et al., 1988). The steps are detailed in the following lines:

In the liquid nitrogen frozen-vacuum drying method (Trzcinski, 2004), SEM Sample was immersed in liquid nitrogen for 1 min before being placed in the vacuum dryer for 20 h (equipment model: ALPHA1-4). As for the critical point drying method (Lawrence, 1979), the sample was dehydrated in a graded ethanol series (5%, 15%, 50%, 75%, 100%) at an interval of 20 min each, and subsequently dried in a critical point drying apparatus for 48 h (equipment model: Quorum-Emitech K850), which is covered with liquid carbon dioxide. For the air drying method (Youn and Tonon, 2010), the SEM sample was kept in the natural environment for 48 h. In the oven drying method, the SEM sample was dried in the oven directly for 12 h at 105 ℃ (Korpa, 2006).

3.3 Image processing for clay microscopic quantification

3.3.1 *A-K* threshold determining method

A novel *A-K* method for determining threshold value has been developed. This method combines the advantages of the theoretical approach and artificial approach and thus can improve the efficiency and reliability of image processing. The method consists of two steps: the first step is the prediction of the threshold range, and the second step is the determination of an optimal threshold value. It should be noted that the threshold range varies from the start-point to the end-point. For example, a threshold range is [20, 100], which means that 20 is the start-point, and 100 is the end-point.

(1) The prediction of the threshold range

This step is aimed at finding the start-point and the end-point within which the threshold value lies. For the convenience of description, the start-point is assumed to be the *A*-value, and the end-point is assumed to be the *K*-value. The *A*-value is confirmed on the principle that image porosity obtained from image processing is close to natural clay porosity, while *K*-value could be calculated by Otsu's method (Otsu, 1979).

Since the SEM image reflected the macroscopic properties of the clay, the three-dimension surface porosity (3-D surface porosity) should be close to the natural clay porosity. Clay is assumed to be homogeneous, so as shown in Fig.4a, the *A*-value can be obtained from the following formula:

$$n_{natural} = n_{3-D},\ n_{3-D} = \frac{V_{pore}}{V_{total}} = \frac{\int_0^a pdg}{\int_0^{255} pdg} \tag{1}$$

where, a is the *A*-value, $n_{natural}$ is the natural clay porosity, n_{3-D} is the 3-D surface porosity, V_{pore} is the volume of the pore, and V_{total} is the total volume of clay sample. p is the number of pixels, g is the grey value.

The natural porosity can be measured from the conventional geotechnical test methods, while the 3-D surface porosity can be obtained from the grey histogram (Fig. 4b) based on the 3-D surface clay model (Fig. 4d) of the original SEM image (Fig. 4e).

In fact, the grey histogram is a kind of statistical analysis of the 3-D surface model. The grey histogram represents the number of pixels with a certain grey value in the SEM image. Besides, in the grey histogram, if the threshold value is a, the grey range of (0, a) represents the pores, and the grey interval of [a, 255] represents the particles (Fig. 5a).

The 3-D surface porosity of the SEM image is the ratio of the pore area to the total area. The pore area can be calculated by integrating the interval of (0, a) in the grey histogram, while the total area of the grey histogram equalled the whole area of the SEM image. In the example shown in Fig. 4, the SEM image (Fig.4e) of a clay sample has a natural porosity of 0.311. If the threshold is 43 (the value of the threshold is accurate to the single-digit), the 3-D surface porosity of the SEM image is 0.319, and 0.319 is the closest to the natural porosity 0.311. Therefore, the A-value is 43. Subsequently, the K-value of the threshold range (Fig.4c) is calculated by Otsu's method. For the SEM image in Fig.4e, the K-value is 76, and hence, for this example considered here, the threshold range would be from 43 to 76.

(2) The determination of an optimal threshold value

The determination of the optimal threshold value within the threshold range calculated above was carried out with the help of the PCAS program. The greyscale in the PCAS program was adjusted within the predicted threshold range above. At the same time, an artificial comparison with naked eyes was conducted to select the greyscale that could best divide pore area and soil particle area.

In terms of the SEM image in Fig.4e, the artificial comparison was conducted in the threshold range of [43, 76] by moving the grey bar of PCAS to show the different binarized images of each threshold value. Typical comparative images for various threshold values within the identified range are shown in Fig.5b. After comparing these binarized images with the original SEM image, results showed that when the grey value was 55, the image division had the optimal effect, since the binarized image matched the original SEM image best. Therefore, the grey value of 55 was determined as the threshold value for this SEM image.

3.3.2 Measurement of micro-parameters

Based on the threshold value calculated above, the SEM image was converted into a binarized image, then the measurement of micro-parameters can be done with the help of PCAS software (Liu et. al, 2011, 2013 and Tang et al, 2012). Finally, the image porosity, mean shape factor, mean fractal dimension, roundness, and orientation probability entropy can be obtained.

3.4 Interval estimation for clay microscopic quantification

The method of interval estimation was proposed by Neyman (Neyman, 1937), aiming at ensuring that the estimated interval had the highest accuracy with certain reliability. In statistics, interval estimation is used to calculate an interval of possible values of an unknown population parameter with sample data. The definition of interval estimation is in the Appendix 1.

3.5 Micro-parameters for clay microscopic quantification

Fig. 6a shows three clay particles in the microstructure. Their particle shapes, particle directions, edge shapes, and edge complexity variations are quite different. These features can be quantified by micro-parameters. In this paper, five statistical parameters, i.e., image porosity n_i, mean shape factor F, mean fractal dimension D_v, roundness R_0, and orientation probability entropy H_m, are introduced to describe the fraction of pore, particle shape, smoothness of particle edge, complexity variation of particle edge and orientation distribution of clay particles, respectively. These micro-parameters are defined in the Appendix 2.

4. An example for characterization of clay microstructure

4.1 Materials

To illustrate the improved quantification method described above, a clay from Hangzhou, China, was employed as an example, which is a typical clay with low permeability and high natural water content. Clay's physical properties are reported in Table 1. The mineralogy and chemical composition of clay are provided in Table 2.

4.2 Methods

The methods mainly include four steps (i.e., Remolded sample preparation, sample preparation for SEM imaging, SEM imaging, and SEM image processing). Their detailed operations are presented in the Appendix 3.

4.3 Test plan

In this paper, four groups of tests were conducted. Each group has its unique test objective, and each group contains several SEM samples. The first group is used to compare different milling directions (Table 3). SEM samples from cross-section (1#, 3#, 5#, 7#) and vertical-section (2#, 4#, 6#, 8#) of 4 different remolded clay samples (I#, II#, III#, IV#) were tested. These SEM samples were air-dried, and 10 SEM images with magnification of 2000x were taken for each sample.

The second group aims at comparing different drying techniques. For this purpose, 4 SEM samples (9#, 10#, 11#, 12#) were prepared. All these SEM samples were made from remolded clay sample V#, whose void ratio and water content were 0.45 and 20% respectively. These SEM samples were dried in air (for 9# sample), hot-air oven (for 10# sample), critical point-device (for 11# sample), and liquid nitrogen frozen-vacuum device (for 12# sample), respectively. The cross-sections of these SEM samples were observed, and 10 SEM images with a magnification of 2000 times were taken for each sample.

The third group is prepared to compare different SEM image magnifications. Only one SEM sample (13#) was required here, which was obtained from clay sample VI#, whose void ratio and water content were 0.51 and 21%, respectively. SEM sample 13# was air-dried, and the cross-section was observed. For different magnifications, SEM images with magnifications of 500×, 1000×, 1500×, 2000×, 3000×, 5000× were obtained and analyzed.

The fourth group is used for comparing different quantities of SEM images. One SEM sample, which was used in the third group (i.e., 13#), along with another SEM sample (14#) was used for this purpose. Sample 14# is milled from the remolded clay sample VII#, whose void ratio and water content were 0.66 and 30%, respectively. Both were air-dried, and the cross-section was observed. 10 images with a magnification of 2000× were taken for both SEM samples.

4.4 Results of SEM imaging

4.4.1 Effect of different milling directions on clay microscopic quantification

The micro-parameters from the SEM imaging for cross-section and vertical-section samples are shown in Table 4, and the relative error in the micro-parameters between the cross-section and the vertical-section of 4 remolded clay samples (I#, II#, III#, IV#) are depicted in Fig. 6b.

The results show that the cross-section and vertical-section had similar image porosity n_i, and the max relative error between the two values was 1.5%. The probability entropy H_m of the two sections were also close, and the max difference was observed to be 0.4%. These similar results were not surprising given that the clay sample is homogeneous, and hence, the arrangement of clay particles is similar in each direction. On the contrary, larger variations were observed in other micro-parameters. The fractal dimension D_v of vertical-section was 3.6% bigger than cross-section, and the mean shape factor F of the cross-section was up to 12.8% higher than that of vertical-section. The max relative error in roundness between cross-section and vertical-section was 5.5%. These relatively big differences are because the three parameters are related to the particle shape. It's well known that most clay particles are flat. Therefore, the difference between the shape projected on cross-section and that on vertical-section is obvious.

4.4.2 Effect of different drying techniques on clay microstructure quantification

Volumetric shrinkage will occur when the clay is dried, and the degree of shrinkage varies under different drying techniques. The quantitative comparison of each method is shown in Table 5 and Fig. 7. Both the micro-parameters and the pore sizes are compared. The pore sizes include pore area, pore length, and pore width.

The data show that there was no significant difference in micro-parameters and in pore sizes between the two methods of liquid nitrogen frozen-vacuum drying and critical point drying, owing to that the original clay structure was maintained by these two methods. Compared with oven drying, the results of air drying were closer to the results of the former two drying methods mentioned above. For example, among all the micro parameters, the biggest difference between air drying and liquid nitrogen frozen vacuum drying was 15% for image porosity n_i. While the biggest difference between oven drying and liquid nitrogen frozen-vacuum drying was 51%, which is also shown by the image porosity n_i. Meanwhile, among all the pore sizes, the biggest difference between air drying and liquid nitrogen frozen was the maximum pore length, and the value was 23%. The biggest difference between oven drying and liquid nitrogen frozen-vacuum drying is the maximum pore length, while the value was 63%. These differences were because oven drying had a larger shrinkage than others.

4.4.3 Determination of optimal image magnification

As shown in Fig.8, the clay sample was placed in the SEM to get the images with different magnifications of 500×, 1000×, 1500×, 2000×, 3000×, 5000×. In these images, if the magnification is too low, the clay particles and pores are hard to be recognised. On the contrary, if the magnification is too high, then the whole image is occupied by large particles, which brings difficulties in recognising the details of microstructural features. Therefore, the optimal magnification should be between 500× and 5000× for this sample. The optimal magnification can be selected from this range with the help of interval estimation. The analysis of interval estimation, whose results are reported in Fig.9, is based on the micro-parameters that are listed in Table 6.

In Fig.9, there are three lines and six dots in each figure. The first line and the third line show the upper limit and lower limit of confidence interval (95% confidence level) for each micro-parameter, respectively, while the line in the middle shows the mean value of the micro-parameter. The dots are the micro-parameter value of different magnifications. The calculation of confident interval and mean value are based on all the SEM images considered in this experiment.

Based on these figures, there exists an optimal magnification for each parameter. For example, in Fig.9a, the magnifications of the dots that are located in the area between the upper line and the lower line are better than the magnifications of dots that are outside this area, since parameter values of dots inside are in the confidence interval. Furthermore, the magnification of the dot that is closer to the middle line is better than that of another dot that is farther, because there is a greater probability that the mean value can represent the real value of the micro-parameter.

Finally, the optimal magnification for the tested clay can be obtained, according to the five commonly used micro-parameters used in this study. If one magnification is always optimal in the five figures, then this magnification can be the best choice. Fig.9 shows that the dots of 1500× and 2000× were inside the confidence interval in all cases, while dots of other magnifications sometimes were out of the confidence interval. Therefore, the magnification of 1500× and 2000× can be considered to be more reliable than others. As for the best one, 1500× should be selected, since, in four of the five cases, the dot of 1500× was closer to the mean line than the dot of 2000×, which means that most micro-parameters were closer to the real value when the magnification is 1500×. Hence, the optimal magnification of Hangzhou clay was considered to be 1500×.

4.4.4 Determination of optimal image quantity

According to the results above, SEM images with a magnification of 1500× were used here to analyze the optimal quantity of images. In theory, when the quantity of SEM images is larger, the mean value of micro-parameters would be closer to the true value. But the time spent taking SEM photos would be more. The more efficient way was to take fewer images to get the most accurate values of the micro- parameter.

10 SEM images had been taken for the SEM sample 13#. Five micro-parameters of each image were calculated by the image processing method introduced above and are listed in Table 7. Later, the interval estimation was done based on these data. For each micro-parameter, interval estimation was carried out for the different quantities of images, such as 2 images, 3 images, or 10 images. These results are presented in Fig.10.

For each parameter in Fig. 10(a-e), there are 9 bars with a short vertical line of different lengths on the top. The height of each bar is the mean value of a certain image quantity, for example, the first bar shows the mean value of 2 images. While the short vertical line on the top of each bar represents the confidence interval, which means that there is a 95% possibility that the true value of the micro-parameter is in this interval. The length of these short vertical lines is useful because if the length is smaller, it will be more credible that the mean value can be used to represent the true value of the micro-parameter.

Fig. 10 shows that the length of confidence interval decreased as the number of images increased when the number of images was less than 5. From 2 to 5 images, there is 78%, 83%, 78%, 81%, 78% decline for n, F, H_m, D_v, R_0, respectively. But the declining trend was less significant when the number of images was more than 5, except for the mean shape factor F, which might be caused by accidental factors. And from 6 to 10 images, there is 23%, 37%, 38%, 36% decline for n, H_m, D_v, R_0, respectively.

Further, the percentage difference between adjacent two confidence interval lengths is defined as error ε. Assuming that when the error ε is less than 25%, the confidence interval is considered acceptable, indicates that the corresponding number of images is selected. The errors ε of each micro-parameters for SEM sample 13# are listed in Table 8. Results show that, with the exception of the mean shape factor F, the error ε is less than 25% when the quantity is equal to 5 or more. The error ε is abnormal in the mean shape factor F, that is, when the quantity is 7, the ε increases to 49%.

To verify these observations, another SEM sample 14# was used. The void ratio and water content of SEM sample 14# were 0.656 and 30%, respectively. The methodology adopted was the same as in the case of SEM sample 13#. Micro-parameters of each image are listed in Table 9, the results of interval estimation are shown in Fig.11, and the error ε are listed in Table 10. The results show that when the number is equal to 5 or more, the error ε are always less than 25% for all the micro-parameters. Therefore, 5 images and more than 5 images meet the requirements, hence, the smallest number of 5 can be chosen as the optimal number.

5. Discussion

To further demonstrate the reliability and efficiency of the proposed improved quantification method, more details about the 3-D surface model are presented to prove the reliability of the *A*-*K* method. Besides, comparison with other researches is discussed.

5.1 Demonstration for reliability and limitation of the 3-D surface model

The 3-D surface model in Fig. 4d is composed of an SEM image and the corresponding grey value of each point in the SEM image. The dimension of grey can reflect the spatial features of the structure surface on the 3rd dimension. This is because the grey value of the SEM image can reflect the distance from the structure surface to the imaging surface due to the imaging principle (as shown in Fig. 12). Fig. 12a shows the imaging principle of secondary electron imaging mode in the SEM, and Fig. 12b illustrates that the spatial features of the structure surface can be reflected by grey values.

In Fig.12a, the high-energy incident electrons were emitted into the clay sample by the electron gun in the SEM, then the secondary electrons in the clay particles can be excited. These secondary electrons can be captured by the detectors in the SEM, and finally, generate an SEM image with different brightness areas inside. Higher brightness area means more electrons accumulation, and more electrons mean a shorter distance from the structure surface to the imaging surface. Therefore, the bright areas in the SEM image represent clay particles because the secondary electrons were generated from clay particles. While the relative dark areas in the SEM image represent pores since the pores cannot generate electrons. Since the degree of brightness in SEM images can be described by the grey value, and the distance from the structure surface to the imaging surface describes the spatial features of the structure surface. Therefore, there is a relation between the grey value and the spatial features of the structure surface, as shown in Fig. 12b.

However, the porosity calculated by the 3-D surface models in Fig. 4 only describes the porosity of the structure surface, and it can't show the pores behind the clay particles on the surface. For instance, the pores may look quite small in the 2D image simply because its appearance is small but inside it may be very big, such as pore 1 in Fig. 12b.

5.2 Comparison with other thresholds determination methods

To verify the proposed *A-K* method is better than the typical theoretical methods. An example is given to compare different thresholds determination methods in the following. Fig. 13a is the original image, which is the same as Fig. 4e. The thresholds for Fig. 13a obtained through the *A-K* method (Fig. 13b), the method of iterative global thresholding (Fig. 13c), the Otsu's method (Fig. 13d), and the method of local properties based thresholding (Fig. 13e) were 55, 106, 76, 52 respectively. The different threshold values from these methods resulted in a great variation in the resultant binary image after processing. The dark area in Fig. 13b and Fig. 13c was much more than the image in Fig. 13a, while the dark area in Fig. 13d was a little lesser than Fig.13a. Meanwhile, Fig. 13b is the most similar to Fig. 13a. Besides, the difference in the binary image could also be characterized by image porosity n_i. As shown in Fig. 13f, n_i varied from 0.44 to 0.86. The natural porosity of Fig. 13a is 0.311, and n_i of Fig. 13b and Fig. 13e are much closer to the natural porosity than Fig. 13c and Fig. 13d.

5.3 Comparison with other quantification methods

Few studies have investigated the quantification method for SEM experiments systematically. Most of the available studies have raised the challenges in microstructure characterization, but still, need to be further studied to find the solutions. For example, the differences between these drying techniques are inconsistent with the research of other studies (Korpa, 2006; Osipov, 1985), and the importance of milling direction for credible quantification is also agreed most with Trzciński (2004). But in their research, no comparison of influence on micro-parameters (image porosity, mean shape factor, mean fractal dimension, roundness, and orientation probability entropy) was presented. In terms of the image magnification, Zhou et al. (2021) compare the magnification of 500×, 800×, 3000× and 20000×. The magnification of 800× is used for further quantitative analysis, but no theoretical reason was given. As for the number of images, Trzciński (2004) proposed that the accuracy of characterization can be improved by a bigger number of the analysed sample parts because the data from one part do not coincide with results from a different portion of the same SEM sample. But no further discussion on the number of images was provided. The improved quantification method in this paper presents solutions to these challenges.

6. Conclusions

This paper aims to propose a practical and economical method to improve the reliability and efficiency in the quantification technique of clay microstructure using SEM. The following conclusions are drawn:

(1) The *A*-*K* threshold determination method is proposed to determine the optimal threshold, combining the advantages of the artificial approach and theoretical approach. An example has been given to show that the binary image of the *A*-*K* method is more fit for the original SEM image than images processed by other typical threshold determining methods.

(2) There exists an optimal magnification and an optimal number of images that can enhance the experimental efficiency. The anisotropy in the clay microstructure should be considered for choosing milling directions. The quantification results are significantly affected by the drying techniques.

(3) This study presents a novel method to deal with the challenge of extracting quantified micro-parameters from SEM images with higher reliability and higher efficiency. As a result, an improved microstructure quantification method for Hangzhou clay can be suggested: liquid nitrogen frozen-vacuum drying, sample milled from the same direction, 1500× magnification, and 5 SEM images.

Data Availability Statement

All data, models, or code generated or used during the study are available from the corresponding author by request, including all raw data from the tests, all used test results, and codes used for the image analysis.

Acknowledgements

The authors would like to acknowledge the financial support provided by National Natural Science Foundation of China (No. 41672264, No. 42177141), the Key Research and Development Program of Zhejiang Province, China (No. 2019C03103) and the China Scholarship Council (201906320246). The authors would also like to thank Yu Peng, Lingling Li, and Jiaqi Jiang for their assistance in the experiment, thank Zhigen Wu for his class in paper writing, thank Chuangzhou Wu for their patient guidance in paper improvement, thank Xiaohui Chen and Yue Ma for their inspiration and suggestions in the environmental geotechnics.

References

1.Alba M D, Chain P ans Orta M M (2009) Chemical reactivity of argillaceous material in engineered barrier: rare earth disilicate formation under subcritical conditions. Applied Clay Science **43(3–4)**: 369-375.

2.Bangaru S S, Wang C, Hassan M, et al. (2019) Estimation of the degree of hydration of concrete through automated machine learning based microstructure analysis–A study on effect of image magnification. Advanced Engineering Informatics **42**: 100975.

3.Bansal S, Maini R (2013) A Comparative Analysis of Iterative and Ostu's Thresholding Techniques. International journal of computer applications, **66(12)**: 45-47.

4.Bennett R, Hulbert M (2012) *Clay microstructure*. D. Reidel publishing company, Dordrecht, Netherlands.

5.Bohor B F, Hughes R E (1971) Scanning electron microscopy of clays and clay minerals. Clays and Clay Minerals **19(1)**: 49-54.

6.Burton G J, Pineda J A, Sheng D et al. (2015) Microstructural changes of an undisturbed, reconstituted and compacted high plasticity clay subjected to wetting and drying. Engng Geol **193**: 63–373

7.Chen Y M, Shi J Y, Zhu W, et al. (2012) A review of geoenvironmental engineering. China Civil Engineering Journal **45(4)**: 165-182.

8.Cheremkhin P A, Kurbatova E A (2019) Comparative appraisal of global and local thresholding methods for binarisation of off-axis digital holograms. Optics and Lasers in Engineering **115**: 119-130.

9.Cox M R and Budhu M (2008) A practical approach to grain fotr quantification. Engineering Geology **96 (1–2)**: 1-16.

10.Dathe A, Eins S, Niemeyer J, Gerold G (2001) The surface fractal dimension of the soil–pore interface as measured by image analysis. Geoderma **103 (1–2):** 203-229.

11.Delage P, Lefebvre G (1984) Study of the structure of a sensitive Champlain clay and of its evolution during consolidation. Canadian Geotechnical Journal **21(1)**: 21-35.

12.Dey S, Basu T S, Roy B et al. (1989) A new rapid method of air-drying for scanning electron microscopy using tetramethylsilane. Journal of Microscopy **156(2)**: 259-261.

13. Di Remigio G., Rocchi I, Zania V (2021) New method for a SEM-based quantitative microstructural clay analysis-MiCA. Applied Clay Science **214**: 106248.

14.Dohrmann R, Kaufhold S and Lundqvist B (2013) The role of clays for safe storage of nuclear waste. In *Developments in Clay Science.*Elsevier, Amsterdam, Netherlands, vol.1, pp. 677-710.

15.Estabragh A R, Kholoosi M, Ghaziani F, et al. (2018) Mechanical and leaching behavior of a stabilized and solidified anthracene-contaminated soil. Journal of Environmental Engineering **144(2)**: 04017098.

16.Hemes S, Desbois G, Urai J L, De Craen M, Honty M (2013) Variations in the morphology of porosity in the Boom Clay Formation: insights from 2D high resolution BIB-SEM imaging and Mercury injection Porosimetry. Netherlands Journal of geosciences **92(4)**: 275-300.

17.Horpibulsuk S, Rachan R, Chinkulkijniwat A, et al. (2010) Analysis of strength development in cement-stabilized silty clay from microstructural considerations. Construction and building materials **24(10)**: 2011-2021.

18.Huang J, Liu X, Cao G (2017) Experimental Study of Chemical Composition of Hangzhou Soft Clay. 2017 International Conference on Advanced Materials Science and Civil Engineering (AMSCE 2017): 116-118.

19.Janecek M and Robert K, eds. (2016) Modern electron microscopy in physical and life sciences. Intech, Rijeka, Croatia. pp.167-177.

20.Kaczyński R and Trzciński J (1997) Ilościowa analiza mikrostrukturalna w skaningowym mikroskopie elektronowym (SEM) typowych gruntów Polski. Przegląd Geologiczny **45**: 721–726.

21.Kjellsen K O, Monsøy A, Isachsen K, Detwiler R J (2003). Preparation of flat-polished specimens for SEM-backscattered electron imaging and X-ray microanalysis—importance of epoxy impregnation. Cement and concrete research **33(4)**: 611-616.

22.Korpa A, Trettin R (2006). The influence of different drying methods on cement paste microstructures as reflected by gas adsorption: Comparison between freeze-drying (F-drying), D-drying, P-drying and oven-drying methods. Cement and Concrete Research **36(4)**: 634-649.

23.Lawrence G P, Payne D, Greenland D J (1979). Pore size distribution in critical point and freeze dried aggregates from clay subsoils. Journal of soil science 3**0(3)**: 499-516.

24.Lin L, Yan J, Chen G, et al. (2018) Does magnification of SEM image influence quantification of particulate matters deposited on vegetation foliage. Micron **115**: 7-16.

25.Lindroth M, Bell Jr P B. and Fredriksson B A (1988) Comparison of the effects of critical point drying and freeze drying on cytoskeletons and microtubules. Journal of Microscopy **151(Pt.2)**: 103-114.

26.Liu C, Shi B, Zhou J, et al.(2011) Quantification and characterization of microporosity by image processing, geometric measurement and statistical methods: application on SEM images of clay materials. Applied Clay Science **54(1)**: 97-106.

27.Liu C, Tang C S, Shi B, et al. (2013) Automatic quantification of crack patterns by image processing. Computers & Geosciences **57**: 77-80.

28.Liu X, Wang J, Ge L, et al. (2017) Pore-scale characterization of tight sandstone in Yanchang Formation Ordos Basin China using micro-CT and SEM imaging from nm-to cm-scale. Fuel **209**: 254-264.

29.Mahaney W C, Stewart A, Kalm V (2001) Quantification of SEM microtextures useful in sedimentary environmental discrimination. Boreas **30(2)**: 165-171.

30.Mazumder M, Ahmed R, Ali AW, et al. (2018) SEM and ESEM techniques used for analysis of asphalt binder and mixture: A state of the art review. Construction and Building Materials **186**: 313-329.

31.Ministry of Water Resources of the People's Republic of China (2019) Standard for geotechnical test method GB/T 50123-2019, National Bureau of Quality and Technical Supervision, Ministry of Construction in China, Beijing.

32.Neyman J (1937) Outline of a theory of statistical estimation based on the classical theory of probability. Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences **236(767)**: 333-380.

33.Osipov B И (1985) The nature of the strength and deformation properties of clay-like soils and rocks. Geological Publishing House, Beijing, China.

34.Otsu N (1979) A threshold selection method from grey-level histograms. IEEE transactions on systems, man, and cybernetics **9(1)**: 62-66.

35. Oztoprak S and Pisirici B (2011) Effects of micro structure changes on the macro behaviour of Istanbul (Turkey) clays exposed to landfill leachate. Engineering Geology **121(3-4)**: 110-122.

36.Prakongkep N, Suddhiprakarn A, Kheoruenromne I et al. (2010) SEM image analysis for characterization of sand grains in Thai paddy soils. Geoderma **156 (1–2)**: 20-31.

37.Pusch R, Weston R (2003) Microstructural stability controls the hydraulic conductivity of smectitic buffer clay. Applied Clay Science **23 (1–4)**: 35-41.

38.Sezer G I, Ramyar K, Karasu B, Goktepe A B, Sezer A (2008) Image analysis of sulfate attack on hardened cement paste. Materials and Design **29**: 224–231.

39.Shaikh S H, Maiti A, Chaki N. (2011) Image binarization using iterative partitioning: A global thresholding approach. In *Proceedings of 2011 International Conference on Recent Trends in Information Systems*. **IEEE**, Kolkata, India, pp: 281-286.

40.Sun W J, Cui Y J (2018) Investigating the microstructure changes for silty soil during drying. Géotechnique **68(4)**: 370-373.

41.Taillon J A, Pellegrinelli C, Huang Y L, et al. (2018) Improving microstructural quantification in FIB/SEM nanotomography. Ultramicroscopy **184**: 24-38.

42.Tang C S, Shi B, Cui YJ, et al (2012) Desiccation cracking behavior of polypropylene fiber–reinforced clayey soil. Canadian Geotechnical Journal **49(9)**: 1088-1101.

43.Terzaghi K (1925) *Erdbaumechanik auf bodenphysikalischer Grundlage*. Franz. Deuticke. Wien.

44.Toprak, E, Olivella, S, Pintado X (2018) Modelling engineered barriers for spent nuclear fuel repository using a double-structure model for pellets. Environmental Geotechnic*s* ***7*(1)**: 72-94.

45.Tournassat C, Steefel C I, Bourg I C, et al (2015) *Natural and engineered clay barriers*. Elsevier, Amsterdam, Netherlands.

46.Trzciński J. (2004). Combined SEM and computerized image analysis of clay soils microstructure: technique & application. In *Advances in geotechnical engineering: The Skempton conference: Proceedings of a three day conference on advances in geotechnical engineering, organised by the Institution of Civil Engineers and held at the Royal Geographical Society*. Thomas Telford, London, UK, pp. 654-666.

47. Wu C, Chu J (2020) Biogrouting method for stronger bond strength for aggregates. Journal of Geotechnical and Geoenvironmental Engineering **146(11)**: 06020021.

48. Yang S, Liu W (2019) Research on Unconstrained Compressive Strength and Microstructure of Calcareous Sand with Curing Agent. Journal of Marine Science and Engineering **7(9)**: 294.

49. Youn H, Tonon F. (2010) Effect of air-drying duration on the engineering properties of four clay-bearing rocks in Texas. Engineering Geology **115(1-2)**: 58-67.

50. Zhou, C., Yu, L., Huang, Z., Liu, Z., Zhang, L. (2021). Analysis of microstructure and spatially dependent permeability of soft soil during consolidation deformation. Soils and Foundations **61**: 708-733.

List of Table Captions

Table 1 Physical properties of Hangzhou clay.

Water content, w (%)	Bulk density, ρ (g/cm^3)	Specific gravity, G_s	Porosity, n	Liquid limit, w_L	Plastic limit, w_P	Particle size distribution		
						0.075~2mm	0.002~0.075mm	<0.002mm
30%	2.01	2.505	0.522	38.0	21.4	25%	56%	19%

Table 2 Mineralogy and chemical composition of Hangzhou clay (Huang et al. 2017).

Mass percentage of main minerals (%)				Mass percentage of main elements (%)			
Quartz	Chlorite	Phlogopite	Anorthite	oxygen(O),	silicon(Si),	aluminum(Al)	others
41.8	34.5	4.4	19.3	46.65	29.85	11.47	12.03

Table 3 Testing program of the first group for comparing different milling directions.

No. of SEM sample	The No. of remolded clay sample that SEM sample obtained from	Void ratio	Water content	Milling direction
1#	Remolded sample I#	0.68	27%	Cross-section
2#	Remolded sample I#	0.68	27%	Vertical-section
3#	Remolded sample II#	0.67	26%	Cross-section
4#	Remolded sample II#	0.67	26%	Vertical-section
5#	Remolded sample III#	0.64	25%	Cross-section
6#	Remolded sample III#	0.64	25%	Vertical-section
7#	Remolded sample IV#	0.78	32%	Cross-section
8#	Remolded sample IV#	0.78	32%	Vertical-section

Table 4 Micro-parameters of different milling directions.

Clay sample No.	SEM sample No.	Milling method	Image porosity n_i	Mean shape factor F	Probability entropy H_m	Fractal dimension D_v	Roundn-ess R_0
I#	1#	Cross-section	0.411	0.293	0.989	1.441	0.421
I#	2#	Vertical-section	0.417	0.284	0.992	1.491	0.410
II#	3#	Cross-section	0.415	0.353	0.992	1.385	0.443
II#	4#	Vertical-section	0.410	0.310	0.993	1.386	0.420
III#	5#	Cross-section	0.374	0.337	0.995	1.439	0.433
III#	6#	Vertical-section	0.368	0.314	0.991	1.491	0.419
IV#	7#	Cross-section	0.428	0.276	0.991	1.529	0.395
IV#	8#	Vertical-section	0.421	0.277	0.994	1.552	0.411

Table 5 Micro-parameters of different drying techniques.

SEM sample No.	Drying techniques	Image porosity n_i	Mean shape factor F	Probability entropy H_m	Fractal dimension D_v	Roundness R_0
9#	Air drying	0.241	0.388	0.995	1.301	0.656
10#	Oven drying	0.138	0.447	0.992	1.201	0.816
11#	Critical point	0.277	0.397	0.986	1.295	0.618
12#	frozen-vacuum	0.284	0.378	0.987	1.324	0.623

Table 6 Micro-parameters of images with different magnifications of 13# SEM sample.

Magnification of image	Image porosity	Mean shape factor	Probability entropy	Fractal dimension	Roundness
500	0.368	0.333	0.982	1.431	0.396
1000	0.377	0.353	0.979	1.362	0.397
1500	0.337	0.393	0.985	1.301	0.416
2000	0.345	0.402	0.983	1.270	0.420
3000	0.299	0.430	0.986	1.242	0.430
5000	0.289	0.445	0.983	1.248	0.410

Table 7 Micro-parameters of 10 SEM images for 13# SEM sample.

Image No.	Image porosity	Mean shape factor	Probability entropy	Fractal dimension	Roundness
1	0.337	0.393	0.985	1.298	0.416
2	0.367	0.402	0.990	1.283	0.429
3	0.376	0.394	0.993	1.266	0.402
4	0.298	0.412	0.995	1.269	0.425
5	0.316	0.409	0.990	1.264	0.395
6	0.325	0.394	0.991	1.281	0.408
7	0.307	0.406	0.995	1.257	0.420
8	0.269	0.444	0.991	1.272	0.433
9	0.304	0.410	0.996	1.267	0.422
10	0.284	0.4278	0.991	1.278	0.426

Table 8 Error ε of each micro-parameter from 10 SEM images for 13# SEM sample.

Adjacent two quantity	Image porosity	Mean shape factor	Probability entropy	Fractal dimension	Roundness
2 and 3	0.730	0.789	0.533	0.582	0.596
3 and 4	0.093	0.099	0.318	0.410	0.436
4 and 5	0.270	0.253	0.322	0.236	0.054
5 and 6	**0.231**	**0.179**	**0.244**	**0.236**	**0.234**
6 and 7	0.133	0.169	0.116	0.049	0.170
7 and 8	0.073	**0.490**	0.163	0.162	0.024
8 and 9	0.122	0.138	0.034	0.125	0.129
9 and10	0.063	0.041	0.114	0.114	0.096

Table 9 Micro-parameters of 10 SEM images for 14# SEM sample.

Image No.	Image porosity	Mean shape factor	Probability entropy	Fractal dimension	Roundness
1	0.362	0.392	0.990	1.281	0.415
2	0.434	0.394	0.994	1.288	0.411
3	0.360	0.392	0.992	1.263	0.422
4	0.446	0.375	0.986	1.283	0.404
5	0.326	0.405	0.989	1.287	0.447
6	0.364	0.396	0.992	1.271	0.397
7	0.440	0.378	0.995	1.304	0.411
8	0.368	0.378	0.996	1.291	0.409
9	0.362	0.395	0.992	1.273	0.414
10	0.291	0.407	0.996	1.261	0.411

Table 10 Error $\boldsymbol{\varepsilon}$ of each micro-parameter from 10 SEM images for 14# SEM sample.

Adjacent two quantity	Image porosity	Mean shape factor	Probability entropy	Fractal dimension	Roundness
2 and 3	0.755	0.787	0.804	0.260	0.463
3 and 4	0.302	0.770	0.139	0.458	0.139
4 and 5	0.116	0.057	0.311	0.277	0.436
5 and 6	**0.230**	**0.230**	**0.208**	**0.171**	**0.102**
6 and 7	0.098	0.052	0.034	0.140	0.190
7 and 8	0.150	0.088	0.063	0.138	0.153
8 and 9	0.123	0.120	0.140	0.101	0.140
9 and10	0.077	0.020	0.051	0.016	0.120

Appendix

Appendix 1

A.1.1 Definition of interval estimation

Let $F = \{f(x, \theta), \theta \in \Theta\}$ be a random distribution, where Θ is the parameter space for θ. $X = (X_1, \cdots X_n)$ is a random sample space from the distribution F. $g(\theta)$ is a real-valued function of θ. $\hat{g}_1(x)$ and $\hat{g}_2(x)$ are two statistics that are defined in the sample space X, and their values are in the Θ. Then the random interval $[\hat{g}_1(x), \hat{g}_2(x)]$ is an interval estimation of $g(\theta)$.

A.1.2 Confidence interval

$P_\theta\big(\hat{g}_1(x) \le g(\theta) \le \hat{g}_1(x)\big)$ is the probability that the value of $g(\theta)$ is in the random interval $[\hat{g}_1(x), \hat{g}_2(x)]$. The random interval $[\hat{g}_1(x), \hat{g}_2(x)]$ is supposed to be an interval estimation of $g(\theta)$, if the value of α $(0 < \alpha < 1)$ is given, and

$$P_\theta\big(\hat{g}_1(x) \le g(\theta) \le \hat{g}_1(x)\big) \ge 1 - \alpha, \ g(\theta) \in \Theta \tag{2}$$

then $[\hat{g}_1(x), \hat{g}_2(x)]$ is the confidence interval of $g(\theta)$, and the confidence level of $g(\theta)$ is $1 - \alpha$. For example, if $\alpha = 0.05$, then $1 - \alpha = 0.95$, then the probability of $g(\theta) \in [\hat{g}_1(x), \hat{g}_2(x)]$ is 0.95, while the probability of $g(\theta) \notin [\hat{g}_1(x), \hat{g}_2(x)]$ is 0.05.

A.1.3 Interval estimation using t-distribution

Let's assume that the parametric distribution $F = \{f(x,\theta), \theta \in \Theta\}$ obeys normal distribution $N(\mu, \sigma^2)$, and $X = (X_1, \cdots X_n)$ is a sample drawn from $N(\mu, \sigma^2)$. μ and σ^2 are the unknown mean and unknown standard deviation of $N(\mu, \sigma^2)$, respectively. $\bar{X}$ and S^2 are defined as the mean and variance of the sample $X = (X_1, \cdots X_n)$, respectively.

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{3}$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \tag{4}$$

If σ^2 is replaced by S^2, then the statistic is known as t distribution:

$$t = \frac{\bar{x}-\mu}{s/\sqrt{n}} \tag{5}$$

Defining

$$T_{n-1} = \frac{\sqrt{n}(\bar{X}-\mu)}{S} \tag{6}$$

T follows the distribution of t_{n-1} , and the distribution of t is symmetrical about the origin point. For any $\alpha \in (0,1)$, the upper quantile $t_{n-1}\frac{\alpha}{2}$ satisfies

$$P_{\mu}(|T| < \varepsilon) = 1 - \alpha \tag{7}$$

where,

$$\varepsilon = \frac{S}{\sqrt{n}} t_{n-1}\frac{\alpha}{2} \tag{8}$$

then the confidence space of μ with confidence coefficient α is

$$\left(\bar{X} - \frac{S}{\sqrt{n}} t_{n-1}\frac{\alpha}{2},\ \bar{X} + \frac{S}{\sqrt{n}} t_{n-1}\frac{\alpha}{2}\right) \tag{9}$$

Appendix 2

A.2.1. Image porosity

In the SEM image, the whole area can be divided into two parts: clay particles and pores between them. The image porosity n_i can be calculated by:

$$n_i = \frac{A_1}{A_0} \tag{10}$$

where A_1 is the pore area; A_0 is the area of total SEM image. The value range of n is (0,1).

A.2.2 Mean shape factor

The shape factor F_i of a single particle describes the particle shape. F_i defines the degree to which the particle is similar to a circle, and can be calculated by the following method (Sezer et al., 2008; Liu, et al., 2011):

$$F_i = \frac{P}{L} \tag{11}$$

where P is the perimeter of a circle that has the same area as the clay particle; L is the actual perimeter of the clay particle. The mean shape factor F is the mean value of all the soil particles' shape factors in the SEM image. F can be calculated by:

$$F = \frac{\sum_{i=1}^{n} F_i}{m} \tag{12}$$

where m is the total number of clay particles in the SEM image. The value range of F is (0, 1). If the particle shape is close to a circle, then F is close to 1.

A.2.3 Fractal dimension

Fractal dimension D_v reflects the complexity of clay particles' edges. D_v can be calculated by the following equation (Dathe et al., 2001):

$$D_v = \lim_{r \to 0} \frac{\ln N(r)}{\ln r} \tag{13}$$

where r is the side length of a square box; $N(r)$ is the number of square boxes. And the calculation of the mean fractal dimension is similar to that of the mean shape factor.

A.2.4 Roundness

Roundness R_0 describes the smoothness or roughness of the perimeter along the particle edge. R_0 can be calculated by the following equation (Cox and Budhu, 2008).

$$R_0 = \frac{A_p}{A'} \tag{14}$$

where A_p is the real area of the particle; A' is the circumcircle area of clay particle, $A' = Q^2/(4\pi)$, where Q is the perimeter of particle circumcircle. And the calculation of the mean roundness is similar to that of the mean shape factor.

A.2.5 Orientation probability entropy

The concept of probabilistic entropy is introduced into the microstructure analysis of clay to define the orientation probability entropy H_m, which is used to represent the orderliness of clay particle arrangement. The definition of H_m is as follows (Liu, et al., 2011):

$$H_m = -\sum_{i=1}^{k} P_i(\alpha_p) log_k P_i(\alpha_p) \quad (15)$$

where α_p is the direction of a soil particle. α_p can be computed directly in PCAS software, and α_p varies between [0°-180°]; k is the number of equally divided areas in the whole particle direction range [0°, 180°], and k can be selected by the user. In this paper, k is set to be 18 (a random number). Thus, each division corresponds to 10°. $P_i(\alpha_p)$ is the percentage of soil particles whose directions α_p belong to a specific range, for example, when $i = 1$, $P_1(\alpha_p)$ means the percentage of soil particles whose directions α_p belongs to the direction range of [0°, 10°].

The value interval of H_m is [0,1]. When $H_m = 0$, all the soil particles are parallel to each other. When $H_m = 1$, all the soil particles are in random directions. With the increase in the value of H_m, the directions of soil particles are more random.

Appendix 3

A.3.1 Remolded sample preparation

The undisturbed clay was sliced into small blocks, and then these clay blocks were oven-dried for 8 h at 105℃, pulverized to pass a 2 mm sieve. After that, the clay powder was mixed with water to form remolded clay, then kept in natural condition for 24 h. The water contents of remolded clay are listed in section 4.3. Later, the remolded clay was transferred to a miniature mould apparatus, which is a split cylindrical mould with 39.1 mm diameter and 80 mm height. The remolded clay was compacted in 3 layers. Each layer was tamped and compacted with the compaction rod. The height and number of blows needed depend on the desired void ratio. According to the wanted void ratio for each remolded clay sample, the weight of the individual clay layers was calculated. Then, the remolded clay sample was slid out from the cylinder. All the experimental operations above meet the requirements of GB/T 50123-1999. The remolded sample is a cylindrical specimen with 39.1 mm diam and 80 mm height.

A.3.2 Sample preparation for SEM imaging

The remolded clay sample above needed to be milled, dried and gold coated before SEM observation. The methods to obtain the cross-section surface and vertical-section surface are shown in section 3.1. For the drying method, the SEM samples were dried with drying methods as described in section 3.2. Then before SEM imaging of the samples, to make clay conductive, the surface of the SEM sample was coated with an additional thin layer of about 20 nm of gold by SBC-12 ion sputterer from KYKY Technology Co., LTD.

A.3.3 SEM imaging

The Scanning Electron Microscope (SEM) of FEG650 type produced by FEI company in the Netherlands was used for SEM observation. The goal was to obtain high-quality images. The accelerating voltage was 5 kV, the working distance was 10 mm, and the secondary electron imaging mode was operated. All the tests were conducted by the same experimenter, which can reduce subjectivity. For example, all the images can keep similar brightness and contrast. Finally, SEM images with suitable magnification and quantity were obtained.

A.3.4 SEM image processing

After SEM imaging, the images should be processed to calculate the micro-parameters. At first, the threshold value for each image is determined by A-K method, which is introduced in section 3.2. Then the threshold value can be used in the program of Pores (Particle) and Cracks Analysis System (PCAS). Subsequently, the micro-parameters of image porosity n_i, probability entropy H_m, fractal dimension D_v, mean shape factor F, roundness R_0 can all be obtained from PCAS software.

List of Figure Captions

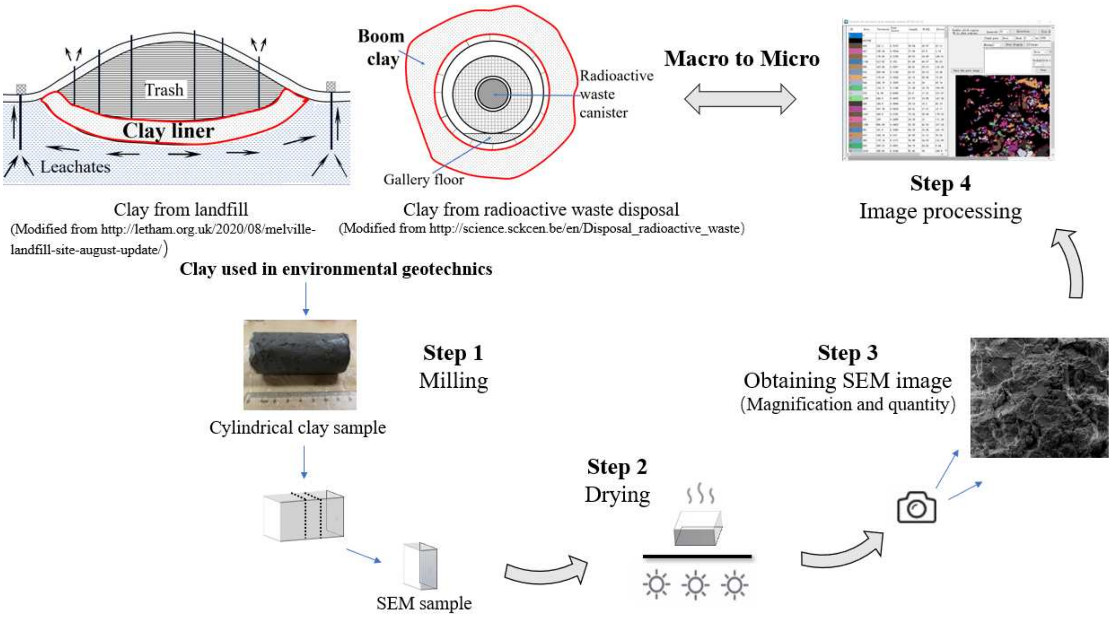


Figure 1 Main difficulties in the quantification procedure for characterization of clay microstructure using SEM.

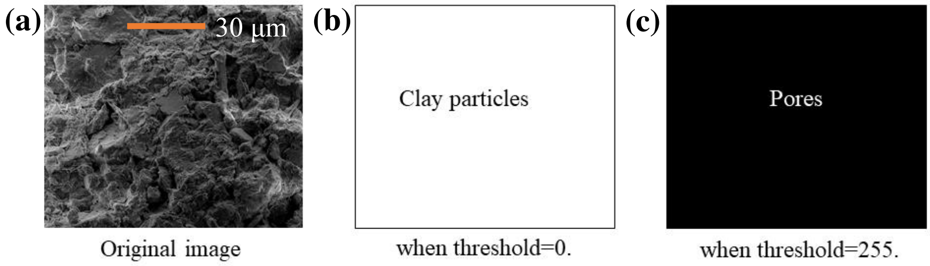


Figure 2 The range of threshold value in SEM image. (a) Original image. (b) The binary image when threshold value is 0. (c) The binary image when threshold value is 255.

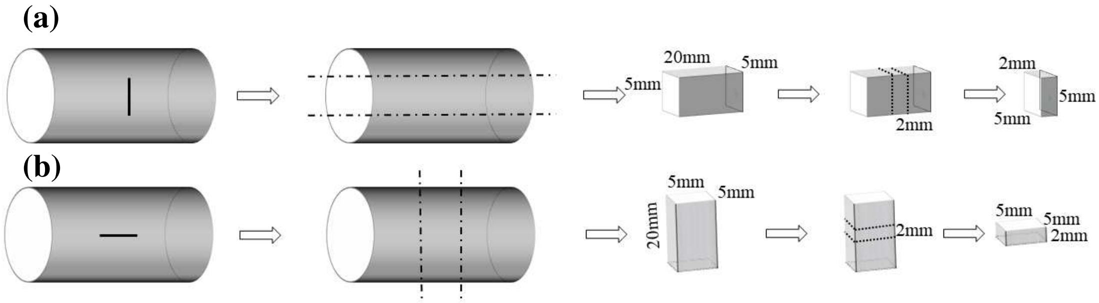


Figure 3 Milling methods. (a) The milling method of cross-section. (b) The milling method of vertical-section.

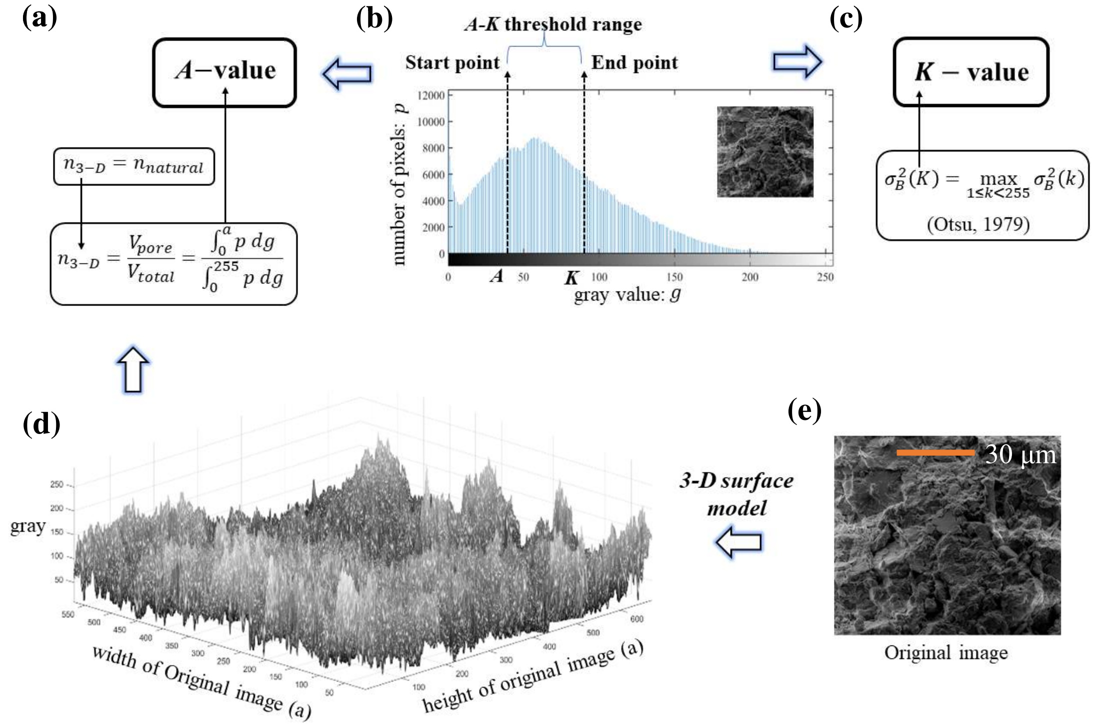


Figure 4 The *A*-*K* threshold determining method in image processing. (a-c) The prediction of threshold range. (d-e) The 3-D surface model.

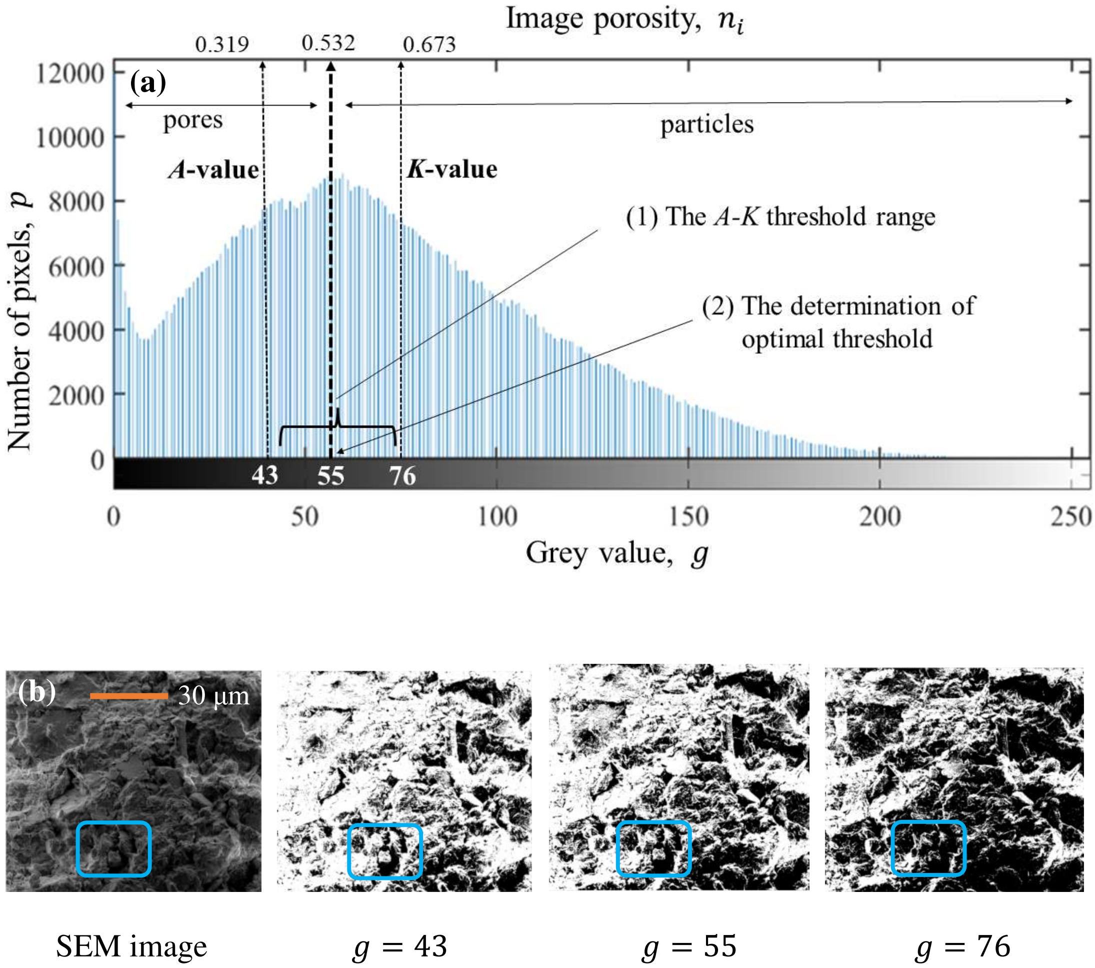


Figure 5 The determination of optimal threshold value in *A*-*K* method. (a) The *A*-*K* threshold range for original SEM image. (b) The optimal threshold value was determined by artificial comparison of binary images with threshold in the *A*-*K* range of [43, 76].

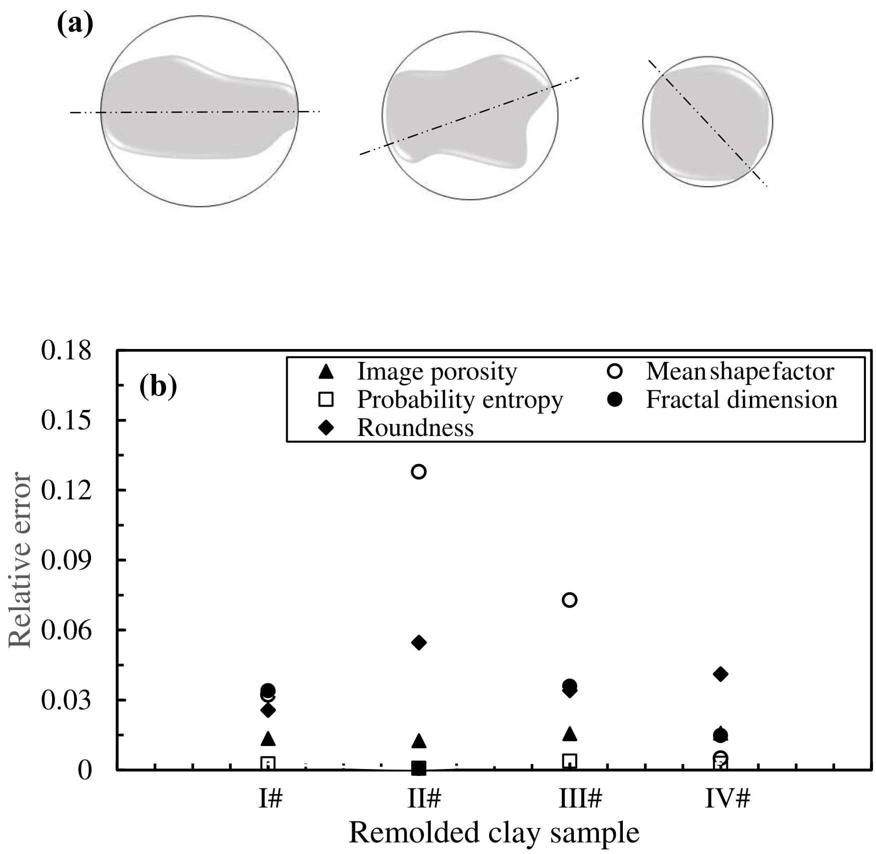


Figure 6 Clay particles and quantitative analysis of different milling directions. (a) Three clay particles with different features. (b) The relative error in the mico-parameters between the cross-section and the vertical-section of 4 remolded clay samples.

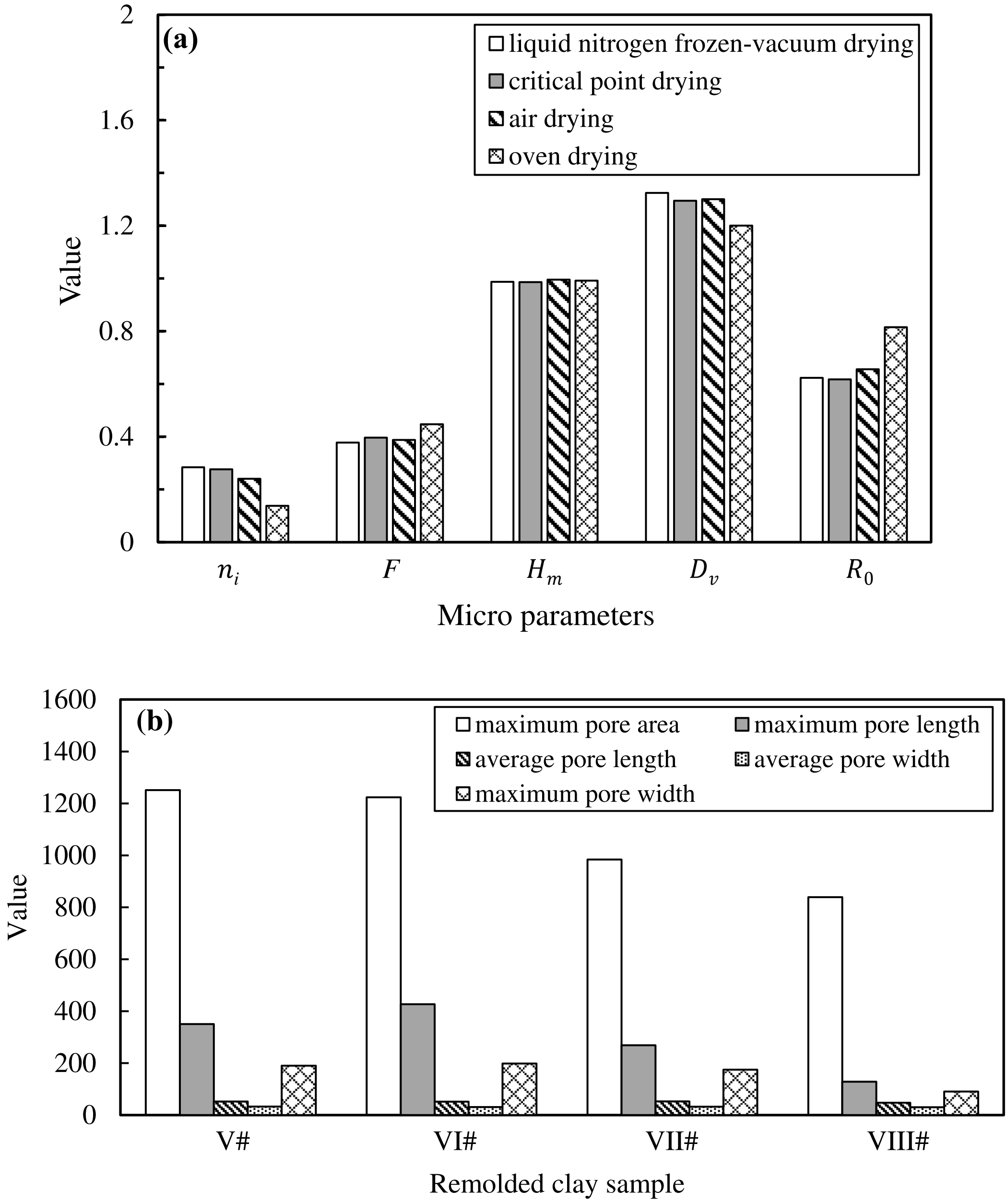


Figure 7 Quantitative analysis of different drying techniques. (a) Comparison of micro parameters for different drying techniques. (b) Comparison of pore sizes for different drying techniques

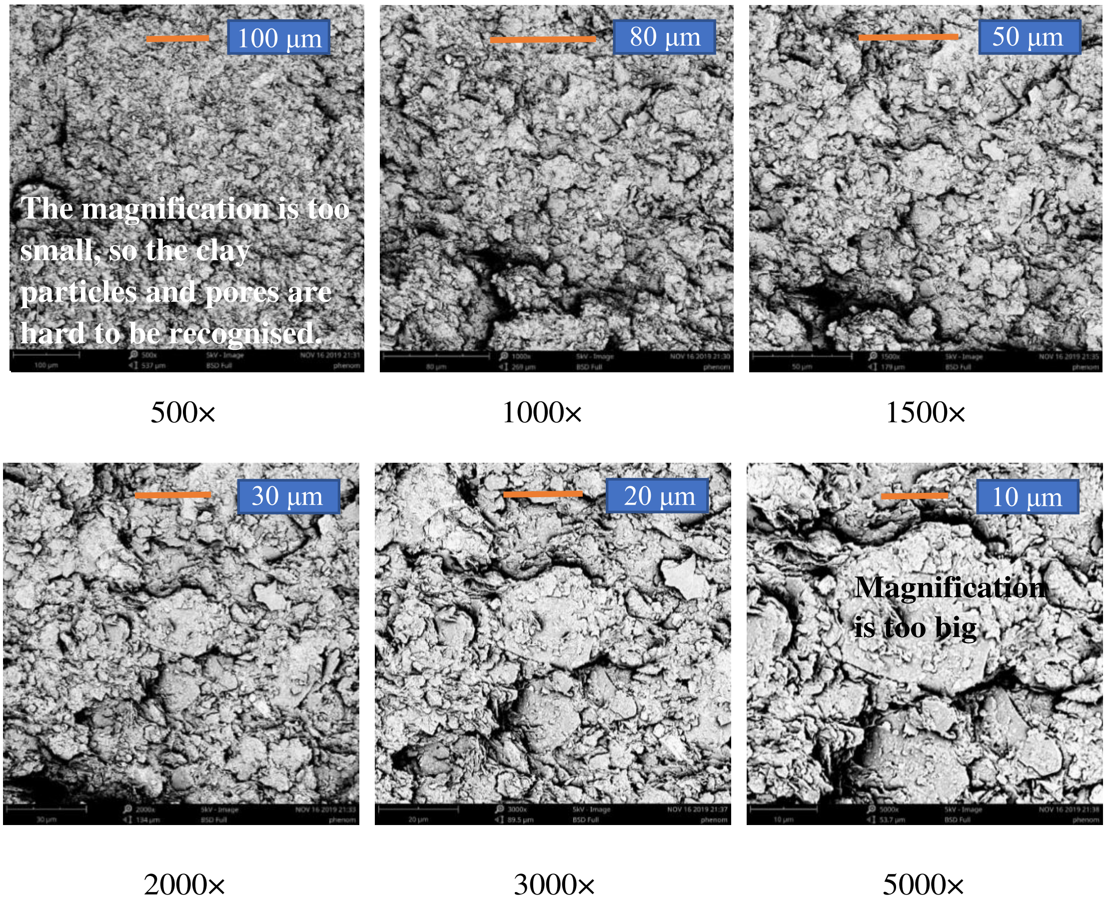


Figure 8 The comparison of magnifications between 500× and 5000× for this sample.

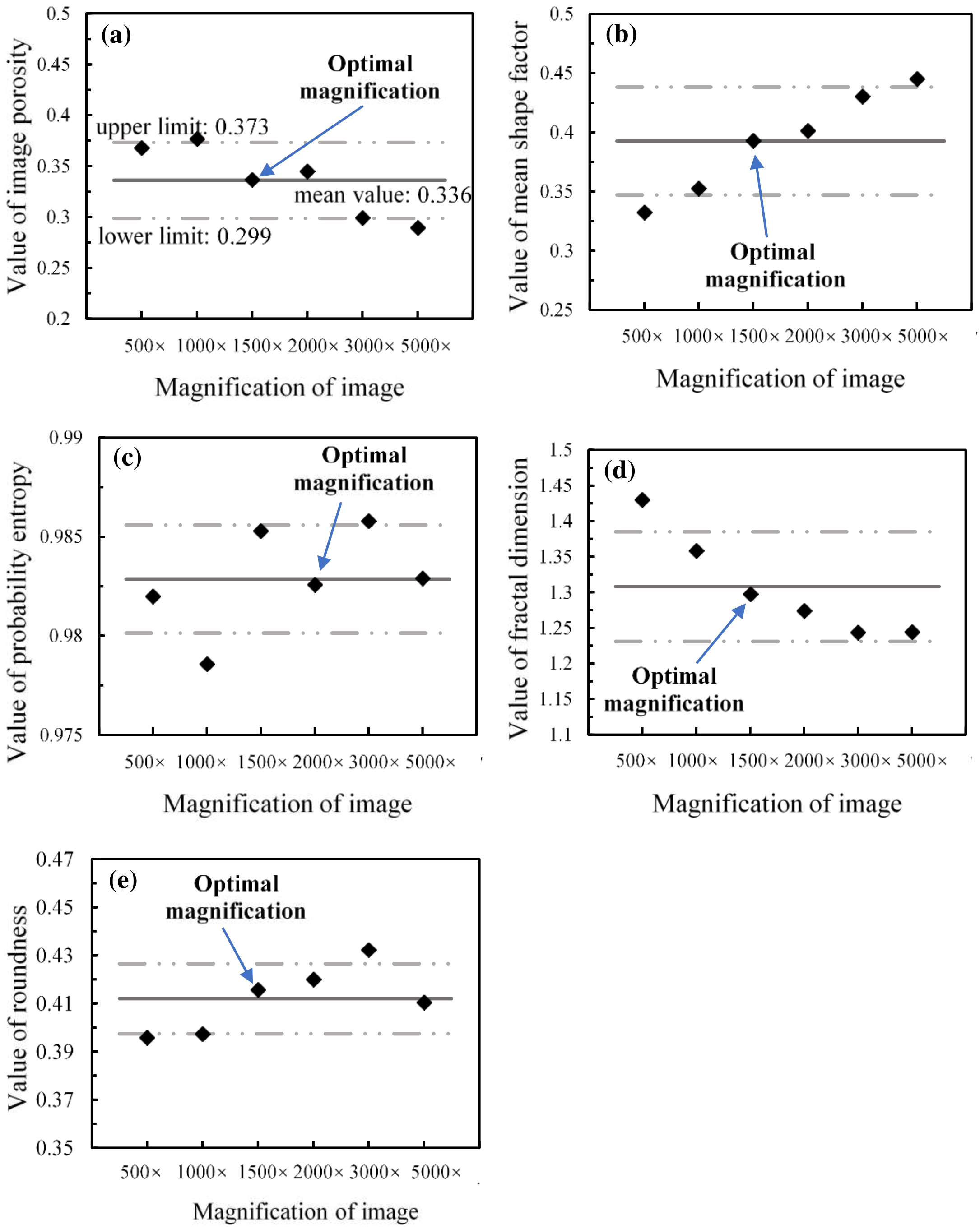


Figure 9 Quantitative analysis for optimal magnification**.** (a) Image porosity. (b) Mean shape factor. (c) Probability entropy. (d) Fractal dimension. (e) Roundness.

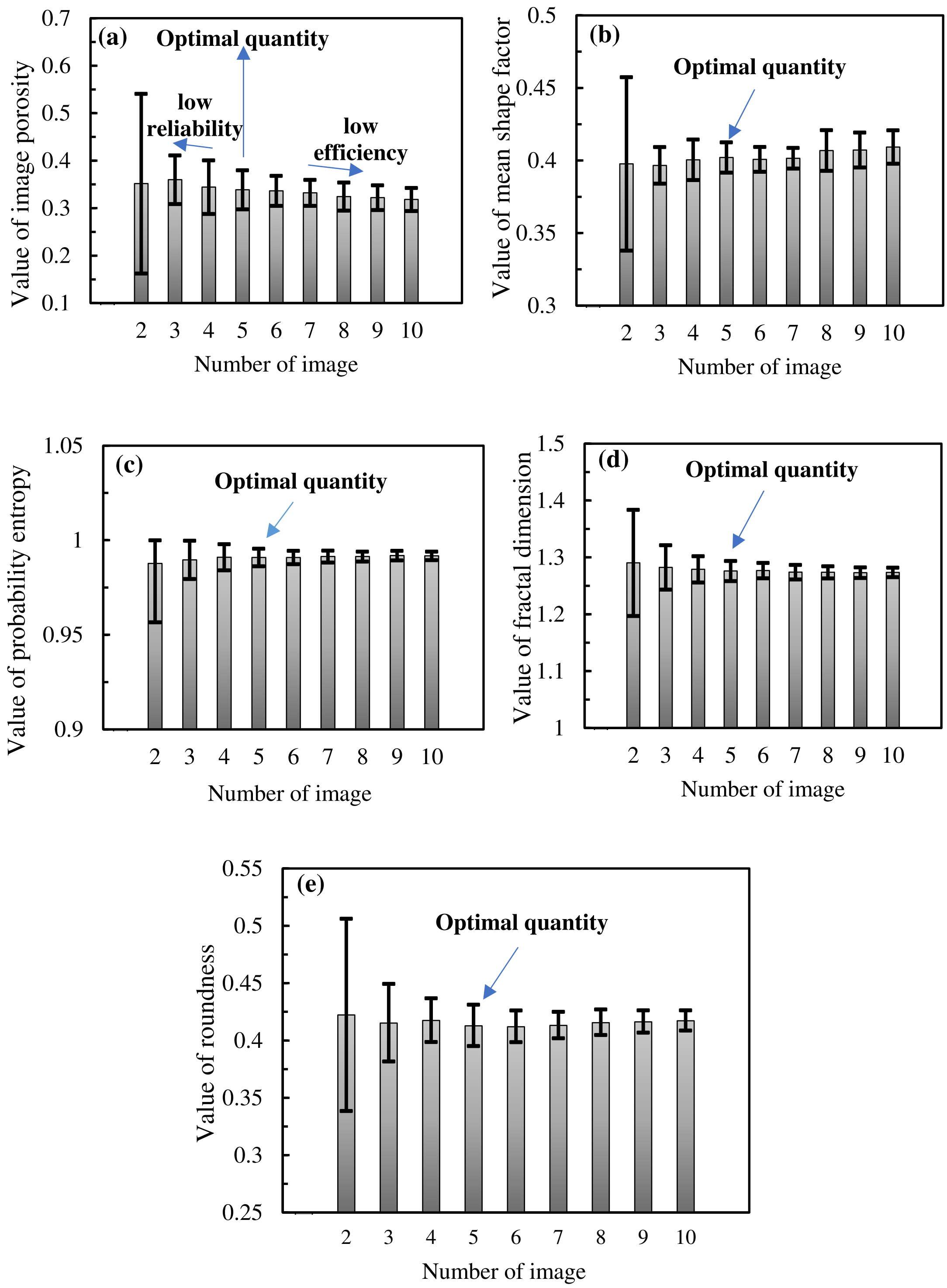


Figure 10 Analysis of optimal image quantity for SEM sample 13#. The height of each bar is the mean value of a certain image quantity, while the short vertical line on top represents the corresponding confident interval. (a) Statistic analysis of image porosity. (b) Mean shape factor. (c) Probability entropy. (d) Fractal dimension. (e) Roundness.

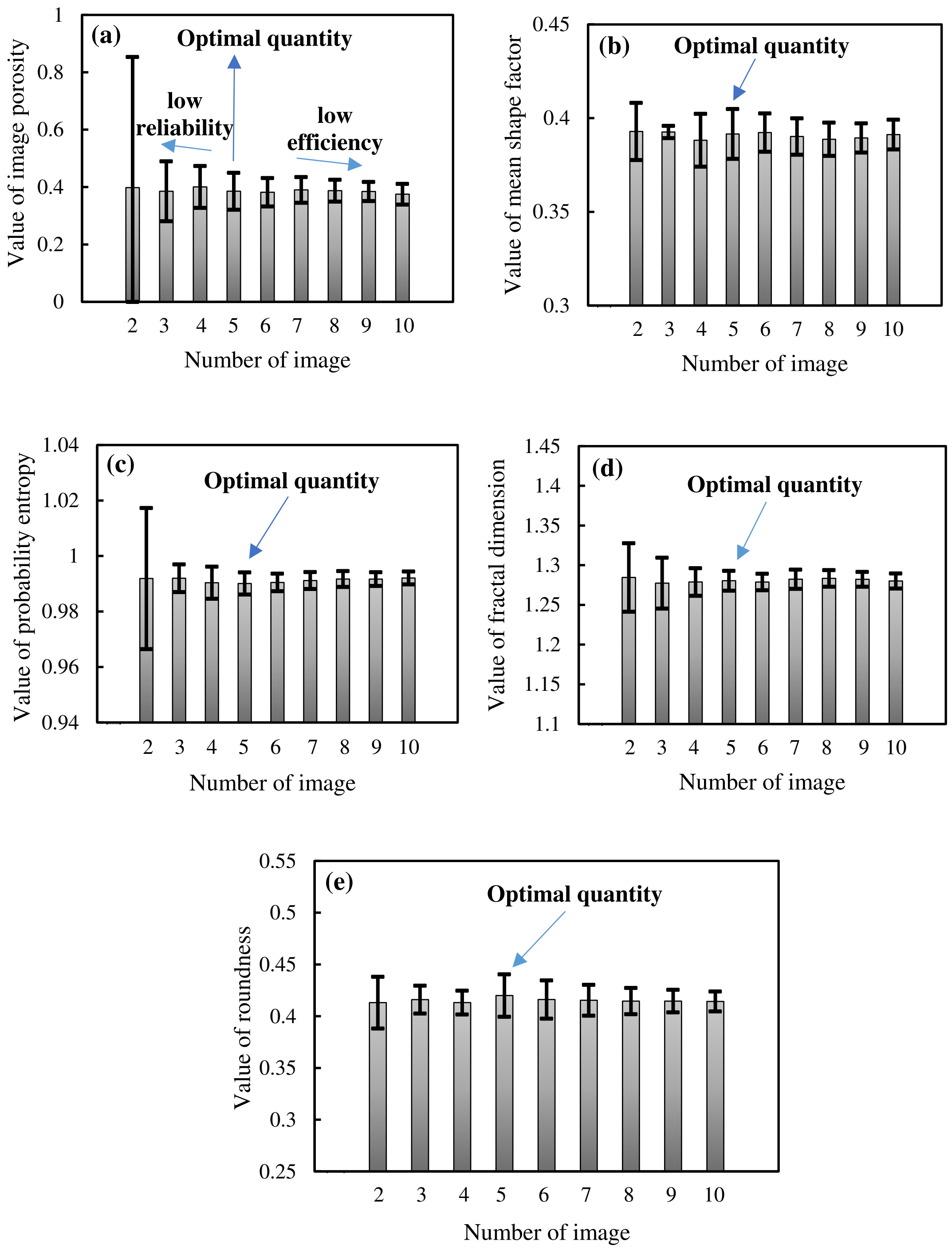


Figure 11 Analysis of optimal image quantity for SEM sample 14#. The height of each bar is the mean value of a certain image quantity, while the short vertical line on the top represents the corresponding confident interval. (a) Statistic analysis of image porosity. (b) Mean shape factor. (c) Probability entropy. (d) Fractal dimension. (e) Roundness.

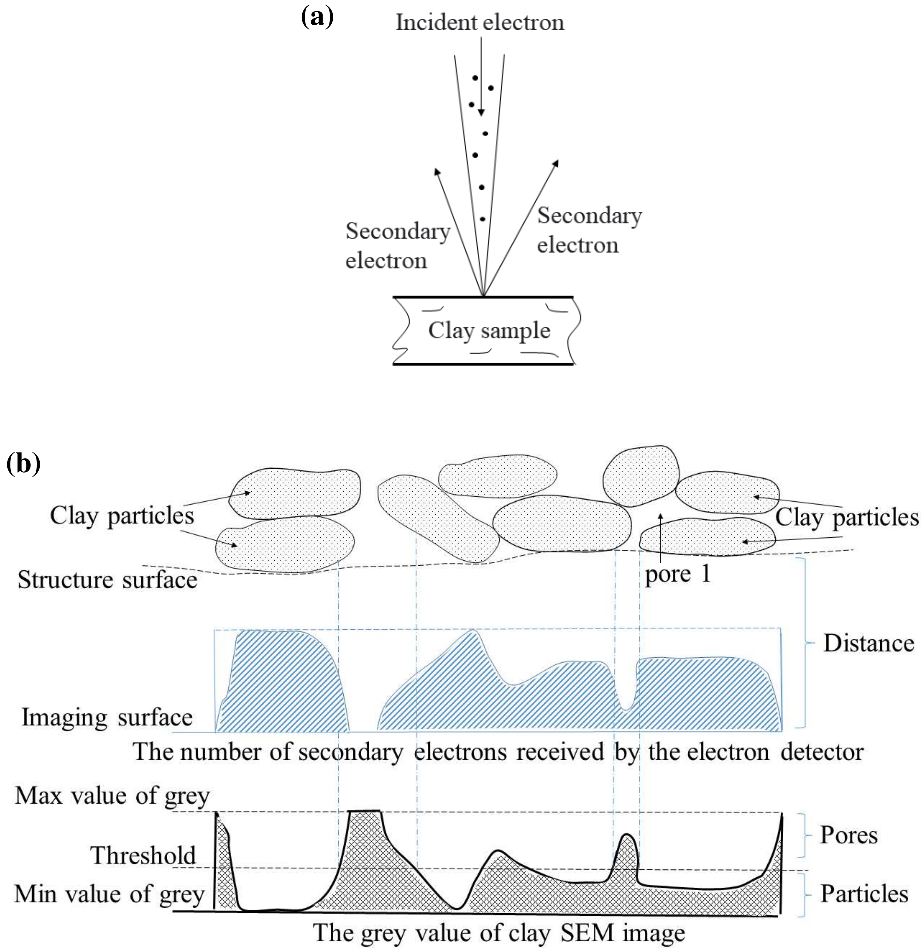


Figure 12 Discussion on 3-D surface model. (a) The imaging principle of SEM. (b) The spatial features of the structure surface of the clay particles can be reflected by the grey value.

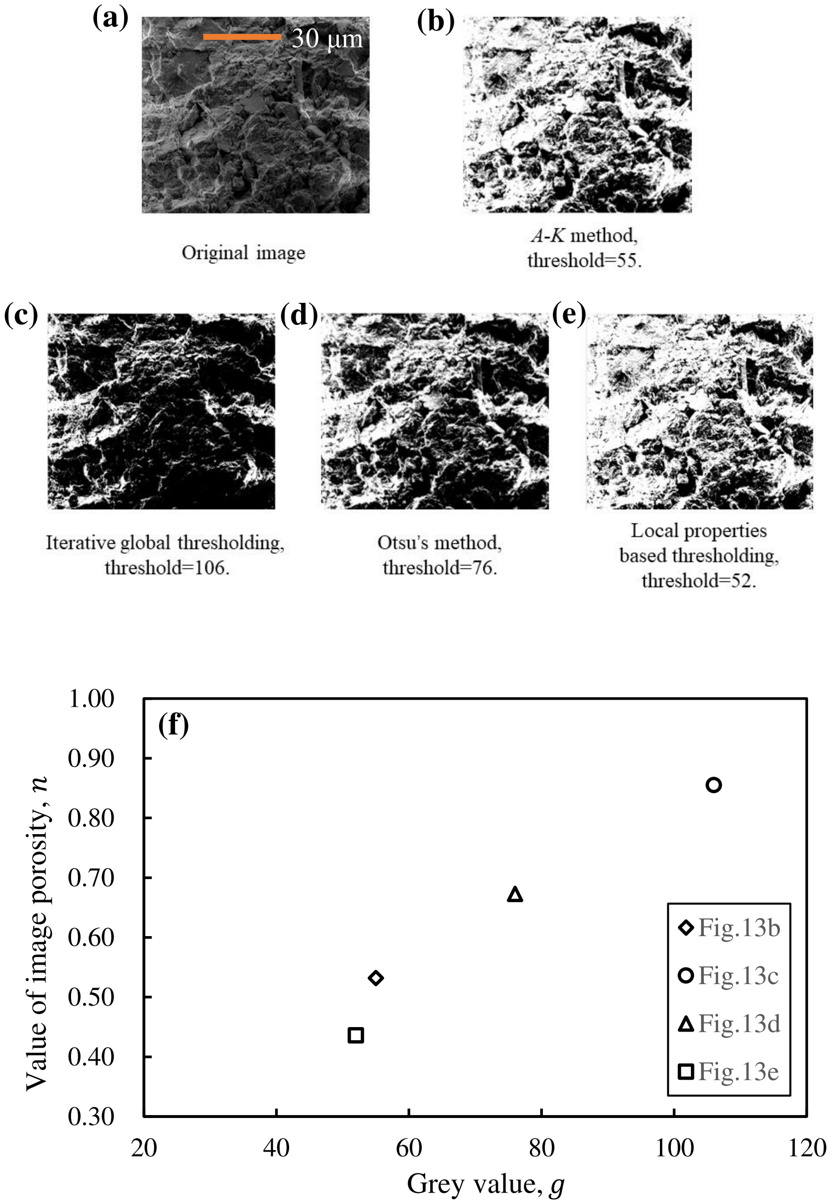


Figure 13 Quantitative comparison of different threshold value determining methods in image processing. (a) Original image. (b-e) The binary image after threshold segmentation with different methods. (f) Threshold and corresponding image porosity.