



UNIVERSITY OF LEEDS

This is a repository copy of *An improved quantification method for characterisation of clay microstructure using SEM*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/183125/>

Version: Accepted Version

---

**Article:**

Xu, L, Xu, R, Shashank, BS et al. (3 more authors) (2022) An improved quantification method for characterisation of clay microstructure using SEM. *Environmental Geotechnics*. ISSN 2051-803X

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

---

This article is protected by copyright. This is an author produced version of an article published in *Environmental Geotechnics*. Uploaded in accordance with the publisher's self-archiving policy.

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>



# Environmental Geotechnics

---

## An Improved Quantification Method for Characterization of Clay Microstructure 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

PDF auto-generated using **ReView**

from



RIVER VALLEY  
TECHNOLOGIES

1 **Information of the authors**

2 **Corresponding author:** Riqing Xu\*

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

6 Email: xurq@zju.edu.cn

7 **Corresponding author:** Xiaohui Chen\*

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

9 Email: x.chen@leeds.ac.uk

10 **First author:** Liyang Xu

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

14 Email: 11712037@zju.edu.cn

15 **Co-author:** B. S. Shashank

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

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

19 **Co-author:** Jianlin Yu

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

23 Email: yujianlin72@126.com

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

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

27 Email: ark\_wu@zju.edu.cn

28

29                   **An Improved Quantification Method for Characterization of Clay**  
30   **Microstructure Using SEM**

31       Liyang Xu<sup>1,2</sup>, Riqing Xu<sup>\*1,2</sup>, B. S. Shashank<sup>3</sup>, Xiaohui Chen<sup>\*4</sup>, Jianlin Yu<sup>1,2</sup>, Chuangzhou Wu<sup>5</sup>

32           <sup>1</sup> *Research Center of Coastal and Urban Geotechnical Engineering, Zhejiang*  
33   *University, China*

34           <sup>2</sup> *Engineering research center of urban underground space development of Zhejiang*  
35   *Province, China.*

36           <sup>3</sup> *Dept. of Civil Engineering, Birla Institute of Technology and Science (BITS) Pilani,*  
37   *India*

38   <sup>4</sup> *School of Civil Engineering, University of Leeds, Leeds, UK*

39           <sup>5</sup> *Institute of Port, Coastal, and Offshore Engineering, Ocean College, Zhejiang*  
40   *University, Zhoushan, China*

41

42   **Abstract:** Clay is a material widely used in environmental geotechnics. Quantification  
43   of clay microstructure properties using SEM images is intuitively the simplest way, but  
44   the challenges remain in extracting more credible quantified micro-parameters that can  
45   represent the clay holistic macroscopic properties from SEM images and with higher  
46   efficiency. In this paper, an improved quantification method was proposed to deal with  
47   this issue. The influence of different milling and drying methods on quantification  
48   results was analyzed to improve reliability and efficiency. The A-K threshold  
49   determination method (A-K method) was proposed to calculate the optimal threshold in  
50   image processing. Also, the interval estimation method was introduced to obtain the  
51   optimal magnification and number of images. Then, this quantification method was

52 applied to Hangzhou clay. Results show that anisotropy in clay microstructure should  
53 be considered when choosing the observed surface. And there is no significant  
54 difference in micro-parameters between liquid nitrogen frozen-vacuum drying and  
55 critical point drying. A comparison among the *A-K* method, artificial method, and  
56 theoretical methods that can determine threshold shows that the proposed *A-K* method  
57 combines the advantages of all methods. The optimal magnification and number of  
58 images can be determined by mathematical theory, which also improves reliability and  
59 efficiency.

60

61 **Keywords:**

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

63

64 **List of notations**

65  $A_p$  real area of the particle  
66  $A'$  circumcircle area of clay particle  
67  $A_1$  pore area  
68  $A_0$  area of total SEM image  
69  $D_v$  fractal dimension  
70  $F$  mean shape factor  
71  $F_i$  shape factor for each clay particle  
72  $G_s$  specific gravity

73	$H_m$	orientation probability entropy
74	$k$	number of equally divided areas in the whole particle direction range
75	$L$	actual perimeter of the clay particle
76	$m$	total number of clay particles in the SEM image
77	$n_i$	image porosity
78	$n$	the porosity of clay
79	$N(r)$	number of square boxes
80	$N(\mu, \sigma^2)$	normal distribution
81	$P$	perimeter of a circle that has the same area as the clay particle
82	$P_i(\alpha)$	percentage of soil particles whose directions $\alpha$ belong to a specific range
83	$r$	side length of a square box
84	$R_0$	roundness
85	$S^2$	variance of the sample $X = (X_1, \dots, X_n)$
86	$t_{n-1} \frac{\alpha}{2}$	upper quantile
87	$n_{3-D}$	3-D surface porosity
88	$V_{pore}$	volume of pore
89	$\bar{X}$	mean value of the sample $X = (X_1, \dots, X_n)$
90	$\alpha$	confidence coefficient
91	$\alpha_p$	orientation of soil particle
92	$\gamma$	unit weight
93	$\mu$	unknown mean value of $N(\mu, \sigma^2)$

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

95  $w$  water content

96  $w_L$  liquid limit

97  $w_P$  plastic limit

98

## 99 **1. Introduction**

100 Clay is a common material widely used in geoenvironmental engineering (Alba et  
101 al., 2009; Dohrmann et al., 2013; Tournassat et al., 2015), ranging from landfill clay  
102 liner to the barrier in the nuclear waste disposal (Chen et al., 2012; Estabragh et al.,  
103 2018; Toprak et al., 2018) (Fig.1). In engineering practice, it's crucial to understand the  
104 physical and mechanical properties of these clay barriers, such as the strength,  
105 mechanical stability, and hydraulic conductivity. The macro-behaviour of clays  
106 depends chiefly on their microstructure. For example, the shear strength of clay can be  
107 improved by increasing inter-cluster cementation bonding and reducing the pore space  
108 (Horpibulsuk et al. 2010). And microstructural changes in the clays may cause  
109 settlement and bearing capacity problems (Oztoprak and Pisirici, 2011).

110

111 Therefore, it's significant to study the micro-mechanisms of clay behaviour by  
112 effectively quantifying clay microstructure parameters such as image porosity, fractal  
113 dimension, mean shape factor, et al. (Pusch and Weston, 2003; Bennett and Hulbert M,  
114 2012).

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

124

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

134



135 In terms of the magnification, too low magnification will make it hard to distinguish  
136 between pores and particles in the SEM image, while too high magnification makes it  
137 difficult to recognise the micro-structural features since the whole image is occupied  
138 only by a few particles or pores (Lin et al. 2018; Bangaru et al. 2019; Wu and Chu,  
139 2020). As for the number, since SEM images can only show the features of clay in the  
140 order of microns, a large number of SEM images are required to improve the  
141 representativeness of the quantification results, but an increased number of SEM  
142 images also costs longer time and higher fee. There are many tools and approaches for  
143 image processing, such as MATLAB®, Image-Pro Plus software, and Pores / Particles  
144 and Cracks Analysis System (PCAS), to extract micro-parameters from the SEM  
145 images (Cox and Budhu, 2008; Prakongkep et al., 2010; Liu et al., 2011; Taillon et al.,  
146 2018). The key for these image processing methods is the thresholding used to identify  
147 the particles and the pores in the binary image. However, it's still a big challenge to  
148 determine the optimal threshold value for SEM images of clay, and there is significant  
149 scope for advancements and refinements of the used techniques (Taillon et al., 2018).

150

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

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

158

## 159 **2. Factors influencing quantification of clay microscopic characterization**

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

164

### 165 **2.1 Milling directions**

166 The milling direction may affect the clay microstructure characterization since clay  
167 is an anisotropic material at the macro scale. The clay sample considered in this study  
168 for SEM analysis was a small cuboid whose length, width, and depth were about 5 mm,  
169 5 mm, and 2 mm, respectively. This SEM sample is milled from a remoulded sample,  
170 which is a cylindrical specimen with 39.1 mm diam and 80 mm height. The method to  
171 mill is a big issue.

172

### 173 **2.2 Drying techniques**

174 The clay sample should be dry when testing in the vacuum environment of SEM.  
175 But if the clay is dried, the pores between clay particles would become smaller (Burton  
176 et al., 2015; Sun and Cui, 2018), then the distance between particles would become

177 smaller, which could change the interparticle forces such as van der Waals and capillary  
178 interactions, and hence the arrangement of the particles. Therefore, the clay  
179 microstructure would be modified, resulting in parameters corresponding to an entirely  
180 different microstructure. The method to dry is also a big issue.

181

### 182 ***2.3 Magnification and quantity of SEM image***

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

189

### 190 ***2.4 Image processing***

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

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

209

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

217

218 **3. Methodology for quantitative analysis of microstructures in SEM**

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

221

### 222   **3.1 The method of milling**

223       The key issue in milling operation is to obtain an undisturbed surface. To study the  
224       influence of milling directions on the quantification result, samples from cross-section  
225       and vertical-section of Hangzhou clay were compared. The cross-section is parallel to  
226       the bottom base of the cylindrical remoulded clay sample, while the vertical-section is  
227       perpendicular to the bottom base. The methods to obtain the cross-section surface and  
228       vertical-section surface are shown in Fig. 3a and Fig. 3b, respectively. At first, the  
229       remoulded clay sample was broken into a strip about 20 mm (length) × 5 mm (width)  
230       × 5 mm (height) from the center. Then two parallel grooves of 0.5mm depth were  
231       carved in the middle of this strip. The distance between the grooves was 2 mm. After  
232       that, the sample was carefully broken by hand along the groove to get the undisturbed  
233       surface for SEM observation (SEM sample).

234

### 235   **3.2 The method of drying**

236       To study the influence of drying techniques on the quantification result, samples  
237       dried by four popular methods, including liquid nitrogen frozen-vacuum drying, critical  
238       point drying, air drying, and oven drying (Delage and Lefebvre G,1984; Dey et al.,

239 1989; Janecek and Robert, 2016; Lindroth et al., 1988). The steps are detailed in the  
240 following lines:

241

242 In the liquid nitrogen frozen-vacuum drying method (Trzcinski, 2004), SEM  
243 Sample was immersed in liquid nitrogen for 1 min before being placed in the vacuum  
244 dryer for 20 h (equipment model: ALPHA1-4). As for the critical point drying method  
245 (Lawrence, 1979), the sample was dehydrated in a graded ethanol series (5%, 15%,  
246 50%, 75%, 100%) at an interval of 20 min each, and subsequently dried in a critical  
247 point drying apparatus for 48 h (equipment model: Quorum-Emitech K850), which is  
248 covered with liquid carbon dioxide. For the air drying method (Youn and Tonon, 2010),  
249 the SEM sample was kept in the natural environment for 48 h. In the oven drying  
250 method, the SEM sample was dried in the oven directly for 12 h at 105 °C (Korpa,  
251 2006).

252

### 253 **3.3 Image processing for clay microscopic quantification**

#### 254 **3.3.1 A-K threshold determining method**

255 A novel A-K method for determining threshold value has been developed. This  
256 method combines the advantages of the theoretical approach and artificial approach and  
257 thus can improve the efficiency and reliability of image processing. The method  
258 consists of two steps: the first step is the prediction of the threshold range, and the  
259 second step is the determination of an optimal threshold value. It should be noted that

260 the threshold range varies from the start-point to the end-point. For example, a threshold  
 261 range is [20, 100], which means that 20 is the start-point, and 100 is the end-point.

262

263 (1) The prediction of the threshold range

264 This step is aimed at finding the start-point and the end-point within which the  
 265 threshold value lies. For the convenience of description, the start-point is assumed to  
 266 be the  $A$ -value, and the end-point is assumed to be the  $K$ -value. The  $A$ -value is  
 267 confirmed on the principle that image porosity obtained from image processing is close  
 268 to natural clay porosity, while  $K$ -value could be calculated by Otsu's method (Otsu,  
 269 1979).

270

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

$$275 \quad n_{natural} = n_{3-D}, \quad n_{3-D} = \frac{V_{pore}}{V_{total}} = \frac{\int_0^a p dg}{\int_0^{255} p dg} \quad (1)$$

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

279

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

283

284 In fact, the grey histogram is a kind of statistical analysis of the 3-D surface model. The  
285 grey histogram represents the number of pixels with a certain grey value in the SEM  
286 image. Besides, in the grey histogram, if the threshold value is  $a$ , the grey range of  $(0,$   
287  $a)$  represents the pores, and the grey interval of  $[a, 255]$  represents the particles (Fig.  
288 5a).

289

290 The 3-D surface porosity of the SEM image is the ratio of the pore area to the total area.  
291 The pore area can be calculated by integrating the interval of  $(0, a)$  in the grey  
292 histogram, while the total area of the grey histogram equalled the whole area of the  
293 SEM image. In the example shown in Fig. 4, the SEM image (Fig.4e) of a clay sample  
294 has a natural porosity of 0.311. If the threshold is 43 (the value of the threshold is  
295 accurate to the single-digit), the 3-D surface porosity of the SEM image is 0.319, and  
296 0.319 is the closest to the natural porosity 0.311. Therefore, the  $A$ -value is 43.  
297 Subsequently, the  $K$ -value of the threshold range (Fig.4c) is calculated by Otsu's  
298 method. For the SEM image in Fig.4e, the  $K$ -value is 76, and hence, for this example  
299 considered here, the threshold range would be from 43 to 76.

300



301 (2) The determination of an optimal threshold value

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

307

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

316

317 3.3.2 Measurement of micro-parameters

318 Based on the threshold value calculated above, the SEM image was converted into  
319 a binarized image, then the measurement of micro-parameters can be done with the help  
320 of PCAS software (Liu et. al, 2011, 2013 and Tang et al, 2012). Finally, the image

321 porosity, mean shape factor, mean fractal dimension, roundness, and orientation  
322 probability entropy can be obtained.

323

### 324 **3.4 Interval estimation for clay microscopic quantification**

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

330

### 331 **3.5 Micro-parameters for clay microscopic quantification**

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

340

## 341 **4. An example for characterization of clay microstructure**

#### 342 **4.1 Materials**

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

347

#### 348 **4.2 Methods**

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

352

#### 353 **4.3 Test plan**

354 In this paper, four groups of tests were conducted. Each group has its unique test  
355 objective, and each group contains several SEM samples. The first group is used to  
356 compare different milling directions (Table 3). SEM samples from cross-section (1#,  
357 3#, 5#, 7#) and vertical-section (2#, 4#, 6#, 8#) of 4 different remolded clay samples  
358 (I#, II#, III#, IV#) were tested. These SEM samples were air-dried, and 10 SEM images  
359 with magnification of 2000x were taken for each sample.

360

361 The second group aims at comparing different drying techniques. For this purpose, 4  
362 SEM samples (9#, 10#, 11#, 12#) were prepared. All these SEM samples were made

363 from remolded clay sample V#, whose void ratio and water content were 0.45 and 20%  
364 respectively. These SEM samples were dried in air (for 9# sample), hot-air oven (for  
365 10# sample), critical point-device (for 11# sample), and liquid nitrogen frozen-vacuum  
366 device (for 12# sample), respectively. The cross-sections of these SEM samples were  
367 observed, and 10 SEM images with a magnification of 2000 times were taken for each  
368 sample.

369

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

376

377 The fourth group is used for comparing different quantities of SEM images. One SEM  
378 sample, which was used in the third group (i.e., 13#), along with another SEM sample  
379 (14#) was used for this purpose. Sample 14# is milled from the remolded clay sample  
380 VII#, whose void ratio and water content were 0.66 and 30%, respectively. Both were  
381 air-dried, and the cross-section was observed. 10 images with a magnification of 2000×  
382 were taken for both SEM samples.

383

#### 384 **4.4 Results of SEM imaging**

##### 385 4.4.1 Effect of different milling directions on clay microscopic quantification

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

390

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

404

## 405 4.4.2 Effect of different drying techniques on clay microstructure quantification

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

410

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

424

#### 425 4.4.3 Determination of optimal image magnification

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

435

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

442

443 Based on these figures, there exists an optimal magnification for each parameter. For  
444 example, in Fig.9a, the magnifications of the dots that are located in the area between  
445 the upper line and the lower line are better than the magnifications of dots that are

446 outside this area, since parameter values of dots inside are in the confidence interval.  
447 Furthermore, the magnification of the dot that is closer to the middle line is better than  
448 that of another dot that is farther, because there is a greater probability that the mean  
449 value can represent the real value of the micro-parameter.

450

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

461

#### 462 4.4.4 Determination of optimal image quantity

463 According to the results above, SEM images with a magnification of 1500× were  
464 used here to analyze the optimal quantity of images. In theory, when the quantity of  
465 SEM images is larger, the mean value of micro-parameters would be closer to the true



466 value. But the time spent taking SEM photos would be more. The more efficient way  
467 was to take fewer images to get the most accurate values of the micro- parameter.

468

469 10 SEM images had been taken for the SEM sample 13#. Five micro-parameters of  
470 each image were calculated by the image processing method introduced above and are  
471 listed in Table 7. Later, the interval estimation was done based on these data. For each  
472 micro-parameter, interval estimation was carried out for the different quantities of  
473 images, such as 2 images, 3 images, or 10 images. These results are presented in Fig.10.

474

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

483

484 Fig. 10 shows that the length of confidence interval decreased as the number of images  
485 increased when the number of images was less than 5. From 2 to 5 images, there is 78%,  
486 83%, 78%, 81%, 78% decline for  $n$ ,  $F$ ,  $H_m$ ,  $D_v$ ,  $R_0$ , respectively. But the declining

487 trend was less significant when the number of images was more than 5, except for the  
488 mean shape factor  $F$ , which might be caused by accidental factors. And from 6 to 10  
489 images, there is 23%, 37%, 38%, 36% decline for  $n$ ,  $H_m$ ,  $D_v$ ,  $R_0$ , respectively.

490

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

498

499 **To** verify these observations, another SEM sample 14# was used. The void ratio and  
500 water content of SEM sample 14# were 0.656 and 30%, respectively. The methodology  
501 adopted was the same as in the case of SEM sample 13#. Micro-parameters of each  
502 image are listed in Table 9, the results of interval estimation are shown in Fig.11, and  
503 the error  $\varepsilon$  are listed in Table 10. The results show that when the number is equal to 5  
504 or more, the error  $\varepsilon$  are always less than 25% for all the micro-parameters. Therefore,  
505 5 images and more than 5 images meet the requirements, hence, the smallest number  
506 of 5 can be chosen as the optimal number.

507

## 508 **5. Discussion**

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

513

### 514 ***5.1 Demonstration for reliability and limitation of the 3-D surface model***

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

523

524 In Fig.12a, the high-energy incident electrons were emitted into the clay sample by the  
525 electron gun in the SEM, then the secondary electrons in the clay particles can be  
526 excited. These secondary electrons can be captured by the detectors in the SEM, and  
527 finally, generate an SEM image with different brightness areas inside. Higher  
528 brightness area means more electrons accumulation, and more electrons mean a shorter

529 distance from the structure surface to the imaging surface. Therefore, the bright areas  
530 in the SEM image represent clay particles because the secondary electrons were  
531 generated from clay particles. While the relative dark areas in the SEM image represent  
532 pores since the pores cannot generate electrons. Since the degree of brightness in SEM  
533 images can be described by the grey value, and the distance from the structure surface  
534 to the imaging surface describes the spatial features of the structure surface. Therefore,  
535 there is a relation between the grey value and the spatial features of the structure surface,  
536 as shown in Fig. 12b.

537

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

542

## 543 **5.2 Comparison with other thresholds determination methods**

544 To verify the proposed *A-K* method is better than the typical theoretical methods.  
545 An example is given to compare different thresholds determination methods in the  
546 following. Fig. 13a is the original image, which is the same as Fig. 4e. The thresholds  
547 for Fig. 13a obtained through the *A-K* method (Fig. 13b), the method of iterative global  
548 thresholding (Fig. 13c), the Otsu's method (Fig. 13d), and the method of local  
549 properties based thresholding (Fig. 13e) were 55, 106, 76, 52 respectively. The different

550 threshold values from these methods resulted in a great variation in the resultant binary  
551 image after processing. The dark area in Fig. 13b and Fig. 13c was much more than the  
552 image in Fig. 13a, while the dark area in Fig. 13d was a little lesser than Fig.13a.  
553 Meanwhile, Fig. 13b is the most similar to Fig. 13a. Besides, the difference in the binary  
554 image could also be characterized by image porosity  $n_i$ . As shown in Fig. 13f,  $n_i$   
555 varied from 0.44 to 0.86. The natural porosity of Fig. 13a is 0.311, and  $n_i$  of Fig. 13b  
556 and Fig. 13e are much closer to the natural porosity than Fig. 13c and Fig. 13d.

557

### 558 ***5.3 Comparison with other quantification methods***

559 Few studies have investigated the quantification method for SEM experiments  
560 systematically. Most of the available studies have raised the challenges in  
561 microstructure characterization, but still, need to be further studied to find the solutions.  
562 For example, the differences between these drying techniques are inconsistent with the  
563 research of other studies (Korpa, 2006; Osipov, 1985), and the importance of milling  
564 direction for credible quantification is also agreed most with Trzciński (2004). But in  
565 their research, no comparison of influence on micro-parameters ( image porosity, mean  
566 shape factor, mean fractal dimension, roundness, and orientation probability entropy )  
567 was presented. In terms of the image magnification, Zhou et al. (2021) compare the  
568 magnification of 500×, 800×, 3000× and 20000×. The magnification of 800× is used  
569 for further quantitative analysis, but no theoretical reason was given. As for the number  
570 of images, Trzciński (2004) proposed that the accuracy of characterization can be

571 improved by a bigger number of the analysed sample parts because the data from one  
572 part do not coincide with results from a different portion of the same SEM sample. But  
573 no further discussion on the number of images was provided. The improved  
574 quantification method in this paper presents solutions to these challenges.

575

## 576 **6. Conclusions**

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

580

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

586

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

591

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

597

598

#### 599 **Data Availability Statement**

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

603

#### 604 **Acknowledgements**

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

613 **References**

- 614 1. Alba M D, Chain P ans Orta M M (2009) Chemical reactivity of argillaceous material  
615 in engineered barrier: rare earth disilicate formation under subcritical conditions.  
616 Applied Clay Science **43(3–4)**: 369-375.
- 617 2. Bangaru S S, Wang C, Hassan M, et al. (2019) Estimation of the degree of hydration  
618 of concrete through automated machine learning based microstructure analysis–A  
619 study on effect of image magnification. Advanced Engineering Informatics **42**:  
620 100975.
- 621 3. Bansal S, Maini R (2013) A Comparative Analysis of Iterative and Ostu's  
622 Thresholding Techniques. International journal of computer applications, **66(12)**:  
623 45-47.
- 624 4. Bennett R, Hulbert M (2012) *Clay microstructure*. D. Reidel publishing company,  
625 Dordrecht, Netherlands.
- 626 5. Bohor B F, Hughes R E (1971) Scanning electron microscopy of clays and clay  
627 minerals. Clays and Clay Minerals **19(1)**: 49-54.
- 628 6. Burton G J, Pineda J A, Sheng D et al. (2015) Microstructural changes of an  
629 undisturbed, reconstituted and compacted high plasticity clay subjected to wetting  
630 and drying. Engng Geol **193**: 63–373
- 631 7. Chen Y M, Shi J Y, Zhu W, et al. (2012) A review of geoenvironmental engineering.  
632 China Civil Engineering Journal **45(4)**: 165-182.



- 633 8.Cheremkhin P A, Kurbatova E A (2019) Comparative appraisal of global and local  
634 thresholding methods for binarisation of off-axis digital holograms. *Optics and*  
635 *Lasers in Engineering* **115**: 119-130.
- 636 9.Cox M R and Budhu M (2008) A practical approach to grain foctr quantification.  
637 *Engineering Geology* **96 (1-2)**: 1-16.
- 638 10.Dathe A, Eins S, Niemeyer J, Gerold G (2001) The surface fractal dimension of the  
639 soil-pore interface as measured by image analysis. *Geoderma* **103 (1-2)**: 203-229.
- 640 11.Delage P, Lefebvre G (1984) Study of the structure of a sensitive Champlain clay  
641 and of its evolution during consolidation. *Canadian Geotechnical Journal* **21(1)**:  
642 21-35.
- 643 12.Dey S, Basu T S, Roy B et al. (1989) A new rapid method of air-drying for scanning  
644 electron microscopy using tetramethylsilane. *Journal of Microscopy* **156(2)**: 259-  
645 261.
- 646 13. Di Remigio G., Rocchi I, Zania V (2021) New method for a SEM-based quantitative  
647 microstructural clay analysis-MiCA. *Applied Clay Science* **214**: 106248.
- 648 14.Dohrmann R, Kaufhold S and Lundqvist B (2013) The role of clays for safe storage  
649 of nuclear waste. In *Developments in Clay Science*.Elsevier, Amsterdam,  
650 Netherlands, vol.1, pp. 677-710.
- 651 15.Estabragh A R, Kholoosi M, Ghaziani F, et al. (2018) Mechanical and leaching  
652 behavior of a stabilized and solidified anthracene-contaminated soil. *Journal of*  
653 *Environmental Engineering* **144(2)**: 04017098.

- 654 16.Hemes S, Desbois G, Urai J L, De Craen M, Honty M (2013) Variations in the  
655 morphology of porosity in the Boom Clay Formation: insights from 2D high  
656 resolution BIB-SEM imaging and Mercury injection Porosimetry. Netherlands  
657 Journal of geosciences **92(4)**: 275-300.
- 658 17.Horpibulsuk S, Rachan R, Chinkulkijniwat A, et al. (2010) Analysis of strength  
659 development in cement-stabilized silty clay from microstructural considerations.  
660 Construction and building materials **24(10)**: 2011-2021.
- 661 18.Huang J, Liu X, Cao G (2017) Experimental Study of Chemical Composition of  
662 Hangzhou Soft Clay. 2017 International Conference on Advanced Materials  
663 Science and Civil Engineering (AMSCE 2017): 116-118.
- 664 19.Janecek M and Robert K, eds. (2016) Modern electron microscopy in physical and  
665 life sciences. Intech, Rijeka, Croatia. pp.167-177.
- 666 20.Kaczyński R and Trzciński J (1997) Ilościowa analiza mikrostrukturalna w  
667 skaningowym mikroskopie elektronowym (SEM) typowych gruntów Polski.  
668 Przegląd Geologiczny **45**: 721–726.
- 669 21.Kjellsen K O, Monsøy A, Isachsen K, Detwiler R J (2003). Preparation of flat-  
670 polished specimens for SEM-backscattered electron imaging and X-ray  
671 microanalysis—importance of epoxy impregnation. Cement and concrete research  
672 **33(4)**: 611-616.
- 673 22.Korpa A, Trettin R (2006). The influence of different drying methods on cement  
674 paste microstructures as reflected by gas adsorption: Comparison between freeze-

- 675 drying (F-drying), D-drying, P-drying and oven-drying methods. Cement and  
676 Concrete Research **36(4)**: 634-649.
- 677 23.Lawrence G P, Payne D, Greenland D J (1979). Pore size distribution in critical  
678 point and freeze dried aggregates from clay subsoils. Journal of soil science **30(3)**:  
679 499-516.
- 680 24.Lin L, Yan J, Chen G, et al. (2018) Does magnification of SEM image influence  
681 quantification of particulate matters deposited on vegetation foliage. Micron **115**:  
682 7-16.
- 683 25.Lindroth M, Bell Jr P B. and Fredriksson B A (1988) Comparison of the effects of  
684 critical point drying and freeze drying on cytoskeletons and microtubules. Journal  
685 of Microscopy **151(Pt.2)**: 103-114.
- 686 26.Liu C, Shi B, Zhou J, et al.(2011) Quantification and characterization of  
687 microporosity by image processing, geometric measurement and statistical  
688 methods: application on SEM images of clay materials. Applied Clay Science **54(1)**:  
689 97-106.
- 690 27.Liu C, Tang C S, Shi B, et al. (2013) Automatic quantification of crack patterns by  
691 image processing. Computers & Geosciences **57**: 77-80.
- 692 28.Liu X, Wang J, Ge L, et al. (2017) Pore-scale characterization of tight sandstone in  
693 Yanchang Formation Ordos Basin China using micro-CT and SEM imaging from  
694 nm-to cm-scale. Fuel **209**: 254-264.

- 695 29. Mahaney W C, Stewart A, Kalm V (2001) Quantification of SEM microtextures  
696 useful in sedimentary environmental discrimination. *Boreas* **30(2)**: 165-171.
- 697 30. Mazumder M, Ahmed R, Ali AW, et al. (2018) SEM and ESEM techniques used for  
698 analysis of asphalt binder and mixture: A state of the art review. *Construction and*  
699 *Building Materials* **186**: 313-329.
- 700 31. Ministry of Water Resources of the People's Republic of China (2019) Standard for  
701 geotechnical test method GB/T 50123-2019, National Bureau of Quality and  
702 Technical Supervision, Ministry of Construction in China, Beijing.
- 703 32. Neyman J (1937) Outline of a theory of statistical estimation based on the classical  
704 theory of probability. *Philosophical Transactions of the Royal Society of London.*  
705 *Series A, Mathematical and Physical Sciences* **236(767)**: 333-380.
- 706 33. Osipov B И (1985) The nature of the strength and deformation properties of clay-  
707 like soils and rocks. Geological Publishing House, Beijing, China.
- 708 34. Otsu N (1979) A threshold selection method from grey-level histograms. *IEEE*  
709 *transactions on systems, man, and cybernetics* **9(1)**: 62-66.
- 710 35. Oztoprak S and Pisirici B (2011) Effects of micro structure changes on the macro  
711 behaviour of Istanbul (Turkey) clays exposed to landfill leachate. *Engineering*  
712 *Geology* **121(3-4)**: 110-122.
- 713 36. Prakongkep N, Suddhiprakarn A, Kheoruenromne I et al. (2010) SEM image  
714 analysis for characterization of sand grains in Thai paddy soils. *Geoderma* **156 (1–**  
715 **2)**: 20-31.

- 716 37.Pusch R, Weston R (2003) Microstructural stability controls the hydraulic  
717 conductivity of smectitic buffer clay. *Applied Clay Science* **23 (1–4)**: 35-41.
- 718 38.Sezer G I, Ramyar K, Karasu B, Goktepe A B, Sezer A (2008) Image analysis of  
719 sulfate attack on hardened cement paste. *Materials and Design* **29**: 224–231.
- 720 39.Shaikh S H, Maiti A, Chaki N. (2011) Image binarization using iterative partitioning:  
721 A global thresholding approach. In *Proceedings of 2011 International Conference*  
722 *on Recent Trends in Information Systems*. **IEEE**, Kolkata, India, pp: 281-286.
- 723 40.Sun W J, Cui Y J (2018) Investigating the microstructure changes for silty soil  
724 during drying. *Géotechnique* **68(4)**: 370-373.
- 725 41.Taillon J A, Pellegrinelli C, Huang Y L, et al. (2018) Improving microstructural  
726 quantification in FIB/SEM nanotomography. *Ultramicroscopy* **184**: 24-38.
- 727 42.Tang C S, Shi B, Cui YJ, et al (2012) Desiccation cracking behavior of  
728 polypropylene fiber–reinforced clayey soil. *Canadian Geotechnical Journal* **49(9)**:  
729 1088-1101.
- 730 43.Terzaghi K (1925) *Erdbaumechanik auf bodenphysikalischer Grundlage*. Franz.  
731 Deuticke. Wien.
- 732 44.Toprak, E, Olivella, S, Pintado X (2018) Modelling engineered barriers for spent  
733 nuclear fuel repository using a double-structure model for pellets. *Environmental*  
734 *Geotechnics* **7(1)**: 72-94.
- 735 45.Tournassat C, Steefel C I, Bourg I C, et al (2015) *Natural and engineered clay*  
736 *barriers*. Elsevier, Amsterdam, Netherlands.

- 737 46. Trzciński J. (2004). Combined SEM and computerized image analysis of clay soils  
738 microstructure: technique & application. In *Advances in geotechnical*  
739 *engineering: The Skempton conference: Proceedings of a three day conference*  
740 *on advances in geotechnical engineering, organised by the Institution of Civil*  
741 *Engineers and held at the Royal Geographical Society*. Thomas Telford, London,  
742 UK, pp. 654-666.
- 743 47. Wu C, Chu J (2020) Biogrouting method for stronger bond strength for aggregates.  
744 *Journal of Geotechnical and Geoenvironmental Engineering* **146(11)**: 06020021.
- 745 48. Yang S, Liu W (2019) Research on Unconstrained Compressive Strength and  
746 Microstructure of Calcareous Sand with Curing Agent. *Journal of Marine Science*  
747 *and Engineering* **7(9)**: 294.
- 748 49. Youn H, Tonon F. (2010) Effect of air-drying duration on the engineering properties  
749 of four clay-bearing rocks in Texas. *Engineering Geology* **115(1-2)**: 58-67.
- 750 50. Zhou, C., Yu, L., Huang, Z., Liu, Z., Zhang, L. (2021). Analysis of microstructure  
751 and spatially dependent permeability of soft soil during consolidation  
752 deformation. *Soils and Foundations* **61**: 708-733.
- 753
- 754
- 755
- 756
- 757
- 758
- 759

760

**List of Table Captions**761 **Table 1** Physical properties of Hangzhou clay.762 **Table 2** Mineralogy and chemical composition of Hangzhou clay.763 **Table 3** Testing program of the first group for comparing different milling directions.764 **Table 4** Micro-parameters of different milling directions.765 **Table 5** Micro-parameters of different drying techniques.766 **Table 6** Micro-parameters of images with different magnifications of 13# SEM sample.767 **Table 7** Micro-parameters of each image for 13# SEM sample.768 **Table 8** Error  $\varepsilon$  of each micro-parameter for 13# SEM sample.769 **Table 9** Micro-parameters of each image for 14# SEM sample.770 **Table 10** Error  $\varepsilon$  of each micro-parameter for 14# SEM sample.

771

772

773

774

775

776

777

778

779

780

781 **Table 1** Physical properties of Hangzhou clay.

Water content, $w$ (%)	Bulk density, $\rho$ (g/cm <sup>3</sup> )	Specific gravity, $G_s$	Porosity, $n$	Liquid limit, WL	Plastic limit, WP	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%

782

783

784

785 **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

786

787

788

789

790

791

792



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

The No. of remolded				
No. of SEM sample	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

794

795

796

797

798

799

800

801

802

803 **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

804

805

806

807

808

809

810

811

812

813

814

815

816 **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

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832 **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

833

834

835

836

837

838

839

840

841

842

843

844

845

846 **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

847

848

849

850

851

852

853

854

855

856

857

858 **Table 8** Error  $\epsilon$  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	<b>0.231</b>	<b>0.179</b>	<b>0.244</b>	<b>0.236</b>	<b>0.234</b>
6 and 7	0.133	0.169	0.116	0.049	0.170
7 and 8	0.073	<b>0.490</b>	0.163	0.162	0.024
8 and 9	0.122	0.138	0.034	0.125	0.129
9 and 10	0.063	0.041	0.114	0.114	0.096

859

860

861

862

863

864

865

866

867

868

869

870

871

872 **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

873

874

875

876

877

878

879

880

881

882

883

884 **Table 10** Error  $\varepsilon$  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	<b>0.230</b>	<b>0.230</b>	<b>0.208</b>	<b>0.171</b>	<b>0.102</b>
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 and 10	0.077	0.020	0.051	0.016	0.120

885



## Appendix

### Appendix 1

#### A.1.1 Definition of interval estimation

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

#### A.1.2 Confidence interval

$P_\theta(\hat{g}_1(x) \leq g(\theta) \leq \hat{g}_2(x))$  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  $\alpha$  ( $0 < \alpha < 1$ ) is given, and

$$P_\theta(\hat{g}_1(x) \leq g(\theta) \leq \hat{g}_2(x)) \geq 1 - \alpha, \quad g(\theta) \in \Theta \quad (2)$$

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  $X = (X_1, \dots, X_n)$  is a sample drawn from  $N(\mu, \sigma^2)$ .  $\mu$  and  $\sigma^2$  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  $X = (X_1, \dots, X_n)$ , respectively.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (3)$$

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

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

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

Defining

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

$T$  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

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

where,

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

then the confidence space of  $\mu$  with confidence coefficient  $\alpha$  is

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

## 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:

$$n_i = \frac{A_1}{A_0} \quad (10)$$

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):

$$F_i = \frac{P}{L} \quad (11)$$

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:

$$F = \frac{\sum_{i=1}^n F_i}{m} \quad (12)$$

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):

$$D_v = \lim_{r \rightarrow 0} \frac{\ln N(r)}{\ln r} \quad (13)$$

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).

$$R_0 = \frac{A_p}{A'} \quad (14)$$

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  $\alpha_p$  is the direction of a soil particle.  $\alpha_p$  can be computed directly in PCAS software, and  $\alpha_p$  varies between  $[0^\circ-180^\circ]$ ;  $k$  is the number of equally divided areas in the whole particle direction range  $[0^\circ, 180^\circ]$ , 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  $10^\circ$ .  $P_i(\alpha_p)$  is the percentage of soil particles whose directions  $\alpha_p$  belong to a specific range, for example, when  $i = 1$ ,  $P_1(\alpha_p)$  means the percentage of soil particles whose directions  $\alpha_p$  belongs to the direction range of  $[0^\circ, 10^\circ]$ .

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 °C, 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.

**Figure 4.** The *A-K* threshold determining method in image processing.

**Figure 5.** The determination of optimal threshold value in *A-K* method.

**Figure 6.** Clay particles and quantitative analysis of different milling directions.

**Figure 7.** Quantitative analysis of different drying techniques.

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

**Figure 9.** Quantitative analysis for optimal magnification.

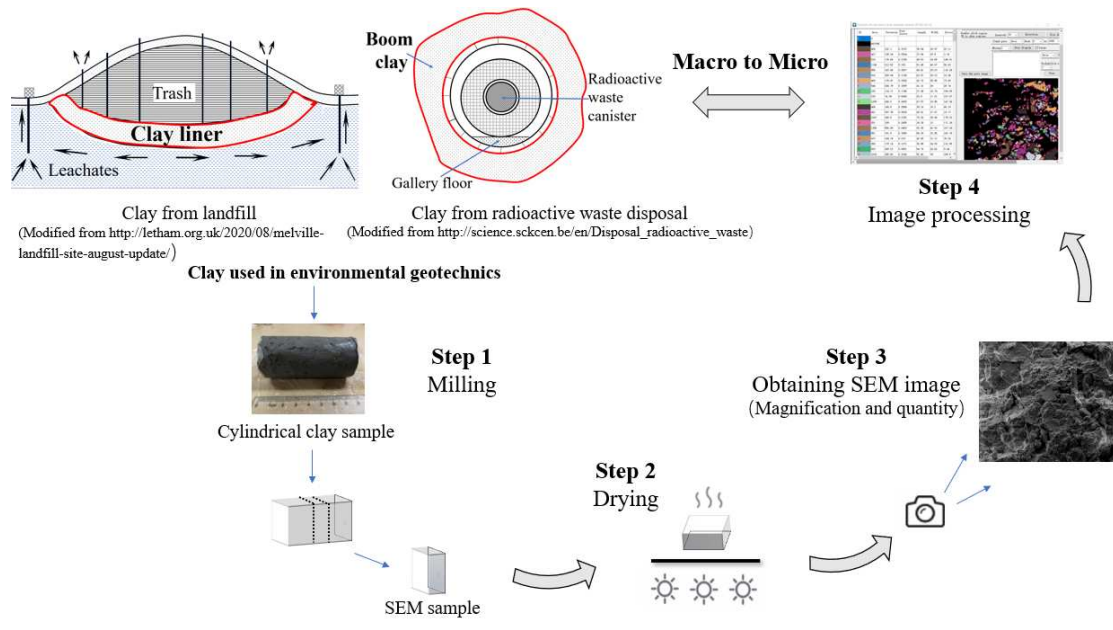
**Figure 10.** Analysis of optimal image quantity for SEM sample 13#.

**Figure 11.** Analysis of optimal image quantity for SEM sample 14#.

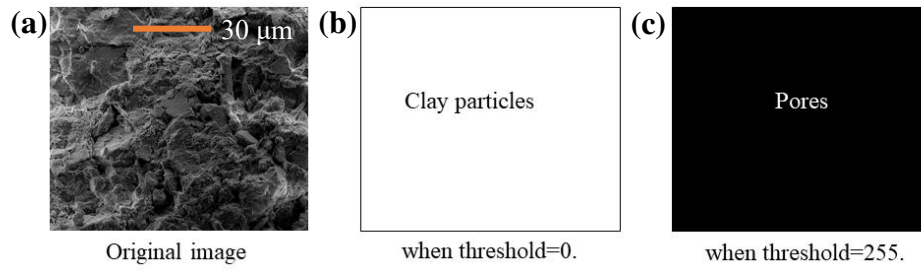
**Figure 12.** Discussion on 3-D surface model.

**Figure 13.** Quantitative comparison of different threshold value determining methods in image processing.

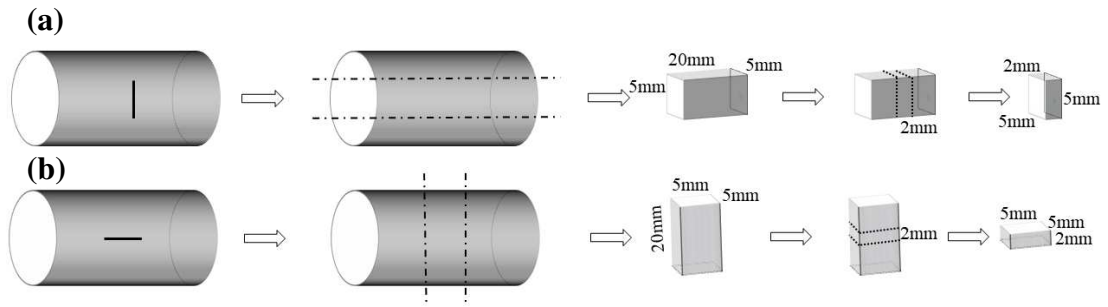




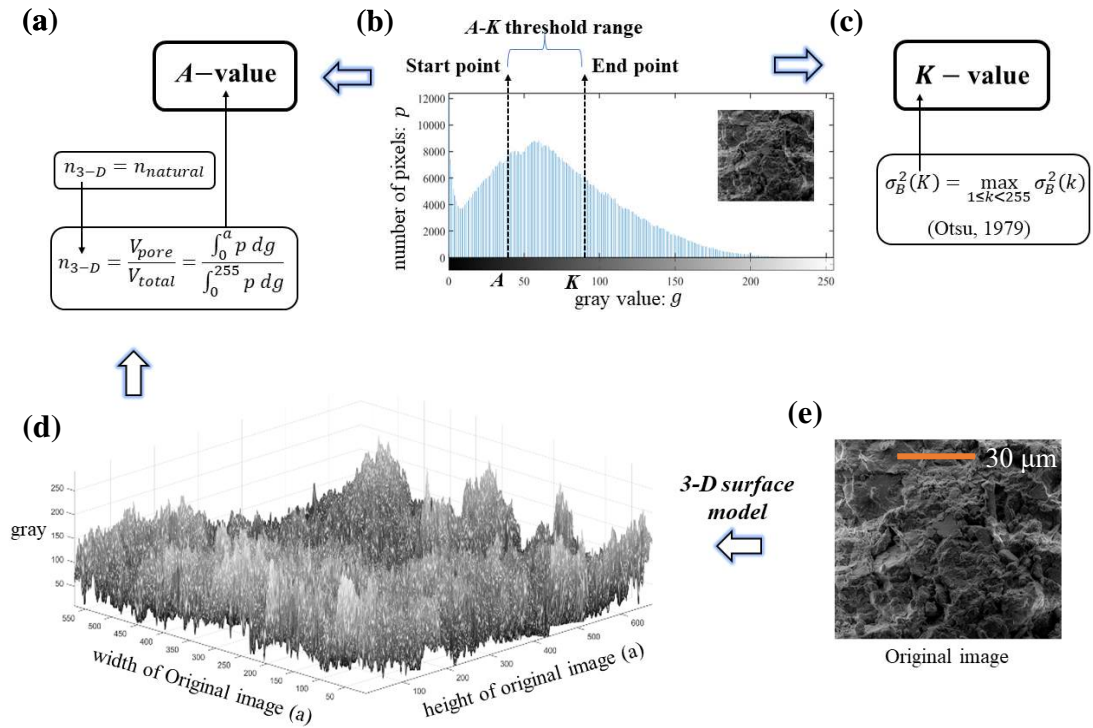
**Figure 1** Main difficulties in the quantification procedure for characterization of clay microstructure using SEM.



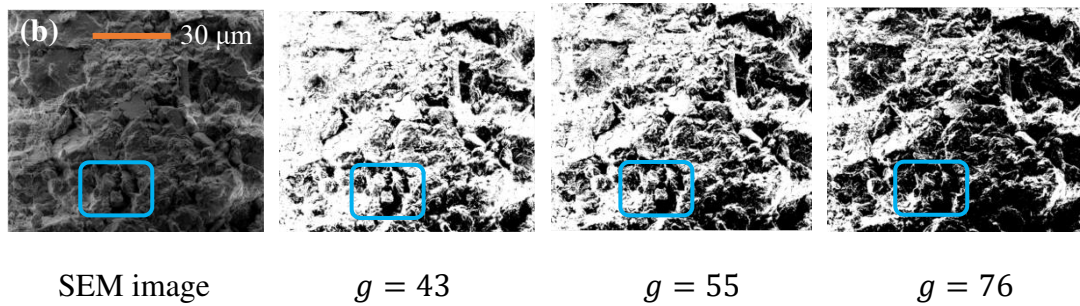
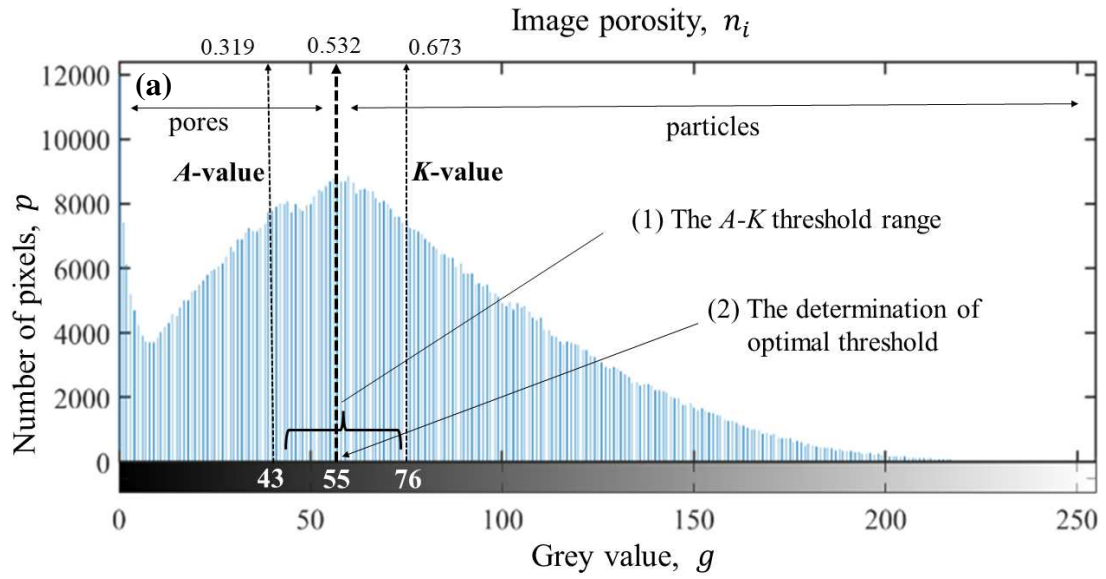
**Figure 2** The range of threshold value in SEM image. (a) Original image. (b) The binary image when threshold value is 0. (c) The binary image when threshold value is 255.



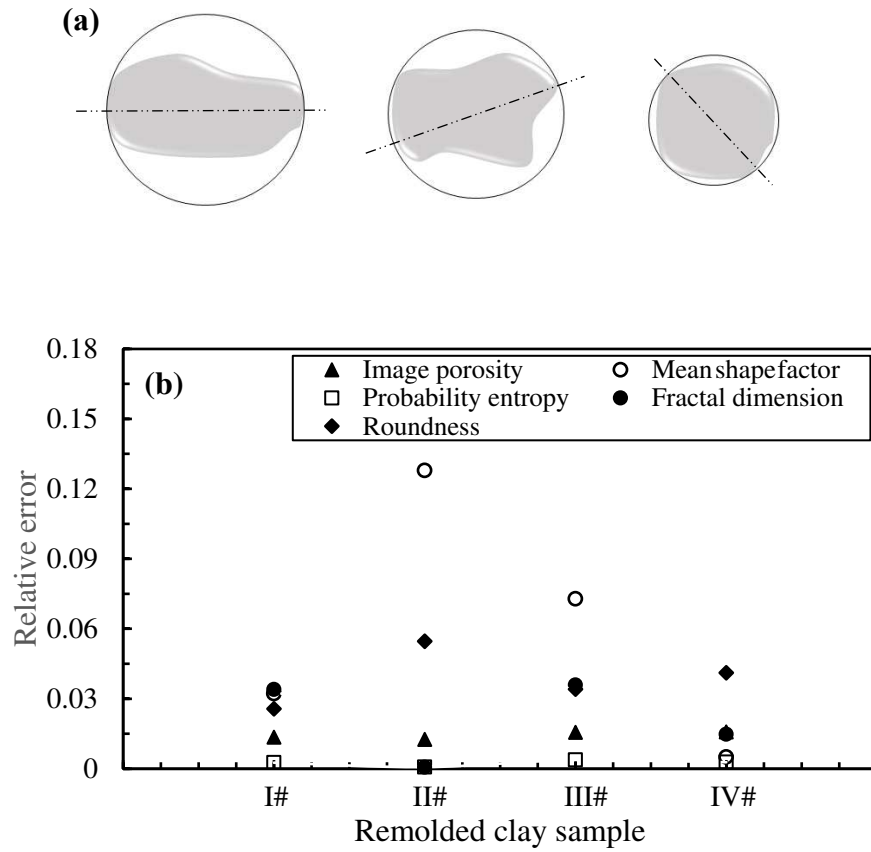
**Figure 3** Milling methods. (a) The milling method of cross-section. (b) The milling method of vertical-section.



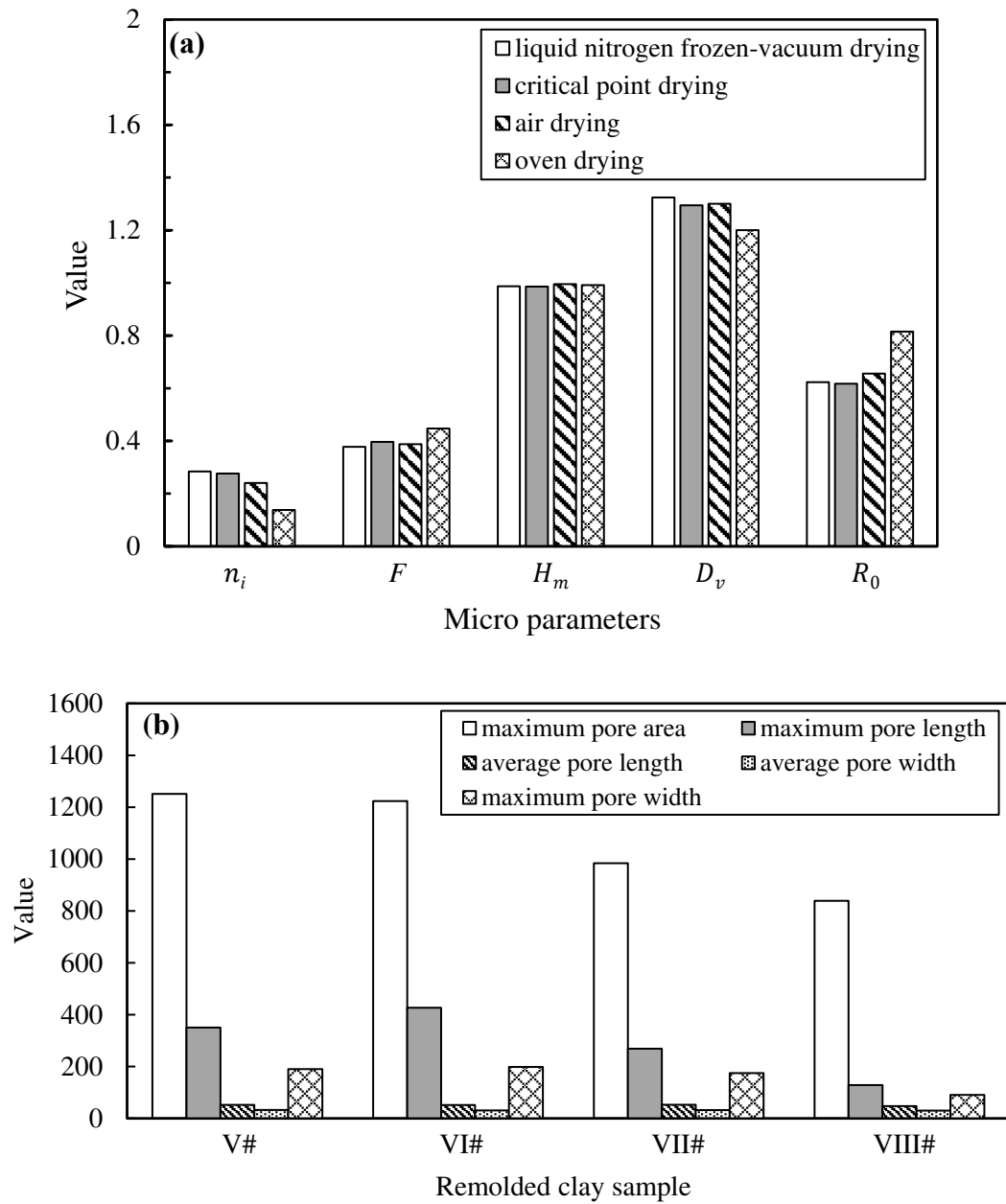
**Figure 4** The A-K threshold determining method in image processing. (a-c) The prediction of threshold range. (d-e) The 3-D surface model.



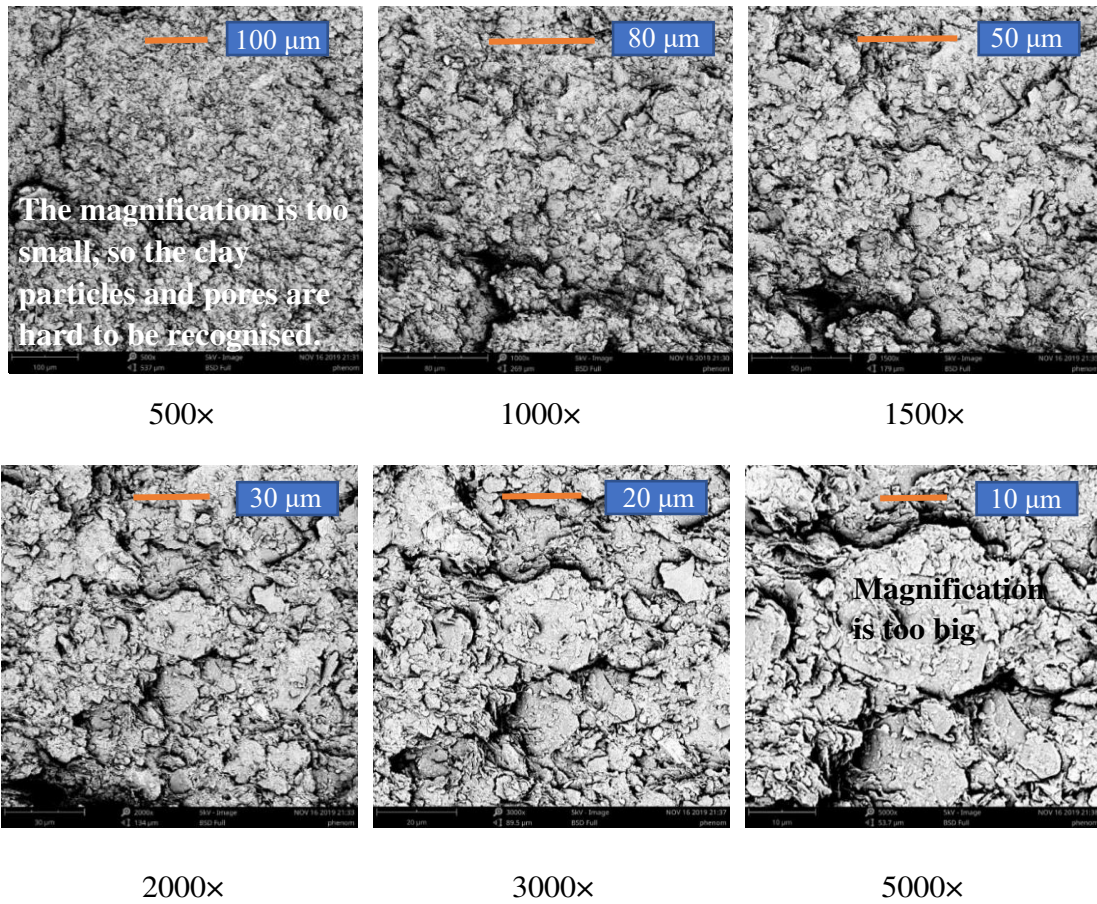
**Figure 5** The determination of optimal threshold value in  $A$ - $K$  method. (a) The  $A$ - $K$  threshold range for original SEM image. (b) The optimal threshold value was determined by artificial comparison of binary images with threshold in the  $A$ - $K$  range of [43, 76].



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

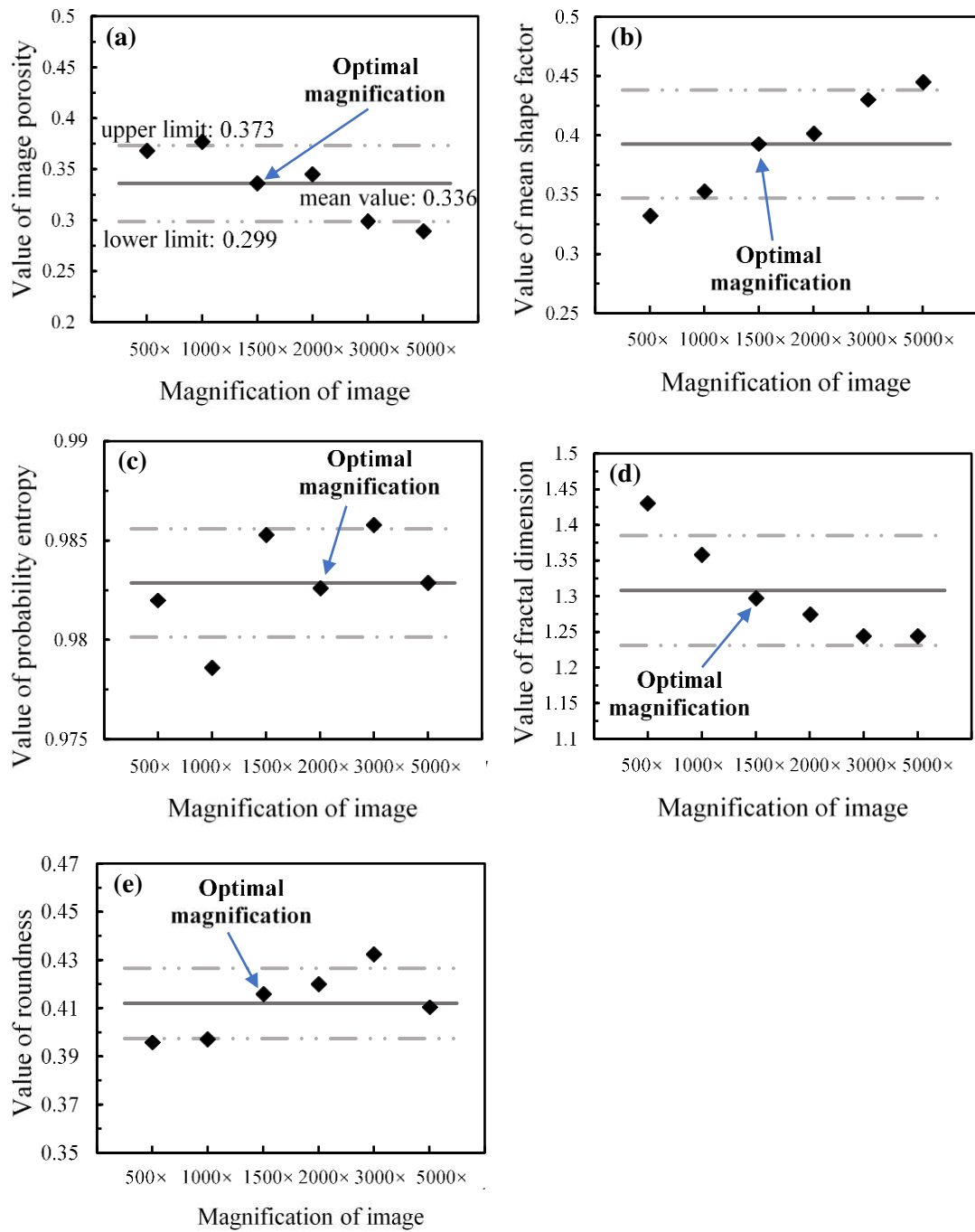


**Figure 7** Quantitative analysis of different drying techniques. (a) Comparison of micro parameters for different drying techniques. (b) Comparison of pore sizes for different drying techniques

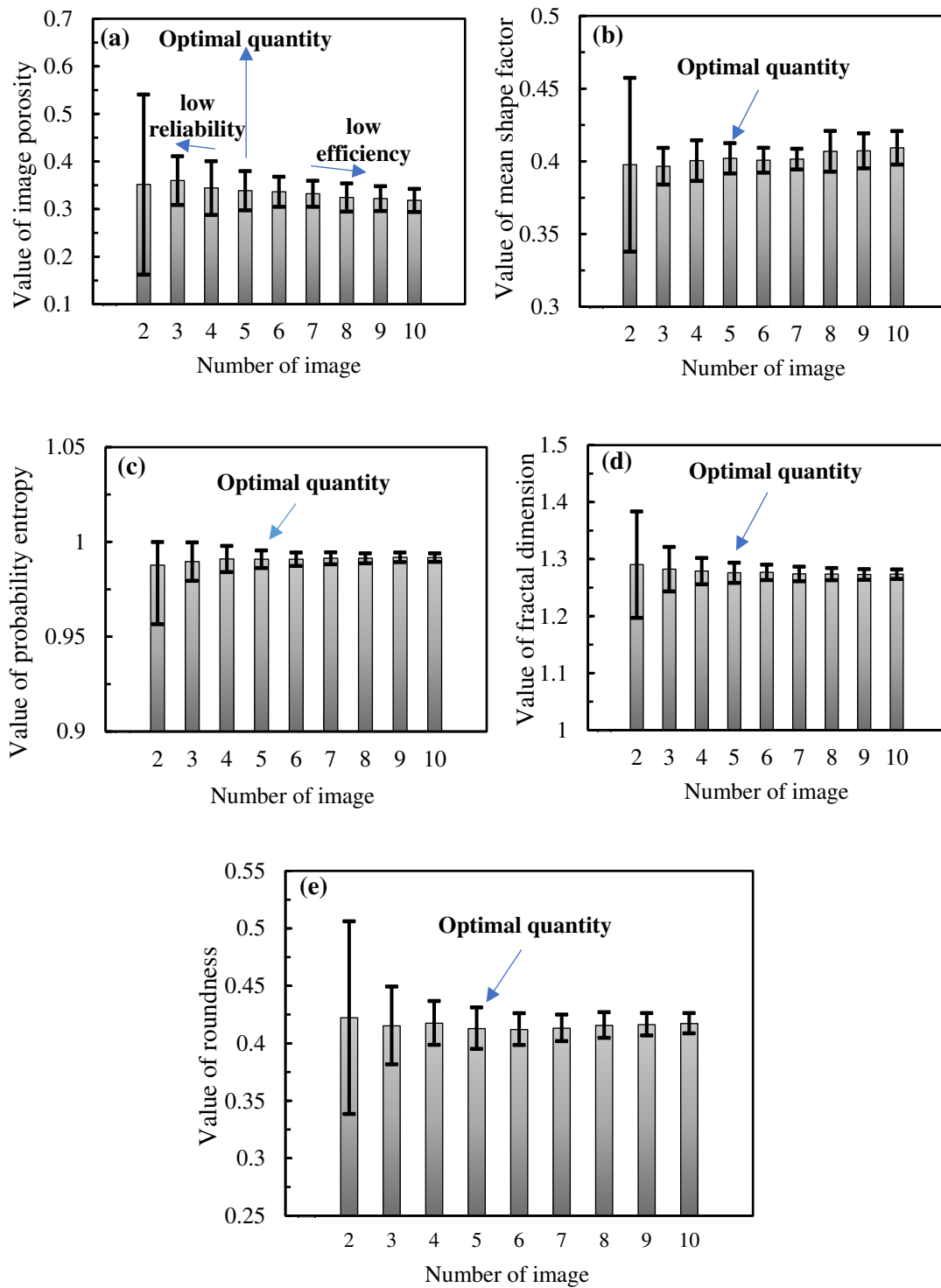


**Figure 8** The comparison of magnifications between 500× and 5000× for this sample.

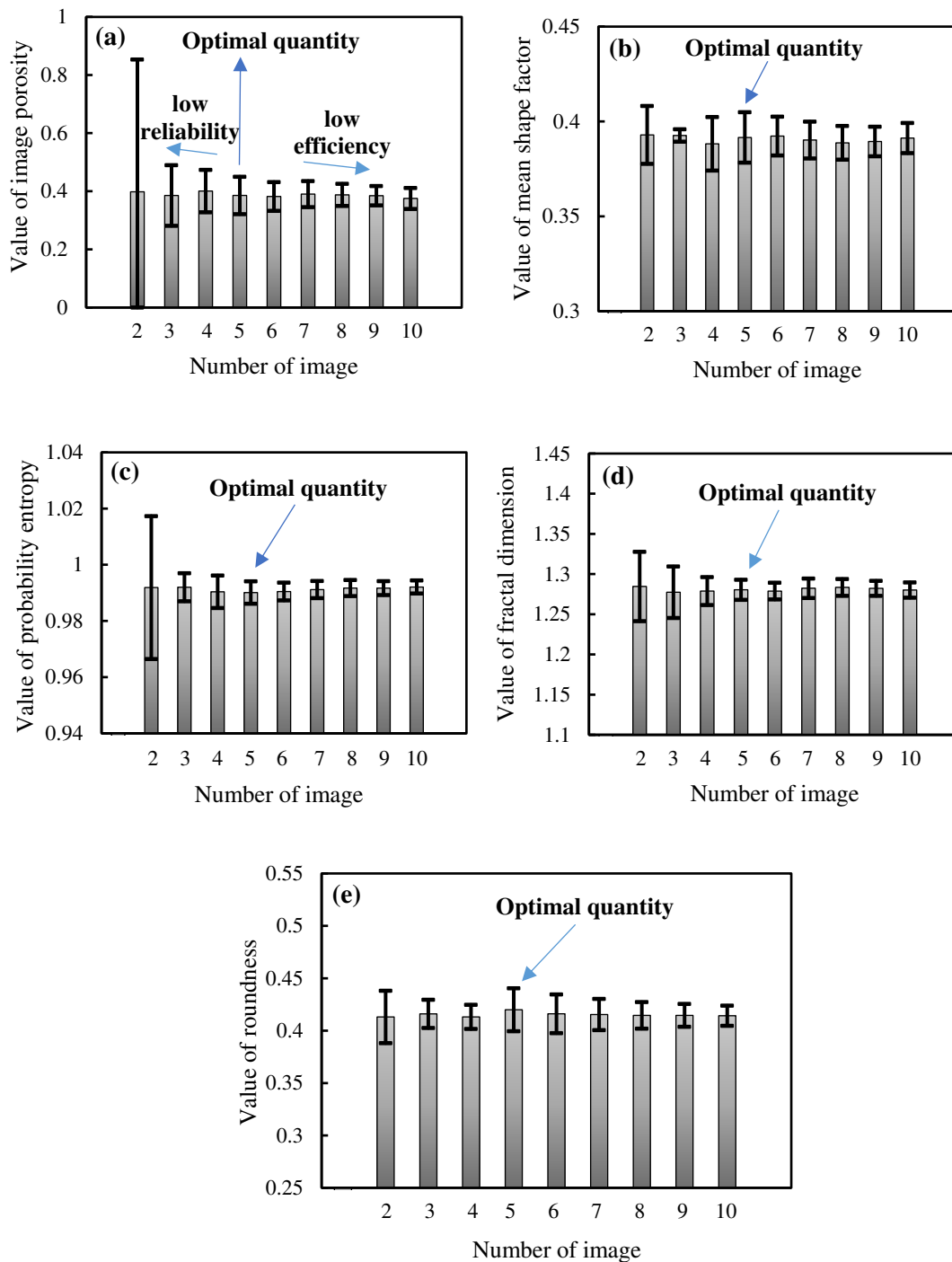




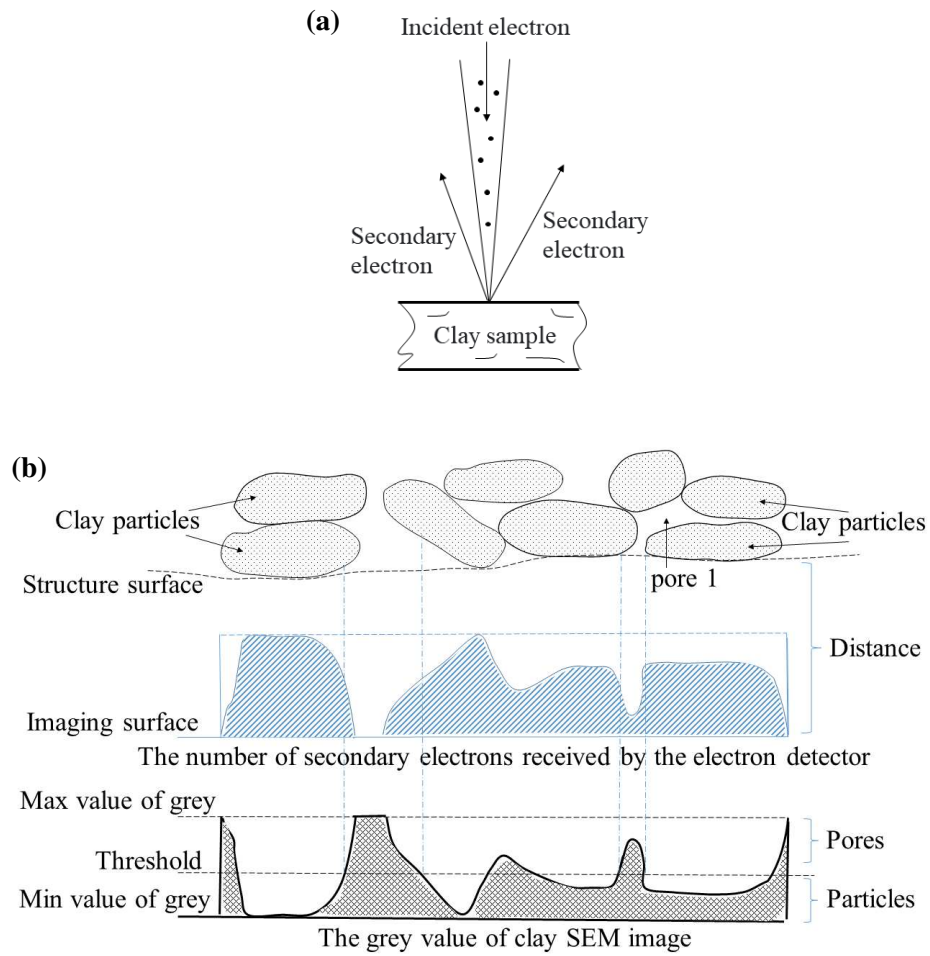
**Figure 9** Quantitative analysis for optimal magnification. (a) Image porosity. (b) Mean shape factor. (c) Probability entropy. (d) Fractal dimension. (e) Roundness.



**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.

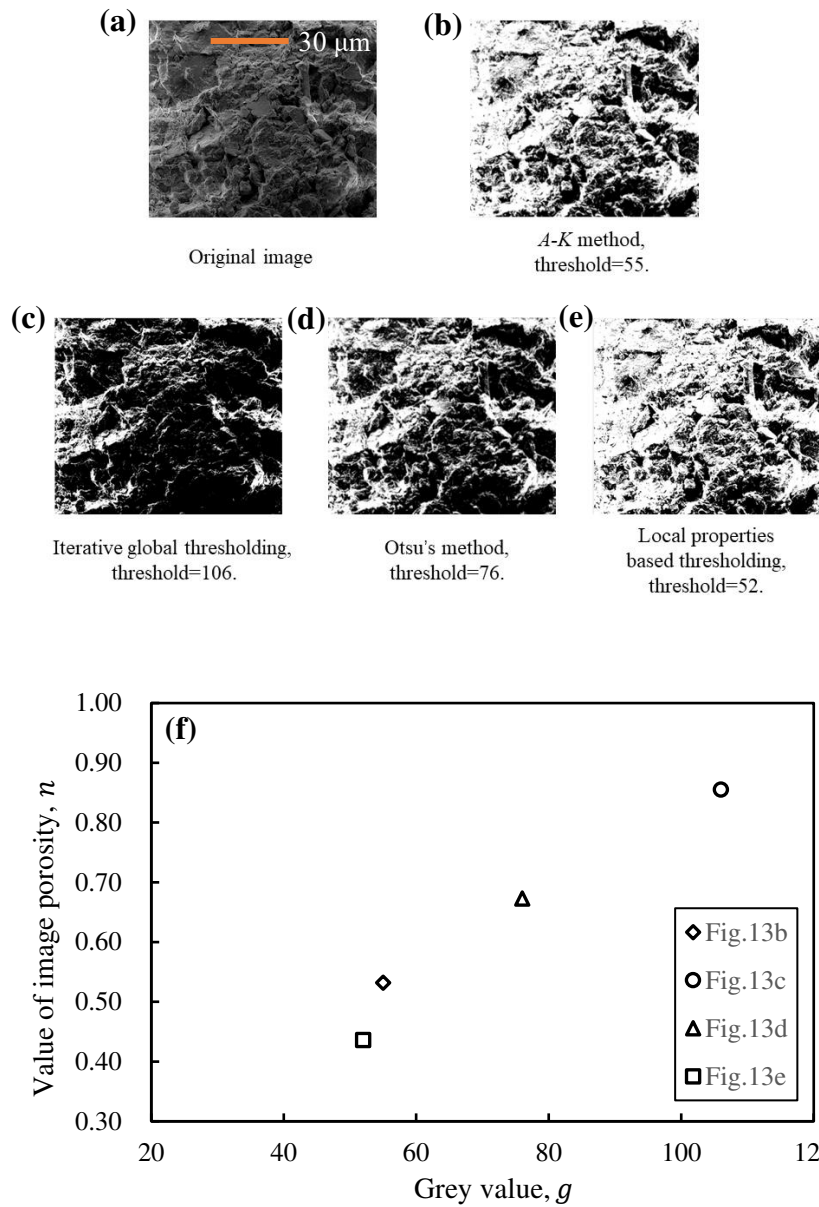


**Figure 11** Analysis of optimal image quantity for SEM sample 14#. The height of each bar is the mean value of a certain image quantity, while the short vertical line on the top represents the corresponding confident interval. (a) Statistic analysis of image porosity. (b) Mean shape factor. (c) Probability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