

# INTERNATIONAL SOCIETY FOR SOIL MECHANICS AND GEOTECHNICAL ENGINEERING



*This paper was downloaded from the Online Library of the International Society for Soil Mechanics and Geotechnical Engineering (ISSMGE). The library is available here:*

<https://www.issmge.org/publications/online-library>

*This is an open-access database that archives thousands of papers published under the Auspices of the ISSMGE and maintained by the Innovation and Development Committee of ISSMGE.*

*The paper was published in the proceedings of the 20<sup>th</sup> International Conference on Soil Mechanics and Geotechnical Engineering and was edited by Mizanur Rahman and Mark Jaksa. The conference was held from May 1<sup>st</sup> to May 5<sup>th</sup> 2022 in Sydney, Australia.*

# Quantitative analysis of SEM micrographs for clay particles orientation by means of computer vision: a sensitivity analysis

Analyse quantitative des micrographies SEM pour l'orientation des particules d'argile au moyen de la vision par ordinateur: une analyse de sensibilité

**Giorgia Di Remigio, Irene Rocchi & Varvara Zania**

*Civil Engineering Department, Technical University of Denmark, Denmark, gidire@byg.dtu.dk*

**ABSTRACT:** Soil being a particulate material, its mechanical behaviour depends on the size and shape of its particles, and how they are arranged, i.e. on its fabric. In the last decades, several studies tried to find a link between the soil microstructure and its behaviour at the macroscopic scale. In the case of fine-grained materials, the soil fabric was often investigated by means of Scanning Electron Microscopy (SEM) micrographs. However, due to the lack of automatized digital tools, most of these studies present only a qualitative interpretation of the particles orientation. The absence of quantifiable parameters hinders the development of an objective constitutive model connecting the micro- and macroscopic behaviour of soil. In this study, the latest developments in computer vision are employed to validate a code quantifying the orientation patterns of the clay particles, together with the main characteristics of their porosity. The aim of this work is to perform a sensitivity study on a custom-made numerical code, showing the effect of the parameters choice on the final results and suggesting effective strategies for their choice.

**RÉSUMÉ :** Le sol étant un matériau particulaire, son comportement mécanique dépend de la taille et de la forme de ses particules, ainsi que de la façon dont elles sont disposées, c'est-à-dire de son tissu. Au cours des dernières décennies, plusieurs études ont tenté de trouver un lien entre la microstructure du sol et son comportement à l'échelle macroscopique. Dans le cas de matériaux à grains fins, le tissu du sol a souvent été étudié au moyen de micrographies par microscopie électronique à balayage (MEB). Cependant, en raison du manque d'outils numériques automatisés, la plupart de ces études ne présentent qu'une interprétation qualitative de l'orientation des particules. L'absence de paramètres quantifiables entrave le développement d'un modèle constitutif objectif reliant le comportement micro- et macroscopique du sol. Dans cette étude, les derniers développements en matière de vision par ordinateur sont utilisés pour valider un code quantifiant les modèles d'orientation des particules d'argile, ainsi que les principales caractéristiques de leur porosité. L'objectif de ce travail est de réaliser une étude de sensibilité sur un code numérique sur mesure, en montrant l'effet du choix des paramètres sur les résultats finaux et en suggérant des stratégies efficaces pour leur choix.

**KEYWORDS:** SEM, Scanning Electron Microscope, microstructure, quantification, clay

## 1 INTRODUCTION

It has been early recognised that microstructure developed in clays because of geological history and physico-chemical interactions greatly influences the observed mechanical behaviour. On the other hand, Skempton (1969) found that regardless of the soil mineralogical content, on a void ratio to effective stress plot, all the curves tend to converge to a unique pattern, by analysing the effect of the sedimentation process on one dimensional behaviour. The research efforts addressing these issues intensified since the seventies, with several studies associating the compressibility and shear strength behaviour with the microstructure characteristics. In one of the pioneering studies on this field, Burland (1990) introduced the notion of intrinsic properties for the mechanical behaviour of reconstituted soils, as a reference state to identify effects of structure in sedimented natural clays. Later on, it was shown that the degree of de structure may depend on the applied reconstitution method (Fearon and Coop, 2000), while in the same study Scanning Electron Microscope (SEM) images illustrated clearly the effect of the remoulding process on the soil microstructure.

SEM micrographs are commonly used in geotechnical studies for the investigation of the soil microstructure. Cotecchia and Chandler (1997) investigated the effect of the microstructure on the pre-failure behaviour of natural clays, by quantifying the clay particles orientation at different stress levels using SEM. The analysis was performed by manually drawing the lines that the authors considered relevant for the problem analysis and, afterwards, calculating their main direction. Great effort was placed in quantitatively correlating the effect of the clay microstructure to its mechanical behaviour, however the absence of an automatic and systematic set of parameters selecting the

lines, introduces a degree of subjectivity that influences the final results. A significant step forward in the quantification process of clay microstructure was achieved by Martinez-Nistal et al. (1999), who introduced a computer vision based code, able to automatically analyse the clay particles orientation. Nevertheless, using this tool it is not possible to compare different images, ranking them from the more to the less ordered, as there had not yet defined relevant parameters that would allow for such operations. Finally, Hattab et al. (2010) performed a series of studies aiming to quantify kaolinite SEM micrographs particles orientations using the software ImageJ. For their analysis, they produce a segmented image of the micrograph and extrapolate the particles inclination compared to the horizontal axis, to build a 360° orientation diagram. However, the edge recognition and the segment identification is not customizable, and therefore, the user does not have a sufficient degree of control over the image analysis.

A custom-made numerical code, able to analyse the clay particles orientation and porosity has been developed, extensively described and validated by Di Remigio et al. (2021). The code allows the user to customize a wide set of parameters, which correct definition is at the base of a high quality set of results. For this work, a sensitivity study is performed, investigating the effect of the settings variation on the final results.

## 2 METHODOLOGY

Hereafter, the structure and main functions of the image analysis custom-made code MiCA, are introduced briefly. All details regarding the code can be found in Di Remigio et al. (2021). The code flowchart and the image settings are presented in Figure 1

in the blue and yellow column respectively. Once all the parameters and the image dataset are defined, each image is singularly filtered and analysed, first characterizing its orientation and then its porosity. To illustrate the different steps within the code a smectitic clay micrograph is selected (Figure 2). During the Image Filtering step, the original image (Figure 2.a) is turned into black and white (Figure 2.b), its colour and contrast are balanced, and then it is smoothed, aiming to reduce the background noise, by means of a Gauss Filter (GF), having customizable size and strength (Figure 2.c). Then, the smoothed image is convoluted with a vertical and a horizontal Sobel filter and processed to find the pixel, normalized gradient as per Sobel (1968). Afterwards, a Threshold for Edge Recognition (TED) is applied, so that the pixels having a gradient greater than the upper bound of TED are considered as Edges, while those which values are below the lower bound of TED are discarded from the edge recognition process (Figure 2.d). The total amount of pixels detected as edges as per Figure 2.d are counted and divided over the total amount of pixels in the image, to be multiplied time 100, in order to obtain the Ed% parameter.

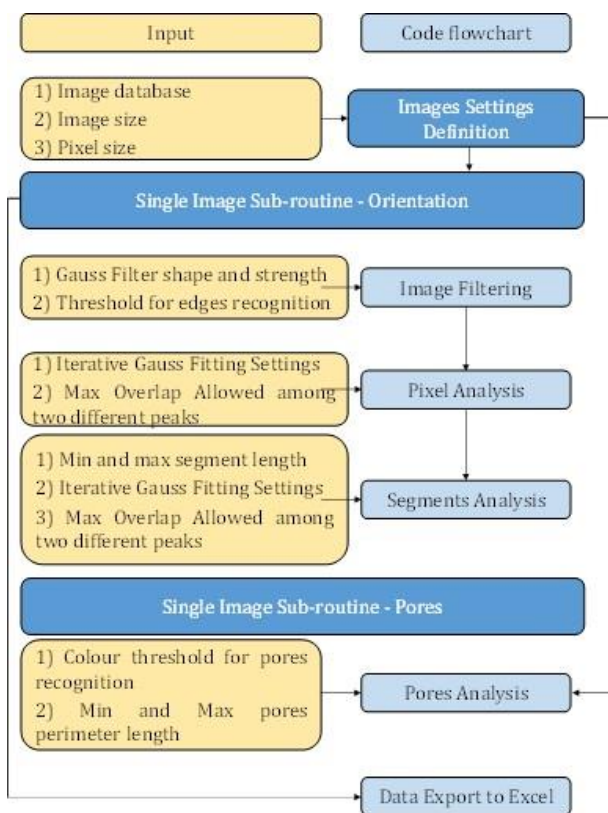


Figure 1. MiCA code flow chart. The yellow column shows the inputs, while the blue column shows the different phases of the image analysis.

This value provides information on the amount of detected edges and it can be used to evaluate the sample density and aggregation level. Then, the Pixel Analysis is performed. For this purpose, the orientation of each pixel recognized as belonging to the particles edges is defined (Sobel, 1968) to create the Orientation Map (OM, Figure 2.e), collected and its distribution plotted in a histogram, which angle interval is user defined ( $\theta$ ). Based on these data, a first set of Full Width at Half Height (FWHM\_% ) values, based on the scatter among the distribution of the particles percentage distribution, is calculated. Specifically, FWHM\_% is calculated as:

$$FWHM\_ \% = 2 * \sqrt{(2 * \log_{10}(2))} * \sigma \quad (1)$$

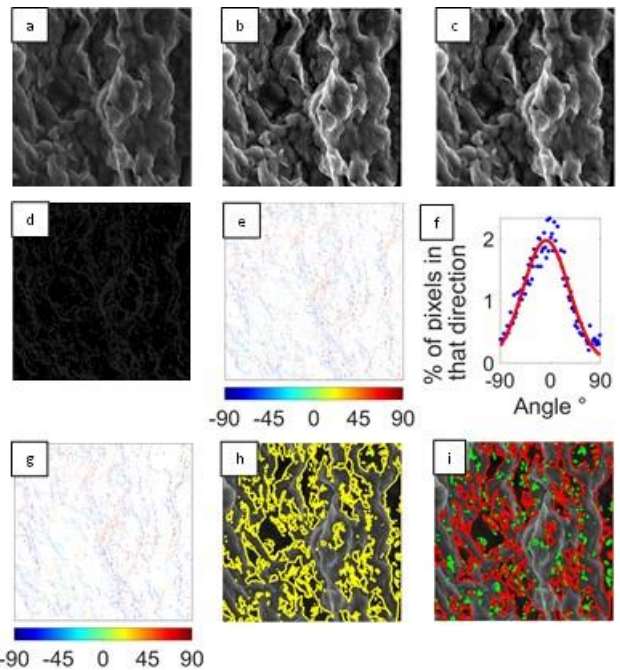


Figure 2. Image Analysis steps. The original image (a) is turned into a black and white image (b) and then smoothed using a Gauss Filter (c). Then, the image edges are recognized (d), the image pixels (e) and segments (g). Orientation Maps are built, together with their histogram of pixels orientation (f). Finally, the porosity of the image is studied (h) and (i).

where  $\sigma$  is the standard deviation of the histogram distribution (Figure 2.f). This is then fitted with a customizable number of Gauss curves ( $G_{max}$ ), until the goal  $R^2$  value, defined by the user, is reached (Figure 2.f). The curves characteristics are then analysed and FWHM of each of them stored as output as FWHM\_  $\theta$ . In order to avoid overfitting, it is possible to apply a Moving Average (MA) to the histogram, in order to smooth the local peaks. At this stage, the number of preferential orientations is found, based on the number of independent Gauss curves fitted. However, due to the presence of background curves, the number of Gauss curves fitted is generally greater than the actual number of particles orientations observed. To avoid data misinterpretations, the curves having a shared area greater than the percentage input (Over%), are excluded from the number of main orientations. Based on these parameters, the Number of Main Orientations ( $N_{mo}$ ) is defined. By combining FWHM\_% and  $N_{mo}$ , it is possible to characterize the degree of orientation of different images, ranking them from the more to the less oriented. Specifically, samples with low  $N_{mo}$  values show a higher orientation degree, while samples with low FWHM\_% values for the same  $N_{mo}$  are obtained for figures for which all the angles have a similar amount of pixels orientated in that direction, so that the image shows a low degree of orientation. The Segment Analysis (Figure 2.g) follows the same steps of the Pixel Analysis, but instead of characterizing the orientation of all the pixels, only those with lengths in the interval between the Maximum and the Minimum Segment Length ( $SL_{max}$  and  $SL_{min}$ ) are selected and analysed. Contrarily to the Pixel Analysis, the Segment Analysis allows the user to exclude cracks and non-clay elements from the image analysis, together with any short elements that were not removed during the Image Filtering phase, responsible for background noise in the histogram analysis (Figure 2.f). The Segment Analysis provides the same outputs as the Pixel Analysis, but based on the pixels laying in the length interval of interest, together with the number of the detected segments and their average length.

Finally, the Porosity of the sample is analysed. First, the black and white image colour and contrast is equalized and a weak Gauss Filter is applied. Then, the image is further filtered with a H-minima transform, in order to find the pores as regional minima, excluding all those which depth is less than the Colour Threshold (T<sub>Po</sub>), and generating a pores map (Figure 2.h). Afterwards, the dark elements are detected, and their shape and extension characterized. As for the Segment Analysis, it is possible to create different pores sizes distributions according to their perimeter length, or to exclude the elements whose number of pixels fall below the Minimum Pore Perimeter length (PP<sub>min</sub>) parameter. Moreover, the pores are divided in two populations, according to whether their perimeter length is larger or smaller than the Maximum Pore Perimeter length (PP<sub>max</sub>) parameter. Therefore, three different sets of output are generated: that referring to all the pores, that with the small porosity and that with the large one. For each of these families, the following information are gathered: void ratio, total and mean pores area and perimeter, together with the main equivalent diameter and the number of detected pores.

### 3 SENSITIVITY STUDY

A sensitivity study showing the effect of different customizable parameters on the final results, is hereby performed. In particular, the effect of GF, TED,  $\theta$ , R<sup>2</sup>, MA, Over%, SL<sub>min</sub>, T<sub>Po</sub>, PP<sub>max</sub> and PP<sub>min</sub> are investigated. The analysis is performed on two micrographs, one obtained from an iso-oriented smectitic clay sample (Figure 3.a), and one from a honeycomb-like kaolinite structure (Figure 3.b). Both the micrographs were obtained at a magnification of 7000x, with a voltage of 15 and a working distance of 7-8 mm.

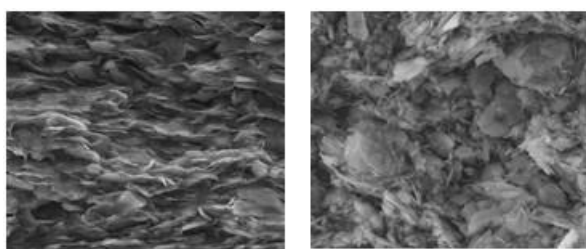


Figure 3. Micrographs of a smectitic iso-oriented (a) and a honeycomb-like kaolinitic (b) structure.

#### 3.1 Gauss Filter (GF)

The Gauss Filter is used to decrease the background noise of the image and smoothen it, therefore, if the value applied is too strong, too much information is averaged and lost, while with a weak GF, non relevant elements would be detected as belonging to the particles edges. GF has two main parameters that can be controlled, namely its size and strength. As a rule of thumb, the size of the square GF, should be equal to the greatest odd number obtained by multiplying the filter strength by 6.

Figure 4 shows the effect of GF on the edge detection, when TED is kept constantly equal to [0.1-0.11]. Specifically, a 49x49 with strength equal to 8 GF was applied to Figures 4.a and b, while a 19x19 with strength equal to 3, to Figures 4.c and d. The analysis was performed on the upper left 500x1000 pixels of Figure 3. The application of a strong GF to Figures 4.a and b resulted in a loss of information and to a consequent misinterpretation of the data. The lines are rounded and isolated, as a consequence of the image averaging performed by the filter and this results in a scattered distribution of the pixels orientation as shown in the histograms below the images. On the contrary,

Figure 4.c and d show a large amount of lines, describing the image in more detail, and allowing the code to better capture the pixel orientation distribution. When using the code for image analysis, it is therefore, recommended to start applying a weak filter, increasing its strength gradually until only the relevant pixels are visible.

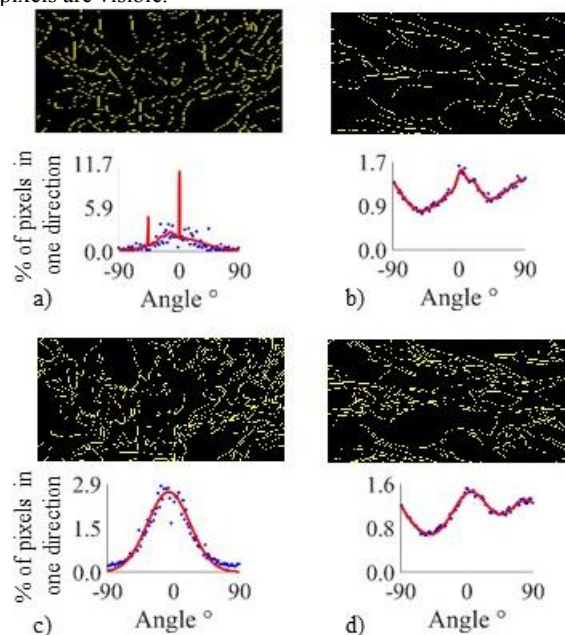


Figure 4 – The effect of GF on the edge detection. 49x49 GF applied to (a) Figure 3a and (b) Figure 3b, 19x19 GF applied to (c) Figure 3a and (d) Figure 3b. The corresponding histograms of pixels orientation is also shown.

#### 3.2 Thresholds for Edges Recognition (TED)

The Thresholds for Edges Recognition (TED) is an interval composed of an upper (UpTED) and a lower (LowTED) value, and it is used to detect the pixels belonging to the clay particle edges. The UpTED and LowTED values range between 1 and 0, specifically, when the UpTED is equal to 1, pixels are not detected as belonging to the edges, while when it is equal to 0, all of them do.

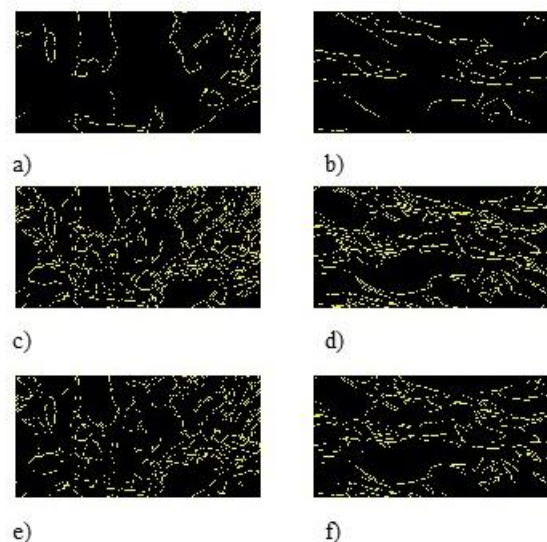


Figure 5 – The effect of TED on the edge detection on Figures 3a and b, with TED [0.1-0.5] in (a) and (b), [0.1-0.25] in (c) and (d) and [0.2-0.25] in (e) and (f).

Figure 5 shows the effect of TED on the edges detection when the GF size is kept constantly equal to 19x19 on the micrographs of Figure 3. Three threshold intervals were used, [0.1-0.5] (Figures 5.a-b), [0.1-0.25] (Figures 5.c-d) and [0.2-0.25] (Figures 5.e-f). Again the analysis was performed on the upper left 500x1000 pixels. Figures 5.a-b show a low amount of lines, insufficient to describe adequately the clay pattern, due to the high UpTED. This value is decreased to 0.25 for Figures 5.c-d, obtaining a better edge characterization of the analysed micrographs. Finally, in Figures 5.e-f LowEd was increased from 0.1 to 0.2, resulting in a significant loss of information, but still allowing for a meaningful analysis. As for GF, also TED influences the edges definition. However, GF averages the surrounding areas, and therefore not only it decreases the amount of detected lines, but it also deforms them, something that does not happen in Figure 5. When modifying TED, the amount of pixels selected changes but not their position. Therefore, if the user aims to reduce the background noise, the use of a higher LowTED should be preferred to the application of a strong GF.

### 3.3 The Iterative Gauss Fitting Parameters

In order to find the number of main orientations in a micrograph, it is necessary to find the number of peaks contained in its histogram distribution. However, it is not sufficient to find the number of maximum values within the distribution, due to the high number of scattered, single points. Therefore, an iterative Gauss curves fitting is performed. The code allows the user to customize different parameters, namely the histogram interval  $\theta$ ,  $G_{max}$ ,  $R^2$ , MA, Over%, which effect on the final orientation pattern is hereby discussed except for the effects of  $G_{max}$  that are self explanatory.

#### 3.3.1 Theta

When generating the histogram describing the pixels orientation distribution (Figure 2.f), the user must set the angle interval ( $\theta$ ). This parameter strongly influences the Gauss curves fitting process, as shown in Figure 6. A  $\theta$  of 0.5, 2 and 10 was used, respectively, on Figure 3.a for Figures 6.a-c and on Figure 3.b for Figures 6.d-f. The analysis was performed with a GF mask of 19x19, and TED of [0.1-0.25].

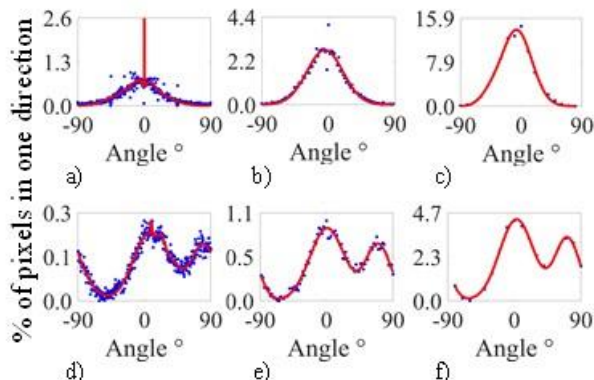


Figure 6 – The effect of  $\theta$  on the edge detection. A  $\theta$  of 0.5, 2 and 10 was used, respectively, on Figure 3.a for a-c and on Figure 3.b for d-f.

Figure 6 indicates that small  $\theta$  values (Figures 6.a and d) show a poor data fitting together with a small y variation, leading to a reduced set of FWHM\_% values. For the purpose of image comparison, it is important to have a good sensitivity of this parameter, therefore, a greater scatter on the y axis is recommended. However, using large  $\theta$  values (Figures 6.c and f) leads to a reduced amount of points that lowers the quality of the FWHM\_% value. Figures 6.b and e, which show the distributions obtained for  $\theta=2$ , offer both y scatter greater than the unit and a large amount of data points (90). It is therefore recommended to

start from a low  $\theta$  and increase it until at least a unit value is obtained for FWHM\_%.

#### 3.3.2 R<sup>2</sup>

The coefficient of determination  $R^2$  defines how well the data are fitted to a regression line; the closer this value is to 1, the better is the fitting. The custom-made image analysis code is structured so that, starting from one Gauss curve, more are added to the pattern in order to reach the goal  $R^2$  value. Figures 7.a-c were obtained applying, respectively, a  $R^2$  value of 0.95, 0.97 and 0.99; while Figures 7.d-f were obtained with  $R^2$  values of 0.5, 0.97 and 0.99. The graphs obtained analysing Figure 3.b having  $R^2$  values between 0.5 and 0.97 are not reported here, as no significant variation compared to Figure 7.e were observed. The analysis is performed with a GF mask of 19x19, Ted=[0.1-0.25],  $G_{max}=8$  and  $\theta=2$ . By looking at Figures 7.a-c it is evident that using higher  $R^2$  values does not necessarily lead to better quality fitting. The presence of the two local minimum and maximum at about 0 degrees forces the code to fit a higher number of curves, in order to reach  $R^2=0.97$  and 0.99. However, the fitting showed in Figure 7.a is already more than sufficient to describe the iso orientation of the sample. For Figure 3.b,  $R^2=0.5$  is sufficient to give a satisfactory fitting of the curves, while an overfitting similar to that observed for Figure 7.c was found for  $R^2$  values greater than 0.98. Also in this case a maximum value of 0.95 is recommended.

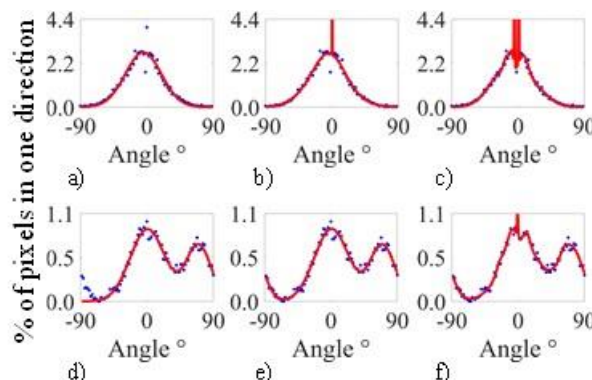


Figure 7 -  $R^2$  value of 0.95, 0.97 and 0.99 applied on Figure 3.a is shown in (a)-(c), respectively; while  $R^2$  values of 0.5, 0.97 and 0.99 applied on Figure 3.b is shown in (d)-(f), respectively.

#### 3.3.3 Moving Average

The third iterative Gauss Parameter is MA. As the presence of local maxima can affect the quality of the final fitting, it is possible to apply a smoothing parameter by performing a Moving Average on the histogram distribution, before fitting it with the Gauss curves. If averaging is not necessary, a unit MA value can be inserted. Figures 8 a-f were obtained by applying MA=1, 2 and 3, respectively, to Figures 3.a and b. The analysis was performed with a GF mask of 19x19, Ted=[0.1-0.25],  $G_{max}=8$ ,  $\theta=2$  and  $R^2=0.99$ . As expected, increasing MA shows results similar to that obtained by lowering  $R^2$  or  $\theta$ . However, it should be considered that one of the code main outcomes is FWHM\_% and that reducing the number of points used for its definition, the sensitivity of this parameter drops consequently. Therefore, using MA>1 or  $R^2<0.97$  instead of a greater  $\theta$  value, allows for a higher precision in the data fitting, without losing the information necessary for a good diversification of FWHM\_%.

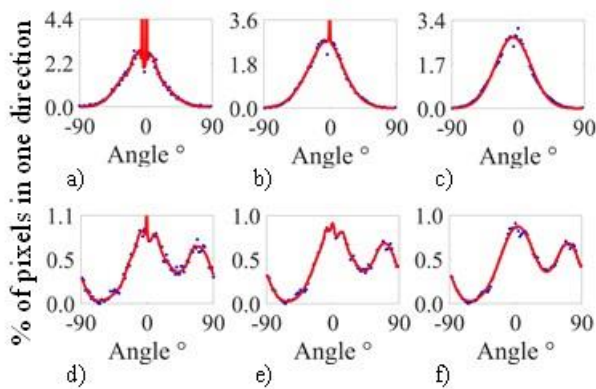


Figure 8 – Effect of moving average. MA=1, 2 and 3 were applied to Figures 3.a and b, and shown in (a) – (f) respectively.

### 3.3.4 Gauss Curves Overlap

Finally, Over% is discussed, based on the results of Figure 7. By setting Over%=40%, the plots of Figures 7.a and b show the presence of a unique main orientation detected, while 2 are found for Figure 7.c. Moreover, for Figures 7.d-f  $N_{mo}=2, 2$  and 3, respectively, was obtained. It may be noted that even though 3 Gauss curves were fitted, the detected curve peaking at values smaller than  $-90^\circ$ , is the same culminating at about  $70^\circ$ , so that the two curves represent the same main orientation, and thus only one of them should be considered for the  $N_{mo}$  counting. On the contrary, the presence of two curves peaking at about  $0^\circ$  in Figure 7.f causes the detection of 3 different orientations. If a greater value of Over% was used, only one of the two directions would be detected.

### 3.4 Minimum and Maximum Segment Length

The effect of the Minimum and Maximum Segment Length ( $SL_{min}$  and  $SL_{max}$ ) is hereby discussed. Figures 9.a-d were obtained by applying  $SL_{min}=3 \mu m$  and  $SL_{max}=10 \mu m$  on Figures 3.a and b, respectively. The analysis was performed with a GF mask of  $19 \times 19$ ,  $Ted=[0.1-0.11]$ ,  $G_{max}=8$ ,  $\theta=2$  and  $R^2=0.90$  and  $MA=1$ .

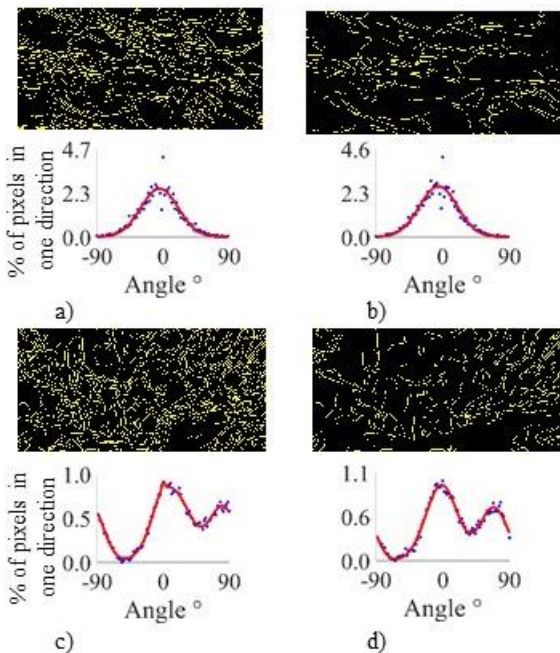


Figure 9 – Comparison of the Pixel Analysis (first column) and the Segment Analysis (second column) results, (a) and (b) were obtained from Figure 3.a, while (c) and (d) from Figure 3.b.

In this case,  $SL_{min}$  and  $SL_{max}$  were chosen to show visible differences in the Edge maps and not to isolate particular elements in the images. In this case, the removal of a selected set of edges does not alter the orientation patterns, as the removed lines did not represent extra elements or cracks, otherwise altering the order pattern. If the user aims to remove external elements, it is recommended to first run a segment analysis with a  $SL_{min}$  value equal to  $\frac{1}{4}$  the average expected length and no  $SL_{max}$ . The obtained average segment length should then be multiplied by 5, and this value used as a reference  $SL_{max}$ .

### 3.5 Pores Threshold

The Pores Threshold (TPo) is the most important parameter for the pores detection. A greater TPo value increases the area of the detected pores, as the local minima are disregarded, and the regional minima basin is extended. This effect is evident in Figure 10. The different graphs are obtained by applying a TPo of 5, 15, 30 and 50 to Figure 3.b, obtaining, respectively, Figures 11.a-d. As forecasted, Figure 10.a shows a high number of small, scattered pores. These progressively disappear or aggregate into bigger pores increasing the TPo value, so that Figure 10.d shows a small number of large pores. It is evident that the use of 5 (Figure 10.a) gives a description of the porosity that does not match that expected, as the larger pores are disregarded. Increasing the value to 15 (Figure 10.b), the pore detection is improved, however several large pores are still ignored. Figure 11.c seems to detect adequately most of the pores, while the small porosity is completely lost for Figure 11.d. The choice of the appropriate TPo value should be based on a visual inspection, with a process similar to that performed for the Thresholds for Edges Recognition.

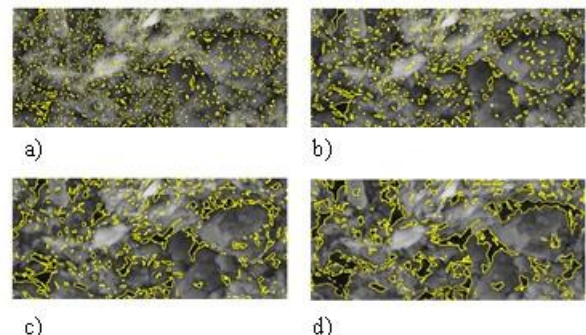


Figure 10 – Effect of TPo in the pores detection. A TPo of 5, 15, 30 and 50 are applied to Figure 3.b, generating (a) - (d).

Also in this case, it is recommended to perform a sensitivity study on a reference clay image, increasing TPo until the pores are satisfactorily recognized. Afterwards, the TPo value selected should be doubled, and progressively decreased to find the optimal value. As the choice of large TPo values leads to a decrease of the amount of the small pores, the same image can be analysed twice, with a high and a low value of the parameter.

### 3.6 Minimum and Maximum Pores Length

As for the Segment analysis, in the Pores Analysis it is possible to differentiate the detected elements according to their length. Therefore, by changing the Maximum and the Minimum Pores Length, the grouping families are altered as well. Figure 11 shows the effect of  $PP_{min}$  and  $PP_{max}$  on Figure 3.b, with Figure 11.a and b having a  $PP_{min}$  of 5 and 100 pixels and a  $PP_{max}$  of 200; and Figures 10.c and d having a  $PP_{min}$  of 5 and a  $PP_{max}$  of 50 and 300. Specifically, the green elements are those which perimeter is smaller than  $PP_{max}$ , while the red elements are the large perimeter pores. Increasing  $PP_{min}$  a reduced amount of pores are detected, so that all the micro-pores of Figure 11.a are not detected in Figure 11.b, and by means of  $PP_{min}$ , the meso-pores

are divided into a small and a large population. By changing  $PP_{max}$ , Figures 11.c and d are generated. If  $PP_{min}$  selects the relevant pores,  $PP_{max}$  differentiates the small and the large elements so that they can be characterized separately. By looking at the code outcome, the small pores area of Figure 11.d is 5.2 times greater than that of Figure 11.a, while the detected Equivalent diameter is 90% smaller. The choice of these parameters solely depends on the porosity targeted by the user. However, it is strongly recommended the use of a  $PP_{min}$  greater than 3 pixels, in order to avoid the detection of single points linked to the image noise.

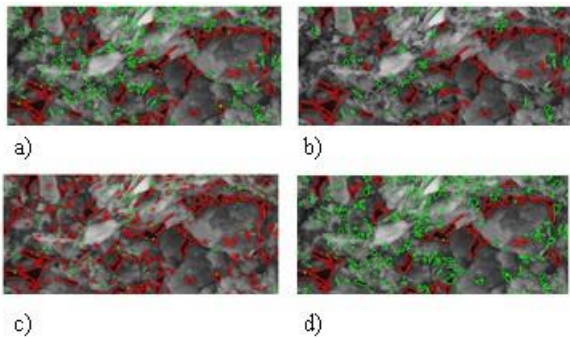


Figure 11 – Effect of  $PP_{min}$  and  $PP_{max}$  on the pore detection and on the population grouping of Figure 3.b.  $PP_{min}=5$  and 100, respectively, with  $PP_{max}=200$  was applied to Figure 11.a and b, while for Figures 11.c and d  $PP_{min}=5$ , and  $PP_{max}=50$  and 300, respectively.

#### 4 CONCLUSIONS

The clay samples mechanical behaviour is expected to be strongly influenced by the particles orientation and the porosity characteristics. Due to the lack of modern, automated tools designed to quantify these features, an algorithm based on the latest Image Analysis advances was developed. The choice of the input parameters strongly influence the final results, therefore the need for a systematic sensitivity study is evident. The single inputs variability and effect were evaluated on the micrograph of an iso-oriented, dispersed smectitic sample and on a multi-directional, aggregated kaolinite sample. Specific instructions were given on how to choose the appropriate values for the parameters, depending on the micrographs analysed. However, for first time users, it is recommended to run a first simulation with the parameters obtained from the sensitivity study performed for this work, and then adjust all the parameters to fit the user needs, following the order outlined here.

Therefore, the first Orientation analysis should be performed with  $GF=19 \times 19$  having a strength 3,  $Ted=[0.25-0.10]$ ,  $G_{max}=8$ ,  $\theta=2$ ,  $R^2=0.95$ ,  $MA=3$ ,  $Over\%=40$  and  $SL_{min}=1/4$  the average expected length,  $SL_{max}=10^5$ . Alternatively,  $GF$  can be set with a unity strength.  $SL_{max}$  should be defined once all the other parameters are defined, and its initial value should be set equal to 5 times the average length value obtained from the Segment Analysis. Moving to the pores analysis,  $PP_{min}$  and  $PP_{max}$  should be defined considering the porosity interval of interest. Regarding  $TPO$ , an initial value of 5 should be used, and it should be increased until all the significant large pores are detected. Afterwards the value should be doubled, and slowly decreased, to incorporate the significant small pores.

It should be taken into consideration that except for some indication on the minimum recommended values, only the strategy for the parameter choice is recommended here, as the algorithm was designed to give the user the greatest possible flexibility.

#### 5 REFERENCES

- Burland, J. B. (1990). On the compressibility and shear strength of natural clays. *Géotechnique* 40(3), 329-378.
- Cotecchia, F., & Chandler, J. (1997). The influence of structure on the pre-failure behaviour of a natural clay. *Geotechnique* 47(3), 523-544.
- Di Remigio, G., Rocchi, I., & Zania, V. (2021). New method for SEM-based quantitative Microstructural Clay Analysis - MiCA. *Applied Clay Science* 214, 106248.
- Fearon, R. E., & Coop, M. R. (2000). Reconstitution: what makes an appropriate reference material? *Geotechnique* 50(4), 471-477.
- Hattab, M., Bouziri-Adrouche, S., & Fleureau, J. (2010). Evolution of microtexture in a kaolinitic matrix on a axisymmetric triaxial path. *Canadian Geotechnical Journal* 47(1), 34-48.
- Martinez-Nistal, A., Veniale, F., Setti, M., & Cotecchia, F. (1999). A scanning electron microscopy image processing method for quantifying fabric orientation of clay geomaterials. *Applied Clay Science* 14(4), 235-243.
- Skempton, A. W. (1969). The consolidation of clays by gravitational compaction. *Quarterly Journal of the Geological Society* 125, 373-411.
- Sobel, I. (1968). A 3x3 Isotropic Gradient Operator For Image Processing. *Stanford Artificial Intelligence Project*.