An Algorithm for the Pore Size Determination Using Digital Image Analysis

Małgorzata Charytanowicz[1,2]

[1] Polish Academy of Sciences,
Newelska 6, PL 01-447 Warsaw, Poland
mchmat@kul.lublin.pl
[2] The John Paul II Catholic University of Lublin,
Konstantynów 1 H, PL 20-708 Lublin, Poland

Abstract. The study is concerned with putting forth a novel method of determination the pore size and its distribution for pores of different shapes. The identification of irregular and branched pores is associated with difficulties in their separation, as well as in the quantification of their size and shape characteristics. Recent developments in digital image processing provides a relatively new technology that allows visualization of the internal structure of objects. Our research was conducted using computer-aided tomography, and in turn, statistical techniques. This seems to be very useful in characterizing the pore volume distribution and in quantifying the differences in pore structures of the investigated materials. In this paper, the methodology is illustrated with a number of soil aggregates which differ in terms of soil fertilization. The approach is universal, and can be successfully applied for many tasks in data mining where pore characteristics are needed.

Keywords: image processing, cumulative distribution function, pore size distribution, pore space, total porosity, statistical methods, soil aggregation.

1 Introduction

The characterization of porous material has been an important topic in the research area. This includes the total porosity, pore size distribution, pore size and shape or interconnectivity [1, 2]. Pore shape is mainly unknown, but it could be approximated by one of the basic pore models: "cylindrical" – with circular cross-section, "ink-bottle" – having a narrow neck and wide body, and "slit-shaped" – with parallel plates. Unfortunately, pore structures do not have any regular or well defined shape, and therefore the use of methods assuming their shape becomes inapplicable. For this reason, some authors have proposed different computational methods to determine the pore size distribution. These techniques can be classified based on the pore size, pore type, amount of pores or nature of measurements. Despite the usefulness of these methods for determining pore size distribution, a combination of different techniques is needed to

E. Pietka, J. Kawa, and W. Wieclawek (eds.), *Information Technologies in Biomedicine, Volume 3*, Advances in Intelligent Systems and Computing 283,
DOI: 10.1007/978-3-319-06593-9_20,

cover the wide range of characteristics describing the pore structure. However, conventional techniques, such as mercury porosimetry, or low-temperature gas adsorption have often been unsatisfactory due to inadequate pore characteristics. Moreover, they are limited in applicably as far as pore size and shape or preferential flow is concerned.

Computed tomography and digital image processing provide new possibilities of identifying pores and quantifying their characteristics. The non-destructive nature of computed tomography scanning allows the same object to be scanned multiple times and provides an opportunity to investigate its particle at any location within a sample. In the last few years, studies indicate that X-ray computer-aided tomography also provides an alternate approach for the non-destructive observing, measuring, and quantifying of the internal microstructure of materials [3, 4]. In our research, the use of image analysis has enabled us to determine pore size distribution.

Classical parametric techniques of density estimation assume that the data can be drawn from one of a known parametric family of distributions, and determined by a few parameters, for example, mean and variance. The density underlying the data could then be estimated by finding the parameters using the data and substituting these estimates into the formula for the chosen density. Such technique requires performing significant tests, such as Chi-square goodness-of-fit, Kolmogorov-Smirnov or Shapiro-Wilk test [5, 6]. The null hypothesis states that our data must follow a specific distribution. A not significant result denoting that the tested hypothesis is not rejected allows us to use the estimated model. Moreover, for skewed distributions, a mathematical transformation, for example, logarithmic, tending the data to normal distribution, is recommended. On the other hand, nonparametric estimation [7–9] assumes that there is no pre-specified functional form for a density function. Some of these methods also have the advantages of being very intuitive and relatively simple to analyze mathematically. Moreover, it is worth noticing that all parameters appearing in the obtained model can be effectively calculated using convenient numerical procedures based on optimizing criteria [10, 11].

In this paper, research was conducted to determine porosity and soil pore size distribution. Pore size distribution is one of the many physical measurements characterizing soil structure as far as plant growth is concerned. Additionally, total porosity, defined by the ratio of the volume of void-space, and the total volume of soil material (including the solid and void components), provides a more useful physical description of a particular soil, such as compaction or the maximum space available for water. A number of scientists have reported studies of employing pore space as a general method for defining soil properties [12]. In this respect, a complete analysis of the soil pore size distribution is used for predicting the gas diffusivity, water conductivity and water availability to plants. The investigated material, suitable for collecting information about the soil porosity, was preprocessed by aggregate preparation, image analysis and statistical methods. Background and information regarding these operations are described in Section 2 and 3. The two packages that were used to conduct and report research are the Aphelion image analysis package and Statistica software.

2 Materials and Methods

The experiments were conducted on three soil samples differ in terms of fertilization (pig manure, mineral fertilization, control group), and were explored at the Polish Academy of Sciences in Lublin. The investigated material was sampled from the cultivated soil layer, classified as silty loam (Word References Base for Soil Resources – Mollic Gleysols). The proportion of each particle size group in the soil was as follows: sand – 46%, silt – 28%, clay – 26%, pH was: H_2O – 5.9, KCl – 5.4. A long-term fertilization trial was executed on the experimental fields. The manure treatment was 80 ton per ha of composed pig manure. The mineral fertilization was according to plant needs. The control was applied neither mineral nor manure treatment, there were plant residues only. The adopted crop rotation was as follows: a cycle of potato – barley – rye from 1955 to 1989, and sugar – beet – barley – rape – wheat from 1990. Aggregate soil organic matter was measured by the Multi N/C 3100 Autoanalyser (Analitic Jena, Germany). The total organic carbon and total nitrogen contents for three fertilizations (pig manure, mineral fertilization, control) were respectively: 21.50, 14.89, 13.54 g/kg, and 2.10, 1.51, 1.35 g/kg. The total organic carbon shows the same tendency as total nitrogen, i.e. they increase in the same order: the lowest – control, middle – mineral fertilization, the highest – pig manure.

2.1 Soil Sample Preparation

The soil samples used for porosity analysis were air dried in room conditions, divided into smaller quantities, and gently sieved through 2 and 10 mm sieves. The remaining at 2 mm sieve soil aggregates of sizes ranging from 2 to 10 mm, were then detected by means of X-ray computational tomography, using a GE Nanotom S device with a molybdenum X-ray source, 230 μA cathode current, 60 kV voltage and voxel-resolution of 2.5 microns per volume pixel. Three 2D sections of 3D objects, uniformly located within them, were performed to characterize the aggregate structure. After this, tomography sections were processed using the Aphelion 4.0.1 package. Pore size measurements of the aggregate sample were collected by means of image processing techniques [13, 15, 16].

2.2 Image Processing and Data Analysis

The entire procedure was composed of the following steps. Firstly, grayscale tomography sections were preprocessed by removing ring artifacts using the ROI (region of interest) method. Next, the automatic Otsu binarization method was employed. After these preprocessing methods were undertaken, a binary morphological closing with increasing size of structuring element SE (starting with 2×2 square SE), was processed subsequently. Operation closing, consisting of a dilation followed by erosion, was used to fill in holes and small gaps without changing the aggregate size and original boundary shape. In each step, the source image was subtracted from the target image, and the result volumes were listed, giving soil aggregate pore distribution. The operation was repeated until all pores were filled. A direct subtraction of two images: the transformed image and the original image, gives the total pore volume in the soil sample. A cumulative aggregate porosity $P(i)$ related to a given and smaller structuring element size can be calculated as:

$$P(i) = \frac{S(i) - S(0)}{S(N)} \text{ for } i = 2, 3, \ldots, N, \tag{1}$$

where:

$S(i)$ – surface area of solids and filled by pores less than or equals to SE size i, $i \geq 2$,

$S(0)$ – surface area of solid phase,

$S(N)$ – total area, equals to surface area of solids and pores,

N – size related to the biggest pore,

$P(N)$ – total porosity.

Next, the soil pore size distribution can be calculated as:

$$\frac{1}{P(N)} \frac{dP}{dr} = \frac{P(i+1) - P(i)}{P(N)} \text{ for } i = 2, 3, \ldots, N-1, \tag{2}$$

where $P(N)$ denotes the total porosity.

Figure 1 shows subsequently obtained aggregate section images related to the described method.

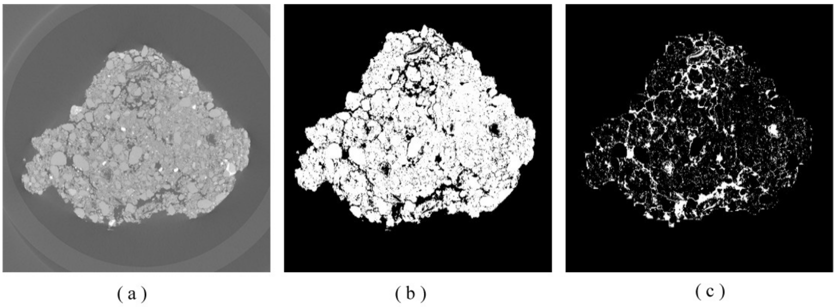

Fig. 1. Cross section of a typical soil with pore space in black (a), a binary soil aggregate section image done by ROI and Otsu methods (b), a binary image of all pores detected by closing operation (c)

The data derived automatically from images was statistically examined using classical methods. For each fertilization (pig manure, mineral fertilization, control), a discrete data set consists of cumulative porosities $P(i)$, $i = 2, 3, \ldots, N$

was explored. Our aim was to present the frequency distribution as a continuous mathematical equation instead of a discrete set of data. We have fitted to the observed data, a known theoretical model of probability distribution function.

Finally, the elaborated procedure for determining the porosity and soil pore size distribution may be explained in the following steps:

The Procedure for Determining the Porosity and Soil Pore Size Distribution

1. Soil sample preparation.
2. X-ray microtomography of soil aggregates.
3. Digital image processing.
   (a) ROI method.
   (b) Otsu thresholding method.
   (c) Binary morphological closing.
4. Frequency analysis.

3 Results

A cumulative porosity of investigated soil aggregates was determined using the image processing methods described in paragraph 2. Figure 2 shows exemplary soil aggregate section images done by ROI and Otsu methods for each type of aggregate.

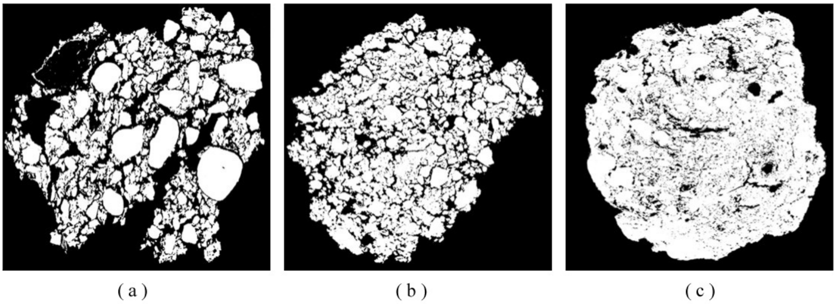

Fig. 2. Exemplary binary images of investigated aggregates: pig manure (a), mineral fertilization (b), control (c)

The total porosity of the investigated aggregates for three fertilizations (pig manure, mineral fertilization, control) calculated as the average of three sections, were equal to 33.28%, 22.95%, 14.2% respectively. The total porosity shows the same tendency as total organic carbon and total nitrogen, i.e. they increase in the same order: the lowest – control, middle – mineral fertilization, the highest – pig manure.

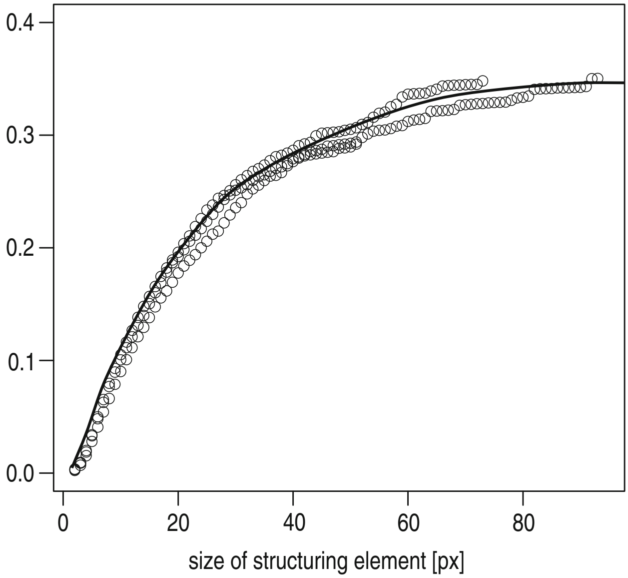

Fig. 3. Cumulative porosity of aggregates treated with pig manure (circle – observed data, line – estimated curve)

Experimental points (x_i, y_i), $i = 2, 3, \ldots, N$ where x_i denotes the size of structuring element (1px = 2.5 microns), and y_i was calculated using formula (1) and denotes cumulative porosity, are shown in Figure 3 – Figure 5.

The function given by the formula:

$$Y = k \cdot ILognorm(x, m, s) \tag{3}$$

where Y denotes the cumulative porosity, x means size of structuring element, k, m, and s constitute the parameters of the models, and $ILognorm(x, m, s)$ denotes the lognormal cumulative distribution function, was fitted to the observed data using nonlinear estimation [14]. The function was selected by means of the minimum square error criterion.

Table 1 shows the parameter values of the model given by the rule (3) and the corresponding coefficients of determination R^2 for the investigated aggregates. The usefulness of such a model is very high, R^2 is shown to be greater than 90% for all the types of fertilization sampled.

The parameter k corresponds to the total porosity of the aggregate, while the larger parameter m corresponds to the higher total porosity of the aggregate.

Pores have traditionally been divided into micropores, mesopores, and macropores, with the division between them being arbitrary. The use of the estimated models allowed us to calculate the cumulative porosity for the arbitrarily given

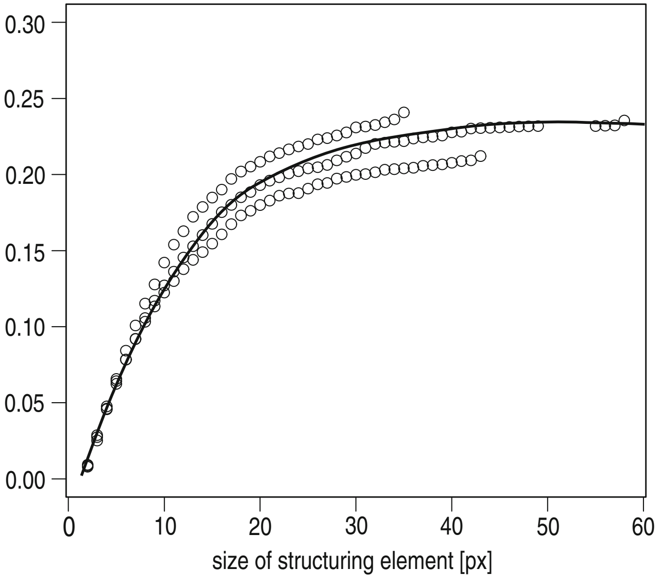

Fig. 4. Cumulative porosity of aggregates treated with mineral fertilization (circle – observed data, line – estimated curve)

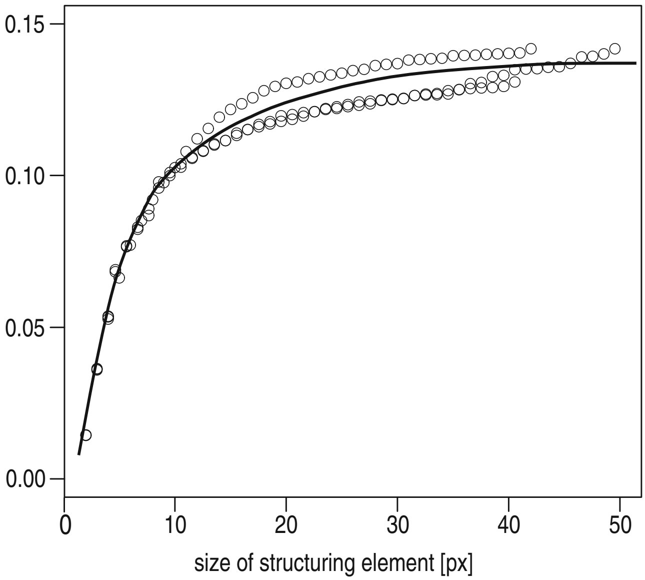

Fig. 5. Cumulative porosity of the control group (circle – observed data, line – estimated curve)

Table 1. The model parameters and the coefficient of determination obtained for the model $Y = k \cdot ILognorm(x, m, s)$

Fertilization	k	m	s	R^2
Pig manure	0.37	2.79	0.89	0.99
Mineral fertilization	0.23	2.16	0.89	0.98
Control	0.14	1.67	1.03	0.95

limits between micropores, mesopores and macropores, equal to 30 microns and 75 microns respectively. This enables to calculate the fractions of micro-, meso-, and macropores for each type of fertilization. The results are given in Table 2.

Table 2. Fractions of micro-, meso- and macropores

Fertilization	Fraction		
	of micropores	of mesopores	of macropores
Pig manure	0.12	0.13	0.09
Mineral fertilization	0.14	0.08	0.02
Control	0.12	0.02	0.001

As can be seen, the largest fraction of mesopores occurs in the soil fertilized with manure, this fraction represents 13% of the total aggregate area. In addition, this sample shows 12% of the total aggregate area to be micropores and 9% to be macropores. This creates the most favorable conditions for plant growth.

Within the aggregate, the second largest fraction of mesopores, equaling 8%, was observed in the soil with mineral fertilization. This soil has a small amount of macropores, equaling 2%, and a quite large amount of micropores, equaling 14%. This creates less favorable conditions for plant growth.

The soil without fertilization (control group) reveals that 12% of the total aggregate area are micropores, 2% are mesopores and 0.1% are macropores. The largest fraction of micropores in comparison with the fractions of meso- and masropores creates the least favorable conditions for plant growth

Next the methodology described above was used to determine the soil pore size distribution. Experimental points (x_i, y_i), $i = 2, 3, \ldots, N$ were done, where x_i denotes the size of structuring element, and y_i was calculated using formula (2). The model fitted takes the form:

$$Y = Lognorm(x, m, s) \tag{4}$$

where

$$Lognorm(x, m, s) = \frac{1}{x\sqrt{2\pi}s} \exp\left(-\frac{(\log(x) - m)^2}{2s^2}\right) \tag{5}$$

is the probability density function of the lognormal distribution with parameters m and s, Y denotes the pore size fraction to be modeled, and positive variable x means the size of structuring element (1px = 2.5 microns).

The results are shown graphically in Figure 6 – Figure 8.

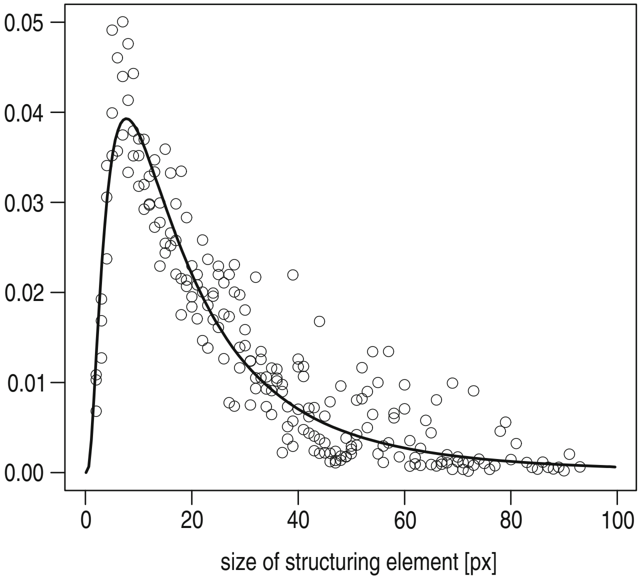

Fig. 6. Frequency distribution of aggregates treated with pig manure (circle – observed data, line – estimated curve)

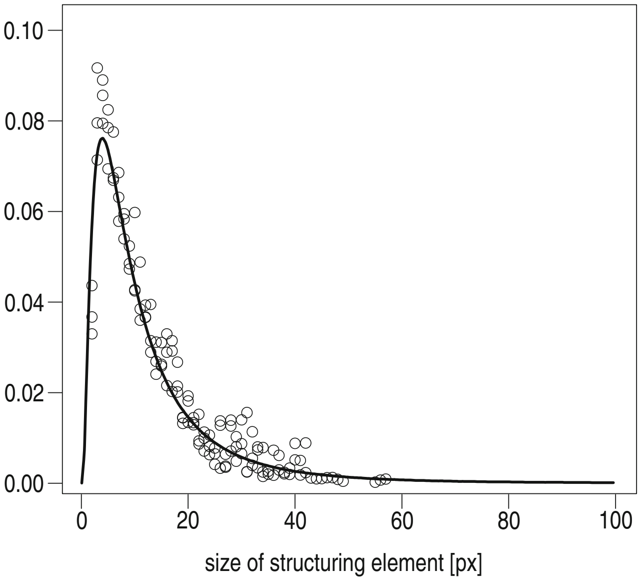

Fig. 7. Frequency distribution of aggregates treated with mineral fertilization (circle – observed data, line – estimated curve)

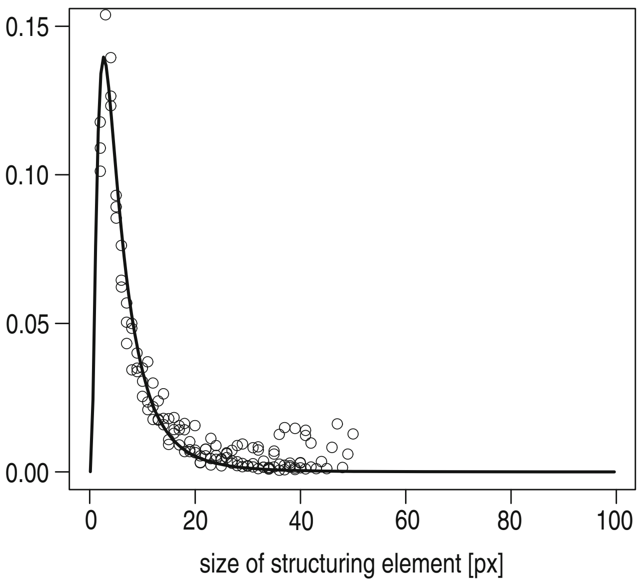

Fig. 8. Frequency distribution of the control group (circle – observed data, line – estimated curve)

Table 3 shows the parameter values of the model (4) and the corresponding coefficients of determination R^2 for the investigated aggregates.

Table 3. The model parameters and the coefficient of determination obtained for the model $Y = Lognorm(x, m, s)$

Fertilization	m	s	R^2
Pig manure	2.83	0.89	0.90
Mineral fertilization	2.17	0.83	0.95
Control	1.59	0.79	0.94

The usefulness of the model is demonstrated to be very high, R^2 is greater than 90% for all the types of sampled fertilization. This indicates that the model has a very good fit to the data. The larger parameter m corresponds to the higher total porosity of the aggregate. The values of parameter m range from 1.59 to 2.83, whilst the values of parameter s are smaller than 1.0 and range from 0.79 to 0.89. As having the same role, the values of these parameters are mostly the same as the corresponding values obtained for the model characterizing the porosity. However, some differences may arise due to the smaller dispersion of the data resulting better fitting the model to cumulative porosity.

The obtained function, characterizing the pore size distribution, reveals their positively skewed and unimodal character. The pore size distribution for pig manure fertilization evidences a high amount of large pores, whilst the pore size distribution for the control group indicates the presence of a high amount of small pores.

4 Summary

Recently, a notable increase in research in the area of image processing methods, has shown that, together with computed tomography, these techniques provide non-destructive tools for studying the internal structures of objects. This seems very useful in exploring soil aggregates and quantifying their characteristics.

In this paper, a systematic procedure was outlined for using X-ray tomography and frequency analysis to quantify measurements of void ratio, porosity, and pore size distribution. The lognormal probability density function was crucial for description of pore size distributions. The presented algorithm is expected to be more objective than classical methods, where arbitrary assumptions concerning the shape of pores are required.

This approach is also motivated by the current rapid growth in computational power. Hence, improved real-time data processing and algorithm efficiency have importance due to the concurrent increase in the quantity and complexity of the data that are being collected.

References

1. Czachor, H., Lipiec, J.: The effects of manure application on soil aggregation. International Agrophysics 18, 217–223 (2004)
2. Dexter, A.R.: Advances in characterization of soil structure. Soil and Tillage Research 11, 198–238 (1988)
3. Papadopoulos, A., Bird, N.R., Whitmore, A.P., Mooney, S.J.: Investigating the effects of organic and conventional management on soil aggregate stability using X-ray computed tomography. European Journal of Soil Science 60, 360–388 (2009)
4. Peth, S., Nellesen, J., Fischer, G., Horn, R.: Non-invasive 3D analysis of local soil deformation under mechanical and hydraulic stresses by uCT and digital image correlation. Soil and Tillage Research 111, 3–18 (2010)
5. Nowak, P., Romaniuk, M.: Pricing and simulations of catastrophe bonds. Insurance: Mathematics and Economics 52, 18–28 (2013)
6. Swanepoel, C.J., Doku, W.O.: New goodness-of-fit tests for the error distribution of autoregressive time-series models. Computational Statistics and Data Analysis 43, 333–340 (2003)
7. Charytanowicz, M., Kulczycki, P.: Nonparametric Regression for Analyzing Correlation between Medical Parameters. In: Pietka, E., Kawa, J. (eds.) Information Technologies in Biomedicine. ASC, vol. 47, pp. 437–444. Springer, Heidelberg (2008)
8. Kulczycki, P., Charytanowicz, M.: Conditional Parameter Identification with Different Losses of Under- and Overestimation. Applied Mathematical Modelling 37, 2166–2177 (2013)
9. Silverman, B.W.: Density Estimation for Statistics and Data Analysis. Chapman and Hall, London (1986)

10. Lange, K.: Numerical Analysis for Statisticians. Statistics and Computing. Springer, New York (2000)
11. Kincaid, D., Cheney, W.: Numerical Analysis. Brooks/Cole, Pacific Grove (2002)
12. Zdravkov, B., Cermak, J., Sefara, M., Janku, J.: Pore classification in the characterization of porous materials: A perspective. Central European Journal of Chemistry 5, 385–395 (2007)
13. Król, A., Niewczas, J., Charytanowicz, M., Gonet, S., Lichner, L., Czachor, H., Lamorski, K.: Water-stable and non-stable soil aggregates and their pore size distributions. In: 20th International Poster Day and Institute of Hydrology Open Day: Transport of Water, Chemicals and Energy in the Soil – Plant – Atmosphere System, Bratislava, pp. 870–871 (2012)
14. Draper, N.R., Smith, H.: Applied regression analysis. John Wiley and Sons, New York (1981)
15. Pratt, W.K.: Digital Image Processing. John Wiley and Sons, New York (2001)
16. Aphelion 4.0.10 user guide: ADCIS S.A. (1997)