MOSTAFA RASTGOU*, HOSSEIN BAYAT*,
MUHARRAM MANSOORIZADEH**

FITTING SOIL PARTICLE-SIZE DISTRIBUTION (PSD) MODELS BY PSD CURVE FITTING SOFTWARE

Abstract. This paper describes a particle-size distribution (PSD) curve fitting software for analyzing the soil PSD and soil physical properties. A better characterization of soil texture can be obtained by describing the soil PSD using mathematical models. The mathematical equations of soil PSD are mainly used as a basis to estimate the soil hydraulic properties. Until now, many attempts are made to represent PSD curves using mathematical models, but selecting the best PSD model requires fitting all models to the PSD data, which would be difficult and time-consuming. So far, no specific program has been developed to fit the PSD models to the experimental data. A practical user-friendly software called "PSD Curve Fitting Software" was developed and introduced to program a simultaneous fitting of all models to soil PSD data of all samples. Some of the capabilities of this software are calculating evaluation statistics for all models and soils and their statistical properties such as average, standard deviation, minimum and maximum for all models, the amount of models' fitting parameters and their statistical properties for all soil samples, soil water retention curve by Arya and Paris (1981) and Meskini-Vishkaee *et al.* (2014) methods, soil hydraulic conductivity by Arya *et al.* (1999) method, different textural and hydraulic properties, specific surface area, and other descriptive statistics of the PSD for all soil samples. All calculated parameters are presented in an output Excel file format by the software. The software runs under Windows XP/7/8/10.

Keywords: mathematical models, soil texture, particle-size distribution (PSD), PSD curve fitting software

* Department of Soil Science, Faculty of Agriculture, Bu Ali Sina University, Hamedan, Iran. E-mail: mostafa.rastgo@gmail.com; Corresponding author's e-mail: h.bayat@basu.ac.ir

** Department of Computer Science, Faculty of Engineering, Bu Ali Sina University, Hamedan, Iran. E-mail: mansoorm@basu.ac.ir

INTRODUCTION

Particle-size distribution (PSD) is considered as a simple test and static physical characteristic of mineral soils that affects many important soil physical and chemical properties (Huang *et al.* 2013), such as Atterberg limits and tensile strength (Bayat *et al.* 2015), growth of plants (Huang *et al.* 2013) and bacterial diversity (Chau *et al.* 2011). In mechanical analysis, soil PSD is expressed as mass percentage of clay, silt and sand fractions. Many investigations have shown that these three fractions are widely applied as predictors of other soil properties such as Atterberg limits, tensile strength (Bayat *et al.* 2015), soil water retention curve (SWRC), saturated as well as unsaturated hydraulic conductivity and available water capacity through mathematical analysis (Botula *et al.* 2013, Gupta and Larson 1979, Haverkamp and Parlange 1986).

Soil PSD is typically applied in the classification systems of soils for engineering and agricultural purposes and also many important soil parameters gained from this curve such as the specific surface area, coefficient of uniformity, coefficient of sorting, the effective particle size and the fines content (Vipulanandan and Ozgurel 2009). Fractal dimension (D) of soil PSD can be used as a practical index for quantifying variations in soil properties and the implications of desertification (Gao *et al.* 2014).

Modeling PSD using parametric models to obtain a continuous curves of PSD is of importance and is used as a basis for estimation of the soil hydraulic properties (Weipeng *et al.* 2015). Many soil databases consist of only several discontinuous points that limit the computation of the soil physical properties. In order to obtain a more complete explanation of soil texture and continuous curves of the PSD, various PSD mathematical models are applied (Shangguan *et al.* 2014). These functions have been suggested to describe a full range of the soil PSD. Some of the equations that have been introduced to describe soil PSD curve, are complicated and difficult to fit to the experimental data (e.g. Fredlund *et al.* (2000) and Nesbitt and Breytenbach (2006) models). In addition, it is extremely time consuming to fit all the PSD models to the experimental data and to select the best performing model, because there was not a full package or software to fit all models to the experimental data of several soil samples and users around the world are confronting a fundamental problem for fitting these models and selecting the most accurate model for their soils. Therefore, the objective of this study is to develop and introduce a type of software to fit all developed PSD models (Table S1 in the supplementary materials) to the experimental data including cumulative mass fraction versus diameter of particles of all soil samples and to estimate the important physical and mechanical properties of soils such as the soil water retention curve by Arya and Paris (1981) and Meskini-Vishkaee *et al.* (2014) methods and also the hydraulic conductivity by Arya *et al.* (1999) method, that are related to soil PSD in a short time. Because

of the large number of models, all of the PSD models that have been contained in the software are shown in the supplementary materials (Table S1).

MATERIALS AND METHODS

Process of developing the PSD curve fitting software

Software design process is as follows:

Definition step

The first step in the problem solving and decision-making process is to identify and define the problem. Many different models have been suggested and developed to describe the PSD data, but, there is no specialized software for fitting them to the experimental PSD data, so far.

Development step

The PSD models were fitted to the PSD experimental data by the nonlinear least squares algorithm that was implemented in MATLAB software (MathWorks 2018) and the accuracy of the results was tested. Nature or type of models is shown in Table S1 in the supplementary materials. The initial values or the ranges (upper and lower) of the parameters have been included in the software. They were determined in two ways: the values reported by the authors who developed their models and by trial and error method (please see Table S1 in the supplementary materials). To make the software more comfortable for using with no need to install the MATLAB software, the final written code was formed as GUI in the MATLAB environment and was extracted from MATLAB and ultimately it was runnable as the solver software. Thus, the users did not need to have MATLAB to use the program. In this software, codes were devised for recalling the input data (the PSD experimental data) in an Excel format and also for exporting the results in an Excel format. This program requires only the cumulative mass fraction versus the diameter of particles to fit the PSD models to the PSD experimental data and calculates the physical, mechanical and hydraulic properties. Moreover, if the soil water retention curve that was predicted by Arya and Paris (1981) and Meskini-Vishkaee *et al.* (2014) methods was needed, then the program requires the bulk density, particle density and saturation water content/porosity. The software would run for each sample individually. The number of the repeated measured points had no effect on the software performance. In addition, if the replicate information was introduced as one sample, it would not have any effects on the software performance and the software would give the best fit to all the data of each sample. However, if the replicate information was introduced as different samples, the software could not handle it and would identify them as the separate individual samples.

The software testing is an activity aimed at evaluating an attribute or capability of a program or system and determining that it meets its required criteria. The testing techniques include the process of executing a program or application with the intent of finding the software bugs (errors or other defects). In this study, the software testing was done at all stages of the software development.

Maintenance step

This software has been programmed in a way that could be installed on Windows operation systems without the user limitation and there were no maintenance costs for it. An overview of the software design process is shown in Fig. 1.
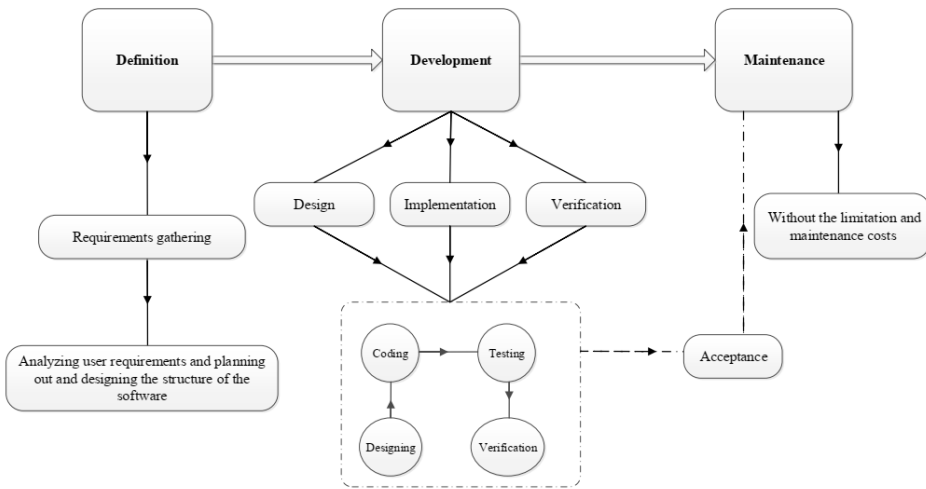
Fig. 1. An overview of the PSD curve fitting software design process

Estimating the soil hydraulic properties by semi-physical methods

Over the past two decades, a variety of PTFs have been developed to estimate soil hydraulic properties based on the soil PSD, parameters and other basic geotechnical properties, but most of them are empirical, not semi-physical. The capillary theory has been used to relate void space between particles to their ability to retain water (Arya and Paris 1981). In other words, the semi-physical methods are mainly based on shape similarity between the SWRC and the PSD curve, implying that the pore-size distribution is closely related to the PSD. The other main purpose of this software was to estimate the SWRC by Arya and Paris (1981) and Meskini-Vishkaee *et al.* (2014) methods and also to estimate the hydraulic conductivity by Arya *et al.* (1999) method in a short time. The Arya and Paris (1981) method is one of the indirect methods which use particle-size distribution curve data to estimate the SWRC. The method provides

some predictions using certain empirical parameters, while the Meskini-Vish-kaee *et al.* (2014) method translates the fitted curve of the PSD into SWRC and does not need any empirical parameters. Also, Arya *et al.* (1999) present-ed a model to calculate the hydraulic conductivity, as a function of soil water content, directly from the PSD. The model is based on the assumption that soil pores can be represented by equal capillary pipes and that the water flow rate is a function of pore size. The pore size distribution is derived from the PSD using the Arya–Paris model. The necessary input data for calculating the soil water retention and the hydraulic conductivity curves are bulk density (g/cm^3), particle density (g/cm^3), the ratio of saturation/porosity, "X" and "C" and their values are determined by the user of arya_paris sheet in the input data file (Fig. S2 in the supplementary materials). In other words, the values of these param-eters are not constant for all soils and should be specified by the user for each soil sample. Arya and Paris (1981) proposed a value of 1.38 for the alpha var-iable, but later research showed that the alpha value is not constant and varies between 1.1 to 2.5 (Vaz *et al.* 2005). On the other hand, "X" and "C" parameters are empirical parameters and are determined by the experimental data for each soil texture. Arya *et al.* (1999) proposed values of 0.482 and 3.602 for the "X" and "C" parameters, respectively, for all soil textures. Therefore, these values were applied to compute Fig. 3 in the manuscript. However, they proposed spe-cific values of "X" and "C" for a few soil textural classes. Also, the PSD Curve Fitting Software can estimate the hydraulic conductivity by Arya *et al.* (1999) method for different values of these parameters (i.e. the values of "X" and "C" can be entered by the users), for each soil.

RESULTS AND DISCUSSION

Review of the PSD curve fitting software

Principles of the output file by the software

In the output Excel file, the results of each model were shown in a separate sheet containing the model name, equation and reference, fitting evaluation statis-tics for all soils which included the coefficient of determination (R^2), root mean square error (RMSE), and the corrected Akaike's information criterion (AIC_c) (Burnham and Anderson 2003), the amount of the estimated parameters for the considered model for all soils, the average, standard deviation, and the minimum and maximum of the amount of the estimated parameters. This software could write two summary sheets, including the "Result" and "Table", allowing a user to compare various models. Moreover, a lot of useful parameters and characteristics have been calculated by this software. An overview of the parameters and charac-teristics that are calculated and provided by the software are shown in Fig. 2.
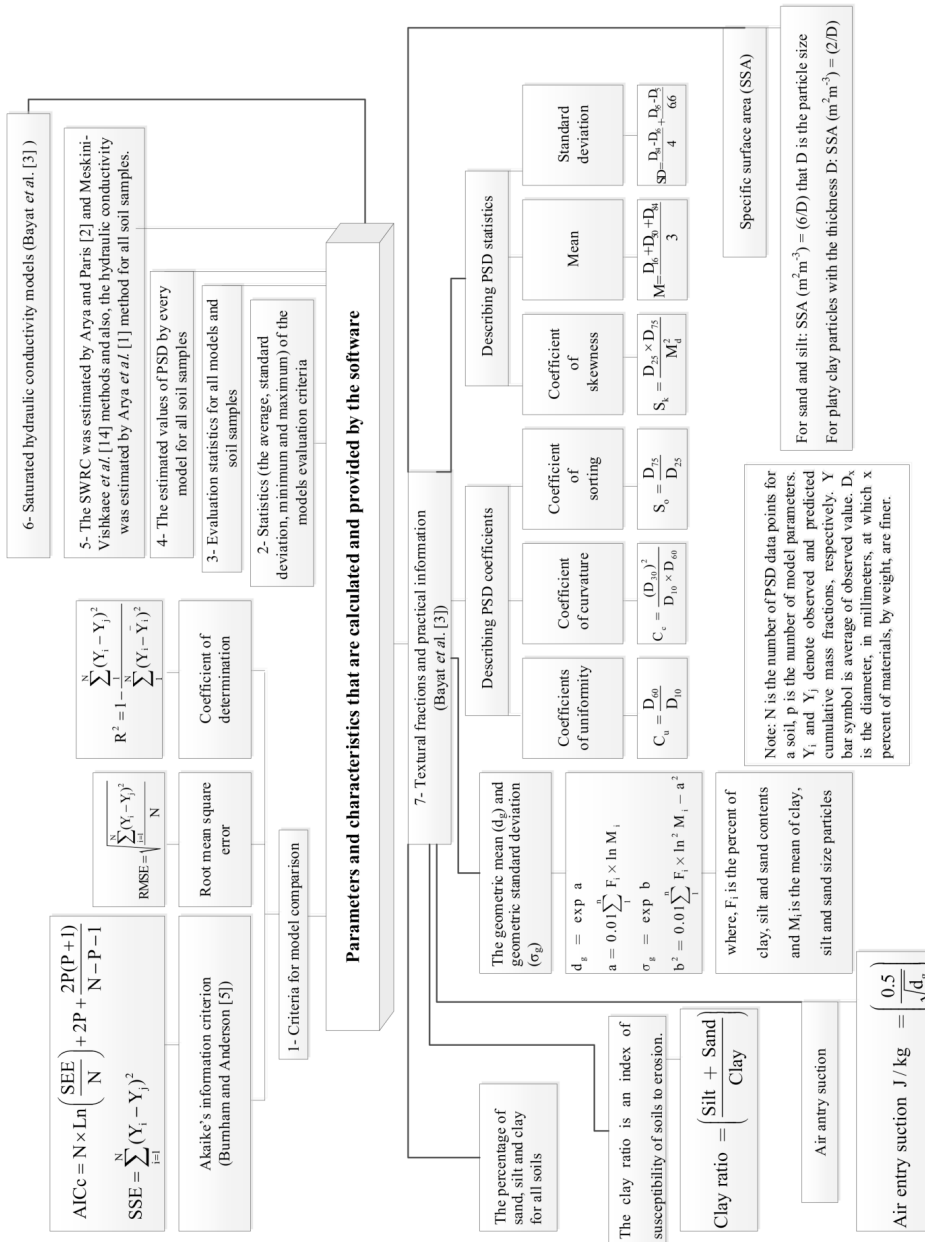
Fig. 2. An overview of the parameters and characteristics that are calculated and provided by the software

Estimating the soil water retention curve (SWRC) and hydraulic conductivity curve (K)

In order to show the capability of the PSD curve fitting software in the estimation of the SWRC and the hydraulic conductivity curve, it was run (the SWRC and the hydraulic conductivity curve were estimated) for nine soil samples from nine soil texture classes and the results were shown in Fig. 3. Arya and Paris (1981) and Meskini-Vishkaee *et al.* (2014) methods could not estimate the SWRC for the studied soil samples precisely. However, Arya and Paris (1981) and Meskini-Vishkaee *et al.* (2014) methods are semi-physical methods and their approach is an attempt to take into account a priori knowledge in the field of soil science (Piron *et al.* 1997). In consequence, the semi-physical methods are less accurate than the mathematical models. Arya *et al.* (1999) developed an easy-to-use spreadsheet-like Corel Quattro Pro 8 for using the Arya–Paris and Arya *et al.* (1999) methods. This spreadsheet has been developed for only one-soil sample, so this is a disadvantage. Also, the Quattro Pro, as a spreadsheet program, is unknown for users, therefore, it is not easy for using. But in the
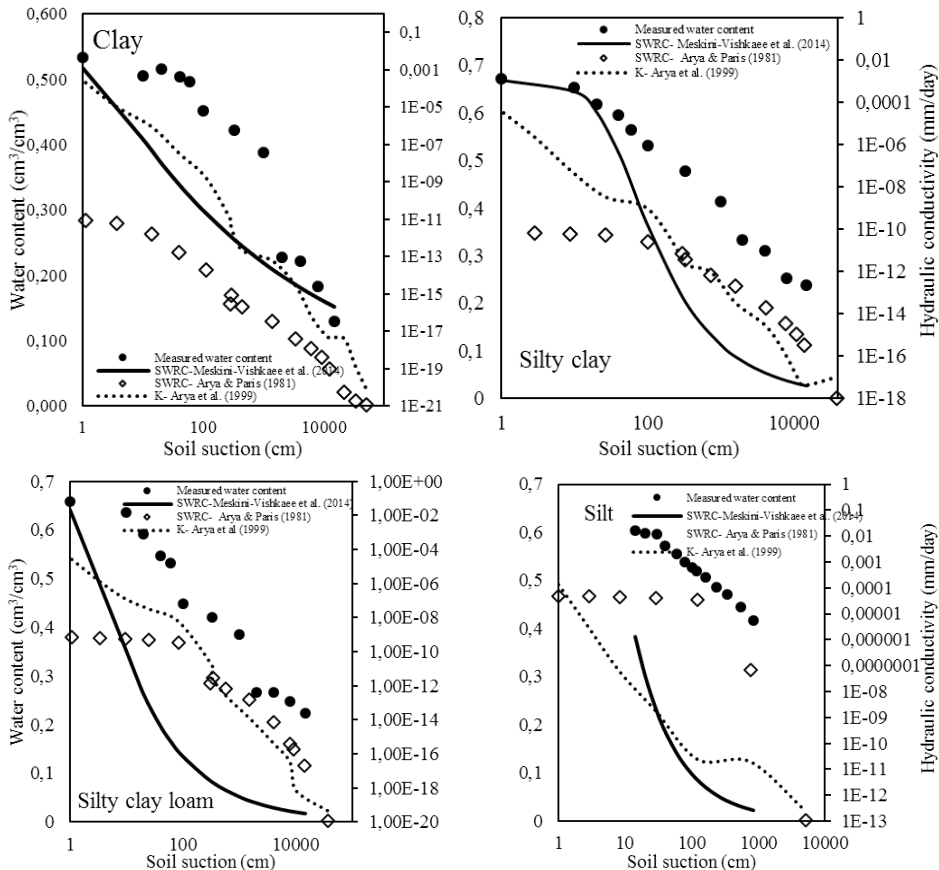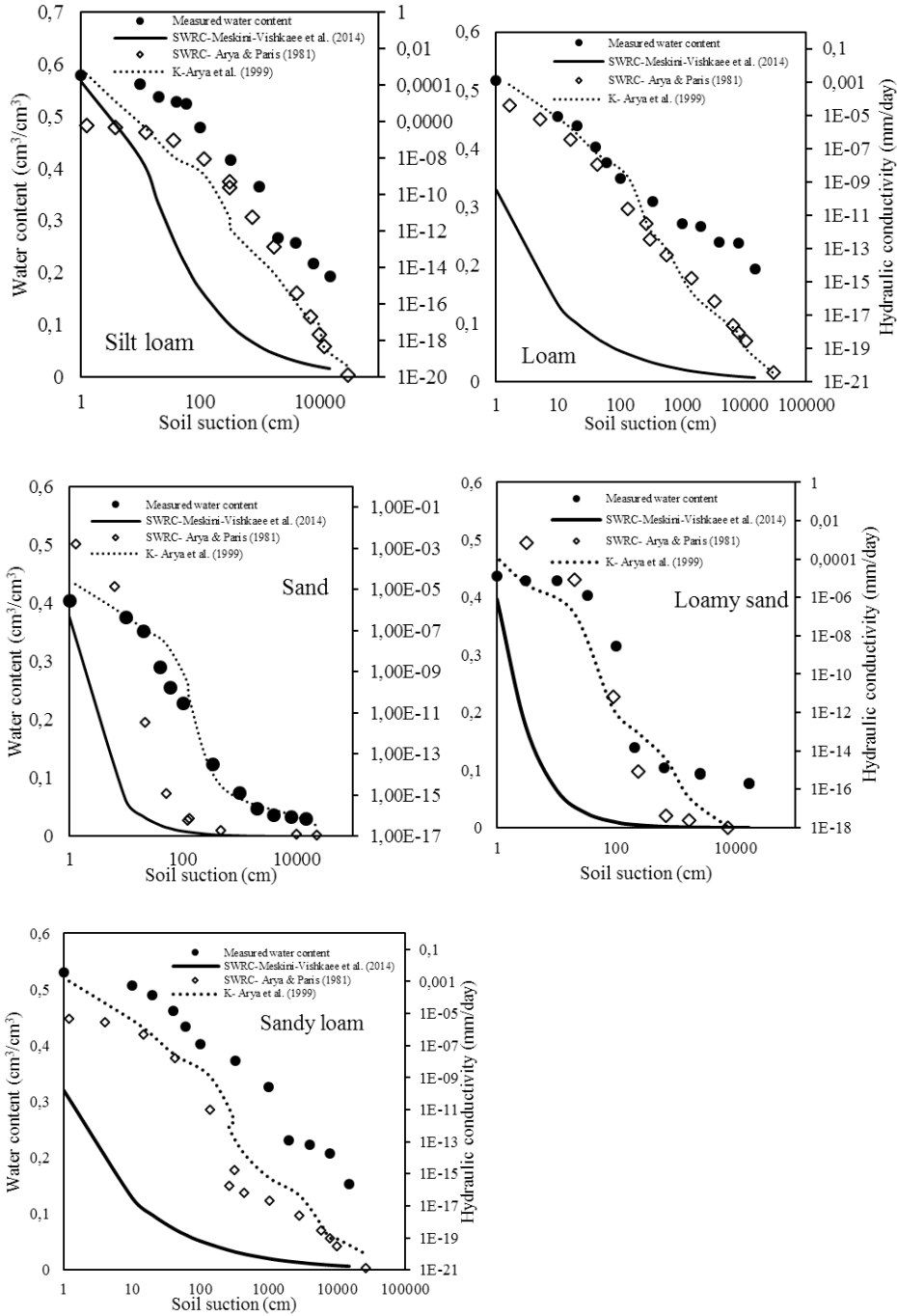
Fig. 3. The results of the accuracy of the Arya and Paris (1981) and Meskini-Vishkaee *et al.* (2014) methods in estimating the SWRC and also the accuracy of the Arya *et al.* (1999) method in estimating the hydraulic conductivity curve for nine soil samples

present study, being easy to use and high calculating speed (in a short time for a large number of soil samples), are the advantages of this software in calculating the SWRC by Arya and Paris (1981) and Meskini-Vishkaee *et al.* (2014) methods and the hydraulic conductivity by Arya *et al.* (1999) method.

Testing the software

Bayat *et al.* (2015) investigated the performances of the PSD models on the PSD data of the soil samples in the unsaturated soil hydraulic database (UNSO-DA) by MATLAB code of the PSD curve fitting software. Moreover, the following investigation was done to evaluate the performance of the software. As mentioned in the study by Bayat *et al.* (2015), 713 soil samples were selected from the UNSODA to fit the 35 PSD models to their PSD experimental data, with different numbers of the measured points. Therefore, 24,955 fittings were performed by the code in 36 min [713 (samples) × 35 (models) = 24,955]. Overall, only in 1,732 cases (6.94%) out of 24,955 fittings, the models were not fitted to the experimental data. In 1,529 cases (6.13%) out of 1,732, the number of the measured points was equal or less than the number of the parameters of the PSD models. Therefore, only in 203 cases (0.81%) out of 24,955 fittings, the number of the measured points was higher than the number of parameters but the code did not perform successfully. In another investigation, 500 soil samples were selected from three provinces of Iran to fit the 35 PSD models to their PSD experimental data. Therefore, 17,500 fittings were performed by the code. Overall, only in 49 cases (0.28%) out of 17,500 fittings, the models were not fitted to the experimental data. The number of experimental data points, N, that is required by the software, is different for various PSD models and should be greater than the number of the model parameters, M, to have a successful fitting ($M \leq N - 1$). In other words, if the number of experimental data points is less than the number of model parameters, the model cannot be fitted to the experimental PSD data (Everitt and Skrondal 2002). There is not any limitation about the maximum number of the measured points for the software. However, by increasing the number of the measured points, accuracy of fitting PSD models to the PSD data would be enhanced. Several reasons could be found for the lack of fitting of the PSD models to the samples with the measured points higher than the number of parameters.

Note that the performance of the code for every PSD model is presented in detail in Table S2 of the supplementary materials. A soil sample was selected to illustrate the fitting accuracy of the different PSD models (Fig. 4). In order to prevent confusing, the PSD curve of 35 models were shown in 4 graphs.
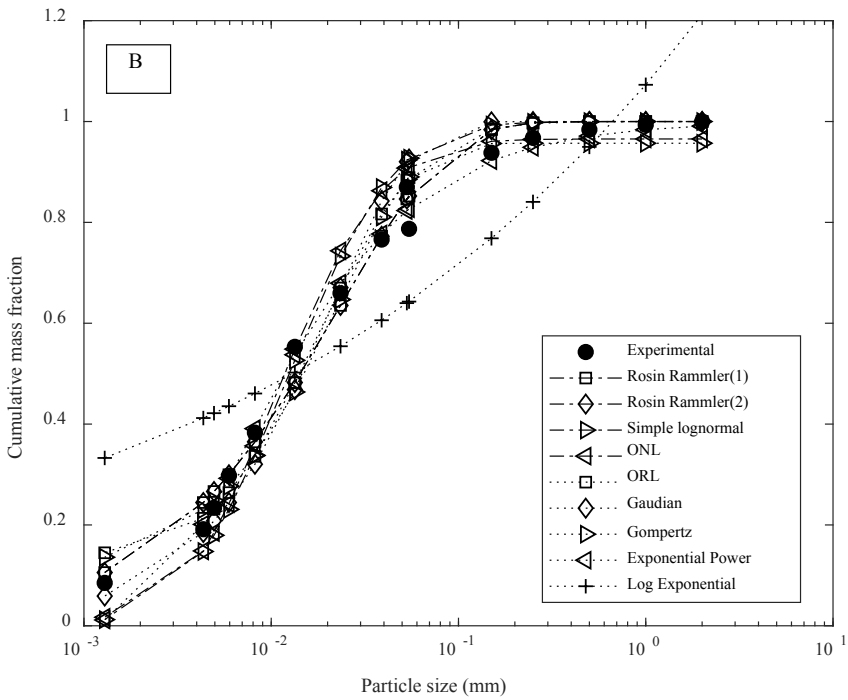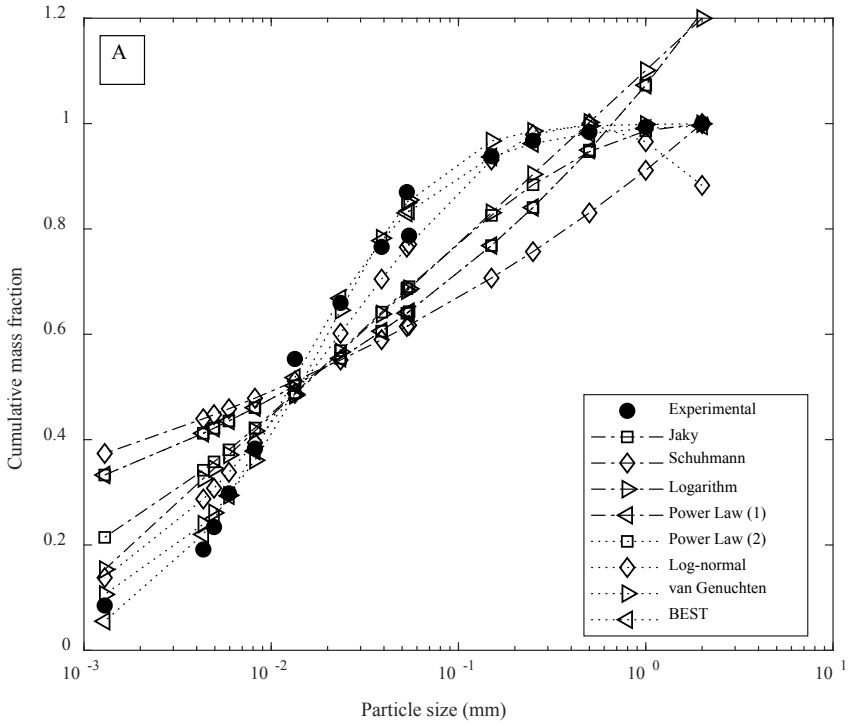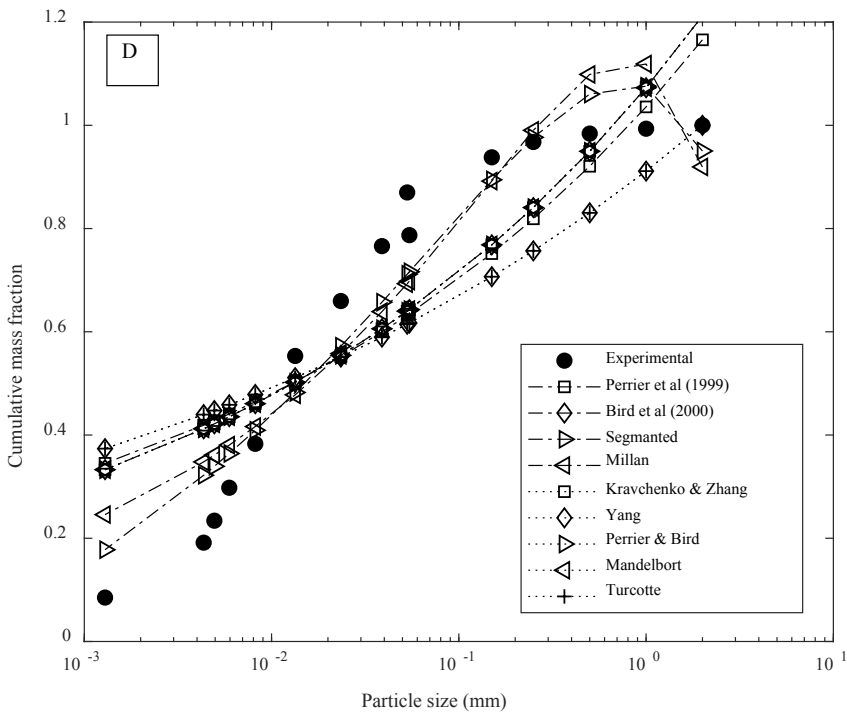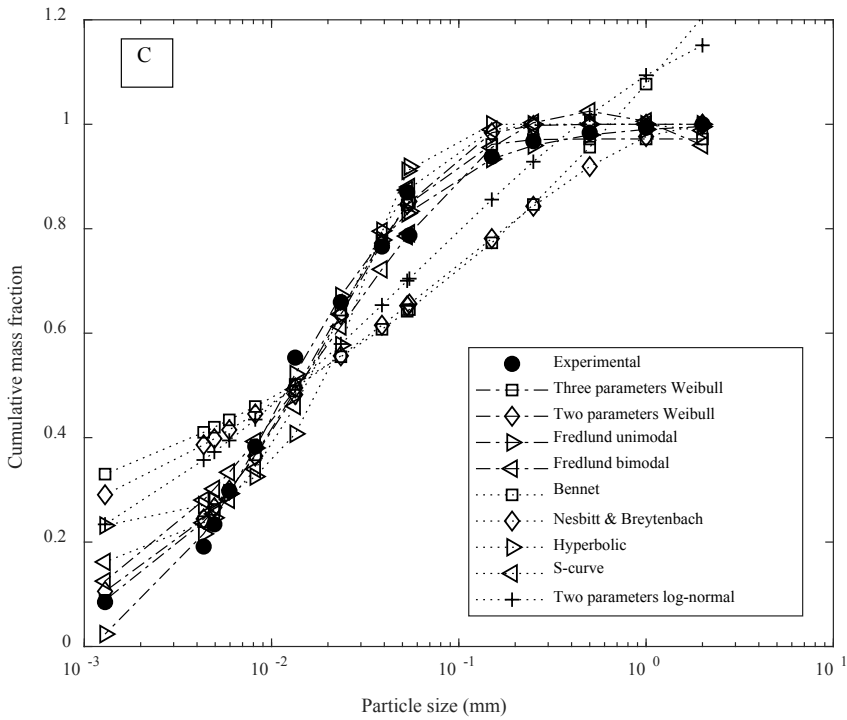
Fig. 4. Comparing fitting accuracy of different PSD models for a soil sample

Software availability

Name of the software: PSD Curve Fitting
Availability and the online documentation: Available upon the request from the corresponding author.
Year of its release: 2019
Hardware required: Laptop and PC
Software required: Windows (tested on Windows 7, 32 and 64 bit)
Programming language: MATLAB
Program size: 254 MB

CONCLUSIONS

There were many different mathematical models which have been developed to describe the PSD curve. A comprehensive and specialized software for fitting the mathematical models of soil PSD on the experimental data has not been developed so far. By understanding the main problems of the model fitting, a comprehensive study was done and all available soil PSD models were collected, and a software for fitting these models to the soil PSD data was developed. The main purpose of this software was the ability to fit several mathematical models of soil PSD of many soils in a short time. Using the Excel files for the input and output data was one of the advantages of this software that made it user-friendly software. In addition, calculating and arranging a huge amount of useful parameters and characteristics, for different areas of agronomy, ecology and environmental science, and their descriptive statistics from the PSD data and fitting the experimental data were additional advantages of this software. Moreover, having a high processing speed and a high standard, could be accounted as other advantages of this software. As a result, the increase in the fitting quality and reduction in servicing and developing costs are the significant features of this software.

REFERENCES

[1] Arya, L.M., Leij, F.J., Shouse, P.J., van Genuchten, M.T., 1999. *Relationship between the hydraulic conductivity function and the particle-size distribution*. Soil Science Society of America Journal, 63(5): 1063–1070.
[2] Arya, L.M., Paris, J.F., 1981. *A physicoempirical model to predict the soil moisture characteristic from particle-size distribution and bulk density data*. Soil Science Society of America Journal, 45(6): 1023–1030.
[3] Bayat, H., Rastgo, M., Zadeh, M.M., Vereecken, H., 2015. *Particle size distribution models, their characteristics and fitting capability*. Journal of Hydrology, 529(3): 872–889.
,

[4] Botula, Y.-D., Cornelis, W.M., Baert, G., Mafuka, P., Van Ranst, E., 2013. *Particle size distribution models for soils of the humid tropics*. Journal of Soils and Sediments, 13(4): 686–698. DOI:10.1007/s11368-012-0635-5.

[5] Burnham, K.P., Anderson, D.R., 2003. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer Science & Business Media, New York, USA.

[6] Chau, J.F., Bagtzoglou, A.C., Willig, M.R., 2011. *The effect of soil texture on richness and diversity of bacterial communities*. Environmental Forensics, 12(4): 333–341.

[7] Everitt, B.S., Skrondal, A., 2002. *The Cambridge Dictionary of Statistics*. Cambridge, Cambridge.

[8] Fredlund, M.D., Fredlund, D., Wilson, G.W., 2000. *An equation to represent grain-size distribution*. Canadian Geotechnical Journal, 37(4): 817–827.

[9] Gao, G.-L., Ding, G.-D., Wu, B., Zhang, Y.-Q., Qin, S.-G., Zhao, Y.-Y., Bao, Y.-F., Liu, Y.-D., Wan, L., Deng, J.-F., 2014. *Fractal scaling of particle size distribution and relationships with topsoil properties affected by biological soil crusts*. PLOS ONE, 9(2): e88559.

[10] Gupta, S., Larson, W., 1979. *Estimating soil water retention characteristics from particle size distribution, organic matter percent, and bulk density*. Water Resources Research, 15(6): 1633–1635. DOI: 10.1029/WR015i006p01633.

[11] Haverkamp, R.T., Parlange, J.-Y., 1986. *Predicting the water-retention curve from particle-size distribution: 1. sandy soils without organic matter*. Soil Science, 142(6): 325–339.

[12] Huang, L., Dong, B.C., Xue, W., Peng, Y.K., Zhang, M.X., Yu, F.H., 2013. S*oil particle heterogeneity affects the growth of a rhizomatous wetland plant*. PLOS ONE, 8(7): e69836.

[13] MathWorks, 2018. *MATLAB: the language of technical computing*. Natick, Massachusetts, United States.

[14] Meskini-Vishkaee, F., Mohammadi, M.H., Vanclooster, M., 2014. *Predicting the soil moisture retention curve, from soil particle size distribution and bulk density data using a packing density scaling factor*. Hydrology Earth System Sciences, 18(10): 4053–4063.

[15] Nesbitt, A., Breytenbach, W., 2006. *A particle size distribution model for manufactured particulate solids of narrow and intermediate size ranges*. Powder Technology, 164(3): 117–123.

[16] Piron, E., Latrille, E., Rene, F., 1997. *Application of artificial neural networks for crossflow microfiltration modelling: "black-box" and semi-physical approaches*. Computers & Chemical Engineering, 21(9): 1021–1030.

[17] Shangguan, W., Dai, Y., García-Gutiérrez, C., Yuan, H., 2014. *Particle-size distribution models for the conversion of Chinese data to FAO/USDA system*. The Scientific World Journal, 2014(2014): 1–11.

[18] Vaz, C.M.P., de Freitas Iossi, M., de Mendonça Naime, J., Macedo, A., Reichert, J.M., Reinert, D.J., Cooper, M., 2005. *Validation of the Arya and Paris water retention model for Brazilian soils*. Soil Science Society of America Journal, 69(3): 577–583.

[19] Vipulanandan, C., Ozgurel, H.G., 2009. *Simplified relationships for particle-size distribution and permeation groutability limits for soils*. Journal of Geotechnical and Geoenvironmental Engineering, 135(9): 1190–1197.

[20] Weipeng, W., Jianli, L., Bingzi, Z., Jiabao, Z., Xiaopeng, L., Yifan, Y., 2015. *Critical evaluation of particle size distribution models using soil data obtained with a laser diffraction method*. PLOS ONE, 10(4): e0125048. DOI: 10.1371/journal.pone.0125048.