The best mathematical models describing particle size distribution of soils

Leila Esmaeelnejad¹ · Fatemeh Siavashi¹ · Javad Seyedmohammadi² ·
Mahmood Shabanpour³

Received: 27 August 2016 / Accepted: 17 September 2016 / Published online: 23 September 2016
© Springer International Publishing Switzerland 2016

Abstract Since particle size distribution (PSD) is a fundamental soil physical property, so determination of its accurate and continuous curve is important. Many models have been introduced to describe PSD curve, but their fitting capability in different textural groups have been rarely investigated. The aim of this study was to evaluate the fitting ability of 15 models on 2653 soil samples from 13 province of Iran, and to determine the best model among them for the PSD of all soil samples as well as for each soil textural group based on evaluation criteria. Results showed that the Weibull model was the most accurate model for all soil samples as well as for the clayey and loamy groups. After the Weibull, Fredlund, Rosin–Rammler and van Genuchten were the most accurate models. However, their differences were not significant ($p \leq 0.05$). Also, for the coarse texture group the S-shape model showed the better fit than the others. These results showed the performance of a particular model varies with the soil textural characteristics.

Keywords Fitting ability · Fredlund · Soil particle size distribution · Textural groups · Weibull model

Introduction

Particle size distribution (PSD) is expressed as mass percentage of clay, silt and sand (Botula et al. 2013) is a fundamentally soil property that is typically used to predict hydraulic properties (Hwang and Hong 2006) and affects soil chemical properties such as adsorption of chemicals, buffering capacity and cation exchange capacity (Rizea et al. 2009; Erashin et al. 2006). Also, soil erodibility, water retention, residual water content, specific surface area, permeability and saturated as well as unsaturated hydraulic conductivity (Vereecken et al. 2010; Nimmo 2004; Vipulanandan and Ozgurel 2009) are affected by PSD. Su et al. (2004) believed that changes in PSD can provide important indications on the impact of land use, desertification processes and soil degradation on soils.

Modeling and mathematical explanation of PSD, in addition to above-mentioned issues, has some more benefits such: (a) classification of soils using the best fitting parameters, (b) facilitating the search for soils with similar characteristics in data bases (Fredlund et al. 2000), (c) using for designing of drainage filters and (d) obtaining useful soil parameters such as effective diameter (d_{10}) and uniformity coefficient (Vipulanandan and Ozgurel 2009).

Traditionally, mechanical analysis of soil particle size is done by mass measurement of three soil fraction, sand, silt, clay and determination of texture class using soil texture triangular. But, soil samples that place in a specified texture class may have different PSD. For example, clay texture class in USDA classification system, contains all soil samples that have clay amount between 40 to 100 percent, while, determination of percentage or proportion of three main soil fraction i.e. sand, silt and clay, don't give us perfect information about soil PSD. So, mathematical models of PSD are used (Bittelli et al. 1999; Posadas et al.

✉ Leila Esmaeelnejad
esmaeelnejad.leila@gmail.com

¹ Department of Soil Science, Faculty of Agricultural Engineering and Technology, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran

² Department of Soil Science, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

³ Department of Soil Science, Faculty of Agriculture, University of Guilan, Rasht, Iran

2001). An appropriate mathematical model must describe PSD such accurate that minimize differences between measured and estimated data (Vipulanandan and Ozgurel 2009). It is essential to say that however the number of model parameters be more, the accuracy of model would be increased, too. In spite of using numerous models for description of PSD, practical application of them is limited because of their intricacy (Vipulanandan and Ozgurel 2009). Soil PSD is often assumed log-normal (Campbell 1985; Shirazi and Boersma 1984), but, Buchan (1989) showed that in half part of soil texture triangular, hypothesis of being lognormal isn't valid, because, soil texture classes consist extensive range of particle size from <0.02 to 2 mm. Fredlund et al. (2000) reported that main limitation of lognormal distribution model for PSD demonstration is its symmetry in logarithmic scale While, PSD of many soils don't show this symmetry.

Soil texture as well as the bimodal character of the soil significantly affects the respective performance of PSD models (Botula et al. 2013). There are only few studies have conducted to compare different mathematical PSD models (see Table 1).

According to the limit of our knowledge, no previous studies have been conducted based on a large dataset of soils which represents various conditions such as climate, texture and parent material. Moreover the performances of different PSD models have not yet been explicitly and comprehensively investigated in the previous studies. Therefore, the objective of this study are (i) to evaluate the performance of 15 various models for describing cumulative PSD of 2653 soils of the different regions of Iran using

a comprehensive database and (ii) to investigate the influence of texture on the fitting performances of these PSD models. Then we would select the best model for all soil samples and each texture group.

Materials and methods

Samples of 2653 soils were collected from 16 provinces of Iran, representing a diversity of parent materials, climates, and times of soil formation, mineralogy and texture. This dataset represents 13 different FAO soil groups. Soil sampling carried out from surface layer (10–35 cm) of surveyed soil series. Samples transferred to laboratory, air dried and passed through 2 mm mesh. The PSDs were determined using conventional methods following H_2O_2 pretreatment to eliminate organic matter. Fine size fractions were determined by hydrometer method, whereas the coarse fractions were obtained by sieving. Particle-size fraction data were classified by USDA standard (Gee and Or 2002). For obtaining 12 points of PSD curve, reading of hydrometer carried out in 30 s, 1, 3, 10, 30, 60, 90, 120, 300, 600, 1440, 2880 min from beginning of test. Then, 15 PSD models (see Fig. 1) fitted on measured PSD data.

All soil samples, were divided in 3 texture categories including clay soils (clay and silty clay), loamy soils (silty loam, clay loam, sandy loam, silty clay loam, loam) and sandy soils (sandy, sandy loam) according to USDA. Accuracy of models evaluated for all samples and each texture category, separately.

Table 1 A review of previous studies on effect of textural groups on performance of different PSD models

Researchers	Studies
Buchan et al. (1993)	Tested five log-normal models for 71 soils from New Zealand
Hwang et al. (2002)	Evaluated fitting ability at 7 models on 1387 Korean soil samples
Hwang (2004)	Studied the effect of soil texture of 1387 Korean soil samples on the performance of nine different PSD models
Lima and Silva (2007)	Conducted a study on PSD of suspended sediments in river water
Bah et al. (2009)	Compared the fitting ability of seven PSD models to data of 55 fine-textured soil samples from New South Wales in Australia
Bagarello et al. (2009)	Evaluated the ability of the hyperbolic and Fredlund models to fit 243 soils
Shangguan (2012)	Investigated ten PSD models for the conversion of soil texture classification from ISSS and Katschineski to FAO/USDA system for soils from china
Botula et al. (2013)	Evaluated the performance of 18 PSD models for 1412 soils from Central Africa which had humid tropics
Weipeng et al. (2015)	Studied critical evaluation of PSD models of 1013 soil samples from different regions of china
Bayat et al. (2015)	Investigated fitting capability of PSD models on the 713 PSD experimental data of the UNSODA database

Fig. 1 An overview of the mathematical models of PSD (See above-mentioned references for further information)

Parameters optimization procedure

Iterative nonlinear regression was used to find the values of fitting parameters, so that, the best fitting would be determined between PSD models and measured data (Hwang et al. 2002; Bah et al. 2009). For this purpose, SOLVER routine operator of the Microsoft Excel software was applied (Wraith and Or 1998). Models comparison carried out via comparison of

evaluation indices means using LSD test in SPSS software (SPSS Inc 2015).

Comparison of models accuracy

We used three statistical indices to determine the performance of the 15 PSD models. One index was Akaike information criterion (AIC) which examines the complexity of the model together with its goodness- of- fit to the

Table 2 The statistics of the clay (0–0.002 mm), silt (0.002–0.05 mm) and sand (0.05–2.0 mm) contents

Texture groups	% clay				% silt				% sand			
	Max	Min	SD	Mean	Max	Min	SD	Mean	Max	Min	SD	Mean
Clay	67.2	40.4	8.3	53.1	46.0	16.0	9.8	33.6	23.5	7.2	5.1	13.2
Loam	39.0	0.4	8.8	18.9	73.2	15.5	12.3	48.3	74.4	7.9	15.1	32.7
Sand	6.60	6.1	0.3	6.3	13.4	12.5	0.67	12.9	81.3	79.8	1.0	80.6
Total samples	67.2	0.45	12.1	20.8	73.2	12.5	13.2	46.9	81.3	7.2	16.3	32.1

Table 3 Statistical analysis of the accuracy of the fitting particle size distribution models on the all soil samples

Models	R^2_{adj}		AIC		RMSE	
	SD	Mean	SD	Mean	SD	Mean
Weibull	0.078	0.972 ^a	17.5	−105 ^a	0.026	0.039 ^a
Rosin–Rummler	0.063	0.970 ^{ab}	17.1	−104 ^a	0.024	0.042 ^{ab}
van Genuchten	0.068	0.967 ^{ab}	16.2	−102 ^{ab}	0.025	0.044 ^{ab}
Fredlund	0.178	0.943 ^{abc}	17.2	−98.9 ^b	0.033	0.045 ^{ab}
Expo-power	0.324	0.932 ^{abcd}	17.6	−99.1 ^b	0.047	0.052 ^b
Jaky	0.069	0.942 ^{abc}	19.7	−93.4 ^c	0.030	0.062 ^c
Logarithmic	0.076	0.918 ^{bcd}	18.3	−84.8 ^d	0.033	0.075 ^d
ORL	0.097	0.912 ^{cde}	12.4	−79.8 ^c	0.029	0.076 ^d
S-shape	0.158	0.878 ^{ef}	20.9	−78.9 ^e	0.051	0.084 ^{de}
hyperbolic	0.099	0.882 ^{def}	10.3	−76.2 ^e	0.027	0.090 ^e
Power law	0.093	0.840 ^{fg}	14.6	−70.9 ^f	0.038	0.110 ^f
Schumann	0.102	0.803 ^g	13.8	−68.8 ^{fg}	0.040	0.122 ^g
SLM	0.210	0.728 ^h	9.9	−62.5 ^h	0.042	0.133 ^h
Gompertz	0.549	0.550 ⁱ	27.2	−66.2 ^{gh}	0.104	0.137 ^h
ONL	0.317	0.551 ⁱ	16.2	−54.6 ⁱ	0.064	0.166 ⁱ

All small letters are signs for significant difference at level 5 %. SD standard deviation, Ave average

sample data, balances between the two, and discourages over fitting:

$$AIC = N \ln(SSE) - N \ln N + 2P, \quad (1)$$

$$SSE = \sum_{i=1}^N (Y_{pi} - Y_{mi})^2$$

where, SSE is the residual sum of squares. N is the number of PSD data points for soil i, P is the number of model parameters, and Y_{pi} and Y_{mi} are predicted and measured cumulative mass fractions, respectively. According to AIC, the model with the smallest AIC value was selected as the best soil description. AIC imposes a penalty for additional fitting parameters (Botula et al. 2013).

Further indexes for testing the quality of fit, was adjusted mean square error, $RMSE_{(adjusted)}$, and R^2 because of the unequal number of parameters in the PSD functions (Weipeng et al. 2015).

$$RMSE_{(adjusted)} = \sqrt{\frac{SSE}{N - P}} \quad (2)$$

$$R^2_{(adjusted)} = 1 - \left[\frac{SSE/N - P}{SST/N - 1} \right] \quad (3)$$

As the number of fitted parameters increases, the fitting performance generally improves. Thus, it is essential to determine this better fitting on measured data is related to its inherent nature or occurs at the expense of a corresponding increase in the possibility of over-parameterization. So, we used adjusted determination coefficient (R^2_{adj}) to eliminate the effect of difference in models parameters number. R^2_{adj} is usually applied to compare models that have different parameters number.

Cluster analysis

Cluster analysis is a multivariate method which aims to classify a sample of subjects (or objects) on the basis of a set of measured variables into a number of different groups such that similar subjects are placed in the same group (Everitt et al. 2001). In this study, selecting the number of clusters was determined automatically by the SPSS 23 software. To classify the PSD models based on their fitting accuracy (using RMSE criterion as a variable), cluster analysis with Ward's method and Euclidean distance between the groups was performed by SPSS 23 software (SPSS Inc 2015).

Fig. 2 Cluster analysis diagram for all soil samples based on accuracy of RMSE

Results and discussion

All soil samples

In Table 2, statistical indexes for clay, silt and sand are given for all soil samples and textural groups, separately. On the basis of Table 2 texture fractions had extended range and showed that soil samples had good distribution in all 12 USDA texture classes.

Table 3 shows statistical analysis of the accuracy of the fitting PSD models on the all soil samples. Models are arranged based on their RMSE values so as each model

Fig. 3 Cluster analysis diagram for clayey texture group based on accuracy of RMSE

which had minimum rate of RMSE is placed upper than the others in all tables.

Also Fig. 2 demonstrates results of clustering. RMSE amounts (Table 3) showed that Weibull model had the best accuracy for PSD curve fitting in all soil samples.

This result was in accordance with Buchan et al. (1993) researches. In cluster analysis diagram, Jaky and Power-exponential models placed in one class, jointly. Rosin–Rammler model had higher accuracy ($p \leq 0.05$) than log-normal (ORL, SLM, ONL) distribution model similar result is reported by Gracza (1962). SLM, Gompertz and ONL had the weakest prediction accuracy ($p \leq 0.05$) in

Table 4 Comparison of models accuracy for prediction of PSD in clay texture group

Models	R^2_{adj}		AIC		RMSE	
	SD	Ave.	SD	Ave.	SD	Ave.
Weibull	0.023	0.973 ^a	21.3	−117.1 ^a	0.010	0.028 ^a
Rosin–Rummler	0.023	0.972 ^a	20.4	−117.7 ^a	0.010	0.029 ^a
van Genuchten	0.025	0.967 ^a	18.8	−114.3 ^a	0.010	0.032 ^a
Fredlund	0.050	0.936 ^{ab}	23.5	−106.6 ^a	0.019	0.045 ^{ab}
Expo-power	0.787	0.730 ^{abc}	25.9	−109.9 ^a	0.076	0.051 ^{abc}
Jaky	0.038	0.901 ^{ab}	11.5	−89.0 ^b	0.009	0.059 ^{abc}
Logarithmic	0.072	0.872 ^{ab}	17.9	−89.9 ^b	0.020	0.066 ^{bcd}
ORL	0.198	0.826 ^{abc}	17.6	−80.8 ^{bc}	0.036	0.069 ^{bcd}
S-shape	0.056	0.856 ^{abc}	11.7	−85.3 ^b	0.010	0.071 ^{bcd}
hyperbolic	0.528	0.680 ^{bc}	25.1	−89.5 ^b	0.073	0.078 ^{bcd}
Power law	0.081	0.807 ^{abc}	15.1	−81.3 ^{bc}	0.021	0.083 ^{cd}
Schumann	0.088	0.749 ^{abc}	13.4	−78.0 ^{bc}	0.022	0.096 ^{de}
SLM	0.218	0.594 ^c	5.4	−67.4 ^c	0.030	0.118 ^e
Gompertz	0.067	0.149 ^d	8.0	−50.8 ^d	0.026	0.179 ^f
ONL	0.031	0.000 ^e	7.4	−43.0 ^d	0.031	0.215 ^g

All small letters are signs for significant difference at level 5 %. SD standard deviation, Ave average

Table 5 Comparison of models prediction accuracy for loamy group

Models	R^2_{adj}		AIC		RMSE	
	SD	Ave.	SD	Ave.	SD	Ave.
Weibull	0.081	0.972 ^a	16.9	−103.9 ^a	0.026	0.039 ^a
Rosin–Rummler	0.110	0.960 ^{ab}	16.4	−99.3 ^{bc}	0.028	0.042 ^a
van Genuchten	0.065	0.970 ^a	16.7	−103.2 ^{ab}	0.025	0.043 ^a
Fredlund	0.070	0.966 ^a	15.8	−101.3 ^{abc}	0.025	0.045 ^a
Expo.-power	0.270	0.945 ^{ab}	16.8	−98.7 ^c	0.045	0.051 ^a
Jaky	0.071	0.943 ^{ab}	18.7	−92.6 ^d	0.030	0.063 ^b
Logarithmic	0.075	0.922 ^{abc}	18.2	−84.7 ^e	0.033	0.075 ^c
ORL	0.100	0.911 ^{bc}	11.5	−78.8 ^f	0.030	0.077 ^{cd}
S-shape	0.156	0.879 ^{cd}	16.9	−77.0 ^f	0.049	0.085 ^{de}
hyperbolic	0.102	0.883 ^{cd}	9.8	−75.4 ^f	0.027	0.092 ^e
Power law	0.093	0.841 ^{de}	14.3	−70.0 ^g	0.037	0.111 ^f
Schumann	0.102	0.804 ^e	13.6	−67.9 ^g	0.040	0.125 ^g
SLM	0.529	0.597 ^g	26.7	−66.9 ^g	0.104	0.133 ^g
Gompertz	0.207	0.734 ^f	9.9	−62.0 ^h	0.042	0.134 ^g
ONL	0.308	0.573 ^g	15.9	−54.3 ⁱ	0.065	0.164 ^h

All small letters are signs for significant difference at level 5 %. *SD* standard deviation, *Ave* average

Fig. 4 Cluster analysis diagram for loamy texture group based on accuracy of RMSE

comparison with the other models. Even, the difference between SLM and Gompertz model with ONL model was significant ($p \leq 0.05$) and in cluster analysis diagram, ONL model, alone placed in a class, and SLM and Gompertz model placed in other class, jointly.

AIC values (Table 3) showed that Weibull model had minimum amount in all studied soil samples, and implied that this model was the most favorable one so that had highest accuracy with reasonable parameter numbers. This difference was significant ($p \leq 0.05$) with all models

except Rosin–Rammmler and van Genukten models. After these models, Fredlund model had the best accuracy than the others, but, hadn't significant differences ($p \leq 0.05$) with van Genukten and power-exponential models. This was in contrast with Hwang (2004) results, which stated Fredlund had better accuracy than Van genukhten model. But Hwang (2004), emphasized that Weibull and van Genukten had reasonable fitting accuracy. In our study, these two later models had good fitting accuracy, too. Significant differences between ORL, S-shape and hyperbolic models were not seen at level 5 %. ORL model had higher prediction accuracy than SLM and ONL models. Also, ONL model had the weakest prediction accuracy among all models. Hagen et al. (1987) showed that log-normal distribution models often failed to provide a close fit on PSD measured data, at the extremes of distribution curves. In other words, extreme tails of PSD curves will often deviate from log-normality. The first limitation associated with using a log-normal type of equation is the assumption that the grain-size distribution is symmetric. In reality, the grain-size distribution is often nonsymmetric and can be better fit by a different type of equation. Second, a method for fitting soils that are bimodal or gap-graded is of value and the four parameter log-normal equations have not been found to be satisfactory for fitting these types of grain-size distribution (Fredlund et al. 2000).

R^2_{adj} values (Table 3) showed that Weibull, Rosin–Rammmler, van Genukten, Fredlund, Power-exponential and Jaky models had no significant difference ($p \leq 0.05$) and so, they can place in one group, jointly. According to R^2_{adj} , these models had similar accuracy. Because, R^2_{adj} isn't

Table 6 Comparison of models accuracy for prediction of PSD in sandy texture category

Models	R^2_{adj}		AIC		RMSE	
	SD	Ave.	SD	Ave.	SD	Ave.
Weibull	0.000	0.997 ^a	11.2	−172.9 ^a	0.007	0.002 ^a
Rosin–Rummler	0.001	0.996 ^{ab}	3.7	−124.8 ^b	0.002	0.018 ^b
van Genuchten	0.001	0.995 ^{ab}	3.4	−123.5 ^b	0.002	0.021 ^{bc}
Fredlund	0.001	0.994 ^{ab}	3.0	−116.8 ^c	0.002	0.024 ^c
Expo-power	0.001	0.990 ^{bc}	1.5	−108.8 ^d	0.001	0.032 ^d
Jaky	0.000	0.985 ^{cd}	0.7	−105.1 ^{de}	0.000	0.038 ^e
Logarithmic	0.001	0.983 ^{cd}	1.0	−103.2 ^{de}	0.001	0.040 ^{ef}
ORL	0.002	0.981 ^d	2.0	−98.4 ^{ef}	0.002	0.043 ^f
S-shape	0.000	0.973 ^e	0.1	−95.3 ^{fg}	0.000	0.051 ^g
hyperbolic	0.001	0.966 ^{ef}	0.7	−93.0 ^{gh}	0.001	0.058 ^h
Power law	0.001	0.962 ^f	0.7	−89.4 ^{gh}	0.001	0.061 ^h
Schumann	0.002	0.951 ^g	0.8	−87.0 ^h	0.002	0.069 ⁱ
SLM	0.003	0.928 ^h	0.9	−78.6 ⁱ	0.002	0.084 ^j
Gompertz	0.004	0.926 ^h	1.2	−77.2 ⁱ	0.003	0.085 ^j
ONL	0.010	0.808 ⁱ	0.8	−62.0 ^j	0.003	0.137 ^k

All small letters are signs for significant difference at level 5 %. *SD* standard deviation, *Ave* average

Fig. 5 Cluster analysis diagram for sandy texture group based on accuracy of RMSE

a powerful tool for relative discrimination of nonlinear models, although it does measure the absolute amount of variability accounted for by the model (Buchan et al. 1993). Thus, we require better relative indicators of model performance, capable of discriminating statistically significant differences between models. Fredlund model had higher R^2_{adj} than S-shape and hyperbolic models. This result is in accordance with Vipulanandan and Ozgurel (2009). SLM, Gompertz and al ONL, had a weak fitting capability rather than other models. Even, difference between ONL

and Gompertz models with SLM one was significant ($p \leq 0.05$). According to Bayat et al. (2015), Weibull, is the one of models had high fitting accuracy over a range of particle sizes.

Clay group

Table 4 shows comparison of models accuracy for prediction of PSD in clay texture group.

Also, cluster analysis diagram is shown in Fig. 3. Results proved that Weibull model had the highest accuracy with a view to all three Validation criterions, and from RMSE point of view, except Rosin–Rummler, van Genukten, jaky, power-exponential, and ORL models, its difference with other models was significant ($p \leq 0.05$). In cluster diagram analysis, Weibull, Rosin–Rummler, and van Genukten placed in one class and Jaky, exponential power and ORL in another class, jointly. ONL and Gompertz models had lower prediction accuracy than other models from all three indicators i.e. AIC, RMSE and R^2_{adj} . They also, settled in a common group. These two models hadn't any significant difference ($p \leq 0.05$).

In clay texture group, one-parametric jaky model, with a view to all three validation criterion, had more acceptable results to predict PSD curve than models with extra parameters such as, four-parametric Gompertz and three-parametric ONL models. Thus, it is concluded that accuracy of a model, doesn't always increase with over-parameterization. This result was different from Hwang (2004) one as well as Fredlund et al. (2000), who reported that the performance of the Fredlund model would be

Fig. 6 A plot of fitting capability comparison for the best five models for a soil sample of different textural groups

increased as the clay content increased. This result may be due to the penalties imposed by the AIC (Burnham and Anderson 2002) for additional fitting parameters of multi-parameter models comparing with the one-parameteric models (Bayat et al. 2015). Hwang (2004) deduced that van Genukhten model wasn't suitable for soils with high clay content. In clay texture group, AIC indicator, demonstrated difference between all models in the best manner while R^2_{adj} criterion, didn't show a good separation capability.

Loam group

Table 5 shows comparison of models prediction accuracy for loamy group.

Cluster analysis diagram is given in Fig. 4 too.

Results showed that Weibull model had the best fitting capability accuracy from view to all three validation criterions. According to RMSE values, difference between Weibull model and other models was significant ($p \leq 0.05$) except Fredlund, Rosin–Rammler, van Genukhten, and Power-exponential models, and these models settled in one group in cluster analysis diagram, too. This result showed that mentioned models had similar fitting capability. ONL model had the lowest fitting accuracy rather than other

models with view to RMSE and AIC amounts and placed in a separate group in cluster analysis diagram. According to R^2_{adj} Gompertz and ONL models, had the weakest results and their differences with other models was significant ($p \leq 0.05$). In loamy texture category, one-parameteric Jaky model, still had higher prediction accuracy than multi-parameter models such as four-parametric Gompertz (according to all three validation criterion i.e. R^2_{adj} , RMSE and AIC), three-parametric ONL (according to tree indicator) and three-parametric ORL (on the basis of RMSE and AIC). Fredlund model had higher R^2_{adj} than S-shape and tangent-hyperbolic models. This result was similar to Vipulanandan and Ozgurel (2009) one.

According to our findings, Weibull model had the best results comparing to the other PSD models for all soil samples, and clay and loamy soil categories. Bayat et al. (2015), with examining the behavior of the fitted parameters for the PSD models concluded that Weibull model had advantages: (a) flexibility and high accuracy which showed this model has high accuracy and flexibility in fitting to different soils, and its parameters could describe the PSD curve adequately, (b) Non-overlap which shows that the effect of the parameters of a PSD model on the shape and

Fig. 7 Percent of the soil samples in which the PSD models performed the best fit on the PSD curve, based on the RMSE criterion. **a** All soil samples, **b** clay texture group, **c** loamy texture group, **d** sandy texture group

position of the PSD curve is distinguishable, (c) Simple form and easy fitting procedure, (d) Non-over-prediction that shows this model have no over-prediction of cumulative mass fraction at the coarse end of the PSD curve. In fact, the models having over-prediction, predicts the values more than 1 for relative cumulative mass fraction that didn't have a physical meaning. Also, according to Bayat et al. (2015), parameters of this model have a physical concept. Also, overall classification of this model is categorized as "good" quality of fitting procedure (Bayat et al. 2015).

Sand group

Table 6 shows comparison of models accuracy for prediction of PSD in sandy texture category.

Also, its cluster analysis diagram is given in Fig. 5. Results of this group are different from two former categories to some extent. In this category, S-Shape model was the most accurate model according to R^2_{adj} , RMSE and AIC indicators, so that settled in a separate group in cluster analysis diagram. After it, Gompertz and Weibull models

had higher fitting capability. Gompertz, Weibull and Fredlund models placed in one group in cluster analysis, jointly. So, we can conclude that these three models are suitable for PSD description of sandy texture group. Various order of accurate in different textural groups showed that fitting capability of models is different in each textural group and texture could affect the performance of PSD models. So, the most accurate model should be selected for each texture category.

Hwang et al. (2002) reported that Gompertz model had the best fit on sandy and sandy loam soils. That is similar to our result. Power-exponential, SLM and logarithmic models were the weakest models for sandy texture group and according to RMSE, R^2_{adj} and AIC criteria; their difference with other models was significant ($p \leq 0.05$). Difference of power-exponential and SLM models with logarithmic model was significant ($p \leq 0.05$) with a view to three evaluation critical i.e., RMSE, R^2_{adj} , AIC.

In Fig. 6, illustration of fitting comparison of the best models for clay, loam and sandy texture groups are given, respectively. Figure 7 showed that each model had the best fitting in how many percent of soil samples (all samples

and each texture group, separately). It must note that in majority of past researches, (Hwang et al. 2002; Hwang 2004; Vipulanandan and Ozgurel 2009), Fredlund model was reported as the best model for PSD description. In our study, acceptable fitting of Fredlund model on measured data was seen, too. So that, in Fig. 7, except sandy texture group, Fredlund model provided the best fitting ability in majority number of all soil samples and also, for loamy and clay texture groups, due to having further parameters than models that have fewer parameters such as lognormal likewise, the relative performance increased with increase of clay content (Fredlund et al. 2000; Hwang 2004). Also, we can consider that Fredlund model is physical one. Investigation of models standard deviation showed that the Fredlund model had higher standard deviation than Weibull model indicates a poor fitting curve. Also, according to Bayat et al. (2015) Fredlund unimodal model had a complex form and difficult fitting procedure which are disadvantages of this model. But advantages of Fredlund model are flexibility and high accuracy as well as non-overlapping. Also, overall classification of this model categorized as “good” quality of fitting procedure (Bayat et al. 2015). According to Fig. 7d, it was found that the fits of PSD models declined with increasing sand content (Hwang et al. 2002). On this basis, it can be noted that model performance deteriorated as the sand proportion of soils increased. This is consistent with results obtained by Weipeng et al. (2015).

Conclusion

Majority of studied models had acceptable accuracy for prediction of soils PSD curves. Our results showed that Weibull, Rosin–Rammler and van Genuchten models had the best prediction accuracy according to all three validation criterion. So, it is probable that accuracy of these models has less dependence to texture than other models. Results indicated that S-Shape and Gompertz models had better fit than other models in coarse texture group. Validation of fitting procedure of 15 models for prediction of soil PSD showed that three-parametric Weibull model had the most accuracy. This ability can't be as a result of number of parameters, because, Gompertz, Fredlund, ONL and ORL models with further or similar parameters number had poorer prediction ability. Thus, increasing in parameter number, can't always guarantee to increase the models ability. Likewise, Fredlund model had the best fit on majority of soil samples. So, we can use Weibull and Fredlund models with acceptable accuracy to predict PSD curve that is an important property in soil physics, mechanics and indirect measurement of soil hydraulic properties.

Acknowledgments The authors would like to thank anonymous reviewers for their thorough and constructive reviews. Also, we acknowledge a group of students of Tehran, Tabriz, Guilan and Ahwaz Universities helped with data collection.

References

- Bagarello V, Provenzano G, Sgroi A (2009) Fitting particle size distribution models to data from Burundian soils for the BEST procedure and other purposes. *Biosyst Eng* 104:435–441. doi:10.1016/j.biosystemseng.2009.07.008
- Bah AR, Kravchuk O, Kirchhof G (2009) Fitting performance of particle-size distribution models on data derived by conventional and laser diffraction techniques. *Soil Sci Soc Am J* 73:1101–1107. doi:10.2136/sssaj2007.0433
- Bayat H, Rastgo M, Mansourizade M, Vereeken H (2015) Particle size distribution models, their characteristics and fitting capability: a review. *J Hydrol* 529:872–889. doi:10.1016/j.jhydrol.2015.08.067
- Bittelli M, Campbell GS, Flury M (1999) Characterization of particle-size distribution in soils with a fragmentation model. *Soil Sci Soc Am J* 63:782–788. doi:10.2136/sssaj1999.634782x
- Botula YD, Cornelis WM, Baert G, Mafuka P, Van Ranst E (2013) Particle size distribution models for soils of the humid tropics. *J Soils Sediments* 13:686–698. doi:10.1007/s11368-012-0635-5
- Buchan GD (1989) Applicability of the simple lognormal model to particle-size distribution in soils. *Soil Sci* 147:155–161. http://journals.lww.com/soilsci/Abstract/1989/03000/APPLICABILITY_OF_THE_SIMPLE_LOGNORMAL_MODEL_TO_1.aspx
- Buchan GD, Grewal KS, Robson AB (1993) Improved models of particle-size distribution: an illustration of model comparison techniques. *Soil Sci Soc Am J* 57:901–908. doi:10.2136/sssaj1993.03615995005700040004x
- Burnham KP, Anderson DR (2002) Model selection and multi-model inference: a practical information-theoretic approach. Springer Publication, p 488. doi:10.1007/b97636
- Campbell GS (1985) *Soil physics with basic: transport models for soil-plant systems*. Elsevier, Amsterdam
- Erashin S, Gunal H, Kutlu T, Yetgin B, Coban S (2006) Estimating specific surface area and cation exchange capacity in soils using fractal dimension of particle-size distribution. *Geoderma* 136(3):588–597. doi:10.1016/j.geoderma.2006.04.014
- Everitt BS, Landau S, Leese M (2001) *Cluster analysis*. 4th edn. Arnold, Taylor and Francis Publication, p 237
- Fredlund MD, Fredlund DG, Wilson GW (2000) An equation to represent grain size distribution. *Can Geotech J* 37:817–827. doi:10.1139/t00-015
- Gee GW, Or D (2002) Particle size analysis. *Methods Soil Anal Part* 4:255–293
- Gimenez D, Rawls WJ, Pachepsky Y, Watt JPC (2001) Prediction of a pore distribution factor from soil textural and mechanical parameters. *Soil Sci* 166:79–88. <http://journals.lww.com/soilsci/toc/2001/02000>
- Gracza R (1962) Average particle size and specific surface of flours and air classified flour fractions. *Cereal Sci Today* 7:272
- Hagen L, Skidmore E, Fryrear D (1987) Using two sieves to characterize dry soil aggregate size distribution. *Trans Am Soc Agric Eng* 30(1):162–165. doi:10.13031/2013.30420
- Haverkamp R, Parlange JY (1986) Predicting the water-retention curve from a particle size distribution: 1. Sandy soils without organic matter. *Soil Sci* 142:325–339
- Hwang SI (2004) Effect of texture on the performance of soil particle-size distribution models. *Geoderma* 123:363–371. doi:10.1016/j.geoderma.2004.03.003

- Hwang SI, Hong SP (2006) Estimating relative hydraulic conductivity from log-normally distributed particle-size data. *Geoderma* 133:421–430. doi:[10.1016/j.geoderma.2005.08.006](https://doi.org/10.1016/j.geoderma.2005.08.006)
- Hwang SI, Lee KP, Powers SE (2002) Models for estimating soil particle-size distributions. *Soil Sci Soc Am J* 66:1143–1150. <http://www.envsci.rutgers.edu/~gimenez/SoilPhysics/Homework02/Hwang2002.pdf>
- Jaky J (1944) Soil mechanics. (In Hungarian.) Egyetemi Nyomada, Budapest
- Lima JEFW, Silva EM (2007) Selecao de modelos para o tracado de curvas granulometricas de sedimentos em suspensao em rios. *Rev Bras Eng Agric Ambient* 11:101–107. <http://www.scielo.br/pdf/rbeaa/v11n1/v11n1a13.pdf>
- Nemes A, Wosten JHM, Lilly A, Voshaar JHO (1999) Evaluation of different procedures to interpolate particle-size distributions to achieve compatibility within soil databases. *Geoderma* 90:187–202. doi:[10.1016/S0016-7061\(99\)00014-2](https://doi.org/10.1016/S0016-7061(99)00014-2)
- Nimmo J (2004) Porosity and pore size distribution. *Encycl Soils Environ* 3:295–303. http://www.rcamnl.wr.usgs.gov/uzf/abs_pubs/papers/nimmo.04.encyc.por.esa.pdf
- Pasikatan M, Steele J, Milliken G, Spillman C, Haque E (1999) Particle size distribution and sieving characteristics of first-break ground wheat. SABE paper MC99-129. St. Joseph, Mich.: ASABE
- Posadas AN, Giménez D, Bittelli M, Vaz CM, Flury M (2001) Multifractal characterization of soils particle-size distributions. *Soil Sci Soc Am J* 65:1361–1367. <http://www.dista.unibo.it/~bittelli/articoli/multifractal.pdf>
- Rizea N, Radu L, Rodica L, Store VM, Aldae MM (2009) Reaction buffering capacity of soils in the Zlatna area. *Sci Pap Ser A LII*, 70–75. http://www.agro-bucuresti.ro/images/Proiecte/2009/11_Reaction_buffering_capacity_of_soils_in_the_Zlatna_Area.pdf
- Rosin P, Rammler E (1933) The laws of governing the fineness of powdered coal. *J Inst Fuel* 7:29–36. <http://ci.nii.ac.jp/naid/10006950361>
- Shangguan W (2012) An investigation of soil particle-size distribution models for the conversion of soil texture classification from ISSS and Katschinskis to FAO/USDA system. In: Paper presented at PEDOFRACT VII Workshop on scaling in particulate and porous media: modeling and use in predictions, A coruna
- Shirazi MA, Boersma L (1984) A unifying quantitative analysis of soil texture. *Soil Sci Soc Am J* 48:142–147. doi:[10.2136/sssaj1984.03615995004800010026x](https://doi.org/10.2136/sssaj1984.03615995004800010026x)
- SPSS Inc. (2015) IBM SPSS Statistics for windows. Version 23
- Su YZ, Zhao WZ, Zhang TH (2004) Fractal features of soil particle size distribution and the implication for indicating desertification. *Geoderma* 122:43–49. doi:[10.1016/j.geoderma.2003.12.003](https://doi.org/10.1016/j.geoderma.2003.12.003)
- Vereecken H, Weynants M, Javaux M, Pachepsky M, Schaap MG, van Cenuchten MTH (2010) Using pedotransfer functions to estimate the van Genuchten-Mualem soil hydraulic properties: a review. *Vadose Zone J* 9(4):795–820. doi:[10.2136/vzj2010.0045](https://doi.org/10.2136/vzj2010.0045)
- Vipulanandan C, Ozgurel HG (2009) Simplified relationships for particle-size distribution and permeation groutability limits for soils. *J Geotech Geoenviron Eng*. doi:[10.1061/\(ASCE\)GT.1943-5606.0000064](https://doi.org/10.1061/(ASCE)GT.1943-5606.0000064)
- Weipeng W, Jianli L, Bingzi Z, Jiabao Z, Xiaopeng L, Yifan Y (2015) Critical evaluation of particle size distribution using soil data obtained with a laser diffraction method. *PLoS One* 10(4):e0125048. doi:[10.1371/journal.pone.0125048](https://doi.org/10.1371/journal.pone.0125048)
- Wraith JM, Or D (1998) Nonlinear parameter estimation using spreadsheet software. *J Nat Resour Life Sci Educ* 27:13–19. <https://www.agronomy.org/files/jnrlse/issues/1998/e96-29.pdf>
- Zhuang J, Jin Y, Miyazaki T (2001) Estimating water retention characteristic from soil particle-size distribution using a non-similar media concept. *Soil Sci* 166:308–321. http://journals.lww.com/soilsci/Abstract/2001/05000/ESTIMATING_WATER_RETENTION_CHARACTERISTIC_FROM.2.aspx