



Sharif University of Technology  
**Scientia Iranica**  
*Transactions A: Civil Engineering*  
<http://scientiairanica.sharif.edu>



# Estimation of van Genuchten SWCC model for unsaturated sands by means of the genetic programming

A. Taban<sup>a,\*</sup>, M. Mirmohammad Sadeghi<sup>b</sup>, and M.A. Rowshanzamir<sup>c</sup>

a. Department of Civil Engineering, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, P.O. Box 158-81595, Iran.

b. Laboratory of Soil Mechanics, Department of Water and Natural Environment, Isfahan Higher Education And Research Institute (IHEARI), Isfahan, Iran.

c. Faculty of Civil Engineering, Isfahan University of Technology, Isfahan, Iran.

Received 24 October 2016; received in revised form 25 December 2016; accepted 1 May 2017

## KEYWORDS

SWCC;  
 van Genuchten model;  
 Unsaturated soils;  
 Genetic programming;  
 Fitting parameters.

**Abstract.** The van Genuchten model (1980) is widely used for the description of the Soil-Water Characteristic Curve (SWCC) of a variety of soils. This study uses the Genetic Programming (GP) for the presentation of equations estimating the van Genuchten (vG) model fitting parameters for unsaturated clean sand soils. Moreover, this study uses the data derived from the valid dataset of Benson et al. [Benson, C.H., et al. “Estimating van Genuchten parameters **a** and **n** for clean sands from particle size distribution data”, *ASCE GSP*, pp. 234-235 (2014).], including 95 measured SWCCs in both drying and wetting phases. The data on the particle size distributions includes the fine-grain percentage (finer %),  $d_{60}$ ,  $d_{10}$ , besides the residual and saturated volumetric water contents ( $\theta_r$  and  $\theta_s$ ), as the GP model inputs of the set of terminals. As for the model outputs of the set of terminals, the fitting parameters for the vG model include **a** and **n**. The functions used in the GP training are ‘plus’, ‘minus’, and ‘times’ taken from the MATLAB default functions; ‘mydivide’ proposed by Silva [Silva, S. “GPLAB: A genetic programming toolbox for MATLAB”, Available at <http://gplab.sourceforge.net>, Coimbra, Portugal (2007)]; and some other new power functions included in this study. Accordingly, new equations are presented for the estimation of vG model fitting parameters for both forms of wetting and drying. Finally, to evaluate the accuracy of the proposed estimation equations, the GP results are evaluated and verified in different procedures.

© 2018 Sharif University of Technology. All rights reserved.

## 1. Introduction

Soil-Water Characteristic Curve (SWCC) is one of the widely used concepts in the earth-related sciences and, particularly, in the mechanics of unsaturated soils.

It describes the relationship between the amounts of water in the soil structure and the matric suction in the soil. Many analyses and numerical studies of unsaturated soils depend on the SWCC. Such expressions as the volumetric water content, the degree of saturation, and the normalized water content are used for the amount of water within the soil [1].

There are various methods for the measurement of SWCC, many of which are costly and time-consuming. However, they are different from each other in terms of the capability of measuring different ranges of the suction. Some measurement methods that directly

\*. Corresponding author. Fax: +98 6633243999  
 E-mail addresses: [am.taban@gmail.com](mailto:am.taban@gmail.com) (A. Taban);  
[msadeghi@ieht.ac.ir](mailto:msadeghi@ieht.ac.ir) (M. Mirmohammad Sadeghi);  
[mohamali@cc.iut.ac.ir](mailto:mohamali@cc.iut.ac.ir) (M.A. Rowshanzamir).

measure the SWCC are the hanging column, the Pressure Plate Extractor (PPT), the chilled mirror hygrometer, and the centrifuge, which directly measure SWCC through the procedure explained in ASTM D6836 [2]. In contrast, there is another group of experiments which measure the values indirectly, e.g., filter paper, heat dissipation, thermal conductivity sensor technique, chilled-mirror, hygrometer technique, etc. [3]. These methods of suction measurement measure different ranges of suction values. In addition to ASTM D6836 [2], more precise details of experiments, such as Pressure Plate Extractor (PPT), have been presented by many researchers, including Wang et al. (2016) [4], in their studies.

It should be noted that the SWCC can be measured under two conditions: wetting (when the water flows into the samples) and drying (when the water flows out of the samples). In this way, there may be two suctions in a single soil for a specific amount of saturation; this will be explained in more details in the next section.

Different models have been introduced in the field of SWCC in recent decades. Among the notable SWCC models, one can mention the ones proposed by Books and Corey (1964) [5], van Genuchten (1980) [6], and Fredlund and Xing (1994) [7]. More recently, tens of models, for example those proposed by Kosugi (1999) [8], Omuto (2009) [9], and Krishnapillai and Ravichandran (2012) [1], among others, have been proposed on the basis of early models. One of the most widely used SWCC models is the one proposed by van Genuchten (1980) [6]. This model can be used for a wide range of disturbed and undisturbed soils from the fine-grained to coarse-grained types of soil. The present study deals with the estimation of fitting parameters of vG model by means of the particle size distribution data and by means of the capabilities of Genetic Programming (GP).

## 2. The van Genuchten SWCC model

The van Genuchten model shows the relationship between the normalized water content and the suction in the form of Eq. (1):

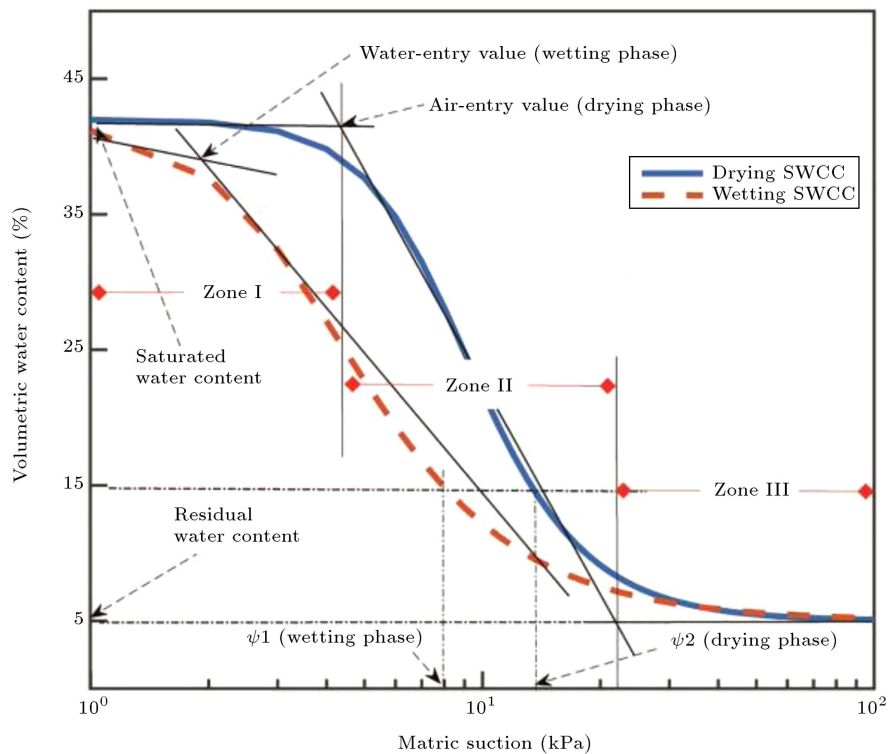
$$S = \frac{(\theta - \theta_r)}{(\theta_s - \theta_r)} = \frac{1}{[1 + (a\psi)^n]^m}, \quad m = 1 - 1/n, \quad (1)$$

where  $S$  is the normalized water content, which is a dimensionless parameter, expressed as  $S = \frac{(\theta - \theta_r)}{(\theta_s - \theta_r)}$ , and  $\theta$  is the volumetric water content. Indices  $r$  and  $s$  in  $\theta_r$  and  $\theta_s$  denote the residual and the saturated volumetric water contents, respectively.  $\psi$  is the suction in the unit of kPa, and  $a$  and  $n$  are the fitting model parameters. Parameter  $a$  is the value related to the air entry value,  $n$  is the value related to pore

size distribution, and  $m$  is the value related to the asymmetry of the model [6].

Figure 1 shows a picture of the typical SWCC related to the sand soil. This figure is for the understanding of many basic concepts of SWCC. It shows two curves: the curve of wetting phase and the curve of drying phase. As it was said above, two different values of suction are obtained (i.e.,  $\psi_1$  and  $\psi_2$ ) for a specific amount of volumetric water content (here,  $\theta = 15\%$ ). This clearly shows that there is a difference between the drying and wetting forms of the SWCC. Although the drying and wetting phases are different, many numerical analyses use only one of the phases for the description of SWCC [1,5-7,10]. However, the present study focuses on both phases for the determination of SWCC.

As Figure 1 shows, the air-entry value in the drying phase is related to the amount of suction, which is concurrent with the air entry into the largest pores in the unsaturated soil. Moreover, this figure shows the water entry value for the wetting phase. The water entry value, on the wetting SWCC, is defined as the matric suction at which the water content of the soil begins to significantly increase during the wetting process [11]. The residual water content is the amount of water where a big change in the suction is needed for the removal of water from the soil. The saturated water content is, in turn, the volumetric water content of the soil at its fully saturated state corresponding to the minimum value of the suction. Figure 1 also shows the three zones (i.e., Zone I, Zone II, and Zone III), which are known as capillary saturation zone, desaturation zone, and residual saturation zone [7,1]. As it was mentioned above, various studies have been conducted on different aspects of SWCC. In order to estimate the SWCC, Arya and Paris developed a model called the physicoempirical model in 1981 [12]. Their model managed to estimate SWCC by means of the Particle-Size Distribution (PSD), the particle density property, and the bulk density of the soil. Simms and Yanful (2002) [13] presented a method, which used the pore-size distribution (POSD) for the estimation of SWCC of compacted plastic soils. Their method took into consideration the influence of pore geometry on SWCC. Romero and Simms (2008) [14] dealt with the relationship between the soil microstructure and the SWCC. Ghanbarian et al. (2015) [15] studied the effects of sample dimensions on the SWCC in their fascinating research. Wang et al. (2016) [4], in turn, focused on the pore-size distribution (POSD) of soils with an emphasis on the role it could play in the SWCC. In recent decades, many efforts have been made for the determination of fitting parameters of the SWCC models. Most methods of SWCC estimations are based on the measurement of water content corresponding to the suction on the basis of soil properties data and particle



**Figure 1.** A typical SWCC and its specifications at the two phases of drying and wetting.

size distribution by means of such common statistical methods as the regression analysis. In order to estimate the fitting parameters of two SWCC models proposed by Brooks and Corey (1964) [5] and van Genuchten (1980) [6], some equations were developed by Rawls et al. (1991) [16] through statistical regressions. These equations could estimate the fitting parameters of van Genuchten by using the percentages of clay, sand, carbon, and the bulk density of soils. The use of such factors as organic matter would limit the use of equations developed by Rawls et al. (1991) [16]. However, these equations have not been designated for a specific soil type, and this is their significant weakness. In 2000, Zapata et al. [17] developed some equations for the estimation of the SWCC model fitting parameters proposed by Fredlund and Xing (1994) [7] through statistical regressions. These equations were divided into two distinct groups on the basis of Plastic Index (PI) of the soils. For the soils with  $PI > 0$ , Zapata et al. (2000) [17] proposed some equations for fitting parameters of the SWCC model proposed by Fredlund and Xing (1994) [7], which were a function of the product of pass percentage of sieve #200 ( $w$ ) and that of PI (i.e.,  $w \cdot PI$ ). However, for granular soils with ( $PI = 0$ ), the equations developed by Zapata et al. (2000) [17] were only functions of  $d_{60}$  (or the grain diameter equal to 60% of the passing in terms of weight or of mass (mm)), but these equations ignored the influence of such significant parameters as the pass percentage of sieve #200. Since the percentage of

fine grains has a significant effect on the behavior of unsaturated soil, there will be an inappropriate prediction for the proposed equations with ( $PI = 0$ ) if one ignores the pass percentage of sieve #200. Based on the characteristics of particle size distribution of the clean sands, Benson et al. (2014) [18] proposed separate equations for the estimations of the fitting parameters of van Genuchten SWCC model for both wetting and drying phases. These equations were just functions of the two parameters of  $d_{60}$  and  $C_u = d_{60}/d_{10}$  of clean sand soils. The procedure adopted by Benson et al. (2014) [18] (Benson's PTF) is as follows: Assuming  $C_u = 1$ , an initial value of a fitting parameter of van Genuchten SWCC model is first estimated by means of two statistical equations (that is,  $a_{1w} = 1.993d_{60}$  for the wetting phase and  $a_{1d} = 1.354d_{60}$  for the drying phase) proposed by Benson et al. (2014) [18]. The obtained initial  $a$  values are then corrected by means of a coefficient in the form of ( $N_a = 0.99C_u^{-0.54}$  for both drying and wetting phases). At the end, using the equations  $a_w = a_{1w}N_a$  and  $a_d = a_{1d}N_a$ , the final estimated values of the fitting parameters will be calculated for both wetting and drying phases, respectively. A similar procedure was also proposed by Benson et al. (2014) [18] for the estimation of the values of  $n$  fitting parameters of van Genuchten SWCC model for both wetting and drying phases.

Among the equations proposed for the estimation of fitting parameters of van Genuchten SWCC model, each had a serious weak point. The weak

points, of course, were separately specified when the equations were introduced. In addition to the weak points specified above, a common weak point in all the equations is that they use the old and ordinary statistical methods of estimation. Compared with the intelligent machine learning methods, these old statistical methods of estimation seemed to be very inefficient. However, some other efforts were based on the SWCC estimations by means of the machine learning methods [19,20]. By means of GP, therefore, the present study uses the particle size distribution data of clean sand soils, including fines%,  $d_{10}$ , and  $d_{60}$  in addition to using  $\theta_r$  and  $\theta_s$ , in order to present some equations for the estimation of fitting parameters **a** and **n** of the vG model. The necessary experiments for obtaining the particle size distribution are simple and economical, and they take less time than the direct and indirect SWCC experiments [21].

### 3. An outline of GP

GP is a generalized branch of the Genetic Algorithm (GA). Inspired by the Darwinian Theory and the Mendelian genetics, GP is capable of solving difficult mathematical problems. GP is different from GA in the way it deals with individuals. In GA, individuals are in the forms of linear strings with fixed lengths, but in the GP, they are nonlinear with different sizes, and they are known as the tree structure. Tree structures consist of a set of functions (such as the mathematical functions used in equations) and a set of terminals (that is, variables of a problem and constants). Before GP is run, the user should do the following basic procedures:

- Determining the set of terminals and the random constants;
- Determining the set of functions used in the equations;
- Selecting the fitness functions for the evaluation of equations;
- Determining the parameters that control the program running, such as the population size, and other details;
- Determining the finishing criterion and the results of program running.

The step-by-step procedure for the running process of the GP is as follows:

- Generation of an initial population from the formulae developed from random combination of the set of functions (the mathematical functions used in the formulae) and set of terminals (variables of a problem and the constants);
- Evaluation of each of the individuals by means of the fitting functions;

- Generation of a new population of formulae, which is followed by a progressive trend for the generation of a new population:

- One of the genetic operators, i.e., crossover, mutation, or reproduction, will be selected (these three genetic operators are the main genetic operators, which are used in GP);
- An appropriate number of individuals in the present population will be selected (the selection of individuals will be on a random basis in which the individual with a better fitting will be preferred to the unfit ones, and this does not mean that the unfit individuals have to be removed);
- The selected genetic operator will be used for the generation of children;
- The generated children (i.e., the new formulae) will go into a new population;
- The resulting model will be evaluated by means of the fitting function.

- Repeating Step 3 until the maximum number of generations is achieved [22,23].

In order to train the GP, one has to select and classify the appropriate datasets, which is done in the following section. Figure 2 shows the flowchart of GP solution process. This figure gives a better knowledge of how GP can solve a problem.

### 4. Dataset

This dataset includes the data of 95 measured SWCCs of fine-to-coarse-grained clean sand soils with different particle size distributions and properties. The measured SWCCs were collected and classified from 17 different reliable sources and, for this reason, the dataset enjoys an appropriate diversity. Precise soil gradation experiments were conducted on the samples, and the values were recorded in terms of  $F_{200}$ ,  $d_{10}$ ,  $d_{60}$ , and  $C_u = d_{60}/d_{10}$  in accordance with ASTM D422 [21]. In order to precisely measure the SWCC of each sample, appropriate standard experiments were performed (ASTM D6836) [2]. For both phases of drying and wetting, separate SWCCs were measured and recorded for each sample of sand soil. In order to fit the above-mentioned values of suction and its corresponding volumetric water content, Eq. (1) was used [18]. In the dataset collected by Benson et al. (2014) [18], some samples have just the drying data, some others have just the wetting data, and some others have the two types of data. For this reason, the present study divided Benson's initial dataset into two parts, namely, the drying phase and the wetting phase. Inspired by different reliable studies, for example, Johari et al. (2006) [19], Garg et al.

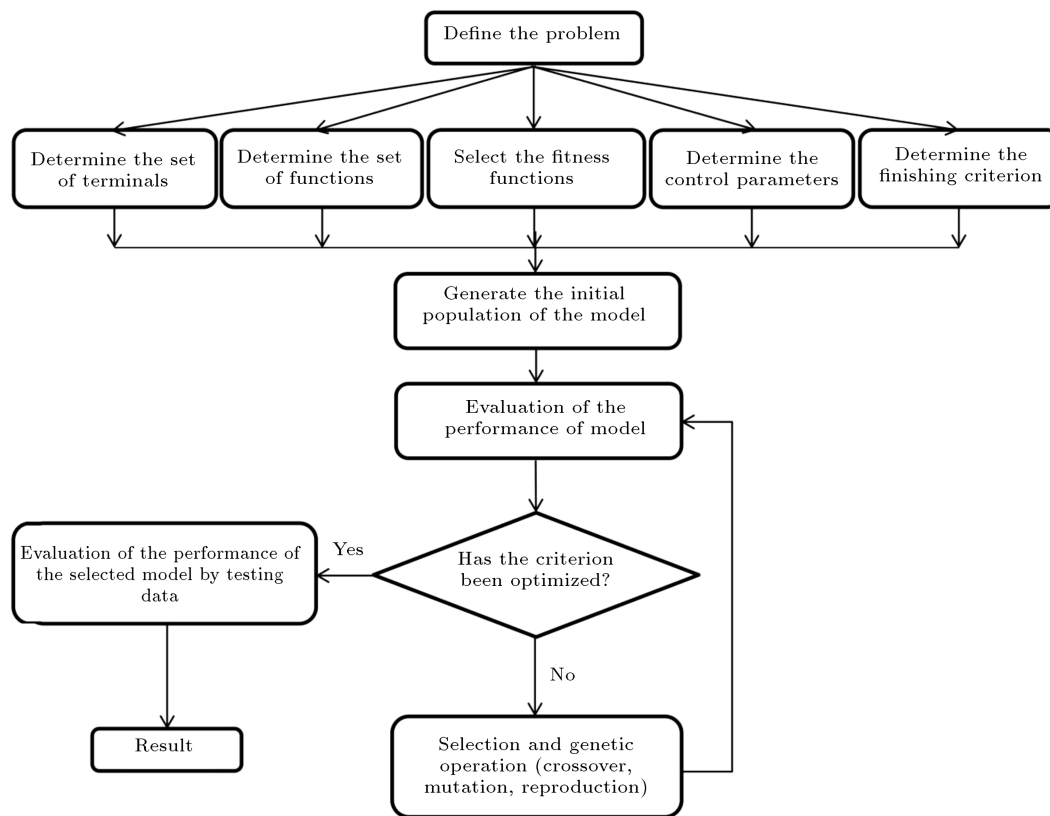


Figure 2. Flowchart of GP solution process.

Table 1. Ranges of the GP model's set of terminals.

SWCC phase	GP inputs					
	$d_{60}$ (mm)	$d_{10}$ (mm)	$C_u^*$	Fines%	$\theta_s$	$\theta_r$
Drying phase	0.14-3.5	0.02-4.83	0.72-30	0-28	0.3-0.55	0-0.13
Wetting phase	0.05-1.52	0.02-1.07	1.0-1.07	0-28	0.33-0.44	0-0.11

\*: Note that  $C_u$  parameters are not considered to be the inputs of GP.

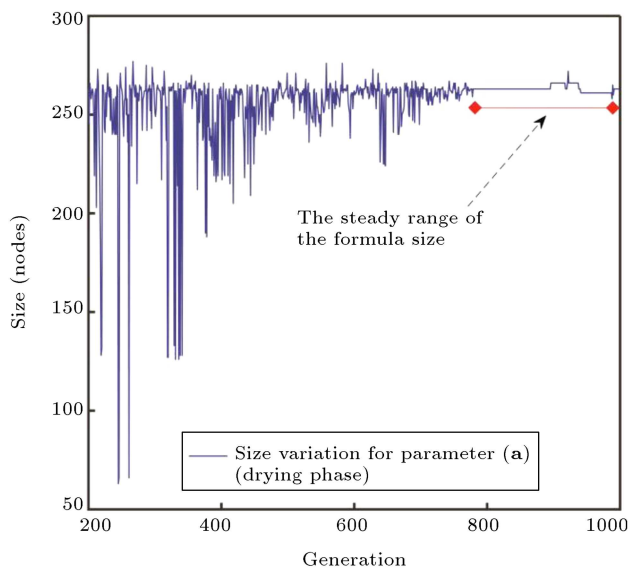
(2014) [20], and Parasuraman et al. (2007) [24], which selected a database and, then, randomly differentiated it into two distinct datasets, i.e., training and testing datasets, the present study divided Benson's database into two distinct datasets consisting of the training and testing ones. In order to evaluate the GP results, about 20% of the initial dataset was separated for evaluation purposes, as shown in Table 3. Clearly, there is no use of any sample data of testing dataset for the training in the present study. For almost all similar studies on development engineering areas, and especially for unsaturated soils, which have considerable limitations of valid databases, it is very common to use the above-mentioned procedure.

Table 1 shows the range of datasets used as GP inputs for the two drying and wetting phases of SWCC, respectively. Given that these GP equations are approximations, Table 1 can give an appropriate perspective for reliable application. By means of the

above separate datasets, the GP was developed for obtaining the vG SWCC fitting parameters.

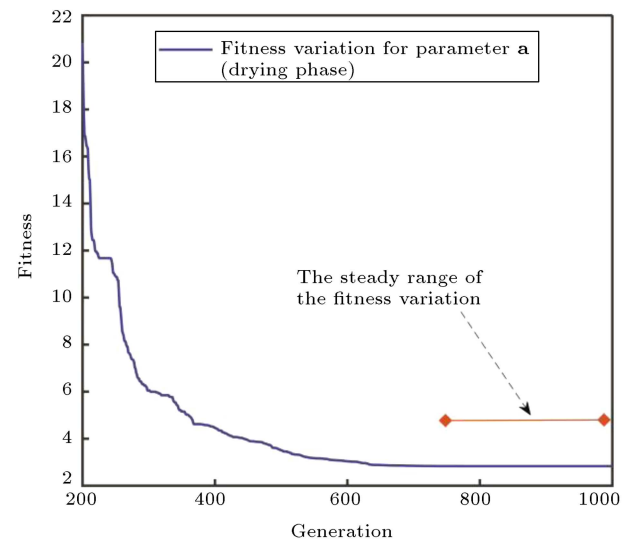
## 5. The running of GP for van Genuchten fitting parameters

After the initial dataset was divided into two separate phases, GP running was also divided into separate wetting and drying phases. Five values ( $d_{10}$ ,  $d_{60}$ ,  $F_{200}$ ,  $\theta_r$ , and  $\theta_s$ ) were used as parts of the set of terminals. Based on previous experiences, moreover, the random function was used for the generation of some fractions and numbers. In order to define the set of functions, the default MATLAB functions (i.e., minus, plus, times) were used. In order to define the division function, the function proposed by Silva (2007) [23] ('mydivide') was used [23]. One of the characteristics of the present study is that it defines new functions for the power under "tpower2" up to



**Figure 3.** Variation of size for fitting parameter **a** (drying phase).

“tpower9”, by means of which the terms of generated formulae will be limited to the powers of the integers from 2 to 9 excluding fractions. These new powers are proposed in the present study in order to avoid the complications of the formulae which use the “mypower” function proposed by Silva (2007) [23]. This became possible after many trials and errors. The new power functions could decrease the complexities and, consequently, the sizes of nodes. The initial population in the present study was assumed to be 100 for the GP running. This selection of initial population was based on the expected decrease in nodes and the decrease in complexity of the generated formula. The present study used the crossover and mutation operators in the process of estimating the **a** and **n** fitting parameters for both phases of drying and wetting. The depth of trees was limited to 17, and the mode of lexic tour was used for the selection of reproduction [23]. After some trials and errors, it became known that the increase in initial population could considerably decrease fitness and increase the precision of the generated formula, although it increased the size of the formula. For this reason, the initial population was designated to be 1000, and the GP was let to run up to 5000 generations. As Figure 3 shows, the fitness was almost fixed at about Generation 770. Clearly, this had a good agreement with the appropriate fixation of the formula size at Generation 770 (Figure 4). For this reason, the run was stopped at Generation 1000 for the process of estimating the **a** fitting parameter at the drying phase. Finally, the best formula with the optimal fitness was selected among the generated formulae in the last 230 generations (that is, generations 770 up to 1000). (Note that both figures show the ranges of variation of size and fitness for generation 200 and beyond. For the



**Figure 4.** Variation of fitness for fitting parameter **a** (drying phase).

calculation of other fitting parameters, a similar trend was adopted. However, the space was limited and, therefore, it is not possible to present the detailed curves, here, showing the variations of size and fitness of other parameters.)

## 6. Results and discussion

### 6.1. The GP model equations for *vG* fitting parameters *a* and *n*

Following the above procedure, separate equations were finally generated by GP for the estimation of each of the **a** and **n** values at the drying and wetting phases. The generated equations were in the forms of lengthy prefix equations including such functions as minus, plus, times, tpower2 up to tpower9, ‘mydivide’, and different constant numbers, which were ultimately presented in the forms of infixes and repeatedly simplified into Eqs. (2) and (3) for the phase of drying and Eqs. (4) and (5) for the phase of wetting (Table 2). As it was said above, the equations seem to be rather lengthy, but one can easily use them through different pieces of software for the estimation of clean sand SWCCs.

### 6.2. Verification of the results of the proposed GP model equations

This section deals with the results of the proposed equations of the GP model (Table 2). To this end, the predicted results of the GP model will be presented in Table 3 for the selected testing datasets. In addition to the **a** and **n** values obtained from precise measurements, the datasets of the two phases include the predicted values of PTF (pedotransfer functions) equations developed by Benson (2014) [18]. Table 3 shows the results of the SWCC fitting parameters measured in the laboratory. The table also shows

**Table 2.** The proposed GP model equations for the vG fitting parameters.

GP of van Genuchten SWCC fitting parameters	
$\begin{aligned} \mathbf{a}_d = & ((0.09305 + (((.1^8 * (((A * (((3 - A)^3) + (B - (.5/A + E)^3)) + ((A * (B - .0001) * (A - B)) \\ & + (B - 0.125)) + (B - (0.50674/A)^4)) * .1 * (A - B))) + (.1 * ((((((3 - A)^3 + A) \\ & * ((A - B + 0.10008) * .5)) + .125 - E + B) * .3))^8 * (0.10096 * (((3 - A)^3 + 5) \\ & * (((4 - E + A) * B * C) * B * C)) + ((0.50674^3 * (A + .1 - 2 * E - E/D)) * C)))))) * D)) \\ & / (D^3 + (((.1 * A^2)^8 * .1 * (A - B))^8 * C) / (((A * (B * (D - B + (.1/B)^4 - .1))) + 3) + 4))) / .1^6)) \\ & + (((.1 * A^2)^8 * (.1 * (A - .1 - B))) + (B - (((D - (E + (B - .1)/A) \\ & + (.5^3 * (A - (.1 + (0.10008 * (A - 2 * B - (B - .1)/.1 + 1))))))^4))) \end{aligned} \quad (2)$	
$\begin{aligned} \mathbf{n}_d = & ((0.52928 / (((.1 / ((A - 2 * .5^7 - 0))) / 0.34447^2) - B/A) + (D))) \\ & + (((E / ((0.37562^3) / D))^8 + D) / (((B/A / (((.1 / (B/A / .1)) / (((E / (0.053/D))^8 + D)) - B/A) + D)) \\ & - ((D / (((D / (B/A / .1)) / .1) - B/A) / A) / B) / A) + (0.37562 + (((B/A / .1) - ((.1/A) / 0.37562)) \\ & + (((.1 / ((2 * D^3 / A) / .1)) / .1) - B/A) + D))) + ((1 / (((.1 / ((.1 / ((B/A / .1) / .1) - B/A) + D)) - E) \\ & + (D + ((0.162 - ((.1/A) / ((1/C)^56)) + (A / ((.1^4 - A - ((.1 + (C/B))^2))^9))) / (((D / ((B/A / .1)) / .1) \\ & - B/A) + B/D)) - E) + (A + ((B - (((D / (B/A / .1)) / .1) - (B/A))) + (.1 / ((E / .1) / (((0.77913 * (A - E)) \\ & + ((C^7 - .5) * B)) * (A - E)) + A^7)))))) + (((B / (A - 2 * D^7)) / .1) - B/D) + (E / (0.053/D^2)))) \end{aligned} \quad (3)$	
$\begin{aligned} \mathbf{a}_w = & (B + (((C * (((2 * B^2 * (A - ((B + C * E^2)^2)))^3 + 2 * C * B^2 + B + .1 + D^6))^2))^3 \\ & + ((C * B^2 + (B + 0.1016 + (0.0979 * ((B + (D^4) * (-A + 0.598 + D * 2) + (B + .1 + 2 * E))^2))) \\ & + (C * ((D - A) * ((B + ((B + D^7)^2 + (E * (((A - 1) / (1 + A)))^3 + (B + B)))) \\ & + (((D * (2 + (D * B)))^8 * (2 + 2 * E^2 * (B^2 - 0.16726)))^8)^3))^2))) \\ & + ((((((A - (-A + .1 + D * 3 + (E * (((D - 1) / (D - A + B))^3 + ((A * (E - C) - 8 + E) \\ & + ((A - 1) / (D + 0.43177))^3)))))) / (D + C))^3 * E)) * .1))) \end{aligned} \quad (4)$	
$\begin{aligned} \mathbf{n}_w = & ((E - .1/A + (E + ((E - B/D + (E + (B^5 - B + (((.5 - B)^3) + (((E + 0.00195) / (1 - B)) \\ & - (B / (((E / 0.9996) * B^2) - B/A))) + (.1 + B/A / .1) + ((E / 0.9996)^3) + E))) \\ & + (((E / 0.9996) * A) * ((((-A - ((A - B/A) / (A - B/A/D))))^3 + (((E/B)^5 + ((E / (1 - 2 * E)) / .1) \\ & * ((E/B)^5 - C)) + (B / .1^3))) + (((.25 / (A - (B/A / (D/2/E)))) - ((E / .1)^3 / (A - (E/B)^5))) / 4) \\ & - (((((E + D^5) / (0.9996 - (C * (E / 1.1112)))) * ((A - ((E + D^5) / B) / A)) - D) \\ & / (((B / (B / (A - ((B/A) / A))) / (A - (B/A/D))) - (B/A/D)))) - D))) + E - B \end{aligned} \quad (5)$	

Note:  $A = d_{60}$  in mm,  $B = d_{10}$  in mm,  $C$  = Fines in percentage,  $D = \theta_{s'}$ , and  $E = \theta_{r'}$ .

the results predicted by the GP model equations. Moreover, it shows the results predicted in Benson's (2014) [19] PTF for drying and wetting phases.

Figures 5 and 6 show the evaluation of the results predicted by the GP model equations versus the measured values in the two phases. The evaluation was conducted on the basis of the training and testing datasets. In each figure, the values of determination coefficient ( $R^2$ ) have been reported for each evaluation. The comparisons presented in the figures show that there is a relatively good agreement between the results predicted by the GP and the values measured in both phases of drying and wetting for training and testing datasets. It is worth mentioning that the equations

presented by the GP for the drying phase are to some extent more precise than those presented for the wetting phase, and this may be due to the fact that the drying dataset is more extensive than the wetting dataset in the present study. For a more precise evaluation, some comparisons were made between the results of Benson's PTF and the measured values. Table 4 shows the evaluation results of Benson's PTF and those of the GP model for the two datasets of training and testing. This comparison indicates that the results obtained from the GP equations are more precise than those predicted by Benson's PTF. This is particularly true for the estimation of  $\mathbf{n}$  parameter (in both phases).

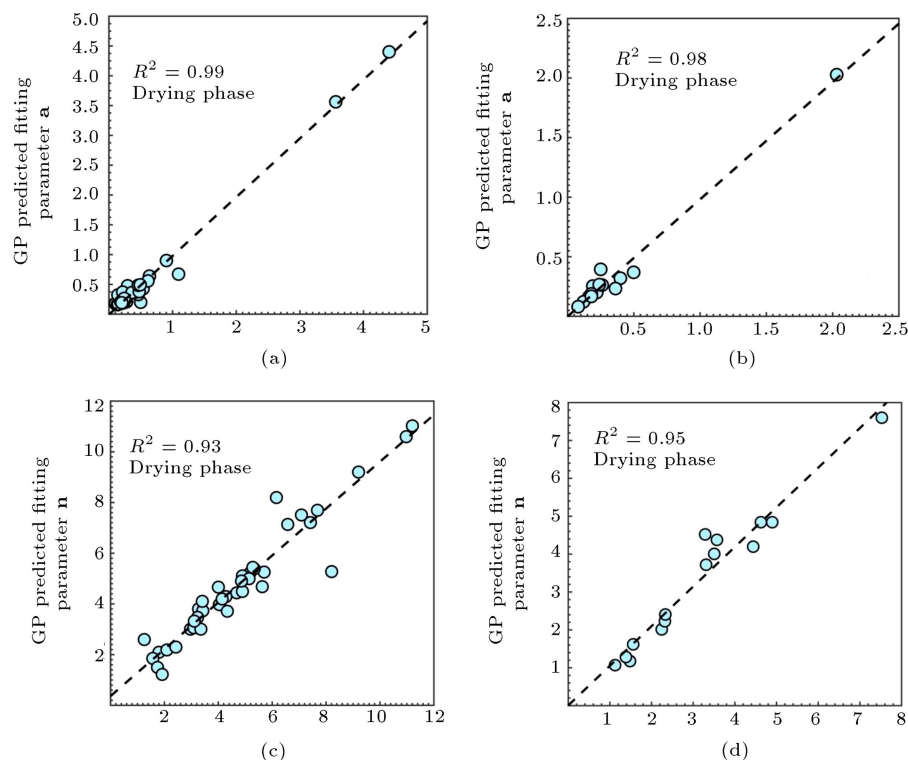
**Table 3.** The dataset of testing input parameters, and the measured values, Benson's PTF prediction, and GP model predictions for **a** and **n** fitting parameters.

Sample name	SWCC phase	Input					Measured		Benson's PTF		GP predicted	
		$d_{60}$	$d_{10}$	$F_{200}$	$\theta_s$	$\theta_r$	<b>a</b>	<b>n</b>	<b>a</b>	<b>n</b>	<b>a</b>	<b>n</b>
							( $\text{kPa}^{-1}$ )		( $\text{kPa}^{-1}$ )		( $\text{kPa}^{-1}$ )	
NW4-coarse sand [25]	Drying	0.46	0.21	1.5	0.374	0.031	0.4	3.57	0.41	4.39	<b>0.32</b>	<b>4.37</b>
NW11-silty sand [25]	Drying	0.19	0.07	11	0.434	0.098	0.16	4.44	0.15	4.91	<b>0.16</b>	<b>4.19</b>
NW19-fine sand [25]	Drying	0.14	0.05	22	0.415	0	0.12	3.31	0.11	5.02	<b>0.12</b>	<b>3.72</b>
NW23-fine-coarse [25]	Drying	0.21	0.09	7	0.381	0.11	0.18	3.29	0.18	4.88	<b>0.18</b>	<b>4.51</b>
Howe concrete sand [26]	Drying	1.09	0.22	1.1	0.349	0.004	0.25	1.13	0.62	3.26	<b>0.39</b>	<b>1.08</b>
Mk96TP15GH 15-16 [26]	Drying	3.5	0.4	4	0.341	0.009	2.03	1.48	1.46	1.1	<b>2.03</b>	<b>1.17</b>
Kidston gold mine [27]	Drying	0.17	0.02	28	0.395	0.05	0.08	2.24	0.07	4.65	<b>0.08</b>	<b>2.01</b>
Esperance soil [28]	Drying	0.3	0.16	0.6	0.39	0.111	0.19	2.32	0.28	6.51	<b>0.25</b>	<b>2.23</b>
Sand [29]	Drying	0.25	0.14	1	0.4	0	0.36	3.51	0.25	7.42	<b>0.24</b>	<b>4.00</b>
Wraith soil 2122 [18]	Drying	0.24	0.11	1	0.45	0.13	0.22	7.53	0.21	4.66	<b>0.20</b>	<b>7.60</b>
Boise sand [18]	Drying	1.5	0.2	2.3	0.373	0.021	0.5	1.56	0.68	2.66	<b>0.37</b>	<b>1.61</b>
Phoenix sand [18]	Drying	0.2	0.1	2.5	0.373	0.027	0.18	4.62	0.19	6.05	<b>0.18</b>	<b>4.84</b>
Green roof sand [18]	Drying	2.2	0.15	5	0.544	0	0.26	1.39	0.69	1.83	<b>0.26</b>	<b>1.28</b>
Newcastle sand [18]	Drying	0.31	0.17	0.1	0.421	0.015	0.24	2.33	0.30	6.97	<b>0.26</b>	<b>2.41</b>
Bluebench sand [18]	Drying	0.17	0.08	6.6	0.42	0.03	0.18	4.9	0.15	5.22	<b>0.17</b>	<b>4.84</b>
Concrete sand [26]	Wetting	0.68	0.2	1	0.35	0.05	0.13	2.4	0.67	3.5	<b>0.19</b>	<b>2.25</b>
Glass beads (0.2 mm) [30]	Wetting	0.2	0.2	0	0.366	0.022	0.25	9	0.38	7.67	<b>0.52</b>	<b>8.72</b>
Sublayer (0.5-0.71) [31]	Wetting	0.62	0.52	0	0.415	0	1.3	6.23	1.07	6.16	<b>1.28</b>	<b>6.28</b>
Sublayer (0.355-0.5) [31]	Wetting	0.44	0.37	0	0.446	0	0.78	7.35	0.76	6.57	<b>0.91</b>	<b>6.72</b>
NW2-coarse sand [25]	Wetting	0.84	0.24	2	0.392	0.067	0.61	8.5	0.81	3.31	<b>0.66</b>	<b>7.73</b>
NW11-silty sand [25]	Wetting	0.19	0.07	11	0.434	0.098	0.28	4.39	0.21	4.17	<b>0.32</b>	<b>4.48</b>
NW19-fine sand [25]	Wetting	0.14	0.05	22	0.415	0	0.28	2.41	0.16	4.24	<b>0.28</b>	<b>2.55</b>
Kidston gold mine [27]	Wetting	0.17	0.02	28	0.395	0.05	0.13	2.11	0.1	4.022	<b>0.16</b>	<b>2.11</b>

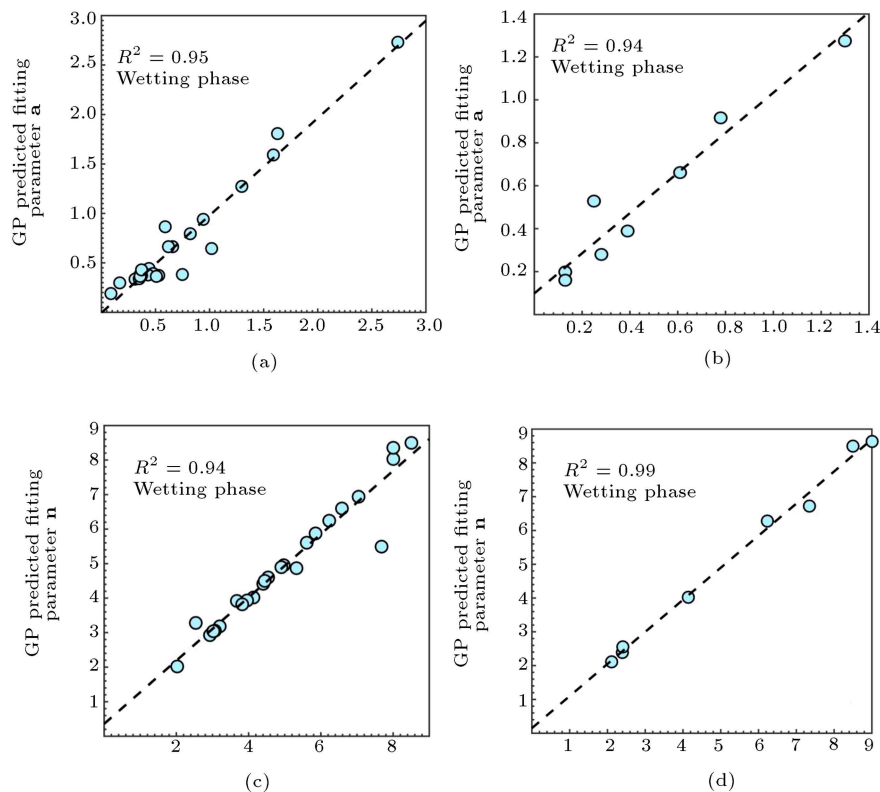
**Table 4.** Values of coefficient of determination ( $R^2$ ) and Mean Squared Error (MSE) on the basis of training and testing datasets for the comparison between the results of GP model predictions versus the measured values and those of Benson's PTF versus the measured values.

vG fitting parameter	SWCC phase	Benson's PTF ( $R^2$ )		Benson's PTF (MSE)		GP predicted ( $R^2$ )		GP predicted (MSE)	
		Training dataset	Testing dataset	Training dataset	Testing dataset	Training dataset	Testing dataset	Training dataset	Testing dataset
$a_d$	Drying	0.928	0.799	0.0775	0.0469	0.985	0.979	0.0108	0.0043
$n_d$	Drying	0.407	0.169	3.7477	5.5333	0.931	0.949	0.4951	0.1909
$a_w$	Wetting	0.896	0.892	0.0685	0.0527	0.948	0.944	0.0175	0.0140
$n_w$	Wetting	0.485	0.077	1.6312	4.6976	0.935	0.991	0.2134	0.1401





**Figure 5.** GP predicted values of vG model versus measured values of vG model in the drying phase: (a) Fitting parameters **a** (the training dataset), (b) fitting parameters **a** (the testing dataset), (c) fitting parameters **n** (the training dataset), and (d) fitting parameters **n** (the testing dataset).

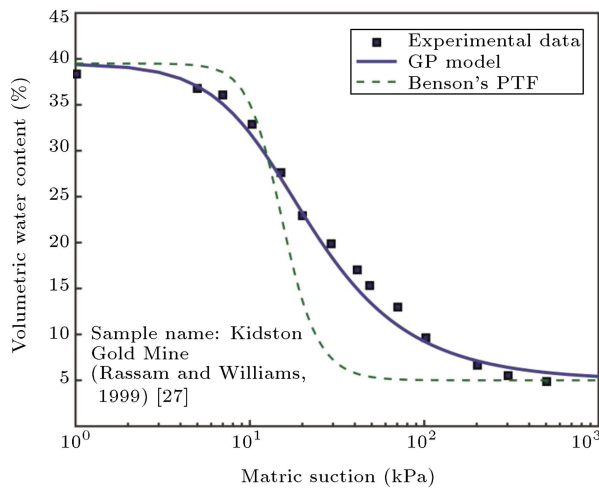


**Figure 6.** GP predicted values of vG model versus measured values of vG model in the wetting phase: (a) Fitting parameters **a** (the training dataset), (b) fitting parameters **a** (the testing dataset), (c) fitting parameters **n** (the training dataset), and (d) fitting parameters **n** (the testing dataset).

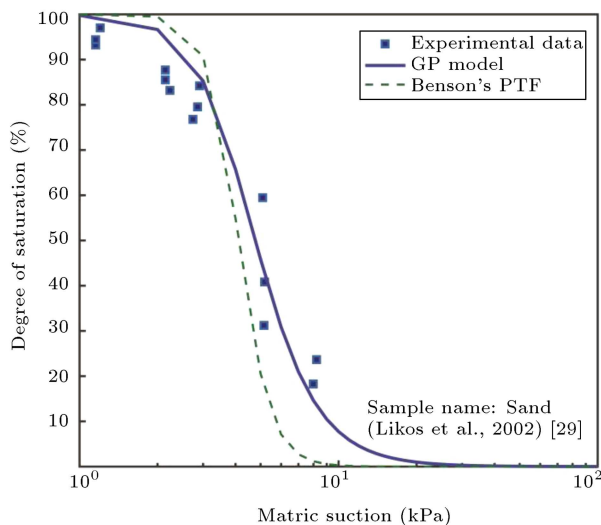
### 6.3. Complementary verification of the results of GP equations

So far, each fitting parameter has been separately evaluated in the previous sections. This section deals with the effects of the vG model fitting parameters on the shape of SWCC. For more verification of the equations presented in the GP for the vG fitting parameters, the results of equations are compared with the SWCC experimental data presented in some well-known references. In Figures 7 to 10, one can see the results of GP equations compared with the SWCC experimental data in those samples. These figures also show the results of Benson's PTF.

As Figure 7 shows, the results of equations presented in the GP in the drying phase have a very good

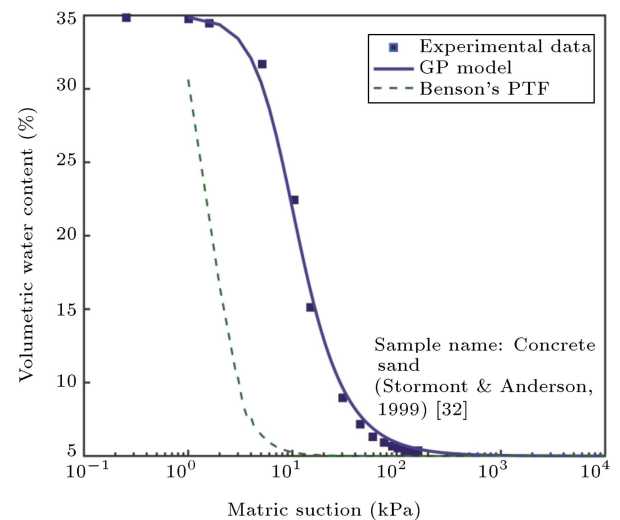


**Figure 7.** SWCC predicted by the GP model equations and Benson's PTF prediction compared with the experimental data for Kidston Gold Mine sample.

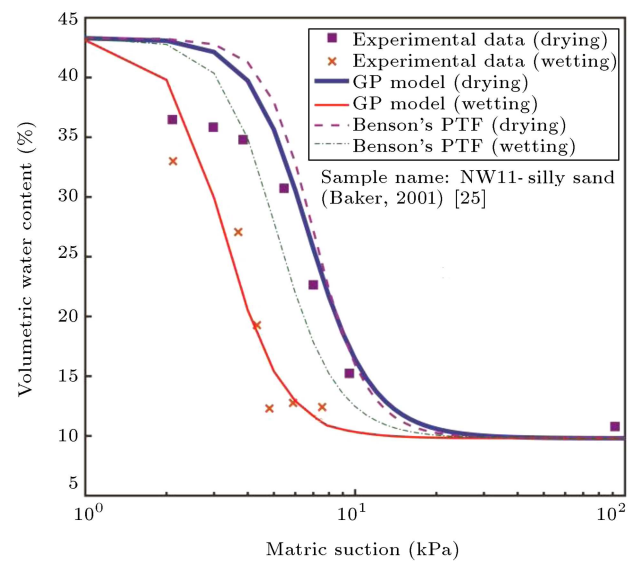


**Figure 8.** SWCC predicted by the GP model equations and Benson's PTF prediction compared with the experimental data for sand sample.

agreement with the results obtained from the SWCC experimental data, and this is very interesting in that the sample has a relatively high percentage of fine grains ( $F_{200} = 28\%$ ). Figure 7 shows that although Benson's PTF is relatively precise in the prediction of parameter  $\alpha$ , it does not have a good agreement with the experimental data since the parameter  $n$  predicted by Benson's PTF does not agree with the experimental data (the measured values in Table 3). The disagreement is more pronounced in Zone I (Figure 1). Figure 8 shows the comparison between the predicted results of the GP equations and the SWCC experimental data of a clean sand soil sample with a low percentage of fine grains. This figure also shows a relatively



**Figure 9.** SWCC predicted by the GP model equations and Benson's PTF prediction compared with the experimental data for concrete sand sample [32].



**Figure 10.** SWCC predicted by the GP model equations and Benson's PTF prediction compared with the experimental data for NW11-silty sand sample.

good agreement between the results obtained from the GP equations and the measured values. (Note that Figure 8 shows suction variations versus the degree of saturation, rather than the volumetric water content.)

Figure 9 shows the comparison between the results predicted by the GP model equations and those obtained from the experimental data in the wetting phase. This figure shows that there is a relatively good agreement between the results predicted by the GP model equations and those obtained from experimental data. Clearly, the figure shows that the Benson's PTF results do not agree with the experimental data. Figure 10 shows the comparison between the results of the experimental data in the drying and wetting phases for a silty sand sample. In this figure, the results predicted by the GP model equations for the SWCC estimation show that there is a relatively good agreement between the results predicted by the GP model equations in the two drying and wetting phases and the experimental data. In this figure, one can see a relatively good agreement between the results predicted by Benson's PTF and the experimental data in the drying phase. However, there is no good agreement between Benson's PTF predictions and the experimental data in the wetting phase. Moreover, the figure shows that there is no agreement between Benson's PTF predictions and experimental data in the drying phase by the few values of suction in Zone I (Figure 1). To some extent, this is true for the comparison between the results predicted by the GP model equations and the experimental data. However, there is a good agreement between the results predicted by the GP model equations and the experimental data in the wetting phase. (Note that the 'matric suction (kPa)' in the present study has been used instead of 'tension head (cm)' in the original reference.)

## 7. Conclusion

Given that SWCC measurements are costly and time-consuming, researchers are increasingly paying attention to the equations which estimate the SWCC. Using the GP capabilities, the present study gave some equations for the estimation of fitting parameters of vG SWCC. The characteristic feature of the present study was that it used GP as a very powerful instrument for solving complicated mathematical problems. The selection as well as the classification of an appropriate dataset can have considerable effects on the whole process of GP training. To this end, the present study selected its dataset from reliable sources, and it distinguished between the two phases of drying and wetting. The initial population was selected to be 1000 through trials and errors, and the operators were designated to be crossover and mutation. Moreover, new power functions (i.e., 'tpower2' up to 'tpower9')

were defined for the GP training. When one compares the results predicted by the GP equations with those obtained from the precise experiments, and compares the GP results with those predicted by Benson's PTF, one will see that there is a relatively good agreement between the proposed GP equations and the real experimental data. This agreement is higher for the SWCC prediction in the drying phase than in the wetting phase. A weakness of the proposed GP equations was that they were relatively lengthy equations as compared to other methods of estimation, such as the PTF developed by Benson (2014) [18]. Generally speaking, one can outline the advantages of the equations proposed in the present study over those presented in previous studies as follows: the use of easy-to-obtain soil parameters, no use of old statistical regression methods, the application of GP machine learning method whose efficacy has been proven in similar studies, the consideration of fine-grain percentage in the soil (as compared to the methods proposed by both Zapata et al. (2000) [17] and Benson et al. (2014) [18] for sand soils), and the focus on a specific class of soil (that is, clean sand). These advantages will increase the accuracy of the GP equations proposed in the present study. Unlike the methods proposed by Rawls et al. (1991) [16] and Zapata et al. (2000) [17], GP equations take both phases of wetting and drying into consideration. Moreover, the GP equations are more appropriate than those presented by Benson et al. (2014) [18] for the prediction of fitting parameters of the SWCC, and this is particularly true for parameter  $n$  of the soils with high  $C_u$ . Moreover, one should expect that the equations work for certain ranges such as those mentioned in the text, for example, for clean sand soils or the sand soils with low fines, and they may not necessarily work for every range of soil sample properties. This does not mean that the GP model equations are not able to eliminate the limitations of previous equations. It has indeed resolved many limitations of other equations. However, it is a fact that the GP model equations have a defined capability, and that one should not expect the proposed model to be efficient for any type of soil.

## Acknowledgment

The authors like to express their gratitude for the invaluable guidance freely offered by Professor van Genuchten. Also, the authors declare that there is no conflict of interests regarding the publication of this paper.

## References

1. Krishnapillai, S.K. and Ravichandran, N. "New soil-water characteristic curve and its performance in the

- finite-element simulation of unsaturated soils”, *Int. J. Geomech.*, **12**(3), pp. 209-219 (2012).
2. “ASTM standard D 6836-02, test methods for determination of the soil-water characteristic curve for desorption using a hanging column, pressure extractor, chilled mirror hygrometer, and/or centrifuge annual book of ASTM standards”, **04.08**, *ASTM International*, West Conshohocken, PA (2003).
  3. Nam, S., Gutierrez, M., Diplas, P., Petrie, J., Waylance, A., Lu, N., and Munoz, J.J. “Comparison of testing techniques and models for establishing the SWCC of riverbank soils”, *Eng. Geol.*, **110**(1-10), pp. 33-47 (2009).
  4. Wang, M., Pande, G.N., Kong, L.W., and Feng, Y.T. “Comparison of pore-size distribution of soils obtained by different methods”, *Int. J. of Geomech.*, **17**(1), pp. 1-6 (2017).
  5. Brooks, R. and Corey, A. “Properties of porous media affecting fluid flow”, *J. Irrigation and Drainage Division, ASCE*, **92**(2), pp. 61-88 (1966).
  6. van Genuchten, M. “A closed-form equation for predicting the hydraulic conductivity of unsaturated soils”, *Soil Science Society of America J.*, **44**(5), pp. 892-898 (1980).
  7. Fredlund, D.G. and Xing, A. “Equations for the soil-water characteristic curve”, *Can. Geotech. J.*, **31**(4), pp. 521-532 (1994).
  8. Kosugi, K. “General model for unsaturated hydraulic conductivity for soils with lognormal pore-size distribution”, *Soil Sci. Soc. Am. J.*, **63**, pp. 270-277 (1999).
  9. Omuto, C.T. “Biexponential model for water retention characteristics”, *Geoderma.*, **149**(16), pp. 235-242 (2009).
  10. Frydman, S. and Baker, R. “Theoretical soil-water characteristic curves based on adsorption, cavitation, and a double porosity model”, *Int. J. Geomech.*, **9**(6), pp. 250-257 (2009).
  11. Yang, H., Rahardjo, H., Leong, E.C., and Fredlund, D.G. “Factors affecting drying and wetting soil-water characteristic curves of sandy soils”, *Canadian Geotechnical Journal*, **41**, pp. 908-920 (2004).
  12. Arya, L.M., and Paris, J.F. “A physicoempirical model to predict the soil moisture characteristic from particle-size distribution and bulk density data”, *Soil Sci. Soc. Am. J.*, **45**(6), pp. 1023-1030 (1981).
  13. Simms, P.H. and Yanful, E.K. “Predicting soil-water characteristic curves of compacted plastic soils from measured pore-size distributions”, *Géotechnique*, **52**(4), pp. 269-278 (2002).
  14. Romero, E. and Simms, P.H. “Microstructure investigation in unsaturated soils: a review with special attention to contribution of mercury intrusion porosimetry and environmental scanning electron microscopy”, *Geotechnical and Geological Engineering*, **26**(6), pp. 705-727 (2008).
  15. Ghanbarian, B., Taslimitehrani A., Dong, G., and Pachepsky, Y.A. “Sample dimensions effect on prediction of soil water retention curve and saturated hydraulic conductivity”, *Journal of Hydrology*, **528**, pp. 127-137 (2015).
  16. Rawls, W.J., Gish, T.J., and Brakensiek, D.L. “Estimating soil water retention from soil physical properties and characteristics”, *Adv. Soil Sci.*, **16**, pp. 213-234 (1991).
  17. Zapata, C., Houston, W., Houston, S., and Walsh, K. “Soil-water characteristic curve variability”, *Advances in Unsaturated Geotechnics*, pp. 84-124 (2000).
  18. Benson, C.H., Chiang, I., Chalermyanont, T., and Sawangsuriya, A. “Estimating van Genuchten parameters a and n for clean sands from particle size distribution data”, *ASCE GSP*, pp. 234-235 (2014).
  19. Johari, A., Habibagahi, G., and Ghahramani, A. “Prediction of soil-water characteristic curve using genetic programming”, *Journal of Geotechnical and Geoenvironmental Engineering*, **132**(5), pp. 661-665 (2006).
  20. Garg, A., Garg, A., Tai, K., Barontini, S., and Stokes, A. “A computational intelligence-based genetic programming approach for the simulation of soil water retention curves”, *Transp. Porous Media*, **103**(3), pp. 497-513 (2014b).
  21. ASTM Standard D422-63 “Standard test method for particle-size analysis of soils, annual book of ASTM standards”, *ASTM International*, **04.08**, West Conshohocken, PA, pp. 129-145 (2003).
  22. Koza, J.R., *A Paradigm for Genetically Breeding Populations of Computer Programs to Solve Problems*, Computer Science Dept., Stanford Univ., Margaret Jacks Hall, Stanford, Calif (1990).
  23. Silva, S. “GPLAB: A genetic programming toolbox for MATLAB”, Available at <http://gplab.sourceforge.net> (verified 25 July 2007), Coimbra, Portugal (2007).
  24. Parasuraman, K., Elshorbagy, A., and Si, B.C. “Estimating saturated hydraulic conductivity using genetic programming”, *Soil Sci. Soc. Am. J.*, **71**, pp. 1676-1684 (2007b).
  25. Baker, K. “Investigation of direct and indirect hydraulic property laboratory characterization methods for heterogeneous alluvial deposits: application to the sandia-tech vadose zone infiltration test site”, MS Thesis, New Mexico Institute of Mining and Technology, Socorro, NM (2001).
  26. RMA “Rocky Mountain Arsenal”, Final RCRA-equivalent cover demonstration project comparative analysis and filed demonstration scope of work”, Dec. 1997, Rocky Mountain Arsenal Remediation Venture Office, Commerce City, CO (1997).

27. Rassam, D. and Williams, D. "A numerical study of steady state evaporative conditions applied to mine tailings", *Canadian Geotechnical J.*, **36**, pp. 640-650 (1999).
28. Lu, N. and Likos, W., *Unsaturated Soil Mechanics*, John Wiley, Hoboken, NJ, p. 556 (2004).
29. Likos, W., Wayllace, A., Godt, J., and Lu, N. "Modified direct shear apparatus for unsaturated sands at low suction and stress", *Geotechnical Testing J.*, **33**(4), pp. 286-298 (2010).
30. Tanahashi, H., Sato, T., and Konishi, J. "Degree of imbibition and residence oil saturation in porous media", *Proc. Creation of a New Geoenvironment, Fourth Kansai International Geotechnical Forum*, Kansai Branch of the Japanese Geotechnical Society, Kyoto, Japan, pp. 167-172 (2000).
31. Baker, R. and Hillel, D. "Laboratory test of a theory of fingering during infiltration into layered soils", *Soil Science Society of America J.*, **54**, pp. 20-30 (1990).
32. Stormont, J. and Anderson, C. "Capillary barrier effect from underlying coarser soil layer", *J. Geotechnical Geoenvironmental Engineering, ASCE*, **125**(8), pp. 641-648 (1999).

## Biographies

**Amir Taban** currently teaches in the Technical and Engineering Department of Islamic Azad University. He received his BSc degree in Civil Engineering in 2005 and his MSc degree in Geotechnical Engineering in 2008 from Arak Branch, Islamic Azad University. He recently finished his PhD dissertation entitled "The estimation of dynamic behavior of driven piles in unsaturated soils," with an 'A' degree from the Islamic Azad University, Isfahan Branch. He was a top student during all his studies in BSc, MSc, and PhD courses. He has presented more than 50 papers in various conferences and journals. Moreover, he is the co-author of a book (with Dr. Masoud Mirmohammad Sadeghi) entitled "Unsaturated soils." His studies revolve around unsaturated soil mechanics, soil

improvement, foundation engineering, and geotechnical earthquake engineering. He has a record of 10 years of professional activities on Geotechnics.

**Masoud Mirmohammad Sadeghi** received his BSc in Civil Engineering in 1987 from the Isfahan University of Technology, and his MSc degree in Geotechnical Engineering from Tehran University, as the first student passing the studies at the university in 1991. He received his PhD degree in Geotechnical Engineering from Tehran University in 2001. He is currently Assistant Professor in the Department of Water and Natural Environment, Laboratory of Soil Mechanics, at Isfahan Higher Education And Research Institute (IHEARI). He is also an Assistant Professor in the Department of Civil Engineering at Najafabad Branch, Islamic Azad University. He has published more than 80 research papers in reputable journals and conferences. His research interests mainly include unsaturated soil mechanics, soil biological treatment, soil improvement, foundation engineering, underground structure, and geotechnical earthquake engineering. He has already supervised more than 20 MSc and PhD students.

**Mohammad Ali Rowshanzamir** received his BSc in Civil Engineering from Isfahan University of Technology, Isfahan, Iran, in 1985; MSc degree in Soil Mechanics and Foundation Engineering from Amirkabir University of Technology, Tehran, Iran, in 1989; and PhD degree in Geotechnical Engineering from The University of New South Wales, UNSW, Sydney, Australia, in 1996. He is currently Associate Professor in Civil Engineering Department at Isfahan University of Technology (IUT). He has published several books on foundation engineering, marine geotechnics, and soil mechanics in Persian. He has received seasonal and annual awards of adorable book for publication of *Marine Geotechnics* (translation). He has 18 years of teaching experience in soil mechanics and foundation engineering.