Consolidation Assessment using Multi Expression Programming

Sohrab Sharifi[a]; Saeed Abrishami[b]; Amir H. Gandomi[c,d*]

[a] Ferdowsi University of Mashhad, Faculty of Engineering, Department of Civil Engineering, Mashhad 91775-1111, Iran, e-mail: sharifi.sohrab@mail.um.ac.ir

[b] Ferdowsi University of Mashhad, Faculty of Engineering, Department of Civil Engineering, Mashhad 91775-1111, Iran, e-mail: s.abrishami@um.ac.ir

[c] (Corresponding author) Faculty of Engineering & Information Technology, University of Technology Sydney, Ultimo, NSW 2007, Australia, e-mail: Gandomi@uts.edu.au

[d] School of Business, Stevens Institute of Technolgy, Hoboken, NJ 07030, USA

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Consolidation Assessment using Multi Expression Programming

Abstract: In this study, new approximate solutions for consolidation have been developed in order to hasten the calculations. These solutions include two groups of equations, one can be used to calculate the average degree of consolidation and the other one for computing the time factor (inverse functions). Considering the complicated nature of consolidation, an evolutionary computation technique called Multi-Expression Programming was applied to generate several non-piecewise models which are accurate and straightforward enough for different purposes for calculating the degree of consolidation for each depth and its average as well for the whole soil layer. The parametric study was also performed to investigate the impact of each input parameter on the predicted consolidation degree of developed models for each depth. Moreover, the results of the consolidation test carried out on four different clays attained from the literature showed the proper performance of the proposed models.

Keywords: Geotechnical engineering; Evolutionary computation; Multi-Expression Programming; Consolidation; Prediction

1. Introduction

One of the parameters that has a crucial role in civil engineering is "settlement". Since the serviceability of many structures relies on the amount of vertical displacement, any designer must be aware of this quantity. Apart from this reason, many local codes determine the maximum of such parameter. Thus, a simple equation which is accurate enough and simple is extremely needed.

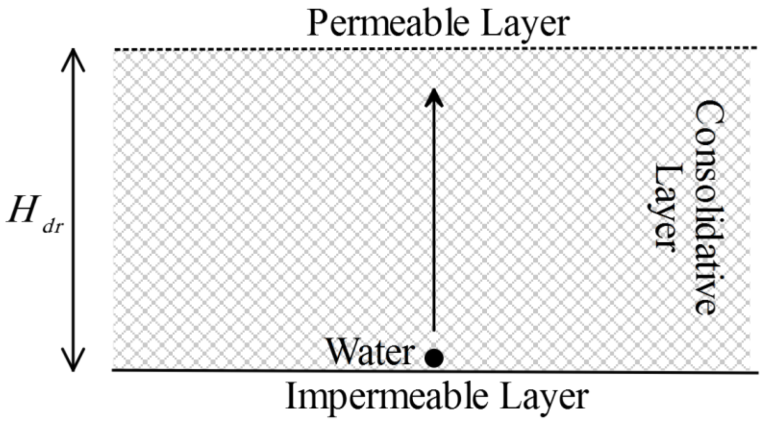
The term "settlement" consists of two concepts. First, an initial or elastic settlement, which is mostly defined for granular soils and immediately occurs after exerting the load. Due to the immediate nature of the initial settlement, researchers and geotechnician usually use elasticity theory to state and formulate it. Second, consolidation settlement is a time-consuming phenomenon, and specially reserved for fine-grained soils. Regarding exerting of load and consequently, increase in stress of fine-grained soils (such as clays), water is forced out because of insignificant compressibility nature of it (it is assumed in many theories like Terzaghi's that water is relatively incompressible which is in this paper as well). With assuming the fact that the soil stays fully saturated, reducing the water content of granular media leads to increasing effective stress since the pore pressure drops after dissipation of pore water. Therefore, both the thickness of the soil layer and void ratio begin to decrease.

From the standpoint of water drainage direction, consolidation usually breaks down into two categories (one-dimensional and three-dimensional). Moreover, water is expelled out from the boundaries of soil stratum as consolidation progresses. These boundaries or drainage faces can be easily the ground surface subjected to air pressure or generally the surface of a high permeable soil layer. However, when the water is drained out both upward and downward, a doubly drained system is said to be available. In some condition, it is possible that only one face allows the pore water to be drained which is referred to as single drainage or one-way drainage (see Fig.1). In this discussion, the planar circumstance is assumed, so term "consolidation" stands for "one-dimensional consolidation" and doubly drained system also is present unless it is mentioned. Although it should be noted that because of the symmetric characteristic of consolidation, single drainage and doubly drained systems are not that much different.

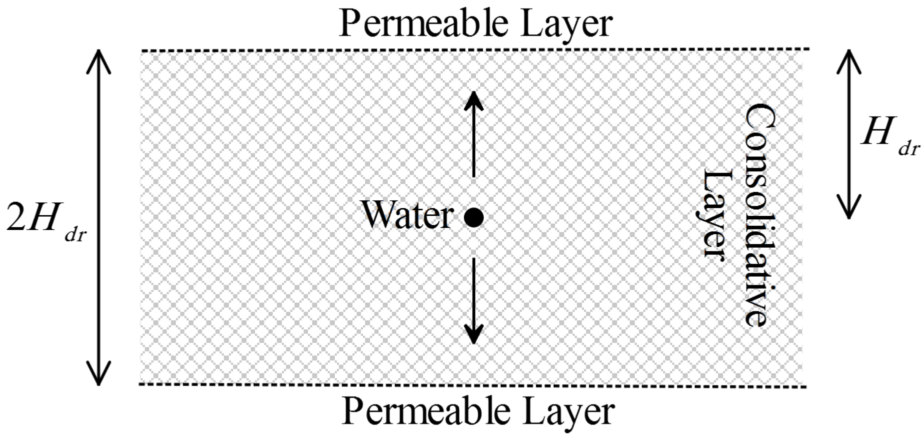
The classical one-dimensional consolidation theory proposed by Terzaghi [1] is thoroughly familiar in geotechnical engineering. Eq.1 is the governing partial differential equation (PDE) of excess pore-water pressure dissipation in a fully saturated clay [1]:

$$\frac{\partial u}{\partial t} = c_v \frac{\partial^2 u}{\partial z^2} \tag{1}$$

Where c_v is consolidation coefficient, u is excess pore-water pressure, t is time, z represents the depth of the desired clayey layer and H_{dr} the longest drainage path length. c_v is the consolidation characteristic of a soil that is dependent to hydraulic conductivity and the variation of void ratio with respect to the change in effective stress. It shows the time rate at which the soil stratum undergoes the consolidation. This parameter can be calculated through a variety of ways that one of them has been introduced further in this study to validate the generated models.



(a)



(b)

Fig. 1. Schematic concept of consolidation in (a) single drainage and (b) double drainage

In this PDE, two boundary conditions and one initial condition are required. Boundary conditions must provide information about excess pore-water pressure at two specific depth of the clay layer. *u* at freely drainage boundaries and impermeable boundaries is obvious, and initial condition in terms of *t* is also given at the onset of consolidation. The initial excess pore-water pressure (u_i) is given by:

$$u(z,0) = f(z) \tag{2}$$

For a sustained loading of pore-water pressure, *f(z)* is constant. Using the boundary conditions specified in Table 1, by means of separation of variables, Eq. 1 can be solved [1]:

$$u(z,t)=\sum_{n=1}^{\infty} A_n \sin\left(\frac{m\pi z}{2H_{dr}}\right)\exp\left(\frac{-m^2\pi^2 T_v}{4}\right) \tag{3}$$

$$T_v=\frac{c_v t}{H_{dr}^2} \tag{4}$$

Where $m = 2n + 1$ and A_n is series coefficient which are determined in order to satisfy initial condition (boundary condition 1 from Table 1). u (z, t) represents the excess pore pressure at depth z and time t, and T_v is a dimensionless factor which is a function of time (Eq. 4). T_v is an important factor to identify transient behaviour of a soil sample under consolidation. After reaching most of the consolidation, major parts of soil layer are not saturated anymore which initiates another phenomenon called "creep" which is also recognized with "secondary consolidation" that should approximately start when T_v is one in an ideal theoretical soil. Although it is absolutely one of the most important phenomena in geotechnical engineering, only primary consolidation is the case of this investigation.

Table 1 Boundary conditions

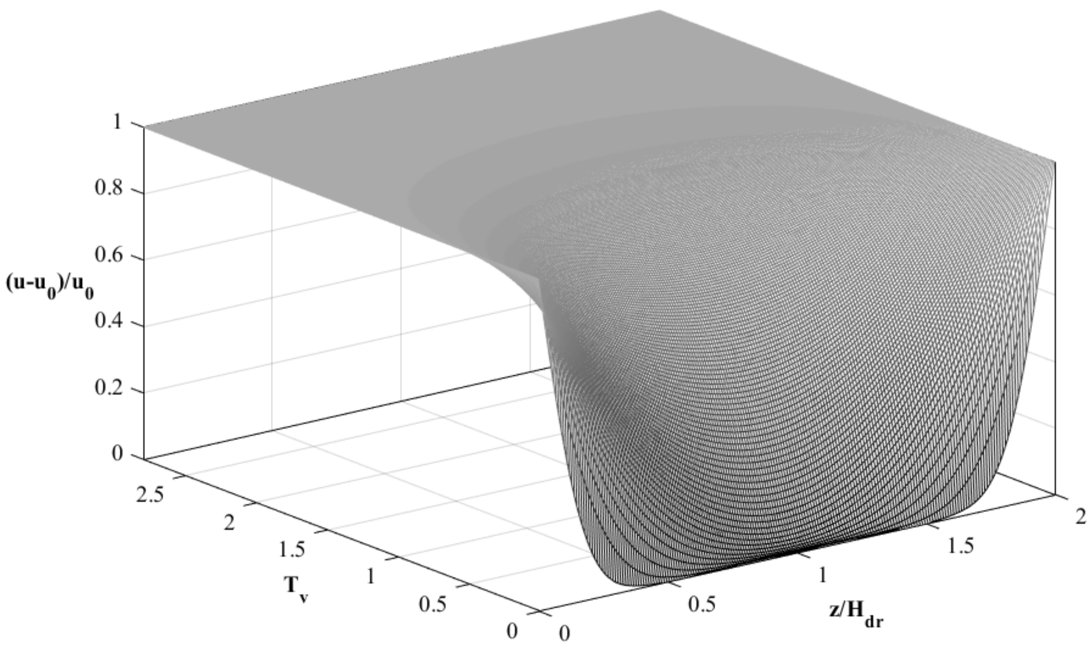
Boundary condition	Mathematical expression
1. There is complete drainage at the top of the soil layer	When z=0, u=0
2. There is complete drainage at the bottom of the soil layer	When z=2H_{dr}, u=0
3. The initial excess pore-water pressure distribution is specified as a constant function of depth	When t=0, u(z,0)=cte.

The average degree of consolidation (U) is the average dissipated pore pressure with respect to the initial status (Eq. 5). U actually represents the percentage amount of ultimate consolidation settlement that has been occurred up to a specific amount of time. As it was mentioned before, creep begin in the proximity of T_v=1. Determination of such a moment can be done by plotting tangent lines of U-T_v curve. The drop in the slope of the middle part of the curve (primary consolidation) mirrors the initiation of creep.

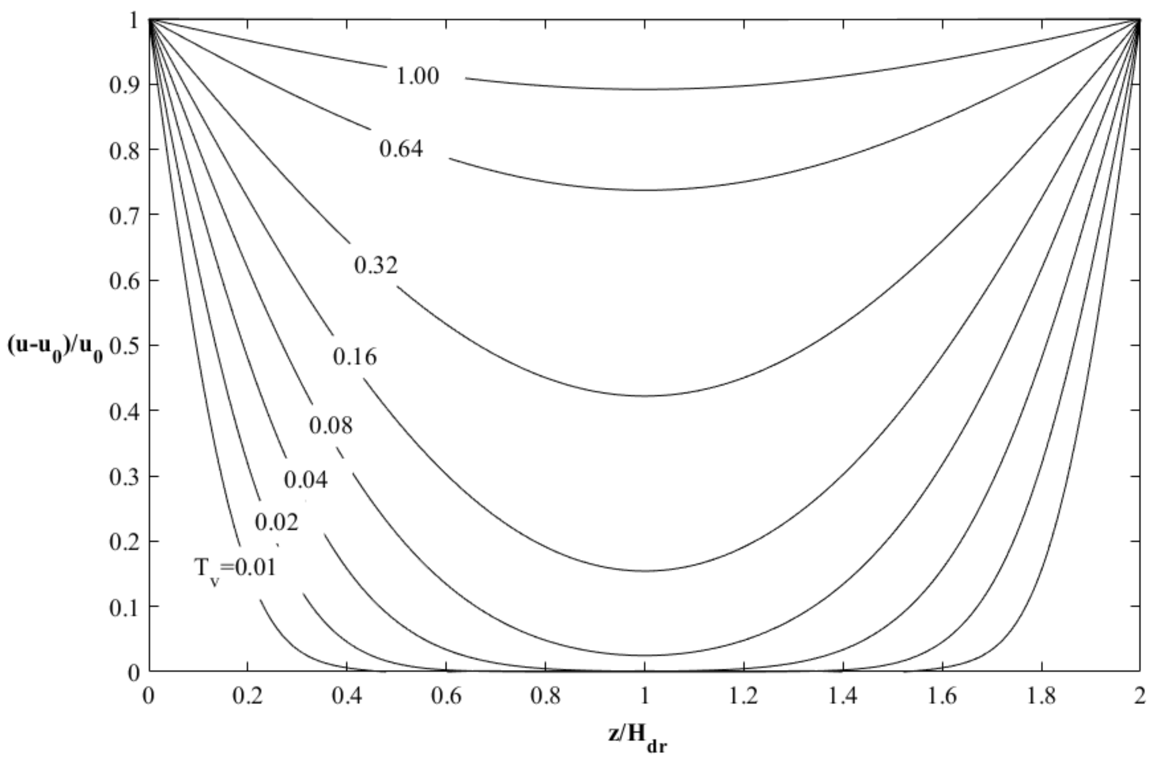
$$U=1-\frac{\int_0^{2H_{dr}} u(z,t)dz}{\int_0^{2H_{dr}} f(z)dz} \tag{5}$$

Hence, the approximate solution of consolidation is presented by Terzaghi as evidenced in Eq. 6 [1]. Fig. 2 shows the solution of Eq. 1 plotted in 3D and 2D views.

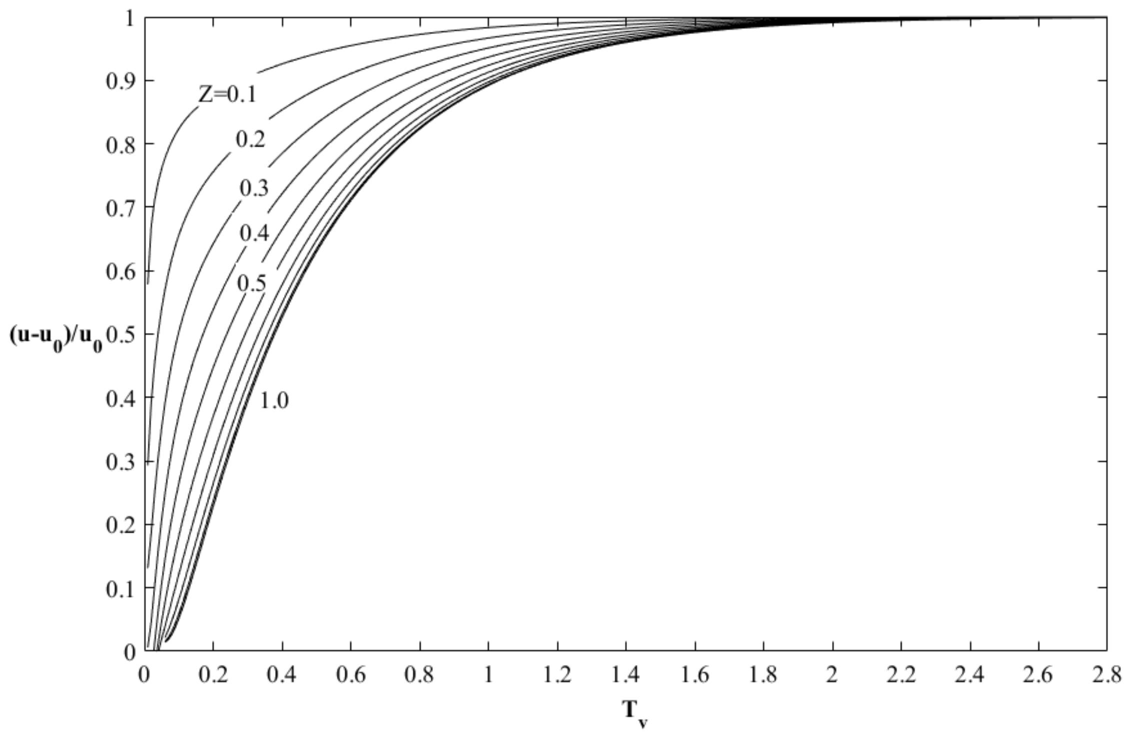
$$.\begin{cases} T_v \approx \frac{\pi}{4} U^2 & 0 \le U \le 0.526 \\ T_v \approx -0.933\log(1-U) - 0.085 & 0.526 \le U \le 1 \end{cases} \tag{6}$$



(a)


1.00
0.64
0.32
0.16
0.08
0.04
0.02
T_v=0.01
(u-u_0)/u_0
z/H_dr

(b)


Z=0.1
0.2
0.3
0.4
0.5
1.0
(u-u_0)/u_0
T_v

(c)

Fig. 2. Demonstration of the analytical solution of consolidation equation from different views

The importance of the T_v-U curve is highly notable in the analysis of transient settlement. The consolidating behavior of soil subjected to constant initial excess pore-water pressure is often described by isochrones and the average degree of consolidation curves [1, 2]. As Fig. 2a illustrates, the variation of $(u\text{-}u_0)/u_0$ lowers when T_v is larger than one which is more demonstrated in Fig. 2b as well. This may be attributed to the fact that was mentioned before about dominancy of creep over consolidation for large amounts of T_v. Further in this study, the upper limit of T_v is selected to be 2.8 for model generating purposes. However, it should be mentioned that it is rather a theoretical assumption since no mathematical boundary exists for Eq. 3 while such matter may not happen in the real condition.

Sivaram and Swamee [3] also gave an approximation of T_v-U function that is available for the sustained loading as Eq. 7.

$$\frac{U\%}{100}=\frac{(4\,T_v/\pi)^{0.5}}{[1+(4\,T_v/\pi)^{2.8}]^{0.179}} \tag{7}$$

Sridharan et al. [4] discovered that $T_v\,/\,U$ is in a linear relationship with T_v for a doubly-drained soil layer under rectangular u_i when 60% < U < 90%. A method called "Rectangular hyperbola" was presented as well to obtain consolidation coefficient which is explained in further sections. Therefore, a mathematical formula was achieved with the following expression:

$$\frac{T_v}{U}(\%)=8.208\times10^{-3}T_v+2.44\times10^{-3} \tag{8}$$

Darrag and El-Tawil [5] adopted the usual assumptions and developed a mathematical model to analyze the consolidation of clays where the u_i distribution was instead represented as a stochastic process. Tewatia et al. [6] proposed Eq.9 for $40\%<U<60\%$ where U is dependent to its rate.

$$U=-\log_{10}\left(\frac{dU}{dT_v}\right)+0.6 \tag{9}$$

Chan [7] reported the following formula, proposed by Hansen that fits the whole range of the analytical solution with a maximum error of 0.0052:

$$T_v \approx \sqrt[3]{\frac{U^6}{2(1-U^6)}} \tag{10}$$

Singh and Swamee [8] studied basic and reverse triangular u_i distributions to give invertible approximate equations for consolidation curves in single drainage condition. McKinley and Sivakumar [9] indicated that a plot of velocity against displacement should be approximately a straight line for a degree of consolidation in the range 50% to 100% by manipulation of piecewise Fox's solution. Lovisa et al. [10] studied various cases under non-uniform and even asymmetric u_i for a doubly drained soil stratum. Lovisa et al. [11] conducted numerical simulation using MATLAB and observed no difference between Terzaghi's traditional formula and mass flux formula which was presented by them. Lovisa et al. [12] proposed a method using a discretization technique for time-dependent loading cases in which the total load was divided into equal smaller increments. Laboratory tests affirmed that parabolic excess pore pressure at the final stages of construction is more realistic than sinusoidal shape. Tewatia et al. [13] derived a new equation for one-dimensional and three-dimensional consolidation as Eq. 11. It is noteworthy that the definition of T_v differs with respect to the vertical or radial consolidation.

$$U=1-e^{-T_v} \tag{11}$$

Lovisa et al. [14] presented several approximation equations that help to acquire the average degree of consolidation and pore pressure at a specific depth for a variety of loading conditions in different U domains. They also proposed different U-T_v approximations for a doubly drained clay layer for different excess pore water pressure distributions. It should be noted that all of these mentioned formulas are piece-wise functions. Al-Zoubi [15] utilized the extended Taylor method (ETM) and compared the U-T_v curves obtained using ETM with Terzaghi's theoretical relationship. Di Francesco made much efforts to attain exact solutions of consolidation equation in one-dimension [16], two-dimension and three-dimension circumstances [17,18]. This led to a nonlinear equation based on the properties of elastic waves and elastic functions using periodic functions, which is able to simulate excess pore pressure.

In this paper, the consolidation of a clay stratum with two-way drainage is studied based on the theory of Terzaghi under the occurrence of uniform pore-water pressure distribution. Both implicit and explicit functions for the average degree of consolidation are derived by utilizing an evolutionary algorithm called MEP. As it was evident in the literature review, almost no studies have been conducted so far on calculating the degree of consolidation for each depth which is another aim of this investigation. Eventually, for validating the generated models, six fine-grained samples were extracted from different locations in Mashhad city, Iran. Oedometer test results revealed good agreements between consolidation coefficient calculated by the proposed models, Terzaghi's approximation (Eq.6) and hyperbola method [4].

2. Evolutionary computation

Most of the problems in the engineering world do not have incisive answers. Thus, it is necessary to accept uncertain methods with tolerable errors. Evolutionary algorithms (EAs) are well-known strategies to conquer complicated challenges. As instances, genetic algorithm (GA) [19] and evolutionary programming (EP) [20] are subsets of EAs. At first place, EAs generate an initial set containing possible but not correct answers called "Generation", then based on a known fitness function these answers are scored and the most ranked ones are selected to be "Parents". During a process, some operations will be applied on these parents to improve their scores, and they will reproduce the next generation called "Children", so the cycle goes on until the convergence is satisfying [21]. Next-generation consists of individuals which possess the highest fitness scores (off-spring) [22]. This mechanism and main idea of EAs is quite similar to the procedure that leads to the elimination of unfit existences and lets the superior ones evolve as known as the Darwinian concept of fitness.

In GA, these individuals usually are converted to the binary form (called "Chromosome"). Therefore, the arithmetic calculation can be practicable and programmable in computers. The intermediate operations between parents and children can be mutation and crossover (Fig. 3). Fig. 3b shows one point cutting type of crossover while it should be mentioned that another type of crossover called “uniform” was utilized in this study (Fig. 4). In this type, chromosomes are not divided and each gene is treated separately. The genes are decided to be part of off-spring totally random.

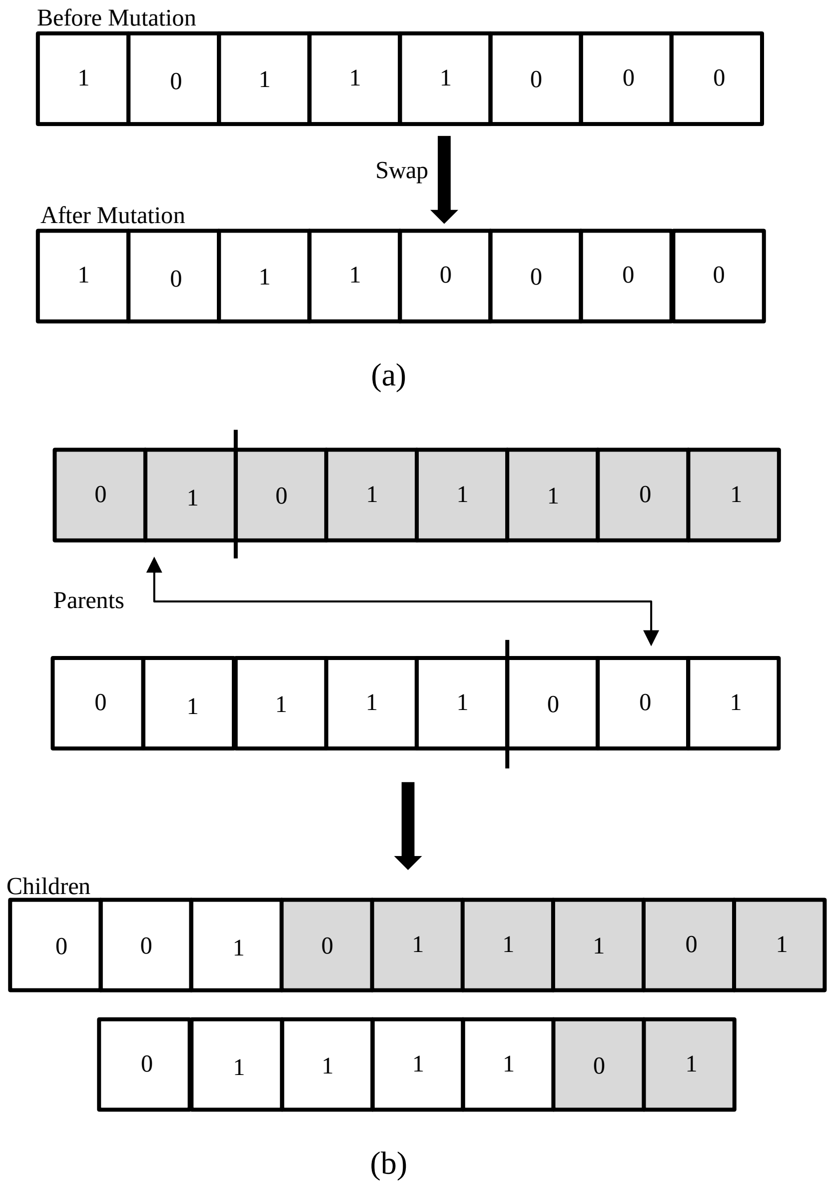


Fig. 3. Concept of (a) mutation and (b) crossover

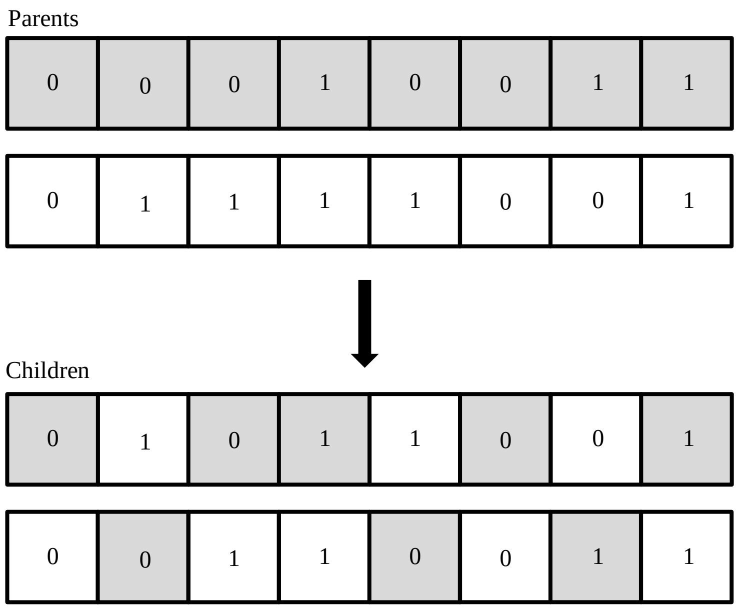


Fig. 4. Concept of uniform crossover

If the aim of GA computation is finding the model itself, a specialization of GA called Genetic Programming (GP) should be utilized. In GP, encoded answers are programs instead of binary strings. Also, inputs and outputs are given, the mathematical rule to relate them is unknown [23].

In the past years, several forms of GP have been developed which linearity is mostly the main factor that distinguishes them from each other. Individuals can be represented as linear entities similar like what happens in linear genetic programming (LGP) [24] and multi-expression programming (MEP) [25] or represented as tree-shaped solutions like traditional genetic programming [21]. Usually, the tree-based GP variants are slower because of using a translator. MEP has two advantages over other methods which improve efficiency and reduces the consuming time of calculation:

1) using a linear-based concept

2) each chromosome contains multiple solutions.

Stages of evolving a computer program in linear-based variants are as follows [26]:

I. Setting up a population with arbitrarily generated programs.

II. Using a binary tournament to select two parents.

III. Acquiring two offspring by cross-over and mutation over the parents.

IV. Replacing the loser parents in the tournament with the winner parents of stage III

V. Repeating steps two through four until convergence.

MEP results include linear strings of instructions which are combinations of variables (terminals) or mathematical operators (functions). The length of the chromosome is defined by the number of MEP genes per chromosome. A terminal or a function symbol is encoded by each instruction line (gene). In order to assure that no false cycle appears, each function contains cursors toward arguments [27]. An example of an MEP chromosome is described below by using the set of terminals T= {*A*, *B*, *C*, *D*} and a set of functions F= {-, ^, *}:

0: *A*

1: *B*

2: ^ 0, 1

3: *C*

4: - 2, 3

5: *D*

6: * 4, 5

To interpret a chromosome into computer programs, it has to be read top-down from the position of the first node. In the present example, simple expressions are encoded by genes 0, 1, 3 and 5 which are single terminal symbols:

$E_0 = A$,

$E_1 = B$,

$E_3 = C$,

$E_5 = D$,

Gene 2 implies the operation ^ on the operands at position 0 and 1 of the chromosome. Therefore, gene 2 encodes the expression:

$$E_2 = A \wedge B \tag{12}$$

Gene 4 implies the operation - on the operands at position 2 and 3. Therefore, gene 4 encodes the expression:

$$E_4 = (A \wedge B) - C \tag{13}$$

Gene 6 implies the operation * on the operands at position 4 and 5. Therefore, gene 6 encodes the expression:

$$E_6 = ((A \wedge B) - C) * D \tag{14}$$

To comprehend the MEP calculation procedure, these chromosomes can be demonstrated in the shape of trees (it is still linear-based) which form a forest of trees. Fig. 5 illustrates the forest of expressions in the above-mentioned MEP chromosomes and possible model can be found in among of these expressions. The fitness of the best-encoded expression defines the fitness of the whole MEP chromosome [28].

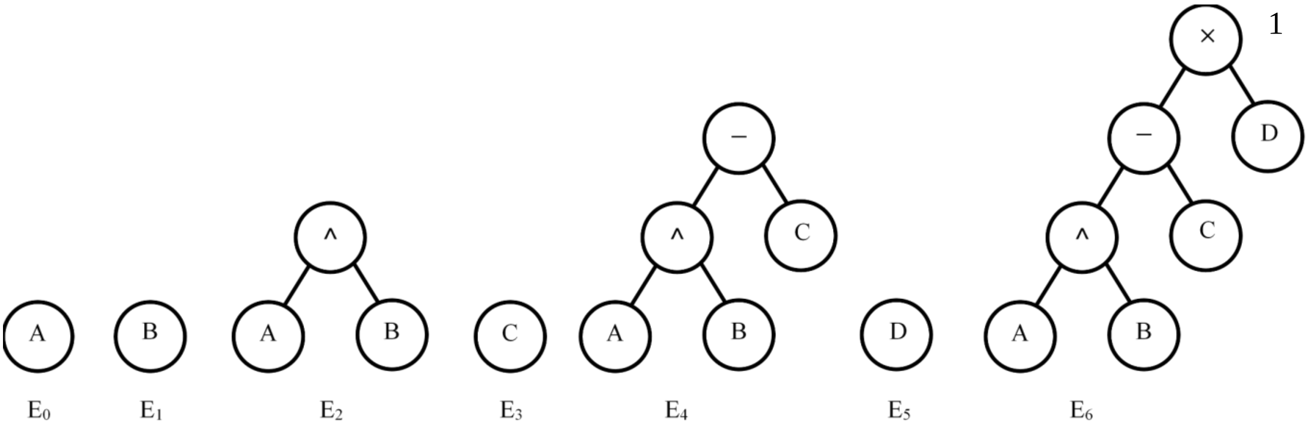


Fig.5. Expressions encoded by a MEP chromosome represented as the shape of trees [27]

To control and select the best expression in an MEP chromosome, the following equation is considered as a fitness function [25]:

$$f = \min_{i=1,m}\left\{\sum_{j=1}^{n}\left|E_j - O_j^i\right|\right\} \tag{15}$$

Where n is the number of fitness cases, E_j is the expected value for the fitness case j, O_j^i is the value returned for the jth fitness case by the ith expression encoded in the current chromosome, and m is the number of chromosome genes [23,27,29].

Sometimes by presenting a new data set on the developed model, the error becomes very large, this incident is called “overfitting”. The complex behavior of data structure can cause this mishap; as an instance, the number of parameters can be greater than the number of observations or sudden changes could happen just between in two observed points. Testing individuals on a validation set is a solution to avoid overfitting and find a more generalized model. In this research, the overall data set generated from Terzaghi solution (Eqs.3 & 5) was classified into three subsets randomly: I) training, II) validation, and III) test subsets. The training subset was used to evolve MEP models, and the validation subset was used to estimate the prediction error. Eventually, the models that pass through validating filter are tested under data of test subset [23]. Here, about 70%, 15% and 15% of the data sets were chosen for the training, validation and testing the obtained model, respectively.

In order to evaluate the capabilities of the proposed MEP models, correlation coefficient (R), root mean squared error (RMSE) and mean of error (MAE) were used as follows [29]:

$$R=\frac{\sum_{i=1}^{n}(a_i-\overline{a_i})(c_i-\overline{c_i})}{\sqrt{\sum_{i=1}^{n}(a_i-\overline{a_i})^2\sum_{i=1}^{n}(c_i-\overline{c_i})^2}} \tag{16}$$

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}(a_i-c_i)^2}{n}} \tag{17}$$

$$MAE=\frac{1}{n}\sum_{i=1}^{n}|a_i-c_i| \tag{18}$$

Where a_i and c_i are actual and estimated output values respectively for the *i*th output and n is the number of points. The best models are selected highest R^2 (usually more than 0.8 is appropriate) and least RMSE and MAE (these indicators are usually evaluated in comparison with others); additionally, another factor for choosing was less complexity since it is not favorable.

Due to the mentioned advantages of MEP, plenty of researches have been carried out in geotechnical engineering throughout recent years. Alavi et al. [30] developed a new constitutive model by formulating secant and reloading module. Using LGP and MEP, Alavi and Gandomi [31] assessed the amount of triggering strain for the occurrence of liquefaction. Later, Alavi et al. [32] took the moduli of pressuremeter tests conducted on over 100 soil samples and presented empirical mathematical relationships for predicting the elastic modulus of other samples. There are also structural instances like a prediction of shear strength of reinforced concrete columns [33], reinforced concrete deep beams [34] and steel fiber-reinforced concrete beams [35]. Moreover, studies have been conducted on assessing the properties of both normal-strength concrete and high-strength concrete using MEP like elastic modulus [36] or the time-dependent total creep [37]. In this study, the MEP computation was handled by MEPx software package [38]. Other prevalent statistical software packages, like SPSS or R, have few disadvantages, namely limited numbers of functions, dependency to user for determining the mathematical expression form of a model, and necessity of specifying initial values for constants that can seriously cause failing to reach the convergence. These reasons along with potent advantages of MEPx are major motives that drove the authors to use it for this study.

3. Results of formulation for average consolidation degree

200 points were generated from $T_v = 0.01$ to 2.00 with 0.01 increment. Using Eq. 3 average degree of consolidation were calculated corresponding to specified time factors. With try and error, the best fitting models for both explicit and implicit forms were found.

As mentioned before, several parameters are required to be determined to obtain a robust model with MEP method which are listed in Table 2 that can be beneficial for applying MEP on other problems as well. They were found and optimized during several runs. For each case, the termination condition was observing no notable improvement in the performance or reaching the run number restriction [29, 39].

Table 2 Parameters setting for MEP

Parameters	Setting
Number of generation	1000,1500,2000,3000
Population size	100, 150, 200
Function set	+, -, *, /, power, Sqrt, Exp, Ln, Sin, Cos, Tan, ASin, ACos, ATan, Tanh
Chromosome length	50, 100, 200
Crossover probability	0.5, 0.9
Crossover type	Uniform
Mutation rate	0.01, 0.5, 0.9

3.1. Explicit form

The outcome-equations of MEP are tabulated in Table 3 with all statistical details. Also, a comparison between formerly proposed equations mentioned in the literature review is provided in Table 3. With a look at Table 3 and Table 4, it can be mentioned that generated models are more accurate than almost all of the previously presented approximations.

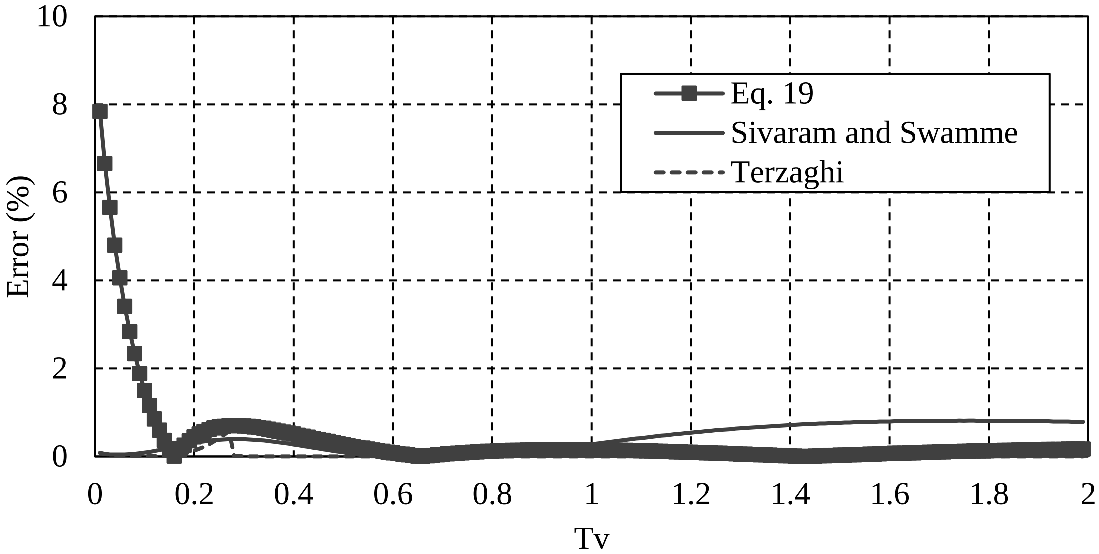
Table 3 Comparison of equations presented by former researchers

Equation No.	Equation	Statistical details		
		R^2	RMSE	MAE
6	Terzaghi [1]	0.9999	0.0002	3.88×10^{-5}
7	Sivaram and Swamee [3]	0.9998	0.0055	0.0046
8	Sridharan et al. [4]	0.9869	0.0543	0.0419
9	Tewatia et al. [6]	0.6346	1.1738	0.9186
10	Chan [7]	0.9508	0.6062	0.3412

The performance of the models was evaluated on the entire database as illustrated in Fig. 6 (The term "Error" is defined as a relative error).

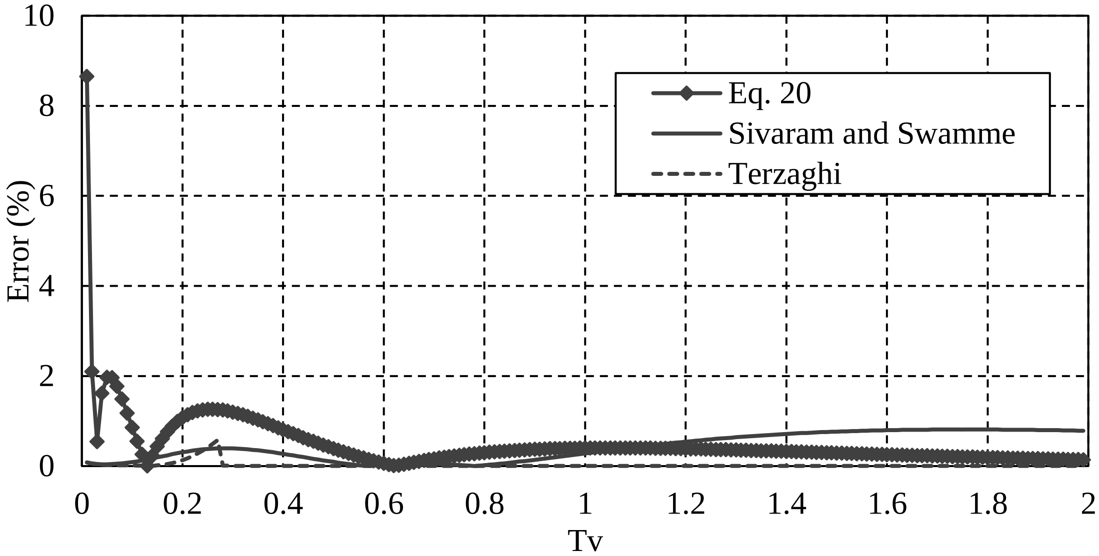
Table 4 Obtained explicit models by MEP for the average degree of consolidation

Equation No.	Equation	Description	Statistical details		
			R^2	RMSE	MAE
19	$U=\left(\tan^{-1}\sqrt{T_v}\right)^{\left(0.301029^{T_v}\right)}$	-	0.9998	0.0027	0.0019
20	$U=1-\tan\left(\frac{0.723228}{10^{T_v}}\right)$	-	0.9997	0.0037	0.0026
21	$U=(T_v\cdot \exp(A))^{0.5\cdot\exp(A+T_v\cdot\exp(A))}$	$A=-3.21097\cdot T_v$	0.9994	0.0060	0.0041
22	$U=\tanh(0.6488T_v^{1.202}+T^{0.5062}+0.01926)$	-	0.9999	0.0005	0.0003



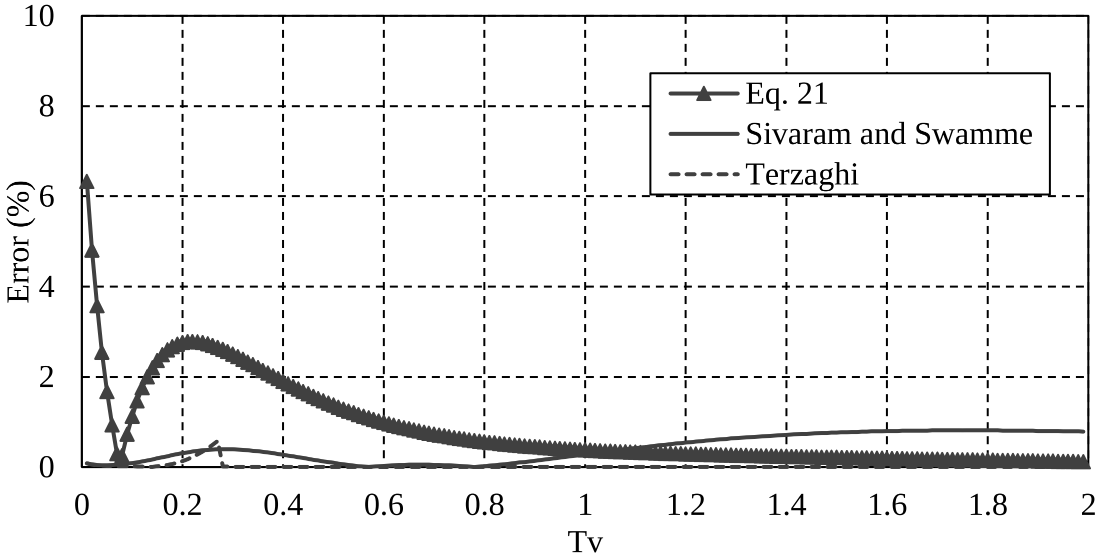
Eq. 19
Sivaram and Swamme
Terzaghi
Error (%)
Tv
0
2
4
6
8
10
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

(a)



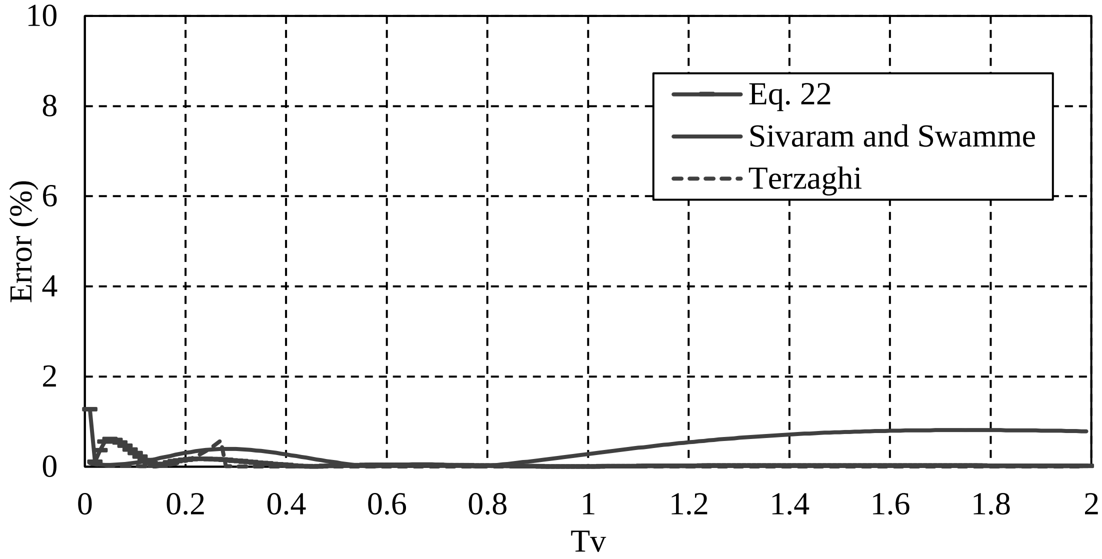
Eq. 20
Sivaram and Swamme
Terzaghi
Error (%)
Tv
0
2
4
6
8
10
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

(b)



Eq. 21
Sivaram and Swamme
Terzaghi
Error (%)
Tv
0
2
4
6
8
10
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

(c)



(d)

Fig. 6. Performance of (a) Eq.19, (b) Eq.20, (c) Eq.21 and (d) Eq.22 in the form of relative error (%) corresponding to T_v

3.2. Implicit form

First, an inverted form of Eq. 22 is obtained as follows:

$$0.6488T_v^{1.202}+T_v^{0.5062}+0.01926=\tanh^{-1}(U) \tag{23}$$

Let "$\tanh^{-1}$ (U) -0.01926" be a parameter called '*K*':

$$f(T_v)=0.6488T_v^{1.202}+T_v^{0.5062}-K=0 \tag{24}$$

With a change of variables as Eq. 25, Eq. 24 can be approximately assumed a quadratic equation:

$$X=T_v^{0.5062} \tag{25}$$

$$0.6488X^{2.37}+X-K=0 \tag{26}$$

$$\Rightarrow X=\frac{-1+\sqrt{1+2.5952K}}{1.2976} \tag{27}$$

Using Eq. 25, Eq. 28 is given by:

$$T_v=\left(\frac{-1+\sqrt{1+2.5952K}}{1.2976}\right)^{\frac{1}{0.5062}} \tag{28}$$

As previously mentioned, Eq. 26 is not accurately a quadratic equation while the Eq. 28 is derived based upon this assumption. Thus, a correction factor to Eq. 28 can be defined as follows:

$$C_c=\frac{(T_v)_{accurate}}{(Tv)_{approximated}} \tag{29}$$

Then, a dataset consists of these correction factors was generated and a curve fitted on it as follows:

$$C_c=0.619U^5+1.882U^3-2.217U^2+0.9089U+0.9872 \tag{30}$$

Eventually, an inverse function of Eq. 22 will be like:

$$T_v=C_c\left(\frac{-1+\sqrt{1+2.5952\ K}}{1.2976}\right)^{1.9755} \tag{31}$$

Statistical details are tabulated in Table 5 along with the other models also developed by MEP method similar to explicit models.

Table 5 Obtained implicit models by MEP for average degree of consolidation

Equation No.	Equation	Description	Statistical details		
			R^2	RMSE	MAE
32	$T_v=C_c\left(\frac{-1+\sqrt{1+2.5952K}}{1.2976}\right)^{\frac{1}{0.5062}}$	Eq. 30 $K=\tanh^{-1}(U)-0.01926$	0.9998	0.0151	0.0080
33	$T_v=\sin\left(U\sqrt{U}\right)\cdot\ln\left(\tan\left(\frac{\sin^{-1}U}{U}\right)\right)$	-	0.9997	0.0441	0.0388
34	$T_v=\log_{10}\left(\frac{0.723228}{\tan^{-1}(1-U)}\right)$	-	0.9998	0.0442	0.0327

4. Results of formulation for consolidation degree of each depth

28280 points were generated from T_v = 0.01 to 2.80 and dimensionless parameter $Z=z / H_{dr}$ from 0.00 to 1.00 both with 0.01 increment. The degree of consolidation was calculated to specified time factors and depth factors by using Eq. 3. The same parameters mentioned in Table 2 were found to be proper to obtain robust models. Outcome equations are tabulated in Table 6 with all statistical details. The performance of generated models is illustrated in Fig. 7. It shows that Eq. 36 is the weakest model as its RMSE and MAE would also implicate. Moreover, Eq. 35 and Eq.37 are intermediate from the viewpoint of performance although, in low amounts of U, the error grows. The others, eventually, are the most robust ones among the generated models like Eq. 40 that is the most robust model.

Table 6 Obtained models by MEP for the degree of consolidation of each depth

Eq. No.	Equation	Description	Statistical details
35	$U=\left[\sin\left(\sqrt{T_v}\right)\right]^{\frac{Z}{\exp\left(\sin\left(\log_{10} T_v\right)\right)}}$	-	R^2=0.9943 RMSE=0.0259 MAE=0.0173
36	$U=\left[\tan^{-1}\left(T_v^{\sqrt{A}}\right)\right]^{Z\cdot A}$	$A=0.594067^{T_v}$	R^2=0.9946 RMSE=0.0415 MAE=0.0278
37	$U=\left[\sin\left(\sqrt{T_v}\right)\right]^{\frac{Z}{T_v^{0.25}}}$	$0.00 \le Z \le 0.69$	R^2=0.9952 RMSE=0.0098 MAE=0.0043
38	$U=\tan^{-1}\left(\tan^{-1}\left(\frac{T_v\cdot\exp(T_v)}{Z^Z}\right)\right)$	$0.70 \le Z \le 1.00$	R^2=0.9985 RMSE=0.0130 MAE=0.0078

39	$U=1.002-1.072\cdot\tanh\left(\frac{4.341\cdot Z\cdot\cos(0.735Z)}{\exp\left(3.231T_v^{0.6}\right)}\right)$	0.00≤Z≤0.69	R^2=0.9959 RMSE=0.0125 MAE=0.0043
40	$U=1.011-1.033\cdot\tanh\left(\frac{5.412\cdot Z\cdot\cos(0.823Z)}{\exp\left(3.48T_v^{0.6}\right)}\right)$	0.70≤Z≤1.00	R^2=0.9994 RMSE=0.0083 MAE=0.0059

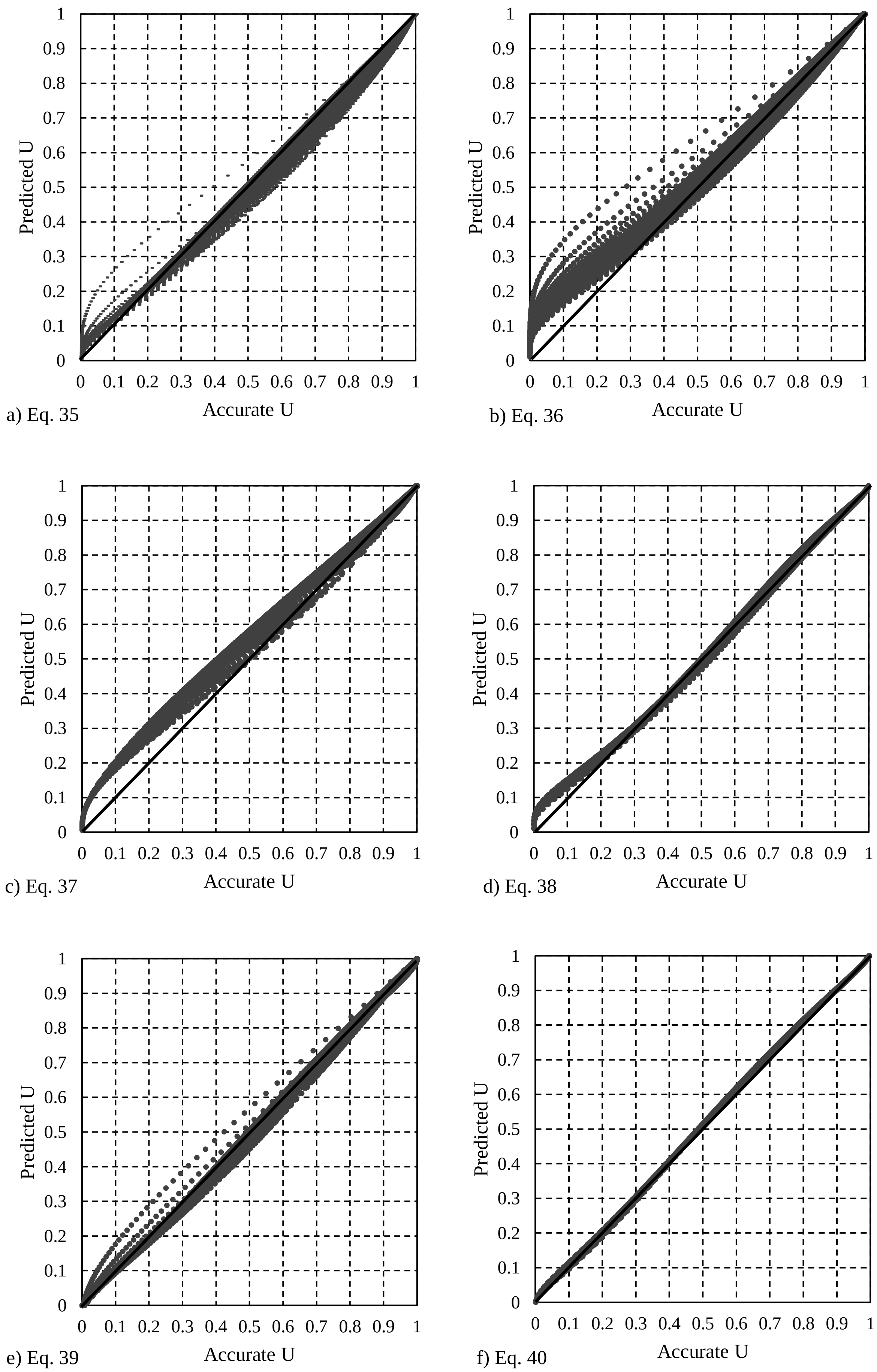


Fig. 7. Performance of the developed models for each depth

4.1. Parametric study

The variation in the degree of consolidation is graphically presented in Fig.2 and regarding the high R^2 of generated models, it is expected that they own the same trend. However, the contribution of each input on the output must be clarified through a parametric study. In this study, a simple procedure was undertaken as reported before [34]. This method takes into account the maximum and minimum of each parameter and calculates their impact based on the difference between these two amounts. The following formulas are used for this purpose:

$$N_i=U_{max}(x_i)-U_{min}(x_i) \tag{32}$$

$$S_i(\%)=\frac{N_i}{\sum_{j=1}^{n} N_j}\times 100 \tag{33}$$

Where $U_{min}(x_i)$ and $U_{max}(x_i)$ are the minimum and maximum of the predicted degree of consolidation over the *i*th input domain respectively, and *n* is the number of all parameters which is two in this case. Considering the fact that minimum and maximum of predicted output for the problem of consolidation is always zero and one, the domain of input parameters was restricted. Therefore, the domains were assumed as: $0\leq Z\leq 0.5$ and $0\leq T_v\leq 0.6$ where the variation of consolidation degree is notable.

The results are tabulated in Table 7 and indicate that no significant difference exists among the models from the viewpoint of the parametric study. However, it is observed that the depth factor has a higher effect on the result rather than the time factor over all of them. This implicates that such factor should be treated more carefully; nevertheless, the impact of both parameters are close, and none of them is predominant over the other one.

Table 7 The results of the parametric study for developed models

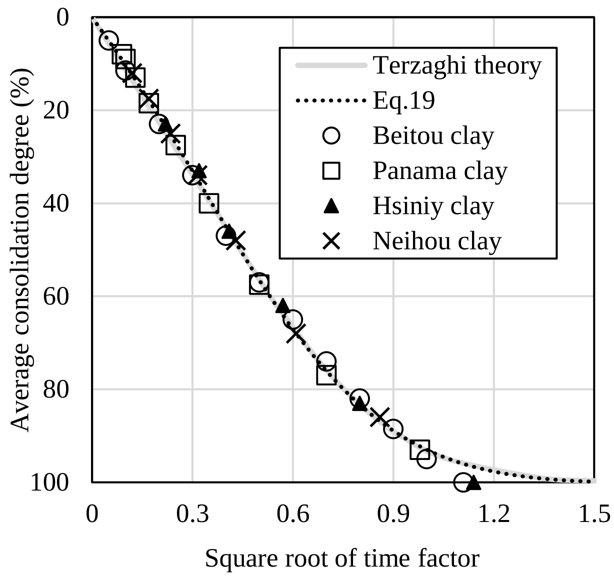
Model		Eq. 35	Eq. 36	Eq. 37	Eq. 39
Impact (%)	Z	54.26	55.50	55.04	55.53
	T_v	45.74	44.50	44.96	44.47

5. The model verification

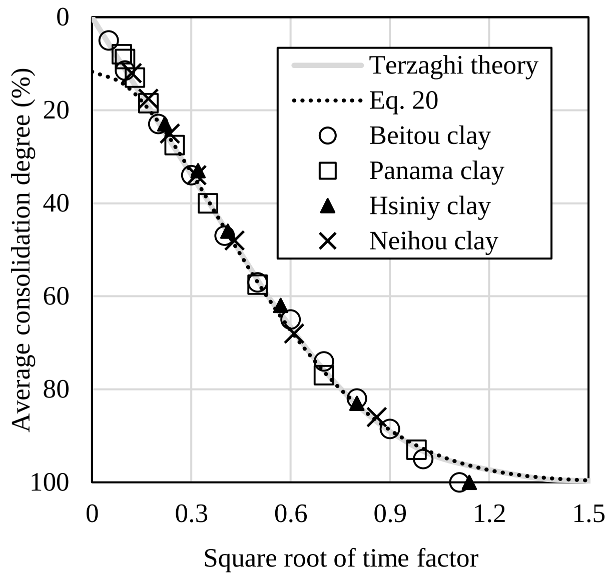
four types of soft clay were selected from a study carried out by Feng and Lee [42] to verify the applicability of generated models. The characteristics of those soils are given in Table 8. Fig. 8 shows a comparison between the results Oedometer test performed on mentioned clays and generated models. As seen here, it is evident that both explicit and implicit models are in favourable accuracy except for Eq. 20 that shows relatively large errors in $\sqrt{T_v}$ <0.15. However, it should be mentioned that this could be predictable since these equations are highly fitted to the Terzaghi theoretical curve and did not originate from fundamentally different assumptions about consolidation. Therefore, these approximations are valid for any clayey soil which Terzaghi theory has been proven accountable for. The same argument can be also alleged to prove the validity of those developed for describing the consolidation degree for each specific depth.

Table 8 Characteristics of soft clays used to verify the models

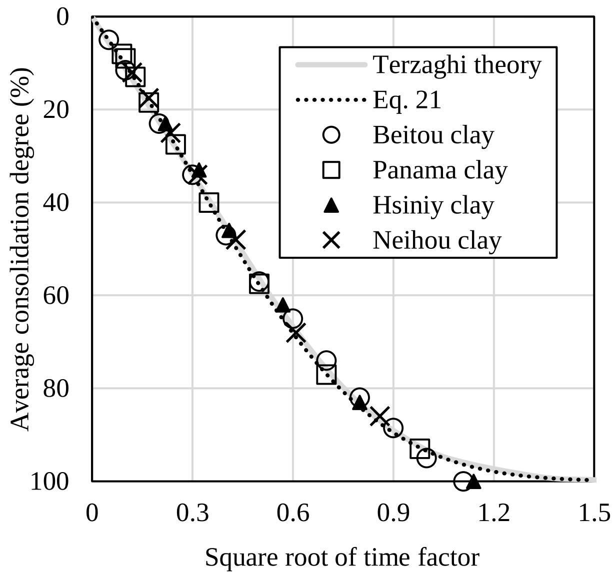
Clay	Location	Natural water content (%)	Liquid limit (%)	Plastic limit (%)	Over-consolidation ratio
Beitou	Taipe, R.O.C.	58	77	35	1.3
Hsiniy	Taipe, R.O.C.	38	51	26	2.8
Neihou	Taipe, R.O.C.	42	41	23	2.2
Panama	Colon, Panama	93	152	53	1.0-1.1


Average consolidation degree (%)
0
20
40
60
80
100
0
0.3
0.6
0.9
1.2
1.5
Square root of time factor
Terzaghi theory
Eq.19
Beitou clay
Panama clay
Hsiniy clay
Neihou clay

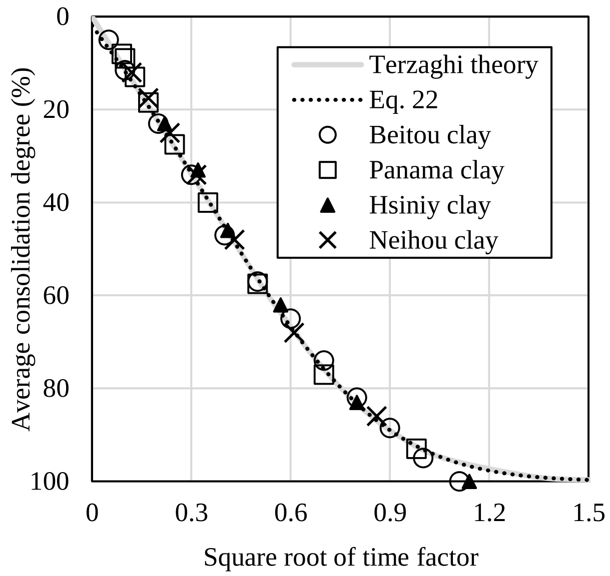
(a)


Average consolidation degree (%)
0
20
40
60
80
100
0
0.3
0.6
0.9
1.2
1.5
Square root of time factor
Terzaghi theory
Eq. 20
Beitou clay
Panama clay
Hsiniy clay
Neihou clay

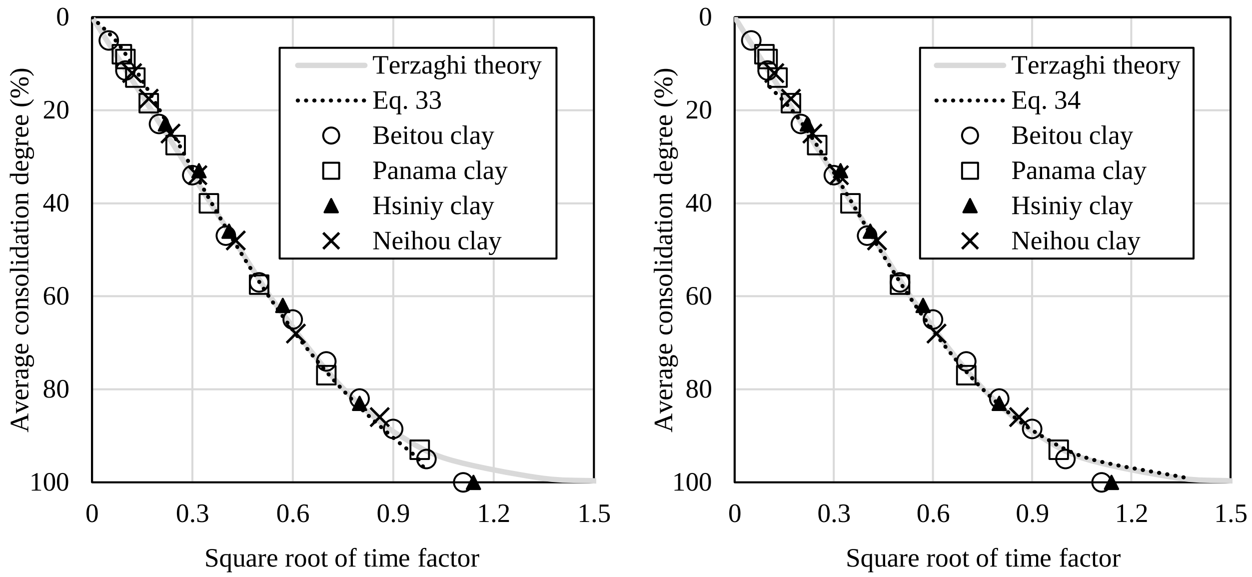
(b)


Average consolidation degree (%)
0
20
40
60
80
100
0
0.3
0.6
0.9
1.2
1.5
Square root of time factor
Terzaghi theory
Eq. 21
Beitou clay
Panama clay
Hsiniy clay
Neihou clay

(c)


Average consolidation degree (%)
0
20
40
60
80
100
0
0.3
0.6
0.9
1.2
1.5
Square root of time factor
Terzaghi theory
Eq. 22
Beitou clay
Panama clay
Hsiniy clay
Neihou clay

(d)



(e) (f)

Fig. 8. Performance of the developed models in comparison with the results of Oedometer test performed on different soft clays

6. External validation

The most common statistical indicators have been discussed and calculated for developed models so far (R^2, RMSE and MAE). However, these are sometimes insufficient and more numerical examination is necessary to increase the reliability of generated mathematical expressions. External validation, therefore, should be conducted through new criteria proposed by Roy and Roy [40] and Golbraikh and Tropsha [41] to this end. Table 11 contains these criteria in conjunction with their acceptable range and the calculated amounts for MEP-generated models. It states that all of the models satisfy the required conditions and show accurate performance within the applicable range.

Conclusion

Simple non-piecewise functions have been developed for rectangular initial excess pore-water pressure. Consolidation degree of a specified time factor can be easily calculated and vice versa (Implicit and explicit forms). Variety of accuracy provides a good opportunity to select the right one which matches with desired precision. A parametric study was also performed, and it was observed that in the developed models for calculating the degree of consolidation for each depth, the depth factor has slightly higher contribution into the prediction compared to the time factor. Furthermore, the results of Oedometer test on four clayey samples in the literature are given and comparison of U-T_v shows that developed MEP models hold well and are also applicable to the real-world problems.

References

[1] Terzaghi, K. 1925. "Principles of soil mechanics, IV—Settlement and consolidation of clay." *Engineering News-Record* 95 (3):874-878.

[2] Clayton, C.R.I., Steinhagen, H.M., Powrie, W., Penman, A., Terzaghi, K. 1997. "Terzaghi's theory of consolidation, and then discovery of effective stress.." *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering* 125 (2):126.

[3] Sivaram, B., Swamee, P.K. 1977. "A computational method for consolidation-coefficient." *Soils and Foundations* 17 (2):48-52.

[4] Sridharan, A., Murthy, N.S., Prakash, K. 1987. "Rectangular hyperbola method of consolidation analysis." *Géotechnique* 37 (3):355-368.

[5] Darrag, A.A., El Tawil, M.A. 1993. "The consolidation of soils under stochastic initial excess pore pressure." *Applied mathematical modelling* 17 (11):609-612.

[6] Tewatia, S.K., Bose, P.R., Sridharan, A., Rath, S. 2007. "Stress induced time dependent behavior of clayey soils." *Geotechnical and Geological Engineering* 25 (2):239-255.

[7] Chan, A.H.C. 2003. "Determination of the coefficient of consolidation using a least squares method." *Géotechnique* 53 (7):673-678.

[8] Singh, S.K., Swamee, P.K. 2008. "Approximate simple invertible equations for consolidation curves under triangular excess pore-water pressures." *Geotechnical and Geological Engineering* 26 (3):251-257.

[9] McKinley, J.D., Sivakumar, V. 2009. "Coefficient of consolidation by plotting velocity against displacement." *Géotechnique* 59 (6):553-557.

[10] Lovisa, J., Read, W., Sivakugan, N. 2010. "Consolidation behavior of soils subjected to asymmetric initial excess pore pressure distributions." *International Journal of Geomechanics* 10 (5):181-189.

[11] Lovisa, J., Read, W., Sivakugan, N. 2011. "A critical reappraisal of the average degree of consolidation." *Geotechnical and Geological Engineering* 29 (5):873-879.

[12] Lovisa, J., Read, W., Sivakugan, N. 2012. "Calculating cυ based on non-uniform initial excess pore pressure." *Geotechnique* 63 (10):885-886.

[13] Tewatia, S.K., Bose, P.R., Sridharan, A. 2012. "Fastest rapid loading methods of vertical and radial consolidations." *International Journal of Geomechanics* 13 (4):332-339.

[14] Lovisa, J., Read, W., Sivakugan, N. 2013. "Time factor in consolidation: Critical review." *International Journal of Geomechanics* 13 (1):83-6.

[15] Al-Zoubi, M.S. 2015. "Consolidation Analysis by the Extended Taylor Method (ETM)." *Jordan Journal of Civil Engineering* 9 (1):253-265.

[16] Di Francesco, R. 2011a. "Exact solution to Terzaghi's consolidation equation." *arXiv preprint arXiv:1102.2060*.

[17] Di Francesco, R. 2011b. "Exact solutions of two-dimensional and tri-dimensional consolidation equations." *arXiv preprint arXiv:1103.6084*.

[18] Di Francesco, R. 2013. "Exact solution of Terzaghi's consolidation equation and extension to two/three-dimensional cases (3th versione)." *arXiv preprint arXiv:1301.3020*.
[19] Holland, J.H. 1992. “Genetic algorithms.” *Scientific american* 267 (1):66-73.
[20] Fogel, L.J., Owens, A.J., Walsh, M.J. 1966. "Artificial intelligence through simulated evolution."
[21] Koza, J.R. 1994. “Genetic programming as a means for programming computers by natural selection.” *Statistics and computing* 4 (2):87-112.
[22] Miller, J.F., Thomson, P. 2000. “Cartesian genetic programming.” *European Conference on Genetic Programming* :121-132, Springer, Berlin.
[23] Mohammadzadeh, D., Bazaz, J.B., Alavi, A.H. 2014. "An evolutionary computational approach for formulation of compression index of fine-grained soils." *Engineering Applications of Artificial Intelligence* 33:58-68.
[24] Brameier, M.F., Banzhaf, W. 2007. “Linear genetic programming.” *Springer Science & Business Media*, New York.
[25] Oltean, M., C. Grosan. 2003. "A comparison of several linear genetic programming techniques." *Complex Systems* 14 (4):285-314.
[26] Gandomi, A.H., Alavi, A.H., Sahab, M.G. 2010. "New formulation for compressive strength of CFRP confined concrete cylinders using linear genetic programming." *Materials and Structures* 43 (7):963-983.
[27] Alavi, A.H., Gandomi, A.H. 2011. "A robust data mining approach for formulation of geotechnical engineering systems." *Engineering Computations* 28 (3):242-274.
[28] Gandomi, A.H., Alavi, A.H., Yun, G.J. 2011. "Formulation of uplift capacity of suction caissons using multi expression programming." *KSCE Journal of Civil Engineering* 15 (2):363-373.
[29] Alavi, A.H., Gandomi, A.H., Sahab, M.G., Gandomi, M. 2010. "Multi expression programming: a new approach to formulation of soil classification." *Engineering with Computers* 26 (2):111-118.
[30] Alavi, A.H., Mollahasani, A., Gandomi, A.H., Boluori Bazaz, J. 2012. “Formulation of secant and reloading soil deformation moduli using multi expression programming.” *Engineering Computations* 29 (2):173-197.
[31] Alvai, A.H., Gandomi, A.H. 2012. “Energy-based numerical models for assessment of soil liquefaction.” *Geoscience Frontiers* 3 (4):541-555.
[32] Alavi, A.H., Gandomi, A.H. Nejad, H.C., Mollahasani, A. Rashed, A. 2013. “Design equations for prediction of pressuremeter soil deformation moduli utilizing expression programming systems.” *Neural Computing and Applications* 23 (6):1771-1786.
[33] Gandomi, A.H. and Alavi, A.H. 2013. “Expression programming techniques for formulation of structural engineering systems.” *Metaheuristic applications in structures and infrastructures*: 439-455.
[34] Gandomi, A.H., Yun, G.J., Alavi, A.H. 2013. “An evolutionary approach for modeling of shear strength of RC deep beams.” *Materials and structures* 46 (12):2109-2119.
[35] Sarveghadi, M., Gandomi, A.H., Bolandi, H., Alavi, A.H. 2015. “Development of prediction models for shear strength of SFRCB using a machine learning approach.” *Neural Computing and Applications*: 1-10.
[36] Gandomi, A.H., Faramarzifar, A., Rezaee, P.G., Asghari, A., Talatahari, S. 2015. “New design equations for elastic modulus of concrete using multi expression programming.” *Journal of Civil Engineering and Management* 21 (6):761-774.
[37] Gandomi, A.H., Sajadi, S., Kiani, B., Huang, Q. 2016 “Genetic programming for experimental big data mining: A case study on concrete creep formulation.” *Automation in Construction* 70:89-97.
[38] Oltean, M. 2004. “Multi-expression programming source code” [online]. Availabe from Internet: http://www.mepx.org/source_code.html

[39] Taji, I., Ghorbani, S., De Brito, J., Tam, V.W., Sharifi, S., Davoodi, A., Tavakkolizadeh, M. 2019. "Application of statistical analysis to evaluate the corrosion resistance of steel rebars embedded in concrete with marble and granite waste dust." *Journal of Cleaner Production* 210:837-846.
[40] Roy, P.P., Roy, K. 2008. "On some aspects of variable selection for partial least squares regression models." *QSAR & Combinatorial Science* 27 (3):302-313.
[41] Golbraikh, A., Tropsha, A. 2002. "Beware of q^2!" *Journal of molecular graphics and modelling* 20 (4):269-276.
[42] Feng, T.W., Lee, Y.J. 2001. "Coefficient of consolidation from the linear segment of the t 1/2 curve" *Canadian Geotechnical Journal* 38 (4):901-909.

Table 11 Results of external validation

No.	Formula	Conditions	Eq. 19	Eq. 20	Eq. 21	Eq. 22	Eq. 35	Eq. 36	Eq. 37	Eq. 38	Eq. 39	Eq. 40
1	$k=\frac{\sum_{i=1}^{n}(a_i\times c_i)}{\sum_{i=1}^{n}a_i^2}$	$0.85<k<1.15$	1.0009	0.9988	1.0036	1.0001	1.0048	1.0235	0.9999	0.9987	1.0021	1.0050
2	$k'=\frac{\sum_{i=1}^{n}(a_i\times c_i)}{\sum_{i=1}^{n}c_i^2}$	$0.85<k'<1.15$	0.9991	1.0012	0.9963	0.9999	0.9611	0.9726	0.9968	1.0010	0.9946	0.9950
3	$m=\frac{R^2-R_O^2}{R^2}$	$\lvert m\rvert<0.1$	-0.0001	-0.0002	-0.0002	0.0000	-0.0287	-0.0302	-0.0047	-0.0015	-0.0040	-0.0005
4	$n=\frac{R^2-R_O'^2}{R^2}$	$\lvert n\rvert<0.1$	-0.0001	-0.0003	-0.0005	0.0000	-0.0250	-0.0413	-0.0047	-0.0015	-0.0040	0.0047
5	$R_m=R^2\times\left(1-\sqrt{\lvert R^2-R_O^2\rvert}\right)$	$R_m>0.5$	0.9881	0.9841	0.9840	0.9974	0.8080	0.7928	0.9269	0.9605	0.9329	0.9763
	where											
	$R_O^2=1-\frac{\sum_{i=1}^{n}(c_i-a_i^o)^2}{\sum_{i=1}^{n}(c_i-\overline{c}_i)^2}$		0.9999	0.9999	0.9997	0.9999	0.9976	0.9837	0.9999	0.9999	0.9999	0.9999
	$R_O'^2=1-\frac{\sum_{i=1}^{n}(a_i-c_i^o)^2}{\sum_{i=1}^{n}(a_i-\overline{a}_i)^2}$		0.9999	0.9999	0.9999	0.9999	0.9940	0.9942	0.9999	0.9999	0.9999	0.9948
	$a_i^o=k\times c_i$											
	$c_i^o=k'\times a_i$											