Effective stress parameter in unsaturated soils: an evolutionary-based prediction model

Alireza Ahangar Asr PhD

Lecturer, School of Science, Engineering, and Environment, Directorate of Civil Engineering, University of Salford, Manchester, UK (corresponding author: a.ahangarasr@salford.ac.uk)

Akbar A. Javadi PhD

Professor of Geotechnical Engineering and Director of Global Development (Engineering), Department of Civil Engineering, College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK

Deformations and failures in unsaturated soils are influenced directly by the effective stress calculated using the stress equation affected by the effective stress parameter. A data-mining-based approach was implemented in this research to develop a prediction model for the effective stress parameter in unsaturated soils. The proposed modelling approach took an evolutionary computing technique to find polynomial models that are structured and explicit. A combination of the genetic algorithm and the least squares approach was implemented to search for the most suitable polynomial structures and the corresponding term parameters in the models. A set of unsaturated soil triaxial test results were used to develop the model. The model was evaluated based on its performance for making predictions using unseen data to validate generalisation capabilities. Model predictions were compared to laboratory and artificial neural-network-based model performance data. A sensitivity analysis assessed the level and form of contributions that input parameters made to the developed model. The developed model accurately captured and redeveloped the intrinsic connections between the introduced input parameters and produced effective stress parameter predictions which competed with the artificial neural network model in terms of accuracy of predictions and outperformed it with regards to the model structure, simplicity and transparency.

Notation

a_j	constant
F	function constructed by the modelling process
f	function defined by the user
h_b	air entry value
m	number of terms of the target expression/polynomial model
N	number of data points/lines on which the coefficient of determination was calculated
$(u_a - u_w)$	suction
X	matrix of input variables
Y_a	actual output parameter value
Y_p	EPR-model-predicted value for the output parameter
y	estimated vector of output
y_a	experimental/actual parameter value
y_p	model prediction
θ_r	volumetric water content under residual conditions
θ_s	volumetric water content under saturated conditions
λ	soil-water characteristic curve slope
$\sigma_3 - u_a$	net confining stress
χ	effective stress parameter

1. Introduction

Understanding the mechanical behaviour and analysing the stability of foundations, natural slopes and geotechnical structures at the first instance is reliant on evaluating the shear strength of soils involved. An important requirement for an economically feasible design is accurate shear strength prediction of geo-materials under both saturated and unsaturated conditions.

Changes in soil properties inevitably affect the effective stress parameter, and this in turn changes the shear strength in unsaturated soils. Öberg and Sällfors (1997) assumed that considering the effective stress parameter to be the same as (equivalent to) the degree of saturation, it can make the prediction of the shear strength of unsaturated soils easier. However, Loret and Khalili (2002) emphasise that, besides the suggestion by Öberg and Sällfors (1997), the effective stress parameter and shear strength of unsaturated soils are dependent on the properties of considered soil and its structure too. The properties of unsaturated soils can also be determined using the soil-water characteristic curve (SWCC). Empirical methods are suggested in the literature to predict the shear strength in unsaturated soils with the SWCC (Fredlund *et al.*, 1996; Vanapalli *et al.*, 1996). A relationship between the plasticity index of the soil and κ was presented by Garven and Vanapalli (2006). Experimental works outcomes from some researchers showed that the net confining pressure noticeably affects the SWCC and the parameters of this curve change by variations in stress (Lee *et al.*, 2005). Khalili and Khabbaz (1998) suggested that the effective stress parameter can be considered first where suction values drop below the bubbling pressure. They also established that the relationship between the matric suction logarithm and the effective stress parameter can be considered linear (Khalili and Khabbaz, 1998).

A similar picture for the effective stress parameter was shown by Xu (2004) by using fractal theory. In the work by Xu (2004), the SWCC was implemented to make reasonable estimations of the surface fractal dimension in soil. Equation 1 was suggested by Russell and Khalili (2006) to calculate the effective stress parameter in sands:

$$\chi = \begin{cases} 1 & \text{for } \frac{(u_a - u_w)}{(u_a - u_w)_b} < 1 \\ \left[\frac{(u_a - u_w)}{(u_a - u_w)_b} \right]^{-0.55} & \text{for } 1 < \frac{(u_a - u_w)}{(u_a - u_w)_b} < 25 \\ 25^{0.45} \left[\frac{(u_a - u_w)}{(u_a - u_w)_b} \right]^{-1} & \text{for } \frac{(u_a - u_w)}{(u_a - u_w)_b} > 25 \end{cases}$$

1.

In Equation 1, $(u_a - u_w)_b$ is the air entry value in the drying process that represents the air expulsion value under wetting conditions. Zargarbashi and Khalili (2011) highlighted the stress state as an influencing factor on the bubbling pressure and emphasised that the corrected value of this parameter needs to be used for estimating the effective stress parameter to a high level of accuracy. A number of empirical relationships are available in the literature to predict shear strength in unsaturated soils, but there is no comprehensive equation/formula that can apply to all unsaturated soil types in predicting shear strength (Fazeli *et al.*, 2009; Garven and Vanapalli, 2006).

A percolation-theory-based numerical method was proposed by Arvin *et al.* (2007) that estimated the effective stress parameter. In the research work by Arvin *et al.* (2007), the SWCC was used to determine the distribution of pore sizes in soil and a conceptual model was created using the percolation theory. The proposed model was used to determine directly the effective stress parameter, but this method would be considerably expensive from a computational point of view, if it was considered to be extended to make predictions in high-suction ranges.

In the past decade, some researchers used artificial neural networks (ANNs) to capture relations between the shear strength of unsaturated soils and the contributing physical properties (Kayadelen, 2008; Lee *et al.*, 2003). These models, however, did not take into consideration the effects of sample preparation methods as well as the stress states during these processes. Results from processing the triaxial unsaturated shear tests revealed that the effective stress parameter varied significantly with changes in net mean stress values under constant suction (Russell and Khalili, 2004, 2006).

An ANN model for making estimations of the effective stress parameter was developed by Ajdari *et al.* (2012), which is a requirement if the shear strength in unsaturated soils is intended to be estimated. The input variables used in the research by Ajdari *et al.* (2012) were the matric suction, net mean stress and SWCC parameters in unsaturated soils that were obtained from various triaxial test results from the literature. Ajdari *et al.* (2012) also investigated the effect that net stress value could have on the effective stress parameter.

This paper applies the evolutionary polynomial regression (EPR) modelling approach to create an innovative and comprehensive model for estimating the effective stress parameter in saturated

soils. The presented model has a simple polynomial structure which is chosen as the optimal model structure based on satisfying accuracy (coefficient of determination (COD) parameter), simplicity, capability in representing the true effective stress variations based on intended intrinsic relations between contributing input parameters as well as the scientific understanding that is available from the literature about the complex problem of effective stress parameter estimation in geotechnical engineering. The proposed model reflects expected sensitivity to contributing/input parameters in line with the expectations based on the available knowledge on the matter.

2. Evolutionary-based polynomial modelling approach

EPR is a data-mining-based modelling technique combining numerical and symbolic regression methods. As a result of this modelling methodology, polynomial models are developed. Polynomial structures have many mathematical specifications that make them very attractive for users. An important specification of this method is that EPR uses genetic algorithm (GA) to conduct evolutionary searches to find suitable exponent values for the terms in the polynomial expressions. This feature facilitates computational implementation of the algorithm as well as efficiently searching for an explicit expression and results in a better capacity to control the complexity of the generated polynomial expression (Giustolisi and Savic, 2006).

EPR is a method based on machine learning and is data driven and relies on evolutionary computing, aiming to find polynomial structures to represent a system/parameter. Assume a physical system with an output value (y) that is dependent on a group of inputs (X) and parameters (θ). This system can then be formulated as follows (Equation 2) from a mathematical point of view:

$$2. \quad y = F(X, \theta)$$

In this equation, F is defined as a function in an m -dimensional space with m being considered the number of inputs. In the EPR modelling process, to control the complexity of the developed polynomial models, that is, to avoid the problem of mathematical expressions that grow fast and continue to become longer and longer in length (increasing number of terms) with time, the evolutionary procedure is arranged to be conducted in such a way that it searches for the exponents of a polynomial function in which the maximum number of terms is fixed. As a result of a single run of the programme, several expressions with increasing numbers of terms up to a limit set by the user to allow the optimum number of terms to be selected will be developed. The form of polynomial equations/expression of EPR-based models can, in general, be presented as (Giustolisi and Savic, 2006)

$$3. \quad y = \sum_{j=1}^m F[X, f(X), a_j] + a_0$$

In this equation form, y is the estimated vector of output of the process, a_j is a constant, F is a function constructed by the process, $\mathbf{X}$ is the matrix of input variables, f is a function defined by the user and m is the number of terms of the target expression/polynomial model.

In general, EPR uses a technique for constructing symbolic models that works in two stages. First, EPR implements standard GA to look for the best structure for the function. This means vectors corresponding to independent inputs, $X_s = 1:k$, are combined, and at the second stage, least squares (LS) regression is conducted with the aim of finding the θ (parameters that are adjustable) for each and every combination of inputs constructed. This strategy ensures that a global and comprehensive search algorithm is used not only for finding the best set of input combinations but also for searching for the most suitable related exponents – all at the same time – considering the cost function that is defined by the model developer/user (Giustolisi and Savic, 2006). To evaluate the adjustable parameters, a_j , the linear LS method is adopted, which operates based on minimisation of the sum of squared errors (SSE) as its cost function. The SSE function is implemented to lead the search process towards the best-fit model and can be expressed as

$$4. \quad SSE = \frac{\sum_{i=1}^N (y_a - y_p)^2}{N}$$

In this function, y_a is the experimental/actual parameter value and y_p is the model prediction aiming to be closest to the actual value to reduce the error margin.

The global search aiming to find the best form for the EPR equation/model is conducted using a standard GA over the values in the user-defined 'exponents vector'. GA operates based on the principles of Darwinian evolution beginning by randomly creating an initial 'solutions population'. Each parameter set in the population is considered to stand for chromosomes of the individuals. Based on how well/poor each individual performs in its relevant environment, a fitness is assigned to that individual. The next generation is then created by way of cross-over and mutation operations, with the probabilities P_c (probability of cross-over) and P_m (probability of mutation), respectively. Fit individuals are then chosen for the 'mating' process, while the weak ones (individuals) are removed/dead. A child (offspring) is then created by 'mated parents' carrying a set of chromosomes that is a mix of the chromosomes of the parents. The EPR modelling process uses integer GA coding with single point cross-over to find where (on which term) the candidate exponents will be located in the polynomial model (Giustolisi and Savic, 2006).

A set of termination criteria including the maximum number of generations, the maximum number of terms in the target

mathematical expression or a particular allowable error is used to bring the EPR modelling process to a stop eventually. The modelling process comes to a halt when either one of the criteria is met. EPR has been used to develop various models so far to represent the complicated behaviour of various geotechnical and civil engineering parameters, systems and materials, including unsaturated soils (Ahangar Asr and Javadi, 2016; Ahangar Asr *et al.*, 2015, 2018; Cuisinier *et al.*, 2013; Hussain *et al.*, 2015; Javadi and Rezaei, 2009). Figure 1 presents a flow diagram of the EPR modelling procedure.

3. Processing and preparation of data used in model development and validation

A comprehensive set of unsaturated triaxial test data from the literature (Ajdari *et al.*, 2012; Bishop and Blight, 1963; Khalili

Figure 1. Flow diagram of the EPR procedure

et al., 2004; Lee *et al.*, 2005; Rahardjo *et al.*, 2004; Rassam and Williams, 1999; Russell and Khalili, 2004, 2006) was used to develop the EPR-based model in this research to predict the effective stress parameter in unsaturated soils.

The developed model was assumed to be affected by some parameters (inputs) proven to be influential on the effective stress parameter in unsaturated soils from the literature. Therefore, these parameters (the air entry value, volumetric water content under residual conditions, volumetric water content under saturated conditions, the SWCC slope and net confining stress and suction, shown in Table 1) were considered input parameters in the model development process with the single output parameter set to the effective stress parameter in unsaturated soil. Statistical analysis was performed on various combinations of training and testing data to make sure that the best representative and closest combination (from the statistical point of view) in terms of key statistical parameter values (mean and standard deviation) was chosen to help develop the most powerful and representative EPR model.

For the model to be representative, there need to be enough data that are used and seen by the software to be able to learn adequately the intended intrinsic relations between contributing input parameters. On the other hand, for the machine-learning-based models, the crucial capability expected of models is their ‘generalisation’ potential, meaning that the developed models have to be able to produce accurate predictions in the case of any data being presented as input to them are unseen by them during the model development stage. For this reason, the available data for model development by the EPR approach in this research were divided into training and testing sets. The training data were used to develop the model; the testing data were kept unseen from the software during the whole model development stage and were used only after the model had been finalised to verify the generalisation capabilities of the developed model.

In the literature, where machine learning techniques are used, division of the data has traditionally been around 80% (used for training) to 20% (used for testing/verification/examining generalisation capabilities) to develop prediction models (Javadi and Rezaia, 2009; Johari *et al.*, 2006a, 2006b). Therefore, a similar approach is taken in this research work.

Just over 120 lines of data (the total number of cases in the database used) were divided into training and testing data sets. By using the created training and testing data sets (finalised after completing a statistical analysis – explained below), the training data were used to train EPR to develop the EPR-based model, while the remaining cases (test data) were kept unseen from EPR during the model development process and were used later on to validate the developed model.

The adopted strategy for testing the generalisation capabilities of the developed model could also be translated as the ‘hold-out’ approach, as 21 out of 121 lines of data (about 18% of the whole data available) were ‘held out’ during the training stage of the model development process. Once the model development had been completed, the ‘left-out’ data lines were used to verify the generalisation capabilities of the model in making predictions based on the left-out (previously unseen by EPR during the model creation process) data.

The data used to develop and verify the model were statistically analysed for the most statistically consistent training and testing sets to be selected and used in the development of the presented models. The aim was to ensure utmost consistency between the training and testing data sets to optimise the learning process to help develop best possible models. In other words, the data analysis process was intended to make sure that the statistical properties of the data in the training and testing data subsets were as close as possible to one another to reflect the fact that they represent the same population from the statistical point of view. Many combinations for testing/verification and training data

Table 1. Parameters involved in developing EPR models and ranges of values of the parameters as well as statistical analysis parameter values for training and testing data sets

			Training data ranges	Testing data range	Mean – training data	Mean – testing data	Standard deviation – training data	Standard deviation – testing data
Input parameters	$u_a - u_w$	Suction	0–612 kPa	76–496 kPa	138	122	131	92
	θ_r	Volumetric water content under residual conditions	0–28.35	5.502–21.54	10.74	8.76	10.38	11.08
	θ_s	Volumetric water content under saturated conditions	23.82–55.95	25.43–52	37.19	43.18	9.26	8.42
	λ	SWCC slope	0.19–11.82	0.94–8.33	1.37	2.41	2.65	2.72
	$\sigma_3 - u_a$	Net confining stress	0–400 kPa	50–300 kPa	113	145	105	98
Output parameter	h_b	Air entry value	1–200 kPa	27–125 kPa	32	38	46	28
	χ	Effective stress parameter	0.091–1	0.24–1	0.676	0.769	0.293	0.153

bunches were considered to start with. The standard deviation as well as the mean values was calculated for all parameters that were considered to be contributing to the final model. Statistical parameter values for training and testing data for the final combination used in the model development and verification processes are shown in Table 1.

Around 18% of data were kept unseen during the model development stage and were used to examine the generalisation capabilities of the developed model. However, while the developed model was created based on a wide range of data and presented very good capabilities in extrapolating when making predictions of the effective stress parameter, due to the fact that extrapolation does not/cannot have any certain limits and to ensure that the developed model is entirely reliable and safe to be implemented using various ranges of unseen data, efforts have been put in place in this research to avoid extrapolation in the developed model predictions (Javadi and Rezanian, 2009). To do this, the training and testing data ranges chosen (before doing the comprehensive statistical analysis detailed earlier) were carefully checked to ensure that all parameter values in the testing data sets were within the ranges of data chosen to be used for training EPR and for developing the EPR model to avoid extrapolation in making predictions to ensure that the predictions were completely reliable from a statistical and engineering point of view. Although (as highlighted earlier) the model predictions seem to be very good for any range of unseen data, to ensure that the reliability factor applies, if the ranges of input data fall outside of the ranges used in the training stage of the model development process (in industry applications of the developed model for instance), retraining of the model using new data that include the new wider ranges is strongly recommended.

4. Model development

A computer programme running on Matlab as the main platform was developed over the years in collaboration between the University of Exeter in the UK and the Technical University of Bari, Italy, to be used to implement the EPR technique (Giustolisi and Savic, 2006; Giustolisi *et al.*, 2008; Rezanian *et al.*, 2008).

To control the length, complexity, type of functions used in developing the model, number of terms in the polynomial model, range of the exponents considered in the model structure and also the number of generations to be used by the programme to complete the evolutionary modelling processes, some constraints were considered. The COD parameter (fitness equation – Equation 5) was used to check model fitness/accuracy level as the model development processes were progressing.

$$\text{COD} = 1 - \frac{\sum_N (Y_a - Y_p)^2}{\sum_N [Y_a - (1/N) \sum_N Y_a]^2}$$

5.

In this equation, Y_a is the actual output parameter value, Y_p is the EPR-model-predicted value for the output parameter and N shows the number of data points/lines on which the COD was calculated.

The COD is a simple parameter used in determining the accuracy of prediction models and has been used in the past years by researchers as an effective parameter for evaluating the level of accuracy of developed models in making predictions of output parameters particularly in data-mining- and machine-learning-based modelling techniques/approaches (Ahangar Asr *et al.*, 2015, 2018).

If the fitness level of the model (accuracy) was not considered acceptable (according to the calculated COD value) or if any other criteria set to terminate the modelling process (user-defined maximum number of generations and/or maximum number of terms) were not met at the end of each model development cycle, the current model will be subjected through another evolution for a new model to be developed. This process continues until a final model is reached and the modelling process is terminated by meeting modelling process termination criteria.

When using modelling techniques that adopt regression as part of their modelling processes, the term ‘fitness’ is usually used as a means of showing how closely the outcomes of equation developed by regression match the actual data points. Also, there is a fact that is accepted by the majority of researchers, and it is that the best model is the one that is the simplest possible satisfying the requirements (i.e. reflects all known contributing parameters). In other words, if there are models developed to represent a system/parameter and all those models could be considered equivalent without considering the simplicity factor, the simplest model must be picked to represent the intended system/parameter (Giustolisi and Savic, 2006). In the EPR approach, in this work, to achieve the best choice of model from among all the ones that were developed in the process, the simplest model that included all the parameters that are proved in the literature that affect the effective stress parameter in unsaturated soils and representing the highest possible level of accuracy (measured by the COD parameter) was chosen.

Thirty-three models were developed in the model development process in this research using the EPR approach with the number of terms ranging between 1 and 15. Some of these models were, however, ruled out on the basis that in each one of them, at least one of the input parameters reported in the literature to be affecting the effective stress parameter in unsaturated soils was excluded. The remaining models were showing various performance levels in terms of accuracy (COD value) and complexity (number of terms).

Equations 6–8 are three example models that included all input parameters, with Equation 8 being chosen as the final model after further analysis stages were completed (details of further model analysis will follow).

$$\begin{aligned}
 \chi &= 3.488 \times 10^{-11}(\sigma_3 - u_a)^{0.5} + 1.775h_b^{-2} \\
 &\quad - 0.479h_b^{-0.5} \times \theta_r^{0.5} - 23.334\theta_s^{-1} \times \lambda^{-0.5} \\
 &\quad + 180.694\theta_s^{-1} + 1.405\theta_s^{0.5} \\
 &\quad - 0.016(u_a - u_w)^{0.5} \times \theta_s^{0.5} + 0.002(u_a - u_w) \\
 &\quad + 9.302 \times 10^{-6}(u_a - u_w)\theta_r^{0.5} \times \theta_s \times \lambda^{-0.5} \\
 &\quad - 1.975 \times 10^{-7}(u_a - u_w)^2 h_b^{0.5} + 6.657 \\
 &\quad \times 10^{-4}(\sigma_3 - u_a)^{0.5} \lambda^{-2} \\
 &\quad + 0.039(\sigma_3 - u_a)^{0.5} - 2.224 \\
 6. \quad &\times 10^{-6}(\sigma_3 - u_a)^2 \lambda^{-0.5} - 11.477
 \end{aligned}$$

$$\begin{aligned}
 \chi &= 2.198 \times 10^{-10}(\sigma_3 - u_a)^{0.5} \times \theta_s^{0.5} + 1.867 \times h_b^{-2} \\
 &\quad + 0.496h_b^{-0.5} \times \theta_r^{0.5} - 23.176\theta_s^{-1} \times \lambda^{-0.5} \\
 &\quad + 167.482\theta_s^{-1} + 1.332\theta_s^{0.5} \\
 &\quad - 0.016(u_a - u_w)^{0.5} \times \theta_s^{0.5} + 0.002(u_a - u_w) \\
 &\quad + 9.215 \times 10^{-6}(u_a - u_w)\theta_r^{0.5} \times \theta_s \times \lambda^{-0.5} \\
 &\quad - 1.949 \times 10^{-7}(u_a - u_w)^2 h_b^{0.5} + 1.61 \\
 &\quad \times 10^{-4}(\sigma_3 - u_a)^{0.5} \lambda^{-3} + 0.04(\sigma_3 - u_a)^{0.5} - 2.291 \\
 7. \quad &\times 10^{-6}(\sigma_3 - u_a)^2 \lambda^{-0.5} - 10.681
 \end{aligned}$$

$$\begin{aligned}
 \chi &= 1.083 \times 10^{-3} \times \lambda^{-4} + 0.197\theta_s^{0.5} - 7.351 \times 10^{-3} \\
 &\quad \times h_b^{0.5} \lambda^{-2} + 0.065h_b^{0.5} \\
 &\quad - 0.077(u_a - u_w)^{0.5} + 1.204 \times 10^{-3}(u_a - u_w) \\
 &\quad + 7.014 \times 10^{-13}(u_a - u_w)^2 \theta_r^4 \lambda^{-3} + 1.419 \\
 &\quad \times 10^{-3}(\sigma_3 - u_a) - 1.142 \times 10^{-6}(\sigma_3 - u_a)^2 \theta_r^{0.5} \\
 8. \quad &- 0.29
 \end{aligned}$$

Table 2 shows, for Equations 6–8, their corresponding COD values calculated using Equation 5 as an indication of the level of accuracy of the models in making predictions of the output parameter (effective stress parameter in unsaturated soils) for both training and testing data as well as some other criteria applied to choose a final model. It presents the selected model with the optimum combination of COD values for testing and training data along with two other models included as examples of either highly complex models (with too many terms in the polynomial

structure) or the ones with low accuracy levels and/or generalisation capabilities (poor performance in terms of accuracy of predictions for unseen data – low testing data COD values).

At the end of the model development process using the training data, a validation process began that was completed by implementing the testing data set that had been kept completely unseen by EPR during the model development process with the aim of creating the potential for a robust testing for the model to examine its generalisation capabilities to unseen cases of data. This was done to test the models developed to see to what extent exactly these models are capable of generalising the training data to the cases not experienced (seen) by EPR before. Two COD values were calculated for every one of the 33 developed models (values for three example models are shown in Table 2).

As can be easily seen in Table 2, Equations 6 and 7 are more complex than the selected final model (Equation 8) and at the same time also show much lower COD values, particularly for the testing/verification data. It can be seen also that Equation 6, for instance, shows a higher COD value for the training data compared with the selected model. On the other hand, it is more complex than other models and more importantly shows a very low COD for the training data, suggesting a low level of generalisation capabilities. Equation 8, however, shows a very high COD value for the training data (although slightly lower than those of Equations 6 and 7) and is much less complex (showing fewer terms) than Equations 6 and 7 and more importantly presents an excellent generalisation capability by outperforming the other two models (Equations 6 and 7) in this by a very high margin. In other words, although the models presented in the form of Equations 6 and 7 show relatively higher training COD values, their generalisation capabilities to unseen cases of data (represented by COD (testing) values) are significantly lower compared with that of Equation 8, which shows a reasonable balance in accuracy levels for predictions made based on both training and testing data as well as the number of terms (relatively lower level of complexity due to smaller number terms in the polynomial structure).

At the end of the model development process, a combination of all criteria considered leads Equation 8 to be chosen as the final

Table 2. Example models and the final selection based on inclusion of expected input parameters, performance and complexity

Model (polynomial equation)	Performance (COD)		Includes all expected contributing parameters	Complexity (number of terms in the equation)	Selected as the final model
	COD (training): %	COD (testing): %			
6	90.24	16.82	Yes	13	No (poor generalisation capability – low testing COD and higher complexity)
7	82.92	6.52	Yes	14	No (poor generalisation capability – low testing COD and higher complexity)
8	85.83	74.05	Yes	10	Yes (best performance with regard to generalisation capability; optimum/balanced number of terms in relation with performance)

model through EPR model development and analysis process to represent the effective stress parameter in unsaturated soils. This model can be considered the strongest and most balanced one among the 33 models developed in the modelling process in terms of its prediction generalisation capabilities to unseen data, which is considered a crucial characteristic for a model as an indication of being able to be applied to the cases that the data connected to them are not previously included in the training phase of the model in any shape or form, level of complexity and also capability in engaging all known contributing parameters in the developed model.

Figures 2 and 3 present comparisons made between the EPR model (Equation 8) predictions for the effective stress parameter in unsaturated soils and the actual results obtained from the experiments (based on literature data – detailed in an earlier section of the paper dedicated to data preparation) for training and testing data sets, respectively.

The COD parameter is a simple, accurate and powerful means of measuring the accuracy levels of models developed using machine learning approaches and has been used in various research works, including recent publications (Ahangar Asr and Javadi, 2016; Ahangar Asr *et al.*, 2015, 2018; Javadi and Rezaia 2009; Rezaia *et al.*, 2008). Table 2 shows the performance of the selected EPR model in this research in comparison with those of two other example models that were ruled out due to complexity as well as law testing/verification/generalisation COD values. The COD values for the training and testing data for the selected model were 86 and 74%, respectively. Ajdari *et al.* (2012)

Figure 2. EPR model predictions against the experimental data for the unsaturated soil effective stress parameter (training data) – selected model (Equation 8)

Figure 3. EPR model predictions against the experimental data for the unsaturated soil effective stress parameter (testing/verification data) – selected model (Equation 8)

developed a black/grey box adaptive learning ANN-based model for predicting the effective stress parameter in unsaturated soils. The proposed network was a multilayer perceptron network with six neurons in the input layer representing the air entry value, the volumetric water content under residual and saturated conditions, the slope of SWCC, the net confining stress and suction, similar to the set of input parameters introduced in this research work. They used the R -squared (R^2) parameter to determine the accuracy level of their model. The R^2 value for the ANN training and testing data were presented as 0.96 and 0.75, respectively.

Figure 4 presents a comparison between the performances of the ANN and EPR models for the testing data that were unseen by EPR during the model development process with the aim of measuring the generalisation capabilities of the model. It can be clearly seen that the models perform similarly, with the EPR model performing slightly better in some of the data points, which is in line with the similar performance measurement parameter (R^2 and COD) values of 0.75 and 74% for ANN and EPR models, respectively. This can be considered a clear testimony to the high capability of the COD parameter as a measure for reflecting the accuracy level of models, including the EPR model developed in this research work.

The comparison of the performances of ANN and EPR models (Figure 4) also reveals that the proposed EPR model is not only is capable of making accurate predictions of the effective stress parameter in unsaturated soils but also has a clear advantage over the ANN model, and it is the fact that EPR – unlike ANN, which provides a black/grey box model – presents an explicit polynomial model that is accessible to the user to be able to scrutinise the model in ensuring that all known assisting

Figure 4. Comparison between EPR and ANN model predictions for unseen EPR validation/generalisation data

parameters are playing a part, and the form and level of contribution of every assisting parameter can easily be examined by conducting sensitivity analysis on the model.

5. Sensitivity analysis

To be able to confirm the capabilities of the EPR model developed in this research with the aim of predicting the effective stress parameter in unsaturated soils, an analysis was conducted to measure the sensitivity of the developed model to contributing parameters with known contribution level and patterns to the output. The aim of this analysis was to ensure that the model had been able to learn adequately and correctly the relationship between the contributing parameters and the output (effective stress parameter for unsaturated

soils). To provide a better understanding on how important parameters affect the model, well-recognised variations of the predictions made by the model in relation to key contributing parameters (with previously recognised behavioural patterns) were considered. The proposed approach not only has the advantage of analysing sensitivity to two parameters on a single graph but also helps provide clearer interpretations of the conducted sensitivity analyses (Ajdari *et al.*, 2012). Figures 5 and 6 show the variations in the effective stress parameter in relation with changing the net mean stress and bubbling pressure, respectively, considering the effect of suction (Figure 5) and net mean stress (Figure 6) while the remaining parameters in both analyses were set to an average/constant value to ensure representativeness, credibility and clarity.

For the sensitivity analysis to have an understandable meaning, a comparison between SWCCs presented by Bishop and Blight (1963) and Khalili *et al.* (2004) was used. The comparison revealed that significant changes in the bubbling pressures of the soils happened in relation to the varying effective stress parameter while other SWCC parameters – which were subjects of the comparisons too – did not vary noticeably (Ajdari *et al.*, 2012). Therefore, the sensitivity analysis of the effective stress parameter to the changes in the bubbling pressure was conducted and is presented in Figure 5, which shows variations in the effective stress parameter obtained from the selected EPR model (Equation 8) for various net mean (confining) stress values as the bubbling pressure changes. It is clear that effective stress parameter increases with the air entry value, which is a confirmation that the developed model is able to reproduce the expected trend correctly (Ajdari *et al.*, 2012). Figure 6 shows changes in the effective stress parameter against the net confining pressure (mean stress) calculated using the developed selected EPR model while other SWCC parameters were kept unchanged for different suction values. The presented results demonstrate that the expected trends have been correctly captured by the developed/chosen final

Figure 5. Changes in the effective stress parameter in relation to the bubbling pressure (suction = 200 kPa, volumetric water content under residual conditions = 20, volumetric water content under saturated conditions = 35, SWCC slope = 1)

Figure 6. Changes in the effective stress parameter in relation to the net confining pressure (bubbling pressure = 50 kPa, volumetric water content under residual conditions = 0, volumetric water content under saturated conditions = 45, SWCC slope = 0.1)

EPR model for the effective stress parameter in unsaturated soils and the model behaviour is consistent with the expected form of variations in the predictions from the literature as the analysed parameters change (Ajdari *et al.*, 2012; Johari, 2006).

6. Conclusions

EPR was used to develop models to predict the effective stress parameter in unsaturated soils considering six contributing input parameters. Experimental triaxial test data from the literature were used to develop and validate the models in this study. From among the developed models, one was selected based on robustness and complexity factors (the most robust and least complex model with the highest possible generalisation capabilities was chosen). After training, the generalisation capabilities of the selected model were evaluated by verification of its performance using a set of data that had been kept unseen from EPR during the model development process. The results revealed that the proposed model was efficient and robust in successfully capturing the complicated underlying relations between the contributing parameters and predicting the effective stress parameter in unsaturated soils directly from a set of raw experimental measurements to a very good level of accuracy.

A sensitivity analysis was also conducted to investigate the capability of the model in capturing the expected behavioural trends in relations between the unsaturated soil effective stress parameter and the key contributing parameters. The sensitivity analysis results confirmed that the expected behaviour (approved by previous knowledge from the literature) had been successfully picked up by the developed model and reflected in the predictions made.

A comparison was made between the performance of the EPR model developed in this research and an ANN-based model from

the literature. The results confirmed the strong capabilities of the proposed model in making accurate predictions to similar levels done by the ANN model with the exceptional advantage of being presented in the form of an explicit and easy-to-interpret polynomial model providing a clear insight into the connections/intrinsic relations between input parameters and the output, the unsaturated soil effective stress parameter.

Another interesting feature of the EPR approach is that as more data become available, the quality of the model predictions can be improved by retraining EPR with the newly available more comprehensive set of data. This feature highlights the flexibility and strength of the methodology in being able to be stretched to include newly generated data in developing stronger, more accurate models.

REFERENCES

- Ahangar Asr A and Javadi A (2016) Air losses in compressed air tunneling: a prediction model. *Proceedings of the Institution of Civil Engineers – Engineering and Computational Mechanics* **169**(3): 140–147, <https://doi.org/10.1680/jenm.15.00023>.
- Ahangar Asr A, Javadi AA and Khalili N (2015) An evolutionary approach to modelling the thermomechanical behaviour of unsaturated soils. *International Journal for Numerical and Analytical Methods in Geomechanics* **39**(5): 539–557, <http://doi.org/10.1002/nag.2323>.
- Ahangar Asr A, Faramarzi A and Javadi AA (2018) An evolutionary modelling approach to predicting stress–strain behaviour of saturated granular soils. *Engineering Computations: International Journal for Computer-aided Engineering* **35**(8): 2931–2952, <https://doi.org/10.1108/EC-01-2018-0025>.
- Ajdari M, Habibagahi G and Ghahramani A (2012) Predicting effective stress parameter of unsaturated soils using neural networks. *Computers and Geotechnics* **40**: 89–96, <https://doi.org/10.1016/j.compgeo.2011.09.004>.
- Arvin M, Veiskarami M, Ajdari M and Habibagahi G (2007) A percolation approach to determination of effective stress parameter in unsaturated

- soils. In *Proceedings of the 3rd Asian Conference on Unsaturated Soils* (Yin ZZ, Yuan JP and Chiu ACF (eds)). Science Press, Beijing, China, pp. 163–168.
- Bishop AW and Blight G (1963) Some aspects of effective stress in saturated and partly saturated soils. *Géotechnique* **13**(3): 177–197, <https://doi.org/10.1680/geot.1963.13.3.177>.
- Cuisinier O, Javadi AA, Ahangar-Asr A and Masroufi F (2013) Identification of coupling parameters between shear strength behaviour of compacted soils and chemical's effects with an evolutionary-based data mining technique. *Computers and Geotechnics* **48**: 107–116, <https://doi.org/10.1016/j.compgeo.2012.10.005>.
- Fazeli A, Habibagahi G and Ghahramani A (2009) Shear strength characteristics of Shiraz unsaturated silty clay. *Iranian Journal of Science and Technology* **33**(4): 327–341.
- Fredlund DG, Xing A, Fredlund MD and Barbour S (1996) The relationship of the unsaturated soil shear to the soil-water characteristic curve. *Canadian Geotechnical Journal* **33**(3): 440–448, <https://doi.org/10.1139/t96-065>.
- Garven EA and Vanapalli SK (2006) Evaluation of empirical procedures for predicting the shear strength of unsaturated soils. In *Unsaturated Soils 2006* (Miller GA, Zapata CE, Houston SL and Fredlund DG (eds)). American Society of Civil Engineers, Reston, VA, USA, pp. 2570–2592.
- Giustolisi O and Savic DA (2006) A symbolic data-driven technique based on evolutionary polynomial regression. *Journal of Hydroinformatics* **8**(3): 207–222, <https://doi.org/10.2166/hydro.2006.020b>.
- Giustolisi O, Doglioni A, Savic DA and di Pierro F (2008) An evolutionary multi-objective strategy for the effective management of groundwater resources. *Water Resources Research Journal* **44**(1): article W01403, <https://doi.org/10.1029/2006WR005359>.
- Hussain MS, Javadi AA, Ahangar-Asr A and Farmani R (2015) A surrogate model for simulation–optimization of aquifer systems subjected to seawater intrusion. *Journal of Hydrology* **523**: 542–554, <https://doi.org/10.1016/j.jhydrol.2015.01.079>.
- Javadi AA and Rezaei M (2009) Intelligent finite element method; an evolutionary approach to constitutive modelling. *Journal of Advanced Engineering Informatics* **23**(4): 442–451, <https://doi.org/10.1016/j.aei.2009.06.008>.
- Johari A (2006) *Predicting Soil Water Characteristic Curve Using Artificial Intelligence*. PhD thesis, Shiraz University, Shiraz, Iran.
- Johari A, Habibagahi G and Ghahramani A (2006a) Prediction of soil–water characteristic curve using genetic programming. *Journal of Geotechnical and Geo-environmental Engineering* **5**: 661–665, [https://doi.org/10.1061/\(ASCE\)1090-0241\(2006\)132:5\(661\)](https://doi.org/10.1061/(ASCE)1090-0241(2006)132:5(661)).
- Johari A, Habibagahi G and Ghahramani A (2006b) Prediction of soil–water characteristic curve using a genetic based neural network. *Scientia Iranica* **13**(3): 284–294.
- Kayadelen C (2008) Estimation of effective stress parameter of unsaturated soils by using artificial neural networks. *International Journal for Numerical and Analytical Methods in Geomechanics* **32**(9): 1087–1106, <https://doi.org/10.1002/nag.660>.
- Khalili N and Khabbaz MH (1998) A unique relationship of χ for the determination of the shear strength of unsaturated soils. *Géotechnique* **48**(5): 681–687, <https://doi.org/10.1680/geot.1998.48.5.681>.
- Khalili N, Geiser F and Blight G (2004) Effective stress in unsaturated soils: review with new evidence. *International Journal of Geomechanics* **4**(2): 115–126, [https://doi.org/10.1061/\(ASCE\)1532-3641\(2004\)4:2\(115\)](https://doi.org/10.1061/(ASCE)1532-3641(2004)4:2(115)).
- Lee S, Lee S and Kim Y (2003) An approach to estimate unsaturated shear strength using artificial neural network and hyperbolic formulation. *Computers and Geotechnics* **30**(6): 489–503, [https://doi.org/10.1016/S0266-352X\(03\)00058-2](https://doi.org/10.1016/S0266-352X(03)00058-2).
- Lee IM, Sung SG and Cho GC (2005) Effect of stress state on the unsaturated shear strength of a weathered granite. *Canadian Geotechnical Journal* **42**(2): 624–631, <https://doi.org/10.1139/t04-091>.
- Loret B and Khalili N (2002) An effective stress elastic–plastic model for unsaturated porous media. *Mechanics of Materials* **34**(2): 97–116, [https://doi.org/10.1016/S0167-6636\(01\)00092-8](https://doi.org/10.1016/S0167-6636(01)00092-8).
- Öberg A and Sällfors G (1997) Determination of shear strength parameters of unsaturated silts and sands based on the water retention curve. *Geotechnical Testing Journal* **20**(1): 40–48, <https://doi.org/10.1520/GTJ11419J>.
- Rahardjo H, Heng OB and Choon LE (2004) Shear strength of a compacted residual soil from consolidated drained and constant water content triaxial tests. *Canadian Geotechnical Journal* **41**(3): 421–436, <https://doi.org/10.1139/t03-093>.
- Rassam DW and Williams DJ (1999) A relationship describing the shear strength of unsaturated soils. *Canadian Geotechnical Journal* **36**(2): 363–368, <https://doi.org/10.1139/t98-102>.
- Rezaei M, Javadi AA and Giustolisi O (2008) An evolutionary-based data mining technique for assessment of civil engineering systems. *Journal of Engineering Computations* **25**(6): 500–517, <https://doi.org/10.1108/02644400810891526>.
- Russell AR and Khalili N (2004) A bounding surface plasticity model for sands exhibiting particle crushing. *Canadian Geotechnical Journal* **41**(6): 1179–1192, <https://doi.org/10.1139/t04-065>.
- Russell A and Khalili N (2006) A unified bounding surface plasticity model for unsaturated soils. *International Journal for Numerical and Analytical Methods in Geomechanics* **30**(3): 181–212, <https://doi.org/10.1002/nag.475>.
- Vanapalli S, Fredlund D, Pufahl D and Clifton A (1996) Model for the prediction of shear strength with respect to soil suction. *Canadian Geotechnical Journal* **33**(3): 379–392, <https://doi.org/10.1139/t96-060>.
- Xu Y (2004) Fractal approach to unsaturated shear strength. *Journal of Geotechnical and Geoenvironmental Engineering* **130**(3): 264–273, [https://doi.org/10.1061/\(ASCE\)1090-0241\(2004\)130:3\(264\)](https://doi.org/10.1061/(ASCE)1090-0241(2004)130:3(264)).
- Zargarbashi S and Khalili N (2011) Discussion of ‘Shear strength equations for unsaturated soil under drying and wetting’ by Goh Shin Guan, Harianto Rahardjo, and Leong Eng Choon. *Journal of Geotechnical and Geoenvironmental Engineering* **137**(12): 1310–1313, [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0000455](https://doi.org/10.1061/(ASCE)GT.1943-5606.0000455).

How can you contribute?

To discuss this paper, please email up to 500 words to the editor at journals@ice.org.uk. Your contribution will be forwarded to the author(s) for a reply and, if considered appropriate by the editorial board, it will be published as discussion in a future issue of the journal.

Proceedings journals rely entirely on contributions from the civil engineering profession (and allied disciplines). Information about how to submit your paper online is available at www.icevirtuallibrary.com/page/authors, where you will also find detailed author guidelines.