Research Article

Prediction of the Void Ratio Parameter in Mineral Tailings Using Gene Expression Programming

Ali Akbar Heshmati R. [iD]**, Hossein Salehzadeh, and Mehdi Shahidi** [iD]

School of Civil Engineering, Iran University of Science and Technology (IUST), Tehran, Iran

Correspondence should be addressed to Ali Akbar Heshmati R.; heshmati@iust.ac.ir

Mineral tailing deposits are one of the most important issues in the field of geotechnical engineering. The void ratio of mineral tailings is an essential parameter for investigating the geotechnical behavior of tailings. However, there has not yet been a comprehensive empirical formulation for initial prediction of the void ratio of mineral tailings. In this study, the void ratio of various types of mineral waste is estimated by using gene expression programming (GEP). Therefore, taking into consideration the effective physical parameters that affect the estimation of this parameter, eight different models are presented. A reliable experimental database collected from different sources in the literature was applied to develop the GEP models. The performance of the developed GEP models was measured based on coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE). According to the results, the model with effective stress (σ'), initial void ratio (e_0), and parameters of $R^2 = 0.92$, MAE = 0.109, and RMSE = 0.180 performed the best. Finally, a new empirical formulation for the initial prediction of the void ratio parameter is proposed based on the aforementioned analyses.

1. Introduction

Understanding tailing behavior is one of the most challenging issues for both geotechnical and environment engineers. Tailings are defined as mineral waste that is crushed and deposited after the extraction of desired minerals. Constructing a tailing dam usually requires a remarkable amount of loan resources. In practice, mineral waste is applied to reduce the construction costs of tailing dams [1]. It is also difficult to work in the laboratory with mineral waste due to the particular conditions of this type of soils. In addition, according to the International Commission on Large Dams, more than 200 failures of tailing dams have occurred since the early twentieth century. A remarkable volume of mineral waste flows down the tailing dams after a failure occurs. This usually leads to both human deaths and severe environmental pollution. Therefore, understanding the strength and consolidation behavior of mineral waste used in the tailing dams can give us an insight into factors that affect failure occurrence and slope stability.

Several research studies have been carried out to investigate the consolidation, strength behavior (monotonic and cyclic), and permeability of the mineral tailings. The void ratio parameter is known as an important parameter for better understanding of the consolidation behavior of tailings. Several studies have been dedicated to the investigation of the effect of different physical parameters on the void ratio parameter for various types of mineral waste. In this regard, there have been many studies that investigated the effect of the effective stress parameter on the void ratio parameter in mineral tailings [1–7]. The results indicated that increasing the effective stress leads to nonlinear reduction of the void ratio for various types of tailings. Quille and O'Kelly investigated the effect of soil classification on the void ratio in zinc/lead mine tailings [8]. The results showed that the void ratio of fine-grained waste is less than the void ratio of coarse-grained waste. Experimental works conducted by Bonin et al. on gold mine tailings revealed that the void ratio parameter increases with increasing initial void ratio parameter [9]. Qiu and Sego studied the effect of physical

parameters on the consolidation features of copper, gold, and coal mineral tailings [10]. It was observed that the void ratio parameter also increases with increasing water content. It can be concluded from the literature review that the void ratio parameter generally depends on the effective stress (σ'), initial void ratio (e_0), water content (w_w), clay percentage ($c\,(\% < 2\eta m)$), and grain size. However, there are no comprehensive relationships for initially estimating the void ratio parameter based on its influential parameters in mine tailings.

In recent years, soft computing methods such as artificial neural networks (ANNs), adaptive network-based fuzzy inference system (ANFIS), and gene expression programming (GEP) have been successfully used to develop predictive models to solve nonlinear and complex problems in various topics of civil engineering, particularly in geotechnical engineering. In these regards, Heshmati et al. used linear genetic programming (LGP) for a formulation of soil classification [11]. Narendra et al. successfully used computational intelligence techniques to estimate the unconfined compressive strength of soft grounds [12]. Furthermore, Heshmati et al. used artificial neural networks (ANNs) to predict the unconfined compressive strength of soil-stabilizer mixes [13]. Soleimani et al. also developed new prediction models for the unconfined compressive strength of geopolymer stabilized soil by employing multigene genetic programming [14]. In addition, several soft computing-based approaches such as neural network [15], decision tree (DT) [16], Bayesian networks (BNs) [17], patient rule induction method (PRIM) [18], and gene expression programming (GEP) have been used to estimate the seismic liquefaction potential. Mozumder et al. predicted the penetrability of microfine cement (MC) grout in granular soil using artificial intelligence techniques [19]. Emamgolizadeh et al. also predicted the soil cation exchange capacity using GEP and multivariate adaptive regression splines (MARS) [20]. Generally, the robustness of soft computing approaches for developing new predictive models in different fields of geotechnical engineering based on the mentioned studies has been confirmed.

The main objective of this study is to develop a predictive model for the void ratio parameter of the mineral tailings using the GEP method. In order to develop the GEP models, a comprehensive database from previous studies including 113 laboratory data was collected. To find a robustness model, eight different GEP models are developed to estimate the void ratio parameter. The performance of the developed GEP models was evaluated using accuracy criteria. The relative importance of influential parameters dealing with the void ratio parameter was also investigated by the sensitivity analysis. In addition, the robustness of predicted GEP models was evaluated through parametric analysis. The GEP models were compared with each other, and finally the most appropriate model was selected.

2. Methods

2.1. Gene Expression Programming (GEP). The gene expression programming method was firstly introduced by Ferreira in 1999 [21]. This method is an evolutionary

algorithm which is closely related to genetic algorithms (GAs) and genetic programming (GP). The significant difference between these methods is in the statement and nature of individuals. In the GEP method, the individuals are stated as linear, symbolic, and fixed length string composed of one or more genes. However, despite their fixed length, the chromosomes are capable of nonlinear state entities with various shapes and sizes, which are known as expression trees, while in GA and GP methods, the individual elements are fixed length linear entities and nonlinear entities with various shapes and sizes, respectively [21]. Figure 1 shows a simple flowchart of the GEP algorithm.

Figure 1 illustrates the main steps of the GEP algorithm. At first, initial chromosomes are randomly generated to create an initial population. After that, these chromosomes are expressed, and fitness function is calculated for each one. Then, the chromosomes are chosen and kept based on their fitness function to develop a new population for the next iteration. This process is continued until the stopping criteria are achieved.

In the GEP model, the chromosomes are expressed as a tree structure (expression tree). In the tree structure, terminals represent leaves and functions represent nodes. The terminal set consists of the independent variables that are considered as input variables of the model. So, the first step to use the GEP method is to define the terminal set. In the present study, the terminal set contains $\{c\,(\% < 2\eta m), D_{50}, e_0, w_w, \sigma'\}$ which are the clay percentage, the mean grain diameter, the initial void ratio, the water content, and the effective stress, respectively. Also, the function set of $\{+, -, \times, /, \sqrt{}, \text{Exp, and Ln}\}$ is considered as nodes of the tree structure, and each chromosome can be expressed according to the terminal set and function set. An evolutionary process is used in the GEP method for finding the best program and individual. The chromosomes are modified and optimized in each iteration based on the fitness function and genetic operators like the genetic algorithm. This process is repeated until the convergence criteria are achieved. In this study, to evaluate the cost of chromosomes, the root relative square error (RRSE) is considered as the fitness function based on the following equation [22]:

$$\text{RRSE} = \sqrt{\frac{\sum_{i=1}^{n}(P_i - T_i)^2}{\sum_{i=1}^{n}(T_i - \overline{T})^2}}, \tag{1}$$

where P_i is the predicted value, T_i is the target value, $\overline{T}$ is the mean of target values, and n is the number of data.

2.2. Dataset. In this study, a suitable laboratory dataset gathered from different references in the literature is used. Different types of tailings including gold, zinc, and so on from various locations in the world are employed in the collected database. More details about the used database are presented in Table 1.

Various physical properties such as the initial void ratio (e_0), effective stress (σ'), water content (w_w), clay content ($c\,(\% < 2\eta m)$), and mean grain diameter (D_{50}) can affect the
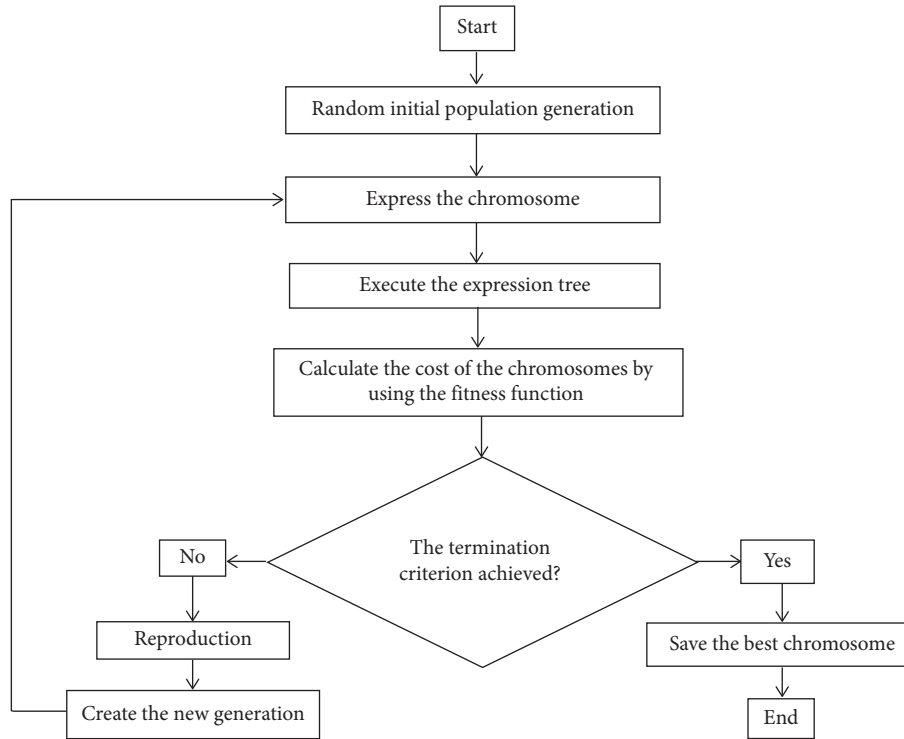
Figure 1: Simple flowchart of the GEP method [22].

Table 1: The sources of the datasets.

Reference	Type of tailing	No. of data
Highter and Vallee [2]	Garnet-zinc	7
Poulos et al. [3]	Aluminum	4
Stone et al. [4]	Gold	8
Aubertin et al. [23]	Sulfide free (not mentioned)	7
Qiu and Sego [10]	Copper-gold-coal-CT	18
Berilgen et al. [24]	Gold	8
Wong et al. [6]	Oil sands	6
Riemer et al. [5]	Not mentioned	5
Jeeravipoolvarn et al. [25]	Oil sands	6
Wickland et al. [26]	Gold	5
Quille and O′Kelly [8]	Zinc	7
Al-Tarhouni et al. [27]	Gold	12
Antonaki et al. [7]	Not mentioned	14
Bonin et al. [9]	Gold	6

void ratio parameter. In the present study, these effective parameters are used as predictor variables. The ranges of input and output variables are presented in Table 2. According to this table, the values of the void ratio parameter appear to be between 0.44 and 4.4. This range includes various waste types such as zinc and gold. It should be mentioned that different methods can be applied to deposit waste in the mentioned range. As a result, the used database contains a comprehensive and practical range related to the e parameter. Furthermore, according to Table 2, the values of the effective stress parameter vary between 0.1 and 7500 kPa which are very common ranges for real issues in geotechnical engineering problems. The initial void ratio changes between 0.68 and 4.4. In the tailing gradation section, the variation of the mean grain diameter is from 0.008 mm to 0.182 mm, which indicates that the waste classification varies from fine to coarse grains. The range of w_w parameter varies between 7% and 141.4%, and the range of clay content parameter is between 1.3% and 35%.

3. Developed GEP Models

To develop a robust formulation for the void ratio parameter based on the GEP model, the database is randomly divided into the training and testing datasets. The GEP model is trained by 70% of the whole database, and the remaining data points are used to test and evaluate the developed GEP model. In fact, the testing dataset is considered to verify the generalization capability of the developed model and also to avoid overfitting problems.

TABLE 2: Statistical properties of the datasets.

Parameter	e	e_0	σ (kPa)	D_{50}	$c\,(\% < 2\eta m)$	w_w
Range	0.44–4.4	0.68–4.4	0.1–7500	0.008–0.182	1.3–35	7–141.4
Minimum	0.44	0.68	0.1	0.008	1.3	7
Maximum	4.4	4.4	7500	0.182	3.5	141.4
Average	0.972	1.540	175.224	0.066	13.513	51.731
Std. deviation	0.505	0.969	843.698	0.058	8.810	31.867

To investigate the effect of each input parameter and also to examine different combinations of input variables, a group of eight GEP models was developed to estimate the void ratio parameter. In Model 1, all input parameters were included in the developing process. In Models 2 to 6, input parameters were excluded one at a time, from the developing process to observe the effect of each parameter on the void ratio parameter. Model 7 contains effective stress (σ'), mean grain diameter (D_{50}), and initial void ratio (e_0); however, Model 8 only consists of effective stress (σ') and initial void ratio (e_0). The eight models are presented below:

Model 1 = $f(\sigma, e_0, D_{50}, w_w, c\,(\% < 2\eta m))$

Model 2 = $f(\sigma, e_0, w_w, c\,(\% < 2\eta m))$

Model 3 = $f(\sigma, e_0, D_{50}, w_w)$

Model 4 = $f(\sigma, D_{50}, w_w, c\,(\% < 2\eta m))$

Model 5 = $f(e_0, D_{50}, w_w, c\,(\% < 2\eta m))$

Model 6 = $f(\sigma, e_0, D_{50}, c\,(\% < 2\eta m))$

Model 7 = $f(\sigma, e_0, D_{50})$

Model 8 = $f(\sigma, e_0)$

To investigate the accuracy of the developed GEP models, three different statistical error criteria, including the coefficient of determination (R^2), the mean absolute error (MAE), and the root mean square error (RMSE) are considered:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(O_i - O_m)^2}, \tag{2}$$

$$\text{MAE} = \frac{\sum_{i=1}^{N}|P_i - O_i|}{N}, \tag{3}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2}, \tag{4}$$

where O_i is the measured value, P_i stands for the prediction values, N is the number of data points, O_m is the mean value for observation, and P_m is the mean value of prediction.

4. Results

4.1. Performance Analysis. The void ratio parameter is one of the effective and practical parameters for specifying the behavior of various types of soils like clay, silt, mineral tailing, and so on. However, there is a need for a reliable predictive model for the estimation of this parameter. In this study, the GEP method is employed to develop new

formulations for estimating the void ratio parameter of tailings. To obtain the best predictive model, different combinations of input parameters are considered. In this aspect, eight models are developed. In Model 1, all input parameters are considered in the model development process. The mean diameter, clay percentage, initial void ratio, effective stress, and water content were not considered in Models 2 to 6, respectively. In Model 7, the clay percentage and water content were not considered. In the last model, the effective stress and initial void ratio were only considered in the developing process. It should be noted that the other combinations of input parameters were also considered but are not presented here due to their weaker performance in comparison with the presented ones.

The statistical error indices related to each developed model are presented in Table 3 for both training and testing datasets. For more illustration, the predicted values of the void ratio parameter by developed models versus the observed ones are shown in Figure 2.

As mentioned, Model 1 contains all input parameters ($\sigma, e_0, D_{50}, w_w, c\,(\% < 2\eta m)$). According to Figure 2(a), the dispersion of Model 1 is less than twenty percent. Also, according to Table 3, the accuracy criteria are $R^2 = 0.92$, MAE = .086, and RMSE = 0.122, and they show high accuracy. Therefore, Model 1 is suitable for the prediction of output parameter e.

The parameters $\sigma, e_0, w_w,$ and $c\,(\% < 2\eta m)$ were considered in Model 2 for estimating the void ratio. As seen in Figure 2(b), the scatter of Model 2 exceeds the twenty percent error in some cases and also the dispersion of Model 2 is more than that in Model 1. In addition, according to the results of Table 3, the accuracy of Model 2 is less than that of Model 1. Therefore, it can be concluded that Model 2 is not suitable for the prediction of the void ratio parameter. In Model 3, all input parameters, except the clay percentage, were considered as predictor variables. Figure 2(c) shows that the scatter diagram of Model 3 is similar to that of Model 1. Also, it can be inferred from Table 3 that the performance of Model 3 and Model 1 is nearly the same. Thus, Model 3 has remarkable accuracy prediction of the output parameter e.

Four input parameters, namely, $\sigma', D_{50}, w_w,$ and $c\,(\% < 2\eta m)$, were used for the prediction of the void ratio in Model 4. According to Figure 2(d), the dispersion of Model 4 is more than that of Model 1. Table 3 also shows that the accuracy of Model 4 is significantly low in comparison with Model 1. As a result, Model 4 cannot be used to accurately predict parameter e. In Model 5, parameters $e_0, D_{50}, w_w,$ and $c\,(\% < 2\eta m)$ were considered, and parameter σ was excluded. As shown in Figure 3(e), the maximum

TABLE 3: Performance of the GEP models in predicting void ratio for training and testing datasets.

Models	Train			Test		
	MAE	RMSE	R^2	MAE	RMSE	R^2
Model 1	0.086	0.122	0.93	0.121	0.24	0.92
Model 2	0.097	0.133	0.93	0.196	0.543	0.90
Model 3	0.09	0.117	0.95	0.127	0.189	0.92
Model 4	0.143	0.183	0.72	0.19	0.388	0.83
Model 5	0.18	0.287	0.31	0.286	0.534	0.55
Model 6	0.091	0.118	0.92	0.136	0.258	0.92
Model 7	0.076	0.104	0.96	0.111	0.17	0.94
Model 8	0.093	0.127	0.93	0.125	0.234	0.91

(a)

(b)

(c)

(d)

FIGURE 2: Continued.

(e)

(f)

(g)

(h)
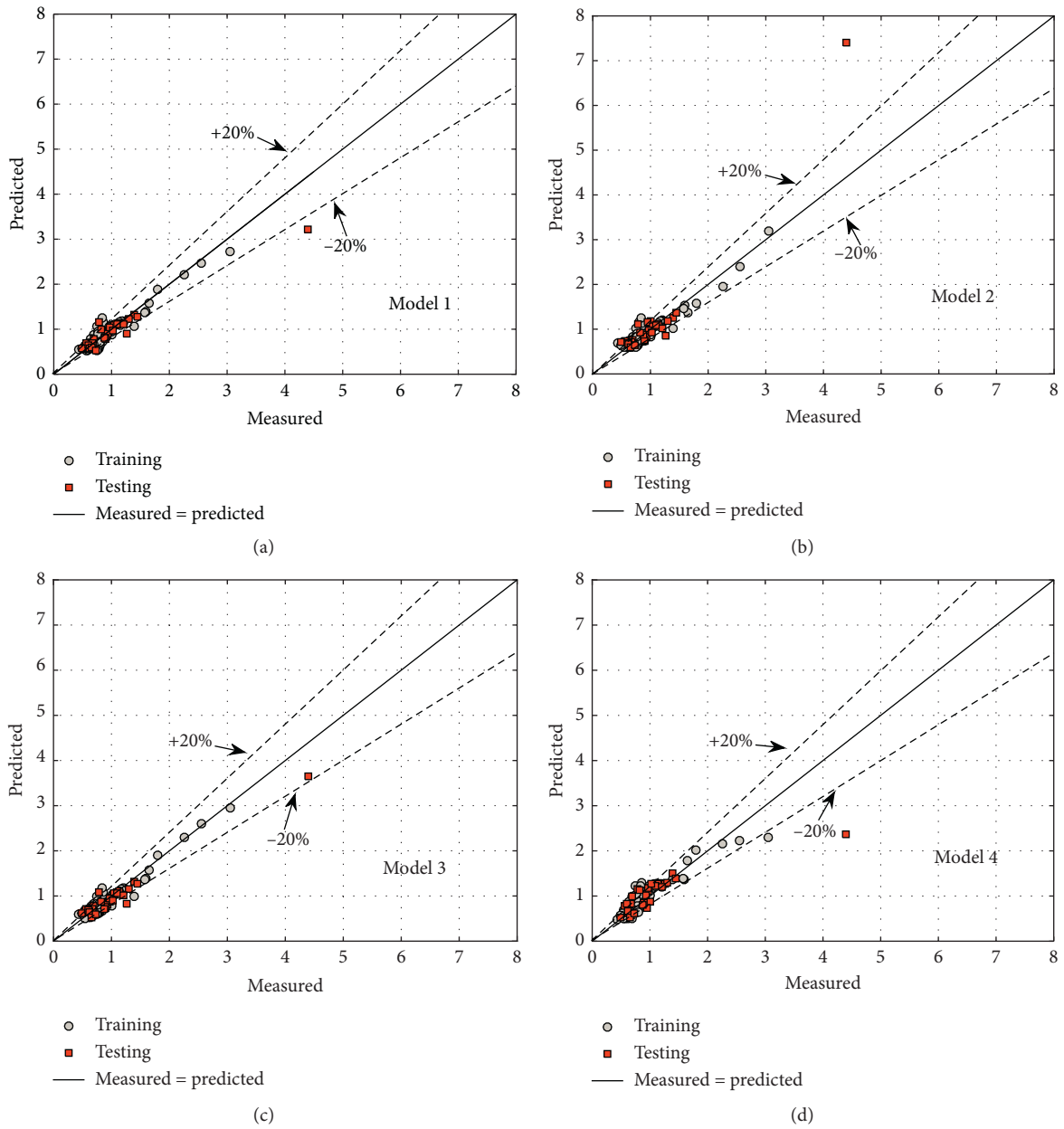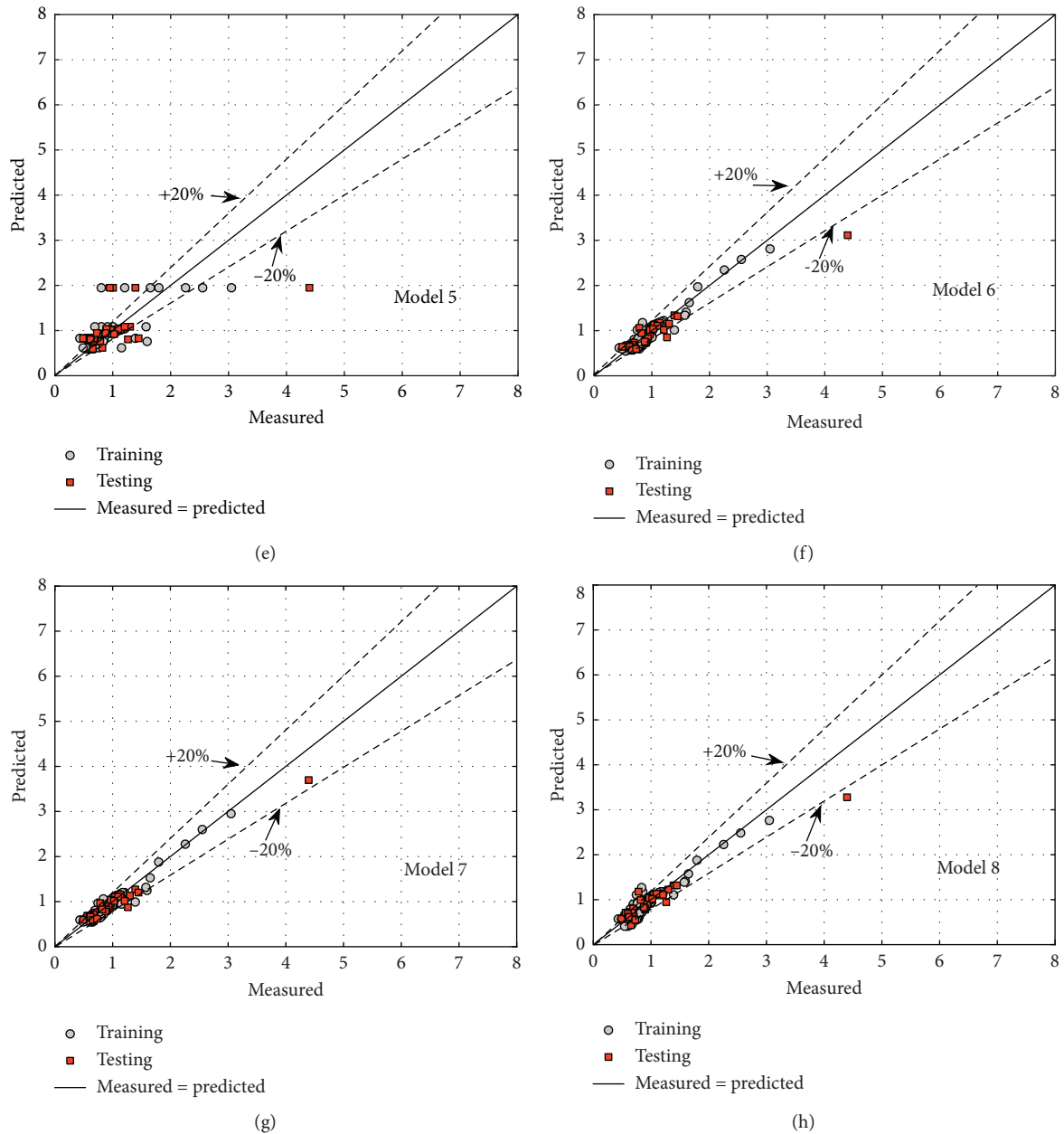
FIGURE 2: Measured versus predicted void ratio of tailings using GEP models. (a) Model 1. (b) Model 2. (c) Model 3. (d) Model 4. (e) Model 5. (f) Model 6. (g) Model 7. (h) Model 8.

dispersion is related to Model 5. Also, Table 3 shows high error levels and a significant drop in the accuracy of Model 5. Therefore, Model 5 can be removed from the list of appropriate predictive models. Parameters $\sigma', e_0, D_{50}, and\, c\,(\% < 2\eta m)$ were applied in Model 6 to estimate the void ratio. Figure 2(f) and Table 3 indicate that the accuracy of Model 6 is similar to that of Model 1. Therefore, Model 6 is appropriate for predicting the output parameter.

Parameters $\sigma', e_0, and\, D_{50}$ were used in Model 7 to develop the appropriate model for predicting the void ratio parameter. According to Figure 2(g), the dispersion of Model 7 is less than that of Model 1. Also, Table 3 indicated that Model 7 outperforms Model 1. As a result, Model 7 is an

effective model for predicting the void ratio. Model 8 presented good results in estimating the void ratio using effective stress and initial void ratio parameters (σ', e_0). The scatter diagram of Model 8 shows an appropriate estimation of the void ratio parameter. According to Table 3, the accuracy of Model 8 is relatively lower than that of Model 7; however, in comparison with Model 1, the conditions are quite similar. Thus, Model 8 is efficient for the prediction of the void ratio.

In general, the results of the performance analysis indicate that Models 1, 3, 6, 7, and 8 can be effectively applied to estimate parameter e. However, it should be noted that the robustness of the developed models must also be physically
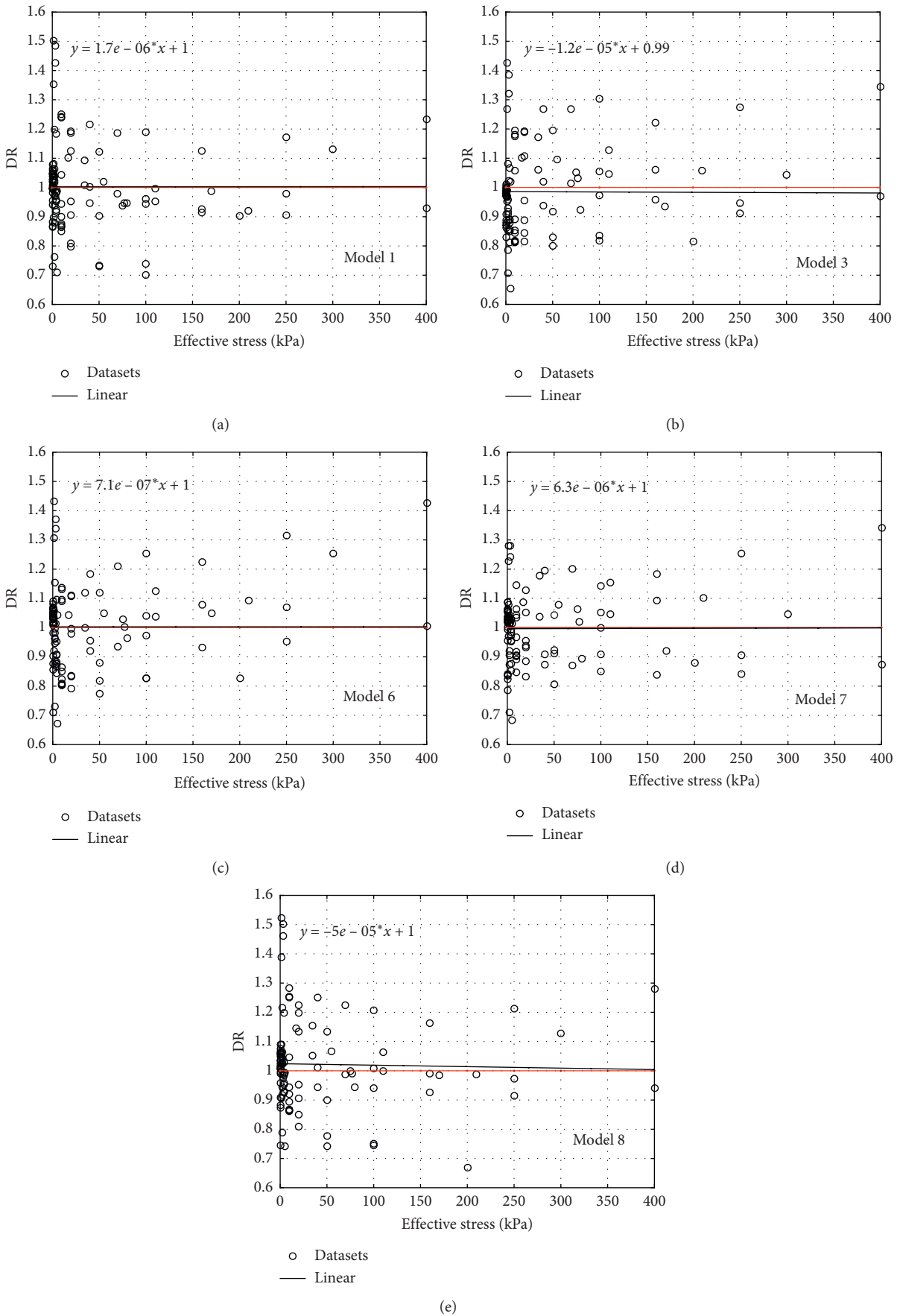
FIGURE 3: Variation of DR against effective stress for (a) Model 1, (b) Model 3, (c) Model 6, (d) Model 7, and (e) Model 8.

TABLE 4: External validation statistical measures for different approaches.

Item	Performance criteria	Condition	Model 1	Model 3	Model 6	Model 7	Model 8
1	R	$R > 0.8$	0.967542	0.97819	0.962684	0.983063	0.965097
2	K	$0.85 < k < 1.15$	1.03594	1.063196	1.044038	1.036733	1.031193
3	k'	$0.85 < k' < 1.15$	0.985769	0.962664	0.98137	0.981004	0.995735
4	R_m	$R_m > 0.5$	0.713235	0.803656	0.695878	0.805972	0.697069
5	M	$|m| < 0.1$	−0.06056	−0.02679	−0.06697	−0.02852	−0.06796
6	N	$|n| < 0.1$	−0.06746	−0.03997	−0.07771	−0.03344	−0.07357

investigated; this is to ensure that their results are also regarded when selecting the best model among Models 1, 3, 6, 7, and 8.

To further confirm the accuracy of the developed GEP models, a new validation criterion introduced by Tropsha et al. was employed [28]. In the mentioned method, several factors including gradients of the regression lines (k and k'), coefficient of the determination for the regression line (m, n) through the origin, correlation coefficient (R), and the condition of cross validation (R_m) are defined. The formulas related to these criteria are presented as follows:

$$k = \sum_{i=1}^{n} \frac{(O_i \times P_i)}{P_i^2}, \tag{5}$$

$$k' = \sum_{i=1}^{n} \frac{(O_i \times P_i)}{O_i^2}, \tag{6}$$

$$m = \frac{R^2 - R_0^2}{R^2}, \tag{7}$$

$$n = \frac{R^2 - R_0'^2}{R^2}, \tag{8}$$

$$R = \frac{\sum_{i=1}^{N} (P_i - P_m)(O_i - O_m)}{\sqrt{(P_i - P_m)^2} \sqrt{(O_i - O_m)^2}}, \tag{9}$$

$$R_m = R^2 \times \left(1 - \sqrt{|R^2 - R_0^2|}\right) > 0.5, \tag{10}$$

where R_0^2 is the squared correlation coefficient between the predicted and measured values and the $R_0'^2$ is the squared correlation coefficient between measured and predicted values. Their formulas are defined as follows:

$$R_0^2 = 1 - \frac{\sum_{i=1}^{n} P_i^2 (1-k)^2}{\sum_{i=1}^{n} (P_i - \overline{P})^2}, \tag{11}$$

$$R_0'^2 = 1 - \frac{\sum_{i=1}^{n} O_i^2 (1-k')^2}{\sum_{i=1}^{n} (O_i - \overline{O})^2}. \tag{12}$$

The values of the R and R_m should be more than 0.8 and 0.5, respectively. The values of the k and k' should be between 0.85 and 1.15, respectively. The values of m and n should also be less than 0.1. The testing dataset was used to obtain the value of the new validation criteria for the developed GEP models (Models 1, 3, 6, 7, and 8). The results of the statistical analysis

are presented in Table 4. According to Table 4, all the GEP models satisfy the condition of the new validation criteria.

In order to complete the performance analysis, the discrepancy ratio (DR) between predicted and measured values is depicted as a function of input parameters in Figures 3 and 4.

$$DR = \log \frac{P_i}{O_i}. \tag{13}$$

As shown in these figures, the DR values of the developed GEP models (Models 1, 3, 6, 7, and 8) are approximately independent of the value of the input parameters. It can be interpreted from these observations that the input parameters are correctly incorporated in the developed GEP models. It should be noted that the same performance is also observed in D_{50}, $c\,(\% < 2\eta m)$, and w_w.

4.2. Sensitivity Analysis.
In the present study, a sensitivity analysis is implemented to determine the most effective parameters in the estimation of the void ratio parameter. To achieve this, seven scenarios are considered. In the first scenario, all input variables are included in the modeling process. In the remaining scenarios, the input parameters are singly excluded from the modeling procedure. According to Table 3, Model 1 includes all of the input variables while the mean diameter, clay percentage, initial void ratio, and water content were not considered in Models 2 to 6, respectively. It can be seen from this table that the errors of the developed models are sensitive to the elimination of each input parameter. Based on Table 3, removing the effective stress parameter from Model 5 remarkably increases the errors of the developed model. This indicates that this parameter is very important in the estimation of the void ratio parameter. The initial void ratio also shows a significant contribution in generating a predictive model for the void ratio parameter. The effectiveness of other parameters in comparison with the initial void ratio and effective stress parameters is negligible. These results are in line with the physical concepts of the problem and also with previous studies in the literature.

4.3. Parametric Analysis.
To investigate the robustness of the developed models, parametric analysis is conducted to ensure that the results of the GEP models are in line with the nature of the problem. To achieve this purpose, parameter e which has been predicted by the developed GEP models is plotted as a function of each input parameter. The results of the parametric analysis are shown in Figure 5. As shown in Figure 5(a), the void ratio predicted by all the GEP models
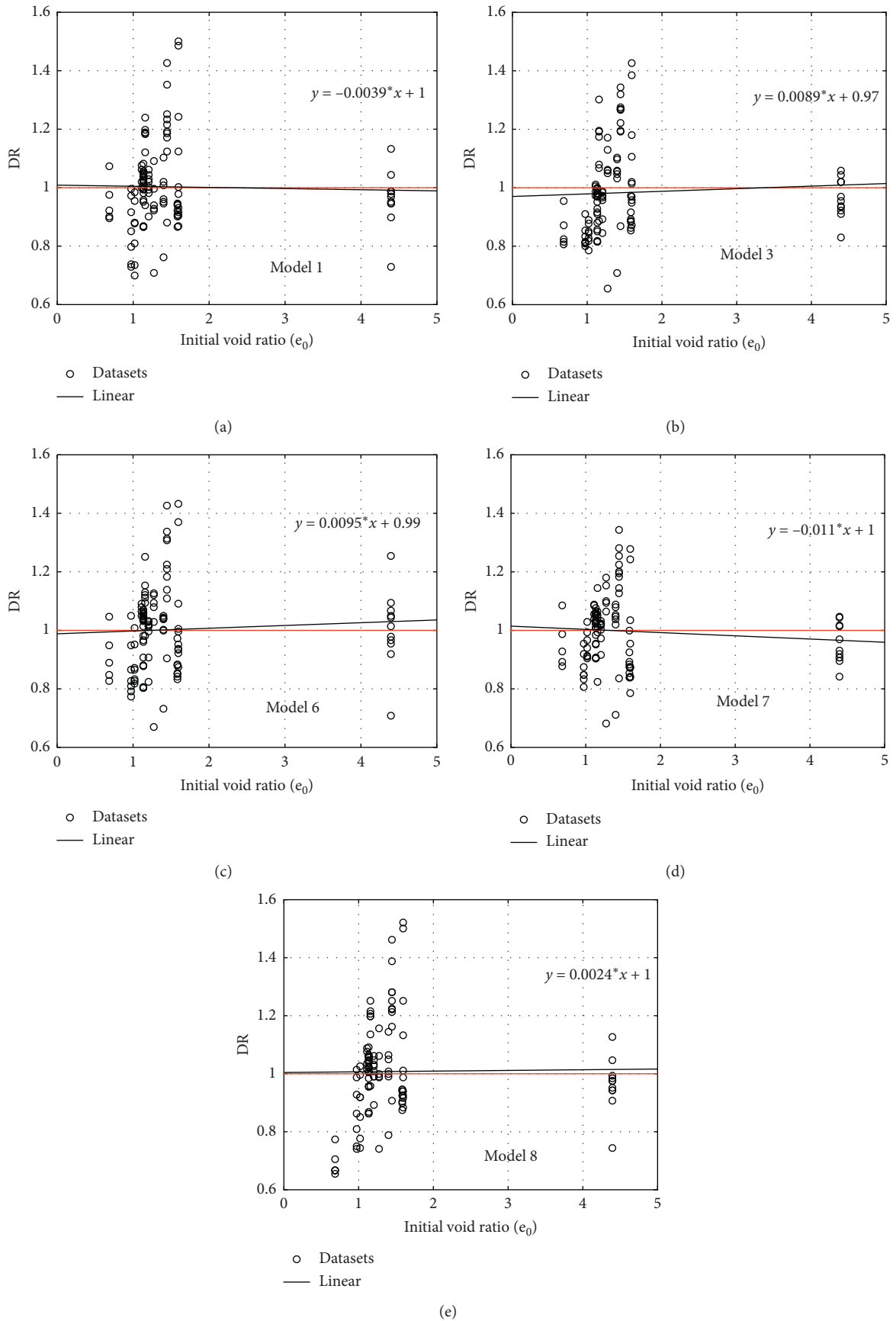
$y = -0.0039 {}^{*}x + 1$

Model 1

$y = 0.0089 {}^{*}x + 0.97$

Model 3

$y = 0.0095 {}^{*}x + 0.99$

Model 6

$y = -0.011 {}^{*}x + 1$

Model 7

$y = 0.0024 {}^{*}x + 1$

Model 8

FIGURE 4: Variation of DR against initial void ratio for (a) Model 1, (b) Model 3, (c) Model 6, (d) Model 7, and (e) Model 8.
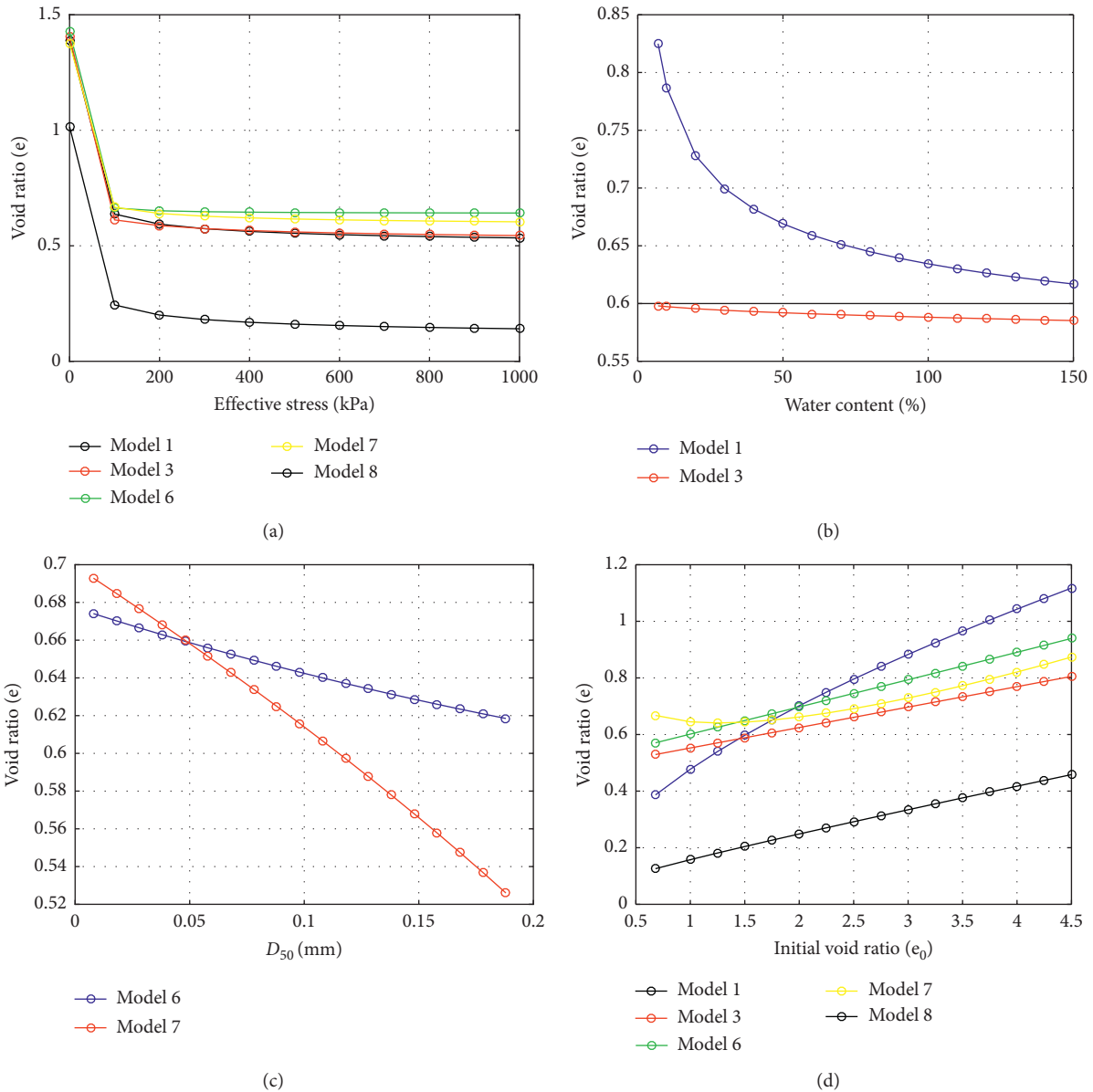
(a)

(b)

(c)

(d)

Figure 5: Variation of the predicted void ratio (e) versus (a) effective stress (kPa), (b) water content (%), (c) D_{50} (mm), and (d) initial void ratio (e_0) for GEP Models.

nonlinearly decreases with the increase of effective stress. This observation is in line with the existing knowledge of soil mechanics. In fact, previous studies have also confirmed that the void ratio parameter is inversely proportional to the effective stress (σ') in consolidation of the soils [2–8, 10].

Variation of the e parameter predicted by Models 1, 3, 6, 7, and 8 versus the e_0 parameter is shown in Figure 5(d). According to Figure 5(d), the e parameter increases with the increase of the e_0 parameter in Models 1, 3, 6, and 8, but this relation is inverse in Model 7. By reviewing the literature, it is inferred that the increase in the initial void ratio leads to the increase of the void ratio parameter [9, 24]. Therefore, Models 1, 3, 6, and 8 are in agreement with the nature of the problem, but Model 7 cannot be verified in the aspect of the parametric analysis. Variations of the e parameter predicted

by Models 1 and 3 versus the w_w parameter are shown in Figure 5(b). According to this figure, the value of the void ratio parameter decreases with the increase of the value of the water content parameter. In this regard, Qiu and Sego investigated the effect of water content on the consolidation properties of mineral deposits of copper, gold, and coal [10]. However, their results showed that the void ratio increases with the increase of the water content. Thus, both Models 1 and 3 may suffer from a lack of physical justification. It should be noted that the w_w parameter was not considered in Models 6 to 8. In Figure 5(c), the e parameter estimated by Models 6 and 7 is also shown as a function of the D_{50} parameter. According to this figure, the e parameter decreases with the increase of the D_{50} parameter. However, in soil mechanics, it has been proved that the value of the void
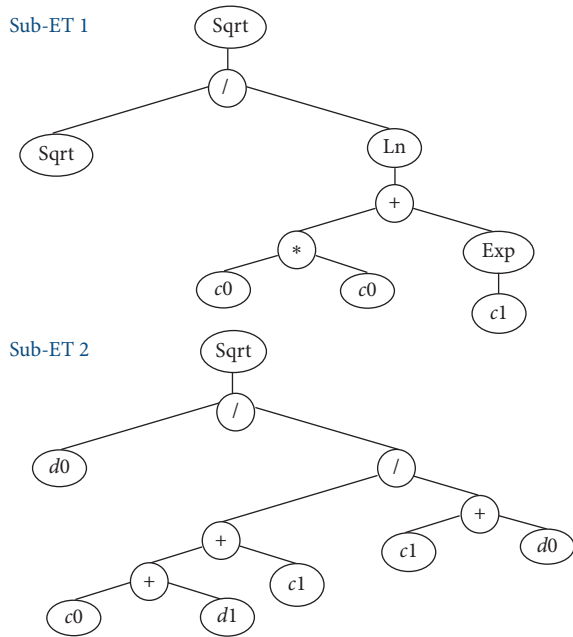
FIGURE 6: Expression tree of optimal GEP for the void ratio (Model 8).

ratio increases with the increase of the value of D_{50}. Also, Quille and O'Kelly showed that the void ratio of the coarse-grained waste is more than that of fine-grained waste [8]. As a result, Models 6 and 7 cannot be approved based on the results of the parametric analysis.

In general, based on the results of performance and parametric analyses, Model 8 is the best model among the other developed GEP models for predicting the void ratio parameter. The formulation of Model 8 is presented in equation (14), and the expression tree of Model 8 is given in Figure 6.

$$e = \frac{\sqrt[4]{e_0}}{2.23} + \sqrt[2]{\frac{e_0 (e_0 - 0.0082)}{(2.712 + \sigma')}}. \tag{14}$$

It is important to note that GEP has high precision in the void ratio (e) estimation. Given the fact that there is no empirical formula for predicting the void ratio, the mathematical expression of the GEP model developed in this study has the advantage of being simple in form and more accurate in predictions of the void ratio. Finally, the predicted value of the e parameter in the design of the tailing dam can be used.

5. Conclusion

Geotechnical behavior of mineral tailings is one of the essential requirements in geotechnical engineering. The consolidation and strength behavior of mineral tailings are important in how to design the types of storage of this type of soil. In both types of tailing behavior, the void ratio parameter is important. Therefore, with the proper estimation of the void ratio, it is easier to understand the performance of mineral tailings. In this study, GEP was used to develop a

robust model for the estimation of the void ratio in tailing dams. A comprehensive laboratory dataset including 113 data vectors gathered from different sources in the literature was used. According to the literature, the initial void ratio (e_0), effective stress (σ'), water content (w_w), clay content ($c (\% < 2\eta m)$), and mean grain diameter (D_{50}) are the most important factors which affect the void ratio parameter. In order to evaluate the effect of each of the input parameters on the output parameter and also to investigate the combination of different input parameters, eight different GEP models were developed to predict the void ratio parameter. The performances of the developed GEP models were evaluated based on the accuracy criteria. Results indicated that five GEP models estimated the void ratio fairly well. Sensitivity analysis was also performed to determine the most effective input parameters in the estimation of the void ratio parameter. The results indicated that e_0 and σ' were the most effective parameters in the estimation of the void ratio parameter. The models without these two parameters showed the least accuracy in comparison with the other models. Finally, the robustness of the developed GEP models was investigated based on a parametric analysis. Results showed that only Model 8, in which e_0 and σ' parameters were considered as input parameters, agreed with the nature of the problem. Model 8 with RMSE = 0.180 and $R^2 = 0.92$ had remarkable accuracy and precision in the prediction of the void ratio parameter.

Abbreviations

σ': The effective stress
e_0: The initial void ratio
w_w: The water content
$c (\% < 2\eta m)$: The clay content
D_{50}: The mean grain diameter
e: The void ratio.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] I. Jantzer, A. Bjelkevik, and K. Pousette, *Material Properties of Tailings from Swedish Mines*, Lulea: Norsk Geoteknisk Forening, ICOLD and UNEP, Lulea, Sweden, 2001.

[2] W. H. Highter and R. P. Vallee, "The liquefaction of different mine tailings under stress-controlled loading," *Engineering Geology*, vol. 16, no. 1-2, pp. 147–150, 1980.

[3] S. J. Poulos, E. I. Robinsky, and T. O. Keller, "Liquefaction resistance of thickened tailings," *Journal of Geotechnical Engineering*, vol. 111, no. 12, 1985.

[4] K. J. L. Stone, M. F. Randolph, S. Toh, and A. A. Sales, "Evaluation of consolidation behavior of mine tailings,"

Journal of Geotechnical Engineering, vol. 120, no. 3, pp. 473–490, 1994.

[5] M. Riemer, Y. Moriwaki, and J. Obermeyer, "Effect of high confining stresses on static and cyclic strengths of mine tailing materials," in *Proceedings of the Geotechnical Earthquake and Engineering and Soil Dynamics IV Congress*, Sacramento, CA, USA, May 2008.

[6] R. C. Wong, B. N. Mills, and Y. B. Liu, "Mechanistic model for one-dimensional consolidation behavior of nonsegregating oil sands tailings," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 134, no. 2, pp. 195–202, 2008.

[7] N. Antonaki, I. Sasanakul, T. Abdoun, M. V. Sanin, H. Puebla, and J. Ubilla, "Centrifuge modeling of deposition and consolidation of fine-grained mine tailings," in *Proceedings of the Geo-Congress 2014 Technical Papers*, Atlanta, Georgia, February 2014.

[8] M. E. Quille and B. C. O'Kelly, "Geotechnical properties of zinc/lead mine tailings from tara mines, Irelands," in *Proceedings of the Geo Shanghai 2010 International Conference*, Shanghai, China, June 2010.

[9] M. D. Bonin, M. Nuth, A. M. Dagenais, and A. R. Cabral, "Experimental study and numerical reproduction of self-weight consolidation behavior of thickened tailings," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 140, no. 12, 2014.

[10] Y. Qiu and D. C. Sego, "Laboratory properties of mine tailings," *Canadian Geotechnical Journal*, vol. 38, no. 1, pp. 183–190, 2001.

[11] A. A. R. Heshmati, H. Salehzadeh, A. H. Alavi, A. Badkobeh, A. H. Gandomi, and A. T. Ghasemi, "On the applicability of linear genetic programming for the formulation of soil classification," *American-Eurasian Journal of Agriculture and Environment Sciences*, vol. 45, pp. 575–583, 2008.

[12] B. S. Narendra, P. V. Sivapullaiah, S. Suresh, and S. N. Omkar, "Prediction of unconfined compressive strength of soft grounds using computational intelligence techniques: a comparative study," *Computers and Geotechnics*, vol. 33, no. 3, pp. 196–208, 2006.

[13] A. A. R. Heshmati, A. H. Alavi, M. Keramati, and A. H. Gandomi, "A radial basis functions neural network approach for compressive strength prediction of stabilized soil," in *Road Pavement Material Characterization and Rehabilitation: Selected Papers from the 2009 Geo Hunan International Conference*, Changsha, China, August 2009.

[14] S. Soleimani, S. Rajaei, P. Jiaoc, A. Sabz, and S. Soheilinia, "New prediction models for unconfined compressive strength of geopolymer stabilized soil using multi-gen genetic programming," *Measurement*, vol. 113, pp. 99–107, 2017.

[15] A. M. Hanna, D. Ural, and G. Saygili, "Neural network model for liquefaction potential in soil deposits using Turkey and Taiwan earthquake data," *Soil Dynamics and Earthquake Engineering*, vol. 27, no. 6, pp. 521–540, 2007.

[16] A. H. Gandomi, M. M. Fridline, and D. Roke, "Decision tree approach for soil liquefaction assessment," *The Scientific World Journal*, vol. 2013, Article ID 346285, 2013.

[17] J. L. Hu, X. W. Tang, and J. N. Qiu, "A bayesian network approach for predicting seismic liquefaction based on interpretive structural modeling," *Natural Hazards and Earth System Sciences*, vol. 18, pp. 1451–1468, 2015.

[18] A. Kaveh, S. M. Hamze-Ziabari, and T. Bakhshpoori, "Patient rrule-induction method for liquefaction potential assessment based on CPT dada," *Bulletin of Engineering Geology and the Environment*, vol. 77, no. 2, pp. 849–865, 2018.

[19] R. M. Mozumder, A. I. Laskar, and M. Hussain, "Penetrability prediction of micro fine cement grout in granular soil using artificial intelligence techniques," *Tunneling and Underground Space Technology*, vol. 72, pp. 131–144, 2018.

[20] S. Emamgolizadeh, S. M. Bateni, D. Shahsavani, T. Ashrafi, and H. Ghorbani, "Estimation of soil cation exchange capacity using genetic expression programming (GEP) and multivariate adaptive regression splines (MARS)," *Journal of Hydrology*, vol. 529, no. Part 3, pp. 1590–1600, 2015.

[21] C. Ferreira, "Gene expression programming: a new adaptive algorithm for solving problems," *Complex Systems*, vol. 13, no. 2, 2001.

[22] I. Ebtehaj, H. Bonakdari, A. H. Zaji, H. Azimi, and A. Sharifi, "Gene expression programming to predict the discharge coefficient in rectangular side weirs," *Applied Soft Computing*, vol. 35, pp. 618–628, 2015.

[23] M. Aubertin, B. Bussiere, and R. P. Chapuis, "Hydraulic conductivity of homogenized tailings from hard rock mines," *Canadian Geotechnical Journal*, vol. 33, no. 3, pp. 470–482, 1996.

[24] S. A. Ber[idot]lgen, P. B[idot]çer, M. M. Ber[idot]lgen, and K. İ. Özayd[idot]n, "Assessment of consolidation behavior of golden horn marine dredged material," *Marine Georesources & Geotechnology*, vol. 24, no. 1, pp. 1–16, 2006.

[25] S. Jeeravipoolvarn, J. D. Scott, and R. J. Chalaturnyk, "Geotechnical characteristics of laboratory in-line thickened oil sands tailings," in *Proceedings of the Tailings and Mine Waste 2009*, Banff, Canada, November 2009.

[26] B. E. Wickland, G. W. Wilson, and D. Wijewickreme, "Hydraulic conductivity and consolidation response of mixtures of mine waste rock and tailings," *Canadian Geotechnical Journal*, vol. 47, no. 4, pp. 472–485, 2010.

[27] M. Al-Tarhouni, P. Simms, and S. Sivathayalan, "Cyclic behaviour of reconstituted and desiccated-rewet thickened gold tailings in simple shear," *Canadian Geotechnical Journal*, vol. 48, no. 7, pp. 1044–1060, 2011.

[28] A. Tropsha, P. Gramatica, and V. K. Gombar, "The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models," *QSAR & Combinatorial Science*, vol. 22, no. 1, pp. 69–77, 2003.