Hindawi
Applied Computational Intelligence and Soft Computing
Volume 2021, Article ID 6686347, 17 pages
https://doi.org/10.1155/2021/6686347

Research Article

Application of Gene Expression Programming to Evaluate Strength Characteristics of Hydrated-Lime-Activated Rice Husk Ash-Treated Expansive Soil

Kennedy C. Onyelowe,[1] **Fazal E. Jalal**,[2] **Michael E. Onyia**,[3] **Ifeanyichukwu C. Onuoha**,[4] **and George U. Alaneme**[5]

[1]*Department of Civil and Mechanical Engineering, Kampala International University, Kampala, Uganda*
[2]*Department of Civil Engineering, Shanghai Jiao Tong University, Shanghai, China*
[3]*Department of Civil Engineering, Faculty of Engineering, University of Nigeria, Nsukka, Nigeria*
[4]*Department of Environmental Technology, Federal University of Technology, Owerri, Nigeria*
[5]*Department of Civil Engineering, Michael Okpara University of Agriculture, Umudike, Nigeria*

Correspondence should be addressed to Kennedy C. Onyelowe; kennedychibuzor@kiu.ac.ug

Received 7 December 2020; Revised 24 March 2021; Accepted 26 March 2021; Published 14 April 2021

Academic Editor: Babak Daneshvar Rouyendegh (B. Erdebilli)

Gene expression programming has been applied in this work to predict the California bearing ratio (CBR), unconfined compressive strength (UCS), and resistance value (R value or R_{value}) of expansive soil treated with an improved composites of rice husk ash. Pavement foundations suffer failures due to poor design and construction, poor materials handling and utilization, and management lapses. The evolution of sustainable green materials and optimization and soft computing techniques have been deployed to improve on the deficiencies being suffered in the abovementioned areas of design and construction engineering. In this work, expansive soil classified as A-7-6 group soil was treated with hydrated-lime activated rice husk ash (HARHA) in an incremental proportion to produce 121 datasets, which were used to predict the behavior of the soil's strength parameters utilizing the mutative and evolutionary algorithms of GEP. The input parameters were HARHA, liquid limit (w_L), (plastic limit (w_P), plasticity index (I_P), optimum moisture content (w_{OMC}), clay activity (A_C), and (maximum dry density ($\delta_{\max}$) while CBR, UCS, and R value were the output parameters. A multiple linear regression (MLR) was also conducted on the datasets in addition to GEP to serve as a check mechanism. At the end of the computing and iterations, MLR and GEP optimization methods proposed three equations corresponding to the output parameters of the work. The responses validation on the predicted models shows a good correlation above 0.9 and a great performance index. The predicted models' performance has shown that GEP soft computing has predicted models that can be used in the design of CBR, UCS, and R value for soils being used as foundation materials and being treated with admixtures as a binding component.

1. Introduction

The design, construction, and monitoring of earthwork infrastructure have been of utmost importance due to the everyday failure civil engineering facilities experience [1–4]. For this reason, composite materials with special properties have been evolved to replace ordinary cement [5–8]. One such technique in the utilization of special binders is the introduction of activators to ash materials to form activated ash with the ability to resist unfavorable conditions and factors that have proven to be averse to constructed infrastructure [9–14]. However, the evolution of soft computing in engineering has added to the efficiency of designing, constructing, and monitoring of the performance of earthworks [15–19]. One such soft computing or machine learning method is gene expression programming (GEP). Invented by Cramer [20], genetic programming (GP) and gene expression programming (GEP) are the branches of genetic algorithm (GA) that is regarded as an evolutionary computing algorithm technique [20–22]. It is based on Darwin's theory of "survival of the fittest" that does not require making prior assumptions about the solution structure [23]. The working procedure of GP comprises various steps [24]: (1) create an initial population in accordance with the function and terminal settings; (2) use two key criteria, fitness function and maximum number of generations, to assess the performance of the generated population; if the performance of this population is according to the requirement or approaches the maximum number of the generation, terminate the program, otherwise, continuously generate a new population using three genetic operations of reproduction, crossover, and mutation for an amount of duration until the threshold criteria are not met. The experimental database was separated into training, validation, and testing set for the GEP analysis. In order to confirm consistent data division, many combinations of the training and testing sets were taken [25].

In Figure 1, it can be seen that input data is fed to either GP or a mathematical model that incorporates GP that yields predicted and observed values. The difference between these is residual errors which are reduced by continuing formulating in the GEP tool until an optimum model is obtained.

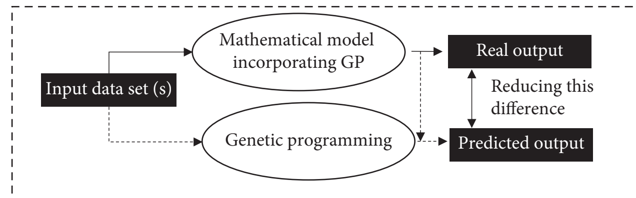

Figure 1: Simple working schematics for gene expression programming.

2. Materials and Methods

2.1. Preparation of Materials. Expansive clay soil was prepared and tests were conducted on both the untreated and the treated soils to determine the datasets presented in Table 1, needed for the evolutionary predictive modeling. The hydrated-lime activated rice husk ash (HARHA) is a hybrid geomaterial binder developed by blending rice husk ash with 5% by weight activator agent, which in this case is hydrated lime ($Ca(OH)_2$) and allowed for 48 hours. At the same time, the rice husk is an agroindustrial waste derived from the processing of rice in rice mills and homes. Through controlled direct combustion proposed by Onyelowe et al. [4], the rice husk mass is turned into ash from rice husk ash (RHA). The HARHA was used in incremental proportions to treat the clayey soil and the response behavior on different properties tested, observed, and recorded (see Table 1).

2.2. Model Method. In Figure 2, the flowchart of the gene expression programming method and execution is presented. The 121 input and output datasets were deployed to the GeneXpro software computing platform to generate the predicted outputs and the models from that operation. Several trials or iterations were carried out to achieve the best fit.

3. Results and Discussion

3.1. Pearson Correlation. Pearson's correlation matrix [26] was generated from the given data comprising seven input and three output parameters using the data analysis capabilities of Microsoft Excel. The correlation matrix is defined as a square, symmetrical $P \times P$ matrix with the $(ij)^{\text{th}}$ element equal to the correlation coefficient *R_ij* among the $(i)^{\text{th}}$ and the $(j)^{\text{th}}$ variable. The diagonal members (correlations of variables with each other) are always equal to one [27]. Thus, the left-hand nine columns of this correlation matrix represent qualitatively the correlations between the input soil hydraulic-prone properties (HARHA, w_L, w_P, I_P, w_{OMC}, A_C, $\delta_{\max}$) and output soil strength properties, i.e., CBR, UCS_{28}, and R_{Value} (Table 2). The range of correlation factors varies from −1 and 1 (0 represents no correlation, whereas ±1 shows greater correlation). A positive value suggests that the respective increase or decrease is linear among the two variables simultaneously. It is indicated in Table 2 that the CBR, UCS_{28}, and R_{Value} have a correlation coefficient above 0.90 for all input parameters with the exception of w_{OMC} for the last two outputs (0.134 and 0.363), respectively. Thus, a high correlation exists in this correlation matrix for the considered input and out parameters. In Figure 3 was presented the frequency histograms of the input variables: (a) HARHA; (b) w_L; (c) w_P; (d) I_P; (e) w_{OMC}; (f) A_C; (g) $\delta_{\max}$; and output variables (h) CBR; (i) UCS_{28}; (j) R_{Value}.

3.2. Gene Expression Programming. The performance of a developed GEP model using a database is affected by the sample size and its variable distributions, which agrees with the findings of Gandomi and Roke [25]. Thus, the frequency histograms for all the input parameters (HARHA, w_L, w_P, I_P, w_{OMC}, A_C, and $\delta_{\max}$) and output values (CBR, UCS_{28}, and R_{Value}) are visualized in Figure 3. It can be seen that the bell-shaped curve indicates even distribution of the data. This diagram is often used for the initial assessment of geochronological data, which involves relatively large sets of data, according to Sircombe [28]. All the data is seen to exhibit even sample distributions and follow a symmetrical pattern such that the display of the histograms straightforward.

The descriptive statistics of the input and out parameters are tabulated in Table 3. This statistical summary shows the minimum and maximum ranges for all input and output parameters. The standard deviation (SD), Kurtosis, and skewness are also given for each parameter, which agrees with Edjabou et al. [29]. A low SD means that most of the values are close to the average (w_P, w_{OMC}, A_C, $\delta_{\max}$, and R_{Value}), whereas a larger SD means that the numbers are more spread out (w_L, I_P, CBR, and UCS_{28}). Skewness quantifies the asymmetry of the probability distribution of a real-valued random variable with respect to its mean. It can be positive, zero, negative, or undefined [30]. The negative values generally suggest that the tail is extended on the left side of the distribution curve (w_L, w_P, I_P, w_{OMC}, A_C, $\delta_{\max}$,

Table 1: 121 datasets of input and output parameters.

HARHA (%)	Input soil hydraulic-prone properties						Output soil strength properties		
	w_L (%)	w_P (%)	I_P (%)	w_{OMC}(%)	A_C	δ_{max} (g/cm^3)	CBR (%)	UCS_{28}(kN/m^2)	R_{Value}
0	66	21	45	16	2.0	1.25	8	125	11.7
0.1	66	21	45	16	1.98	1.25	8.1	125	11.7
0.2	65.7	20.9	44.8	16.1	1.96	1.27	8.2	126	11.7
0.3	65.6	20.9	44.7	16.3	1.96	1.27	8.2	126	11.7
0.4	65.3	20.8	44.5	16.3	1.93	1.28	8.3	126	11.8
0.5	65	21	44	16.4	1.9	1.30	8.5	128	12.0
0.6	64.8	20.8	44	16.4	1.88	1.31	8.55	128	12.2
0.7	64.5	20.8	43.7	16.45	1.88	1.31	8.6	128	12.2
0.8	64.1	20.8	43.3	16.47	1.87	1.33	8.6	130	12.3
0.9	63.5	20.9	42.6	16.49	1.85	1.33	8.85	130	12.6
1	63	21	42	16.5	1.8	1.35	9.2	132	13.1
1.1	62.5	20.6	41.9	16.6	1.8	1.35	9.25	132	13.3
1.2	62.1	20.3	41.8	16.7	1.81	1.36	9.4	133	13.5
1.3	61.9	20.2	41.7	16.8	1.8	1.37	9.5	133	13.6
1.4	61.7	20.1	41.6	17	1.81	1.38	9.7	134	13.8
1.5	61.5	20	41.5	17.2	1.8	1.38	9.8	134	14.2
1.6	61.4	20	41.4	17.2	1.8	1.39	9.8	136	14.4
1.7	61.3	20	41.3	17.3	1.79	1.39	9.85	137	14.8
1.8	61.3	20.1	41.2	17.5	1.81	1.4	9.92	137	14.8
1.9	61.2	20.1	41.1	17.7	1.8	1.41	9.96	138	15
2	61	20	41	17.8	1.8	1.41	10.4	138	15.3
2.1	60.9	19.9	41	17.9	1.8	1.42	10.4	139	15.6
2.2	60.7	19.7	41	17.9	1.8	1.42	10.7	139	15.7
2.3	60.6	19.6	41	18	1.8	1.425	11	140	15.8
2.4	60.4	19.4	41	18.2	1.8	1.43	11.6	141	16
2.5	60	19	41	18.3	1.8	1.43	12.0	141	16.2
2.6	59.8	19	40.8	18.35	1.79	1.435	12.1	142	16.5
2.7	59.7	19.1	40.6	18.4	1.77	1.45	12.4	142	16.8
2.8	59.5	19.1	40.4	18.45	1.75	1.455	12.9	142	17
2.9	59.2	19	40.2	18.5	1.72	1.46	13.3	143	17.1
3	59	19	40	18.5	1.7	1.46	13.8	143	17.3
3.1	58.8	19.2	39.6	18.55	1.7	1.47	13.9	144	17.4
3.2	58.4	18.9	39.5	18.6	1.7	1.475	14.2	145	17.7
3.3	57.9	19.1	38.8	18.7	1.71	1.48	14.5	146	18
3.4	57.4	19	38.4	18.75	1.69	1.484	14.7	147	18.3
3.5	57	19	38	18.8	1.7	1.49	14.8	148	18.5
3.6	56.8	18.9	37.9	18.85	1.69	1.5	15	148	18.7
3.7	56.7	19	37.7	18.9	1.65	1.51	15.3	150	18.9
3.8	56.5	18.9	37.6	18.93	1.64	1.51	15.7	151	19.1
3.9	56.3	19	37.3	18.98	1.61	1.52	15.9	152	19.2
4	56	19	37	19.0	1.6	1.52	16.0	153	19.4
4.1	55.7	19	36.7	19.0	1.59	1.53	16.3	154	19.5
4.2	54.9	18.7	36.2	19.0	1.57	1.54	16.8	156	19.6
4.3	54.1	18.5	35.6	19.0	1.55	1.55	17.5	157	19.7
4.4	53.6	18.4	35.2	19.0	1.52	1.56	17.8	158	19.7
4.5	53	18	35	19.0	1.5	1.57	18.0	159	19.8
4.6	52.8	18	34.8	18.98	1.5	1.58	18.1	160	20
4.7	52.7	18	34.7	18.96	1.5	1.59	18.3	160	20
4.8	52.6	18.1	34.5	18.93	1.5	1.60	18.8	162	20.1
4.9	52.3	18	34.3	18.91	1.5	1.61	19.5	163	20.2
5	52	18	34	18.9	1.5	1.61	19.8	164	20.4
5.1	51.5	17.7	33.8	18.88	1.48	1.62	19.9	165	20.4
5.2	51.1	17.7	33.4	18.86	1.46	1.63	20	166	20.5
5.3	50.8	18.1	32.7	18.84	1.43	1.64	20.3	167	20.5
5.4	50.3	18	32.3	18.82	1.41	1.65	20.9	168	20.6
5.5	50	18	32	18.8	1.4	1.65	21.7	168	20.6
5.6	49.9	18	31.9	18.78	1.4	1.66	21.9	169	20.7
5.7	49.6	17.9	31.7	18.75	1.41	1.67	22.1	170	20.8
5.8	49.4	17.9	31.5	18.71	1.42	1.67	22.3	171	20.8

Table 1: Continued.

HARHA (%)	Input soil hydraulic-prone properties						Output soil strength properties		
	w_L (%)	w_P (%)	I_P (%)	w_{OMC}(%)	A_C	δ_{max} (g/cm^3)	CBR (%)	UCS$_{28}$(kN/m^2)	R_{Value}
5.9	49.1	17.7	31.4	18.65	1.41	1.68	22.5	172	20.9
6	49	18	31	18.6	1.4	1.69	22.8	172	20.9
6.1	48.6	17.8	30.8	18.55	1.38	1.7	23.1	173	21
6.2	48.3	17.6	30.7	18.48	1.37	1.71	23.3	173	21.1
6.3	47.7	17.3	30.4	18.6	1.35	1.72	23.7	174	21.2
6.4	47.2	17	30.2	18.44	1.33	1.73	23.8	175	21.4
6.5	47	17	30	18.4	1.3	1.74	24.0	175	21.5
6.6	46.8	17.1	29.7	18.4	1.31	1.75	24.3	176	21.6
6.7	46.5	16.8	29.7	18.41	1.31	1.76	24.9	177	21.8
6.8	45.6	15.9	29.7	18.4	1.3	1.77	25.2	177	21.9
6.9	45.2	15.9	29.3	18.41	1.3	1.78	25.5	178	22.0
7	45	16	29	18.4	1.3	1.78	25.9	179	22.0
7.1	44.8	16.3	28.5	18.39	1.29	1.79	26.2	180	22.1
7.2	44.3	16.1	28.2	18.37	1.27	1.8	26.6	181	22.3
7.3	43.7	15.9	27.8	18.35	1.26	1.81	27	182	22.4
7.4	43.4	16	27.4	18.32	1.23	1.83	27.3	183	22.5
7.5	43	16	27	18.3	1.2	1.84	27.6	183	22.6
7.6	42.8	15.9	26.9	18.29	1.19	1.85	27.7	184	22.7
7.7	42.4	16	26.4	18.28	1.18	1.86	28.3	184	22.8
7.8	41.8	15.4	26.4	18.26	1.16	1.87	28.5	183	22.8
7.9	41.5	15.4	26.1	18.23	1.14	1.87	28.7	184	22.9
8	41	15	26	18.2	1.13	1.88	29.0	185	22.9
8.1	40.7	14.9	25.8	18.2	1.12	1.88	29.3	186	23
8.2	40.3	15	25.3	18.2	1.11	1.89	29.9	187	23.2
8.3	39.8	15.1	24.7	18.2	1.11	1.90	30.4	188	23.3
8.4	39.3	15	24.3	18.21	1.1	1.90	30.7	189	23.5
8.5	39	15	24	18.2	1.0	1.91	31.2	190	23.6
8.6	38.8	15	23.8	18.2	1.0	1.92	31.5	191	23.7
8.7	38.3	14.9	23.4	18.2	1.0	1.93	32.1	192	23.8
8.8	37.9	15.2	22.7	18.2	1.0	1.94	32.4	193	23.9
8.9	37.5	15.2	22.3	18.2	1.0	1.95	33.5	194	24
9	37	15	22	18.2	1.0	1.96	34.0	195	24.0
9.1	37	15	22	18.19	1.0	1.962	34.5	196	24.1
9.2	37	15	22	18.18	1.0	1.964	34.8	197	24.2
9.3	37	15	22	18.16	1.0	1.966	35.2	198	24.3
9.4	37	15	22	18.13	1.0	1.969	35.8	199	24.4
9.5	37	15	22	18.1	1.0	1.97	36.0	200	24.5
9.6	36.8	15.1	21.7	18	0.99	1.972	36.5	202	24.6
9.7	36.7	15.1	21.6	17.92	0.98	1.973	36.9	204	24.7
9.8	36.5	15.1	21.4	17.93	0.97	1.975	37.6	208	24.8
9.9	36.3	15.2	21.1	17.91	0.94	1.977	37.8	208	24.8
10	36	15	21	17.9	0.9	1.98	38.0	210	24.9
10.1	35.7	14.9	20.8	17.88	0.88	1.98	38.3	213	25.1
10.2	35.5	15.1	20.4	17.84	0.86	1.982	38.5	214	25.3
10.3	34.6	14.9	19.7	17.79	0.84	1.984	38.9	215	25.4
10.4	33.3	14	19.3	17.73	0.82	1.987	39.6	218	25.4
10.5	33	14	19	17.7	0.8	1.99	40.0	220	25.5
10.6	32.8	14	18.8	17.7	0.79	1.99	41.1	222	25.8
10.7	32.4	13.9	18.5	17.71	0.78	1.99	42.4	223	26.2
10.8	31.5	13.9	17.6	17.71	0.75	1.99	43.2	225	26.3
10.9	31.1	14	17.1	17.7	0.72	1.99	43.5	228	26.5
11	31	14	17	17.7	0.7	1.99	44.0	230	26.8
11.1	30.7	13.9	16.8	17.68	0.7	1.99	44.0	231	26.8
11.2	30.3	13.7	16.6	17.63	0.71	1.99	44.5	232	26.8
11.3	29.8	13.4	16.4	17.57	0.71	1.99	44.6	232	26.9
11.4	29.4	13.2	16.2	17.53	0.71	1.98	44.6	232	26.9
11.5	29	13	16	17.5	0.7	1.97	43.8	225	26.9
11.6	28.7	12.8	15.9	17.5	0.69	1.97	43.8	224	26.9
11.7	28.5	13	15.5	17.4	0.67	1.96	43.7	223	27.0

Table 1: Continued.

HARHA (%)	Input soil hydraulic-prone properties						Output soil strength properties		
	w_L (%)	w_P (%)	I_P (%)	w_{OMC}(%)	A_C	δ_{max} (g/cm^3)	CBR (%)	UCS_{28}(kN/m^2)	R_{Value}
11.8	27.8	13	14.8	17.3	0.65	1.96	43.6	222	27.0
11.9	27.6	13.2	14.4	17.2	0.62	1.95	43.5	221	27.0
12	27	13	14	17.1	0.6	1.95	43.4	221	27.0

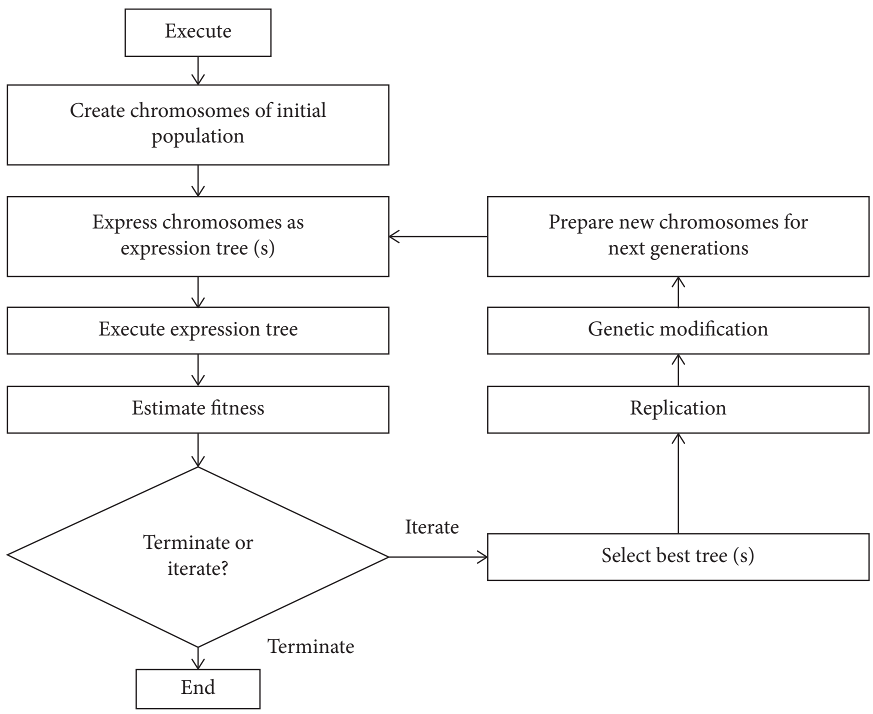

Figure 2: Flowchart of the GEP model execution.

and R_{Value}), while positively skewed shows that the tail is on the right side (CBR and UCS_{28}), which is reflected from the frequency histograms given in Figure 3 and the variable importance presented in Figures 4–6. Like skewness, kurtosis explains the shape of a probability distribution [31]. The Pearson measure of kurtosis of a given univariate normal distribution is generally taken as 3. Kurtosis values below 3 are called platykurtic, meaning that the distribution produces fewer and less extreme outliers than does the normal distribution, for instance, a uniform distribution, that is reflected in Figure 3.

To select the most appropriate GEP estimation model for HARHA treated expansive soils, several models with a varying number of genes were generated by employing a set of genetic operators (mutation, transposition, and crossover). Originally, a model composed of two genes with additional linking functions and head sizes of four (head size, $H = 4$) was selected and run a number of times. After that, the parameters were altered, in a stepwise order, by increasing the number of genes to three, head size to eight (head size, $H = 8$), number of chromosomes to 50, and weights of function sets. The program was run various times for different models, and the predicted final models were checked and compared with regard to their performance. Furthermore, the parameters such as mutation rate, inversion, and points of recombination were chosen on the basis of past studies [32–34] and then assessed to obtain their optimum impact. After running several trials, the final mathematical model was obtained, for which the selected parameters including detailed information of the general, numerical constants, and the genetic operators, are listed in Table 4. The final prediction model was chosen on the basis of criteria of the best fitness and lesser complexity of the mathematical formulation, while the expression trees (ETs) are illustrated in Figures 7–9 for the model outcomes CBR, $UCS_{28,}$ and R_{Value}, respectively.

In order to formulate the three models for the respective output parameters, initially, the input parameters were selected from the extensive experimental study, which is given below:

$$\mathrm{CBR}, \mathrm{UCS}_{28}, R_{\mathrm{Value}} = f\left(\mathrm{HARHA}, w_L, w_P, I_P, A_C, w_{\mathrm{OMC}}, \delta_{\max}\right), \quad (1)$$

Table 2: Pearson correlation matrix for input and output parameters.

	HARHA	w_L	w_P	I_P	w_{OMC}	A_C	δ_{max}	CBR	UCS_{28}	R_{Value}
HARHA	1									
w_L	−0.99724	1								
w_P	−0.98926	0.991515	1							
I_P	−0.99652	0.999411	0.986472	1						
w_{OMC}	0.201388	−0.1435	−0.17491	−0.1348	1					
A_C	−0.99388	0.997543	0.984584	0.998142	−0.12039	1				
δ_{max}	0.985771	−0.98176	−0.97696	−0.98026	0.23936	−0.97417	1			
CBR	0.991609	−0.99425	−0.98026	−0.99514	0.097679	−0.9951	0.969326	1		
UCS_{28}	0.990886	−0.99098	−0.97628	−0.99206	0.134931	−0.99283	0.967127	0.996459	1	
R_{Value}	0.984407	−0.9721	−0.96953	−0.97003	0.363941	−0.96588	0.972762	0.96009	0.967161	1

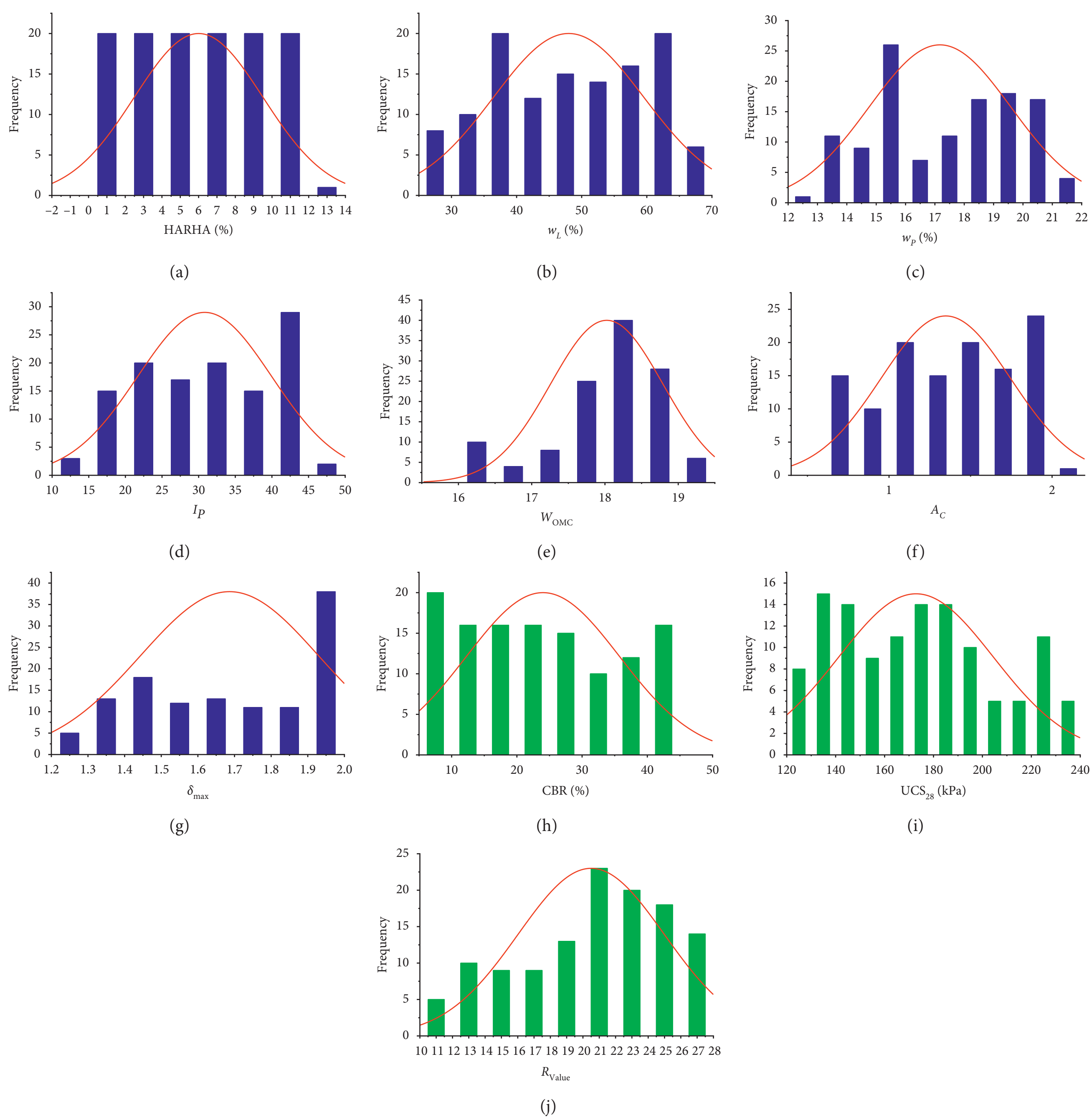

Figure 3: Frequency histograms of the input and output variables: (a) HARHA; (b)w_L; (c)w_P; (d)I_P; (e)w_{OMC}; (f) A_C; (g) δ_{max}; (h) CBR; (i) UCS_{28}; (j) R_{Value}. Note: the input parameters are illustrated in blue while the output variables are expressed as green histograms.

Table 3: Statistical parameters of the input and output parameters.

	Input soil hydraulic-prone properties						Output soil strength properties		
	w_L	w_P	I_P	w_{OMC}	A_C	δ_{max}	CBR	UCS_{28}	R_{Value}
Min.	27.00	12.80	14.00	16.00	0.60	1.25	8.00	125.00	11.70
Max.	66.00	21.00	45.00	19.00	2.00	1.99	44.60	232.00	27.00
Sum	5808	2078	3730	2181	163	204	2904	20917	2481
Mean	48.00	17.17	30.82	18.02	1.35	1.69	24.00	172.87	20.50
Median	49.00	17.70	31.00	18.20	1.40	1.69	22.80	172.00	20.90
SD	11.49	2.40	9.11	0.77	0.40	0.24	11.69	31.53	4.46
Kurtosis	−1.25	−1.24	−1.25	0.24	−1.18	−1.42	−1.17	−1.04	−0.79
Skewness	−0.13	−0.06	−0.14	−0.94	−0.21	−0.17	0.30	0.26	−0.44

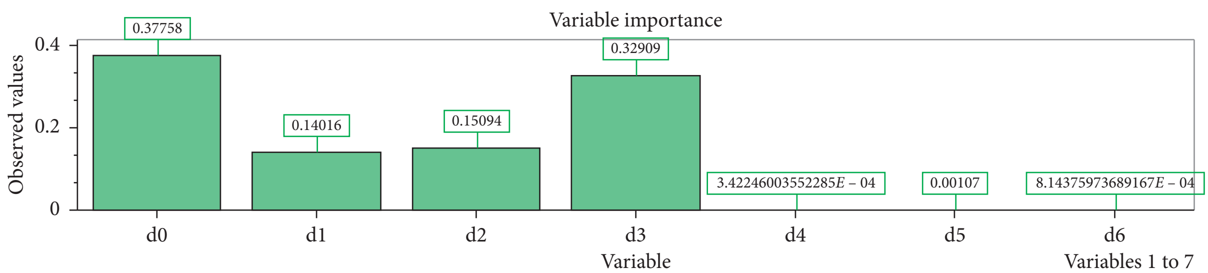

Figure 4: CBR variable importance of observed values and the input variable; (d0) HARHA; (d1) w_L; (d2) w_P; (d3) I_P; (d4) w_{OMC}; (d5) A_C; (d6) δ_{max}.

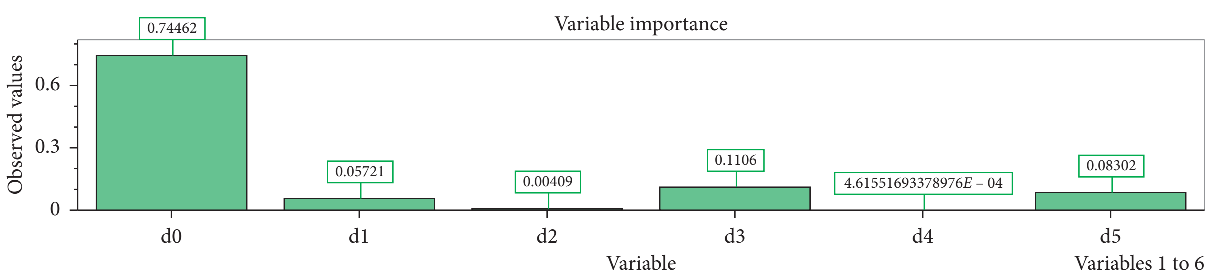

Figure 5: UCS variable importance of observed values and the input variable; (d0) HARHA; (d1) w_L; (d2) w_P; (d3) I_P; (d4) w_{OMC}; (d5) A_C; (d6) δ_{max}.

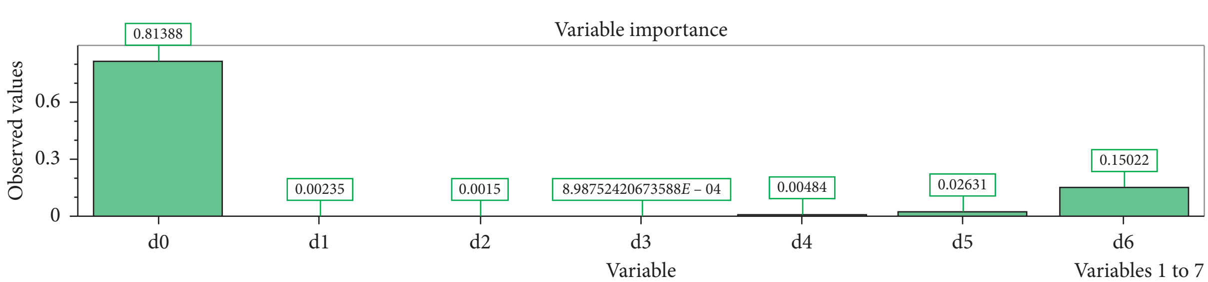

Figure 6: R value variable importance of observed values and the input variable; (d0) HARHA; (d1) w_L; (d2) w_P; (d3) I_P; (d4) w_{OMC}; (d5) A_C; (d6) δ_{max}.

where CBR is California bearing ratio, UCS_{28} is unconfined compression strength after 28 days, R_{Value} is resistance value, HARHA is hydrated lime activated rice husk ash, w_L is the liquid limit, w_P is the plastic limit, I_P is the plasticity index, w_{OMC} is the optimum moisture content, A_C is the activity value, and δ_{max} is the maximum dry density.

Table 4: Parameters setting for GEP algorithms.

Parameters	Settings CBR, UCS_{28}, R_{Value}
General	
Training records	81
Validation/Testing records	40
Number of chromosomes	50
Number of genes	3
Head size	8
Linking function	Addition
Function set	+, −, ×, ÷, exp, Sqrt, ln
Numerical constants	
Constants per gene	10
Type of data	Floating point
Maximum complexity	10
Ephemeral random constant	[−10, 10]
Genetic operators	
Mutation	0.00138
Inversion rate	0.00546
IS transposition rate	0.00546
RIS transposition rate	0.00546
One-point recombination rate	0.00277
Two-point recombination rate	0.00277
Gene recombination rate	0.00277
Uniform recombination	0.00755

The *K*-expressions and the genes nodal values for the ET of the modeled parameters of strength are presented as follows.

3.2.1. California Bearing Ratio.

Sqrt.Sqrt.+.−.−.−.+. ∗ .d5.d5.d4.c6.c1.d4.d1.d6.d3

+

∗ .Sqrt.−.d0.+.+.−.+.c1.d3.d0.d2.d1.c1.d2.d1.d6

+

∗ .+.d6./.Exp./.c1.Ln.d2.d1.d5.d1.d2.d2.d5.d5.d4

Numerical Constants:

Gene 1

c0 = 6.01733451338237
c1 = 5.82940372479169
c2 = 11.2892741508225
c3 = −1.38096255378887
c4 = −7.16238898892178
c5 = 6.36524552140873
c6 = 438.770447855123
c7 = −3.76850684316538
c8 = −3.92196417126987
c9 = 5.34226508377331

Gene 2

c0 = 5.61693166905728
c1 = −33451.121590902
c2 = 9.04538102359081
c3 = 4.02193288475646
c4 = 7.06854457228309
c5 = −5.52471996798914
c6 = 9.28254036072878
c7 = −9.37192907498398
c8 = 7.87691579943236
c9 = 7.84859767448958

Gene 3

c0 = 9.3145542771691
c1 = 0.683142490481803
c2 = 0.65507980590228
c3 = 2.23527237769707
c4 = 2.1560127041438
c5 = −3.4600786347084
c6 = −0.443433942686239
c7 = 6.32145146031068
c8 = −243.307901242103
c9 = 3.60334589462102

3.2.2. Unconfined Compressive Strength.

∗ .−.c9.+.d0.c1.Exp. ∗ .d5.c6.d0.d0.d0.d3.d4.d2.c4

+

∗ ./.c3.Sqrt.d5./.−.+.d4.d0.c8.d5.d6.d0.c8.c1.d0

+

−.+.+.+.−.d3.Sqrt.−.c4.d2.d4.d0.d1.c1.d5.d0.d1

Numerical Constants:

Gene 1

c0 = 9.40635120700705
c1 = −9.52207061952574
c2 = -6.06555375835444
c3 = 8.41547898800623
c4 = 6.96584978789636
c5 = 4.43152256843776
c6 = −4.66996057039345
c7 = −1.44721823786126
c8 = 2.64381847590564
c9 = −9.17752843515198

Gene 2

c0 = −6.69023712881863E-02
c1 = 1.7045835749382
c2 = 3.74612759288614
c3 = 5.99579447574825
c4 = −4.96296086938292
c5 = −3.58989226966155
c6 = −0.914639728995636
c7 = −6.71803949095126
c8 = 7.91580822137299E-02
c9 = −0.480693990905484

Gene 3

c0 = 8.17865535447249
c1 = 3.47497553241407
c2 = −6.28205053865169
c3 = −7.01719634907071
c4 = 5.34816290007036
c5 = 6.77358317819758
c6 = −4.4777053132725

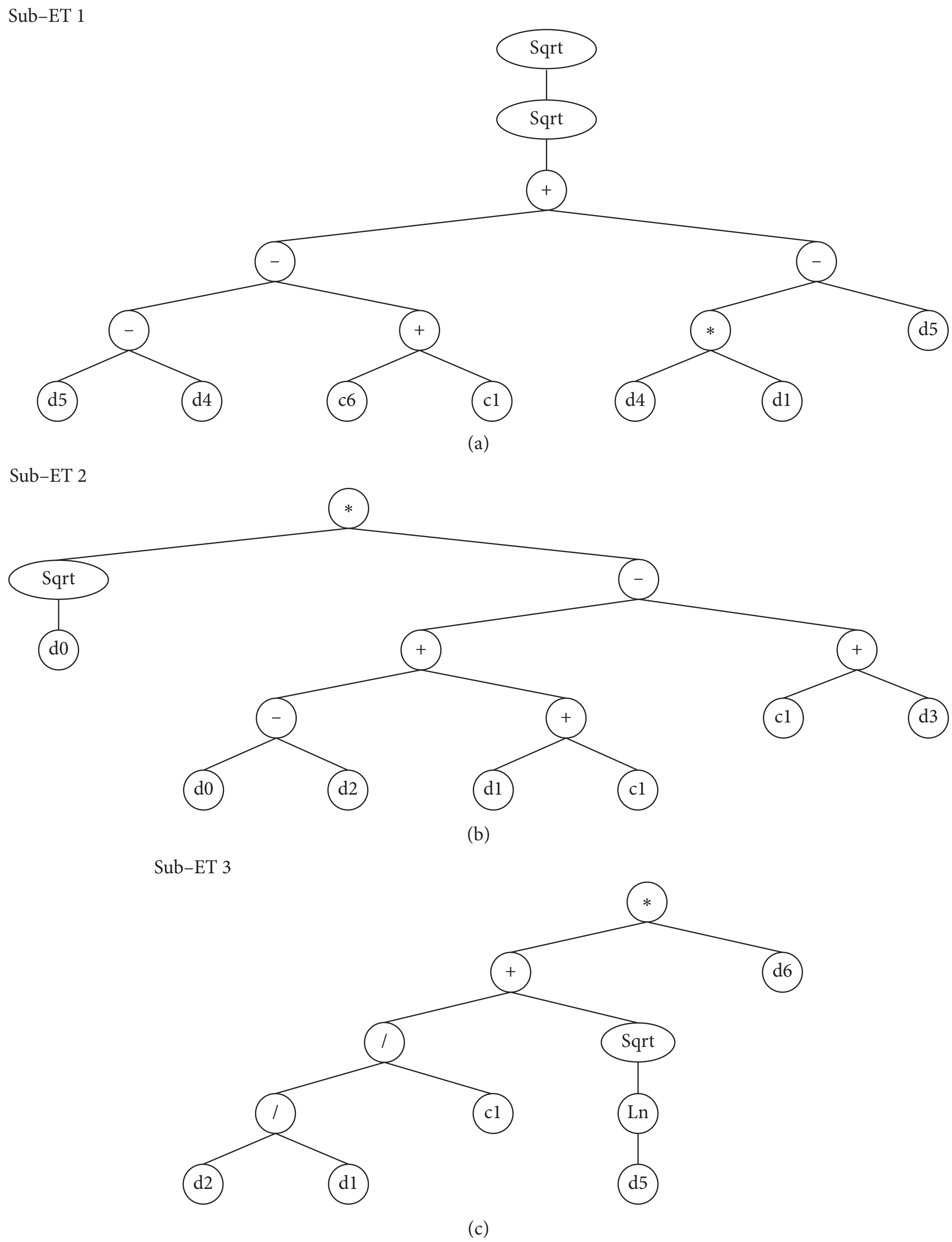

Figure 7: Expression tree of the model formulated for CBR (where d0: HARHA, d1: w_L, d2: w_P, d3: I_P, d4: w_{OMC}, d5: AC, d6: δmax, Sub-ET1 C6: 438.77, Sub-ET1 C1: 5.829, Sub-ET2 C1: -3341.12, Sub-ET3 C1: 0.6831).

c7 = −9.76500747703482
c8 = 8.85799737540819
c9 = 2.08953825495163

3.2.3. Resistance Value (R value)

Sqrt. ∗ ./.Sqrt.Exp.d6.+.d5.d1.d4.d5.c5.d4.c5.d1.c7.d1

+

/. ∗ .d6.d0.Ln.+.−.+.d4.d5.c7.d4.c4.c1.d2.d1.d3

+

+.Ln./.+./.−.+. ∗ .d4.d3.d0.d2.c2.d4.d4.c6.c7

Numerical Constants:

Gene 1

c0 = −5.76100955229347
c1 = 4.89717612231819
c2 = −3.93536179692984
c3 = 3.23796197393719
c4 = −6.77412671285134
c5 = 3.29407635731071

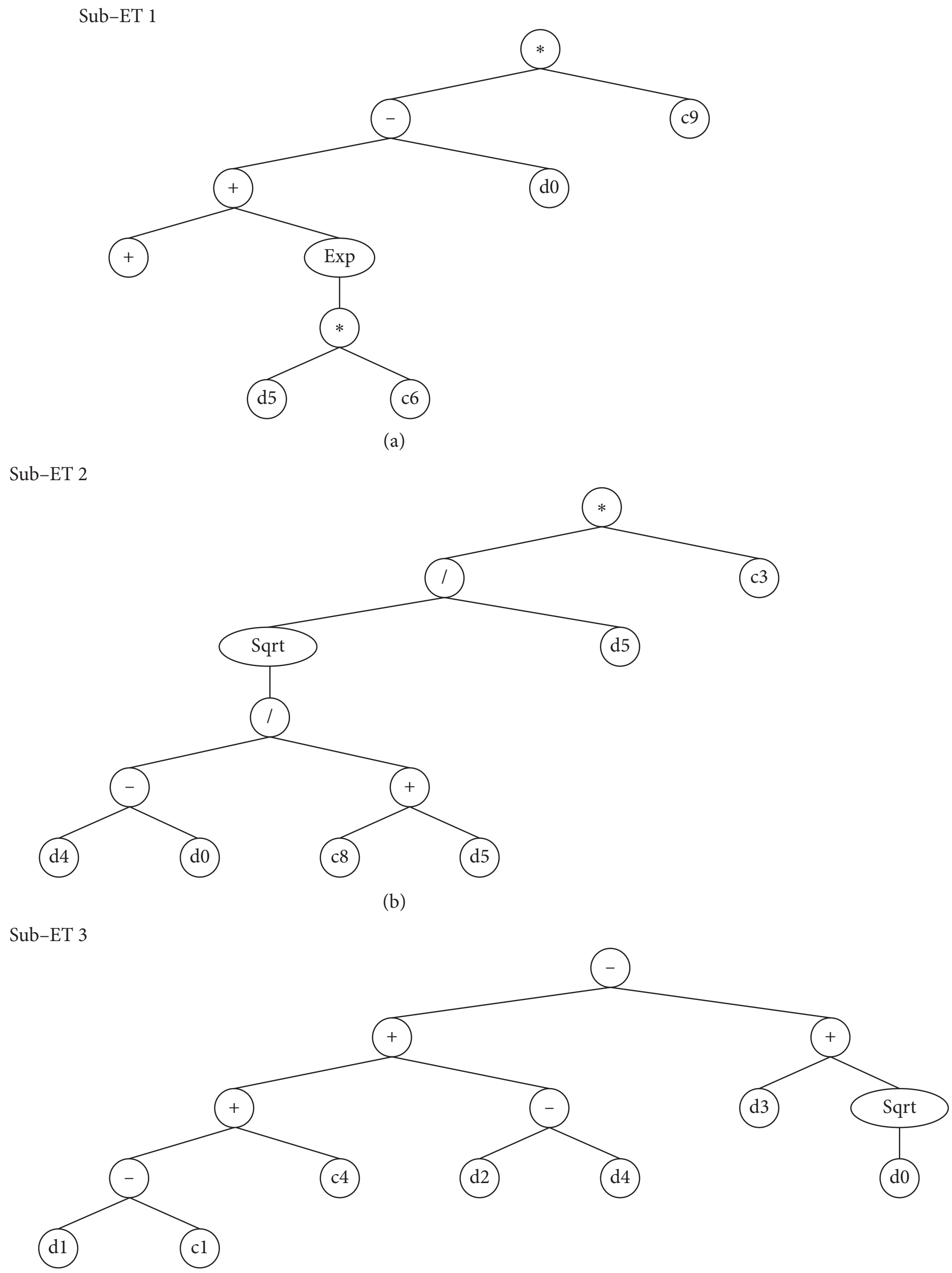

FIGURE 8: Expression tree of the model formulated for UCS (where d0: HARHA, d1: w_L, d2: w_P, d3: I_P, d4: w_{OMC}, d5: AC, d6: δmax, Sub-ET1 C9: -9.177, Sub-ET1 C1: -9.522, Sub-ET1 C6: -4.669, Sub-ET2 C3: 5.995, Sub-ET2 C8: 0.07915, Sub-ET3 C4: 5.3481, and Sub-ET3 C1: 3.479).

c6 = 2.38074892422254
c7 = −3.36100344859157
c8 = 7.98272652363659
c9 = 3.71135593737602

Gene 2

c0 = -5.65450864340739
c1 = −7.65190588091678E-02
c2 = 0.593482469817356
c3 = −0.21698660237434
c4 = −7.5964995269631
c5 = −6.84987945188757
c6 = 3.66069521164586
c7 = 1.44131669080772
c8 = −7.00961638233589
c9 = 8.11291842097232

Gene 3

c0 = 2.97519449316012

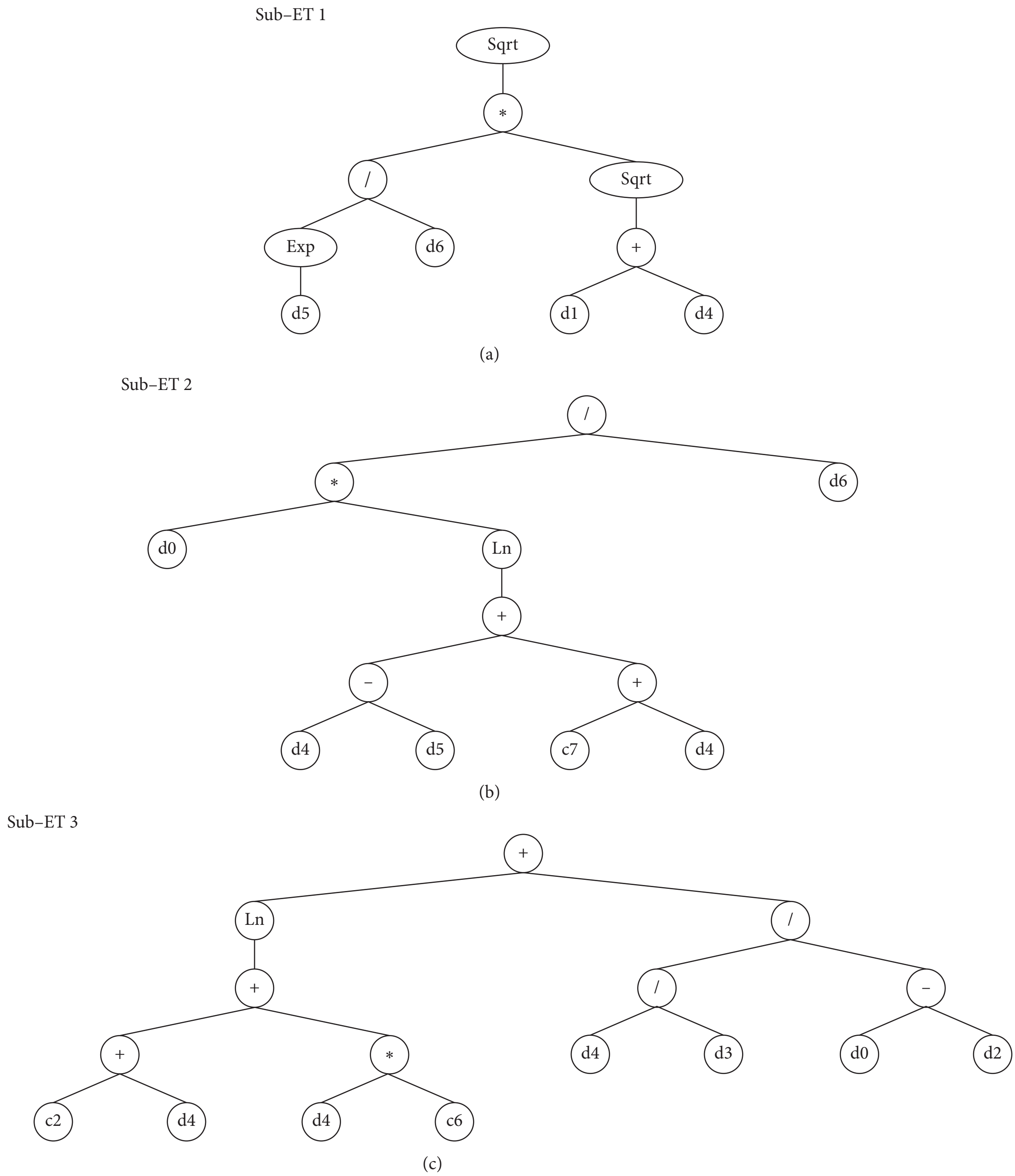

Figure 9: Expression tree of the model formulated for R value (where d0: HARHA, d1: w_L, d2: w_P, d3: I_P, d4: w_{OMC}, d5: AC, d6: δmax, Sub-ET2 C7: 1.4413, Sub-ET3 C2: -12.398, Sub-ET3 C6: 3.212).

c1 = −2.45399334696493
c2 = −12.3985913762825
c3 = 3.00576799829096
c4 = −6.60390026551103
c5 = 5.46067690054018
c6 = 3.21220500714347
c7 = 3.68913754692221
c8 = −10.8087886989959
c9 = −6.13330484939116

It has been reported earlier that multilinear regression (MLR) was conducted to evaluate quantitatively the relationships between the input soil hydraulic-prone properties and output soil strength properties, i.e., CBR, $\mathrm{UCS}_{28,}$ and R_{Value}. Each output value was defined as a combination of the six soil parameters (HARHA, w_L, w_P, I_P, w_{OMC}, A_C, and $\delta_{\max}$, respectively), and the following equations were derived:

$$\mathrm{CBR} = 21.423 + 4.314\,\mathrm{HARHA} + 0.352 w_L + 0.379 w_P - 1.592 w_{\mathrm{OMC}} - 7.27 A_C - 4.871 \delta_{\max}, \quad (2)$$

$$\mathrm{UCS}_{28} = 126.54 + 12.275\,\mathrm{HARHA} + 3.330 w_L - 1.519 I_P - 2.21 w_{\mathrm{OMC}} - 46.67 A_C - 22.30 \delta_{\max}, \quad (3)$$

$$R_{\mathrm{Value}} = 3.219 + 1.07\,\mathrm{HARHA} - 0.0694 w_L - 0.015 w_P + 1.163 w_{\mathrm{OMC}} - 1.199 A_C - 3.225 \delta_{\max}. \quad (4)$$

These are useful tools to estimate the soil strength properties based on easily determinable geotechnical indices for HARHA treated expansive soils. However, these MLR equations can only be employed in the case when the points show linearly changing behavior [27]. These equations were derived from making a comparison with the developed GEP models for CBR, $\mathrm{UCS}_{28,}$ and R_{Value}.

Using the expression trees given from Figures 6–9 for evaluating the CBR, $\mathrm{UCS}_{28,}$ and R_{Value} of soils, respectively, decoding was done to derive the three simple mathematical expressions (equations (5)–(7)) as follows:

$$\mathrm{CBR} = \left(\left(w_{\mathrm{OMC}} \left(w_L - 1 \right) - 444.6 \right)^{1/4} \right) + \left(\sqrt{\mathrm{HARHA}} + \mathrm{HARHA} \right) + \left(\left(\frac{1.46 w_P}{w_L} + \exp\left(\ln\, A_C \right) \right) * \delta_{\max} \right), \quad (5)$$

$$\mathrm{UCS}_{28} = \left(\left(9.18\,\mathrm{HARHA} - 9.18\, \exp\left(-4.67 A_C \right) \right) \right) + \left(\sqrt{\left(\frac{w_{\mathrm{OMC}} - \mathrm{HARHA}}{A_C + 7.916} \right)} * \frac{5.995}{A_C} \right) + \left(2 w_P - w_{\mathrm{OMC}} - \sqrt{\mathrm{HARHA}} + 89.25 \right), \quad (6)$$

$$R_{\mathrm{Value}} = \left(\sqrt{\left(\frac{\exp\, A_C}{\delta_{\max}} \right) * \sqrt{w_L + w_{\mathrm{OMC}}}} \right) + \left(\frac{\mathrm{HARHA} * \ \ln\left(2 w_{\mathrm{OMC}} - A_C + 1.441 \right)}{\delta_{\max}} \right) + \left(\left(\ln\left(4.212 w_{\mathrm{OMC}} - 12.40 \right) \right) + \left(\frac{w_{\mathrm{OMC}}}{I_P \left(\mathrm{HARHA} - w_P \right)} \right) \right). \quad (7)$$

The comparisons between the predicted and the observed expansive soil parameters are shown in Figure 10. The indicators indicate high accuracy can be observed for CBR, UCS_{28}, and R_{Value}, with higher R^2 values for GEP formulated models. This suggests that the prediction of the output parameters using the proposed model is in good agreement with the testing data.

It can be seen in Figure 11 that the range of error distribution for CBR and R_{Value} is significantly lower in contrast to that of UCS_{28}. It could be attributed to the larger SD value and range of data for the UCS_{28}, as reflected in Table 1. In addition, the GEP proposed models exhibit superior performance for CBR and R_{Value} cases in comparison with the respective MLR plots. However, the results of GEP are not better than that of the MLR model in terms of error distribution which is shown in Figures 7(c) and 7(d), respectively.

Finally, the summary of statistical performance is listed in Table 5. Variety of performance indices have been determined, including root mean square error (RMSE), mean absolute error (MAE), root square error (RSE), Nash–Sutcliffe efficiency (NSE), relative root mean square error (RRMSE), coefficient of correlation (R), performance index (ρ), and objective function (OBF) to evaluate the performance of developed CBR, UCS, and R value GEP models. The following equations were used to calculate the performance indices. The RMSE errors are squared, implying that relatively a much larger weight is assigned to the larger errors. High R values and low RRMSE values achieve a high degree of accuracy, which agrees with the results of Gandomi and Roke [25]. The proposed models indicate that the MAE, RMSE, RSE, and RRMSE values are significantly lower while the NSE and R values are larger for the CBR and $R_{\mathrm{value,}}$ which shows superior model performance. However, these values are vice versa in the case of UCS_{28} that leads to lower performance. Similarly, the performance indices and OBF values are well within allowable limits in the literature [32, 35, 36]. These results further show that the proposed models of CBR and R_{Value} using GEP were much better than for the case of UCS_{28}, thereby achieving reliable and accurate results. The range of data for the input parameters of UCS_{28} is several times greater than those of CBR and R_{Value}, which is also reflected in Table 2. So, GEP models were used to formulate simple mathematical equations which can be readily employed to predict CBR, $\mathrm{UCS}_{28,}$ and R_{Value} values, as mentioned earlier in detail.where

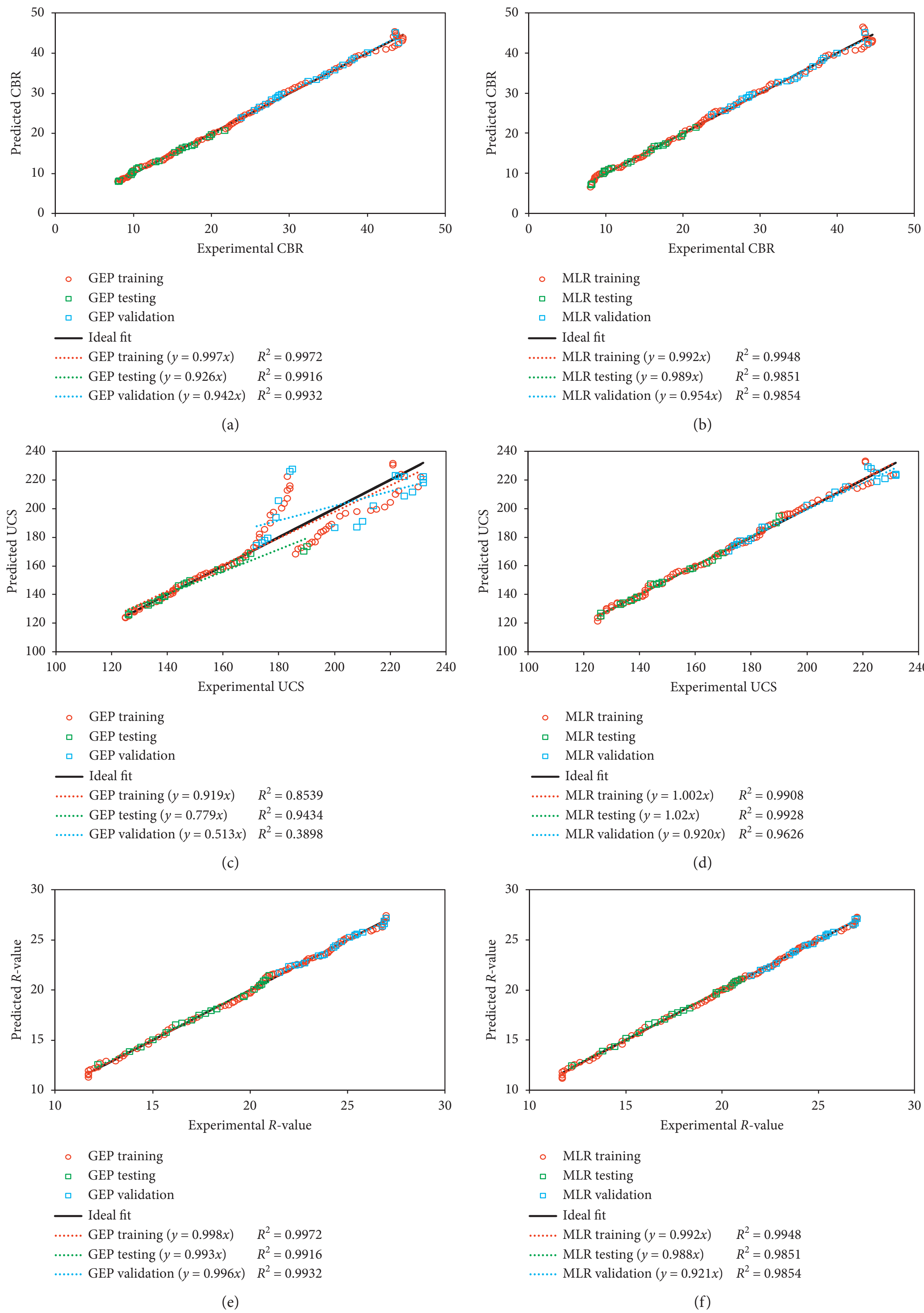

Figure 10: Experimental versus predicted values of CBR, UCS, and R values for models formulated using GEP and MLR, respectively.

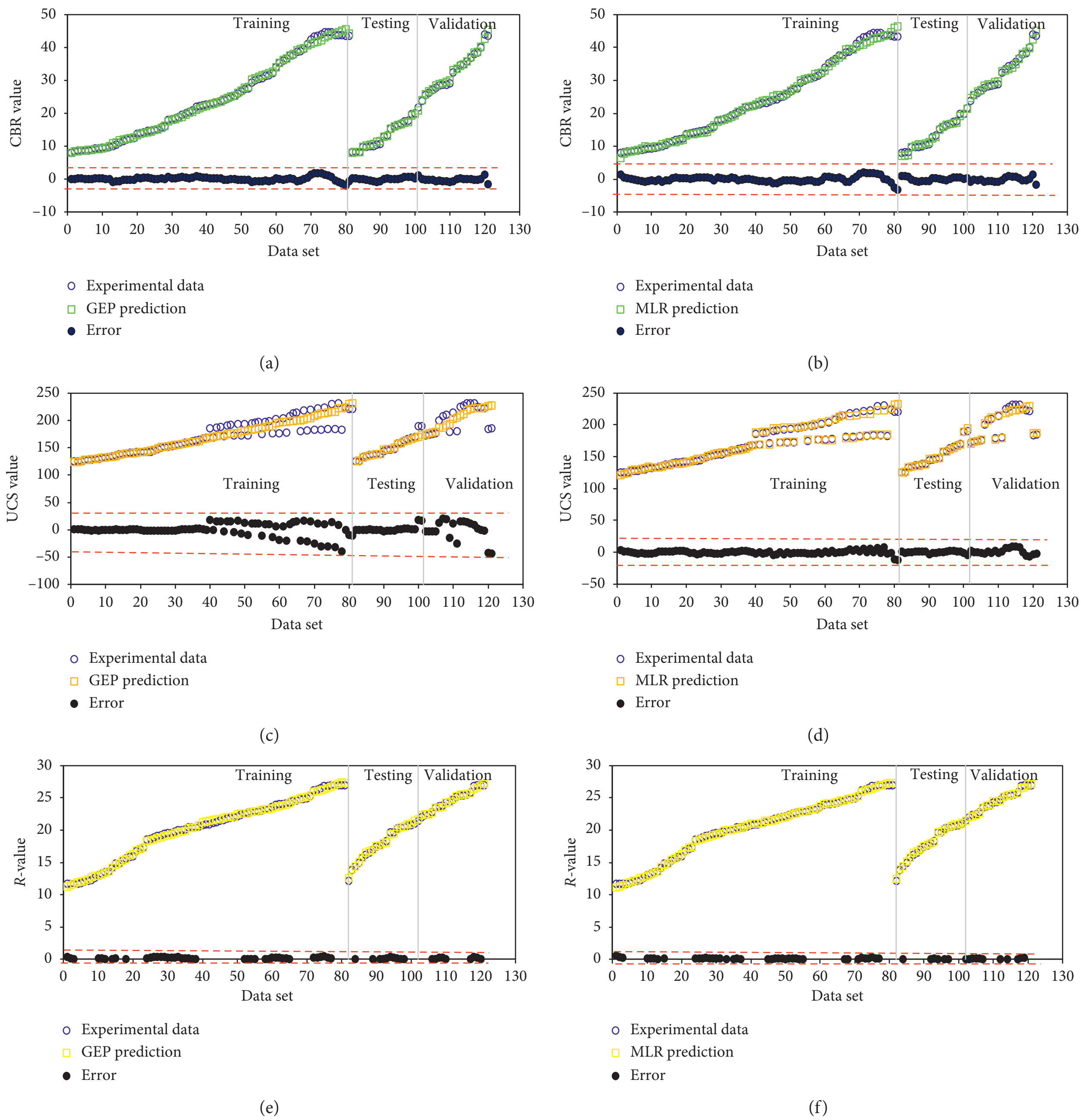

Figure 11: Error distribution diagram for CBR, UCS, and R value generated models using GEP and MLR, respectively.

Table 5: Statistical calculations, performance indices, and objective functions of the GEP models for CBR, UCS, and R value.

GEP model	Dataset	Statistical parameters							
		MAE	RSE	RMSE	NSE	R	RRMSE	ρ	OBF
CBR	Training	0.5	0.003	4.94	0.997	0.998	0.202	0.101	0.028
	Testing	0.3	0.011	3.69	0.989	0.996	0.271	0.136	
	Validation	0.5	0.011	5.49	0.989	0.996	0.167	0.084	
UCS_{28}	Training	7.8	0.151	13.06	0.849	0.924	0.076	0.040	0.013
	Testing	2.7	0.970	12.21	0.903	0.971	0.081	0.041	
	Validation	13.7	0.647	13.61	0.353	0.624	0.067	0.041	
R_{Value}	Training	0.2	0.003	4.49	0.997	0.998	0.222	0.111	0.032
	Testing	0.2	0.007	4.22	0.992	0.997	0.238	0.119	
	Validation	0.1	0.012	4.72	0.988	0.994	0.193	0.097	

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (e_i - p_i)^2}{n}},$$
$$\text{MAE} = \frac{\sum_{i=1}^{n} |e_i - p_i|}{n},$$
$$\text{RSE} = \frac{\sum_{i=1}^{n} (p_i - e_i)^2}{\sum_{i=1}^{n} (\bar{e} - e_i)^2},$$
$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n} (e_i - p_i)^2}{\sum_{i=1}^{n} (p_i - \overline{p_i})^2}, \tag{8}$$
$$\text{RRMSE} = \frac{1}{|\bar{e}|} \sqrt{\frac{\sum_{i=1}^{n} (e_i - p_i)^2}{n}},$$
$$R = \frac{\sum_{i=1}^{n} (e_i - \bar{e}_i)(p_i - \overline{p}_i)}{\sqrt{\sum_{i=1}^{n} (e_i - \bar{e}_i)^2 \sum_{i=1}^{n} (p_i - \overline{p}_i)^2}},$$
$$\rho = \frac{\text{RRMSE}}{1 + R},$$
$$\text{OBF} = \left(\frac{n_T - n_v}{n}\right)\rho_T + 2\left(\frac{n_v}{n}\right)\rho_v,$$

where e_i and p_i are the i number of experimental and predicted outputs, respectively; $\bar{e}_i$ and $\overline{p}_i$ are the average values of the experimental and predicted output values, respectively, and n is total the number of samples.

4. Conclusions

From the gene expression programming of California bearing ratio, unconfined compressive strength and resistance value of hydrated-lime modified expansive soil with input parameters; HARHA, liquid limit (w_L), (plastic limit (w_P), plasticity index (I_P), optimum moisture content (w_{OMC}), clay activity (A_C), (maximum dry density ($\delta_{\max}$), CBR, UCS, and R value generated from series of laboratory exercise which produced 121 datasets, the following can be concluded:

(1) The A-7-6 expansive soil and hydrated-lime activated rice husk were blended in varying proportions of the additive to the soil, and the modified blend specimens were tested to get the liquid limit, plastic limit, plasticity index, optimum moisture content, clay activity, maximum density, California bearing ration, unconfined compressive strength, and resistance value responses.

(2) The responses were deployed to both MLR and GEP evolutionary operations to model the output parameters: CBR, UCS, and R value.

(3) The outcome of the GEP training, testing, and validation of the datasets showed a consistent agreement between the MLR and GEP.

(4) Three model equations were formed, each of MLR and GEP under optimized conditions, and the agreement between the predicted models and the generated datasets is above 0.9.

(5) Generally, the GEP showed that design, construction, performance, and infrastructure management could be predicted with perfect accuracy using the gene expression programming soft computing method for sustainable earthworks and other engineering operations. This can be easily implemented when the treatment materials for construction are similar in properties to the ones used in this project and also when similar numbers of predictor parameters are used in proposing the model.

(6) Lastly, it can be recommended to have more multiple experiments to generate upwards of a thousand datasets for a perfect and more reliable outcome.

Data Availability

The data used in the study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] K. C. Onyelowe, T. Amhadi, C. Ezugwu et al., "Strength of pozzolan soil blend in chemically improved lateritic soil for pavement base material purpose," *International Journal of Low-Carbon Technologies*, vol. 14, no. 3, pp. 410–416, 2019.

[2] K. C. Onyelowe, E. Onukwugha, A. B. Salahudeen et al., "Microstructural and mineralogical analysis of weak erodible soil for gully site study and solutions," *NIPES Journal of Science and Technology Research*, vol. 1, no. 3, pp. 24–37, 2019.

[3] K. Onyelowe et al., "Sorptivity, swelling, shrinkage, compression and durability of quarry dust treated soft soils for moisture bound pavement geotechnics," *Journal of Materials Research and Technology*, vol. 8, no. 4, pp. 3529–3538, 2019.

[4] K. C. Onyelowe, F. D. A. Onyelowe, D. Bui Van et al., "Valorization and sequestration of hydrogen gas from biomass combustion in solid waste incineration NaOH oxides of carbon entrapment model (SWI-NaOH-OCE Model)," *Materials Science for Energy Technologies*, vol. 3, pp. 250–254, 2020.

[5] K. C. Onyelowe, D. Bui Van, M. O. Idrees et al., "An experimental study on compaction behaviour of lateritic soils treated with quarry dust based geopolymer cement," *The Journal of Solid Waste Technology and Management*, In press, 2021.

[6] K. C. Onyelowe, M. E. Onyia, E. R. Onukwugha et al., "Atterberg limits of modified compacted clayey soil for sustainable green subgrade structure," *Jurnal Kejuruteraan*, In press, 2021.

[7] K. C. Onyelowe, D. Bui Van, L. Dao-Phuc et al., "Evaluation of index and compaction properties of lateritic soils treated with quarry dust based geopolymer cement for subgrade purpose," *Epitőanyag– Journal of Silicate Based and Composite Materials*, vol. 72, no. 1, pp. 12–15, 2020.

[8] K. C. Onyelowe, M. E. Onyia, F. D. A. Onyelowe et al., "Critical state desiccation induced shrinkage of biomass treated compacted soil as pavement foundation Epitőanyag–,"

Journal of Silicate Based and Composite Materials, vol. 72, no. 2, pp. 40–47, 2020.

[9] K. C. Onyelowe, E. J. Fazal, E. O. Michael, C. O. Ifeanyichukwu, U. A. George, and I. Chidozie, "Artificial intelligence prediction model for swelling potential of soil and quicklime activated rice husk ash blend for sustainable construction," *Jurnal Kejuruteraan*, In press, 2021.

[10] K. C. Onyelowe, M. E. Onyia, D. Bui Van et al., "Shrinkage parameters of modified compacted clayey soil for sustainable earthworks," *Jurnal Kejuruteraan*, vol. 33, no. 1, pp. 133–140. In press, 2021.

[11] K. C. Onyelowe, F. A. Deborah, and D. Bui Van, "Overview of ash as supplementary cementitious silicate-based composite and construction material," *Epitőanyag–Journal of Silicate Based and Composite Materials*, vol. 72, no. 3, pp. 80–85, 2020.

[12] K. C. Onyelowe, D. Bui Van, C. Ikpa et al., "Resilient modulus and deviatoric stress of cemented soils treated with crushed waste ceramics (CWC) for pavement subgrade construction," *Epitőanyag–Journal of Silicate Based and Composite Materials*, vol. 72, no. 3, pp. 86–90, 2020.

[13] K. C. Onyelowe, M. E. Onyia, E. R. Onukwugha, O. C. Nnadi, I. C. Onuoha, and F. E. Jalal, "Polynomial relationship of compaction properties of silicate-based RHA modified expansive soil for pavement subgrade purposes," *Epitőanyag–Journal of Silicate Based and Composite Materials*, vol. 72, no. 6, pp. 223–228, 2020.

[14] D. Bui Van, K. C. Onyelowe, M. Onyia et al., "Strength and consistency behaviour of replacement of cement with silicate-based geopolymer cement modified soft soil treated with crushed waste glasses for pavement underlain," *Epitőanyag–Journal of Silicate Based and Composite Materials*, vol. 72, no. 6, pp. 186–197, 2020.

[15] A. B. Salahudeen, J. A. Sadeeq, A. Badamasi, and K. C. Onyelowe, "Prediction of unconfined compressive strength of treated expansive clay using back-propagation artificial neural networks," *Nigerian Journal of Engineering*, vol. 27, no. 1, pp. 45–58, 2020.

[16] G. U. Alaneme, K. C. Onyelowe, M. E. Onyia et al., "Modelling of the swelling potential of soil treated with quicklime-activated rice husk ash using fuzzy logic," *Umudike Journal of Engineering and Technology (UJET)*, vol. 6, no. 1, pp. 1–12, 2020.

[17] G. U. Alaneme, K. C. Onyelowe, M. E. Onyia et al., "Modeling volume change properties of hydrated-lime activated rice husk ash (HARHA) modified soft soil for construction purposes by artificial neural network (ANN)," *Umudike Journal of Engineering and Technology (UJET)*, vol. 6, no. 1, pp. 1–12, 2020.

[18] G. U. Alaneme, K. C. Onyelowe, M. E. Onyia et al., "Comparative modelling of strength properties of hydrated-lime activated rice-husk-ash (HARHA) modified soft soil for pavement construction purposes by artificial neural network (ANN) and fuzzy logic (FL)," *Jurnal Kejuruteraan*, In press), 2021.

[19] K. C. Onyelowe and M. E. Onyia, "Analytics of swelling potential on highly expansive (plastic) clayey soils behavior for sustainable admixture stabilization," *Jurnal Kejuruteraan*, In press, 2021.

[20] R. S. Faradondeh, D. J. Armaghani, M. Monjezi, and E. T. Mohamad, "Genetic programming and gene expression programming for flyrock assessment due to mine blasting," *International Journal of Rock Mechanics and Mining Sciences*, vol. 88, pp. 254–264, 2016.

[21] C. Ferreira, "Gene expression programming in problem solving," *Soft Computing and Industry*, Springer, Berlin, Germany, 2002.

[22] J. R. Koza and J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT press, Cambridge, MA, USA, 1992.

[23] Z. L. Cheng, W. H. Zhou, and A. Garg, "Genetic programming model for estimating soil suction in shallow soil layers in the vicinity of a tree," *Engineering Geology*, vol. 268, Article ID 105506, 2020.

[24] A. Garg, A. Garg, and K. Tai, "A multi-gene genetic programming model for estimating stress-dependent soil retention curves," *Computational Geosciences*, vol. 18, pp. 45–56, 2014.

[25] A. H. Gandomi and D. A. Roke, "Assessment of artificial neural network and genetic programming as predictive tools," *Advances in Engineering Software*, vol. 88, pp. 63–72, 2015.

[26] K. Pearson, "X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Sciences*, vol. 50, pp. 157–175, 1900.

[27] J. Gravier, V. Vignal, S. Bissey-Breton, and J. Farre, "The use of linear regression methods and Pearson's correlation matrix to identify mechanical-physical-chemical parameters controlling the micro-electrochemical behavior of machined copper," *Corrosion Science*, vol. 50, pp. 2885–2894, 2008.

[28] K. N. Sircombe, "Age display: an EXCEL workbook to evaluate and display univariance geochronological data using binned frequency histograms and probability density distributions," *Computers and Geosciences*, vol. 30, no. 1, pp. 21–31, 2004.

[29] M. E. Edjabou, J. A. Martin-Fernandez, C. Scheutz, and T. F. Astrup, "Statistical analysis of solid waste composition data: arithmetic mean, standard deviation and correlation coefficients," *Waste Management*, vol. 68, pp. 13–23, 2017.

[30] C. Sharma and C. Ojha, "Statistical parameters of hydrometeorological variables: standard deviation, SNR, skewness and kurtosis," *Advances in Water Resources Engineering and Management*, Springer, Berlin, Germany, 2020.

[31] M. K. Cain, Z. Zhang, and K. H. Yuan, "Univariate and multivariate skewness and kurtosis for measuring nonnormality: prevalence, influence and estimation," *Behavior Research Methods*, vol. 49, pp. 1716–1735, 2017.

[32] M. F. Iqbal, Q. F. Liu, I. Azim et al., "Prediction of mechanical properties of green concrete incorporating waste foundry sand based on gene expression programming," *Journal of Hazardous Materials*, vol. 384, Article ID 121322, 2020.

[33] G. B. Jumaa and A. R. Yousef, "Predicting shear capacity of FRP-reinforced concrete beams without stirrups by artificial neural networks, gene expression programming and regression analysis," *Advances in Civil Engineering*, vol. 2018, Article ID 5157824, , 2018.

[34] M. F. Javed, M. N. Amin, M. I. Shah et al., "Applications of gene expression programming and regression techniques for estimating compressive strength of bagasse ash-based concrete," *Crystals*, vol. 10, p. 737, 2020.

[35] M. I. Shah, M. F. Javed, and T. Abunama, "Proposed formulation of surface water quality and modeling using gene expression, machine learning and regression techniques," *Environmental Science and Pollution Research*, vol. 28, pp. 13202–13220, 2021.
[36] K. C. Onyelowe, M. Onyia, E. R. Onukwugha, D. Bui Van, J. Obimba-Wogu, and C. Ikpa, "Mechanical properties of fly ash modified asphalt treated with crushed waste glasses as fillers for sustainable pavements," *Epitőanyag–Journal of Silicate Based and Composite Materials*, vol. 72, no. 6, pp. 219–222, 2020.