Neural Computing and Applications
https://doi.org/10.1007/s00521-024-09919-0

ORIGINAL ARTICLE

Predictive models for treated clayey soils using waste powdered glass and expanded polystyrene beads using regression analysis and artificial neural network

E. Akis[1] · O. Y. Cigdem[2]

Received: 28 May 2022 / Accepted: 28 April 2024

Abstract
Waste materials contribute to a wide range of environmental and economic problems. To minimize their effects, a safe strategy for reducing such negative impact is required. Recycling and reusing waste materials have proved to be effective measures in this respect. In this study, an eco-friendly treatment is investigated based on using waste powdered glass (*WGP*) and *EPS* beads (*EPSb*) as mechanical and chemical admixers in soils. For this purpose, Atterberg limit, standard proctor, free swell, and unconfined compression tests are performed on soil samples with different ratios of waste materials at their optimum moisture contents. The obtained test results indicate that adding *WGP* to cohesive soils increases the unconfined compressive strength (*UCS*) and reduces free swell (*FS*). In contrast, using *EPSb* reduces both *FS* and *UCS* of the treated soil samples. An optimum combination of both waste materials is determined for the improvement of the properties of high plasticity clay used in this study. Furthermore, multiple linear regression (*MLR*) and artificial neural network (*ANN*) methods are used to predict the *FS* and *UCS* of the clayey soils based on the data obtained here and the experimental test results reported in the literature. Once the *FS* and *UCS* values of untreated soil and additive percentages are defined as independent variables, both methods are shown to predict the *FS* and *UCS* values of the treated soil samples on a satisfactory level with the coefficient of correlation (R^2) values greater than 0.926. Additionally, when only the index properties (liquid limit, plastic limit, and plasticity index) of the soil samples with waste materials are used as dependent variables, the R^2 values obtained by the *ANN* method are 0.968 and 0.974 for *FS* and *UCS*, respectively. The results of the untreated soil samples' *FS* and *UCS* tests are known, and the linear regression and ANN techniques yield similar results. Lastly, the *ANN* method is used to predict the *FS* and *UCS* of the treated samples in accordance to the limited predictors (e.g., only the Atterberg limits of the soil sample).

Keywords Expanded polystyrene beads treatment · Waste glass powder treatment · Free swell · Unconfined compressive strength · Regression analysis · Artificial Neural Network (*ANN*)

1 Introduction

Climate change is one of the most serious environmental problems that has been affecting the world in the last decades. In order to find a global solution to this issue, sustainable development is a must and includes waste management. Waste management not only consists of the collection, transportation, and recovery of natural resources, but it also covers the disposal of waste through recycling or reuse of the materials [1]. One of the solutions is the use of waste materials for soil improvement, which leads to the reduction of the settlement of structures, increase in the bearing capacity of the soil, reducing the voids in the soil, increasing the safety factor against slope failure, and controlling shrinkage and swelling.

There is a large amount of expanded polystyrene (*EPS*) waste since it cannot be degraded in nature. Globally, 350 million tons of plastic waste are produced every year, and

✉ E. Akis
ebru.akis@atilim.edu.tr

[1] Department of Civil Engineering, Atilim University, Ankara 06830, Turkey

[2] Graduate School of Natural and Applied Sciences, Atilim University, Ankara, Turkey

Published online: 21 May 2024

EPS accounts for 6.26% of this plastic waste [2]. The use of expanded polystyrene beads (*EPSb)* for soil improvement purposes may directly reduce the amount of *EPS* waste in nature since soil improvement requires large amounts of material. Several researchers, such as Shirazi et al. [3], Illuri [4], Rocco [5], and Soundara and Selvakumar [6] investigate the effects of *EPSb* on the engineering properties of different soil types. These studies show that the addition of *EPSb* improves the swelling potential of the treated soils by reducing the free swell (*FS*) values.

Another material that is commonly discarded, glass has the lowest recycling percentage (about 4.43%) [7]. The main benefits of recycling glass are a 25% reduction in energy consumption, a 20% reduction in air pollution, an 80% reduction in mine waste, and a 50% reduction in water consumption. These values can be doubled by expanding the use of waste glass [8]. There are several studies in the literature that investigate the use of waste glass as an additive for soil improvement through laboratory tests [9–14]. These studies show that the addition of glass powder leads to an increase in the unconfined compressive strength (*UCS*) of the treated specimens and a decrease in the free swell (*FS*) values. Therefore, glass powder can be considered as a strength-swell modifier, while *EPSb* is a swell modifier only.

Although *EPSb* has the potential to improve the *FS* behavior of soils, there has been few research investigating the effect of both *EPSb* and *WGP* on the *FS* and *UCS* values. In the first part of this study, waste glass powder (*WGP*) and/or expanded polystyrene beads (*EPSb*) are used as additives, and their effects on the engineering properties of high plasticity clay (*CH*) are investigated both individually and together by performing a series of experiments. In these experiments, the liquid limit, plastic limit, plasticity index, optimum moisture content, maximum dry density, free swell, and unconfined compressive strength of the treated soil samples are obtained.

It is known that the *ANN* method can be widely and successfully used to estimate the mechanical properties of engineering materials [15–25]. Therefore, in the second part of the study, the *FS* and *UCS* of the treated soil samples are predicted by artificial neural network (*ANN*) methods. In addition, multiple linear regression analyses are performed to provide empirical equations for the prediction of *FS* and *UCS*. The training and testing datasets based on the test results and those obtained by relevant studies in the literature are gathered. These datasets consist of the Atterberg limits, additive amounts, and test results of untreated soils as the independent variables, as well as the *FS* or *UCS* of treated samples as dependent variables. Firstly, multiple linear regression (*MLR*) analyses are carried out and empirical correlations are proposed. Then, a previously developed *ANN* algorithm [15] is used to predict the *FS* and *UCS* of the treated samples. Finally, since the *ANN* models may heavily depend on the selection of datasets, a k-fold cross-validation analysis is carried out.

2 Materials

2.1 Bentonite

Bentonites, often referred to as clay rocks, are primarily composed of clay minerals. Renowned for their remarkable physical and chemical properties, they find extensive use across diverse industries. These properties encompass their crystal structure, chemical composition, high ion exchange capacity, hydration, and swelling behavior. Due to these attributes, bentonites are applicable in a broad spectrum of uses, ranging from clarification processes to applications in civil engineering [26]. The concentration of ions within these clay rocks determines the type of application that is most suitable for any given bentonite. Bentonites, in general, are divided into three main groups: sodium bentonite, calcium bentonite and sodium-calcium bentonite. Sodium bentonite exhibits the highest swelling capacity among these three types, expanding up to 10–15 times its volume. It finds primary use in the drilling industry. Sodium-calcium bentonite represents an intermediate type, with a swelling capacity of approximately 5–7 times their volume. In contrast, calcium bentonite has a lower swelling capacity, expanding only about 2–4 times its volume. It is commonly utilized in cat litter and bleaching agents [27].

In this study, calcium bentonite is used, and the present section provides the information related to the physical and chemical properties of the used soil. The soil is obtained from Eczacıbaşı Esan Industrial Raw Material Company. It is commercially available bentonite supplied in milled and packaged form. The soil has an *FS* value of 30.41% and is classified as fat clay (*CH*) according to the unified soil classification system (*USCS*). The chemical composition of the bentonite is determined by an X-ray fluorescence test at the Earth Science Applications and Research Centre of Ankara University. The particle size distribution and index properties of the bentonite are given in Table 1, while its chemical composition is depicted in Table 2.

2.2 Glass powder

The disposal of waste materials—such as tires, plastics, and glass—poses a significant challenge for many countries due to climate change. Currently, only a small fraction of these materials is recycled globally, with just 10% of tires, 19.5% of plastics, and 21% of glass being reused. The substantial amount of the remaining waste is typically sent

Table 1 Physical properties of the bentonite

Property	Value	Standard
Specific gravity (G_s)	2.50	ASTM D854-14 [28]
Optimum water content (w_{opt}) (%)	52.5	ASTM D698 [29]
Maximum dry density ($\gamma_{dry,max}$) (g/cm^3)	0.994	ASTM D698 [29]
Sand (%)	0	ASTM D6913 [30]
Fines content (%)	100	ASTM D6913 [30]
Silt (%)	19	ASTM D7928-21 [31]
Clay (%)	81	ASTM D7928-21 [31]
LL (%)	161.7	BS:1377: Part 2 [32]
PL (%)	47.6	ASTM D4318-17 [33]
PI (%)	114.1	ASTM D4318-17 [33]
Free Swell (%)	30.41	ASTM D4546-14 [34]
USCS classification	CH	ASTM D2487-17 [35]

Table 2 Chemical composition of the bentonite

Element	Value (%)	Element	Value (%)
SiO_2	73.57	MnO	0.0691
Al_2O_3	12.68	Na_2O	0.039
LoI	7.84	SO_3	0.00687
MgO	3.359	V_2O_5	0.00093
CaO	1.318	Cr_2O_3	0.00071
K_2O	1.076	Cl	0.0002
Fe_2O_3	0.741	P_2O_5	0.0014
TiO_2	0.052		

Table 3 Physical properties of the *WGP*

Property	Value	Standard
Specific gravity (G_s)	2.56	ASTM D854-14 [28]
Sand (%)	73.01	ASTM D6913 [30]
Fines content (%)	27.0	ASTM D6913 [30]
Silt (%)	26.4	ASTM D7928-21 [31]
Clay (%)	0.6	ASTM D7928-21 [31]

Table 4 Chemical composition of the *WGP*

Element	Value (%)	Element	Value (%)
SiO_2	66.75	MnO	0.00504
Al_2O_3	0.009	Na_2O	10.35
LoI	9.75	SO_3	0.1977
MgO	3.533	V_2O_5	0.00091
CaO	8.258	Cr_2O_3	0.00891
K_2O	0.2205	Cl	0.00131
Fe_2O_3	0.2952	P_2O_5	0.0021
TiO_2	0.0511		

to landfills, contributing to environmental pollution. Consequently, it is crucial to develop new strategies for utilizing these landfill materials more effectively [36]. Given that glass consists the largest percentage among all waste materials, this study focuses on utilizing waste glass and examining its impact on a high plastic clayey soil. The waste glass powder is provided by Boyabat Osmanlı Cam Mozaik glass and ceramics factory. As a sieve analysis is carried out on the material, it is seen that almost all (99.7%) of the material passes through the No. 30 sieve. Therefore, it is decided to use the waste glass powder (*WGP*) that passes through this sieve. The results of sieve analyses, hydrometer test, and the specific gravity of the *WGP* are given in Table 3. Additionally, an X-ray fluorescence test is performed at the Earth Science Applications and Research Centre of Ankara University. The results of this test are shown in Table 4.

2.3 Expanded polystyrene beads

Expanded polystyrene (*EPS*) is a petroleum-derived foamy, closed-cell thermoplastic material that is made up of 98% dry air and 2% polystyrene [37]. Even though the material is considered to be environmentally friendly and 100% recyclable, its negative impact on nature should not be ignored. EPS can accumulate in landfills and marine environments, posing risks to wildlife and ecosystems due to its slow decomposition rate and potential for leaching harmful chemicals [37]. In addition to its conventional uses in packaging and insulation, EPS has gained attention in geotechnical engineering for its capacity to improve the mechanical properties of soils**.** Besides, research has revealed that using *EPS* as an additive, particularly in bead

form (*EPSb)* as an additive has a positive effect on the *FS* of expansive soils. In this study, the use of waste *EPS* in geotechnical applications is investigated by employing *EPSb* as an additive for the treatment of bentonite. The *EPSb* used in the tests is obtained from the Istanbul Strafor Factory. The size of the *EPSb* used in the treatment of soil samples is in the range of 1.00–4.75 mm as shown in Fig. 1.

3 Testing and results

In this study, the effects of *EPSb* and *WGP* on the improvement of fat clay are examined separately. Then, the properties of the treated samples, when both of the waste materials are used as additives, are investigated. The bentonite samples are mixed with additives (*EPS* and/or *WGP*) at different percentages by considering the dry weight of the soil. Firstly, the index properties, optimum moisture content (*OMC*), and the maximum dry density (*MDD*) of both the treated and the untreated soil samples are determined. Next, the *FS* and *UCS* tests are performed on both samples.

3.1 Sample preparation

For the *FS* and *UCS* tests, the soil samples and additives are initially mixed with distilled water according to the *OMC* of related admixtures. Then, they are placed in plastic bags for at least 36 h to allow the material to moist as per ASTM 3080 [38]. The soil samples are compacted at *OMC* and at *MDD* by using a standard proctor test apparatus according to ASTM D698 [29]. The samples for the *FS* and *UCS* tests are prepared from the standard proctor mold. The abbreviation '*BGiEj'* is used to represent treated bentonite (*B*) with *i* percent waste powdered glass (*G*) and *j* percent expanded polystyrene beads (*E*) as shown in Table 5. The percentage of the additives is determined by considering the dry weight of the bentonite. As an example, the specimen symbol '*BG2E0.9*' indicates a bentonite sample containing 2% of waste glass powder and 0.9% of *EPSb*.

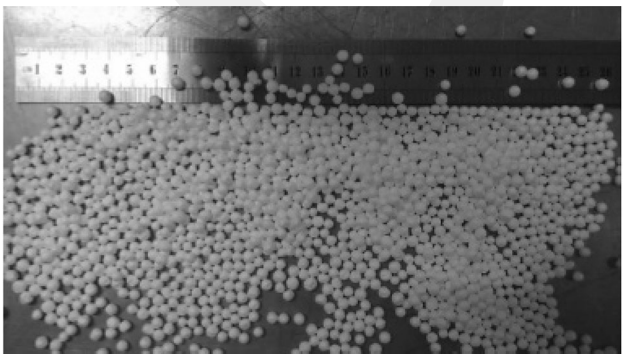


Fig. 1 Sizes of *EPSb*

Table 5 Abbreviations used to represent the additives of bentonite

Abbreviations	*WGP* (%)	*EPSb* (%)
BG0E0(untreated)	0	0
BG2E0	2	0
BG4E0	4	0
BG6E0	6	0
BG0E0.3	0	0.3
BG0E0.9	0	0.9
BG0E2	0	2
BG2E0.9	2	0.9
BG4E0.9	4	0.9
BG6E0.9	6	0.9

3.2 Atterberg limits

In order to determine the percentage of the stabilizing agents, a common range is selected that varies between 0–6% and 0–2% of the dry weight of the soil sample for *WGP* and *EPSb*, respectively [3–6, 9–11, 13, 14, 39].

In the first part of the experiments, only *WGP* is used as a chemical stabilizer where the fat clay samples are mixed with it at a ratio of 2, 4, and 6% of the dry soil weight. As the percentage of *WGP* is increased, both liquid limits (*LL*) and plastic limits (*PL*) decline, resulting in a drop in the plasticity indexes (*PI*). Since glass is a non-plastic material, a decrease in the plasticity index could occur, which might be considered as an improvement. In the second part of the experiments, different *EPSb* percentages (0.3, 0.6, 0.9, and 2.0%) are used as additives. The result is that the *PL* values with increasing *EPSb* do not show a definite relation with the percentage values; however, it is observed that as the percentage of *EPSb* increases, the *PI* values climb slightly as well. The main consequence of this behavior might be the detachment of *EPSb* (especially *EPSb* = 2%) from the samples during the rolling phase of the *PL* tests (Fig. 2). The test results obtained in [40] are presented in Table 6 and Fig. 3.

3.3 Standard proctor tests

The *OMC* and *MDD* values are obtained for untreated and treated bentonite specimens having various percentages of *WGP* and/or *EPSb* according to the ASTM D698-12, Method A [29]. The standard proctor, *FS,* and *UCS* tests are carried out simultaneously to determine the optimum

Fig. 2 Examples of detached *EPSb* during plastic limit tests

Table 6 Atterberg Limits of the treated and untreated bentonite [40]

Sample Name	Liquid Limit (%)	Plastic Limit (%)	Plasticity Index (%)
BG0E0(untreated)	161.7	47.6	114.1
BG2E0	157.2	45.3	111.9
BG4E0	153.5	44.4	109.1
BG6E0	148.1	43.8	104.3
BG0E0.3	154.2	44.5	109.7
BG0E0.9	159.3	45.1	114.2
BG0E2.0	159.4	43.4	116.0
BG2E0.9	157.0	44.8	112.2
BG4E0.9	152.6	45.9	106.7
BG6E0.9	146.9	43.5	103.4

combination of two wastes, and it is decided to investigate the use of *WGP and EPSb* percentages as 2, 4, and 6% and 0.3, 0.9, and 2%, respectively.

In view of the results, *OMC* shows a modest decrease by the addition of *WGP*, whereas *MDD* shows a slight increase as depicted in Table 7 and Fig. 4. The addition of *WGP* slightly reduces *OMC,* which stays unchanged even upon replacing the fines in the sample with coarse particles. This latter action, in turn, reduces the surface area. As a result, less water is required for compaction; what is more the inclusion of *WGP* increases *MDD* which could be attributed to a chemical reaction occurring in *WGP*-treated samples [13]. These findings are consistent with those of Canakci et al. [9], Fauzi et al. [10], Bilgen [11], Mujtaba, et al. [14], and Olufowobi et al. [41]. As in the studies of Rocco [5], Soundara and Selvakumar [6], and Silveria et al. [42], the OMC values decline slightly with the rise in *EPS* and remain constant though a decline is observed in the *MDD* values. The low density and limited moisture absorption of *EPSb* may explain this fact [42].

3.4 Free swell

The one-dimensional swell, or free swell (*FS*), of the treated and untreated samples is determined according to the ASTM D4546-14, Method A [34] since the tests are conducted on reconstructed samples. In the previous studies, it is shown that using glass powder as a stabilizing agent has a positive effect on unconfined compressive strength values and swelling potential, whereas the addition of *EPSb* reduces the free swell of the soil samples [3, 4, 11, 13, 14]. Due to this reason, firstly, the optimum percentage of *EPSb* is investigated by treating bentonite with only *EPSb,* yielding a value of 0.9. Then, the *FS* of the bentonite is examined for both *WGP* and *WGP-EPSb* treatment. The *FS* test results are shown in Table 8 and Fig. 5.

In view of these results, it can be concluded that the reduction in the *FS* percentage indicates an approximate 27, 20, and 26% of reduction with the usage of 0.9% *EPSb*, 6% *WGP,* and $0.9\%EPSb + 4\%WGP$ contents, respectively. Soil treated with *WGP* exhibits less swell, which could be attributed to the change in the water absorption of the mixture. The swelling of clay is generally influenced by

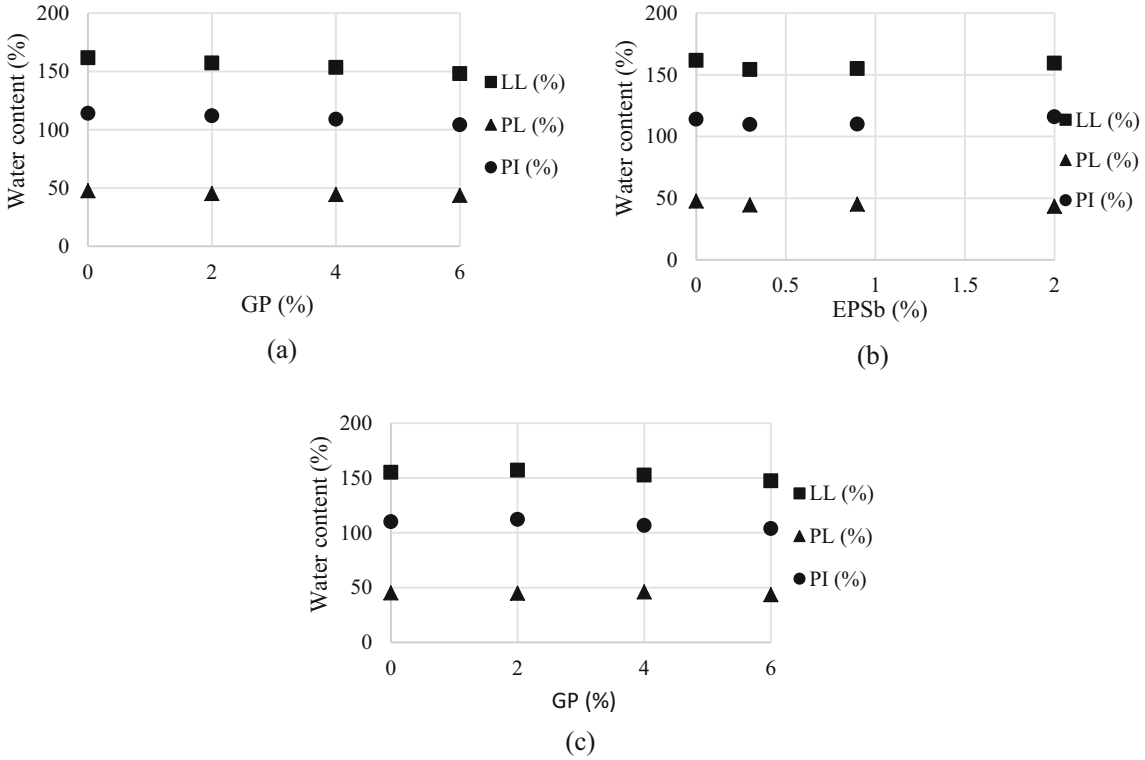


Fig. 3 Atterberg Limits of treated and untreated fat clay with (**a**) *WGP* only, (**b**) *EPSb* only, and (**c**) with *WGP* and 0.9% *EPSb*

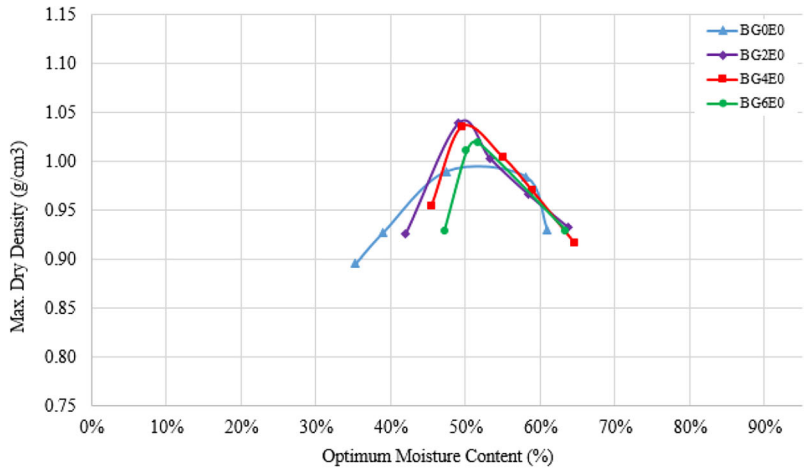
Table 7 *OMC* and *MDD* values obtained from standard proctor test

Soil Sample	*OMC* (%)	*MDD* (g/cm^3)
BG0E0(untreated)	52.00	0.994
BG2E0	50.20	0.990
BG4E0	51.22	1.003
BG6E0	51.42	1.012
BG0E0.3	52.00	0.914
BG0E0.9	51.95	0.864
BG0E2	48.00	0.745
BG2E0.9	48.30	0.842
BG4E0.9	50.60	0.843
BG6E0.9	50.80	0.846

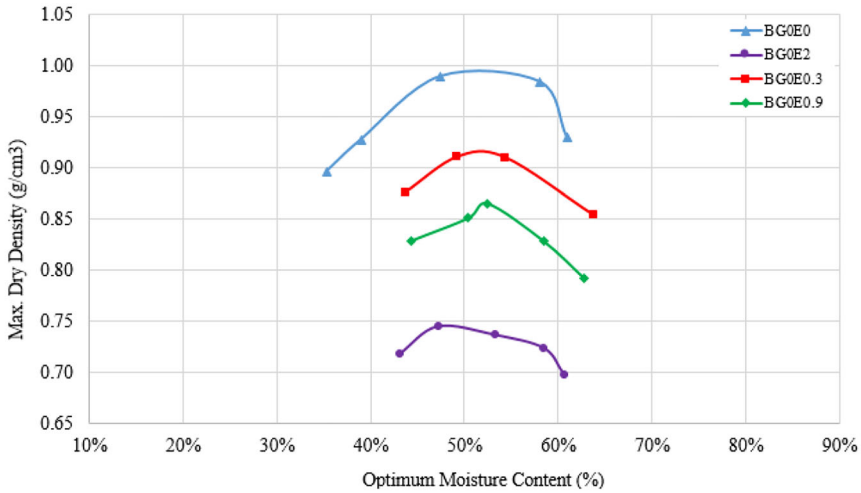
the amount of water in the soil and the minerals in the clay that try to absorb water. When a granular material, such as *WGP*, is used as a stabilizing agent for cohesive soils, the addition of siliceous *WGP* may lead to tighter packing of clay particles, resulting in a drop in swell potential [14]. It can also be concluded that the amount of *FS* declines as the ratio of *EPSb* increases. When *EPSb* is used for treatment, it serves as an inclusion or barrier to the passage of water into the soil, thereby reducing the swelling potential [4, 6]. It is also observed that, if the bentonite is treated with high *EPS* (2%) or 6%*WGP* + 0.9%*EPSb* contents, the stabilizing agents may not act as a barrier and, as such, may result in a higher swell potential. Finally, it can be concluded that the optimum additive percentages to treat the bentonite used in this study are 4% and 0.9% for *WGP* and *EPSb*, respectively.

3.5 Unconfined compressive strength tests

Unconfined compressive strength (*UCS*) tests on the 0-day cured soil samples are performed according to ASTM D2166-16 [43]. The photographs taken before and after the tests and the test results are given in Figs. 6, 7, and Table 9. All tests are duplicated to ensure that identical samples are tested consistently. From the test results, it can be seen that the *UCS* increases with the addition of *WGP* and decreases with the addition of *EPSb*. These results are in agreement with the previous studies conducted on treated clays [3, 11, 13, 14]. Since the cohesion between *EPSb* and clay particles is less than that of clay–clay particles, the cohesion of the mixture decreases [3]. Furthermore, the increase in strength with the addition of *WGP* could be explained by the pozzolanic reaction. The value of *UCS* increases as more *WGP* is added until reaching a peak strength value [13]. With further increase in WGP, a decline in *UCS* is observed in the treated samples. This finding is in line with the previous studies conducted by İbrahim et al. [13],



(a)



(b)

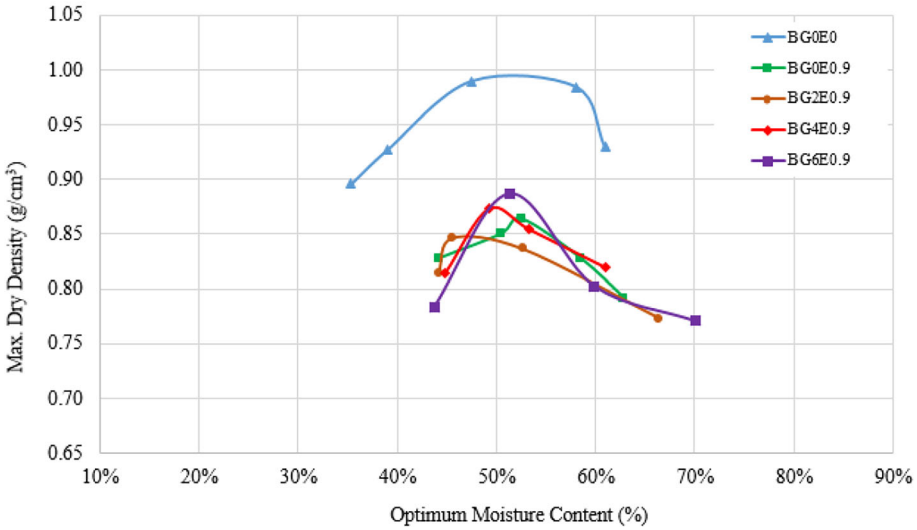


Fig. 4 Compaction curves for treated and untreated bentonite

Table 8 One-dimensional swell test results

Soil Sample	Free swell (%)
BG0E0(untreated)	30.41
BG0E0.3	29.63
BG0E0.9	22.08
BG0E2	23.89
BG2E0	31.31
BG4E0	25.25
BG6E0	24.46
BG2E0.9	24.02
BG4E0.9	22.63
BG6E0.9	23.63

Mujtaba et al. [14] on fat clays, and Canakci et al. [9] on lean clay. Such a declining trend trend could be attributed to a drop in the adhesive strength between waste glass surface and clay soil [9, 13]. Hence, it is noted that the percentage of *WGP* should be limited to 4% in view of the test results obtained in this study.

4 Predictive models for *FS* and *UCS*

The mechanical characteristics of the specimens (*FS* and *UCS)* are predicted using multiple linear regression (*MLR*) and artificial neural network (*ANN*) algorithms based on the test results obtained in this study and those reported in previous studies. In this part, the computational details are presented.

4.1 Regression analysis

Regression analysis is a statistical method for determining the relationship between one or more independent variables (predictors) and dependent (outcome) variable [44]. Multiple regression is a type of regression analysis that uses more than one independent variable to predict the dependent variable. The general equation is given in Eq. 1.

$$Y' = A + B_1X_1 + B_2X_2 + \ldots + B_nX_n \quad (1)$$

Here, Y' is the dependent variable, A is the constant value, X_s are the independent variables, and B_s are the coefficients assigned to each independent variable. The aim of regression analysis is to obtain the regression coefficients that bring the estimated Y' values as close to the actual Y values as possible. The existence of the statistically significant linear relationship between two variables (predictor and outcome) is controlled by the Pearson correlation coefficient (r) [44]. It is an indicator of a number between -1 and $+1$ where an r value of 0.0 shows no relationship and a value that is greater than 0.2 indicates dependency [45]. Additionally, the level of statistical significance of variables is presented using a p value. The relation between the variables can be defined as statistically significant when the values of p are less than 0.05 (i.e., confidence interval at a level of 95%). Furthermore, intercorrelations (especially multicollinearity state) between the independent variables should be evaluated. The variance inflation factor (*VIF*), an indicator of the effects of independent variables on the regression coefficient standard error, may be used in this evaluation. The *VIF* values greater than 10 show a high degree of collinearity (or multicollinearity) between these variables [46]. In this study, the relationships between the

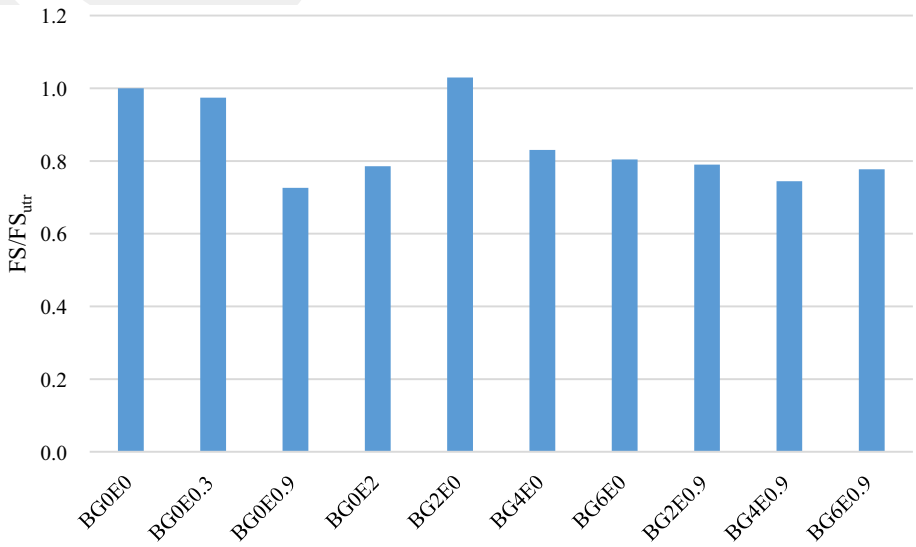


Fig. 5 Effects of *EPSb* and *WGP* on the *FS* values

Fig. 6 Clay sample with *EPSb* before and after *UCS* test

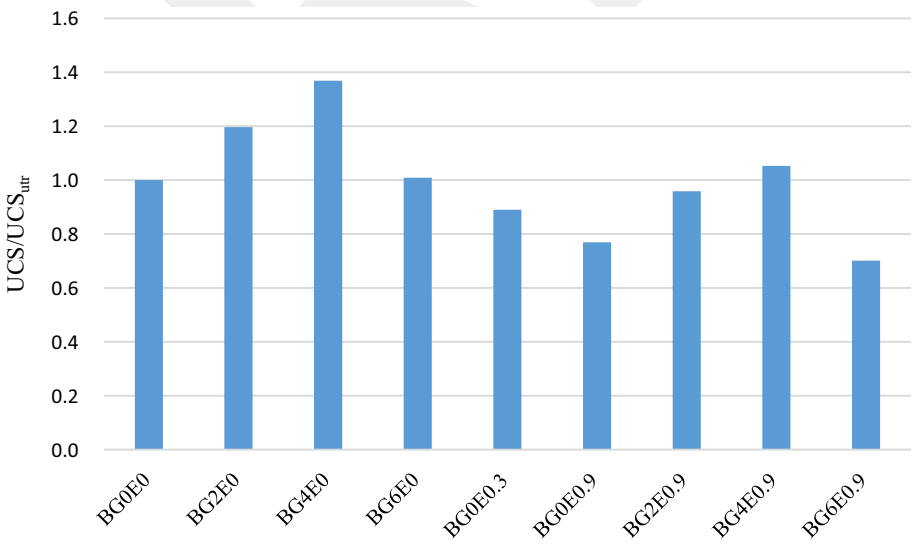


Fig. 7 Effects of *EPSb* and *WGP* on the *UCS* values

characteristics of untreated samples and the *FS* and *UCS* of the treated soil samples are studied using the *MLR* model.

In this study, the parameters that have the greatest impact on *FS* and *UCS* are identified through statistical analysis. The index properties of the soil, *EPSb,* and *WGP* contents in percentage and the *FS* or *UCS* of the untreated soil are defined as independent variables. A number of linear regression analyses are performed, and two datasets based on the test results of this study and the results of the relevant studies [3–6, 11, 13, 14] in the literature are used to predict *FS* and *UCS*. In the first dataset, the Atterberg limits and additives are defined as predictors, while the test results of untreated samples in addition to the first dataset are also considered as independent variables in the second dataset. The training datasets consist of 36 and 32 samples in the determination of the *FS* and *UCS*, respectively. The effect of predictors on *FS* and *UCS* is shown in Table 10. In this table, *EPSbp* and *WGPp* represent the percentages

Table 9 Strength parameters of treated and untreated bentonite

Soil Sample	UCS (kPa) 0-day
BG0E0(untreated)	337.4
BG2E0	403.8
BG4E0	461.9
BG6E0	340.4
BG0E0.3	300.3
BG0E0.9	259.6
BG2E0.9	323.5
BG4E0.9	355.3
BG6E0.9	236.7

of *EPSb* and *WGP,* respectively. Depending on the *r* values, it can be concluded that tests performed on the untreated samples (FS_{utr} and UCS_{utr}) are statistically significant for the prediction of *FS* and *UCS*. Then, *MLR* analyses are carried out. The results of these analyses with high R^2 values are presented in Table 11. The prediction equations are considered as functions of Atterberg limits as well as other independent variables such FS_{utr} and UCS_{utr}. As a result of the assessments, it is concluded that untreated test results significantly affect the *FS* and *UCS* of the treated soils. Furthermore, a modest increase in the R^2 values is observed if only the Atterberg limits are used as predictors. The proposed empirical correlations and the R^2 values for training and testing datasets obtained by linear regression analyses are given in Table 12.

4.2 The *ANN* model

Artificial neural network *(ANN)* is a computing method that is built on artificially simulating the operational principles of the human brain. An *ANN* could be trained to learn a pattern automatically using input and output data [47]. Artificial multilayer perception (*MLP*) neural networks are the most commonly utilized *ANN* model and consist of input, multiple hidden, and output layers [48]. The activation function determines a neuron's output and the output signal (o_j), which is the outcome of the activation function and can be determined by Eq. 2. In this study, the Sigmoid function is used as the activation function. The input to this function is net_j, which is the total of bias (b_j) and multiplications of scalar inputs $(i_1, i_2, \ldots, i_n)$ with weights $(w_{1j}, w_{2j}, \ldots, w_{nj})$.

$$o_j = f\left(net_j\right) = f(i_1 w_{1j} + i_2 w_{2j} + \ldots + i_n w_{nj} + b_j) \tag{2}$$

ANN iteratively adjusts the weights of neurons during the training process to reduce prediction errors. The input parameters are processed via hidden and output layers, and the output is estimated. Next, the estimated and real experimental values are compared with each other to calculate errors by using the root mean square of errors (*RMSE*) given in Eq. 3.

$$RMSE = \frac{\sqrt{\sum_d \sum_i |t_{ip} - o_{ip}|}}{N} \tag{3}$$

Here, t_{ip} is the real experimental value and o_{ip} is the model estimated output over all the data (d). Then, since

Table 10 Effect of predictors on *FS* and *UCS*

Dependent Variable	Data set	Predictor	$\bar{X}$	S	N	r
FS	1	*EPSbp*	0.375	0.505	36	0.081
		WGPp	3.444	7.101	36	-0.361
		PL	30.807	9.891	36	0.039
		LL	103.176	76.55	36	0.052
		PI	72.369	70.879	36	0.050
	2	FS_{utr}	23.829	19.784	36	0.953
UCS	1	*EPSbp*	0.096	0.268	32	0.017
		WGPp	9.063	10.398	32	0.427
		PL	29.463	12.256	32	0.149
		LL	85.513	65.055	32	0.142
		PI	56.051	54.473	32	0.136
	2	UCS_{utr}	232.119	115.822	32	0.841

Dependent: *FS, UCS*

$\bar{X}$: Mean

S: Standard deviation

N: Number of samples

r: Pearson correlation coefficient

Table 11 Model summary of *FS* and *UCS* based on the training data

Data set	Predictor	R^2	p	*VIF*
FS	*EPSbp, WGPp, FS$_{utr}$*	0.936	0.000	1.133
				1.220
				1.167
	EPSbp, WGPp, FS$_{utr}$, PL	0.942	0.000	1.192
				1.228
				1.176
				1.073
	EPSbp, WGPp, FS$_{utr}$, PL, LL	0.942	0.000	1.201
				1.266
				1.179
				1.643
				1.697
UCS	*EPSbp, WGPp, UCS$_{utr}$*	0.926	0.000	1.188
				1.065
				1.120
	EPSbp, WGPp, UCS$_{utr}$, PL	0.931	0.000	1.487
				1.076
				1.150
				1.429
	EPSbp, WGPp, UCS$_{utr}$, PL, LL	0.931	0.000	1.496
				1.117
				1.160
				5.852
				4.995

forward and backward processes are part of the training, the error is fed back into the network and the weights and biases of the neurons are adjusted as a consequence. The effect of the error on these variables is calculated.

Equations 4 and 5 are used to update the weights, w_{ij}, and biases, b_i^l [49].

$$w_{i,j}^l = w_{i,j}^l - \alpha \frac{\partial E(w,b)}{\partial w_{i,j}^l} \tag{4}$$

$$b_i^l = b_i^l - \alpha \frac{\partial E(w,b)}{\partial b_i^l} \tag{5}$$

In the above equations, *E(w, b)* is the error value due to weight, w, and bias, b, and α is the learning rate. This process is performed numerous times in order to minimize the error and improve the network accuracy, and it is referred to as the iteration number in the training process [50]. In this model, the Levenberg–Marquardt (*LM*) learning algorithm is utilized which is a type of supervised learning algorithm with different learning rates and without bias terms. The *LM* algorithm modifies the weights to reduce the errors and obtain the closest results to the experimental outcomes. The maximum and minimum values of each input dataset are used to normalize the data by using Eq. 6.

$$x_{in} = \chi + \beta \frac{x_i - x_{min_i}}{x_{max_i} - x_{min_i}} \tag{6}$$

In this equation, the i^{th} input data are x_{in}, and the maximum and minimum values of that input data are x_{max_i} and x_{min_i}, respectively. The constants χ and β are taken as 0.1 and 0.8, respectively, to prevent zero values during the normalization. To achieve the lowest error, several learning rates, iterations, layers, and neurons are tested. The coefficient of correlation (R^2), mean squared errors (*MSE*), and mean absolute percentage error (*MAPE*) is used to evaluate the performance of models by using Eqs. 7–9.

$$R^2 = 1 - \left(\frac{\sum_i |t_i - o_i|^2}{\sum_i (t_i - \bar{t})^2} \right) \tag{7}$$

$$MAPE = \frac{100}{p} \sum_i \left| \frac{t_i - o_i}{o_i} \right| \tag{8}$$

$$MSE = \frac{1}{p} \sum |t_i - o_i|^2 \tag{9}$$

In Eqs. 7–9, the target value achieved by the experiments is t_i, the model output is o_i, the mean value of experimental real data is $\bar{t}$, and the total amount of data available in the datasets is shown by p.

The satisfactory performance of *ANN* models makes them a popular and efficient tool for overcoming various engineering challenges. As stated earlier, the *FS* and *UCS* of treated soil samples are estimated based on index properties and test results of untreated soil samples using a previously developed *ANN* algorithm [15]. The flow chart of the algorithm is provided in Fig. 8.

4.2.1 Results of *ANN* model

Two different datasets that consist of 45 and 40 data are used to predict the *FS* and *UCS*, respectively. The datasets are generated based on the data obtained from both the literature and this experimental study. Subsequently, the data are randomly divided into training and testing datasets. The sources from which the randomly selected data are sourced and the numbers of the data points in both the testing and training datasets are as follows.

In the case of the *FS* dataset, comprising a total of 45 data points, ten data points are obtained from this experimental study, with eight designated for training and two for testing (Table 8). Additionally, six data points are obtained by İbrahim et al. [13], with five allocated for training and one for testing. Similarly, eight data points are sourced from Mujtaba et al. [14], with six for training and two for

Table 12 Prediction equations of the *FS*, *UCS*, and R^2 values for training and testing datasets

Prediction Equations	*Training* R^2	*Testing* R^2
$FS = 1.591 - 5.965EPSbp - 0.180WGPp + 0.844FS_{utr}$	0.936	0.981
$FS = -2.751 - 6.616EPSbp - 0.163WGPp + 0.850FS_{utr} + 0.142PL$	0.942	0.980
$FS = -2.661 - 6.65EPSbp - 0.158WGPp + 0.851FS_{utr} + 0.128PL + 0.003LL$	0.942	0.980
$UCS = -58.46 - 102.948EPSbp + 5.889WGPp + 1.203UCS_{utr}$	0.926	0.928
$UCS = -84.175 - 124.618EPSbp + 5.996WGPp + 1.187UCS_{utr} + 1.035PL$	0.931	0.931
$UCS = -82.743 - 124.371EPSbp + 5.982WGPp + 1.188UCS_{utr} + 0.916PL + 0.024LL$	0.931	0.930

FS_{utr}: the free swell (%)

UCS_{utr}: the unconfined compressive strength (kPa)

LL: liquid limit (%)

PL: plastic limit (%)

WGPp: glass powder used as additive (%)

EPSbp: *EPS* beads used as additive (%)

testing. Five data points are collected from Soundara and Selvakumar [6], with four for training and one for testing, while four data points are gathered from Rocco [5], with three for training and one for testing. Finally, twelve data points are obtained from İlluri [4], with ten for training and two for testing.

For predicting unconfined compressive strength (*UCS*), a dataset of 40 data points is employed, with subsets selected based on the following studies:

From the present experimental study, nine datasets are obtained, with seven for training and two for testing (Table 9). Additionally, twelve datasets are taken from Bilgen [11], with ten for training and two for testing, whereas six datasets are provided by İbrahim et al. [13], with four for training and two for testing. Eight datasets are collected from Mujtaba et al. [14], with seven for training and one for testing. Lastly, five datasets are supplied by Shirazi et al. [3], with four for training and one for testing.

For training the *ANN* model, 36 and 32 data out of 45 and 40—that is, 80% of the datasets—are selected for training, and the remaining 20% of the data is chosen for testing purposes. The *FS* and *UCS* of the treated soils are estimated using three different combinations of inputs:

- Three inputs including *WGPp* (waste glass powder percentage), *EPSbp* (waste *EPSb* percentage), and test results performed on untreated samples (FS_{utr} or UCS_{utr}).
- Five inputs including *WGPp* (waste glass powder percentage), *EPSbp* (waste *EPSb* percentage), *LL* (liquid limit), *PL* (plastic limit), and *PI* (plasticity index) of untreated samples.
- Six inputs including *WGPp* (waste glass powder percentage), *EPSbp* (waste *EPSb* percentage), *LL* (liquid limit), *PL* (plastic limit), *PI* (plasticity index), and test results performed on untreated samples (FS_{utr} or UCS_{utr}).

Various combinations of input parameters, such as *LL* (liquid limit), *PL* (plastic limit), *PI* (plasticity index), test results conducted on untreated samples (FS_{utr} or UCS_{utr}), and the proportions of waste glass powder (*WGPp*) and waste *EPS* beads (*EPSbp*), are considered as model input parameters. Meanwhile, the model outputs, represented by *F*S and *UCS* of the treated soils, are illustrated in Fig. 9.

It should be noted that the *LL*, *PL*, and *PI* of the soil can be determined quickly and economically, making sample supply more easily since they are the tests performed on disturbed samples. Whereas, to determine FS_{utr} or UCS_{utr}, undisturbed samples should be collected from the site with special equipment and, for this reason, the duration of these tests may be relatively long depending on the type of clay (e.g., *FS* tests duration for the bentonite used in this study is approximately 15 days).

Firstly, numerous analyses are carried out to determine the best network structure for various input combinations based on statistical performance metrics ($R^2, MSE, and MAPE$). The best results are obtained from runs with a learning rate of 0.2 for 3 hidden layers with 5, 10, and 5 neurons. The 35,000 iterations designated for the *UCS* analysis involving six variables and 50,000 iterations allocated for all other analyses. The process times, testing and training the R^2 values mean squared errors (*MSE*), and mean absolute percentage error (*MAPE* with selected learning rates) for different iterations (ranging from 5000 to 50,000) are presented in Tables 13 and 14 for *FS* and *UCS*, respectively.

Figures 10, 11, 12, 13, 14, 15, 16, 17, 18,19, 20, and 21 and Tables 15, 16, and 17 show the results of the *ANN* models with three, five, and six input variables for training

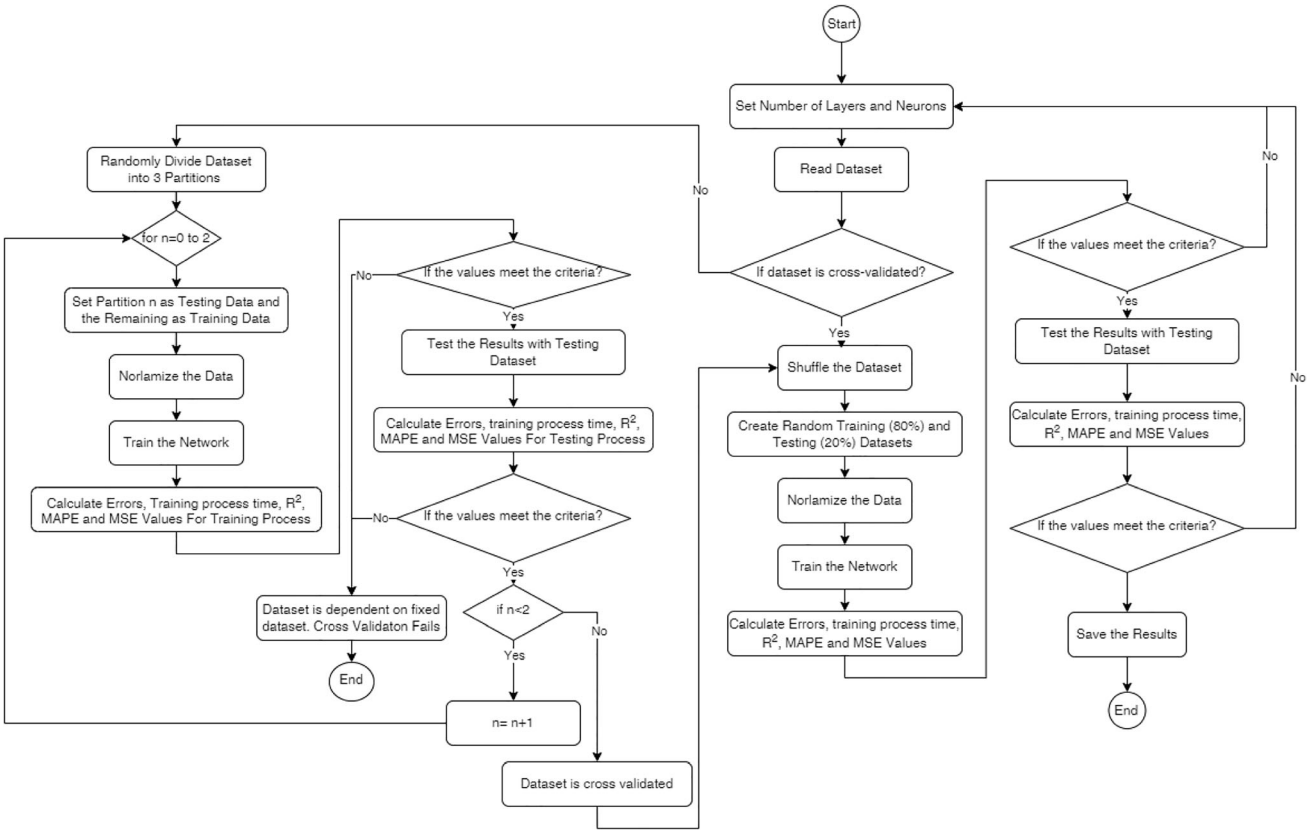


Fig. 8 Flow chart of the *ANN* algorithm

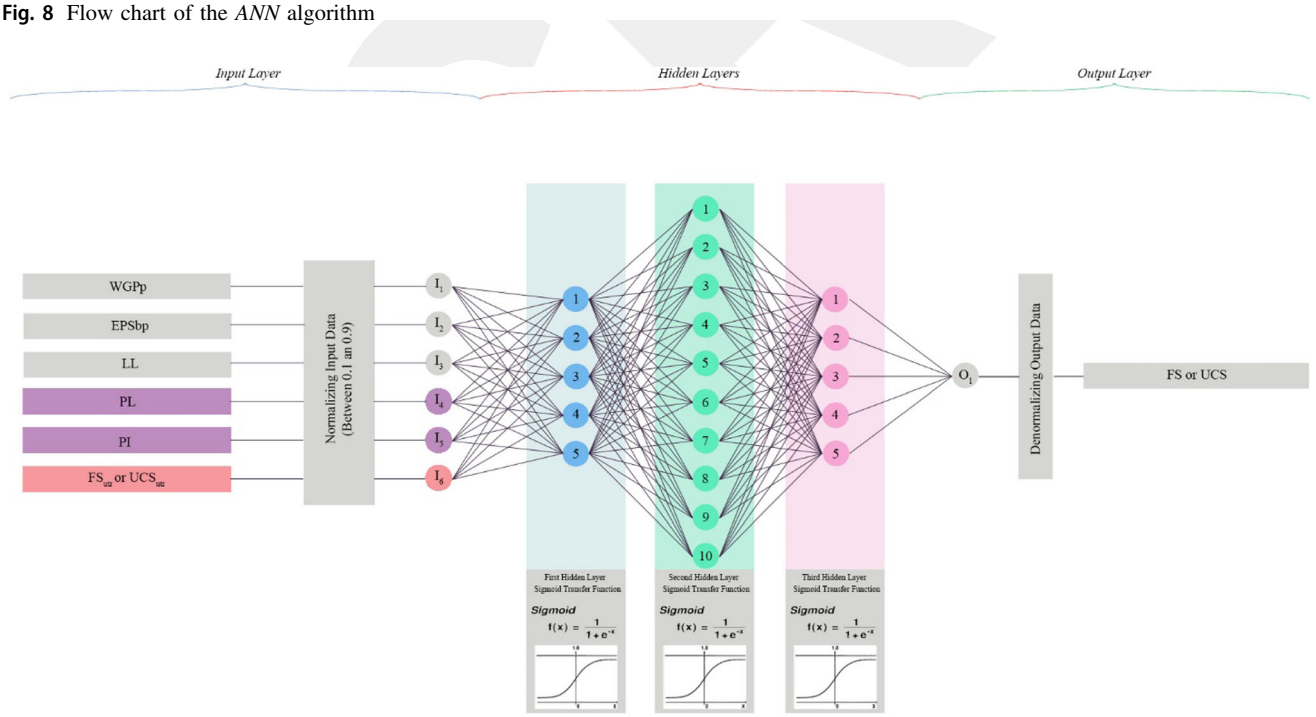


Fig. 9 *ANN* structure used in this research

and testing processes. The real experimental data and estimated values for *FS* and *UCS* are shown in Figs. 10a, 11, 12, 13, 14, and 15a and 10d, 11, 12, 13, 14, and 15d for the training process, respectively. Besides, the residual plots (Figs. 10b, 11, 12, 13, 14, and 15b and 10d, 11, 12, 13, 14, and 15d) indicate that the residuals are distributed

Table 13 Comparison between the *ANN* models of *FS* with six, five, and three inputs with different iteration numbers

Test Name	Neurons	Iteration number	Training *MSE*	Training *MAPE*	Training R^2	Elapsed time(s)	Testing *MSE*	Testing *MAPE*	Testing R^2
FS_6V	[6, 5, 10, 5, 1]	5000	10.22	10.51	0.989	0.91	4.18	67.44	0.992
	[6, 5, 10, 5, 1]	10,000	7.60	9.07	0.991	1.72	3.66	58.44	0.993
	[6, 5, 10, 5, 1]	15,000	5.71	7.86	0.992	2.48	3.64	53.43	0.994
	[6, 5, 10, 5, 1]	20,000	4.37	6.88	0.993	3.23	3.66	49.26	0.994
	[6, 5, 10, 5, 1]	25,000	3.30	5.97	0.993	3.95	3.74	45.38	0.995
	[6, 5, 10, 5, 1]	30,000	2.34	5.03	0.993	4.80	3.96	41.49	0.995
	[6, 5, 10, 5, 1]	35,000	1.34	3.81	0.994	6.31	4.47	37.26	0.995
	[6, 5, 10, 5, 1]	40,000	0.99	3.26	0.994	7.15	4.62	33.93	0.995
	[6, 5, 10, 5, 1]	45,000	0.91	3.13	0.994	7.87	4.57	31.03	0.995
	[6, 5, 10, 5, 1]	50,000	0.88	3.08	0.995	8.61	4.54	28.05	0.995
FS_5V	[5, 5, 10, 5, 1]	5000	15.07	12.76	0.909	0.91	31.74	54.25	0.987
	[5, 5, 10, 5, 1]	10,000	18.15	14.01	0.942	1.70	20.86	40.37	0.986
	[5, 5, 10, 5, 1]	15,000	14.44	12.49	0.952	2.44	17.39	37.12	0.984
	[5, 5, 10, 5, 1]	20,000	6.68	8.50	0.958	3.17	13.63	34.77	0.982
	[5, 5, 10, 5, 1]	25,000	4.21	6.75	0.961	3.89	11.97	34.79	0.981
	[5, 5, 10, 5, 1]	30,000	3.37	6.03	0.963	4.67	11.24	36.17	0.980
	[5, 5, 10, 5, 1]	35,000	3.05	5.74	0.964	5.40	10.83	38.16	0.979
	[5, 5, 10, 5, 1]	40,000	2.92	5.62	0.965	6.14	10.53	40.38	0.978
	[5, 5, 10, 5, 1]	45,000	2.88	5.58	0.967	6.87	10.26	42.54	0.978
	[5, 5, 10, 5, 1]	50,000	2.88	5.58	0.968	7.60	10.00	44.44	0.978
FS_3V	[3, 5, 10, 5, 1]	5000	0.21	1.50	0.973	0.89	4.65	73.21	0.991
	[3, 5, 10, 5, 1]	10,000	0.17	1.37	0.975	1.69	4.60	69.11	0.992
	[3, 5, 10, 5, 1]	15,000	0.09	0.97	0.977	2.42	4.50	64.11	0.994
	[3, 5, 10, 5, 1]	20,000	0.04	0.64	0.979	3.16	4.40	57.02	0.995
	[3, 5, 10, 5, 1]	25,000	0.09	0.96	0.981	3.87	4.59	48.30	0.997
	[3, 5, 10, 5, 1]	30,000	0.21	1.52	0.983	4.64	4.98	39.42	0.998
	[3, 5, 10, 5, 1]	35,000	0.29	1.77	0.983	5.37	5.28	33.57	0.998
	[3, 5, 10, 5, 1]	40,000	0.30	1.80	0.984	6.11	5.42	30.22	0.998
	[3, 5, 10, 5, 1]	45,000	0.29	1.78	0.984	6.84	5.47	28.18	0.998
	[3, 5, 10, 5, 1]	50,000	0.29	1.77	0.984	7.56	5.49	26.90	0.998

Table 14 Comparison between the *ANN* models of *UCS* with six, five, and three inputs with different iteration numbers

Test Name	Neurons	Iteration number	Training *MSE*	Training *MAPE*	Training R^2	Elapsed time(s)	Testing *MSE*	Testing *MAPE*	Testing R^2
UCS_6V	[6, 5, 10, 5, 1]	5000	25.57	1.50	0.959	0.98	672.28	8.59	0.949
	[6, 5, 10, 5, 1]	10,000	251.71	4.70	0.964	2.22	663.87	10.95	0.941
	[6, 5, 10, 5, 1]	15,000	433.44	6.17	0.967	2.91	599.05	10.39	0.943
	[6, 5, 10, 5, 1]	20,000	578.48	7.13	0.968	3.61	578.83	10.61	0.944
	[6, 5, 10, 5, 1]	25,000	531.48	6.83	0.970	4.31	582.56	11.44	0.942
	[6, 5, 10, 5, 1]	30,000	481.79	6.51	0.971	5.05	650.86	12.47	0.938
	[6, 5, 10, 5, 1]	35,000	453.29	6.31	0.972	5.74	711.01	13.03	0.934
	[6, 5, 10, 5, 1]	40,000	396.74	5.90	0.973	6.42	751.91	13.38	0.930
	[6, 5, 10, 5, 1]	45,000	334.05	5.42	0.975	7.14	805.85	14.02	0.925
	[6, 5, 10, 5, 1]	50,000	287.94	5.03	0.976	7.83	891.75	15.17	0.918
UCS_5V	[5, 5, 10, 5, 1]	5000	0.15	0.11	0.913	0.91	4317.05	52.32	0.821
	[5, 5, 10, 5, 1]	10,000	0.87	0.28	0.951	1.67	2970.45	48.24	0.874
	[5, 5, 10, 5, 1]	15,000	90.43	2.82	0.964	2.39	1558.39	32.38	0.911
	[5, 5, 10, 5, 1]	20,000	284.36	5.00	0.968	3.12	919.78	20.01	0.927
	[5, 5, 10, 5, 1]	25,000	444.55	6.25	0.970	3.84	730.33	13.50	0.934
	[5, 5, 10, 5, 1]	30,000	492.56	6.58	0.971	4.62	693.41	10.87	0.935
	[5, 5, 10, 5, 1]	35,000	468.42	6.41	0.972	5.36	697.12	9.62	0.934
	[5, 5, 10, 5, 1]	40,000	464.51	6.39	0.973	6.15	692.94	8.16	0.934
	[5, 5, 10, 5, 1]	45,000	478.19	6.48	0.973	6.92	686.94	7.07	0.935
	[5, 5, 10, 5, 1]	50,000	496.98	6.61	0.974	7.69	686.18	8.00	0.935
UCS_3V	[3, 5, 10, 5, 1]	5000	25.57	1.50	0.956	0.91	586.37	8.88	0.940
	[3, 5, 10, 5, 1]	10,000	238.07	4.57	0.959	1.71	762.95	11.77	0.937
	[3, 5, 10, 5, 1]	15,000	281.24	4.97	0.960	2.43	770.49	12.64	0.937
	[3, 5, 10, 5, 1]	20,000	315.06	5.26	0.961	3.13	779.42	13.20	0.936
	[3, 5, 10, 5, 1]	25,000	363.76	5.65	0.962	3.86	788.73	13.73	0.936
	[3, 5, 10, 5, 1]	30,000	695.83	7.82	0.963	4.63	1001.37	15.85	0.933
	[3, 5, 10, 5, 1]	35,000	783.83	8.30	0.964	5.33	1007.45	16.32	0.931
	[3, 5, 10, 5, 1]	40,000	859.90	8.69	0.964	6.05	1007.27	16.73	0.931
	[3, 5, 10, 5, 1]	45,000	927.21	9.02	0.965	6.77	1009.31	17.17	0.930
	[3, 5, 10, 5, 1]	50,000	985.37	9.30	0.965	7.46	1015.45	17.61	0.930

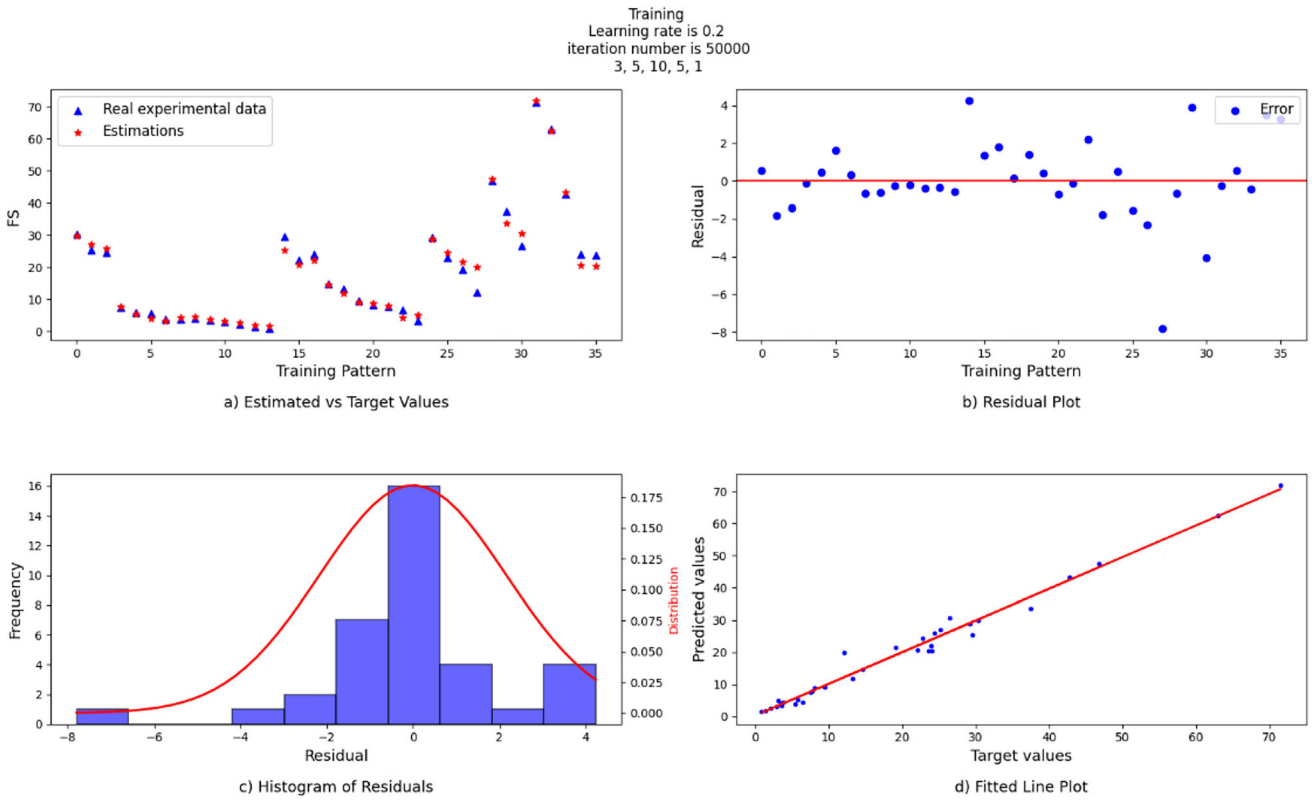


Fig. 10 Training results for the three-input *ANN* model of *FS* tests (Inputs: *WGPp, EPSbp, and* FS_{utr})

randomly and unbiasedly throughout the training pattern in all cases. The histogram of residuals for all *ANN* models is produced to check the variance distribution. These findings are shown in Figs. 10c, 11, 12, 13, 14, and 15c. It is seen that the inaccuracy of estimations for all models is regularly distributed and all the histograms have symmetric bell-shaped forms, which is indicative of an accurate normality assumption [51]. When the training R^2 values for models with three, five, and six input variables are assessed, it can be concluded that all models are successfully trained with $R^2 \geq 0.965$ (Tables 15, 16, 17).

Figures 16, 17, 18, 19, 20, and 21 show a comparison between the model estimation for *FS*, *UCS*, and the obtained values from the tests in the form of fitted plot lines for testing procedures. The R^2 values for the testing models of three, five, and six inputs are 0.998, 0.978, and 0.995 for *FS* and 0.930, 0.935, and 0.934 for *UCS*, respectively. These results can be regarded as very satisfactory.

4.2.2 K-fold cross-validation

The performance of data-based methods may be sensitive to variations in data. Therefore, it is necessary to engender variations in experiments by selecting multiple test datasets. In order to check if the *ANN* estimations are dependent or not on pre-defined training and testing datasets, a threefold cross-validation technique is used. The supplied dataset is stratified and partitioned into three equal folds for this purpose. Stratification is used to obtain folds that are good representatives of the whole data. The training and testing process is repeated three times, with one of the three folds chosen as the test data and the remaining two as training data for each run [52]. The performance metrics (*MSE*, *MAPE,* and R^2) for the training and testing processes are calculated for each run, and the model's predictive performance is calculated by averaging the performance values of three runs. The average performance metrics derived from threefold cross-validation are shown in Table 18. This procedure is done for each output (*FS* and *UCS*) as well as for each *ANN* model individually. As seen from the results, the changing training and testing data have no effect on the performance of the model, and the model is not dependent on a fixed dataset.

4.3 Results and discussion

The prediction of the *FS* and *UCS* of the treated samples is crucial in ground improvement applications. In order to predict the empirical equations, datasets are formed based on the test result of this study and the relevant studies in the literature. Considering the results of multiple linear regression analyses, it is found that all the *FSs* and *UCSs* of

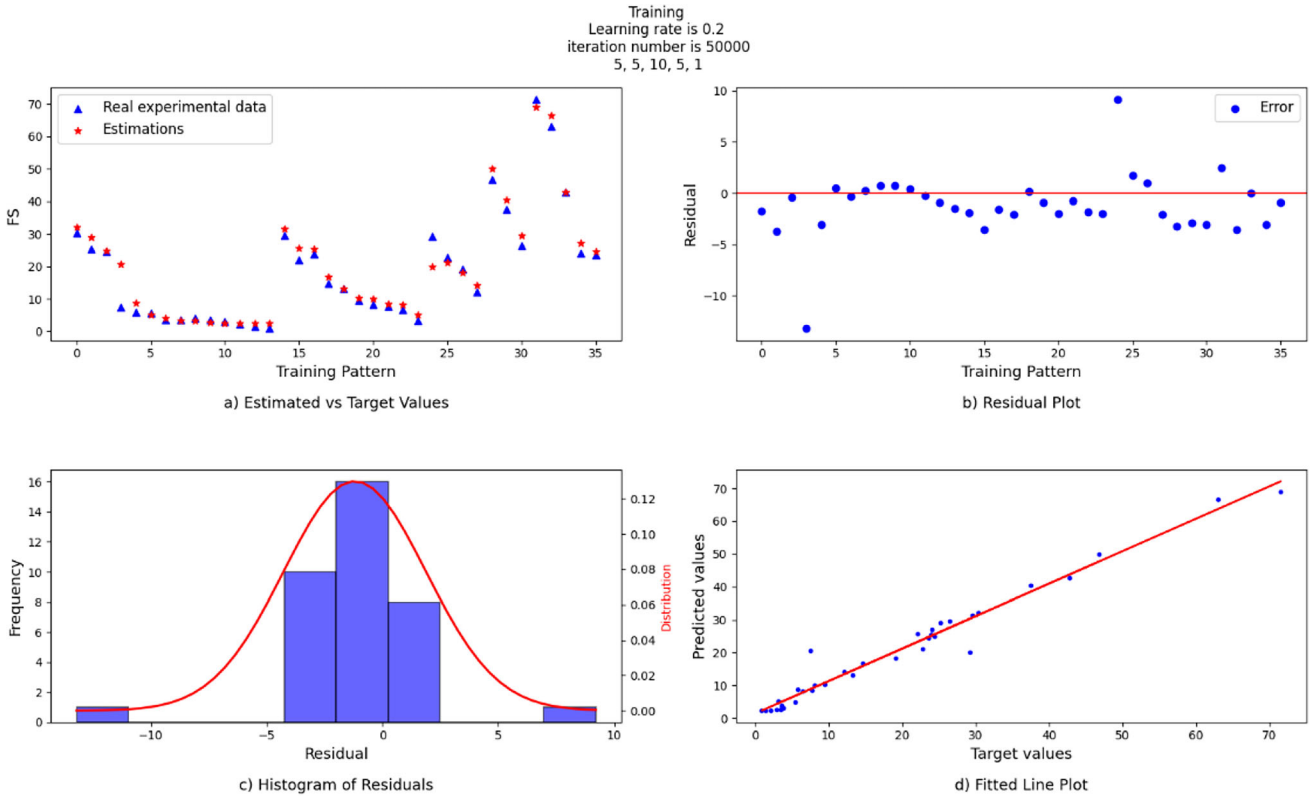


Fig. 11 Training results for the five-input *ANN* model of *FS* tests (Inputs: *WGPp, EPSbp, LL, PL, and PI*)

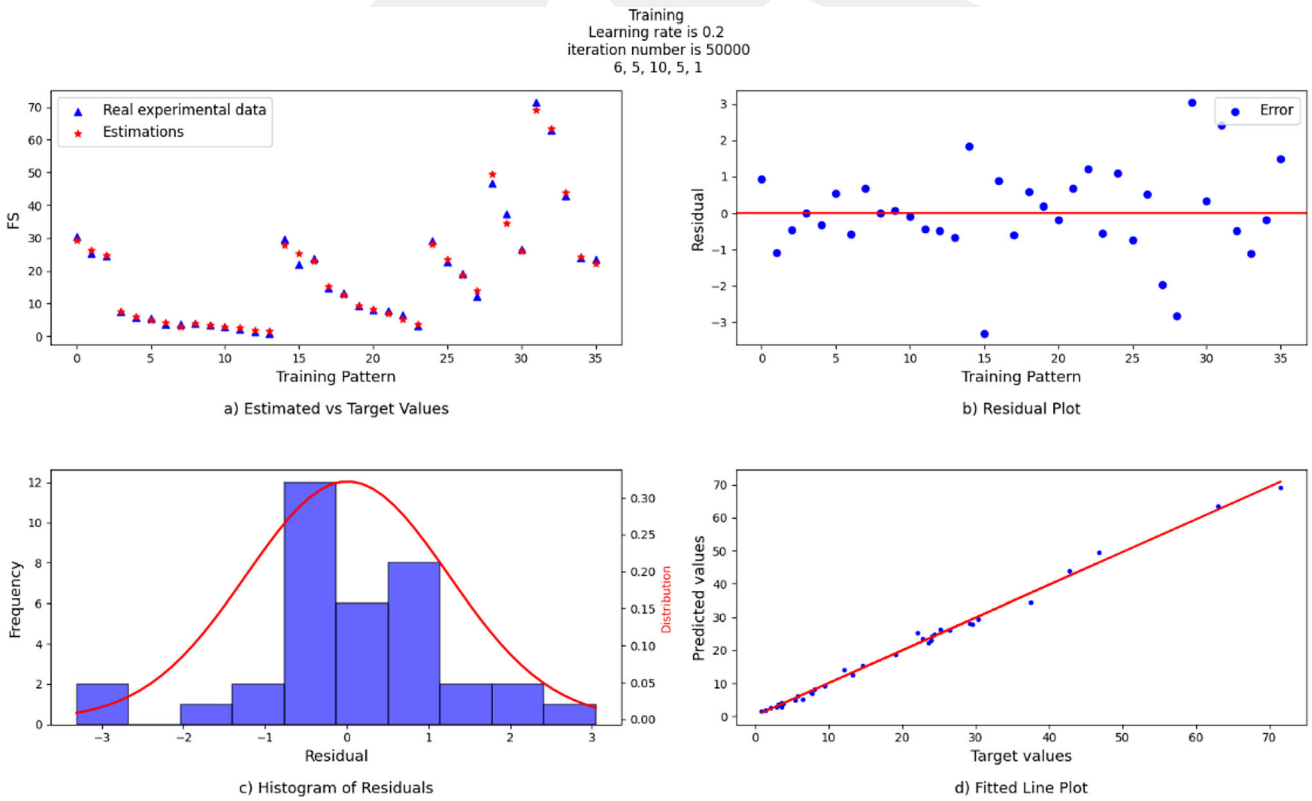


Fig. 12 Training results for the six-input *ANN* model of *FS* tests (Inputs: *WGPp, EPSbp, LL, PL, PI, and* FS_{utr})

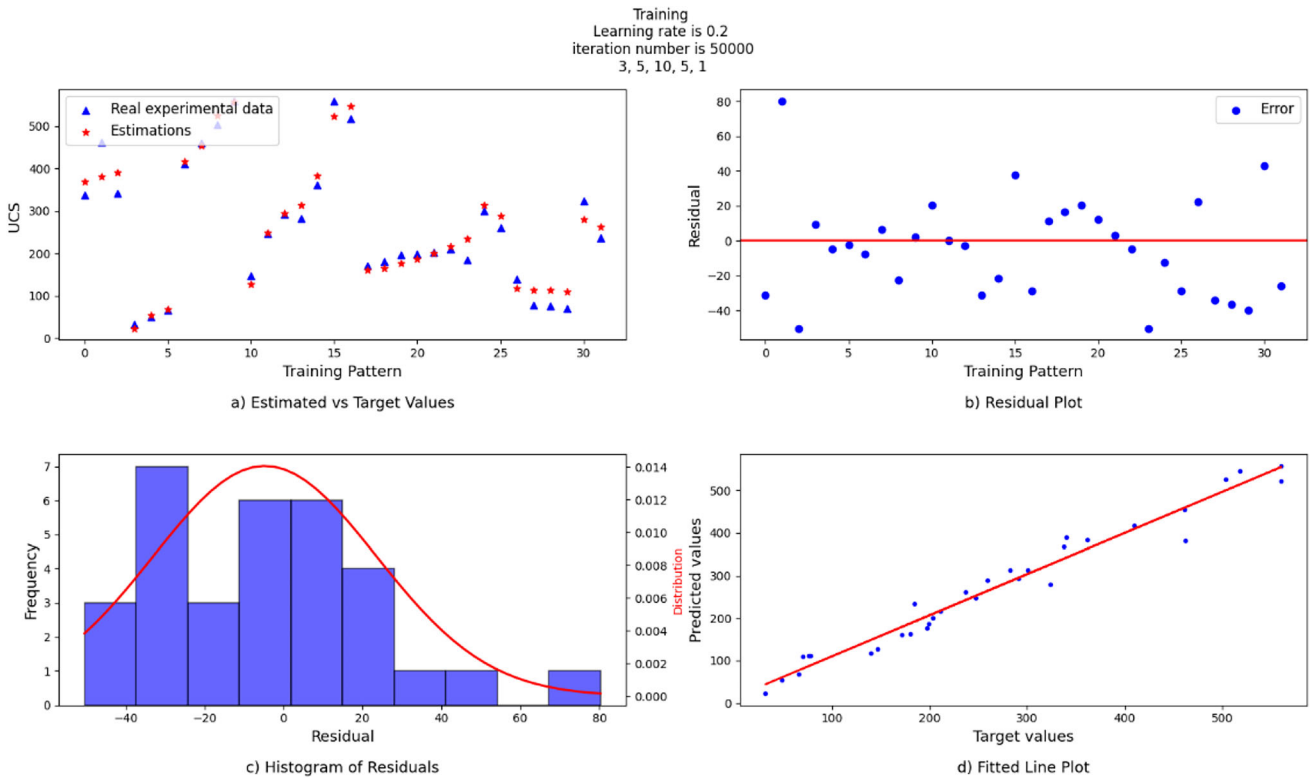


Fig. 13 Training results for the three-input *ANN* model of *UCS* tests (Inputs: *WGPp, EPSbp, and* UCS_{utr})

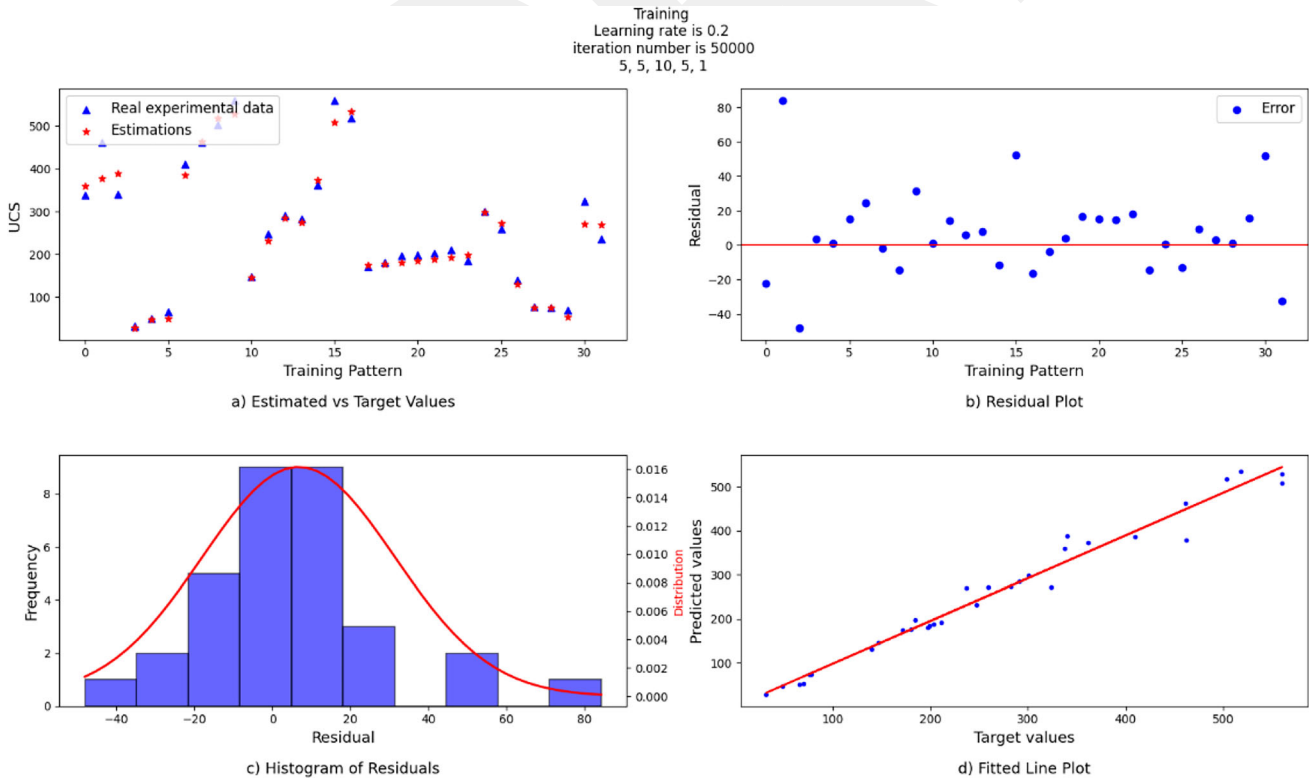


Fig. 14 Training results for the five-input *ANN* model of *UCS* tests (Inputs: *WGPp, EPSbp, LL, PL, and PI*)

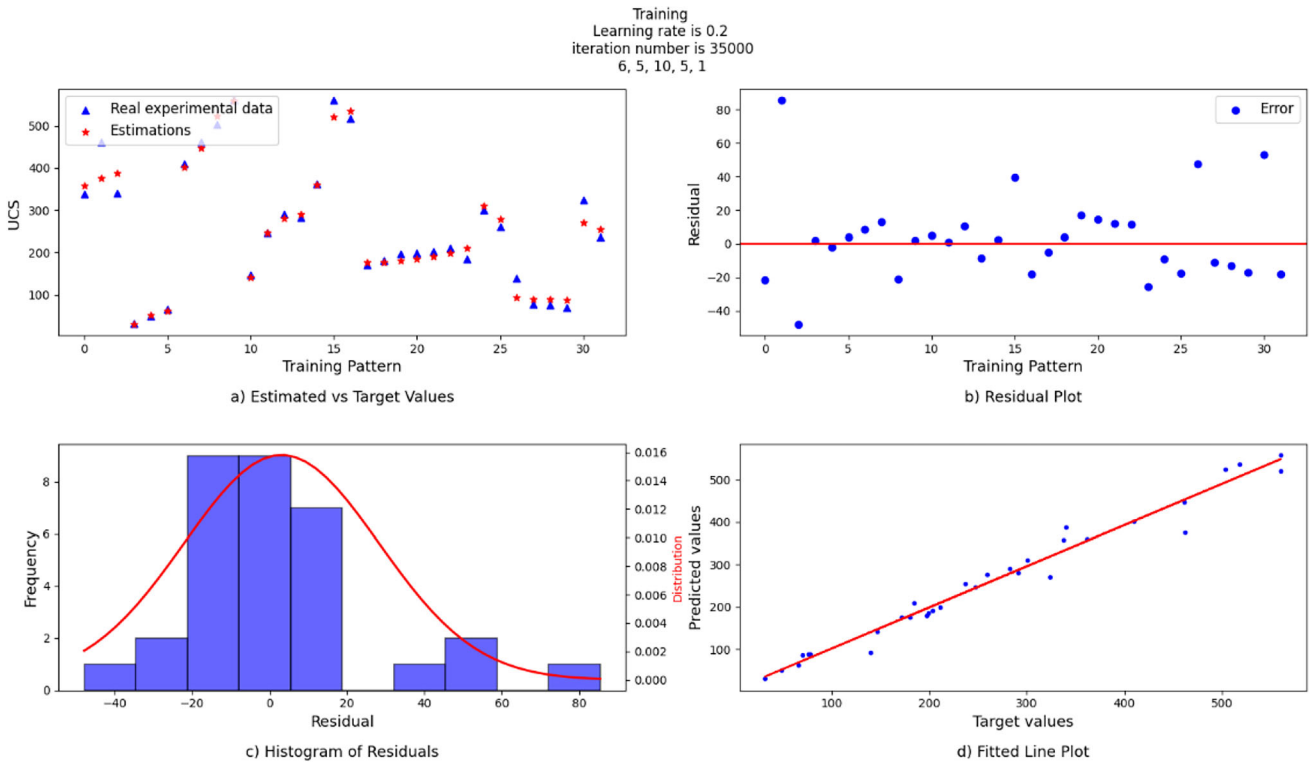


Fig. 15 Training results for the six-input *ANN* model of *UCS* tests (Inputs: *WGPp, EPSbp, LL, PL, PI, and* UCS_{utr})

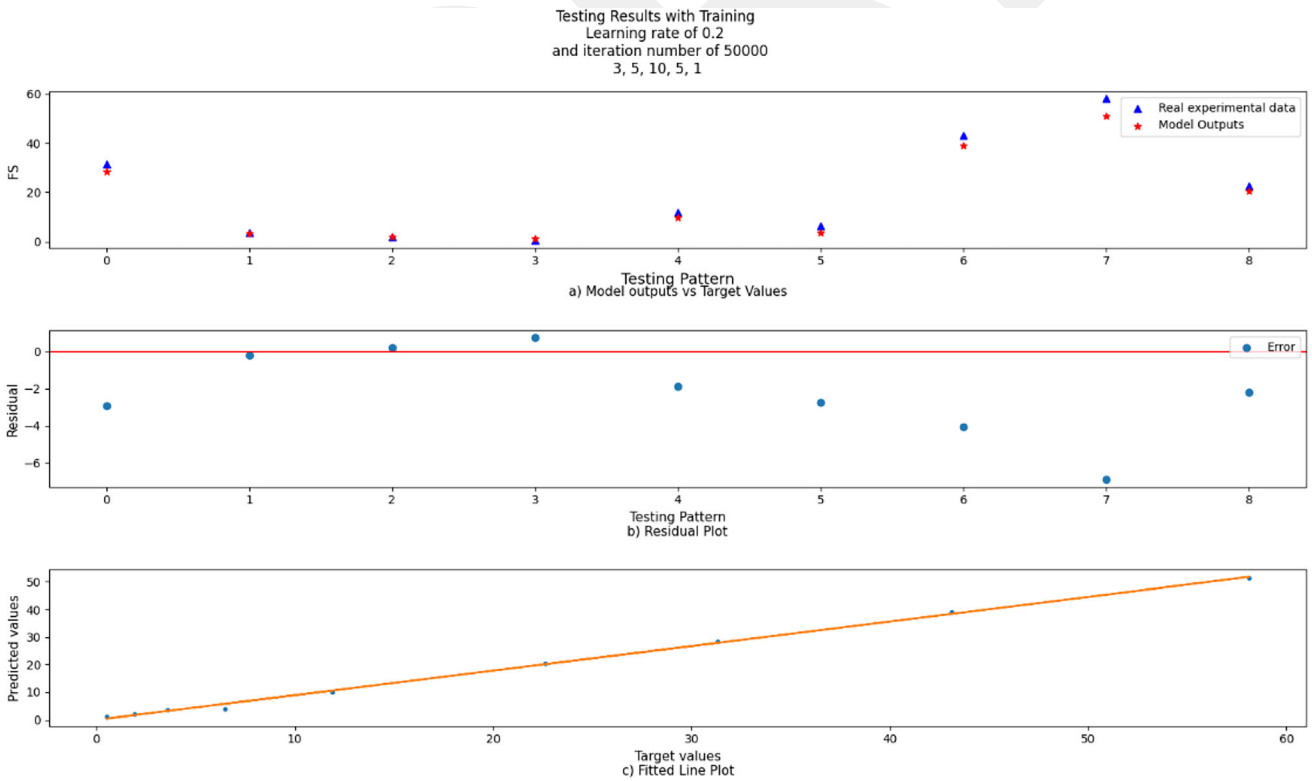


Fig. 16 Testing results for the three-input *ANN* model of *FS* tests (Inputs: *WGPp, EPSbp, and* FS_{utr})

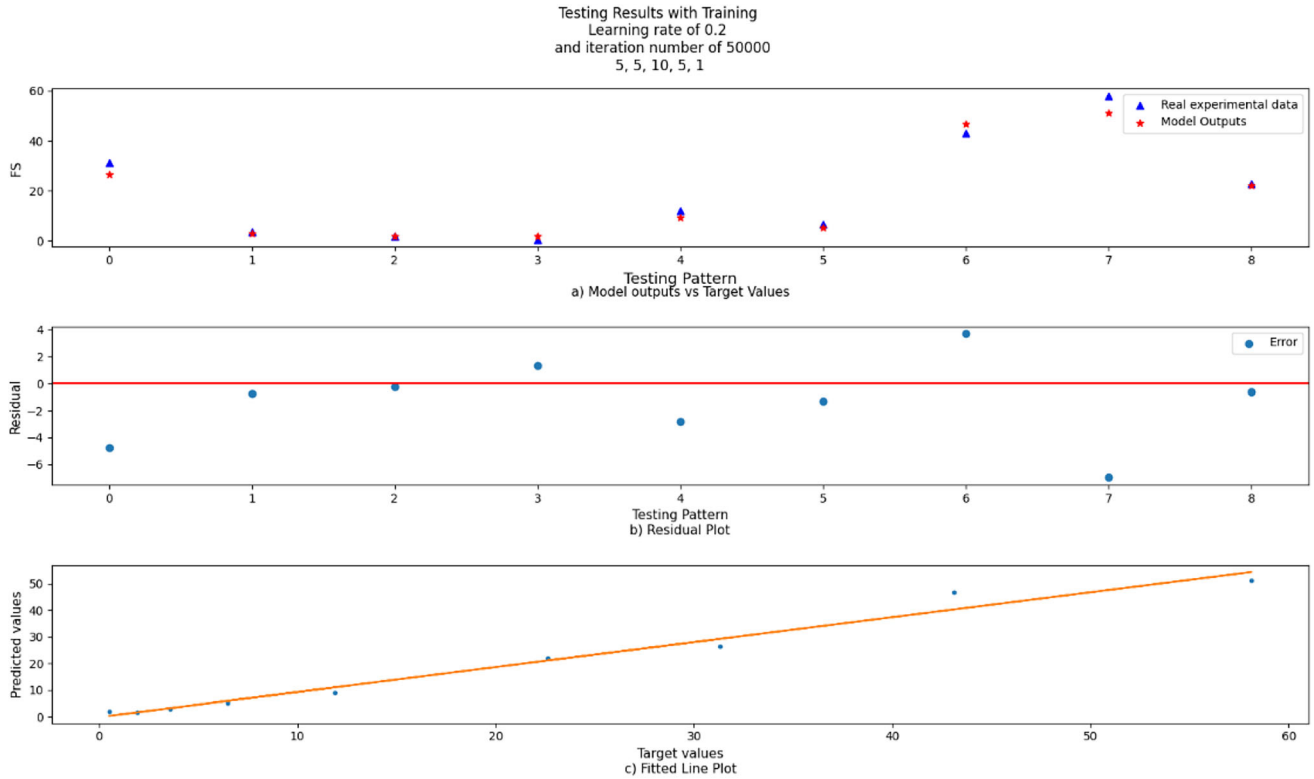


Fig. 17 Testing results for the five-input *ANN* model of *FS* tests (Inputs: *WGPp, EPSbp, LL, PL, and PI*)

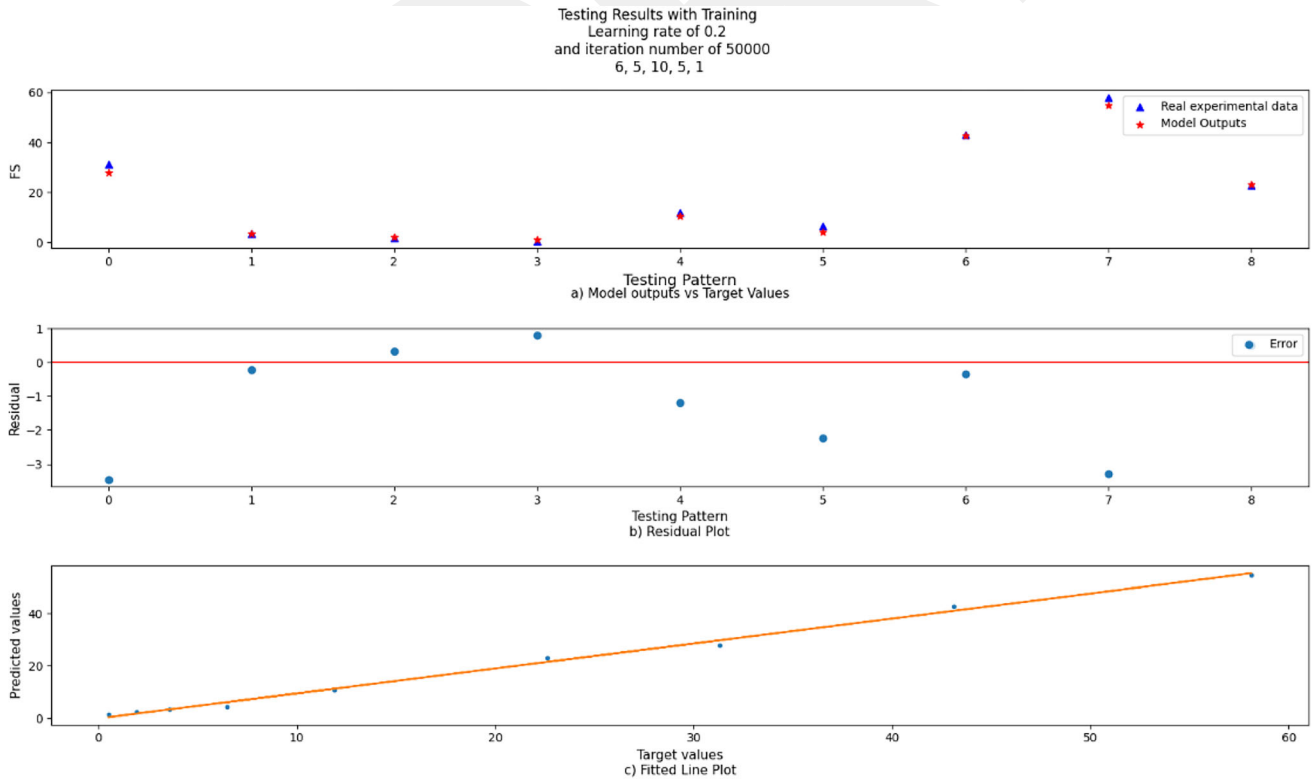


Fig. 18 Testing results for the six-input *ANN* model of *FS* tests (Inputs: *WGPp, EPSbp, LL, PL, PI, and* FS_{utr})

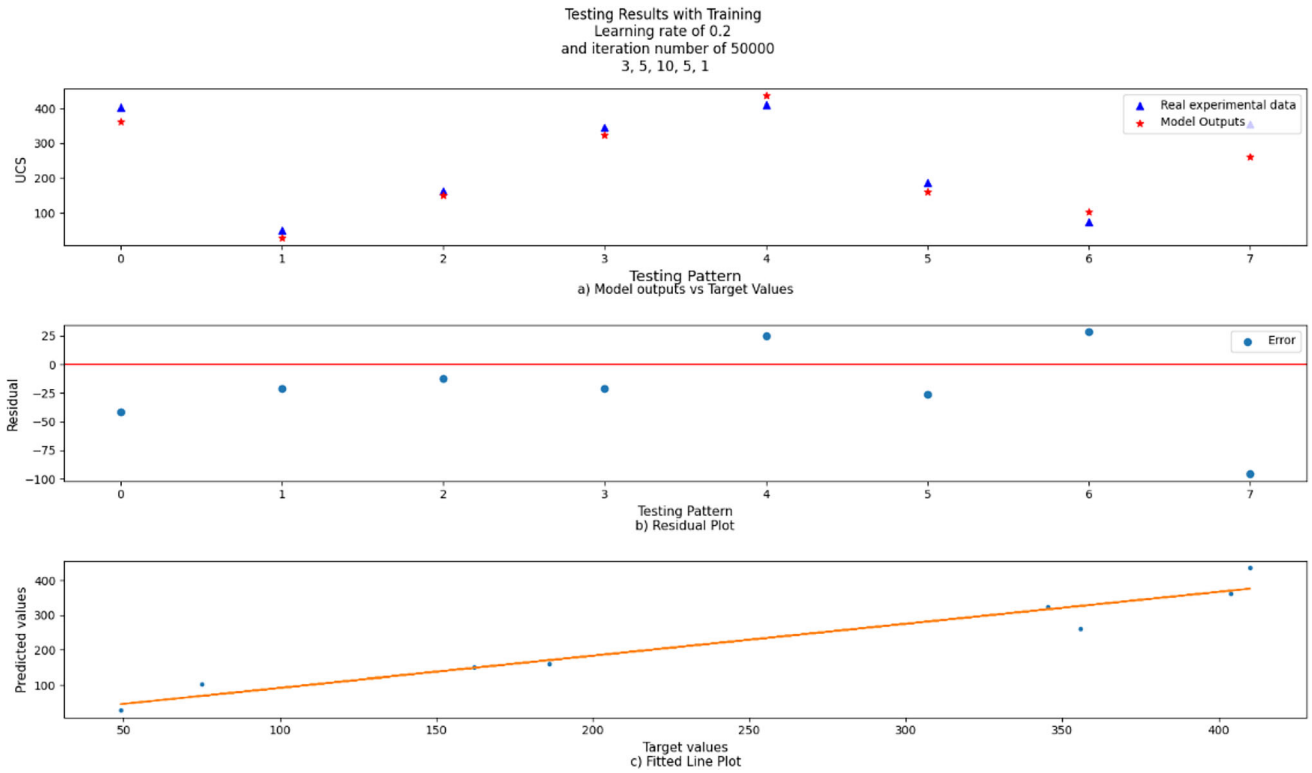


Fig. 19 Testing results for the three-input *ANN* model of *UCS* tests (Inputs: *WGPp, EPSbp, and* UCS_{utr})

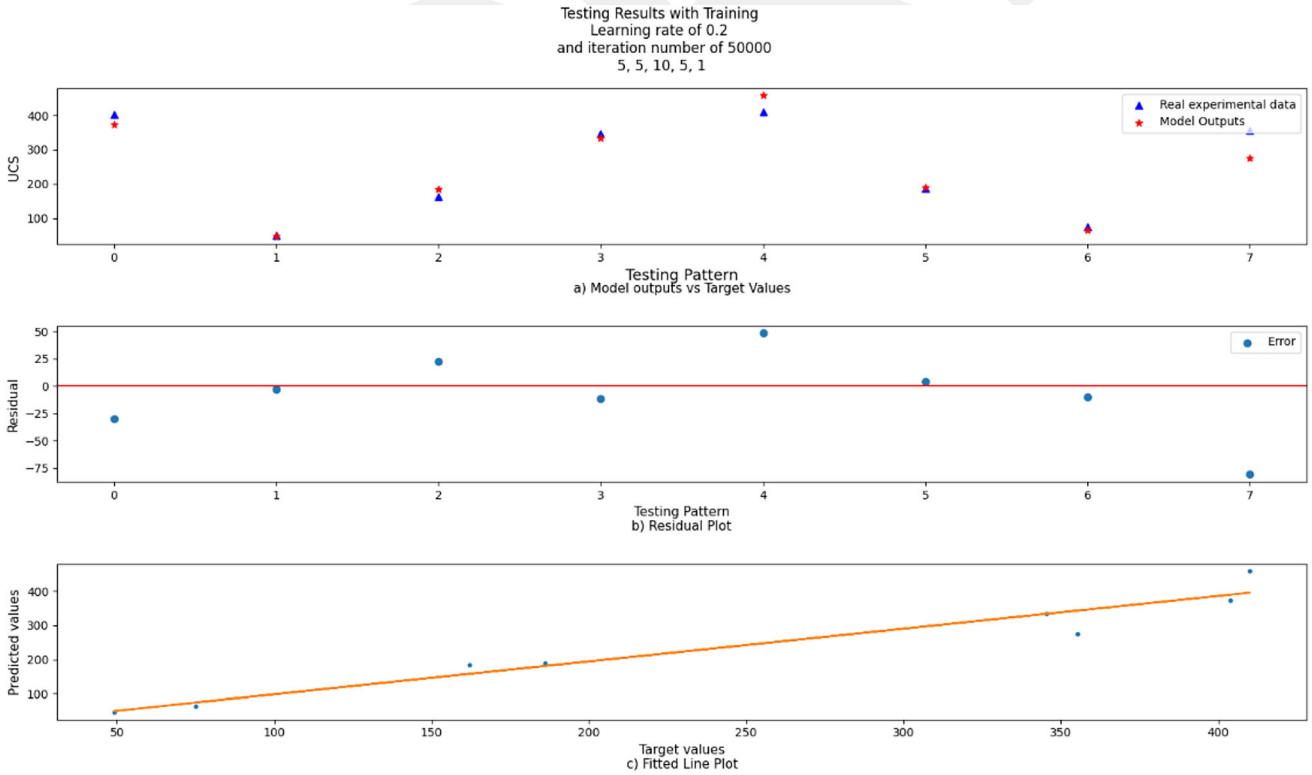


Fig. 20 Testing results for the five-input *ANN* model of *UCS* tests (Inputs: *WGPp, EPSbp, LL, PL, and PI*)

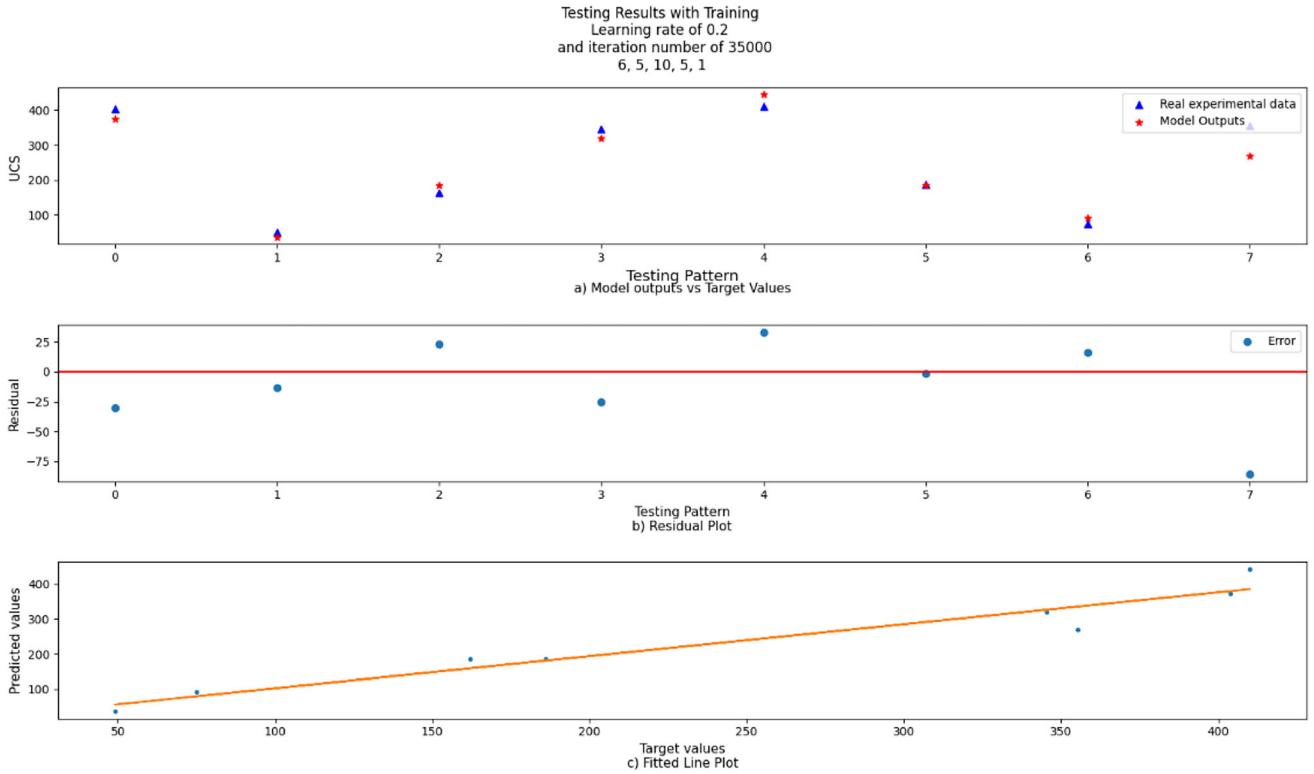


Fig. 21 Testing results for the six-input *ANN* model of UCS tests (Inputs: *WGPp*, *EPSbp*, LL, PL, PI, and UCS_{utr})

Table 15 Statistical data obtained for training and testing processes with three-input parameters

Output	*ANN* (Three Inputs)							
	Inputs	Iteration No	Training			Testing		
			MSE	*MAPE*	R^2	*MSE*	*MAPE*	R^2
FS	*WGPp, EPSbp,* FS_{utr}	50,000	0.29	1.77	0.984	5.49	26.90	0.998
UCS	*WGPp, EPSbp,* UCS_{utr}	50,000	985.37	9.30	0.965	1015.45	17.61	0.930

Table 16 Statistical data obtained for training and testing processes with five-input parameters

Output	*ANN* (Five Inputs)							
	Inputs	Iteration No	Training			Testing		
			MSE	*MAPE*	R^2	*MSE*	*MAPE*	R^2
FS	*WGPp, EPSbp, LL, PL, PI*	50,000	2.88	5.58	0.968	10.00	44.44	0.978
UCS	*WGPp, EPSbp, LL, PL, PI*	50,000	496.98	6.61	0.974	686.18	8.00	0.935

Table 17 Statistical data obtained for training and testing processes with six-input parameters

Output	*ANN* (Six Inputs)							
	Inputs	Iteration No	Training			Testing		
			MSE	*MAPE*	R^2	*MSE*	*MAPE*	R^2
FS	*WGPp, EPSbp, LL, PL, PI,* FS_{utr}	50,000	0.88	3.08	0.995	4.54	28.05	0.995
UCS	*WGPp, EPSbp, LL, PL, PI,* UCS_{utr}	35,000	453.29	6.31	0.972	711.01	13.03	0.934

Table 18 Threefold cross-validation results

Outputs	Inputs	Network	LR	Iteration	Training			Testing		
					MSE	*MAPE*	R^2	*MSE*	*MAPE*	R^2
FS	*WGPp, EPSbp,* UCS_{utr}	[3, 5, 10, 5, 1]	0.2	50,000	0.65	4.72	0.988	19.57	24.46	0.924
	WGPp, EPSbp, LL, PL, PI	[5, 5, 10, 5, 1]	0.2	50,000	0.08	1.55	0.982	25.49	31.93	0.899
	WGPp, EPSbp, LL, PL, PI, UCS_{utr}	[6, 5, 10, 5, 1]	0.2	50,000	0.87	4.86	0.996	11.99	21.33	0.964
UCS	*WGPp, EPSbp,* UCS_{utr}	[3, 5, 10, 5, 1]	0.2	50,000	28.97	2.56	0.963	2790.67	15.69	0.942
	WGPp, EPSbp, LL, PL, PI	[5, 5, 10, 5, 1]	0.2	50,000	147.56	13.18	0.977	4463.18	31.88	0.860
	WGPp, EPSbp, LL, PL, PI, UCS_{utr}	[6, 5, 10, 5, 1]	0.2	50,000	20.76	1.44	0.983	4595.57	26.05	0.884

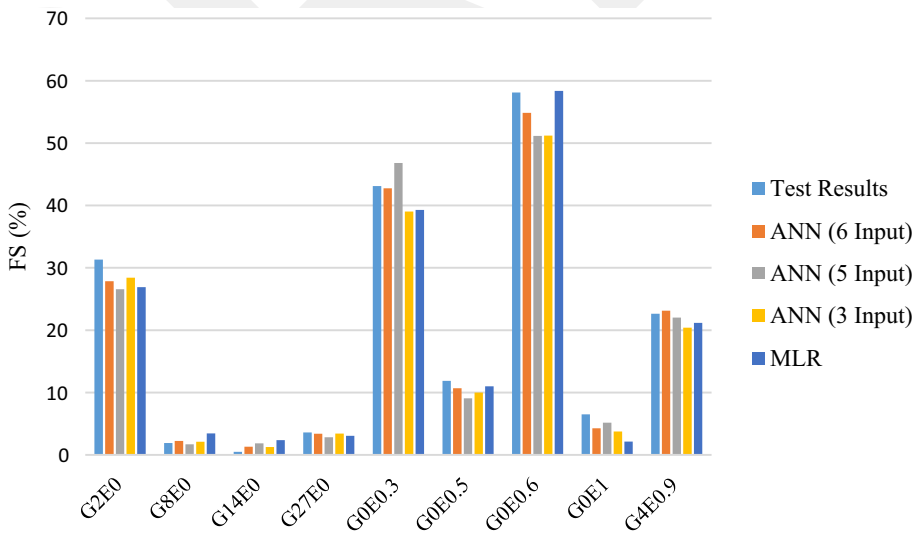


Fig. 22 Comparison of the estimated *FS* values with the testing data obtained from experimental results of soils in this study and previous researches

the treated soils can be evaluated comprehensively when the relevant test result (FS_{utr} or UCS_{utr}) performed on the undisturbed samples is the only independent variable of the prediction equation. However, slightly better results can be achieved if a second variable, such as the plastic limit (*PL*) of the untreated sample, is added to the prediction function to estimate *UCS*. Finally, it can be concluded that, if the *FS* and *UCS* values of the untreated soil samples (FS_{utr} or UCS_{utr}) are available, the *FS* and *UCS* values of the treated soils could be predicted by using Eqs. 10 and 11 given below where the R^2 values are more than 0.926.

$$FS = 1.591 - 5.965EPSbp - 0.180WGPp + 0.844FS_{utr} \quad (10)$$

$$UCS = -58.460 - 102.948EPSbp + 5.889WGPp + 1.203UCS_{utr} \quad (11)$$

A developed *ANN* model is also used to estimate the *FS* and *UCS* of treated soil samples, which can be used with or

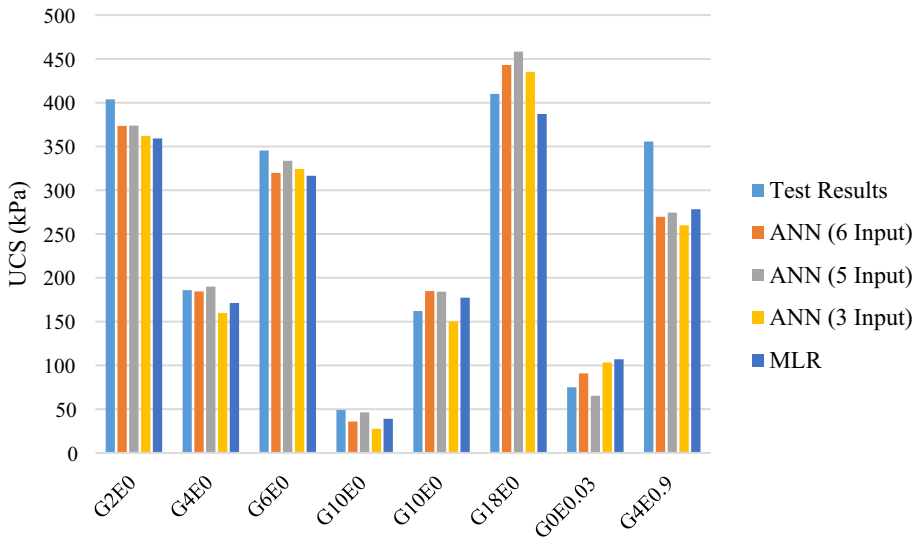


Fig. 23 Comparison of the estimated *UCS* values with the testing data obtained from experimental results of soils in this study and previous researches

without FS_{utr} and UCS_{utr} depending on available data. The results show that the model slightly outperforms linear regression in the prediction of *FS* and *UCS* by using additives, Atterberg Limits, FS_{utr} and, UCS_{utr} ($R^2 = 0.996$ and 0.975). Besides, this alternative strategy produces similar results in the estimation of *FS* and *UCS* ($R^2 = 0.987$and0.964) by using limited input data, such as additives and the Atterberg Limits. Similar datasets are employed for training and testing purposes in order to compare the results obtained from the *MLR* and *ANN* algorithms. Using the testing data, the performance of the proposed equations and the trained *ANN* model is tested. Figures 22 and 23 show the experimentally obtained *FS* and *UCS* values from this and prior studies, as well as the results of the prediction equations (Eqs. 10 and 11) and the *ANN* method.

As stated earlier, the first part of this study investigates both the individual and combined effects of *WGP* and/or *EPSb* on the engineering properties of high plasticity clay through a series of experiments. In the second part, the study focuses on utilizing the effectiveness of *ANN* methods in estimating the mechanical properties of the clay. It is shown that the algorithm satisfactorily predicts the *FS* and *UCS* of the treated soil samples. On top of this, empirical equations are also derived through *MLR* analyses. This study not only examines the individual effects of waste glass powder (*WGP*) and expanded polystyrene beads (*EPSb*) on high plasticity clay, but also investigates their combined impact—a novel approach not previously explored in the literature. While individual regression analyses and *ANN* models are already available specifically for each waste material in the literature, this research focuses on the effects of these methods and models when used in a combined form. Therefore, the regression and *ANN* models are developed in accordance with this approach, thus representing a meaningful contribution to the field.

5 Conclusion

In this study, the change in the index and the mechanical properties of treated soil samples with waste materials are studied. In the experimental part, the fat clay sample is treated with different percentages of additives (waste glass powder and/or *EPS* beads), and the Atterberg limits, optimum moisture content, maximum dry density, free swell, and unconfined compressive strength values are obtained. Both the liquid limits (*LL*) and the plastic limits (*PL*) decrease as the percentage of *WGP* addition increases, resulting in a decline in the plasticity indexes (*PI*). The *OMC* and *MDD* values for untreated and treated bentonite with various percentages of *WGP* and/or *EPSb* are determined. The soil mixtures are compacted at these values by using a standard proctor test apparatus; for the *FS* and *UCS* tests, the samples are extracted from the proctor mold. It is observed that the addition of *EPSb* reduces the *FS* and *UCS* of the soil samples; however, employing *WGP* as a stabilizing agent has a significant effect on the *FS* and *UCS*. In conclusion, the ideal percentage of additives is decided as 4% *WGP* and 0.9% *EPSb* for the fat clay used in this study.

In the computational part, two datasets containing 45 and 40 sets of experimental test results obtained by this study as well as reported findings in the literature are used to develop and verify the models for the estimation of *FS* and *UCS* of the treated samples, respectively. As much as 80% of this data is used for training, whereas the remaining 20% is used for testing purposes. Based on these datasets, the empirical prediction equations are developed by using *MLR*. The *ANN* models are trained to estimate the *FS* and

UCS of the treated clays. The performances of the *MLR* and *ANN* models are also compared by using the same datasets and input variables. The results indicate that both models have the ability to predict the *FS* and *UCS* of the treated soil samples with an acceptable level of accuracy in cases where the properties of the untreated soil samples together with waste material percentage (with three inputs-*EPSbp*, *WGPp, FS_{utr} or UCS_{utr}*) are used. The obtained R^2 values of *FS* and *UCS* using the ANN model with three inputs are 0.984 and 0.965 for training, and 0.998 and 0.930 for testing, respectively. These values are reasonably close to those of *MLR* for *FS* and *UCS*, namely the R^2 values of 0.936 and 0.926 for training, as well as 0.981 and 0.928 for testing, respectively. Furthermore, the *ANN* models with five and six inputs are successfully implemented. It is observed that, if there are limited known input parameters, the *ANN* model with five inputs should be used. Furthermore, the model with six inputs can also be used in case the test results carried out on untreated soil samples are available. The numerical results indicate that the training R^2 values for *FS and UCS* are 0.968 and 0.995 and and 0.974 and 0.972 for five and six inputs, respectively. On the other hand, the testing R^2 values for these tests are 0.978 and 0.935 and 0.995 and 0.934, respectively. Based on these findings, it can be concluded that the performance of the *ANN* models in all the scenarios investigated is promising, and that this method can be used to predict *FS* and *UCS* if only limited parameters are available.

Finally, in order to assess the independency of the *ANN* models on the selection of datasets, a threefold cross-validation technique is used to validate the performance of the models. The results of this validation reveal that the *ANN* models used in this study to achieve similar results over all the employed datasets. To conclude, this methodology can be adapted to various datasets without overfitting, and both of the proposed *MLR* and *ANN* methods can successfully predict the *FS* and *UCS* of improved soils.

Acknowledgements The authors would like to thank Asst. Prof. Dr. Bahram Lotfi for his constructive comments and support on this study.

Author's contribution EA contributed to the conceptualization, methodology, data curation, draft preparation, regression analyses, ANN modeling, writing, and editing. OYC was involved in the data curation, draft preparation, regression analyses, and writing.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK). Not Applicable.

Data availability Data are available upon request.

Code availability Code is available upon request.

Declarations

Conflict of interest Not Applicable.

Consent to participate Hereby, we as the authors of the article "Predictive Models for Treated Clayey Soils Using Waste Powdered Glass and Expanded Polystyrene Beads Using Regression Analysis and Artificial Neural Network", declare our consent to participate in this research on our free will.

Consent for publication Hereby, we as the authors of the article "Predictive Models for Treated Clayey Soils Using Waste Powdered Glass and Expanded Polystyrene Beads Using Regression Analysis and Artificial Neural Network", declare our consent to publish this paper in the Journal of Neural Computing and Applications.

Ethical approval Not Applicable.

References

1. Bacinschi Z, Rizescu CZ, Stoian EV and Necula C (2010) Waste management practices used in the attempt to protect the environment. Latest Trends Eng Mech Struct Eng Geol 378–382
2. Hidalgo-Crespo J, Jervis FX, Moreira CM, Soto M, Amaya JL (2020) Introduction of the circular economy to expanded polystyrene household waste: a case study from ecuadorian plastic manufacturer. Procedia CIRP 90:49–54. https://doi.org/10.1016/j.procir.2020.01.089
3. Shirazi AN, Haydarian H, Nasehi SA (2018) Shear and compression behaviors of sandy and clayey soils mixed with different sizes of expanded polystyrene beads. Geotech Geol Eng 36:3823–3830. https://doi.org/10.1007/s10706-018-0575-y
4. Illuri HK (2007) Development of soil-EPS mixes for geotechnical applications. School of Urban Development, Australia
5. Rocco NT (2012) Characterization of Expanded Polystyrene (EPS) and Cohesive Soil Mixtures. Missouri University of Science and Technology, Missouri
6. Soundara B, Selvakumar S (2019) Swelling behavior of expansive soils randomly mixed with recycled geobeads inclusion. SN Appl Sci 1:1253. https://doi.org/10.1007/s42452-019-1324-4
7. United States Environmental Protection Agency (2021) National Overview: Facts and Figures on Materials: Wastes and Recycling
8. Subbarao GVR, Siddartha D, Muralikrishna T, Sailaha KS, Sowmya T (2011) Industrial wastes in soil improvement. Int Scholar Res Netw. https://doi.org/10.5402/2011/138149
9. Canakci H, Al-Kaki A, Celik F (2016) Stabilization of clay with waste soda lime glass powder. Proc Eng 161:600–605. https://doi.org/10.1016/j.proeng.2016.08.705
10. Fauzi A, Djauhari Z, Fauzi UJ (2016) Soil engineering properties improvement by utilization of cut waste plastic and crushed waste

glass as additive. IACSIT Int J Eng Technol. https://doi.org/10.7763/IJET.2016.V8.851
11. Bilgen G (2020) Utilization of powdered glass as an additive in clayey soils. Geotech Geol Eng 38:3163–3173. https://doi.org/10.1007/s10706-020-01215-7
12. Bilgen G (2020) Utilization of powdered glass in lime-stabilized clayey soil with sea water. Environ Earth Sci 79:437. https://doi.org/10.1007/s12665-020-09195-w
13. Ibrahim HH, Mawlood YI, lshkane YM (2019) Using waste glass powder for stabilizing high plasticity clay in Erbil City-Iraq. Int J Geotech Eng 15(2):1–8. https://doi.org/10.1080/19386362.2019.1647644
14. Mujtaba H, Khalid U, Farooq K, Elahi M, Rehman Z, Shahzad HM (2020) Sustainable utilization of powdered glass to improve the mechanical behavior of fat clay 24, 3628–3639. KSCE J Civ Eng. https://doi.org/10.1007/s12205-020-0159-2
15. Akis E, Guven G, Lotfisadigh B (2022) Predictive Models for mechanical properties of expanded polystyrene (EPS) geofoam using regression analysis and artificial neural network. Neural Comput Appl. https://doi.org/10.1007/s00521-022-07014-w
16. Yaprak H, Karacı A, Demir I (2013) Prediction of the effect of varying cure conditions and w/c ratio on the compressive strength of concrete using artificial neural networks. Neural Comput Appl 22:133–141. https://doi.org/10.1007/s00521-011-0671-x
17. Bal L, Buyle-Bodin F (2014) Artificial neural network for predicting creep of concrete. Neural Comput Appl 25:1359–1367. https://doi.org/10.1007/s00521-014-1623-z
18. Belalia Douma O, Boukhatem B, Ghrici M et al (2017) Prediction of properties of self-compacting concrete containing fly ash using artificial neural network. Neural Comput Appl 28:707–718. https://doi.org/10.1007/s00521-016-2368-7
19. Adil M, Ullah R, Noor S et al (2020) Effect of number of neurons and layers in an artificial neural network for generalized concrete mix design. Neural Comput Appl. https://doi.org/10.1007/s00521-020-05305-8
20. Armaghani DJ, Asteris PG (2021) A comparative study of ANN and ANFIS models for the prediction of cement-based mortar materials compressive strength. Neural Comput Appl 33:4501–4532. https://doi.org/10.1007/s00521-020-05244-4
21. Ahmad A, Ostrowski KA, Maślak M, Farooq F, Mehmood I, Nafees A (2021) Comparative study of supervised machine learning algorithms for predicting the compressive strength of concrete at high temperature. Materials 14(15):4222. https://doi.org/10.3390/ma14154222
22. Al Khazaleh M, Bisharah M (2023) ANN-based prediction of cone tip resistance with Tabu-Search optimization for geotechnical engineering applications. Asian J Civ Eng 24:3037–3054. https://doi.org/10.1007/s42107-023-00693-3
23. Nguyen TH, Chau TL, Hoang T et al (2023) Developing artificial neural network models to predict corrosion of reinforcement in mechanically stabilized earth walls. Neural Comput Appl 35:6787–6799. https://doi.org/10.1007/s00521-022-08043-1
24. Nguyen DK, Nguyen TP, Ngamkhanong C et al (2023) Bearing capacity of ring footings in anisotropic clays: FELA and ANN. Neural Comput Appl 35:10975–10996. https://doi.org/10.1007/s00521-023-08278-6
25. Shimobe S, Karakan E, Sezer A (2023) Evaluation of dependency of compression index on toughness limit for fine-grained soils. Neural Comput Appl 35:11183–11205. https://doi.org/10.1007/s00521-023-08292-8
26. Christidis G, Huff W (2009) Geological aspects and genesis of bentonites. Elements 5(2):93–98. https://doi.org/10.2113/gselements.5.2.93
27. Kok E, Erdogan Y, Ozdemir A (2023) Bentonitlerin Oluşumu, Sınıflandırılması ve Kullanım Alanları. Mühendislikte Güncel Araştırmalar (1st ed.), Turkey: Gece Kitaplığı
28. ASTM D854–14 (2014) Standard Test Methods for Specific Gravity of Soil Solids by Water Pycnometer. ASTM International, West Conshohocken, PA, USA
29. ASTM D698 (2012) Standard Test Methods for Laboratory Compaction Characteristics of Soil Using Standard Effort. ASTM International, West Conshohocken, PA, USA
30. ASTM D6913 (2004) Standard Test Methods for Particle-Size Distribution (Gradation) of Soils Using Sieve Analysis. ASTM International, West Conshohocken, PA, USA
31. ASTM D7928-21 (2021) Standard Test Method for Particle-Size Distribution (Gradation) of Fine-Grained Soils Using the Sedimentation (Hydrometer) Analysis. ASTM International, West Conshohocken, PA, USA
32. BS:1377-2 (1990) Methods of Test for Soils for Civil Engineering Purposes- Classification Tests.
33. ASTM D4318-17 (2017) Standard Test Methods for Liquid Limit, Plastic Limit, and Plasticity Index of Soils. ASTM International, West Conshohocken, PA, USA
34. ASTM D4546 (2014) Standard Test Methods for One-Dimensional Swell or Collapse of Soils. ASTM International, West Conshohocken, PA, USA
35. ASTM D2487-17 (2017) Standard Practice for Classification of Soils for Engineering Purposes (Unified Soil Classification System). ASTM International, West Conshohocken, PA, USA
36. Ferdous W, Manalo A, Siddique R, Mendis P, Zhunge Y, Wong H, Lokuge W, Aravinthan T, Schubel P (2021) Recycling of landfill wastes (tyres, plastics and glass) in construction – a review on global waste generation, performance, application and future opportunities. Resour Conserv Recycl 173:105745. https://doi.org/10.1016/j.resconrec.2021.105745
37. Sayanthan R, Ilamaran S, Rifdy M and Nanayakkara SMA (2013) Development of Interlocking Lightweight Cement Blocks. In: Special Session on Construction Materials & Systems, Internal Conference on Structural Engineering and Construction Management, vol 13/53, pp 194–202
38. ASTM D3080/D3080M-11 (2011) Standard Test Method for Direct Shear Test of Soils Under Consolidated Drained Conditions. ASTM International, West Conshohocken, PA, USA
39. Al-Kaki AK (2016) Clay soil stabilization with waste soda lime glass powder. University of Gaziantep, Turkey
40. Akis E, Çiğdem ÖY (2023) EPS Daneciklerinin ve/veya Cam Tozunun Killi Zeminlerin Kıvam Limitlerine Etkisi ve Limitlerin YSA ve Regresyon ile Tahmin Edilmesi. J Inst Sci Technol 13(1):385–398. https://doi.org/10.21597/jist.1173024
41. Olufowobi J, Ogundoju A, Micheal B, Aderinlewo O (2014) Clay Soil stabilizaiton using powdered glass. J Eng Sci Technol 9(5):541–558
42. Silveria MV, Calheiros AV, Casagrande MDT (2018) Applicability of the expanded polystyrene as a soil improvement tool. J Mater Civ Eng (ASCE) 30(6):06018006. https://doi.org/10.1061/(ASCE)MT.1943-5533.0002276
43. ASTM D2166 (2000) Standard Test Method for Unconfined Compressive Strength of Cohesive Soil. ASTM International, West Conshohocken, PA, USA
44. Tabachnick BG, Fidell LS and Ullman JB (2019) Using Multivariate Statistics (7th ed.), New York: Pearson
45. Salkind NJ (2016) Statics for People Who (think they) Hate Statistics. (4th ed.), LA, London, New Delhi, Singapore, Ishington D.C., Melbourne, SAGE

46. Hair JF, Anderson RE, Tatham RL and Black WC (1995) Multivariate Data Analysis (3rd ed.), New York: Macmillan
47. Zurada JM (1992) Introduction to Artificial Neural Systems. West St. Paul
48. Nguyen H, Bui XN, Bui HB, Mai NL (2020) A comparative study of artificial neural networks in predicting blast-induced air-blast overpressure at Deo Nai Open-Pit Coal Mine, Vietnam. Neural Comput Appl 32(8):3939–3955. https://doi.org/10.1007/s00521-018-3717-5
49. Yaguo L (2017) Individual intelligent method-based fault diagnosis. In Yaguo Lei (Ed) Intelligent Fault Diagnosis and Remaining Useful Life Prediction of Rotating Machinery, pp 67–174
50. Zerguine A (2001) Multilayer perceptron-based DFE with lattice structure. IEEE Trans Neural Netw 12(3):532–545. https://doi.org/10.1109/72.925556
51. Ruppert D (2004) Statistics and finance: an introduction. Springer, New York, USA
52. Han J, Kamber M and Pei J (2012) Data Mining: Concepts and Techniques (3rd ed). Elsevier/Morgan Kaufmann