Developing new Adaptive Neuro-Fuzzy Inference Systems to predict granular soil groutability

Mostafa Asadizadeh a,*, Abbas Majdi b

a Department of Mining Engineering, Hamedan University of Technology, Mardam Street, Hamedan, Iran
b School of Mining Engineering, University of Tehran, Tehran, Iran

Abstract

Three Neuro-Fuzzy Inference Systems (ANFIS) including Grid Partitioning (GP), Subtractive Clustering (SCM), and Fuzzy C-means clustering Methods (FCM) have been used to predict the groutability of granular soil samples with cement-based grouts. Laboratory data from related available in the literature was used for the tests. Several parameters were taken into account in the proposed models; water:cement ratio of the grout, the relative density of the soil, grouting pressure, soil and grout particle size dimensions named $D_{15\text{soil}}$, $D_{15\text{grout}}$, $D_{45\text{soil}}$, and $D_{45\text{grout}}$, and the percentage of the soil particles passing through a 0.6 mm sieve. The accuracy of ANFIS models was examined by comparing these models with the results of experimental grout-ability tests. A sensitivity analysis showed that the ratios of $D_{15\text{soil}}/D_{45\text{grout}}$ and $D_{15\text{soil}}/D_{45\text{grout}}$ were the most effective parameters on the groutability of granular soil samples with cement-based grouts and the grout water:cement ratio of the grout was determined as the least effective parameter.

Keywords: Groutability, ANFIS, Clustering Algorithm, Granular soil

1. Introduction

Grouting is a process whereby some external materials are injected into soil pores and rock fissures to improve the mechanical and hydraulic properties of the media for short- and long-term engineering needs. Generally, the process involves injecting a viscous grout mixture under the pressure into the porous media or joints until they are blocked by the larger grout particles. Permeation grouting has been commonly used in ground improvement techniques, it has also been used extensively to increase the liquefaction resistance of existing structures because it has a low grouting pressure (P) [1].

Groutability (N) of grouts is defined in terms of its capability for injection into a target soil or rock to improve the mechanical properties of materials or to reduce their permeability [2–6]. An important issue for application of cement-based grout is the trustworthy prediction of groutability in different media. The grouting process relies on complicated time-dependent transportation of cement particles right through the other side of the soil or rock materials. This prediction process requires several variables such as the distribution of grain size in the grout and soil materials, pore size and hydraulic conductivity of the soil, the injection pressure, and the viscosity of the grout suspension. This complicates the predictions meaning that no universally comprehensive set of criteria or methodology have been determined. Many researches have been carried out on tests for soil groutability predictions. Some of these have presented basic empirical equations related to the size of the soil and cement particles [7,8,9–20]. In practice, empirical criteria are regularly applied as the main tools to determine the groutability ratio of soil samples. A primary research on groutability predictions of granular soil samples includes an analogy between the particle dimension of the host soil and cement grout. It was reported that the penetration zone by the grout mixture in a soil sample is limited by the grain size distribution of the soil [8]. However, the research on grout penetrability has reported that an important consideration is the size of the voids being grouted in relation to the size of the solid particles in the grout [12]. Groutability was defined as the ratio (GR) for grouting natural soil formations as Equation 1.

$$GR = \frac{D_{15\text{soil}}}{D_{45\text{grout}}} > 25$$

Where $D_5$ corresponds to the sieve diameter through which 15 wt % of the soil sample passes or 15 wt% of the soil sample is finer than the grain size diameter, and $D_6$ represents the grain size diameter below which 85 wt % of the grout mixture is finer. According to [16], the first requirement for the selection of grout is that its particle size should be smaller than the dimensions of the voids to be filled, which is determined by Equation 2.

$$\frac{D_{15\text{soil}}}{D_{45\text{grout}}} > N$$

Where N ranges from 5 to 20 proposed for clay grouts and depends on local conditions. It was further added that the second and third (the last) requirements are stability and pumpability of a grout sample, respectively. However, in [13], the major controllable variables affecting the efficiency of injection are as follow: pumping rate, setting time of the grout, pumping time at a given pipe location, the distance-time schedule for pulling or driving the pipe, grout viscosity, grouting pattern, and possible inter-relationships among the parameters. Additionally, in [15], the major uncontrollable variables were groundwater flow and stratification. According to [14], grouting with cement is not possible when the sand:grout ratio, $D_{15\text{soil}}/D_{15\text{grout}}$ is below 11. Grouting is only possible when the sand:grout ratio, $D_{15\text{soil}}/D_{15\text{grout}}$ is more

* Corresponding author. E-mail address: m.asadizadeh@hut.ac.ir (M. Asadizadeh).
than 24. For an effective injection, the soil pore space should be three times of the diameter of the grout particle to avoid the blockage by bridging (Kennedy, 1962; King & Bush, 1963). Based on experimental data, the U.S. Army Corps of Engineers Waterways Experiment Station proposed that grouting could be successfully accomplished if the grout ability ratio was greater than \( N = 15 \) [15], as shown in Equation 3.

\[
N = \frac{D_{15\text{soil}}}{D_{15\text{grout}}} \geq 15
\]

According to [21], the groutability ratio is defined as the ability of grout particulates (soil, cement, clay) to penetrate a soil formation. For successful grouting of a soil:

\[
GR = \frac{D_{30\text{soil}}}{D_{15\text{grout}}} \geq 25
\]

The soil grain size limits the penetrability of the grout mixture [22]. So that, in a soil sample, if more than 10 wt% of particles passes through sieve number 200, the soil sample is not considered groutable. According to [9], the dimensions of the voids exist within the soil particles compared to those of the grout mixture particles and is considered as a logical criterion. Furthermore, some ‘groutability ratios’ that have been proven useful for soil samples were presented as follows:

\[
\begin{align*}
\text{N} & = \frac{D_{10\text{soil}}}{D_{10\text{grout}}} \\
\text{N} & = \frac{D_{20\text{soil}}}{D_{20\text{grout}}}
\end{align*}
\]

(4)

When \( N > 24 \), grouting is permanently feasible and when \( N < 11 \), grouting is not practicable. As well, when \( N > 11 \), grouting is permanently feasible and when \( N < 6 \), grouting is not practicable [23]. A method was developed and used to measure the complicated structure of pore spaces [24]. According to [25], permeation is controlled by the size of particle than by viscosity and cohesion. It was added that the size of the pores of granular soil and aperture of rock fissures were dominant controls over groutability. While the geometry of a fissure is relatively simple to model, the pore system of loose soils is complex. Based on the results of cement injected sands and gravel samples [11], groutability ratios are suggested to be as shown in Equation 6.

\[
\begin{align*}
\text{N} & = \frac{D_{30\text{soil}}}{D_{15\text{grout}}} > 10 \\
\text{N} & = \frac{D_{20\text{soil}}}{D_{15\text{grout}}} > 3
\end{align*}
\]

In which, \( D_m \) is the size of soil voids. An empirical the groutability ratio of granular soil was proposed for consolidation grouting as a function of grain size, relative density and fine contents of soil, the dimension of cement particulates, water:cement ratio of the grout mixture and the grouting pressure [7]. The performance of conventional groutability criteria which are based on the groutability ratios can be considered as optimistic. According to [17], test groutability formulas do not consider parameters such as the characteristic grain sizes of cement and soil, water-cement ratio, type and percentage of used additives, and sand density. Well-designed formulations through the proper application of additives reduce the cost of a grouting operation via decreasing the cohesion and increasing the penetrability of the grout [19]. Decreasing the cohesion of a grout mixture, whenever keeping a consistent grout mixture, which boosts its penetration, has been reported in other researches [26]. However, it is reported in [27] that groutability was also enhanced by using flash, while permeability showed a significant decrease. Moreover, the possible cement grout penetration, investigation of the influence of grout viscosity variation and filtration were studied in [28]. Artificial inference systems have been recently developed to solve complicated problems. Systems such as neural networks and fuzzy logic have been used to solve many geotechnical difficulties in recent years. Such systems and their associated methods have advantages and some disadvantages. The advantage of artificial neural networks is pattern recognition and the capacity of adapting a method to cope with changing environments. Fuzzy logic has the advantage of incorporating human knowledge and expertise to deal with uncertainty and imprecision. Therefore, many efforts have been made to take the advantage of both of these approaches. As a result of these studies, many investigators have recently suggested the application of a combination of these approaches termed as the ANFIS method [29-40]. In the case of permeation grouting, some researchers utilized an Artificial Neural Network (ANN) for predictions of soil groutability [41]. The researchers reported that classical groutability prediction formulas, which are mainly contingent upon the grain-size of the soil and the grout, were not suitable for semi-parametric scale grout. They found that the accuracy of the proposed formulas varied from 45% to 68%, a domain that is not suitable for practical purposes. Unlike this, an ANN model was proposed by [42] to predict the groutability of granular soils by cement-based grouts, utilizing grouting pressure, water:cement ratio of the grout, the diameter of sieve through which 15% of soil particles and 85% of the grout pass, and the relative density of the soil. It was reported that high success rates exceeding 90% for some existing empirical methods and a highly successful prediction ratio (95.4%) was obtained using the ANN models. According to [43], the water-cement ratio of the grout was the most effective parameter on the dynamic response of grouted sands and the influence of cement grain size and cement pozzolan ingredient were secondary, but not insignificant.

In this paper, three known ANFIS models are developed to evaluate the groutability of granular soil samples with cement grout; these include GP, SCM, and FCM. To fulfill this goal, datasets of 87 laboratory-grouting tests were employed using the data available in the related literature. The efficiency of ANFIS models was compared with the test results. Since it was not clear which of the above-mentioned ANFIS methods had the better performance in terms of addressing the grouting problems, a comparison was carried out to compare the performance of the three models to determine the best performing model. Furthermore, a sensitivity analysis was performed to distinguish the most effectual input variables on the groutability of granular soil samples.

2. Background study

2.1 Takagi–Sugeno fuzzy system

The theory of fuzzy sets was presented by Zadeh to cope with problems that have uncertainty due to ambiguity and imprecision [44]. The fuzzy set theory accurately investigates ambiguous conceptual phenomena using a precise mathematical framework. This theory is an appropriate modeling language for imprecision and ambiguous theoretical criteria, phenomena, and relations [45]. The process of formulating the mapping from an input to an output utilizing fuzzy logic is called fuzzy inference. The fuzzy set theory can decode the enigmatic states of reasoning utilized in an environment defined by uncertainty and ambiguity. Fuzzy logic is the system of concepts, rules, and approaches applied to rough reasoning using the fuzzy set theory [46]. Fuzzy logic uses a list of ‘if-then’ statements called rules to map an input domain to an output domain. The general form of rules is as follows:

\[
\text{if } w \text{ is } C \text{ then } z \text{ is } D
\]

in which \( w \) and \( z \) are variables in domains \( W \) and \( Z \); \( C \) and \( D \) are fuzzy sets based on \( W \) and \( Z \), respectively. In this rule, the antecedent ‘is if’ part of the sentence and the consequence is the ‘then’ part of the sentence [46]. Applying the information verbalized in the form of natural linguistic statements is the prominent feature of rule-based fuzzy logic. Membership function (MF) is the main concept of fuzzy logic, and numerically, it states the degree to which a given element belongs to a fuzzy set. Some methods such as expert judgment or data analysis can be used to apply the number of MF, location, and shape to the fuzzy model [46]. A subcategory of model designation that copes with creating a fuzzy logic is fuzzy modeling. The response of an unfamiliar system specified using some sample data can be foreseen and explained using a fuzzy inference system [47]. A fuzzy inference system is a world-renowned computing system, which is contingent upon notions of fuzzy logic. Different parts of a Fuzzy Inference System (FIS) are presented in Table 1. Furthermore, a schematic diagram of FIS is presented in Fig. 1 [48].

Different applications of FIS have been published so far. The most prevalent fuzzy models used are the Mamdani, Takagi–Sugeno–Kang (TSK), Tsukamoto Singleton [49]; and the TS fuzzy model introduced by Takagi and Sugeno [50].
2.2 Basic Concept of ANFIS

FIS is able to model the inference procedure and linguistic features of human understanding without applying accurate quantitative investigations. ANNs are a combination of many interdependent processing components that are comparable to neurons. A collection of data is imported to the training algorithm of ANN and the output is checked for the desired result through this algorithm. In this method, the human process of decision-making is intelligently imitated by a combination of ANN and FIS. In conventional ANN, just weight quantity alters throughout the learning phase, whereas in a neuro-fuzzy decision-making system, the learning capability of ANN is coupled with the reasoning process of FIS [53]. ANFIS has been introduced as an Adaptive Neuro-Fuzzy Inference System. Basically, ANFIS utilizes a FIS and adjusts it using a backpropagation algorithm and employing a set of input-output data. The combination of FIS and ANN enables FIS to learn. The structure of an adaptive neural network includes several nodes joined via oriented links. A node function with unchangeable or adaptable parameters defines each node. Neural network algorithms, when FIS is loaded, can be applied to calculate unknown factors and this decreases the error values, as traditionally described for every parameter of the model and this optimization process makes the model adaptive [54].

ANFIS usually utilizes amalgamation of backpropagation for the purpose of learning the presupposition parameters and the least mean square for determination of resulting parameters. A stage of the learning process has two phases: the first phase involves the propagation of input patterns and applies the iterative least mean square process to evaluate the optimal ending parameters, whilst ancestor parameters are supposed not to change for the present phase throughout the training set. Furthermore, the second phase is included in the repeated propagation of patterns and in this epoch, backpropagation is utilized to adjust the ancestor parameters. This process is followed by iteration (Fig. 2a). For instance, we considered a FIS that has two inputs x and y and one output z. Therefore, two fuzzy ‘if-then’ rules of Takagi and Sugeno's type are presented in Equation 8 [52].

**Rule 1:** If x is A1 and y is B1, then \( f_1 = p_1x + q_1y + r_1 \)  
**Rule 2:** If x is A2 and y is B2, then \( f_2 = p_2x + q_2y + r_2 \)  

The adaptive neural network and its operationally identical to FIS are presented in Figs. 2a and 2b respectively. In this study, a neuro-fuzzy model was utilized containing five layers [55], as follow:

- In the first layer; every node i creates a membership grade of a lingual ticket. For example, the node function of the ith node could be as shown in Equation 9, [52]:

\[
Q_i^1 = \mu_{A_i}(x) = \exp\left(-\frac{x - c_i}{a_i}\right)  
\]

in which, x is the input to node i, and \( A_i \) is the lingual ticket (small, large, etc.) introduces this node and \(\{a_i, b_i, c_i\} \) is the parameter collection that alters the MF configuration. The parameters of the first layer are named “premise parameters”.

The outputs of the third layer are named “normalized firing strengths”.

In the fourth layer, the node function of each node i is presented in Equation 12, [52].

\[
Q_i^4 = \bar{w}_i f_i = \frac{w_i}{\sum w_i} f_i \quad i = 1, 2  
\]

in which \(\bar{w}_i\) is the third layer output. Parameters in the fourth layer are called ‘consequent parameters’.

In the fifth layer, an individual node shown by a circle calculates the overall output as the aggregate of all interring signals. The node function is represented by Equation 13.

\[
Q_i^5 = \text{Overall Output} = \Sigma \bar{w}_i f_i = \frac{\Sigma w_i f_i}{\Sigma w_i}  
\]
parameter identification can be applied to boost the forestall-ability of the model. Therefore, a set of information including experimental tests is employed to develop the well-made ANFIS models. The tests consist mainly of water to cement proportion of the grout mixture (W/C), the amount of soil which is less than 0.6 mm (FC), the relative density of soil (Dr), the pressure under which grout is injected (P) and the grain-size distribution of the host soil and the grout. Based on [57], an investigation was done to determine the groutability of sand samples with various grain-size distribution curves using micro fine cement samples. In that study, fifteen experiments were accomplished by injecting the grout mixture under 100 kPa pressure into the sand with a given relative density. Furthermore, in [42] the same test setup was used to carry out sixteen grouting tests. Based on [7], 38 grouting tests were performed on granular soil samples. However, eighteen experimental grouting tests were accomplished on granular soil samples [58]. All datasets were collected from the literature, and the ANFIS models were established according to a total of 87 experiments involving W/C, Dr, P, FC, NI = D_{15\text{ soil}}/d_{45\text{ grout}}, N2 = D_{15\text{ soil}}/d_{45\text{ grout}} and the results of the grouting experiments. All data sets are presented in Table 2 and in order to investigate the exact details of the experimental procedure, related references are presented as follows [58]:

<table>
<thead>
<tr>
<th>No</th>
<th>W/C</th>
<th>Dr(%)</th>
<th>P(kPa)</th>
<th>FC(%)</th>
<th>NI = D_{15\text{ soil}}/d_{45\text{ grout}}</th>
<th>N2 = D_{15\text{ soil}}/d_{45\text{ grout}}</th>
<th>Experimental Groutability</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>80</td>
<td>50</td>
<td>1</td>
<td>87.10</td>
<td>41.67</td>
<td>1</td>
<td>[19]</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>80</td>
<td>100</td>
<td>1</td>
<td>45.16</td>
<td>21.33</td>
<td>1</td>
<td>[19]</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>80</td>
<td>100</td>
<td>1</td>
<td>20.97</td>
<td>10.00</td>
<td>1</td>
<td>[19]</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>80</td>
<td>100</td>
<td>100</td>
<td>11.29</td>
<td>5.53</td>
<td>0</td>
<td>[19]</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>80</td>
<td>150</td>
<td>100</td>
<td>11.29</td>
<td>5.53</td>
<td>0</td>
<td>[19]</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>80</td>
<td>250</td>
<td>100</td>
<td>11.29</td>
<td>5.53</td>
<td>0</td>
<td>[19]</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>30</td>
<td>100</td>
<td>100</td>
<td>11.29</td>
<td>5.53</td>
<td>0</td>
<td>[19]</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>30</td>
<td>200</td>
<td>100</td>
<td>11.29</td>
<td>5.53</td>
<td>0</td>
<td>[19]</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>80</td>
<td>100</td>
<td>34</td>
<td>14.19</td>
<td>6.17</td>
<td>0</td>
<td>[19]</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Data sets of groutability collected from [7,42,57,58].
4. Application of ANFIS to predict groutability

As discussed in section 4, the Adaptive Neuro-Fuzzy Inference System is a FIS that has to be initialized one at a time. There are techniques for structure recognition to establish a prime ANFIS structure prior to the application of any parameter-adjusting mechanism.

Structure recognition in fuzzy modeling includes the following parts [52]:
- Selecting pertinent input parameters;
- Dividing of input domain;
- Quantifying MFs for every input variable;
- Quantifying if-then rules;
- Ancestor statement of fuzzy rules;
- Result statement of fuzzy rules.
- Selecting primary factors of MFs

In this paper, in order to recognize the ancestor MFs, the following three ANFIS models have been utilized:
1. Grid partitioning (GP);
2. Subtractive Clustering Method (SCM);

4.1 Grid Partitioning (GP)

Autonomous divisions of every ancestor parameter are proposed in the Grid Partitioning (GP) method [52]. For the purpose of developing a model, the MFs of all ancestor parameters can be defined by an expert and by applying former experience and understanding. The essence of linguistic phrases in an apparent text is represented by means of designed MFs. However, in plenty of organizations, particular understanding is not accessible in these divisions. In this method, the spaces of ancestor parameters are easily divided into a number of MFs with equal space and shape. The MFs parameters can be optimized by means of the available input-output data.

4.2 Subtractive Clustering Method (SCM)

SCM was originally suggested by [54]. In this method, data are considered as nominations for the central point of a cluster. The SCM algorithm is presented below:

First, imagine a set of \( n \) data points \( \{x_1, x_2, ..., x_n\} \) in an M-dimensional domain. SCM calculates this matter utilizing data points as nominations for cluster centers. In view of the fact that every data point is a nomination for a cluster center, the function of density measure at a given data point \( x_i \) is presented as Equation 14.

\[
D_i = \sum_{j=1}^{x} \exp \left( -\frac{\|x_i - x_j\|^2}{\tau_j} \right)
\]

(14)

in which \( \tau_j \) is a positive constant. If a data point has lots of neighboring data, it will have a high-density value. The neighborhood is defined by the radius \( \tau_j \); all data points beyond this radius grant just moderately to the density measure. When the density function of every data point is calculated, the maximum density will be chosen as the first cluster center. Following the calculation of the density function of every data point, with the maximum density measure is chosen as the first cluster center. If \( x_{i_1} \) - the first cluster center- is selected as the data containing a higher density amount \( D_{i_1} \), the density measurement of every data point \( x_i \) will be updated as Equation 15.

\[
D_i = D_{i_1} - D_{i_1} \exp \left( -\frac{\|x_i - x_{i_1}\|^2}{\tau_{i_1}} \right)
\]

(15)

in which \( \tau_{i_1} \) is an affirmative fixed amount. When the density computation was updated for every data point, the subsequent cluster center \( x_{i_2} \) will be chosen. Finally, the whole density computations for data points are updated reiteratively. This calculation is iterated up to create an adequate quantity of cluster centers.

4.3 Fuzzy C-means clustering Method (FCM)

It should be noted that FCM is based on Hard C-means clustering (HCM). The main difference between FCM and HCM is that in FCM all data is placed in a cluster with a degree of membership, provided that in HCM all data is rightly placed in a specific cluster or not. This algorithm was proposed by [18]. FCM divides a set of \( n \) vector \( x_i \), \( i = 1, 2, ..., n \) into fuzzy sets, and identifies the center of each cluster by minimizing the cost function of incongruity measure. \( i = 1, 2, ..., c \) are arbitrarily chosen from the \( n \) points. The FCM algorithm is explained as follows: at first, the cluster center \( c_i \), \( i = 1, 2, ..., c \) are selected arbitrarily out of \( n \) data points \( \{x_1, x_2, ..., x_n\} \). The elements of membership
matrix U can vary from 0 to 1. However, as presented in Equation 16, the aggregate of membership degree of a data point to all clusters is unity.

$$\sum_{j=1}^{n} u_{ij} = 1 \quad \forall j = 1, \ldots, n$$  \hspace{1cm} (16)

Then, the membership matrix U is computed utilizing Equation 17.

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} d_{ik}^{2m-1}}$$ \hspace{1cm} (17)

In which $d_{ik} = \lVert c_i - x_j \rVert$ is the Euclidean interval of the $i$th cluster center from the $j$th data point, $d_{ik} = \lVert c_i - x_j \rVert$ is the Euclidean interval of the $k$th cluster center from the $j$th data point, and $m$ is the fuzziness index.

Subsequently, the cost function is computed based on Equation 18. If the cost function is less than a specific limit, the process will stop.

$$J(U, c_1, \ldots, c_c) = \sum_{j=1}^{n} I_j = \sum_{j=1}^{n} \sum_{i=1}^{c} u_{ij}^m d_{ij}^2$$ \hspace{1cm} (18)

Finally, new fuzzy cluster centers $c_i$, $i = 1, 2, \ldots, c$ are computed utilizing Equation 19.

$$c_i = \sum_{j=1}^{n} u_{ij}^m x_j \quad \sum_{j=1}^{n} u_{ij}^m$$ \hspace{1cm} (19)

5. Groutability prediction utilizing ANFIS model

In this study, three types of modeling producer were carried out to foretell groutability for granular soil samples. ANFIS models including GP, SCM, and FCM were applied to predict the groutability of granular soil samples with cement-based grouts. According to the GP model, the fuzzy rule-base for the Sugeno model was specified in MATLAB software utilizing the ANN model presented in [52]. The schematic design structure of ANFIS is illustrated in Fig. 3. Calculations for this study were conducted using Matlab R2010a software and all runs were carried out on a Laptop with Intel Core 2 Duo @ 2 GHz CPU and 2 GB RAM. Depending on the chosen method for the ANFIS models, the run time of each model varied from 2 to 5 minutes.

![ANFIS Diagram](image)

**Fig. 3.** Shows a schematic of design structures for the GP, SCM and FCM ANFIS models.

As illustrated in Fig. 3, the parameters of W/C, Dr, P, FC, N1, and N2 were introduced to the ANFIS models as input parameters and just the ‘a’ output parameter, i.e. groutability, was acquired. Specifications of groutability data are presented in Table 3.

**Table 3.** Specifications of groutability data applied to ANFIS models.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Symbols</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>Water/cement ratio of the grout</td>
<td>W/C</td>
<td>1-6</td>
</tr>
<tr>
<td></td>
<td>Relative density of the host soil (%)</td>
<td>Dr</td>
<td>27-80</td>
</tr>
<tr>
<td></td>
<td>Grouting pressure (kPa)</td>
<td>P</td>
<td>50-690</td>
</tr>
<tr>
<td></td>
<td>Content of soil passing through a 0.6 mm sieve (%)</td>
<td>FC</td>
<td>1-100</td>
</tr>
</tbody>
</table>

**Table 4.** Types of parameters and the amounts applied to train the ANFIS models.

<table>
<thead>
<tr>
<th>ANFIS parameter type</th>
<th>ANFIS (GP)</th>
<th>ANFIS (SCM)</th>
<th>ANFIS (FCM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MF type</td>
<td>Gaussian</td>
<td>Gaussian</td>
<td>Gaussian</td>
</tr>
<tr>
<td>Output MF</td>
<td>Linear</td>
<td>Linear</td>
<td>Linear</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>161</td>
<td>345</td>
<td>177</td>
</tr>
<tr>
<td>Number of linear parameters</td>
<td>448</td>
<td>168</td>
<td>84</td>
</tr>
<tr>
<td>Number of nonlinear parameters</td>
<td>24</td>
<td>288</td>
<td>144</td>
</tr>
<tr>
<td>Total number of parameters</td>
<td>472</td>
<td>456</td>
<td>228</td>
</tr>
<tr>
<td>Number of training data pairs</td>
<td>62</td>
<td>62</td>
<td>62</td>
</tr>
<tr>
<td>Number of testing data set</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Number of fuzzy rules</td>
<td>64</td>
<td>24</td>
<td>12</td>
</tr>
</tbody>
</table>

The training phase for each model applied 62 data collection, randomly selected from a database. For the purposes of having a more effective training phase, the normalization of datasets was carried out to the domain of [0,1] by Equation 20.

$$x_{normalized} = \frac{x - x_{min}}{x_{max} - x_{min}}$$ \hspace{1cm} (20)

in which $x$ is an input variable and $x_{min}$ and $x_{max}$ are minimum and maximum values of each variable, respectively.

In the grid partitioning ANFIS model, for every input parameter, two Gaussian fuzzy collections were adapted and 64 fuzzy rules were created in its network structure. Detailed properties of the GP model are presented in Table 4.

**Fig. 4.** Ultimate MFs of the grid partitioning method.

The ANFIS model based on SCM, includes 24 rules and its specification is presented in Table 4. The ultimate (trained) MFs of input variables are illustrated in Fig. 5.
In order to calculate the prediction capability of the proposed models, an efficiency evaluation of the ANFIS models has been made using the test data. The test data are almost 25 data that were not utilized in the learning process of the model. It should be noticed that the test data were completely different in each arrangement. The results of applying these three datasets to the models are presented in Table 6.

Table 6. Cross-validation and the average performance of the ANFIS models.

<table>
<thead>
<tr>
<th>Data sets</th>
<th>ANFIS Model</th>
<th>Case Records</th>
<th>Incorrect prediction</th>
<th>Correct prediction</th>
<th>Failure rate (%)</th>
<th>Success rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SCM</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>SCM</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>SCM</td>
<td>25</td>
<td>2</td>
<td>25</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>SCM</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>SCM</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>SCM</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>SCM</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>SCM</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td>SCM</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>SCM</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

Cross-validation indicated that the models worked very well with different input data. Therefore, the overfitting could not be a problem and the models could be considered as comprehensive. In general terms, the outcomes presented in Table 6 and Fig. 7 display the remarkable possibility of the Subtractive Clustering Method with an average success rate of 100% in correctly foretelling the groutability of granular soil samples with the cement grout. However, other methods including Grid Partitioning and Fuzzy C-means clustering Method predicted the groutability of granular soil with success rate averages of 97.3% and 93.3%, respectively.
As seen in Fig. 8, several terminologies are used to calculate confusion matrix as follows: True Positive (TP) means that the model predicted positive values and it is true. True Negative (TN) indicates that the model predicted negative values and it is true. False Positive (FP) reveals that the model predicted positive values and it is false and finally, False Negative (FN) means that the model predicted negative values and it is false. The confusion matrix is extremely useful for measuring Recall, Precision, Accuracy, and F-Means according to the following formulas:

Recall = \frac{TP}{TP + FN} \quad (23)

Precision = \frac{TP}{TP + FP} \quad (24)

Accuracy = \frac{TP + TN}{Total} \quad (25)

F - Measure = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (26)

The calculated parameters are presented in Table 7.

<table>
<thead>
<tr>
<th>ANFIS Model</th>
<th>Recall</th>
<th>Precision</th>
<th>Accuracy</th>
<th>F-means</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP</td>
<td>0.63</td>
<td>1.00</td>
<td>0.96</td>
<td>0.77</td>
</tr>
<tr>
<td>SCM</td>
<td>0.60</td>
<td>1.00</td>
<td>1.00</td>
<td>0.75</td>
</tr>
<tr>
<td>FCM</td>
<td>0.64</td>
<td>0.93</td>
<td>0.88</td>
<td>0.76</td>
</tr>
</tbody>
</table>

As shown in Table 7, the results of confusion matrix analyses are in a high agreement with the previous results of the succession rate formula.

7. Sensitivity analysis

Sensitivity analyses were performed to determine the most effective input parameters on the output parameter. To achieve this goal, the cosine amplitude method (CAM) was used [46,47,59–62]. To apply this method, all data were declared in an X-space. The data pairs were utilized to establish a data array X as described in Equation 27.

\[ X = \{X_1, X_2, \ldots, X_m\} \quad (27) \]

Every member \( X_i \) in the data order X is a vector with the length of m, which is presented in Equation 28.

\[ X_i = \{x_{i,1}, x_{i,2}, \ldots, x_{i,m}\} \quad (28) \]

Therefore, each data pair was considered as a point in an m-dimensional domain, whereby every data requires m-coordinates for a comprehensive explanation. Every element of \( \tau_{ij} \) in Equation (25), presents a pairwise comparison of two data pairs.

The strength of the relation between the data pairs, \( x_i \) and \( x_j \) is described by the membership amount introducing the strength and is presented in Equation (29).

\[ \tau_{ij} = \sum_{k=1}^{m} x_{ik} \cdot x_{jk} / \left( \sum_{k=1}^{m} x_{ik}^2 \cdot \sum_{k=1}^{m} x_{jk}^2 \right) \quad (29) \]

These strengths of relations' quantities, i.e. \( \tau_{ij} \), between the input parameters and groutability for three types of the ANFIS models are illustrated in Fig. 9.

As illustrated in Fig. 8, the most effective variables on groutability were N1, N2, and P respectively. Furthermore, the least effective parameter on groutability was W/C.

8. Conclusion

Three new ANFIS models were improved to predict the groutability in geoscience operations based on laboratory tests. The models were boosted on the strength of experts’ knowledge and the datasets were collected from the literature. In this analogical investigation, three Takagi–Sugeno Fuzzy Inference Systems; GP, SCM, and FCM methods were applied to evaluate the groutability of granular soils. Some parameters included the proportion of water to cement in the grout (W/C), the relative density of soil (Dr), the soil percentage moving through a 0.6 mm sieve (FC), the injection pressure of grout (P), N1 = D10 soil / d10 grout and N2 = D60 soil / d60 grout were considered as inputs for models. To achieve favorable outputs, the training of ANFIS models using the datasets was carried out and the neuro-fuzzy parameters were obtained utilizing an ANN. The results of this study indicated that Subtractive Clustering Method with an average success rate and accuracy of 100% was the most effective method to foretell the groutability of granular soil samples in comparison with the other two methods. However, other methods including Grid Partitioning and Fuzzy C-means clustering with average success rates of 97.3% and 93.3%, respectively, were able to reasonably well foretell the groutability of granular soils. The comparison displayed the excellence of the ANFIS model on the strength of the SCM algorithm. In addition, it showed that the GP algorithm had better results in relation to the FCM algorithm. Furthermore, it was observed that the simulation results of the ANFIS models were close to the real measured values. Moreover, the three-fold cross validation indicated that the models worked very well with different input data. Therefore, overfitting cannot be considered as a problem and the models could be regarded as comprehensive. Consequently, sensitivity analyses were carried out using the cosine amplitude method (CAM) and indicated that based on the Grid Partitioning Method, the most effective factors in order of importance are N1, N2, and P respectively and W/C, FC and Dr respectively relate to the least effective parameter on groutability of granular soil. Moreover, according to the Subtractive Clustering Method, the most effective parameters in order of significance are N2, N1, and P respectively and W/C, Dr and FC are the least effective parameters on groutability, respectively. Finally, on the strength of the Fuzzy C-means clustering Method N1, N2, and P are respectively the most effective parameters and W/C, Dr and FC were the least effective parameters on the groutability granular soils. Although the three tested methods indicated the same parameters as the most/least effective parameters, the order of importance in each was different. The order of the most effective parameters in the GPM and FCM methods is N1, N2, and P respectively; however, the order of the least effective parameters in the SCM and FCM methods was W/C, Dr and FC.
REFERENCES

[39] Yilmaz İ, Kaynar O. Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils.


