



Sharif University of Technology
Scientia Iranica
Transactions A: Civil Engineering
<http://scientiairanica.sharif.edu>



Research Note

Liquefaction prediction using rough set theory

M. Arabani and M. Pirouz*

Department of Civil Engineering, University of Guilan, Rasht, P.O. Box 3756, I.R. Iran.

Received 7 January 2015; received in revised form 15 July 2017; accepted 16 October 2017

KEYWORDS

Earthquake;
 Liquefaction;
 Ground failure;
 Data classification;
 Rough sets;
 Uncertainties;
 Decision rules.

Abstract. Evaluation of liquefaction is one of the most important issues of geotechnical engineering. Liquefaction prediction depends on many factors, and the relationship between these factors is non-linear and complex. Different authors have proposed different methods for liquefaction prediction. These methods are mostly based on statistical approaches and neural network. In this paper, a new approach based on rough set data mining procedure is presented for liquefaction prediction. The rough set theory is a mathematical approach to the analysis of imperfect knowledge or unclear description of objects. In this approach, decision rules are derived from conditional attributes in rough set analysis, and the results are compared with actual field observations. The results of this study demonstrate that using this method can be helpful for liquefaction prediction and can reduce unnecessary costs in the site investigation process.

© 2019 Sharif University of Technology. All rights reserved.

1. Introduction

Liquefaction is one of the major causes of destruction during an earthquake. This phenomenon occurs when soil loses much of its strength or stiffness for a quite short time, but long enough to cause many disasters. This can happen when pore water pressure increases rapidly as a result of the reduction of pore volume and soil densification when saturated soil is subjected to seismic shaking. When the excess pore pressure becomes equal to the initial effective stress, the inter-granular effective stress is negligible and the liquefaction occurs [1].

The evaluation of liquefaction potential of a site is an important issue in geotechnical earthquake engineering. Liquefaction potential is generally determined through in situ and laboratory investigations that are usually expensive and time consuming.

Seed and Idriss [2] developed a simplified procedure from empirical evaluations' field studies and laboratory tests. In this method, the number of cycles and the shear stress level are the most important requirements. In the study by Seed et al. [3], the equivalent shear stress was selected as 65% of the maximum shear stress, while Ishihara and Yasuda [4] considered the equivalent shear stress as 57% for 20 cycles of loading. This method gradually proceeded towards completion in other works by Seed [5], Seed and Idriss [6], and Seed et al. [7]. This method later became completed more by Youd et al. [8] and was presented as a summary report, which is commonly referred to as NCEER workshop method.

Dobry et al. [9] presented the cyclic strain method for prediction of pore water pressure buildup and liquefaction of sands during earthquakes. Many researchers have studied energy-based liquefaction assessment approaches [10-13].

Researchers later adopted artificial neural network approaches. Rahman and Wang [14] developed a fuzzy neural network model for liquefaction prediction. Chen et al. [15] assessed the liquefaction probability using an energy-based method with a neural network

*. Corresponding author. Tel.: +98 1333690270
 E-mail addresses: Arabani@guilan.ac.ir (M. Arabani);
 mpr643@msc.guilan.ac.ir (M. Pirouz)

procedure. Baziar and Jafarian [16] developed an ANN model to correlate some of the soil parameters with the strain energy required for liquefaction triggering. Hanna et al. [17] proposed a neural network model for the two major earthquakes in Turkey and Taiwan.

Khozaghi and Choobbasti [18] predicted the liquefaction potential in the southeast of Tehran city in Iran using neural network approach. A trained ANN based on ant colony algorithm was proposed by Özbakır et al. [19]. Abbaszadeh Shahri [20] predicted liquefaction potential using different artificial neural network models during a case study.

In this paper, a new method is developed based on rough set theory, which is a powerful mathematical approach to data mining. Rough set theory is a mathematical approach to the analysis of a vague description of objects presented by a well-known mathematician, Pawlak [21]. Using this procedure, we can simply determine the decision rules by data mining. The decision rules are derived from conditional attributes in rough set analysis to account for data vagueness and uncertainty and to potentially reduce data collection needs. The advantage of this method is that it does not need any additional knowledge about the data, and the results of this study indicate that using this method can be effective and useful for the liquefaction prediction.

2. Rough set theory

Pawlak [21] developed the rough set theory for data mining and knowledge discovery. Rough set theory provides a powerful tool to discover rules in data tables; by eliminating the insignificant attributes, the computational complexity of data processing reduces.

The rough set method is based on the assumption that a certain amount of information expressed by some attributes can be associated with any object of the universe. Objects with the same description are indiscernible or similar with respect to the available information. The indiscernible relation forms a mathematical basis of the rough set theory. It partitions the universe with blocks of elementary sets that can be used to build knowledge about a world. In rough set method, the input information can be quantitative or qualitative data. The final results in this method are in the form of “if..., then...” decision rules using the most significant attribute [22,23].

The lower and upper approximations of a set (Figure 1) can be described as follows:

Let us consider an approximation space:

$$apr = (U, A), \quad (1)$$

where U is the universe that is a finite and non-empty set, and A is the set of attributes. The upper

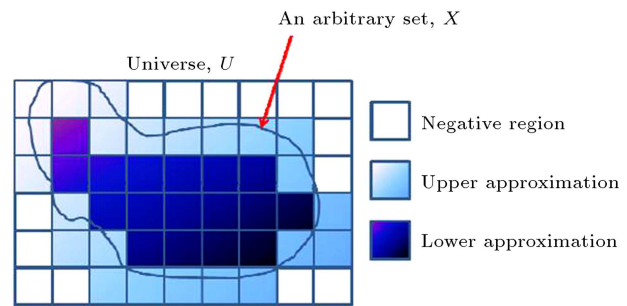


Figure 1. A set approximation of an arbitrarily set X in U [23].

approximation of X in A is:

$$\overline{apr}(A) = \{x|x \in U, U/ind(A) \cap X \neq \emptyset\}. \quad (2)$$

In addition, the lower approximation of X in A is:

$$\underline{apr}(A) = \{x|x \in U, U/ind(A) \subset X\}, \quad (3)$$

where,

$$\begin{aligned} U/ind(a) &= \{(x_i, x_j) \in U \times U, f(x_i, a) \\ &= f(x_j, a), \forall a \in A\}, \end{aligned} \quad (4)$$

where X is a subset of U .

The boundary can be represented as follows:

$$BN(A) = \overline{apr}(A) - \underline{apr}(A). \quad (5)$$

According to the approximation space, the reducts and the decision rules can be calculated. Given an information system, $I = (U, A)$.

The reduct, $RED(B)$, is a minimal set of attributes $B \subseteq A$, such that $r_B(U) = r_A(U)$ where:

$$r_B(U) = \frac{\sum \text{card}(\underline{B}(X_i))}{\text{card}(U)}, \quad (6)$$

denotes the quality of approximation of U by B .

A decision rule can be expressed as $\varphi \Rightarrow \theta$, where φ denotes the elementary condition and θ denotes the elementary decision. The advantage of the induction-based approaches is that it can provide the intelligible rules for decision-makers. These intelligible rules can help decision-makers realize the contents of data sets.

3. Application of the rough set theory in liquefaction prediction

In data analysis of the rough set theory, the main computational attempt to determine the relationship between the attributes and to find the deterministic rules in the information system. In this study, two different databases are used based on SPT [7] and CPT tests [24]. Application of the rough set theory needs to define the conditional attributes.

The attributes defined in Table 1 are shown to be the most appropriate for liquefaction study [14].

Table 1. Conditional attributes for liquefaction prediction.

Conditional attributes	Classification of attributes	Conditional attributes	Classification of attributes
(a) Earthquake magnitude	1-very high	(f) Cyclic shear stress ratio, τ/s'_{vo}	1-very high
	2-high		2-high
	3-fairly high		3-fairly high
	4- medium		4- medium
	5-fairly low		5-fairly low
	6-low		6-low
	7-very low		7-very low
(b) Vertical total overburden pressure, σ_{vo}	1-very high	(g) Fine content of the soil, F (%)	1-very high
	2-high		2-high
	3-fairly high		3-fairly high
	4- medium		4- medium
	5-fairly low		5-fairly low
	6-low		6-low
	7-very low		7-very low
(c) Vertical effective overburden pressure, σ'_{vo}	1-very high	(h) Median grain diameter of soil, D_{50}	1-very high
	2-high		2-high
	3-fairly high		3-fairly high
	4- medium		4- medium
	5-fairly low		5-fairly low
	6-low		6-low
	7-very low		7-very low
(d) Corrected SPT value (N_1) ₆₀ or q_c value from CPT	1-very high	(i) Critical depth of liquefaction, D_{cr}	1-very high
	2-high		2-high
	3-fairly high		3-fairly high
	4- medium		4- medium
	5-fairly low		5-fairly low
	6-low		6-low
	7-very low		7-very low
(e) Acceleration ratio, a_{\max}/g	1-very high	(j) Water table depth, D_w	1-very high
	2-high		2-high
	3-fairly high		3-fairly high
	4- medium		4- medium
	5-fairly low		5-fairly low
	6-low		6-low
	7-very low		7-very low

Table 2(a). SPT-based data inspection for prediction of liquefaction.

Sites	Decision table										Decision levels**	Results using simplified procedure [8]	Fuzzy neural network [14]
	Conditional attributes*												
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>			
<i>S</i> ₁	5	5	5	6	6	6	6	7	5	6	0	No	0.3
<i>S</i> ₂	5	1	2	7	6	6	5	7	1	6	0	No	0.3
<i>S</i> ₃	6	6	5	1	7	7	5	7	6	5	0	No	0.3
<i>S</i> ₄	6	6	5	7	7	7	5	7	6	5	0	No	0.3
<i>S</i> ₅	6	5	5	4	7	7	5	7	5	5	0	No	0.3
<i>S</i> ₆	6	7	7	6	6	5	2	7	7	7	0	No	0.7*
<i>S</i> ₇	6	5	6	5	6	5	6	7	5	7	0	No	0.3
<i>S</i> ₈	6	6	6	6	5	5	6	7	6	6	1	Yes	0.7
<i>S</i> ₉	6	6	5	4	5	5	6	7	6	5	0	No	0.3
<i>S</i> ₁₀	6	6	5	5	5	5	6	7	6	5	0	No	0.4
<i>S</i> ₁₁	6	5	5	7	5	5	1	7	5	5	1	Yes	0.7
<i>S</i> ₁₂	6	6	6	7	5	5	3	7	6	5	1	Yes	0.7
<i>S</i> ₁₃	6	6	5	6	5	5	1	7	6	5	1	Yes	0.7
<i>S</i> ₁₄	6	7	5	7	6	6	3	7	6	5	1	Yes	0.3*
<i>S</i> ₁₅	6	6	6	5	6	6	6	7	6	5	0	No	0.3
<i>S</i> ₁₆	6	6	6	6	6	6	5	7	6	6	0	No	0.3
<i>S</i> ₁₇	2	3	3	4	6	5	7	7	3	6	0	No	0.7*
<i>S</i> ₁₈	2	5	3	5	4	5	7	7	5	6	1	Yes	0.7
<i>S</i> ₁₉	2	5	6	5	6	5	7	7	5	6	1	Yes	0.3*
<i>S</i> ₂₀	2	5	5	5	6	5	6	7	5	6	0	No	0.7*
<i>S</i> ₂₁	2	2	3	2	6	5	6	7	3	6	0	No	0.3
<i>S</i> ₂₂	2	4	4	3	6	5	6	7	4	6	0	No	0.3
<i>S</i> ₂₃	2	5	5	6	5	5	6	7	5	6	1	Yes	0.7
<i>S</i> ₂₄	2	3	4	5	6	5	6	7	4	6	1	Yes	0.7
<i>S</i> ₂₅	2	4	4	5	6	5	6	7	4	6	1	Yes	0.7
<i>S</i> ₂₆	2	3	4	5	6	5	6	7	3	6	1	Yes	0.7
<i>S</i> ₂₇	2	6	6	4	6	5	6	7	7	6	1	Yes	0.7

* The conditional attributes are defined in Table 1.

** Decision level = 1: liquefaction likely; Decision level = 0: Liquefaction unlikely.

Consequently, each conditional attribute is provided with 7 classes: 1) very high, 2) high, 3) fairly high, 4) medium, 5) fairly low, 6) low, and 7) very low. The decision attribute is classified in two levels: level 1 shows that the liquefaction potential exists and level 2 indicates that there is no liquefaction potential.

The classification of all attributes has been carried

out by defining the specific levels and assigning a code to each specified attribute in the rows of the Table 1 [25].

Tables 2(a) and 2(b) show the class numbers of conditional attributes and decision levels [14] for two different databases for liquefaction potential based on SPT [7] and CPT tests [24]. These tables, which are called the decision tables, show the relationship

Table 2(b). CPT-based data inspection for prediction of liquefaction.

Sites	Decision table									Decision levels**	Results using simplified procedure [8]	Fuzzy neural network [14]
	Conditional attributes*											
	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>h</i>	<i>i</i>	<i>j</i>			
<i>S</i> ₁	3	6	6	7	7	6	7	6	6	1	Yes	0.7
<i>S</i> ₂	3	6	6	7	7	7	7	6	6	1	Yes	0.7
<i>S</i> ₃	3	5	5	7	7	6	7	4	6	1	Yes	0.7
<i>S</i> ₄	3	6	6	7	7	7	7	6	6	1	Yes	0.7
<i>S</i> ₅	3	3	4	7	7	6	7	3	6	0	No	0.6*
<i>S</i> ₆	3	4	5	7	7	6	7	4	6	1	Yes	0.7
<i>S</i> ₇	3	6	6	6	6	6	7	6	6	1	Yes	0.7
<i>S</i> ₈	3	5	6	6	6	6	7	5	6	1	Yes	0.7
<i>S</i> ₉	3	5	6	7	6	6	7	5	6	1	Yes	0.7
<i>S</i> ₁₀	3	5	5	7	6	6	7	4	6	1	Yes	0.7
<i>S</i> ₁₁	3	4	5	6	6	6	7	4	6	1	Yes	0.7
<i>S</i> ₁₂	5	6	6	3	4	3	7	6	6	0	No	0.3
<i>S</i> ₁₃	5	6	6	7	4	3	7	6	6	1	Yes	0.7
<i>S</i> ₁₄	5	6	6	6	4	3	7	6	6	0	No	0.7*
<i>S</i> ₁₅	5	7	7	7	6	5	7	7	7	1	Yes	0.7
<i>S</i> ₁₆	5	6	7	6	6	5	7	5	7	1	Yes	0.7
<i>S</i> ₁₇	5	6	6	2	2	2	7	6	6	0	No	0.3
<i>S</i> ₁₈	5	6	6	1	2	2	7	6	6	0	No	0.3
<i>S</i> ₁₉	5	5	6	6	2	1	7	6	6	1	Yes	0.7
<i>S</i> ₂₀	5	6	6	5	2	2	7	6	6	0	No	0.7*
<i>S</i> ₂₁	5	6	6	6	2	2	7	6	6	0	No	0.7*
<i>S</i> ₂₂	1	6	6	5	6	6	7	6	6	0	No	0.3
<i>S</i> ₂₃	1	6	6	4	6	6	7	6	6	0	No	0.3
<i>S</i> ₂₄	1	5	6	4	6	6	7	5	6	0	No	0.3
<i>S</i> ₂₅	1	6	6	7	6	6	7	6	6	1	Yes	0.7
<i>S</i> ₂₆	1	6	6	6	6	6	7	6	6	1	Yes	0.6
<i>S</i> ₂₇	1	5	6	6	6	6	7	5	6	1	Yes	0.6
<i>S</i> ₂₈	1	5	5	5	6	6	7	5	6	1	Yes	0.4*

* The conditional attributes are defined in Table 1.

** Decision level = 1: Liquefaction likely; Decision level = 0: Liquefaction unlikely.

between the class numbers of the conditional attributes of each site and its decision attribute [26].

4. Minimal decision algorithm and the decision rules

In order to find the minimal decision table and extract the decision rules, the compatibility of the decision levels with the conditional attributes should be evaluated [27].

The next step is to find the non-deterministic rules in the table or the sites with the same class in all conditional attributes, yet different in decision levels. The non-deterministic rules in the table imply that the number of conditional attributes is not sufficient, and new conditional attributes should be added [28,29].

There are not any non-deterministic rules in Tables 2(a) and 2(b), and the decision levels are subordinate to the conditional attributes.

The process of extracting the decision rules from

Table 3(a). Arrangements of conditional attributes (SPT-based data).

Case number	Conditional attributes				
1	<i>d</i>	<i>f</i>	<i>g</i>	<i>i</i>	-
2	<i>d</i>	<i>e</i>	<i>g</i>	<i>i</i>	-
3	<i>a</i>	<i>d</i>	<i>g</i>	<i>i</i>	-
4	<i>b</i>	<i>d</i>	<i>e</i>	<i>g</i>	-
5	<i>a</i>	<i>b</i>	<i>d</i>	<i>g</i>	-
6	<i>c</i>	<i>d</i>	<i>f</i>	<i>g</i>	-
7	<i>c</i>	<i>d</i>	<i>e</i>	<i>g</i>	-
8	<i>c</i>	<i>d</i>	<i>f</i>	<i>j</i>	-
9	<i>c</i>	<i>d</i>	<i>e</i>	<i>j</i>	-
10	<i>a</i>	<i>c</i>	<i>d</i>	<i>e</i>	-
11	<i>a</i>	<i>c</i>	<i>d</i>	<i>f</i>	-
12	<i>a</i>	<i>c</i>	<i>d</i>	<i>g</i>	-
13	<i>b</i>	<i>d</i>	<i>f</i>	<i>g</i>	<i>j</i>

Table 3(b). Arrangements of conditional attributes (CPT-based data).

Case number	Conditional attributes			
1	<i>d</i>	<i>f</i>	<i>i</i>	-
2	<i>c</i>	<i>d</i>	<i>f</i>	-
3	<i>b</i>	<i>d</i>	<i>e</i>	-
4	<i>b</i>	<i>d</i>	<i>f</i>	-
5	<i>a</i>	<i>b</i>	<i>d</i>	<i>j</i>
6	<i>a</i>	<i>b</i>	<i>d</i>	<i>i</i>
7	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>

the data tables is a trial and error process in which the conditional attributes are removed one by one, and then the decision table is checked for any contradiction. In this way, the significant conditional attributes in the diagnoses would be determined (Tables 3(a) and 3(b)). This process is used for Tables 2(a) and 2(b); considering the extracted rules (Tables 4(a) and 4(b)), it can be seen that the most significant attributes are *d*, *e*, *c*, *a*, *b*, *g*, and *f*, respectively.

5. Results of fuzzy neural network models

The results of fuzzy neural network model from Rahman and Wang [14] are shown in Tables 2(a) and 2(b). to compare with the results of the rough set theory for the same databases. In the fuzzy neural network developed by Wang and Rahman [14], a range of linguistic labels were used as follows:

0.0-0.2: liquefaction *highly unlikely*;

0.2-0.4: liquefaction *unlikely*;

0.4-0.6: not enough information to decide;

0.6-0.8: liquefaction *likely*;

0.8-1.0: liquefaction *highly likely*.

The results of the liquefied case histories appear to have an output range of 0.6-0.7, and the liquefaction prediction indices for liquefaction unlikely cases have a range of 0.3-0.4; however, there were some misclassifications in 5 cases in each table. The decision levels used in rough set analysis in Tables 2(a) and 2(b) are a result of real data from past earthquakes with the same classifications as the simplified method proposed by Youd et al. [8].

6. Comparison of rough set with stepwise regression analyses

The regression analysis is a simple method for finding the relationship between variables. In this method, the relationship is expressed in the form of an equation connecting the dependent variable to the predictor variables [30].

The response variable is denoted by Y and the set of predictor variables is shown by X_1, X_2, \dots, X_p , where p indicates the number of predictor variables.

The relationship between Y and X_1, X_2, \dots, X_p , is approximated by the following equation:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 x_2 + \dots + \beta_p X_p + \varepsilon, \quad (7)$$

where $\beta_0, \beta_1, \dots, \beta_p$ are called the regression coefficients, and ε represents the error of approximation.

In this paper, attempt is made to find the effective attributes in liquefaction prediction; therefore, the dependent parameters should be the indicators of the decision levels and the independent parameters should represent the characteristics of the sites.

It is practically difficult or impossible to use all the parameters to form a regression equation to predict the liquefaction potential of a site. Consequently, the stepwise method is used to select the most suitable combinations of attributes. In the stepwise procedure, different parameters are used to develop the best correlation with the highest value of R^2 .

For this reason, first, the value of the correlation coefficient between each independent parameter and dependent variables is estimated.

The independent parameters should be added one by one to the first parameter in the form of a regression equation with two independent variables; in each step, the value of R^2 is evaluated. This process continues until the addition of another independent parameter to the model has a negligible effect on R^2 . Accordingly, the presented parameters in the linear regression equation acquired by the stepwise procedure are regarded as the most significant defined parameters

Table 4(a). Rules generated by rough set analysis (based on SPT testing data).

	Deterministic rules	Satisfactory cases
Rule 1	If σ'_{vo} is medium or fairly high and (N_1) ₆₀ is fairly low, then the liquefaction occurs c in ($\{3,4\}$) and ($d = 5$) \Rightarrow (Decision = 1)	$\{S_{18}, S_{24}, S_{25}, S_{26}\}$
Rule 2	If (N_1) ₆₀ is low or very low and a_{\max}/g is fairly low, then the liquefaction occurs d in (6,7) and ($e = 5$) \Rightarrow (Decision = 1)	$\{S_8, S_{11}, S_{12}, S_{13}, S_{23}\}$
Rule 3	If the earthquake magnitude is high and σ'_{vo} is low, then the liquefaction occurs ($a = 2$) and ($c = 6$) \Rightarrow (Decision = 1)	$\{S_{19}, S_{27}\}$
Rule 4	If fine content of the soil (F (%)) is fairly high, then the liquefaction occurs ($g = 3$) \Rightarrow (Decision = 1)	$\{S_{14}\}$
Rule 5	If σ_{vo} is very high or high, then the liquefaction does not occur b in (1,2) \Rightarrow (Decision = 0)	$\{S_2, S_{21}\}$
Rule 6	If σ_{vo} is fairly low and a_{\max}/g is low and F (%) is low, then the liquefaction does not occur ($b = 5$) and ($e = 6$) and ($g = 6$) \Rightarrow (Decision = 0)	$\{S_1, S_7, S_{20}\}$
Rule 7	If (F (%)) is fairly low or low and water table depth (D_w) is fairly low, then the liquefaction does not occur g in (5,6) and ($j = 5$) \Rightarrow (Decision = 0)	$\{S_3, S_4, S_5, S_9, S_{10}, S_{15}\}$
Rule 8	If (N_1) ₆₀ is fairly high, then the liquefaction does not occur ($d = 3$) \Rightarrow (Decision = 0)	$\{S_{22}\}$
Rule 9	If (F (%)) is low and a_{\max}/g is low, then the liquefaction does not occur ($d = 6$) and ($e = 6$) \Rightarrow (Decision = 0)	$\{S_1, S_6, S_{16}\}$
Rule 10	If σ_{vo} is fairly high and σ'_{vo} is fairly high, then the liquefaction does not occur ($b = 3$) and ($c = 3$) \Rightarrow (Decision = 0)	$\{S_{17}\}$
Quality of classification: 1		
Accuracy of approximation: 1		

that can be used for liquefaction prediction. The results of regression equations are shown in Tables 5(a) and 5(b).

The results show that the linear regression procedure cannot make an accurate equation for liquefaction prediction. The highest values of R^2 using all the attributes in the regression analysis are 58.9% for SPT-based data and 53.6%. Besides, the quality and the

accuracy of approximation in the rough set procedure are higher than those of the other procedures are.

7. Conclusions

The prediction of liquefaction is very important in geotechnical engineering. In this paper, a new approach was introduced to predict the liquefaction po-

Table 4(b). Rules generated by rough set analysis (based on CPT testing data).

	Deterministic rules	Satisfactory cases
Rule 1	If q_c is low or very low and a_{\max}/g is low, then the liquefaction occurs d in $(\{6, 7\})$ and $(e = 6) \Rightarrow (\text{Decision} = 1)$	$\{S_7, S_8, S_9, S_{10}, S_{11}, S_{15}, S_{16}, S_{25}, S_{26}, S_{27}\}$
Rule 2	If the earthquake magnitude is fairly high and σ_{vo} is medium, fairly low or low, then the liquefaction occurs $(a = 3)$ and $(b$ in $(\{4, 5, 6\})) \Rightarrow (\text{Decision} = 1)$	$\{S_1, S_2, S_3, S_4, S_6, S_7, S_8, S_9, S_{10}, S_{11}\}$
Rule 3	If τ/s'_{vo} is very high, then the liquefaction occurs $(f = 1) \Rightarrow (\text{Decision} = 1)$	$\{S_{19}\}$
Rule 4	If σ'_{vo} is fairly low, then the liquefaction occurs $(c = 5) \Rightarrow (\text{Decision} = 1)$	$\{S_{28}\}$
Rule 5	If the earthquake magnitude is fairly low and q_c is very low, then the liquefaction occurs $(a = 5)$ and $(d = 7) \Rightarrow (\text{Decision} = 1)$	$\{S_{13}\}$
Rule 6	If τ/s'_{vo} is high, then the liquefaction does not occur $(f = 2) \Rightarrow (\text{Decision} = 0)$	$\{S_{17}, S_{18}, S_{20}, S_{21}\}$
Rule 7	If q_c is fairly high or medium, then the liquefaction does not occur d in $(\{3, 4\}) \Rightarrow (\text{Decision} = 0)$	$\{S_{12}, S_{23}, S_{24}\}$
Rule 8	If σ'_{vo} is low and q_c is fairly low, then the liquefaction does not occur $(c = 6)$ and $(d = 5) \Rightarrow (\text{Decision} = 0)$	$\{S_{20}, S_{22}\}$
Rule 9	If σ_{vo} is fairly high or low, then the liquefaction does not occur $(b = 3) \Rightarrow (\text{Decision} = 0)$	$\{S_5\}$
Rule 10	If q_c is low and τ/s'_{vo} is fairly high, then the liquefaction does not occur $(d = 6)$ and $(f = 3) \Rightarrow (\text{Decision} = 0)$	$\{S_{14}\}$
Quality of classification: 1		
Accuracy of approximation: 1		

Table 5(a). Stepwise regression equations for SPT-based data.

Data	Regression equations	R^{2*}
SPT-based	Decision = $2.421 - 0.342e$	22.2%
	Decision = $2.665 - 0.332e - 0.070a$	29.5%
	Decision = $3.397 - 0.290e - 0.136a - 0.134g$	42.7%
	Decision = $2.519 - 0.234e - 0.148a - 0.107g + 0.092d$	48%
	Decision = $2.023 - 0.205e - 0.184a - 0.102g + 0.097d + 0.087b$	52.3%
	Decision = $2.957 - 0.194e - 0.205a - 0.093g + 0.111d + 0.082b - 0.176j$	55.7%
	Decision = $2.864 - 0.190e - 0.203a - 0.090g + 0.119d + 0.090i - 0.183j$	56.3%
	Decision = $3.578 - 0.238e - 0.215a - 0.091g + 0.102d - 0.028i - 0.267j + 0.161c^{**}$	58.7%
	Decision = $3.559 - 0.238e - 0.213a - 0.090g + 0.104d - 0.027b - 0.268j + 0.159c^{***}$	58.8%

* R^2 value resulting from the linear regression of all parameters (all 10 parameters) is 59.8%.

**The algorithm resulting from stepwise regression analysis.

***The regression analysis of attribute in rough set analysis decision rules.

Table 5(b). Stepwise regression equations for CPT-based data.

Data	Regression equations	R^{2*}
CPT-based	Decision = $-0.516 + 0.203d$	44.6%
	Decision = $-0.609 + 0.174d + 0.052f$	47.5%
	Decision = $-1.524 + 0.177d + 0.063f + 0.145c$	50.6%
	Decision = $-1.595 + 0.167d + 0.056f + 0.299c - 0.135i$	52.5%
	Decision = $-1.345 + 0.176d + 0.216f + 0.263c - 0.146i - 0.161e$	53.4%
	Decision = $-0.964 + 0.177d + 0.205f + 0.312c - 0.165i - 0.150e - 0.094j^{**}$	53.5%
	Decision = $-1.463 + 0.035a - 0.098b + 0.194c + 0.173d - 0.163e + 0.251f^{***}$	52.4%

* R^2 value resulting from the linear regression of all parameters (all 9 parameters) is 53.6%.

**The algorithm resulting from stepwise regression analysis.

***The regression analysis of attribute in rough set analysis decision rules

tential of sites. This method is based on the rough set theory, which is a powerful mathematical data mining procedure. In the presented approach, a data sampling of the major liquefaction records of different sites was used to extract appropriate rules for liquefaction prediction.

The advantage of the rough set procedure over other methods is that the decisions generated by this method are explicit, and the modeling process is not limited to restrictive assumptions.

The rough set approach reduces the complexity of the attribute space by finding the reducts, and this provides engineers with an additional valuable tool for finding essential attributes to predict the liquefaction potential of a site. This paper showed that the rough set theory could provide an appropriate algorithm to generate simple rules to predict liquefaction that can be helpful for engineers in primary judgments to make decisions about needed laboratory and in situ tests for accurate studies.

References

- Madabhushi, G., Knappett, J., and Haigh, S., *Design of Pile Foundation in Liquefiable Soils*, published by imperial collage press London WC2H 9HE (2010).
- Seed, H.B. and Idriss, I.M. "Simplified procedure for evaluating soil liquefaction potential", *Journal of the Soil Mechanics and Foundations Division, ASCE*, **97**, pp. 1249-1273 (1971).
- Seed, H.B., Idriss, I.M., Makdisi, F., and Banerjee, N. "Representation of irregular stress time histories by equivalent uniform stress series in liquefaction analyses", Report No. UCB/EERC-75/29, Earthquake Engineering Research Centre, U.C. Berkeley (1975).
- Ishihara, K., and Yasuda, S. "Sand liquefaction in hollow cylinder torsion under irregular excitation", *Soils and Found.*, **15**(1), pp. 45-59 (1975).
- Seed, H.B. "Soil liquefaction and cyclic mobility evaluation for level ground during Earthquakes", *Journal of Geotechnical Engineering, ASCE*, **105**, pp. 201-255 (1979).
- Seed, H.B., and Idriss, I.M., *Ground Motions and Soil Liquefaction During Earthquakes*, Earthquake Engineering Research, Institute Monograph (1982).
- Seed, H.B., Tokimatsu, K., Harder, L.F., and Chung, R.M. "The influence of SPT procedures in soil liquefaction resistance evaluations", *Journal of Geotechnical Engineering, ASCE*, **111**(12), pp. 1425-1445 (1985).
- Youd, T.L., Idriss, I.M. Andrus, R.D. Arango, I., Castro, G., Christian, J.T., Dobry, R., Liam Finn, W.D.L., Harder, L.F. Jr., Hynes, M.E., Ishihara, K., Koester, J.P., Liao, S.S.C., Marcuson, W.F., III, Martin, G.R., Mitchell, J.K., Moriwaki, Y., Power, M.S., Robertson, P.K., Seed, R.B., Stokoe, K.H., II, "Liquefaction resistance of soils: Summary report from the 1996 NCEER and 1998 NCEER/NSF workshops on evaluation of liquefaction resistance of soils", *J. of Geotech. Geoenviron. Eng.*, **127**(10), pp. 817-833 (republishation 137 with correct authors) (2001).
- Dobry, R., Ladd, R.S., Yokel, F.Y., Chung, R.M., and Powell, D. "Prediction of pore water pressure buildup and liquefaction of sands during earthquakes by the cyclic strain method", *Building Science Series*, National Bureau of Standards, US Department of Commerce, US Governmental Printing Office, **138**, Washington, DC (1982).
- Okada, N., and Nemat-Nasser, S. "Energy dissipation in inelastic flow of saturated cohesionless granular media", *Géotechnique*, **44**(1), pp. 1-19 (1994).
- Trifunac, M. "Empirical criteria for liquefaction in sands via standard penetration tests and seismic wave energy", *Soil Dynamics and Earthquake Engineering*, **14**, pp. 419-426 (1995).
- Kayen, R.E., and Mitchell, J.K. "Assessment of liquefaction potential during earthquakes by arias intensity", *ASCE, Journal of Geotechnical and Geoenvironmental Engineering*, **123**(12), pp. 1162-1174 (1997).
- Green, R. "Energy based evaluation and remediation of liquefiable soils", PhD Dissertation, Virginia Polytechnic Institute and State University (2001).

14. Rahman, M.S., and Wang, J. "Fuzzy neural network models for liquefaction prediction", *Soil Dynamics and Earthquake Engineering*, **22**(8), pp. 685-694 (2002).
15. Chen, Y.R., Hsieh, S.C., Chen, J.W., and Shih, C.C. "Energy-based probabilistic evaluation of soil liquefaction", *Soil Dynamics and Earthquake Engineering*, **25**(1), pp. 55-68 (2005).
16. Baziari, M.H., and Jafarian, Y. "Assessment of liquefaction triggering using strain energy concept and ANN model: Capacity energy", *Soil Dynamics and Earthquake Engineering*, **27**, pp. 1056-1072 (2007).
17. Hanna, A.M., Ural, D., and Saygili, G. "Neural network model for liquefaction potential in soil deposits using Turkey and Taiwan earthquake data", *Soil Dynamics and Earthquake Engineering*, **27**(6), pp. 521-540 (2007).
18. Khozaghi, S.S.H., and Choobasti, A.J.A. "Predicting of liquefaction potential in soils using artificial neural networks", *Electronic Journal of Geotechnical Engineering*, **12**, Bundle C (2007).
19. Baykasoglu, A., Cevik, A., Ozbakir, L., and Kulluk, S. "Generating prediction rules for liquefaction through data mining", *Expert Systems with Applications*, **36**, pp. 12491-12499 (2009).
20. Abbaszadeh Shahri, A. "Assessment and prediction of liquefaction potential using different artificial neural network models: A case study", *Geotechnical and Geological Engineering*, **34**(3) p. 807 (2016).
21. Pawlak, Z., *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Dordrecht, Boston (1991).
22. Hung, C.T., Chang, J.R., Lin, J.D., and Tzeng, G.H., *Rough Set Theory in Pavement Maintenance Decision*, Springer-Verlag Berlin Heidelberg (2009).
23. Lee, C.S. "A framework of adaptive T-S type rough-fuzzy inference systems (ARFIS)", PhD Thesis, The University of Western Australia (2009).
24. Stark, T.D., and Olson, S.M. "Liquefaction resistance using CPT and field case histories", *J. Geotech. Eng. Div, ASCE*, **121**(12) pp. 856-869 (1995).
25. Arabani, M. and Lashteh Neshaei, M.A. "Application of rough set theory as a new approach to simplify dams location", *Scientia Iranica*, **13**(2), pp. 152-158 (2006).
26. Lashteh Neshaei, M.A., and Pirouz, M. "Rough Sets theory in site selection decision making for water reservoirs", *Computational Methods in Civil Engineering*, **1**(1) pp. 85-94 (2010).
27. Arabani, M., and Pirouz, M. "Water treatment plant site location using rough set theory", *Environmental Monitoring and Assessment*, **188**(10) p. 552 (2016).
28. Arabani, M., Haghi, A.K., and Amani, B. "Making a decision between the rehabilitation and reconstruction of asphalt pavements using the rough-set theory", *Scientia Iranica, Transaction A: Civil Engineering*, **16**(2) pp. 116-125 (2009).
29. Arabani, M., and Amani, B. "Evaluating the parameters affecting urban trip-generation", *Iranian Journal of Science & Technology, Transaction B, Engineering*, **31**(B5), pp. 547-560 (2009).
30. Chattefuee S., and Hadi A.S., *Regression Analysis by Example*, 4th edition, John Wiley & Sons (2006).

Biographies

Mahyar Arabani holds a PhD in Civil Engineering, Road and Transportation from Iran University of Science and Technology (IUST), and an MSc in Civil Engineering, Road and Transportation Engineering from Iran University of Science and Technology (IUST), and a BSc in Civil Engineering from Iran University of Science and Technology (IUST). Professor Arabani is the author of over 200 original articles and several chapter books in English, and he has written several academic books in Persian. He has several registered inventions such as manufacturing device for asphalt adhesion evaluation for the first time in Iran, invention of asphalt with waste tire thread mesh reinforcement. It is apparent from his works that he has made valuable contributions to Civil Engineering sciences. Dr. Arabani is a Professor of Civil Engineering in Faculty of Engineering at the University of Guilan, Rasht, Iran.

Mojgan Pirouz holds an MSc in Geotechnical Engineering and a BSc in Civil Engineering both from the University of Guilan Rasht, Iran. She has written several papers published in English and Persian. Her research interests include computational methods in civil engineering, geotechnical engineering, road engineering, and earthquake engineering. Ms. Pirouz is a member of Guilan Organization for Engineering Order of Building (I.R. Iran) and has several years of experience in civil engineering as a consultant engineer.