SPT-BASED SOIL-LIQUEFACTION MODELS USING NONLINEAR REGRESSION ANALYSIS AND ARTIFICIAL INTELLIGENCE TECHNIQUES

MODELI UTEKOČINJENJA ZEMLJIN NA OSNOVI SPT Z UPORABO REGRESIJSKE ANALIZE IN TEHNIK UMETNE INTELIGENCE

Mehmet Cemal Acar *(corresponding author)*
Vocational College of Kayseri University,
Department of Construction
Kayseri, Turkey
E-mail: acarc@kayseri.edu.tr

Tülay Hakan
Civil Engineer (Geotechnic) DSİ 12. Bölge Müdürlüğü
Kayseri, Turkey

DOI https://doi.org/10.18690/actageotechslov.19.2.33-45.2022

Keywords

liquefaction, standard penetration test (SPT), ANN, ANFIS, NMRA

Ključne besede

utekočinjenje, standardni penetracijski preizkus (SPT), ANN, ANFIS, NMRA

Abstract

Saturated, cohesionless soils can temporarily lose their shear strength due to increased pore-water pressure under the effect of repetitive dynamic loads such as earthquakes. This event is defined as soil liquefaction and causes significant damage to structures. The liquefaction potential of soils depends on many soil parameters obtained in the field and from laboratory tests. In this study new models have been developed to estimate the liquefaction potential of cohesionless soils. For this purpose, 837 soil data sets were collected to calculate the liquefaction potential with nonlinear multiple regression and artificial intelligence in the cities of Kayseri and Erzincan. The models based on Nonlinear Multiple Regression Analysis, Artificial Neural Networks, and Adaptive Neuro-Fuzzy-Inference System techniques were compared with the results of the simplified method. Determination coefficients (R^2) and various error rates were calculated for the performance-evaluation criteria of the models. The proposed ANN model effectively found the complex relationship between the soil and the input parameters and predicts the liquefaction potential more accurately than other methods. It has an overall success rate of 90 percent and the lowest mean absolute error rate of 0.024. With the improvement of existing methods, new models have been introduced to estimate the liquefaction probability of soils.

Izvleček

Zasičene nekoherentne zemljine lahko zaradi povečanega pritiska vode v porah pod vplivom ponavljajočih se dinamičnih obremenitev, kot so potresi, začasno izgubijo svojo strižno trdnost. Ta primer je opredeljen kot utekočinjenje zemljine in povzroči znatno škodo na konstrukcijah. Potencial utekočinjanja zemljin je odvisen od številnih parametrov zemljin, pridobljenih s terenskimi in laboratorijskimi preiskavami. V pričujoči študiji so bili razviti novi modeli za oceno potenciala utekočinjenja nekoherentnih zemljin. V mestih Kayseri in Erzincan je bilo zbranih 837 nizov podatkov o zemljinah za izračun potenciala utekočinjenja z nelinearno multiplo regresijo in umetno inteligenco. Modele, ki temeljijo na tehnikah nelinearne multiple regresijske analize (NMRA), umetnih nevronskih mrež (ANN) in sistema prilagodljivega nevro-mehkega sklepanja (ANFIS), smo primerjali z rezultati poenostavljene metode. Za kriterije ocenjevanja uspešnosti modelov so bili izračunani determinacijski koeficienti (R^2) in različne stopnje napak. S predlaganim modelom ANN smo našli kompleksno razmerje med zemljino in vhodnimi parametri ter napovedali potencial utekočinjenja natančneje kot z drugimi metodami. Model ANN ima skupno stopnjo uspešnosti 90 odstotkov in najnižjo srednjo absolutno stopnjo napake 0,024. Z izboljšanjem obstoječih metod so bili uvedeni novi modeli za oceno verjetnosti utekočinjenja zemljin.

1 INTRODUCTION

When natural disasters are evaluated in terms of loss of life and property, earthquakes come first. An earthquake has negative effects on building structures. In particular, when the saturated sandy and silty soils are liquefied, they cannot bear the weight of the structures standing on them, which causes the structures to sink and tilt [1].

The liquefaction of soils can be expressed as the liquid-like behaviors of saturated, cohesionless or low-cohesive soils that lose their shear strength because of the vibrations of cyclic, earthquake shock waves. Liquefaction is the increase of the pressure of the water in the soil void spaces (pores) and the deterioration of the soil's structure under repeated loads from the effect of the earthquake. The increase in pore-water pressure reduces the effective stress in the soil and then leads to a loss of shear strength and the soil starts to act like a liquid.

When earthquakes with a moment magnitude (M_w) greater than 5 are examined, liquefaction occurs most frequently on loose, saturated sandy and silty soils. These deformations caused buildings to collapse or be severely damaged. For example, in the Niigata, Japan earthquake with a moment magnitude of 7.5 occurred in 1964, concrete buildings sank and tilted laterally. In 1995, in the 7.2 magnitude earthquake in Kobe, Japan, bridges, buried pipelines, port facilities, the retaining walls on the coasts were damaged by tilting and the buildings sank due to liquefaction. Similarly, a 7.5 magnitude earthquake occurred in Gölcük-Marmara, Turkey in 1999. Many structures sank into the soil or tilted to one side [1-3]. Assessments of liquefaction risk started with the 1964 Niagata earthquake and became much more important for the 1971 San Fernando, 1985 Mexico City, 1989 Loma Prieta, 1995 Kobe and 1999 Marmara earthquakes.

Laboratory and field tests are used in the liquefaction assessment based on the simplified method. Competent people should prepare the samples representing the field conditions to obtain the correct results from laboratory experiments. It is a challenging and laborious job in practice. Damage often occurs during the collection of the samples from the field and the transportation and preparation for the experiment. Therefore, a determination of the liquefaction potential according to just the results of a laboratory experiment causes errors. Instead, the results of the Standard Penetration Test (SPT) carried out in the field are often used to estimate a soil's liquefaction resistance [4].

Some factors affecting liquefaction include ground water, earthquake moment size, soil type, corrected soil SPT penetration resistance (N_{60}) value, relative density (D_r), depth of the obtaining SPT N_{60} value (d), fines content (FC), which is defined as the ratio of soils passing a No. 200 (75 μm) sieve, average grain diameter (D_{50}), unit weight (UW), groundwater level (GWL), effective stress (σ'_{vo}), total stress (σ_{vo}), peak ground acceleration (a_{max}) and depth of the earthquake from the surface [1, 2, 5-10]. Seed and Idriss [4] carried out the first studies on liquefaction based on SPT data. They suggested that liquefaction could be estimated with graphs and equations, depending on SPT. A liquefaction assessment using SPT data, which was developed by Seed and Idriss, is referred to as the "simplified method"[4]. Versions of this method improved by other authors are used worldwide [11],[12],[13]. The Turkish Building Earthquake Code (TBEC-2018) accepts the "simplified method", which consists of the empirical equations dependent on the SPT published by Seed and Idriss[4] as the standard method for soil-liquefaction analyses.

The cyclic stress approach is used to evaluate the liquefaction potential. In this approach, the Cyclic Stress Ratio (CSR) represents the earthquake load, i.e., the earthquake's soil liquefaction effect or demand. Depending on the results of the SPT test in the field, the resistance of the soil to liquefaction is represented by the Cyclic Resistance Ratio (CRR). The fact that the liquefaction factor of safety, i.e., the CRR over CSR ratio, is less than 1.10 usually means that soil will liquefy according to TBEC-2018 [38]. The cyclic-stress approach in assessing the liquefaction potential expresses both the earthquake effect (CSR) and the soil-liquefaction resistance (CRR) in cyclical stresses. The cycle number for the CSR, which is a function of the duration of the earthquake movement, is proportional to the magnitude of the earthquake. The cyclic liquefaction resistance (CRR) is obtained in the laboratory by cyclic triaxial and simple shear tests or most often in the field by the SPT. The CRR is expressed in terms of the number of cycles required for the occurrence of collapsing in a soil exposed to cyclic shear stress at a certain level. However, the CRR is affected by the stress and unit deformation history, age, and soil texture, which are disturbed during sampling and are very difficult to simulate in the laboratory. Therefore, field tests are preferred for a liquefaction assessment. The SPT is a widely used field test used to calculate the CRR of the soil.

In recent years the number of scientific studies using numerical methods based on statistical and artificial intelligence techniques for estimating the liquefaction of soils has increased [9, 10, 14-23]. Finn, Dowling and Ventura[22] developed methods that estimate the liquefaction potential and lateral expansion displacements. Boulanger and Idriss[13] studied the probability of triggering liquefaction based on SPT. In their study,

they obtained the maximum probability approach and the liquefaction trigger correlation related to SPT. Keramatikerman, Chegenizadeh, and Nikraz [24] conducted a series of repeated triaxial experiments to determine the effect of fly ash on the liquefaction resistance of sands, and they observed that the resistance to liquefaction increases with the increasing ash ratio and time.

Yang et al. developed an SPT-based empirical equation to assess sand liquefaction [19]. Rahman and Siddiqua [21] estimated the liquefaction resistance of soils using the standard penetration test, cone penetration test, and shear wave velocity data for the cities of Dhaka, Chittagong and Sylhet in Bangladesh. The effects of *FC* on the liquefaction of soils were also investigated [25]. Anwar et al. obtained a model to find the *CRR* for MRA-based soil-liquefaction analysis using SPSS and the MATLAB program [26]. Fei-hong[27] investigated the statistical relationship between the liquefaction index and the depth using SPT data to assess soil liquefaction in the port area of Tianjin city, and they showed a significant relationship between the liquefaction index, the depth, and the SPT N-value. Another study claimed that liquefaction is a complex ground-degradation problem involving soil and earthquake parameters, and ground deformations caused by liquefaction should be investigated by nonlinear methods [28]. In a study investigating fuzzy neural network models for the prediction of liquefaction, integrated fuzzy neural network models were developed to evaluate the liquefaction potential [5]. Muduli and Das developed an empirical model using multi-gene genetic programming (MGGP), which is an SPT-based artificial intelligence technique, to determine the *CRR* of the soil [23]. In a study estimating the safety coefficient against liquefaction with artificial neural networks, the liquefaction potential of the soils in the adjacent area of Gümüşler-Denizli province was evaluated, and the safety coefficient against liquefaction was estimated with the help of ANN [29].

In this study, standard penetration tests were carried out in 63 drill holes in Erzincan and 60 drill holes in Kayseri. Seed and Idriss's simplified liquefaction analysis [4] was used to determine the liquefaction safety coefficient and the effects of various soil parameters on calculating the *CRR* were investigated.

A data set containing these parameters and the liquefaction safety coefficients was prepared. First, a quadratic nonlinear multiple regression model (NMRA) that predicts the soil liquefaction and reflects the nonlinear behavior of the soil was developed with this data set. Consequently, models that predict the liquefaction of the soils were created with Artificial Neural Networks (ANN) and then with the Adaptive Neuro-Fuzzy Inference System (ANFIS) using the same training and test data. Randomly selected training and test data were used in the development of the NMRA, ANN, and ANFIS models. To develop the best model, CRR_{act} values obtained as a result of a simplified liquefaction analysis were compared with the predicted CRR_{pred} values using the NMRA, ANN, and ANFIS models. This study examines the estimation methods against the liquefaction hazard that an earthquake can cause. By determining the liquefaction potentials and comparing the estimation methods, the soil improvements and geotechnical designs will be more secure. They can help to prevent the devastating consequences of earthquake-induced liquefaction.

2 METHOD OF MODELLING

The standard penetration test is widely used in the calculation of liquefaction analysis. SPT is a simple and relatively low-cost field test for the evaluation of liquefaction potential due to easy data acquisition, the presence of a database prepared from the data obtained in previous earthquakes, and revealing a good correlation of these data with new earthquakes. Ref [4] proposed equation (1) for the liquefaction analysis.

$$CSR = \left(\frac{\tau_{av}}{\sigma'_{vo}}\right) = 0.65\left(\frac{a_{max}}{g}\right)\left(\frac{\sigma_{vo}}{\sigma'_{vo}}\right)r_d \qquad (1)$$

where

a_{max} = peak horizontal ground acceleration on the soil surface (m/s^2)
g = acceleration due to gravity (m/s^2)
σ_{vo} = total vertical soil stress (kPa)
σ'_{vo} = effective vertical soil stress (kPa)
r_d = stress-reduction coefficient from equations (2) and (3)

The largest (*CSR*) in the formula is the ratio of the mean shear stress ($\tau_{av} = 0.65 * \tau_{max}$) to the effective vertical stress. The effective stress-reduction coefficient r_d is a value that considers the flexibility of the soil column (e.g., $r_d = 1$ corresponds to rigid mass behavior).

$$\begin{aligned} r_d &= 1.0 - 0.00765z \quad z \le 9.15\ m\ ; \\ r_d &= 1.174 - 0.0267z \quad 9.15\ m < z \le 23\ m \end{aligned} \qquad (2)$$

$$\begin{aligned} r_d &= 0.744 - 0.008z \quad 23\ m < z \le 30\ m; \\ r_d &= 0.5 \quad z > 30\ m \end{aligned} \qquad (3)$$

Ref [12] proposed equation (4) for calculating the recurrent resistance ratio ($CRR_{M7.5}$) for clean sands with an *FC* of less than 5 % and earthquakes with a magnitude of $M_w = 7.5$. The next step is to find the dynamic cyclic resistance ratio (*CRR*) for the ground, based on the calculated clean sand equivalent.

$$CRR_{M7.5} = \frac{1}{34-(N_1)_{60cs}} + \frac{(N_1)_{60}}{135} + \frac{50}{(10(N_1)_{60}+45)^2} - \frac{1}{200} \quad (4)$$

$CRR_{M7.5}$ = cyclic resistance ratio to soil liquefaction for M_w = 7.5 earthquake

Considering the effect of the *FC* on the liquefaction resistance, the corrected SPT-N values used in the liquefaction analysis are suggested to be corrected as follows. Ref [12] proposed using the $(N_1)_{60}$ value after converting to the clean-sand equivalent $(N_1)_{60cs}$. They wanted to reduce the impact of *FC* on the soil on the *CRR*. Equation (5) is given as follows.

$$(N_1)_{60cs} = \alpha + \beta(N_1)_{60} \quad (5)$$

Where α and β are the fines-content correction coefficients.

The $(N_1)_{60cs}$ value is calculated using equations (6), (7) and (8) according to the *FC* ratio.

$$for\ FC \leq 5\ \% \quad \alpha = 0 \ \text{and} \ \beta = 1.0 \quad (6)$$

$$for\ 5\,\% < FC < 35\,\% \quad \alpha = \exp\left[1.76 - \left(\frac{190}{FC^2}\right)\right] \quad \text{and} \quad \beta = \left[0.99 + \left(\frac{FC^{1.5}}{1000}\right)\right] \quad (7)$$

$$for\ \text{FC} \geq 35\ \% \quad \alpha = 5 \ \text{and} \ \beta = 1.2 \quad (8)$$

where SPT-N_{field} is the value adjusted to 60 % of the energy ratio and $(N_1)_{60cs}$ is the number of SPT blow-count values with the fines-content correction. Liquefaction occurs when the *CRR*, which shows the soil's resistance to liquefaction, exceeds the liquefaction resistance (*CSR*) caused by earthquakes. If this situation is explained in terms of safety factor, the Safety Coefficient (F_S) is given by equation (9).

$$F_S = \frac{CRR}{CSR} \quad (9)$$

$F_S \leq 1.1$ is considered as there being a liquefaction potential and $F_S \geq 1.1$ is considered as there being no liquefaction potential [38]. The equations and curves given for the calculation of *CRR* are valid for an earthquake with a moment magnitude M = 7.5. The

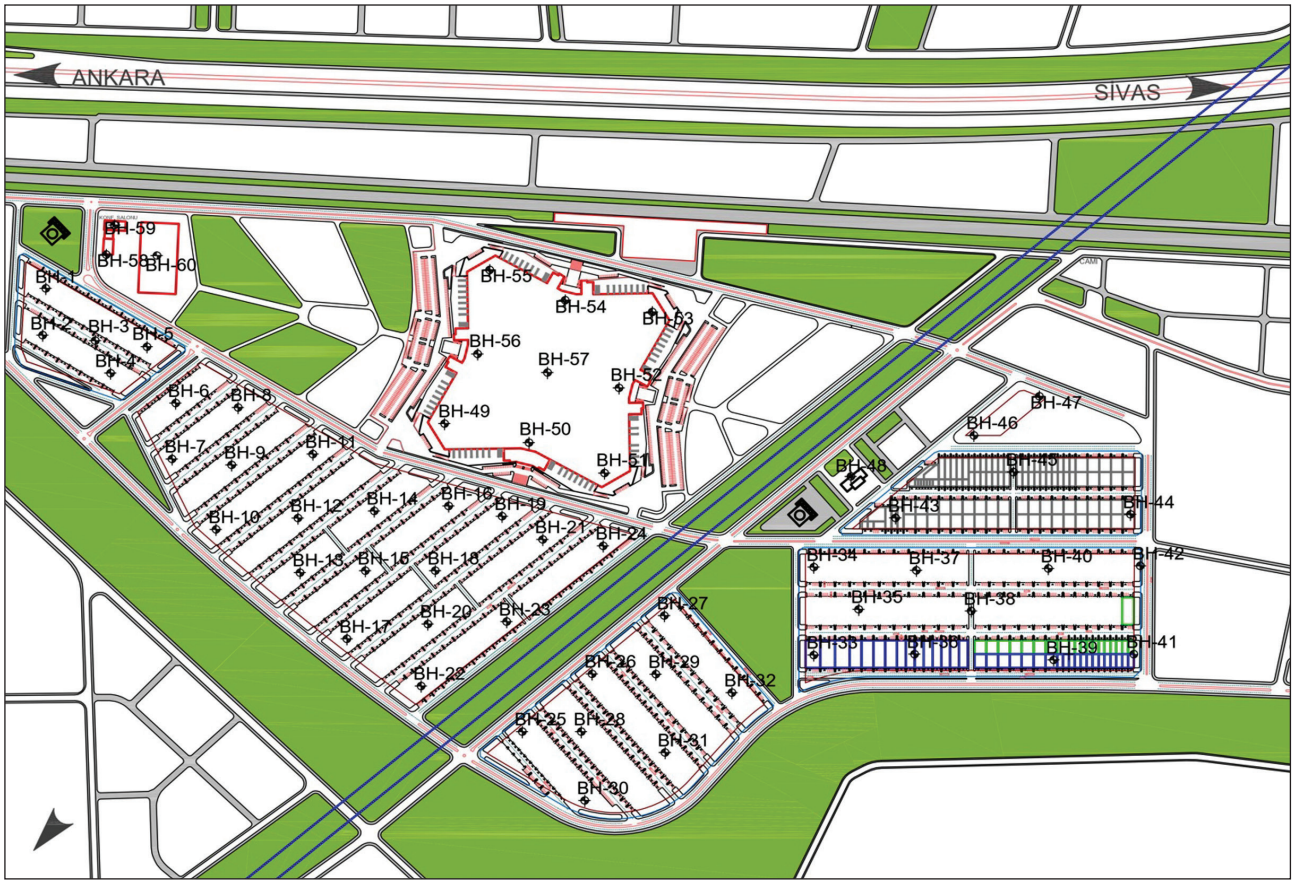

Figure 1. Site layout and borehole location plan in Kayseri province.

earthquake-magnitude correction factor specified in equations (10) and (11) is proposed to use in different earthquake magnitudes.

$$F_s = \frac{CRR}{CSR} * MDF \qquad (10)$$

MDF = earthquakmagnitude correction factor

$$MDF = \frac{10^{2.24}}{M^{2.56}} \qquad (11)$$

M = earthquake moment magnitude

2.1 Study Area Description

The analysis was performed using data from two different cities (Kayseri and Erzincan). In Kayseri, there are generally sand and silt soil layers in the area under investigation (approximately 1.5 million square meters) (Figure 1). However, the soil properties of the region have variable soil conditions, and it is silty sand in some places and silty clay with sand inter bands at some other sites. The 60 borehole drillings, shown in Figure 1, were performed in the study area, and SPT was achieved every 1.5 m in drillings between SK1-SK48 and SK49-SK60. The peak horizontal ground acceleration at the ground surface (a_{max}) value can be expressed in gravitational acceleration (g). In the study area of Kayseri, the a_{max} value changes between 0.190 g and 0.200 g (TBEC-2018). The earthquake that is thought to affect Kayseri Province and its surroundings is the movement that might occur in the Ecemiş fault, which is a strike slip fault. The second study area is in the Erzincan plain in the city center of Erzincan. The soil types in Erzincan are generally non-plastic, silty sand and clayey sand. The peak ground acceleration, a_{max} value changes between 0.600 g and 0.615 g in the

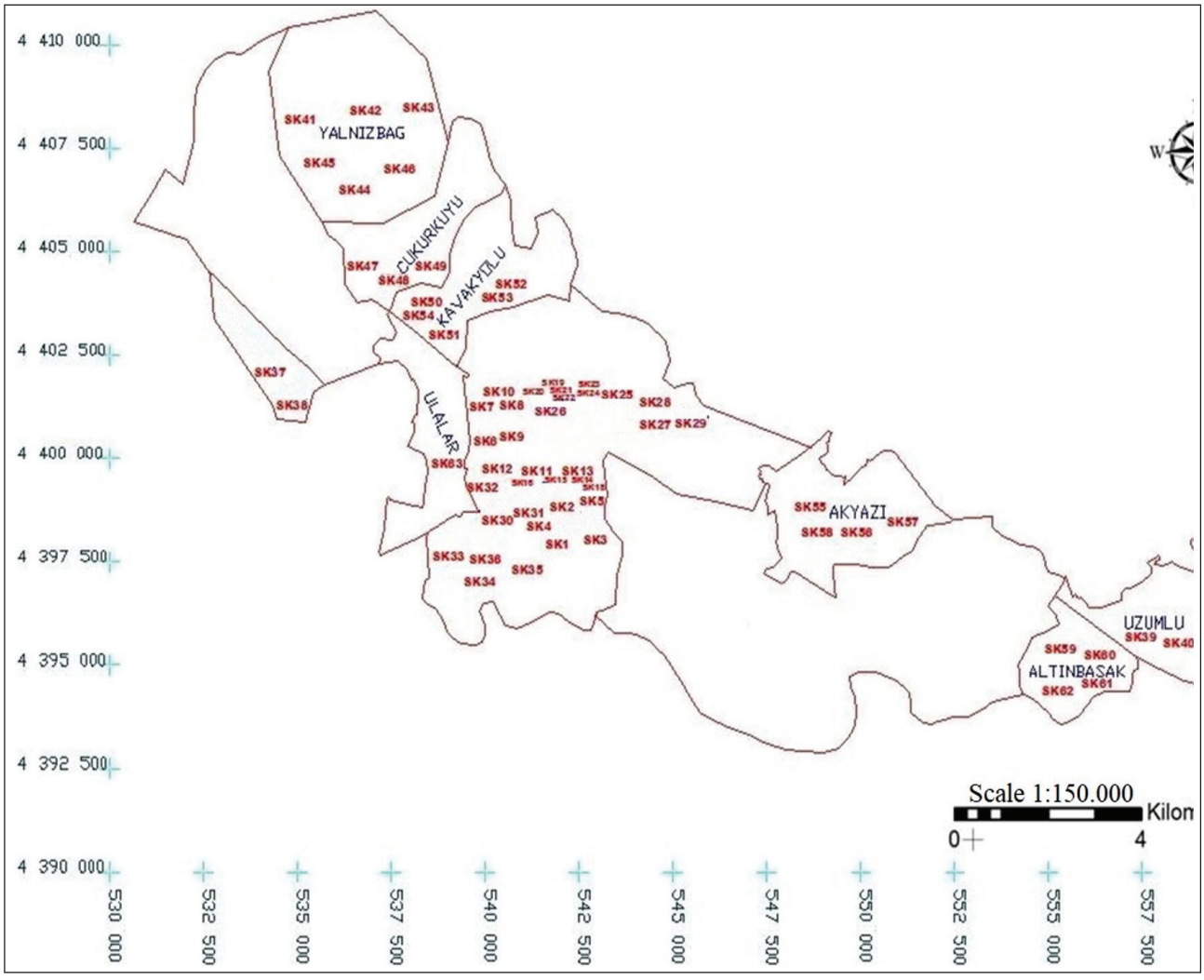

Figure 2. Bore holes in the study area in Erzincan province[33]. *Note: SK = Borehole*

study area in Erzincan. This region is located on the KAFZ (North Anatolian Fault Zone), which is the most effective fault zone of Turkey.

The earthquake parameters were calculated separately using the geographical location data entry for Erzincan and Kayseri with the interactive web application (https://deprem.afad.gov.tr). According to the soil class in the study areas, the earthquake ground-motion level with a probability of exceeding 10 % in 50 years (recurrence period of 475 years) was taken into account (this is called DD-2, TBEC-2018). According to the earthquake ground-motion level, the peak acceleration values varied between 0.600 *g* and 0.615 *g* and 0.190 *g* and 0.200 *g* in the study area of Erzincan and Kayseri, respectively (TBEC-2018). According to these acceleration values, it was estimated as M_w 7.5 for Erzincan and M_w 6 for Kayseri. In the estimation of these moment-magnitude values, the approaches given in the literature were used [39].

In the city of Erzincan, 63 borehole drillings with depths of 1.5 m to 20 m, shown in Figure 2, were made. The soil parameters and SPT data were collected, and soil profiles were created for these drillings.

3 PROCESSING AND ANALYSIS OF DATA

This study explains the *CRR* values obtained from Simplified Liquefaction Analysis [4], NMRA, ANFIS, and ANN methods for Kayseri and Erzincan. Groundwater levels varied between 1.7 and 2.8 m in the study area in Kayseri. The Unit Weight Test to determine the mass properties of the soil, the Water Content Test to determine the amount of water in the unit volume of the ground, the Liquid Limit and Plastic Limit Tests to determine the consistency characteristics, the Sieve Analysis Test to determine the grain diameter and distribution of the soil and Hydrometer test to determine the *FC* were carried out in the samples taken from the investigated area. As the result of the tests, it was observed that the unit-weight values varied between 15 and 20 kN/m^3, the *FC* ranged from 12 to 45 %, and the water contents ranged between 13 % and 47 %. The soils' liquid limit (LL) values varied between 30 % and 47 %, and the plastic limit (PL) values ranged from 20 % to 27 %.

In Erzincan, 63 boreholes were drilled in 16 different locations with depths ranging from 1.5 to 20 m. The soil types are SM (silty sands), SC (clayey sands), and CL (inorganic clay of low plasticity), which are characterized as primarily coarse-grained soils. Therefore, tests were made on the samples and the following results were obtained; the unit-weight values varied between 16.93 and 19.96 kN/m^3, *FC* ranged from 12 to 77 %, and the water-content (w) value varied between 10 and 30 %. The soil types in the study area were generally non-plastic (NP) according to the PL and LL test results. According to the results of the SPT obtained from both regions and the laboratory experiments, 837 data sets were obtained. The reason for choosing these data is that the groundwater table is close to the ground surface and the soil properties also show the liquefaction potential in both regions. These data were randomly divided into two groups: a training group composed of 70 % (586) and a test group consisting of 30 % (251) of the data. Data of the parameters used in *CRR* calculation were the earthquake magnitude (M_w), depth (d), corrected soil SPT penetration resistance (N_{60}), saturated Unit Weight (UW), peak ground acceleration (a_{max}), fines content (FC), cyclic stress ratio (CSR), groundwater level (GWL), total stress (σ_{vo}) and average grain diameter (D_{50}). There are the effects of 9 different independent variables on the calculation of *CRR*.

In liquefaction, the groundwater level is generally crucial up to the first 3 m from the surface. Although liquefaction rarely occurs in environments where the groundwater level is deeper than 10 m from the surface, liquefaction is not expected in environments where the groundwater level is deeper than 20 m, in general [36]. In a complex hydrogeological environment, the groundwater level is variable due to hydraulic transitions that affect the hydraulic properties of the soil. Moss et al. (2017) investigated the effect of the groundwater level on the liquefaction potential and the effect of changing the depth of the water table on the liquefaction according to seasonal variations. The water level rises to its maximum during the rainy season due to rain. The study highlights the need for seasonal liquefaction-sensitivity studies [37].

The groundwater levels varied between 1 and 2.7 m and 1.7 and 2.8 m in the study area of Erzincan and Kayseri, respectively. The groundwater table is assumed to be at the surface for the worst-possible scenario for both regions in determining the liquefaction hazard by considering seasonal and global climate change. Besides, since the groundwater-depth values measured in the field show minor local variations according to the regions, these parameters did not make a significant difference in the training of the prediction models compared to the effect of other parameters. For this reason, it was not preferable to use the groundwater-level depth parameters as variables in the models.

In estimating the CRR, N_{60}, d, FC, D_{50}, UW, GWL, effective stress (σ'_{vo}), and total stress (σ_{vo}) parameters were considered. The soils are completely saturated below the water table since the groundwater layer is assumed to be at the surface. Saturated unit weights were used

for a point below the groundwater table to calculate the effective unit weight. However, since the soil depth and *UW* are used to calculate the total stress values, only the depth and saturated unit weight represent the total stress to avoid over-learning in the estimation methods.

In addition, a Variance Influence Factor (VIF) analysis was performed to see the effect of independent variables on *CRR*. All the estimation models used in this study are based on the five most influential variables for the saturated condition of the soils as a result of a VIF analysis. These are, namely, the SPT value (N_{60}), fines content (*FC*), saturated unit weight (*UW*), depth from which SPT is obtained (*d*), and average grain diameter (D_{50}) parameters.

Regression analysis coefficients, T-test values, and VIF values, which were obtained with the analysis performed to determine the relationship between the dependent and independent variables, are given in Table 1. In Table 1, the N_{60}, *UW*, D_{50}, *d*, and *FC* parameters are preferred as independent variables, since the VIF values were less than 5. It is clear that the *t* values obtained were not within the range -1.645 < *t* < 1.645, which were determined for $t_{critical}$ = 1.645. Thus, the N_{60} , *UW*, D_{50} , *d*, and *FC* parameters were significant in the *CRR* estimation and were used in the analysis.

Table 1. Results of regression analysis according to *t* test and VIF analysis.

Independent variables	Regression coefficients	*t*	VIF
Constant	2.7806	6.88	
N_{60}	0.0212426	26.39	1.581
FC	0.17645	2.03	1.268
UW	−0.16398	−7.08	1.795
d	0.0018631	2.24	1.112
D_{50}	0.6052	2.78	1.257
$t_{critical}$ = 1.645			

Table 2. Data statistics.

Parameters	*n*	Min	Max	Mean	Standard deviation
N_{60}	837	2	53	16.86	9.873
FC	837	0.12	0.77	0.36	0.076015
UW	837	16.93	19.96	18.16	0.36116
d	837	1.50	40.50	11.23	7.86592
D_{50}	837	0.01	0.334	0.042	0.023233
CRR	837	0.084	1.24	0.30	0.1426768

It was reported that the test performance of fuzzy logic-based models such as ANFIS decreases with an increase in the number of independent variables [5]. Therefore, the number of independent variables was limited to 5 in this study. Table 2 shows the statistical data of this dependent (*CRR*) and the independent variables.

3.1 Performance Criteria

In estimating the *CRR* value, MAE, MSE, RMSE, MARE, and R^2 are taken into account to compare the performance of the models. The model error rate occurs because it does not fully represent a proper relationship between the predicted and the actual parameters. As a result of this incomplete relationship, different error-rate indices can be expressed. The mean absolute error (MAE) is the measured difference between two variables. The MAE is also the average horizontal distance between each data point and the best-fit line. Since the MAE value is easily interpretable, it is frequently used in regression and artificial intelligence techniques. The MAE value can vary from zero to infinity. The mean square error (MSE) measures the performance of the model, the estimator, telling how close the prediction curve is to a set of points. When the MSE value is zero, the model has the best-possible performance. The RMSE (root-mean-square error) is the standard deviation of the estimation errors. The RMSE is a measure of the distribution of residues. The RMSE value can range from zero to infinity. A zero RMSE value means that the model made no errors. The MARE expresses the difference between the estimated value and the observed value. The MARE is a non-negative error rate that can take a value from zero to infinity. When the MARE value is zero, the considered model has the best-possible performance. The performance criteria used for model evaluation in this study are given in Table 3. These are (R^2), MAE, MARE, MSE and RMSE. Here, the value of R^2 indicates the closeness of our model (as a percentage) to the real values. In Table 3, CRR_{act} , $\overline{CRR}_{act}$, CRR_{pred} , $\overline{CRR}_{pred}$ are the real values of CRR_{act} , calculated by simplified liquefaction analysis, the calculated real mean $\overline{CRR}_{act}$, the predicted CRR_{pred} , and the predicted mean $\overline{CRR}_{pred}$, respectively.

Equation (12) is used to normalize the data to transfer to the MATLAB program.

$$X_n = \frac{(X_0 - X_{min})}{(X_{max} - X_{min})} \qquad (12)$$

Here, X_n is normalized data, X_0 is original data, X_{min} is minimum data and X_{max} is maximum data. All data were scaled between 0 and 1.

Table 3. Performance-evaluation criteria.

Evaluation criteria	Definition
Coefficient of determination (R^2)	$R^2=\left[\frac{\sum_{i=1}^{n}(CRR_{act}-\overline{CRR}_{act})(CRR_{pred}-\overline{CRR}_{pred})}{\sqrt{\sum_{i=1}^{n}(CRR_{act}-\overline{CRR}_{act})^2}\sqrt{\sum_{i=1}^{n}(CRR_{pred}-\overline{CRR}_{pred})^2}}\right]^2$
Mean absolute error (MAE)	$MAE=\frac{1}{n}\lvert CRR_{Pred}-CRR_{act}\rvert$
Mean absolute relative error (MARE)	$MARE=\frac{1}{n}\lvert 100\text{x}\frac{CRR_{act}-CRR_{pred}}{CRR_{pred}}\rvert$
Mean square error (MSE)	$MSE=\frac{1}{n}\sum_{i=1}^{n}(CRR_{act}-CRR_{pred})^2$
Root mean square error (RMSE)	$\text{RMSE}=\sqrt{\frac{\sum_{i=1}^{n}\left(CRR_{act}-CRR_{pred}\right)^2}{n}}$

3.2 Nonlinear Multiple Regression Analysis (NMRA) model

Nonlinear multiple regression analysis (NMRA) is used to detect two or more correlations. NMRA is a statistical method that can reveal the relationship between more than one independent variable and a single dependent variable, make predictions, and create a mathematical model. In this study, quadratic regression was used to estimate the *CRR* value. The basic equation of the regression model is relatively simple, as given by Equation 13. *CRR* represents the dependent variable; N_{60}, *FC*, *UW*, *d*, and D_{50} are the independent variables. The ability of the estimations to give reliable results depends on the coefficient of determination (R^2) being the largest value and the error rates being the smallest value.

Then, with the help of the SPSS program, various functions were tested with these independent variables, and the best fit for the distribution is the nonlinear quadratic equation. The quadratic NMRA equation was chosen, which gave the highest R^2 and the lowest RMSE values. The nonlinear regression equation obtained from the NMRA analysis is shown in Equation 13.

Thus, the R^2 was 0.718 for the training data and 0.681 for the test data. Other error statistics of the NMRA model are shown in Table 6. As shown in Figure 3, the CRR_{act} and CRR_{pred} values were close to each other. However, as shown in Figure 3, the *CRR* values diverged from the calculated *CRR* values after 0.80, which results in a reduction of the determination coefficient.

$$CRR_{pred}=0.233+0.022\cdot N_{60}-0.322\cdot FC+0.622\cdot FC^2+0.267\cdot UW-0.016\cdot UW^2-0.035\cdot d+0.001\cdot d^2+1.633\cdot D_{50}-0.614\cdot D_{50}^2 \quad (13)$$

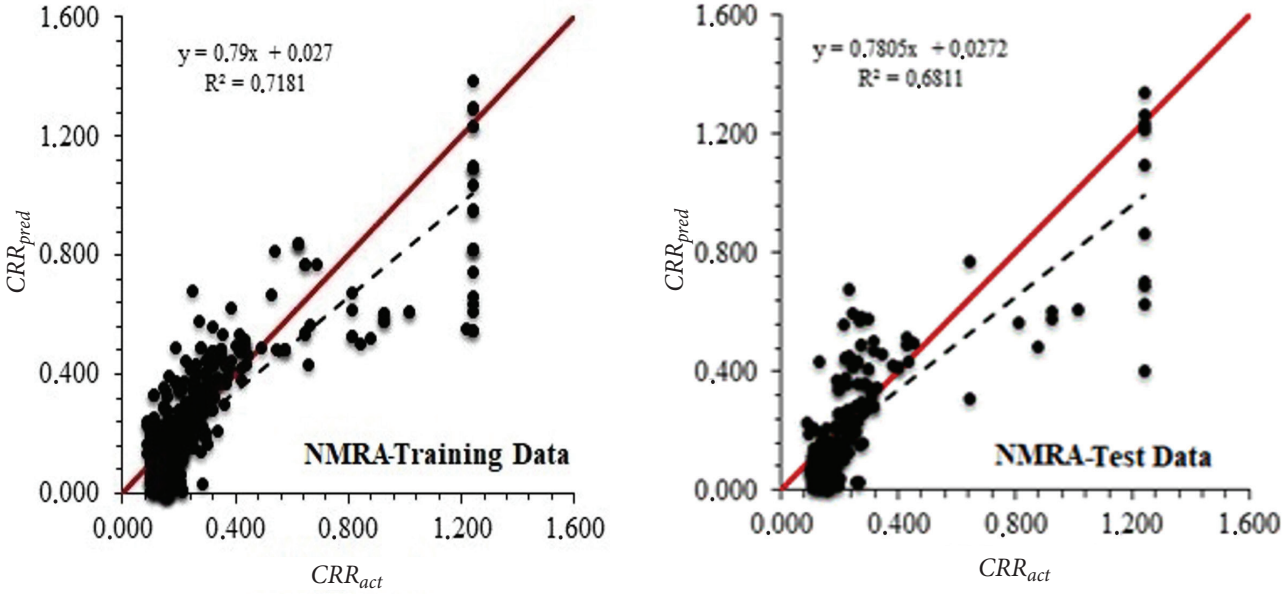

Figure 3. Comparison of predicted and actual CRR values for NMRA model in training and test data (Nonlinear Multiple Regression Analysis).

3.3 Artificial Neural Network (ANN) model

The estimation method, called ANN (Artificial Neural Networks), is the most well-known and widely used method among the artificial intelligence techniques. This method estimates the dependent variable by finding linear or nonlinear relationships between the parameters that represent many independent variables. In this technique, the working system imitates the human brain. ANN makes routing with multi-layer sensor networks and can learn and generalize between the input and output layers. The ANN structure consists of an input layer, an output layer, many hidden layers, and a large number of neurons corresponding to each independent variable. An ANN is very successful in finding nonlinear relationships between independent variables about the dependent variable. The output layer also corresponds to the predicted dependent variable. The system updates the weight values, moving from the output layer to the input layer, and the error value is minimized [5,10]. The *CRR* was estimated with the ANN model. In the prediction model, the input parameters are N_{60}, *FC*, *UW*, *d*, D_{50}, and the output parameter is the *CRR* (Figure 4).

The feed-forward ANN model inputs with five variables consisting of N_{60}, *FC*, *UW*, *d* and D_{50}, and a single output system *CRR* was obtained, as shown in the diagram in Figure 4. In the training of the models, a random selection of 586 parameters was used, and 251 parameters were used to test the prediction model's performance. First, the most appropriate and widely used tansig, logsig and purelin functions from 11 member functions were used in the multi-layered ANN method in MATLAB [34]. The input data were trained with the Levenberg-Marquardt algorithm, due to its ease of use, convergence rate and predictive success in linear and nonlinear models. The numbers of neurons in a hidden layer ranged from 2 to 10, and the numbers of iterations ranged from 1 to 100. Using a trial-and-error method, a model was determined from the network structures obtained. In this model, the input membership function was logsig, which gives the lowest all error statistic values and the highest Determination Coefficient (R^2) value, and its output membership function was purelin.

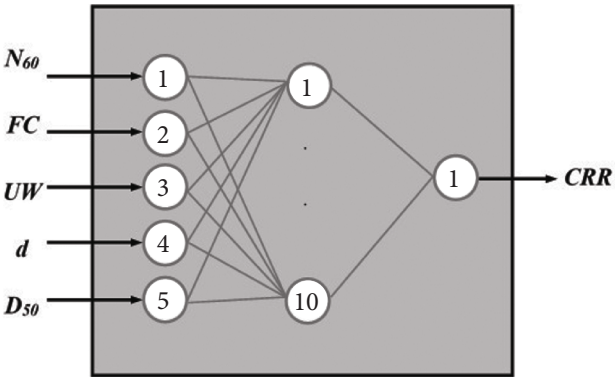

Figure 4. ANN block diagram for the designed system.

Table 4. Best ANN model for predicting CRR.

Membership function		Membership function number	Iteration number
Input	Output		
Logsig	Purelin	10	95

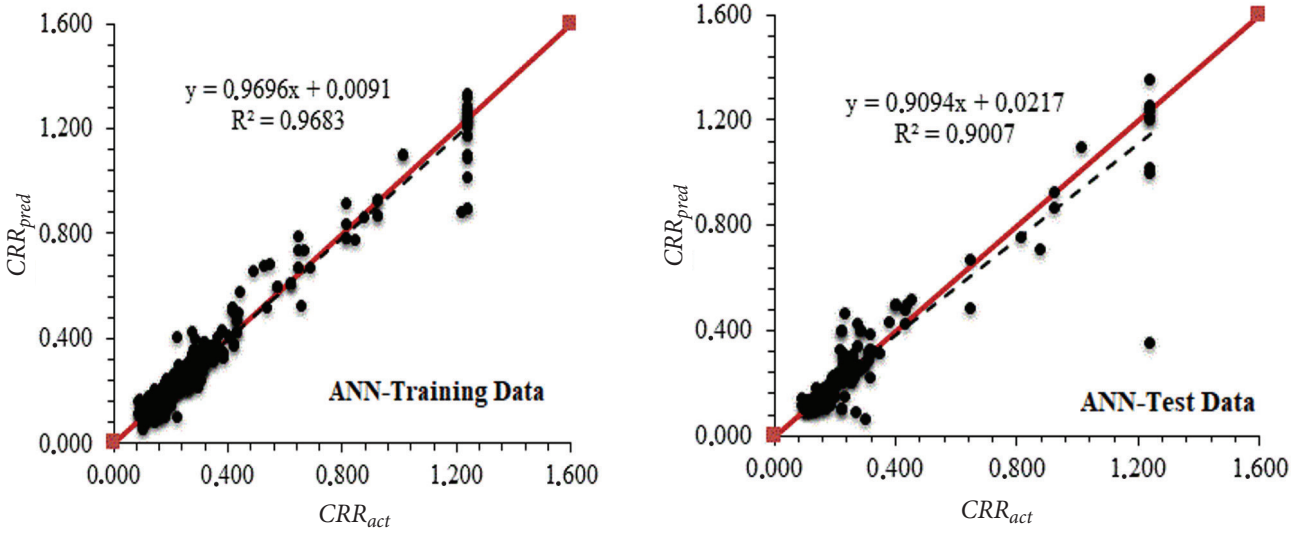

Figure 5. Comparison of predicted CRR and actual CRR values for ANN model.

ANN characteristics of the best ANN model obtained after various trial-and-error attempts are given in Table 4. All the statistical error values of the best ANN model are given in Table 6. According to Table 6, the error statistic values were within acceptable values. As can be seen in Table 6, and Figure 5, the ANN model revealed successful results. However, it can be seen in Figure 5 that the predictive *CRR* values diverged between 1.0 and 1.5 from the calculated *CRR* values.

3.4 Adaptive neuro-fuzzy inference system (ANFIS) model

The adaptive ANN-based fuzzy inference system (ANFIS) is one of the essential artificial intelligence techniques that can optimize parameters with an inference system. ANFIS provides for the optimization of rule-base and membership function values to model systems with known input and output values with fuzzy logic. The optimization process involves th learning methods of ANN. In this way, fuzzy systems, which normally cannot learn, gain a learning ability for the data sets to be modeled. ANFIS uses the backpropagation method, as a learning method, or a combination of the backpropagation method and the least-squares estimation method. The ANFIS architecture consists of six layers. The first layer (input layer) transmits the incoming input signals to the other layers. The second layer is called the fuzzy layer, the third layer is the rule layer, and the fourth layer is the normalization layer. The fifth layer is the annotation layer, and finally, in the sixth layer, the values from the annotation layer are aggregated within this layer to obtain the actual output value of the ANFIS system [5]

In the ANFIS model, since the increase in the membership function numbers, as mentioned above, decreased the performance of the ANFIS model, the input membership function numbers 2 and 3 were taken. After determining the Gaussian membership function (gaussmf) and the triangular (trimf) membership function as the input membership functions and the constant and linear functions as the output functions, the best ANFIS model was determined by trial and error in iteration numbers ranging from 1 to 5. Depending on the type of input and output functions, Gaussmf-constant, Gaussmf-linear, trimf-constant and trimf-linear combinations were determined. The lowest errors criteria and the highest R^2 are used to select the best model among the four different ANFIS combinations. The ANFIS features of the best ANFIS model obtained from various trial-and-error attempts are given in Table 5. The input membership function is Trimf, and the output membership function is linear. Allthe error-statistics values of the best ANN model and successful results of the ANFIS model are given in Table 5, Figure 6 and Table 6.

Table 5. Best ANFIS model for CRR prediction.

Membership function		Membership function number	Iteration number
Input	Output		
Trimf	Linear	2	4

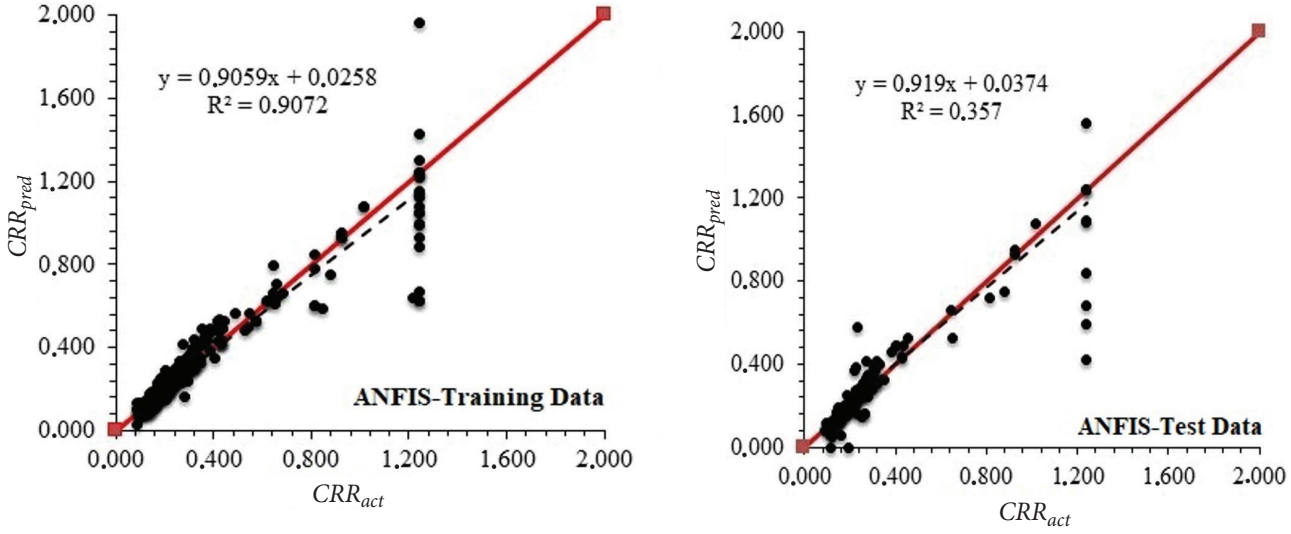

Figure 6. Comparison of predicted CRR and actual CRR values for the ANFIS model.

However, as can be seen in the scatter chart (Figure 6), the CRR_{pred} values between 1.0 and 1.5 diverged from the CRR_{act} values. This situation, which decreases the R^2 coefficient value, indicated that the model failed between 1.0 and 1.5 values.

4 RESULTS AND DISCUSSION

As shown in Table 6, the accuracy of the model's results was accepted as satisfactory, with a determination coefficient greater than 70 % obtained for all the methods. However, considering all the error statistics, the ANN model seems the best among these three methods in terms of R^2 value and the lowest MAE, highest MARE, and MSE ratios. (Table 6).

The results of the models based on the Nonlinear Multiple Regression Analysis (NMRA), Artificial Neural Networks (ANN), and Adaptive Neural Fuzzy Inference System (ANFIS) were compared with simplified analysis results in order to develop the best model for estimating the liquefaction of soils. In all models developed using statistical and artificial intelligence techniques, N_{60} , *FC*, *UW*, *d*, D_{50} were used as the input parameters, and the *CRR* value was estimated as the output parameter.

Table 6. Performance statistics of all models.

		ANN	ANFIS	NMRA
Training	MAE	0.024	0.0350	0.095
	MARE	95.214	11.0330	41.900
	MSE	0.0019	0.0068	0.018
	RMSE	0.0432	0.0822	0.134
	R^2	0.968	0.885	0.718
Test	MAE	0.034	0.036	0.098
	MARE	12.675	10.798	0.441
	MSE	0.006	0.010	0.021
	RMSE	0.0777	0.1002	0.145
	R^2	0.901	0.838	0.681

The R^2 and various error ratios were calculated by comparing the performance of the models created with the training data. It was concluded that the most suitable model is the ANN model based on the success rates and consistency of the models. Liquefaction analysis was carried out for different drilling depths and soil properties. The liquefaction status was calculated by loading data into the Excel spreadsheet. The VIF analysis was performed to see which parameters influence the *CRR* calculation. The numerical studies in the *CRR* calculation showed that the N_{60}, *FC*, *UW*, *d*, and D_{50} parameters affect the *CRR* estimation after the VIF analysis. Thus, new models with higher accuracy were produced using these most influential parameters. The determination coefficients were 0.681, 0.838, and 0.901 in the NMRA, the ANFIS, and the ANN methods, respectively. These values show that a desired level of estimation is achieved. When the error criteria in the NMRA, ANN, and ANFIS methods were evaluated, it was observed that the ANN method was superior to the other methods, with an overall success rate of 90 % and the lowest mean absolute error rate of 0.024.

5 CONCLUSIONS

In this study, a nonlinear multiple regression analysis was performed between the *CRR* value and the other soil parameters. Then, the *CRR* estimation was made with fuzzy logic and artificial neural networks for the N_{60} , *FC*, *UW*, *d* and D_{50} variables, which gave high correlation coefficients. The most sensitive parameters to the soil's liquefaction are d, *FC*, and N_{60}, while the least sensitive ones are the *UW* and D_{50} soil parameters. The ANN model used in the *CRR* estimation has more successful performance criteria when the comparing R^2 and errors. The ANN model has lower error values and a higher correlation than the NMRA and ANFIS models. Improving the existing methods for predicting the liquefaction of soils and estimating the probability of liquefaction with the new models to be developed will enable civil engineers to take precautions against liquefaction. In addition, this study has demonstrated the successful and rapid use of artificial intelligence techniques in solving geotechnical problems, especially in modeling nonlinear complex soil behavior, such as liquefaction. This study is the basis for further studies on liquefaction. In future studies that can be performed to obtain better results in the calculation of the *CRR*, the number of model data can be increased by adding new data. Also, new models and higher accuracy results can be obtained using different artificial intelligence techniques.

REFERENCES

[1] Yoshida, N., Tokimatsu, K., Yasuda, S., Kokusho, T., and Okimura, T. (2001). Geotechnical aspects of damage in Adapazari city during 1999 Kocaeli, Turkey earthquake. Soils and foundations, 41(4), 25-45.

[2] Youd, T. L. (1993). Liquefaction-induced damage to bridges. Transportation Research Record, 1411, 35-41.
[3] Finn, W. L., Byrne, P. M., Evans, S., and Law, T. (1996). Some geotechnical aspects of the Hyogo-ken Nanbu (Kobe) earthquake of January 17, 1995. Canadian Journal of Civil Engineering, 23(3), 778-796.
[4] Seed, H. B. and Idriss, I. M. Simplified procedure for evaluating soil liquefaction potential. Journal of Soil Mechanics Foundations Div, 97, No SM9, PROC PAPER 8371, ,(1971) 1249-1273.
[5] Rahman, M. S. and Wang, J. Fuzzy neural network models for liquefaction prediction. Soil Dynamics and Earthquake Engineering, 22, 8,(2002) 685-694.
[6] Boulanger, R. W. and Idriss, I. M. Probabilistic Standard Penetration Test-Based Liquefaction-Triggering Procedure. Journal of Geotechnical and Geoenvironmental Engineering, 138, 10,(2012) 1185-1195.
[7] Anwar, A., Jamal, Y., Ahmad, S. and Z Khan, M. Assessment of liquefaction potential of soil using multi-linear regression modeling. International Journal of Civil Engineering & Technology (IJCIET), 7,(2016) 373-415.
[8] Muduli, P. and Das, S. "Evaluation of liquefaction potential of soil based on standard penetration test using multi-gene genetic programming model". In Proceedings of the Indian Geotechnical Conference IGC2016 (Chennai, India,(2016).
[9] Kurnaz, T. F. and Kaya, Y. SPT-based liquefaction assessment with a novel ensemble model based on GMDH-type neural network. Arabian Journal of Geosciences, 12, 15,(2019) 456.
[10] Sabbar, A. S., Chegenizadeh, A. and Nikraz, H. Prediction of Liquefaction Susceptibility of Clean Sandy Soils Using Artificial Intelligence Techniques. Indian Geotechnical Journal, 49, 1,(2019) 58-69.
[11] Seed, H. and Idriss, I. Ground motions and soil liquefaction during earthquakes: engineering monographs on earthquake criteria, structural design, and strong motion records. MNO-5. Earthquake Engineering Research Institute, Oakland, CA, (1982).
[12] Youd, T. L. and Idriss, I. M. Liquefaction resistance of soils: summary report from the 1996 NCEER and 1998 NCEER/NSF workshops on evaluation of liquefaction resistance of soils. Journal of geotechnical geoenvironmental engineering, 127, 4,(2001) 297-313.
[13] Boulanger, R. W. and Idriss, I. Probabilistic standard penetration test–based liquefaction–triggering procedure. Journal of Geotechnical Geoenvironmental Engineering, 138, 10, (2012) 1185-1195.
[14] Geyin, M., Baird, A. J. and Maurer, B. W. Field assessment of liquefaction prediction models based on geotechnical versus geospatial data, with lessons for each. Earthquake Spectra, (2020) 1-26.
[15] Farrokhi, F., Firoozfar, A. and Maghsoudi, M. S. Evaluation of liquefaction-induced lateral displacement using a GMDH-type neural network optimized by genetic algorithm. Arabian Journal of Geosciences, 13, 1,(2020) 4.
[16] Moayedi, H., Mosallanezhad, M., Rashid, A. S. A., Jusoh, W. A. W. and Muazu, M. A. A systematic review and meta-analysis of artificial neural network application in geotechnical engineering: theory and applications. Neural Computing Applications, (2018) 1-24.
[17] Kırbaş, İ. Short-term multi-step wind speed prediction using statistical methods and artificial neural networks. Sakarya University Journal of Science, 22, (2018) 24-38.
[18] Asvar, F., Shirmohammadi Faradonbeh, A. and Barkhordari, K. Predicting potential of controlled blasting-induced liquefaction using neural networks and neuro-fuzzy system. Scientia Iranica 25, 2,(2018) 617-631.
[19] Yang, Y., Chen, L., Sun, R., Chen, Y. and Wang, W. A depth-consistent SPT-based empirical equation for evaluating sand liquefaction. Engineering Geology, 221, (2017) 41-49.
[20] Şenol, Ü. and Musayev, Z. Estimating wind energy potential by artificial neural networks method. Bilge International Journal of Science and Technology Research, 1, 1, (2017) 23-31.
[21] Rahman, M. Z. and Siddiqua, S. Evaluation of liquefaction-resistance of soils using standard penetration test, cone penetration test, and shear-wave velocity data for Dhaka, Chittagong, and Sylhet cities in Bangladesh. Environmental Earth Sciences, 76, 5, (2017) 207.
[22] Finn, W. D. L., Dowling, J. and Ventura, C. E. Evaluating liquefaction potential and lateral spreading in a probabilistic ground motion environment. Soil Dynamics and Earthquake Engineering, 91, (2016) 202-208.
[23] Muduli, P. K. and Das, S. K. Model uncertainty of SPT-based method for evaluation of seismic soil liquefaction potential using multi-gene genetic programming. Soils and Foundations, 55, 2, (2015) 258-275.
[24] Keramatikerman, M., Chegenizadeh, A. and Nikraz, H. Experimental study on effect of fly ash on liquefaction resistance of sand. Soil Dynamics and Earthquake Engineering, 93,(2017) 1-6.
[25] Maurer, B. W., Green, R. A., Cubrinovski, M. and Bradley, B. A. Fines-content effects on liquefaction

hazard evaluation for infrastructure in Christchurch, New Zealand. Soil Dynamics and Earthquake Engineering, 76, (2015) 58-68.
[26] Anwar, A., Jamal, Y., Ahmad, S., Z Khan, M. and Publication, I. Assessment of liquefaction potential of soil using multi-linear regression modeling. International Journal of Civil Engineering & Technology (IJCIET), 7,(2016) 373-415.
[27] Fei-hong, G. Evaluation of Soil Liquefaction in Harbor District in Tianjin City. The Open Civil Engineering Journal, 10, (2016) 293-300.
[28] Siyahi, B., Akbas, B. and Dogan Onder, N. "Evaluation Of Liquefaction-Induced Lateral Spreading By A Neural Network (Nn) Model". In Proceedings of the 14th World Conference on Earthquake Engineering (Beijing, China, 2008).
[29] Şen, G., Akyol, E. and Fırat, M. Sıvılaşmaya Karşı Güvenlik Katsayısının Yapay Sinir Ağları İle Tahmin Edilmesi: Denizli-Gümüşler Örneği. Selçuk Üniversitesi Mühendislik, Bilim Ve Teknoloji Dergisi, 22, (2007) 177-184.
[30] Kramer, S. L. Geotechnical Earthquake Engineering. (United Stated of America. Prentice Hall) 1996.
[31] Sönmezer, Y. B., Çeliker, M. and Kılınç, M. Y. Kırıkkale İli Bahçelievler ve Fabrikalar Mahallelerinin Sıvılaşma Potansiyelinin Coğrafi Bilgi Sistemlerinde Analizi. International Journal of Engineering Research and Development, 4, 1, (2012) 33-40.
[32] Seed, H. B. and Idriss, I. M. Ground motions and soil liquefaction during earthquakes. (Berkeley. Earthquake Engineering Research Institute) 1982.
[33] Duman, E. S. Erzincan il merkezi ve çevresindeki zeminlerin standart penetrasyon deneyi verileri kullanılarak sıvılaşma potansiyelinin belirlenmesi. (Determination of the liquefaction potential of soils in and around the centre of the city of Erzincan using standard penetration-test data) Karadeniz Teknik Üniversitesi/Fen Bilimleri Enstitüsü, 2013.
[34] Beale, M. H., Hagan, M. T. and Demuth, H. B. J. T. M. Neural network toolbox™ user's guide,(2010).
[35] Subaşı Duman, E. Erzincan il merkezi ve çevresindeki zeminlerin standart penetrasyon deneyi verileri kullanilarak sivilaşma potansiyelinin belirlenmesi (Determination of the liquefaction potential of soils in and around the centre of the city of Erzincan using standard penetration-test data). MSc. Thesis, Karadeniz Technical University Science Institute, Trabzon, Turkey, 2013.
[36] Youd, T. L. 1984. Geological effects-liquefaction and associated ground failure. Geological and Hydrogeological Hazards Training Program, United States Geological Survey Open-File Report 87-76, 210-232.
[37] Moss RES, Baise LG, Zhu J, Kadkha D (2017) Examining the discrepancy between forecast and observed liquefaction from the 2015 Gorkha, Nepal, earthquakes. Earthq Spectra 33(S1): S73–S83. https://doi.org/10. 1193/120316EQS220M
[38] General Directorate for Foundations, Turkish Building Earthquake Code, TBEC-2018, Ankara, Turkey, (2018).
[39] Wang, B. (2020). Geotechnical investigations of an earthquake that triggered disastrous landslides in eastern Canada about 1020 Cal BP. Geoenvironmental Disasters, 7(1), 1-13.