# Application of Soft Computing Techniques and Multiple Regression Models for CBR prediction of Soils

**Fatimah Khaleel Ibrahim**
*College of Engineering, University of Basrah*
Fatimahkhaleel6@gmail.com

## Abstract

The techniques of soft computing technique such as Artificial Neutral Network (ANN) have improved the predicting capability and have actually discovered application in Geotechnical engineering. The aim of this research is to utilize the soft computing technique and Multiple Regression Models (MLR) for forecasting the California bearing ratio (CBR) of soil from its index properties. The indicator of CBR for soil could be predicted from various soils characterizing parameters with the assist of MLR and ANN methods. The data base that collected from the laboratory by conducting tests on 86 soil samples that gathered from different projects in Basrah districts. Data gained from the experimental result were used in the regression models and soft computing techniques by using artificial neural network. The liquid limit, plastic index , modified compaction test and the CBR test have been determined. In this work, different ANN and MLR models were formulated with the different collection of inputs to be able to recognize their significance in the prediction of CBR. The strengths of the models that were developed been examined in terms of regression coefficient ($R^2$), relative error (RE%) and mean square error (MSE) values. From the results of this paper, it absolutely was noticed that all the proposed ANN models perform better than that of MLR model. In a specific ANN model with all input parameters reveals better outcomes than other ANN models.

**Keywords:** MLR, OMC, LL,PI,CBR, ANN, Prediction, Soils.

## الخلاصة

لقد حسنت تقنيات الحوسبة المرنة مثل تقنية الشبكات العصبية الصناعية (ANN) من القدرة على التنبؤ وقد تم تطبيقها فعليا في مجال الهندسة الجيوتقنية. الهدف من هذا البحث هو الاستفادة من تقنية الحوسبة المرنة ونماذج الانحدار المتعدد (MLR) للتنبؤ بقيمة نسبة التحمل الكاليفورني (CBR) للتربة من خلال مؤشرات خصائص التربة الاخرى. يمكن التنبؤ بقيمة نسبة تحمل كاليفورنيا (CBR) من معاملات خصائص التربة بمساعدة طرق الانحدار المتعدد MLR والشبكات العصبية الصناعية ANN. تم جمع البيانات المطلوبة مختبريا من خلال فحص 86 نموذج من التربة التي جمعت من مشاريع مختلفة في محافظة البصرة. البيانات التي جمعت من نتائج التجارب المختبرية استخدمت في ANN وكذلك في MLR . وقد تم اجراء فحوصات حد السيولة، دليل اللدونة وفحص بروكتر المعدل وكذلك فحص الـــ CBR على نماذج التربة المختلفة. في هذا البحث تم صياغة نماذج متعددة من ANN وMLR مع مجموعة مختلفة من المدخلات لتكون قادرة على تمييز تاثيرهم على قيمة الـــ CBR. وقد اخذت قيمة معامل الانحدار ($R^2$) والخطأ النسبي (RE%) ومتوسط مربع الخطأ (MSE) كمقياس لقياس قوة النماذج. وقد كان من الواضح ان جميع نماذج الـــ ANN كانت أفضل من نماذج MLR. وتحديدا فان نموذج الشبكات العصبية الصناعية ANN والذي استخدمت فية جميع المتغيرات كمدخلات اعطى أفضل النتائج من نماذج ANN الاخرى.

**الكلمات المفتاحية:** نماذج الانحدار المتعدد، محتوى الرطوبة الامثل، حد السيولة، دليل اللدونة، نسبة تحمل كاليفورنيا، تنبأ، الترب

## 1. Introduction

The CBR test had been firstly used in 1920s in the department of California State Highway division. In the 1940s, the CBR test was adopted by US Corp of Engineer for military airfield (Purwana *et.al.,* 2012). The results of the CBR test have become essential for Geotechnical engineering as well as for earth structures such highway, earth dams, abutments, embankments, bridge, additionally the fills at the rear of retaining walls (Yildirim and Gunaydin ,2011). The CBR tests can either be executed in the laboratory or in the field. The CBR test is typically carried out on compacted soil samples within the laboratory, but in the field, the CBR test is conducted on a ground surface, or level surface excavated in a test pit, bulldozer slice, or trench (Day ,2001). To determine the CBR value, the CBR test in the laboratory must commonly be carried out on molded samples of soils. CBR test is time-consuming and laborious; but occasionally the results from the test are not valid as a result of the laboratory that is in poor conditions. Moreover, if the soil that is available is of low quality, suitable additives are combined with soil and the strength of the soil that is the result will be evaluated by CBR value, which is cumbersome. The type of soil as well as the different soil properties effects on the CBR value (Zumrawi, 2011). Many researchers (Kin ,2006; Agarwal and Ghanekar,1970;Linveh,1989; Satyanarayana and Pavani, 2006;Patel and Desai,2010; Taskiran ,2010; Vinod and Reena,2008; Alawi and Rajab 2013; Al-Refeai and Al-Suhaibani,1997;Ramasubbarao and Siva,2013; Gunaydin, 2009; Yildirim and Gunaydin,2011; Venkatasubramanian and Dhinakaran,2011; Black 1962;Stephens,1990) have actually conducted studies to present that the CBR values affected by the characteristics and soil types.

Artificial Neutral Network (ANN) provides an interesting approach for the behavior of soil modeling (Shahin *et.al.*, 2001). ANN is a simulation that is oversimplified of human brain and is accepted as a trusted data modeling tool to fully capture and portray connections that are complex between inputs and outputs (Cabalar and Cevik,2009) . ANNs have already been successfully used to very nearly every Geotechnical engineering problem, since the early 1990s (Shahin,2008). Several studies already been been done to predict the values of CBR for soils. ANN model develop by (Taskiran,2010) based on CBR test results on fine-grained soils in Turkey. The proposed ANN model was successfully discovered to be able for the prediction of the CBR values for fine-grained soils. Another ANN model for the prediction of CBR of soil was developed by Kaur (Kaur,2011). Different parameters were used as inputs when establishing the ANN model. The outcomes attained from the model showed that the ANN model predicts CBR values rather effectively. (Gunaydin and Gunaydin,2011) showed that the results obtained from their constructed ANN models, in predicting the CBR values, showed a higher efficiency than other traditional models.

In this research, an MLR model and an ANN model, with respect to the benefits which realized above, were applied for some soils to predict it's CBR values. To do this, the results of CBR tests performed on 86 different soil samples with varying soil properties were utilized. Both of MLR and ANN models had four input parameters, namely dry unit weight, liquid limit, plastic index and optimum moisture content respectively, and an output parameter, CBR. The results of the MLR and ANN models were compared with the results attained from the experiments. The coefficient of determination, mean absolute error, and root-mean-square error, were utilized to assess the prediction performance of the MLR and ANN models. As a result, It is found that the predicted CBR values of the ANN model are much closer to the experimental values (measured) than those obtained from the MLR model. Furthermore, ANN prediction performs better than MLR based model for different soil samples taken from various locations in the Basrah's districts.

## 2. Database of testing soils

To be able to get data for the models establishment, 86 CBR test data that belong to different soil groups were selected among other tests. Tests were carried out for the viability evaluation of soils for using it as base material. Therefore, tested soils are collected from different districts in Basrah and tested in the laboratory of soil and roads, in the college of engineering, University of Basrah. These samples of soil are tested for CBR value, OMC, Maximum Dry Density, LL, PL, and PI. These tests were conducted as per standard procedure. Table 1 shows the details of the parameters used in the modeling . The properties consist of maximum dry density (MDD) and optimum moisture content (OMC) from the proctor compaction test (PCT) , the results includes also the index properties of soil such as liquid limit (LL) and plasticity index (PI). California Bearing Ratio (CBR) carried out for the arrived optimum moisture contents shown in table 1.

**Table 1, details of the parameters used in the modeling**

| Parameter used | Minimum | Maximum | Mean | SD |
|---|---|---|---|---|
| Input parameters | | | | |
| OMC | 13.0 | 19.0 | 15.5000 | 1.40243 |
| MDD | 1.62 | 1.96 | 1.7990 | 0.05762 |
| LL | 22.0 | 50.6 | 37.8965 | 4.88085 |
| PI | 4.0 | 30.0 | 15.5349 | 5.32419 |
| Output parameter | | | | |
| CBR | 6.0 | 25.0 | 12.3128 | 4.09059 |

### 2.1. Multiple regression model

The multiple regression (MLR) is an approach that is time-honored returning to Pearson's usage in 1908. It is applied to predict the variance in an interval dependent, dedicated to linear combinations of interval or independent variables that are dummy (Yilmaz and Yuksek,2008). The general-purpose of the MLR is to discover the relationship between some predictor variables or independent and a dependent or qualification variable (Yilmaz and Yuksek,2008). MLR analysis was carried through the application of SPSS 18 package to match the measured CBR value with the four soil index properties. The database used during developing the ANN models was used in the formation of the MLR model.

### 2.2. Artificial neural network Model

In the past decade, as a result of difficulties in solving the complex engineering systems, artificial neural network (ANN) has been applied by many researchers to analyze the behavior human brain and additionally nervous system. ANN model can be in a different prepared dependent on same basic structure. ANN structure has three main layers; a collection of input nodes, a layer or layers of hidden nodes, and a collection of output nodes. The possibility of using ANNs for the estimation of CBR was investigated by building various appropriate ANN models. Variables (parameters) that belong to two categories of soil index properties which reflect compaction properties and plasticity are implemented. Therefore, entirely four basic soil parameters such as liquid limit, plasticity index, dry unit weight, and optimum moisture content were taken into consideration as input parameters for the ANN models. The toolbox of ANN in MATLAB computer aided Software (Demuth and Beale, 2001) was used to perform the computations being essential. To develop the best proper ANN architecture in each model, the neurons's number in the hidden layer and different transfer functions was attempted for the purpose of getting a better prediction of CBR values. They had been, consequently, varied until the convergence had been achieved in the mean squared error.

## 3. Construction of artificial neural network model

A three-layer of back-propagation artificial neural network (BP-ANN) model could comprehend any continuous mapping. The three-layer BP-ANN model was shown in Fig. 1.

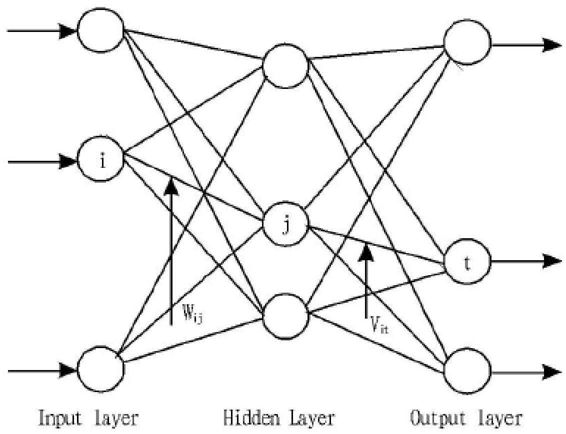



**Fig.1. Three-layer BP neural network model**

Where, $w_{ij}$ was the connection weight between the $i^{th}$ neuron of the input layer and the $j^{th}$ neuron of the hidden layer and $v_{jt}$ was the connection weight between the $j^{th}$ neuron of the hidden layer and $t^{th}$ neuron of the output layer. If the threshold value of the $j^{th}$ neuron in the hidden layer was supposed as θj and the threshold value of $t^{th}$ neuron in the output layer was supposed as $\gamma_t$, the input of the $j^{th}$ neuron in the hidden layer was as follows:

$$S_j = \sum_{i=1}^{n} w_{ij} x_i - \theta_j \quad \ldots\ldots(1)$$

Where $x_i$ was the input value of the $i^{th}$ neuron in the input layer and n was the neuron number in the input layer. The output was as follows:

$$b_j = f(s_i) \quad j = 1,2,3,\ldots\ldots, p \quad \ldots\ldots(2)$$

Where p was the neural number in the hidden layer and f was the activation function, its form was as follows:

$$f(x) = 1/(1+e^{-x}) \quad \ldots\ldots(3)$$

Its effect was to stimulate the nonlinear characteristics of biological neurons. The input of $t^{th}$ neuron in the output layer was as follows:

$$L_t = \sum_{i=1}^{p} \upsilon_{jt} b_j - \gamma_j \quad \ldots\ldots(4)$$

Its output was as follows

$$C_t = f(L_t) \quad \ldots\ldots(5)$$

In the calculation process, the stimulary function of the neuron in the hidden layer was adopted as S type and the stimulary function of the neuron in the output layer was adopted as linear type. In the interest of enhancing the performance of the network, the improved BP algorithm momentum method was usually used to reduce its possibility of falling into the local minimum value and increase the convergence speed, namely:

$$wij(t+1) = w_{ij}(t) + \Delta w_{ij}(t+1) + \mu \Delta w_{ij}(t) \quad \ldots\ldots(6)$$

$$\Delta w_{ij} = \eta(\partial E / \partial w) \quad \ldots\ldots(7)$$

Where μ was momentum factor and E was error function. The error function of moment t network was defined as follows:

$$E(t) = \frac{1}{2}\sum_{i=1}^{q}[y_i(t) - d_i(t)]^2 \quad \text{.......................................................}(8)$$

Where $y_j(t)$ was the actual output of the $j^{th}$ neural in the output layer at t moment, $d_j(t)$ was expected output of this moment, q was the neural number in the output layer. When E(t) was equal to or less than ε ( ε was the given error in advance), the network stopped training and here the network model was required.

# 4. Results and discussion

## 4.1. Structure of artificial neural network model

Due to different dimensions of original data and the obvious differences in order of magnitude for numerical values, standardization process should be made on original data first of all. In order to compare the imitative precision of the model under different reference factors , CBR was calculated in three instances: the first was only considering four factors of data which are soil index properties and compaction properties (OMC, MDD, LL, and PI) factors (denoted by CBR1), the second was considering soil index properties and a part of compaction properties ( MDD, LL, and PI) factors (denoted by CBR2), the third was considering soil index properties and the other part of compaction properties (OMC, LL, and PI) factors (denoted by CBR3). As for this model, three types of combination elements were taken as the input vectors of network model being 4, 3 and 3. The CBR requirement of was determined by weighing method was taken as the network output vector. According to three instances, the nodes in hidden layer were confirmed as 10 after many times training and comparison, which was to say the topological structures of the network model were 4-10-1, 3-10-1 and 3-10-1.

## 4.2 Comparison of coefficient of determination in models

The fitting equation of the fitting values, actual values, and the coefficient of determination ($R^2$) in three models were obtained by comparing the fitting values and actual values of the models. Table 2 shows the fitting equation and the coefficient of determination of the predictive values and the actual values in the models

**Table 2, Fitting equation and the coefficient of determination ($R^2$) of the predictive values and the actual values in each model**

| Model | Fitting equation | Coefficient of determination ($R^2$) |
|---|---|---|
| CBR1 | y = 1.10 x - 1.2 | 0.8672 |
| CBR2 | y = 0.65 x + 4.0 | 0.7624 |
| CBR3 | y = 0.45 x + 7.6 | 0.6439 |

From Table 2, it was shown that the coefficient of determination of model CBR2 and CBR3 were lower, and the coefficient of determination obtained from the predictive values and the actual values in the three models were increased in turn. Comparison of the coefficient of determination in the three models showed that the coefficient of determination of CBR1 was much greater than that of the other two models. With the increasing of the referential factors in constructing the model, the fitting precision of the model was gradually increased (CBR1 >CBR2 >CBR3). The coefficient of determination of model CBR1 was 0.8672, which indicated that when the effects of the soil index properties and compaction properties were considered enough, so it was a kind of the best model construction mode.

The multiple regression model for the soil samples revealed the following correlations:

CBR=42.538-22.625MDD+0.843OMC-0.101LL+0.079PI $\quad R^2=0.191$ ………..(9)

It is noticeable from the equation (9) that, the prediction equation for CBR according to multiple regression model cannot used as prediction equation due to its very low value of coefficient of determination ($R^2=0.191$) . Figure 2, represent the fitting error in the MLR model. The relative error of less than 5% in the MLR model was 24.41%.

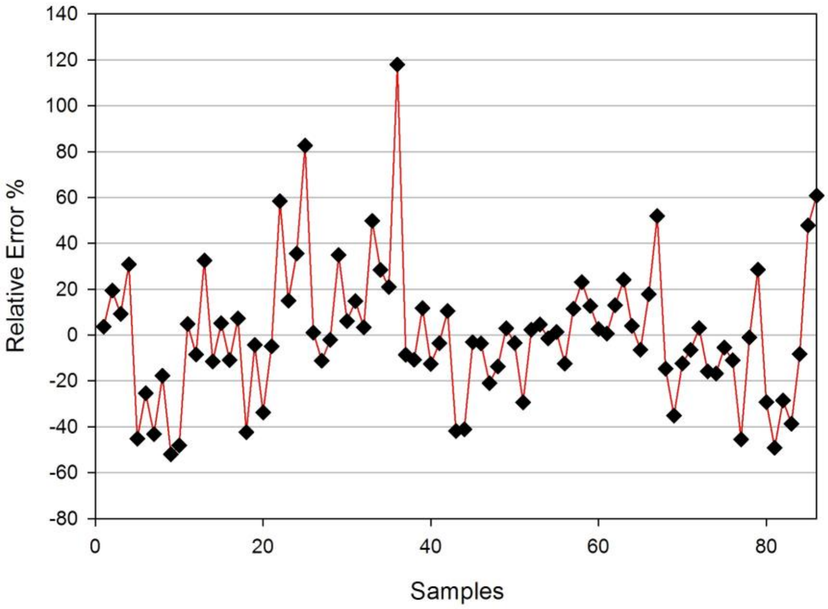
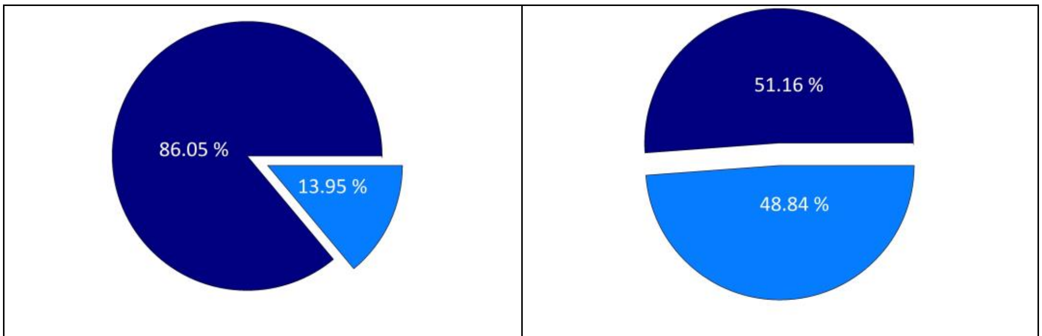


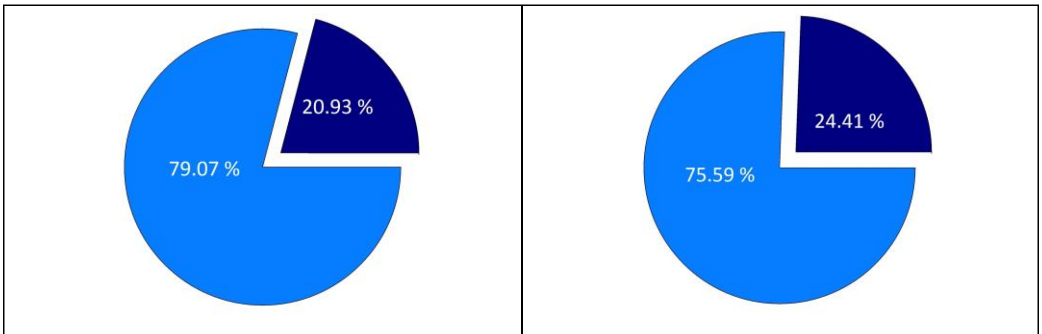

**Fig.2. Fitting error in MLR model**

**4.3. Comparison of relative error in ANN models**

An artificial neural network model was constructed by using the elements in two groups, which imitated the California bearing ratio. The imitative values were compared with the actual California bearing ratio in the same term, the results were shown in Fig.4 a-c. From the relative error percentage of less than 5% (Fig. 3 a-d), it was known that among 86 predictive values, the number with its relative error less than 5 % in CBR1 model was 74 (occupying 86.05% ), which in CBR2 model was 44 (occupying 51.16%) and that in CBR3 was 18 (occupying 20.93% ). The data showed that BP-network models after training could all be used in predicting of California bearing ratio. It was known from the analysis that the number with the relative error less than 5 % in three models had little differences, but their imitative precisions were CBR1 >CBR2 >CBR3. Considering the compaction soil properties, it was noticed that the imitative precision of CBR2 was higher than that of model CBR3.




*a- Relative Error (%) in CBR1 Model* | *b- Relative Error (%) in CBR2 Model*




*c- Relative Error (%) in CBR3 Model* | *d- Relative Error (%) in CBR Model*

**Fig.3. Relative error percentage of less than 5% in the models (dark blue color)**

## 5. Conclusion

The artificial neural network models were constructed by using three types of combination elements, including CBR1 (MDD, OMS, LL, and PI), CBR2 (MDD, LL, and PI) and CBR3 (OMC,LL, and PI) as well as make an MLR model (denoted by CBR) and the coefficient of determination ($R^2$) of the models were CBR1>CBR2 >CBR3>CBR. The artificial neural network model constructed by different elements of soil index properties and compaction properties.

The main objective of this study is to investigate the applicability of three Soft Computing Techniques, methods which is an artificial neural network (ANN) and Multiple Regression Models (MLR) for prediction of California bearing ratio. To achieve this, CBR test results of different soils, was conducted. Three ANN and one MLR models which have different input parameters were trained to establish the best interrelationship between basic soil properties and the parameter CBR. Performances of the models were examined in terms of some statistical verification criteria. The best results were produced for both MLR and ANN of CBRl model which have four input parameters. In addition, several performance indices (coefficient of determination, mean square error and relative error) were used to assess the prediction performance of the MLR and ANN models. In the ANN model, the $R^2$ values were obtained as 0.8672, 0.7624, and 0.6439, respectively, for the training samples and RE% less than 5%, obtained as 86.05%, 51.16%, and 20.93%, respectively, from the samples. In the MR model, the $R^2$ and RE% less than 5% values were obtained as 0.191 and 24.41%, respectively, for all samples. Based on the performance indices, the ANN model has shown higher prediction performance than the MLR model, which demonstrates the usefulness and efficiency of the ANN model. Therefore, the ANN model can be used to predict the CBR value of the soils included in this study as an inexpensive substitute for the laboratory testing, quite easily and efficiently.

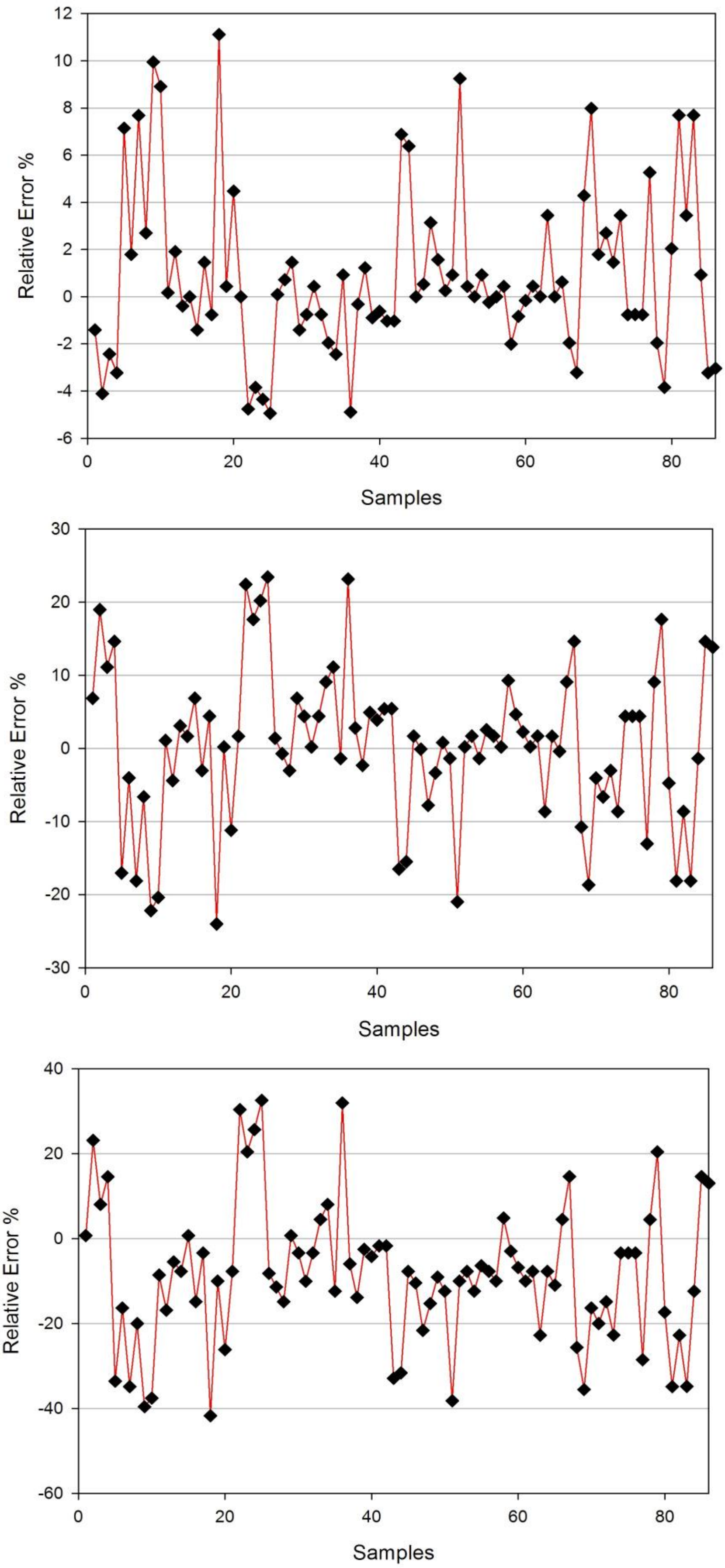



**Fig.4. Fitting error in BP neural network models (a-CBR1, b-BR2, and c-CBR3)**

## References

Agarwal KB, Ghanekar KD ,1970, Prediction of CBR from plasticity characteristics of soil. In: Proceeding of the 2nd south- east Asian conference on soil engineering, Singapore, June 11–15. Asian Institute of Technology, Bangkok, pp 571–576.

Alawi MH, Rajab MI ,2013. Prediction of California bearing ratio of sub-base layer using multiple linear regression models. Road Mater Pavement Des 14(1):211–219.

Al-Refeai T, Al-suhaibani A ,1997. Prediction of CBR using dynamic cone penetrometer. King Saud U J Eng Sci 9(2):191–204.

Black WPM ,1962. A method of estimating the CBR of cohesive soils from plasticity data. Geotechnique 12:271–272.

Cabalar AF, Cevik A ,2009. Modeling damping ratio and shear modulus of sand-mica mixtures using neural networks. Eng Geology 104:31–40.

Day WR ,2001. Soil testing manual ‘procedures, classification data, and sampling practices, USA, p 619.

Demuth H, Beale M.,2001. Neural network toolbox for use with MATLAB. Natick, Mass: The MathWorks Inc.; 840 pp.

Gunaydin O ,2009.Estimation of compaction parameters by using statistical analyses and artificial neural networks. Environ Geol 57:203–215.

Kaur S, Ubboveja VS, Agarwal A ,2011.Artificial neural network modeling for prediction of CBR. Indian Highw 39(1):31–37.

Kin MW ,2006. California bearing ratio correlation with soil index properties. Master degree Project, Faculty of Civil Engi- neering, University Technology Malaysia.

Linveh M ,1989.Validation of correlations between a number of penetration test and in situ California bearing ratio test. Transp Res Rec 1219:56–67.

Patel SR, Desai MD,2010.CBR predicted by index properties for alluvial soils of South Gujarat, Dec. 16–18. In: Proceedings of the Indian Geotechnical conference, India, pp 79–82.

Purwana YM, Nikraz HR, Jitsangiam P ,2012.Experimental study of suction-monitored CBR test on sand-kaolin clay mixture. Int J Geomate 3(2):419–422.

Ramasubbarao GV, Siva Sankar G,2013. Predicting soaked CBR value of fine grained and compaction characteristics. Jordan J Civil Eng 7(3):354–360.

Satyanarayana Reddy CNV, Pavani K ,2006.Mechanically stabilised soils-regression equation for CBR evaluation. In: Pro- ceedings of the Indian Geotechnical conference, Chennai, India, pp 731–734.

Shahin MA, Jaksa MB, Maier HR ,2008.State of the art of artificial neural networks in Geotechnical engineering. EJGE Special Volume Bouquet 08. http://www.ejge.com/Bouquet08/ Shahin/Shahin_ppr.pdf.

Shahin MA, Jaksa MB, Maier HR ,2001.Artificial neural net- work applications in Geotechnical engineering. Aust Geomech 36(1):49–62.

Stephens DJ ,1990. Prediction of the California bearing ratio. J Civil Eng S Afr 32(12):523–527.

Taskiran T,2010. Prediction of California bearing ratio (CBR) of fine grained soils by AI methods. Adv Eng Softw 41(6):886–892.

Taskiran T,2010. Prediction of California bearing ratio (CBR) of fine grained soils by AI methods. Adv Eng Softw 41(6):886–892.

Venkatasubramanian C, Dhinakaran G ,2011.ANN model for predicting CBR from index properties of soils. Int J Civil Struct Eng 2(2):605–611.

Vinod P, Reena C ,2008. Prediction of CBR value of lateritic soils using liquid limit and gradation characteristics data. Highw Res J IRC 1(1):89–98.

Yildirim B, Gunaydin O,2011. Estimation of California bearing ratio by using soft computing systems. Expert Syst Appl 38:6381–6391.

Yildirim B, Gunaydin O ,2011.Estimation of California bearing ratio by using soft computing systems. Expert Syst Appl 38:6381–6391.

Yildirim B, Gunaydin O ,2011.Estimation of California bearing ratio by using soft computing systems. Expert Syst Appl 38:6381–6391.

Yilmaz I, Yuksek AG ,2008. An example of artificial neural network application for indirect estimation of rock parameters. Int J Rock Mech Rock Eng 41(5):781–795.

Zumrawi M ,2012.Prediction of CBR from index properties of cohesive soils. In: Chang S-Y, Al Bahar SK, Zhao J (eds) Advances in civil engineering and building materials. CRC Press, Boca Raton, pp 561–565.