American International Journal of Research in Science, Technology, Engineering & Mathematics

Available online at http://www.iasir.net

ISSN (Print): 2328-3491, ISSN (Online): 2328-3580, ISSN (CD-ROM): 2328-3629

AIJRSTEM is a refereed, indexed, peer-reviewed, multidisciplinary and open access journal published by International Association of Scientific Innovation and Research (IASIR), USA

(An Association Unifying the Sciences, Engineering, and Applied Research)

Prediction of California Bearing Ratio of Soils Using Artificial Neural Network

Sudhir Bhatt[1], Pradeep K. Jain[2],

[1]Research Scholar, [2]Professor, Department of Civil Engineering,
Maulana Azad National Institute of Technology,
Bhopal, Madhya Pradesh, INDIA

***Abstract**: California Bearing Ratio (CBR) value is an important parameter in indexing the resistance offered by soils in the sub grade layers or in the foundation of a structure. The laboratory and field tests are extensively used for its determination to assess the strength of soils. An attempt has been made in the present study to estimate the soaked CBR from the test result values of sieve analysis, Atterberg limits, optimum moisture content (OMC) and maximum dry density (MDD). Simple, multiple linear regression (SRA, MLR) and artificial neural networks techniques were applied for analysis of data to develop the most suitable model. The data were collected from road construction works of Madhya Pradesh region of India. Simple linear regression analysis gave satisfactory correlation (R^2 =0.62) with MDD variable only, however artificial neural networks and multiple linear regression analysis resulted in strong correlations (R^2 =0.88-0.98) of CBR with sieve analysis, Atterberg limits, optimum moisture content and maximum dry density. The obtained correlation equation as a result of regression analysis is in satisfactory agreement with the test results. Gravel, sand content, OMC and MDD indicated to be the governing factors and developed model can be used for determination of CBR for a preliminary design of a project having financial and time limitations.*

***Keywords**: California bearing ratio (CBR), sieve analysis, Atterberg limits, Optimum Moisture Content,*

I. Introduction

California Bearing Ratio test is an important field/laboratory test in geotechnical engineering. It is performed to assess the resistance offered by sub grade layer of soil or in the foundation of a structure viz. earth dams, highway embankments, bridge abutments and retaining wall fills. The strength of soil can be considered to be indexed by its CBR values.

California Bearing Ratio (CBR) is defined as a ratio expressed in percentage of force per unit area required to penetrate a soil mass with a circular plunger of 50 mm diameter at the rate of 1.25 mm/min to that required for corresponding penetration in a standard material. The ratio is usually determined for penetration of 2.5 and 5 mm. Where the ratio at 5 mm is consistently higher than that at 2.5 mm, the ratio at 5 mm is used. The load value/corrected load value is taken from the load penetration curve and the CBR is calculated as follows (IS: 2720-Part XVI-1987).

$$CBR = \frac{Applied\ Stress\ in\ Experiment\ (or\ Load)}{Standard\ Stress\ (or\ Load)} \text{ x } 100 \tag{1}$$

The test can be performed in the laboratory on undisturbed or compacted remoulded specimens in water soaked or unsoaked conditions, however CBR values are highly dependent on the condition of the material at the time of testing. In the field, the test can be performed at ground surface or in a test pit, trench, on a level surface. The test on crushed stone is defined to have a CBR value of 100 percent and the corresponding load is called Standard Load.

To predict CBR value of soils, estimation models were developed by researchers and correlations were established relating various soil parameters. Reference [19] stressed on the changes of the obtained experimental values, which were caused by changing in the geographical area all over the world. For this he made to verify of correlations between a series of penetration tests and in situ California bearing ratio tests. Reference [1] has done a study on the estimation of CBR by using conic penetrometer experiment. Reference [14] calculated the CBR values by correlating the soil index properties and measured CBR values. Reference [8] studied the estimation of the compaction parameters with soil index properties by using statistical analysis and artificial

neural networks. Reference [26] estimated CBR from sieve analysis, Atterberg limits, maximum dry unit weight and optimum moisture content of the soils.
Artificial neural networks (ANNs) have been widely used for the estimation of various parameters in geotechnical engineering and other disciplines. In recent years, ANNs have found their way into the geotechnical areas ([2],[4],[6],[9],[10],[12],[16],[17],[18],[20],[21],[23],[25],[27]).

In the present course of this study simple linear regression (SLR), multiple linear regression (MLR) and artificial neural networks (ANNs) analyses were used to predict the soaked CBR values. Test result values of sieve analysis, Atterberg limits, maximum dry unit weight (MDD) and optimum moisture content (OMC) of soils were used for analysis. Soil classification and heavy compaction test results of 124 soil samples, consisting of three varieties of soils (CL, CI, SC) were used. The soils have wide range of gradation and index properties and are classified as per Indian Standard code IS 1498 -1970. The liquid limit, plastic limit, gradation, heavy compaction and soaked CBR tests were carried out as per relevant IS code of practices. The samples were collected from different locations during the construction of road and test result data were procured from M/s EDMG consultants of state of M.P. of India. Out of the 124 test results 114 test results were used for analysis for comparison and development of models of SLR, MLR and ANNs and ten (10) test results were used for cross validation to check the efficiency of the developed model. The ranges of various parameters of data are shown in Table 1.

Table1: Statistical parameter of Data used for analysis

Description	GRAVEL	SAND	FINE GRAIN	LL	PL	PI	OMC	MDD	CBR
Min	2.75	12.61	28.18	29	10.67	11.85	10	1.42	1.55
Max	31.14	51.5	74.73	47.16	26.56	30	21	2.03	22.4
Mean (μ)	22.10	30.66	47.20	38.53	17.66	20.88	14.49	1.63	3.77
S.D. (σ)	6.17	9.20	11.68	4.28	3.27	3.36	2.30	0.11	2.87

Simple linear regression analyses and multiple linear regression analysis were performed to establish relationship between soaked CBR and other variables and the analysis was carried out using MS Excel software. Artificial neural networks (ANNs) method was applied for the prediction of CBR values. The values of CBR obtained as output (estimated) from ANN models are compared with targeted values i.e. measured values and coefficient of regressions were determined. Levenberg-Marquardt back propagation ANN model of MATLAB was used for the computation of data and to determine the best model for prediction of CBR with classification, index properties and compaction parameters of soils.

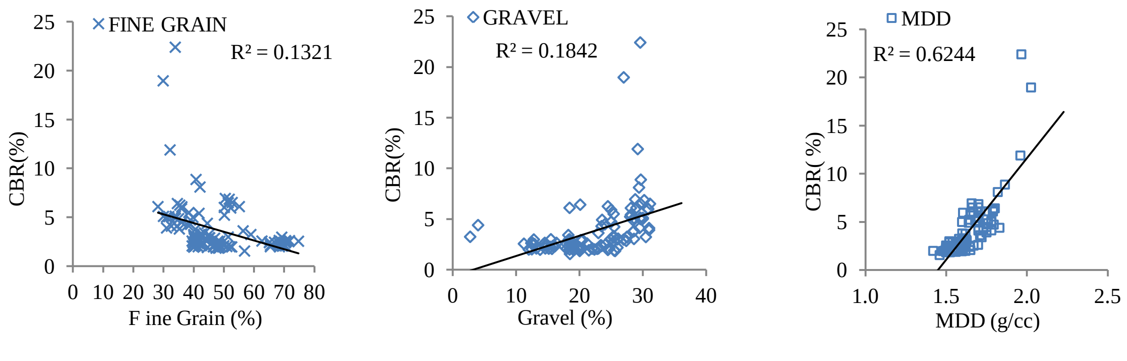
II. Statistical Analyses

A. *Simple regression analysis (SRA)*

All the test results consisting of gravel, sand, fine contents, liquid limit, plastic limit, OMC, MDD and CBR were analyzed by statistical method of least regression. The best linear fitting approximation equations having maximum R square values are determined and are shown below.

$$CBR = -0.0892\ FG + 7.9851 \qquad R^2 = 0.1321 \qquad (2)$$
$$CBR = 0.1996\ G - 0.6365 \qquad R^2 = 0.1842 \qquad (3)$$
$$CBR = 21.101 MDD - 30.56 \qquad R^2 = 0.6244 \qquad (4)$$

Figure1. a. Fine Grain vs. CBR   b. Gravel vs. CBR   c. MDD vs. CBR



The CBR values were correlated with all the variables independently and it is observed that the increase in fine grain content, liquid limit, plastic limit, plasticity index and OMC values of soils causes decrease in CBR values however increase in gravel, sand and MDD values resulted in improvement in CBR values. The SLR analysis gives the best $R^2 = 0.6244$ value when MDD is taken as an independent variable. All variables other than MDD resulted in low R-square values and therefore further analysis may be resorted to for development of better equations to make them useful for practical purposes.

B. Multiple Linear Regressions (MLR)

The multivariate linear regression analysis identifies the effect of two or more independent variables on dependent variable. The MLR analysis was carried out by taking all the independent variables in consideration at first and thereafter eliminating one by one forming various combinations to get the best correlation/determination coefficients. The results of MLR are given in Table 2. The independent variables were fine grained (FG), sand (S), gravel (G), liquid limit (LL), plastic limit (PL), plasticity index (PI), optimum moisture content (OMC) and maximum dry density (MDD) and CBR was taken as dependent variable.

It has been observed from the MLR analysis that the most important factors affecting the CBR are gravel, sand, fine grain contents, OMC and MDD (Model-E). It can be deduced from the following that the inclusion of other factors viz. liquid limit, plastic limit and plasticity index in the model 'E' resulted in a marginal improvement in regression values. However, elimination of any of these factors resulted in decrease in regression coefficients considerably as can be seen in models F, G, H and I.

Table 2: Correlation and determination coefficients after MLR analysis

Model	Dependent Variable	Independent Variable	R^2 Value	R Value
A	CBR	G,S,FG,LL,PL,PI,OMC,MDD	0.88460	0.94053
B		G,S,FG,LL,PL,OMC,MDD	0.88310	0.93974
C		G,S,FG,LL,PL,OMC,MDD	0.88309	0.93973
D		G,S,FG,PI,OMC,MDD	0.88250	0.93942
E		G,S,FG,OMC,MDD	0.88192	0.93910
F		S,FG,OMC,MDD	0.82473	0.90815
G		G,FG,LL,PI,OMC,MDD	0.79987	0.89435
H		G,FG,OMC,MDD	0.78518	0.88610
I		G,S,OMC,MDD	0.75258	0.86751

The predictive model for CBR containing the minimum variables and giving significant value of coefficient of determination derived by MLR analysis is given below, where MDD is in g/cc and all other parameters are in %.

$$CBR = -0.3776\ G - 0.4528\ S - 0.4094\ FG + 0.3487\ OMC + 24.7518\ MDD \qquad R^2 = 0.8819 \qquad (5)$$

The above equation can be taken as satisfactory for prediction of CBR and more reliable equations need to be evolved for better regression coefficient.

C. Artificial Neural Network (ANN)

Since early 1990s, artificial neural networks have been in use in analysing the geotechnical problems and demonstrated to be a superior predictive performance as compared to traditional methods. ANNs need no prior knowledge regarding the nature of the relationship between the input and output variables. This is one of the main benefits of ANNs when compared with most empirical and statistical methods. ANNs are a form of artificial intelligence and mimics the nervous system of the human brain. It consists of a series of processing elements (PEs) called nodes which are arranged in input, output and one or more hidden layers.

The propagation of information in an ANN starts at the input layer where the input data are presented. The network adjusts its weights on the presentation of a training data set and uses a learning rule to find a set of weights that will produce the input/output mapping that has the smallest possible error. This process is called 'learning' or 'training.' Once the training phase of the model has been successfully accomplished, the performance of the trained model needs to be validated using an independent testing set.

As the ANN is an alternate statistical method, the results should be compared in terms of statistical performance criteria. The correlation coefficient (R^2) and root means square error (RMSE) are mostly used for performance criteria evaluation of ANN models. The RMSE indicates the accuracy of approximation as overall, without indicating the individual data points. The successful application of ANN is that the network needs to be equally efficient for new data during testing or validation, which is called as generalization. The over fitting ratio (OR) is defined as the ratio of MSE for testing and training data and its value close to 1.0 shows good generalization of the model [7].

There are different methods for generalization like early stopping or cross validation ([3],[5],[6]). In cross validation an independent test set is used to assess the performance of the model at various stages of learning.

C.1 ANN Model

The available data set is normalised prior to training to obtain better convergence. The data set has a wide range of values for inputs and targets and is scaled between 0 and 1 using the following Eq. 6 ([21],[13]).

$$\text{Unormalized} = \frac{\text{Uactual} - \text{Umin}}{\text{Umax} - \text{Umin}} \tag{6}$$

Where, $U_{normalized}$ is the normalized value of the observed variable, U_{actual} is the actual value of the observed variable, U_{max} is the maximum observation value of the data set and U_{min} is the minimum observation value of the data set. The normalized data set was then used to train neural networks to obtain the final weights and in the end of the analyses, the network outputs were post processed to convert the data back into non-normalized units.
For developing the optimum model the available experimental data is randomly divided into two separate data sets; the training data set and the testing data set. Among 114 data sets 90% of the total data sets were randomly used in the training and 10% of it was used for the testing stage.
The remaining independent 10 nos. of test data set is kept aside for cross validation purpose to assess the performance of the optimal model. The network architecture involves the selection of input parameters, input layers, the number of hidden layer nodes and a combination of transfer functions between the layers. In the present study, weight percentages of fine grains i.e. silt and clay (FG), sand (S), gravel (G), liquid limit (W_L), plastic limit (W_P), plasticity index, optimum moisture content (OMC), maximum dry density (MDD) in different combinations forming different sets made up the input layer; and soaked CBR made up the output layer.
The number of hidden layers and number of neurons are varied to find the optimal structure with the goal to achieve convergence in the mean sum of squared errors and testing/training ratios of their MSE values near to one. The model having MSE ratio value one or near to one is good for generalisation. Linear, tan-sigmoid, log-sigmoid are the most commonly used transfer functions between the layers. In the present analysis a tan-sigmoid function is used for the hidden layer neurons and a linear function is used for the output neurons.
The feed forward neural networks with Levenberg - Marquardt back propagation algorithm consisting of multilayer perceptions were employed to estimate CBR of soil. The ANN toolbox and a written script in MATLAB environment has been implemented using Matlab (Math Works Inc., V7.11- R2010b).
The developed network models, their performance, over fitting ratio and determination coefficient are shown in Table 3.

Table 3: Performance of LM neural networks

Model	Independent Variable	Structure	Performance (MSE)		Over fitting Ratio	Correlation Coefficient (All)	Determination Coefficient
			Training	Testing			
I	G,S,FG,LL, PL, OMC, MDD	8-2-1	0.000778	0.000907	1.165	0.97874	0.9579
II	G,S,PI, OMC,MDD	5-4-1	0.000720	0.000752	1.044	0.98055	0.9615
III	G,S,FG, OMC, MDD	5-3-1	0.000930	0.00105	1.127	0.97472	0.9501
IV	G,S,OMC, MDD	4-4-1	0.000389	0.000398	1.023	0.98956	0.9792
V	OMC,MDD	2-3-1	0.00214	0.00193	0.902	0.94184	0.8871
IV (Cross Validation)	G,S,OMC, MDD	4-4-1	------	-----	-----	0.9426	0.8885

In these five different models the number of input independent variables change from eight to two and the output dependant variable is CBR. Several networks with different number of hidden neurons were trained and results for predicted and desired values were compared to determine the optimal neuron structure.
As seen from Table 3, Model IV seems to be the best model in terms of MSE values, over fitting Ratio (OR=1.023) and R^2 (0.9792) values. The value of over fitting Ratio (OR), very close to 1.0 for Model IV shows good generalisation of the model. Model IV has the four input variables viz. gravel (G), sand (S), optimum moisture content (OMC) and maximum dry density (MDD) and California Bearing Ratio (CBR) as output variable. The training, testing and All regression value data of model IV is shown in Fig. 2.
Model IV which showed the best performance is presented with new kept aside independent data (10 nos.) for evaluation of its efficiency. The results obtained for estimated and measured CBR values on this cross validation has a coefficient of correlation (R=0.9426). Figure 3 shows the regression plots for these data. If a suggested model gives R value, which is greater than 0.80, a strong correlation is assumed between the measured and estimated values ([24]). The model may thus be considered to be good and can be used for the prediction of soaked CBR values by presenting data viz. Gravel content, Sand content, OMC and MDD of soils.

Figure 2 Training, Testing and All R-value data of Model IV

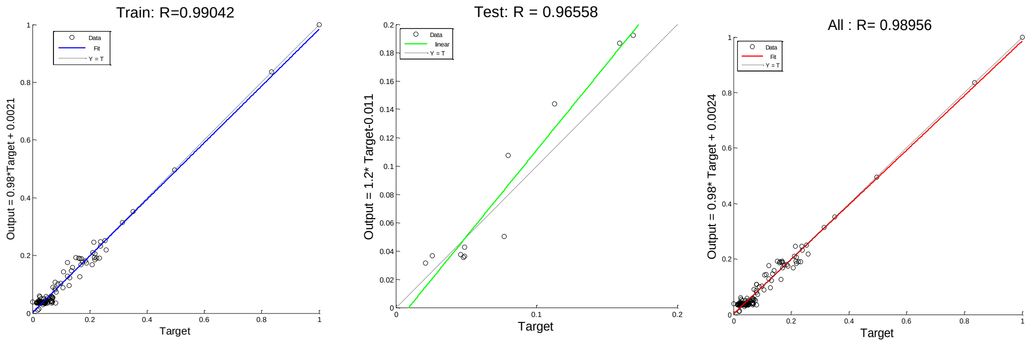
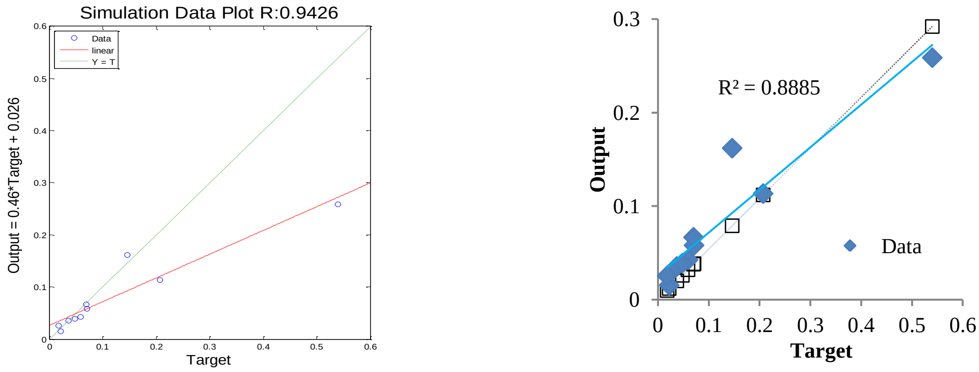


Figure 3 Regression plots for CBR values of cross validation data



The statistical parameters such as coefficient of regression (R^2), standard deviation (σ), standard error (SE) and mean (μ) of estimated and measured values obtained after ANN analysis were determined. The determination coefficient (R^2=0.8885) is higher than that obtained in MLR. The other statistical parameters are given below:

Table 4: Statistical parameters of cross validation output data

Model	Regression (R^2)	Standard Deviation (σ)	Standard Error (SE)	Mean (μ)
IV	0.8885	0.0765	0.0271	0.0817

III. Result Analysis and Conclusion

In the present paper an attempt has been made to estimate the value of soaked CBR using other parameters like gravel, sand, fine grain contents, Atterberg limits, OMC and MDD of soils. The available road data (124 nos.) was used to develop the SRA, MRA and ANN models. In the SRA analysis a satisfactory correlation coefficient for soaked CBR (R^2 = 0.6244) was obtained when MDD was used as independent variable. The MLR analysis gave higher correlation coefficients equal to or greater than 0.8819 for the models A to E having five or more independent variables. It is observed that independent variables less than five in numbers resulted in reduction in R^2 value. The best equation for soaked CBR value developed with minimum number of independent variables viz. gravel, sand, fine grain contents, OMC, MDD is given as equation (5) and this has a good regression value of 0.8819 (Model-E). All the other models in MLR also showed the satisfactory performance and inclusion of more input variables however did not result in any appreciable change in regression coefficient.

The ANN analysis showed an appreciable enhancement in regression values. Model IV gave the best determination coefficient R^2 = 0.9792 with an independent variables of gravel, sand contents, OMC and MDD as input and soaked CBR as the dependent output variable. The model proved to be working efficiently when an independent data set was presented to it and resulted in satisfactory correlation coefficient (R^2 = 0.8885). Thus the ANN model exhibited a higher performance in comparison to traditional statistical models viz. SRA and MLR and can therefore be used for the purpose of prediction of soaked CBR values reasonably. ANN technique has also proved its superiority over the other techniques of analysis.

References

[1] Al-Refeai, T., & Al-Suhaibani, A. (1996). Prediction of CBR using dynamic cone penetrometer. *Journal of King Saud University (Engineering Sciences), 2, 191–204 (A.H. 1417/1997).*

[2] Bai, J., Wild S., Ware, J. A. & Sabir, B. B. (2003). Using neural networks to predict workability of concrete incorporating metakaolin and fly ash. *Advances in Engineering Software, 34(11–12), 663–669.*

[3] Basheer, I. A. (2001). Empirical modeling of the compaction curve of cohesive soil. *Canadian Geotechnical Journal, 38(1): 29-45.*

[4] Basma, A. A., & Kallas, N. (2004). Modeling soil collapse by artificial neural networks. Geotechnical and Geological Engineering, 22, 427–438.

[5] Das, S.K., Basudhar, P.K. (2006). Undrained lateral load capacity of piles in clay using artificial neural network. *Computers and Geotechnics,* ***33****, 454–459.*

[6] Das, S. K., & Basudhar, P. (2008). Prediction of residual friction angle of clay artificial neural network. Engineering Geology, 142–145.

[7] Das Sarat Kumar, Sabat Akshaya Kumar (2008) Using Neural Networks for Prediction of Some Properties of Fly Ash, *EJGE Vol. 13, Bund. D*

[8] Gunaydin, O. (2009). Estimation of compaction parameters by using statistical analyses and artificial neural networks. *Environmental Geology Journal, 57,203–215.*

[9] Gunaydin, O., Gokoglu, A., & Fener, M. (2010). Prediction of artificial soil's unconfined compression strength test using statistical analyses and artificial neural networks. *Advances in Engineering Software, 41, 1115–1123.*

[10] Hurtado, J. E., Londono, J. M., & Meza, M. A. (2001). On the applicability of neural networks for soil dynamic amplification analysis. *Soil Dynamics and Earthquake Engineering, 21, 579–591.*

[11] Jaksa,M.B., Maier,H.R., Shahin, M.A. (2008). Future Challenges for Artificial Neural Network Modelling in Geotechnical Engineering. *The 12th International Conference of International Association for Computer Methods and Advances in Geomechanics (IACMAG) 1-6 October, 2008 Goa, India.*

[12] Kahraman, S., Gunaydin, O., Alber, M., & Fener, M. (2009). Evaluating the strength and deformability properties of Misis fault breccia using artificial neural Networks. *Experts Systems with Applications, 6874–6878.*

[13] Kayadelen C (2007) Estimation of effective stress parameter of unsaturated soils by using artificial neural networks. *Int J Numer Anal Meth Geomech. doi:10.1002/nag.660.*

[14] Kin, M. W. (2006). California bearing ratio correlation with soil index properties. *Master degree Project, Faculty of Civil Engineering, University Technology Malaysia.*

[15] Maier, H.R., Dandy, G.C. (1998). The effect of internal parameters and geometry on the performance of back-propagation neural networks: An empirical study. *Environmental Modelling & Software,* ***13****, 193-209.*

[16] Lai, S., & Sera, M. (1997). Concrete strength prediction by mean of neural networks. *Construction and Building Materials, 11(2), 93–98.*

[17] Lee, I. M., & Lee, J. H. (1996). Prediction of pile bearing capacity using artificial neural networks. *Computer and Geotechnics, 18(3), 189–200.*

[18] Lee, S. J., Lee, S. R., & Kim, Y. S. (2003). An approach to estimate unsaturated shear strength using artificial neural network and hyperbolic formulation. *Computer and Geotechnics, 30, 489–503.*

[19] Linveh, M. (1989). Validation of correlations between a number of penetration test and in situ california bearing ratio test. *Transportation Research Record(1219),56–67.*

[20] Najjar, Y. M., Basheer, I. A., & Naouss, W. A. (1996). On the identification of compaction characteristics by neuronets. *Computer and Geotechnics, 18(3), 167–187.*

[21] Rafiq, M. Y., Bugmann, G., & Easterbrook, D. J. (2001). Neural network design for engineering applications. *Computers and Structures, 79, 1541–1552.*

[22] Shahin Mohamed A, Jaksa Mark B. & Maier Holger R. (2001). Artificial neural network applications in geotechnical engineering, *Australian Geomechanics-March 2001, 49-62.*

[23] Sinha, S. K., & Wang, M. C. (2007). Artificial neural network prediction models for soil compaction and permeability. *Geotechnical and Geological Engineering, 26, 47–64.*

[24] Smith GN. Probability and statistics in civil engineering. London: Collins; 1986.

[25] Yang, Y., & Zhang, Q. (1998). The applications of neural networks to rock engineering systems (RES). *International Journal of Rock Mechanics and Mining Science, 35(6), 727–745.*

[26] Yildirim B., Gunaydin O. (2011). Estimation of California bearing ratio by using soft computing systems. *Expert Systems with Applications 38 (2011) 6381–6391 Elsevier.*

[27] Yuanyou, X., Yanming, X., & Ruigeng, Z. (1997). An engineering geology evaluation method based on an artificial neural network and its application. *Engineering Geology, 47, 149–156.*