Prediction of Standard Penetration Test (SPT) Value in Izmir, Turkey using General Regression Neural Network

Y. Erzin and Y. Tuskan

***Abstract*— Site exploration, characterization and prediction of soil properties by in-situ test are key parts of a geotechnical preliminary process. In-situ testing is progressively essential in geotechnical engineering to recognize soil type and stratigraphy alongside. In this study, a general regression neural network (GRNN) model was developed for estimating standard penetration resistance (SPT-N) value. In order to develop the GRNN model, 121 SPT-N values have been collected from 13 boreholes spread over an area of 17 km^2 of Izmir, Turkey including fine-grained deposits of mostly silt and clay with weathered gravel and sand. While developing the model, borehole location coordinates and soil type percentages were used as input parameters. The results obtained from the models were compared with those obtained from the field tests. Moreover, several performance indices, such as determination coefficient, variance account for, mean absolute error, root mean square error, and scaled percent error were calculated to check the prediction capacity of the model developed. The obtained indices make it clear that the ANN model has shown high prediction performance.**

***Keywords*— Standard penetration test, general regression neural network, in-situ test.**

I. Introduction

The complex geotechnical behavior and uncertainty of soil are a challenge for a simplified geotechnical model. In the case of nonlinear modeling, phase of using traditional modeling techniques of MR in predicting SPT; association between independent variables such as soil types and borehole coordinates cannot be performed and thus the development of a comprehensive model of SPT-N estimation is almost impossible for this technique. An alternative approach using artificial neural networks (ANN) is developed based on field SPT-N data to determine the organization and parameters of the model. The process is suitable to model complex problems where the relationships between the model variables and parameters are unknown for the model process of the complex problem. The use of the different numerical and graphical procedures can be served the geotechnical engineer as the exceptionally compelling instruments [1].

Bozbey and Toğrol [2], and Kayabaşı [3] studied the correlations of SPT-N value and pressuremeter parameters Schmertmann [4], Chang (1988) and Akca (2003) investigated empirical equations between cone penetration test (CPT) and SPT-N value in literature. Ahmed et al. (2013) proposed a unified approach correlating CPT and SPT readings for both crushable-calcareous and non-crushable-siliceous sands. The studied soil was qualified by the compressibility related to CPT in terms of the behavior index (Ic), and the compressibility related to SPT in terms of the mean diameter (D50). Mohamed and Vanapalli (2015) investigated the bearing capacity of shallow foundations of sands from CPT-SPT correlations.

In this work, a detailed study of the geological and geotechnical behavior of Zeytindagi Formation soil was conducted. The topography and geologic history of the Zeytindağı formation slope influences the behavior of the Manisa-Izmir State Road. The roadway is located near the crest of an ancient valley slope and is situated below the top of a plain between Izmir and Bornova fault zone. Theses slopes consist of ancient landslides created by the unpredictable earthquake effects in the post glacial period. Gediz Basin is bordered in the north by the Bornova Fault and in the south by the İzmir Fault. Sedimentary rocks with ages ranging from Lower Miocene to middle Miocene are exposed in İzmir and its environs. The Miocene aged deposits in the test area were affected by these two faults with the developed alluvium deposits. Historical earthquakes and morphological traces of the faults show that the Bornova and İzmir faults are active faults. A general aspect of the study area is depicted in Figure 1.

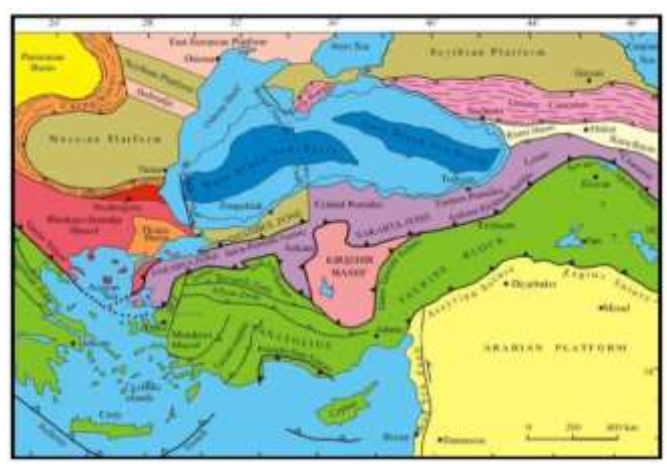

Fig. 1. Tectonic structure of the study area

A total of 13 boreholes in in the central station were drilled and 121 SPT were performed in the boreholes. A series of laboratory tests were applied for the slope stability project of Izmir-Manisa State Road. Sieve tests, Atterberg limit tests, natural moisture content tests were conducted on the borehole samples. The soil samples of the boreholes consist of fine-grained material 69 % and 51 % silt sized material; 11 of

Manuscript received Feb. 8, 2016.

Yusuf Erzin, Celal Bayar University, Engineering Faculty Civil engineering Department.

Yesim Tuskan, Celal Bayar University, Engineering Faculty, Civil engineering Department.

84 samples are mainly low plastic (CL), with the exceptions being 61 sample silts with high plasticity to low plasticity (ML and MH) and 12 samples that are high plastic clay (CH). All of the samples of the boreholes separate nearly parallel and under to the A line of the plasticity chart. The mean values of the liquid limit (LL) and plastic limit (PL) are 51 and 30 %, respectively with a mean plasticity index value of 19 %.

II. Standard Penetration Test

SPT is a commonly used method of investigating soil properties such as bearing capacity, liquefaction and site characterization. The test is applicable to a widely ranged soil conditions. Although the use of this test is prevalent in subsurface investigations, it has some majör drawbacks [1]. With the advance of modern geotechnical engineering the actual driving energy of the SPT entering the rods was measured easily as described in ASTM D4633 by energy measurement devices.

SPT-N was normalized to an overburden pressure of 100 kPa as part of semi-empirical procedure using the correction factor (C_N) proposed by [9]. The C_N value was limited to a maximum value of 1.70 as suggested by [10] and this factor commonly calculated from the following equation (1):

$$C_N = \frac{2.2}{1.2 + \frac{\sigma'_{v0}}{P_0}} \quad (1)$$

where, σ'_{v0} is effective overburden pressure and P_0 is 100 kPa (Kayen et al. 1992). Field SPT-N value was corrected for C_N using equation (2) proposed by [11] for measured values used in GRNN model:

$$N_{cor} = SPTN \times C_N \quad (2)$$

The negative pore pressure of the SPT sampler into saturated sand and silts may result in higher shearing resistance. Therefore, N_{cor} values greater than 15 and under ground water level, obtained from equation (2) were corrected using the equation of [12]:

$$N'_{cor} = 15 + 0.5 \times (N_{cor} - 15) \quad (3)$$

Figure 2 shows soil profile of the study area and SPT values with soil classification for borehole 1 (BH-1).

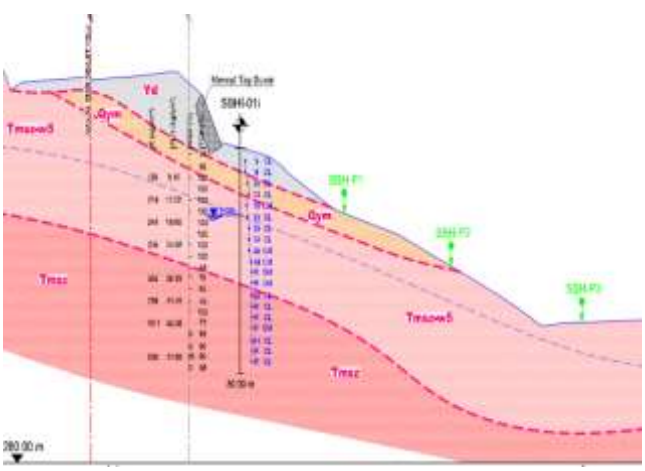

Fig. 2. SPT blow counts for Borehole 1 and a soil profile of study area

III. Generalized Regression Neural Network Model

Artificial Neural Networks (ANN) governs the learning of the connection between input and output variables in a way similar to the human brain, without preconditioned or optional assumptions (McCulloch and Pitts, 1943). The human brain resembles systems of learning, association, classification, making generalizations, estimation and optimization were provided by ANN as an information system (Haykin 1994).

The limitations of various numerical modeling techniques and fails of many mathematical models for highly non-linear behavior of soils are also considered to be complex, time-consuming and not always practical for geotechnical approaches. The limitations of various numerical modeling techniques and fails of many mathematical models for highly non-linear behavior of soils are also considered to be complex, time-consuming and not always practical for geotechnical approaches. In geotechnical engineering problems as with many areas of civil engineering, ANN has been used with great accuracy to predict and model the field tests.

A GRNN model with soil type and borehole location is designated for predicting the SPT-N values of the soils by using the neural network toolbox written in Matlab environment [13]. In order to develop the GRNN model, 121 SPT-N values have been collected from 13 boreholes spread over an area of 17 km^2 of Izmir, Turkey including fine-grained deposits of mostly silt and clay with weathered gravel and sand. In GRNN model, sand %, gravel %, silt % and clay % with borehole X,Y,Z coordinates from the reference point in the three geometric directions were used as the input parameters, while the measured SPT-N value was the only output parameter. The descriptive statistics of eight field and coordinate parameters are shown in Table I. These parameters are used by GRNN as input and output parameters for estimation of SPT-N values.

TABLE I: DESCRIPTIVE STATISTICS FOR COLLECTED FIELD DATA OF SPT-N.

	Minimum	Maximum	Mean	Std. Deviation
X	3	44	35.08	17.08
Y	3	44	11.85	17.07
Z	0.50	12	5.11	2.32
Gravel(%)	0	67	13.28	17.61
Sand(%)	3	72	32.18	16.18
Silt(%)	3	66	32.44	16.49
Clay(%)	0	67	22.27	16.61
SPT-N	1	50	19.45	14.56

The overall block diagram of the GRNN in its adaptive form for SPT-N estimation is depicted in Figure 3. GRNN provides estimation of continuous variables and converges to the underlying linear or nonlinear regression surface [14]. The advantages of GRNN can be summarized as follows: simple and fast training; independence from initial condition; nonlinear interpolation based on a sound mathematical foundation; training time required only for loading training matrix; it predicts very well for a large quantity of data because the estimate surface converges to the optimal regression surface for increasing sample size [15]

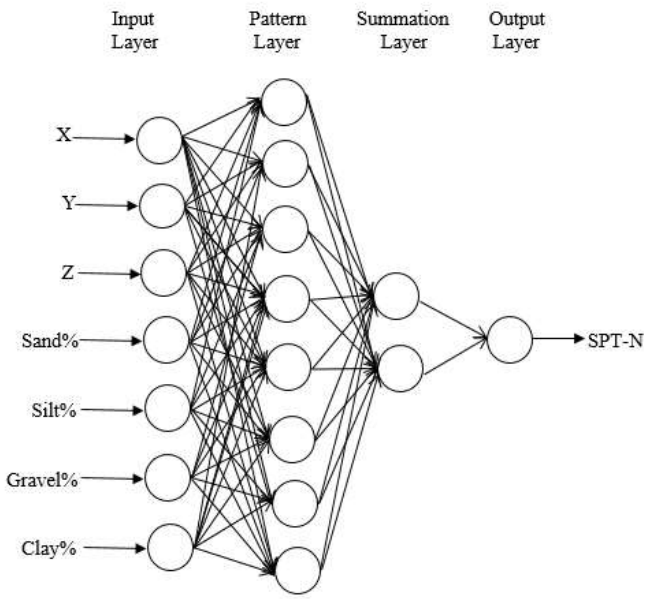

Fig. 3. GRNN model architecture for SPT-N estimation

A variation of the radial basis neural networks, GRNN, is based on kernel regression networks [16]. An iterative training procedure as back propagation network is not required in GRNN algorithm [17]. GRNN is a three-layer network where there must be one hidden neuron for each training pattern [18]. The first layer is connected to the pattern layer and each neuron presents a training pattern and its output. The pattern layer is connected to the summation layer. The summation layer has two different types of summation, which are a single division unit and summation units [19]. A radial basis and linear activation functions are used in hidden and output layers of the training process [19].

From probabilistic theory, for a model formed of the independent vector x and the dependent scalar y, the best prediction of y value consist of conditional expectation of x and is expressed as equation (4):

$$E(y|x) = \frac{\int_{-\infty}^{\infty} y f(x,y) dy}{\int_{-\infty}^{\infty} f(x,y) dy} \tag{4}$$

where $f(x, y)$ is the probability density function of GRNN. A sample of observations X_i and Y_i of x and y are approximated by given equation (5):

$$f(X,Y) = \frac{1}{2\pi^{0.5(d+1)}\sigma^{d+1}} \cdot \frac{1}{n}\sum_{i=1}^{n} exp\left[-\frac{D_i^2}{2\sigma^2}\right] \cdot exp\left[-\frac{(Y-Y_i)^2}{2\sigma^2}\right] \tag{5}$$

where d is the dimension of the vector x, n is the number of observation, and

$$D_i^2 = (X - X_i)^T \cdot (X - X_i) \tag{6}$$

A sample probability of width σ is connected to X_i and Y_i and the sum of those sample probabilities is assigned to $f(X, Y)$.

$\hat{Y}(X)$ can be expressed as equation (7) by substituting equation (5) into equation (4):

$$\hat{Y}(X) = \frac{\sum_{i=1}^{n} Y_i exp\left(-\frac{D^2}{2\sigma^2}\right)}{\sum_{i=1}^{n} exp\left(-\frac{D^2}{2\sigma^2}\right)} \tag{7}$$

In the GRNN algorithm known as supervised learning, the learning ability of a neural network depends on its architecture and algorithmic method of training process, whose optimal value of GRNN, spread parameter observed experimentally in Figure 4:

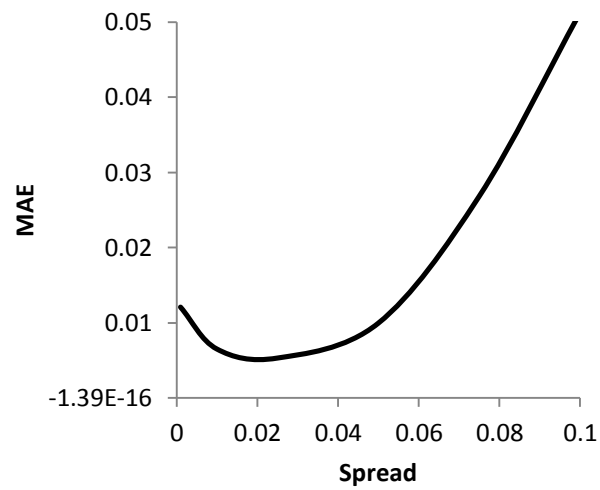

Fig. 4. The effect of spread parameter on GRNN performance

IV. Results And Discussions

A comparison of SPT-N values obtained from the field test with the SPT-N values predicted from the GRNN model is depicted in Figs. 5 and 6 for training and testing samples, respectively. It can be noted from the figures that predicted SPT-N values are quite close to the measured SPT-N values. An overall good agreement between predicted GRNN values and measured SPT-N has been found.

From here, it can be concluded that the SPT-N value of soils in this study could be predicted from easily determined soil graining properties and borehole coordinate using trained GRNNs values, with acceptable accuracy for a specific study area. It can be noticed from the Figure 3 and 4 that predicted SPT-N values from GRNN model are in good agreement with the measured SPT-N values, as R^2 of 0.9738 and 0.9348 for training and testing of ANN model respectively.

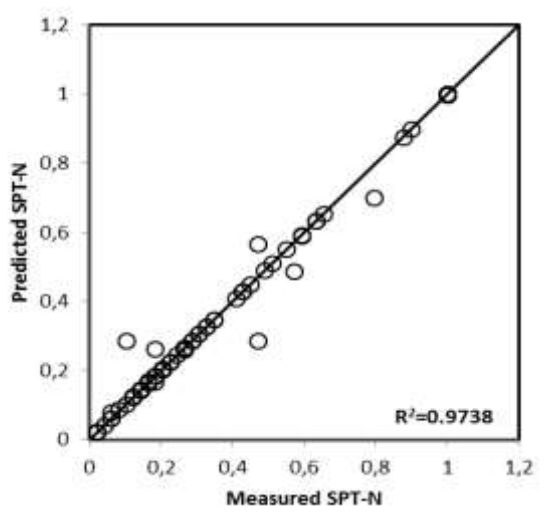

Fig. 5. The comparison of the measured versus predicted SPT-N values for training samples

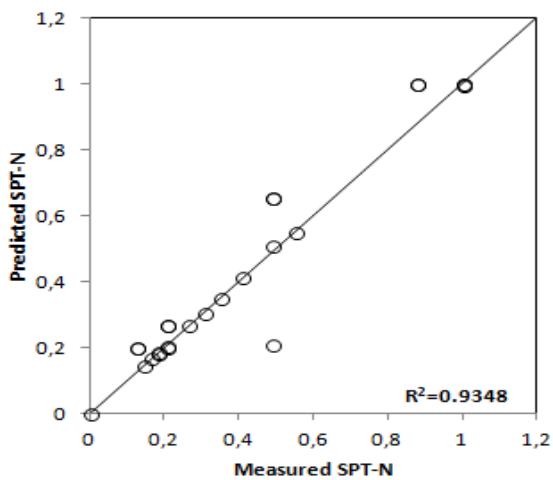

Fig. 6. The comparison of the measured versus predicted SPT-N values for testing samples

In fact, the coefficient of correlation between the measured and predicted values is a good indicator to check the prediction performance of the model. In this study, different performance indices, namely, variance *VAF*, defined by Eq. (8), the root mean square error *RMSE*, defined by Eq. (9), and mean absolute error *MAE*, defined by Eq. (10) were also computed to check the performance of the prediction capacity of predictive GRNN model developed in the study.

$$VAF = \left[1 - \frac{var(y-\hat{y})}{var(y)}\right] \times 100 \quad (8)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(yi - \hat{y}i)^2} \quad (9)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|yi - \hat{y}i| \quad (10)$$

where *var* describes the variance, *y* is the measured value, $\hat{y}$ is the predicted value, and *N* is the number of the sample. If *VAF* is 100 % and *RMSE* and *MAE* is 0, the model is considered as excellent. The performance indices calculated for the GRNN model developed in this study are given in Table 2. The ANN model exhibited high prediction performance based on the performance indices in Table 2.

TABLE II: PERFORMANCE INDICES (R^2, RMSE, MAE, AND VAF) OF THE GRNN MODEL

Model	Data	R^2 (%)	*MAE*	*RMSE*	*VAF* (%)
GRNN	Training set	97.38	0.01	0.04	98.59
	Testing set	93.48	0.05	0.08	92.69

In addition to the performance indices, to gain an insight into the capabilities of the proposed correlations, a graph between the scaled percent error (SPE) defined by (Eq. (11)), and cumulative frequency was depicted in Fig 7 for GRNN model developed.

$$SPE = \frac{(SPT-N_p - SPT-N_m)}{((SPT-N_m)_{max} - (SPT-N_m)_{min}} \quad (11)$$

where SPT-N_p and SPT-N_m are the predicted and measured SPT-N values, respectively; and $(SPT\text{-}N_m)_{max}$ and $(SPT\text{-}N_m)_{min}$ are the maximum and minimum measured SPT-N values, respectively. About 93% of SPT-N value predicted from the GRNN model fall into ± 10% of the SPE, indicating a perfect estimate for the SPT-N value of soil from the GRNN model. It can be concluded that the GRNN model developed in this study can be used for the estimation of the SPT-N values and so by using predicted SPT-N values, the bearing capacity, liquefaction potential and site characterization of selected study areas can be determined.

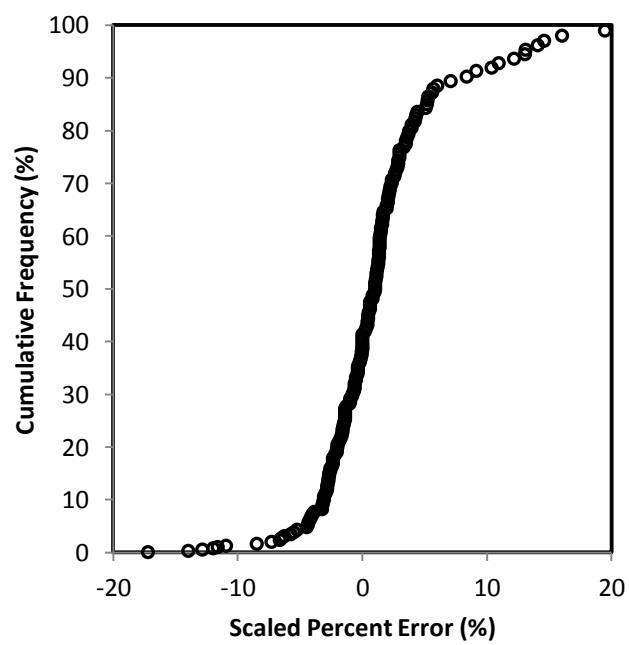

Fig. 7. Scaled percent error of the SPT-N predicted from the GRNN model

V. Conclusion

Standard methods of standard penetration resistance (SPT-N) determination are time-consuming and involve several

steps. In this study, a general regression neural network (GRNN) model was developed for estimating SPT-N value. In order to develop the GRNN model, 121 SPT-N values have been collected from 13 boreholes spread over an area of 17 km^2 of Izmir, Turkey including fine-grained deposits of mostly silt and clay with weathered gravel and sand. While developing the model, borehole location coordinates and soil type percentages were used as input parameters. The results obtained from the models were compared with those measured from the field tests**.** It was found that predicted SPT-N values are quite close to the measured SPT-N values.

In order to check the prediction performance of the GRNN model developed, several performance indices, such as R^2, VAF, MAE, MAPE, RMSE, and SPE were also calculated. The GRNN model has shown high prediction performance based on the performance indices. Thus, the developed GRNN model can be used to predict SPT-N from the soil graining parameters and borehole coordinates.

The performance of the GRNN model has also shown that the neural network is a useful tool to minimize the uncertainties encountered during the soil engineering projects. Thus, the usage of artificial neural network may supply new approaches and methodologies, and minimize the potential inconsistency of correlations.

REFERENCES

[1] D. Sarker, M.Z. Abedin,. "Applicability of Standard Penetration Test in Bangladesh and Graphical Representation of SPT-N Value," *International Journal of Science and Engineering Investigations*, vol. 4 pp. 55-59, 2015.

[2] I. Bozbey,E. Togrol,"Correlation of standard penetration test and pressuremeter data: a case study from Istanbul, Turkey," *The Bulletin of Engineering Geology and the Environment*, vol. 69, pp 505–515, 2010. http://dx.doi.org/10.1007/s10064-009-0248-4

[3] A. Kayabaşı, "Prediction of pressuremeter modulus and limit pressure of clayey soils by simple and nonlinear multiple regression techniques: a case study from Mersin, Turkey," *Environmental Earth Sciences* vol.66, pp 2171–2183, 2012. http://dx.doi.org/10.1007/s12665-011-1439-4

[4] J.H. Schmertmann, "Static cone to compute static settlement over sand," *ASCE Journal of Soil Mechanics and Foundations Divisions*, vol. 96 pp 1011–1043, 1970.

[5] M.F. Chang, "In-situ testing of residual soils in Singapore," *Proceedings 2nd International Conference Geomechanics in Tropical Soils*, pp 97-108, vol.1, Singapore, 1988.

[6] N. Akca, "Correlation of SPT-CPT data from the United Arab Emirates," *Engineering Geology*, vol. 67, pp 219–231, 2003. http://dx.doi.org/10.1016/S0013-7952(02)00181-3

[7] S. A. Ahmed, S. Agaiby, A.H. Abdel-Rahman, "A unified CPT-SPT correlation for non-crushable and crushable cohesionless soils," Ain Shams Engineering Journal, vol.5, pp 63-73, 2013. http://dx.doi.org/10.1016/j.asej.2013.09.009

[8] F.M.O. Mohamed, S.K. Vanapalli, "Bearing capacity of shallow foundations in saturated and unsaturated sands from SPT-CPT correlations," International Journal of Geotechnical Engineering, vol.9 pp 2-12, February, 2015. http://dx.doi.org/10.1179/1939787914Y.0000000082

[9] R.E. Kayen, J.K. Mitchell, H.B. Seed, A. Lodge, S. Nishio, R.Coutinho," Evaluation of SPT, CPT, and shear wave-based methods for liquefaction potential assessment using Loma Prieta data," National Center for Earthquake Engineering Research (NCEER) pp 177-204,vol.1, 1992.

[10] T.L. Youd, I.M. Idriss "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 And Geoenvironmental Engineering, ASCE, vol. 127(10), pp 817-833, 2001. http://dx.doi.org/10.1061/(ASCE)1090-0241(2001)127:10(817)

[11] R.B. Peck,W.E. Hanson, T.H. Thorburn, Foundation engineering, John Wiley & Sons, New York, 1974.

[12] K. Terzaghi, R.B. Peck, Soil Mechanics in Engineering Practice. 2nd edn. Wiley, New York, 1968.

[13] MathWorks, Neural Network Toolbox 7.0,MathWorks Announces Release 2010a of the MATLAB and Simulink Product Families, MathWorks, Inc. www.mathworks.com/trademarks, 2010.

[14] D. Specht, "A general regression neural network " IEEE transactions on neural network,1991. http://dx.doi.org/10.1109/72.97934

[15] G. Sarkar,S. Siddiqua, R. Banik and Md. Rokonuzzaman," Prediction of soil type and standard penetration test (SPT) value in Khulna City, Bangladesh using general regression neural network," Quarterly Journal of Engineering Geology and Hydrogeology, vol.48 pp 190-203, 2015. http://dx.doi.org/10.1144/qjegh2014-108

[16] H.K. Cigizoglu and M. Alp, "Generalized regression neural network in modelling river sediment yield," Advances in Engineering Software, vol.8 pp 37- 63, 2005.

[17] H.B. Celikoglu, H.K. Cigizoglu, "Public transportation trip flow modeling with generalized regression neural Networks," Advances in Engineering Software, vol.9 pp 38-71, 2007. http://dx.doi.org/10.1016/j.advengsoft.2006.08.003

[18] S.S. Salem and K. El-Zahaby, "Application of general regression neural networks (Grnns) in assessing liquefaction susceptibility," International Conferences on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics, pp 1-8 vol.4.18a, California, 2010.

[19] S.W. Hannan, R.R. Manza and R.J. Ramteke, "Generalized regression neural network and radial basis function for heart disease diagnosis," International Journal of Computer Applications, vol. 7 pp 7-13, October 2010. http://dx.doi.org/10.5120/1325-1799

Yusuf Erzin is Associate Professor in the Department of Civil Engineering at Celal Bayar University, Manisa, Turkey. He received his B.S. degree in Civil Engineering from Selçuk University, Konya, Turkey, and his M.S. and Ph.D. degrees in Civil Engineering from Middle East Technical University, Ankara, Turkey. His research interests include artificial neural networks, swelling soils, slope stability and crushing strength of sands. He has 20 papers published in many International journals.

Yesim Tuskan is Research Assistant in the Department of Civil Engineering at Celal Bayar University, Manisa, Turkey. She received B.S and M.S. degrees in Civil Engineering from Celal Bayar University, Manisa, Turkey. She is currently a Ph.D. student in the same university. Her research interests include computer programming, artificial neural networks, slope stability, liquefaction.