Multiple Regression and ANN (MLP) Model for Predicting
Swelling index of Ramadi Cohesive Soil

Ahmed H Abdulkareem¹ and Dalal O Aziz²

¹Assist. Prof. in Civil Engineering Dept., Engineering College, University Of Anbar, Ramadi, Iraq, ahm1973ed@gmail.com
²Civil Engineering, Ramadi, Iraq, Gullanhar0@gmail.com

Abstract. The response of expansive soils in the form of swelling and shrinkage due to changes in water content is frequently expressed as heaving and settling of lightly loaded structures such as roads, buildings, and canal linings. Compression and/or swelling indices are used for the calculation of the consolidation settlement of fine grained soils. They are conventionally determined by laboratory oedometer tests. Because of the time and expense involved in performing consolidation tests, it is often desirable to obtain approximate values of swelling index, Cₚ, by using other soil properties which are more easily determined. A database consisting of 102 consolidation tests from different parts of Ramadi city were compiled, identified, and used to conduct and utilized a statistical study to estimate suitable relations to determine the swelling index. this study develops the artificial neural networks based multi-layer perceptron, ANN-MLP and multiple regression, MR models. Neural model offers significant improvements in prediction accuracy of the statistical models.

Keyword: Swelling index, ANN-MLP, multiple regression, expansive soils, neural networks

1. Introduction
Expansive soils are considered one of the most difficult soils where foundations are constructed. Such structures include residential buildings, highway pavements, canal linings, clay liners, etc. In the United States, $9 billion are spent each year to repair the damage to the buildings, roads, pipelines, airports, and other facilities associated with expansive soils [1]. Buildings are exposed to risks when constructed on expansive soils due to the volume change of these soils for it causes major damage to the structure of the buildings and related services[2]. The response of expansive soils in the form of swelling and shrinkage due to the changes in water content is frequently expressed superficially as heaving and settlement of lightly loaded structures[3]. High plasticity of swelling soils that consist of clay minerals such as Montmorillonite absorb large amount of water [4]. The physical properties such as liquid limit, plasticity index and activity as well as the swelling characteristics of expansive soils are controlled depending on the type and amount of mineral, and the percentage of clay fraction [5].

Various researchers have proposed many empirical models to predict the swelling characteristics of expansive soils based on physical properties. The evaluation of swelling parameters like swell potential and swell pressure, includes both direct (physical measurements through laboratory tests) and indirect measurements involving the use of empirical models and correlations formulated based on basic soil properties. Table 1 shows several empirical equations of swelling potential predicting by authors for a variety of expansive soils [5].

To design approaches for estimating the surface heave and swelling pressure acting on a building, the numerical and analytical methods for swell potential of the soil is mainly used [6].

The main objective of this study is to develop a model that enables the prediction of swelling potential, Cₚ, of soils. Multiple regression and artificial neural network models were used to estimate and compare the models with the measured data capabilities.
Table 1. Swelling potential predicting by various researchers[5].

| Correlation                        | Author       | Reference |
|------------------------------------|--------------|-----------|
| $C_s (%) = Be^{A(PI)}$             | Chen         | [7]       |
| $C_s (%) = 7.518+0.323(C_f)$       | Muntohar     | [4]       |
| $C_s (%) = 60K (PI)^{2.44}$        | Holtz et al.,| [8]       |

$PI$: plasticity index; $A$: activity; $C_f$: clay fraction; $B$ and $K$ are empirical constants.

2. Experimental work

The collected data for various places in Ramadi city were obtained from field investigations conducted by laboratory testing (including standard one-dimensional consolidation). Testing Laboratories collected 102 data sets for the analyses and construction of the predictive model. Standard one-dimensional consolidation tests were carried out by ASTM D2435 and used to calculate the swelling index of the soil samples.

The Unified System (USCS) was used for classification of soil samples, used. The percentages of each soil type used in this study shown in Figure 1(a). The plasticity chart shown in Figure 1(b), indicates that the nature of the soil samples is inorganic with high plasticity.

The statistical parameters for $C_s$ shown in Table 2 that were determined by SPSS software package [9]. The independent value of $C_s$ shown in Figure 2 represents the normal distribution of histogram chart. In this case the analysis are acceptable and the values of skewness is 0.239 and kurtosis is 0.474 which is consider very low.

Table 2. The soil properties for all samples by the statistical parameters.

| Parameter | No. of Samples | Mean       | Median      | Mode  | Std. Deviation | Variance | Skewness | Std. Error of Skewness | Kurtosis | Std. Error of Kurtosis |
|-----------|----------------|------------|-------------|-------|----------------|----------|----------|------------------------|----------|------------------------|
| $C_s$     | 102            | 0.1108     | 0.1050      | 0.09  | 0.04337        | 0.002    | 0.501    | 0.239                  | 0.004    | 0.474                  |

3. Data analysis and interpretation

Simple regression analysis was performed in the first stage to establish the predictive models among the parameters obtained in this study. The linear, logarithmic and exponential functions were analyzed to find the relations between $C_s$ with other parameters. The linear relation for $C_s$ are highly correlated with the compression index, $C_c$ and initial void ratio, $e_0$ and to a lesser extent with the liquid limit, LL and plasticity index, PI. Therefore, the best linear regression models can be found by used these four variable as in Table 3.

In Figure 3, the swelling index relationship , $C_s$ with LL, PI, $e_0$ and $C_c$ respectively are shown. It was found that the relationship in each case justify the use of the linear model.
Figure 1. The soil classification of samples.

Figure 2. Histogram chart of $C_s$ values of samples.
3.1 The model of Multiple regression

In this study, multiple regressions were performed to obtain the predictive models among the parameters. The swelling index, $C_s$ predicted by multiple regression model were used more than one parameters and as shown in Table 4. To confirm the prediction performance of the model, the correlation coefficient is a good indicator. A good correlation coefficient of the relationship between the measured and predicted values were obtained, as shown in Figure 4.

In Table 5, the values account for (VAF) and the root mean square error (RMSE) were calculated to control the performance of predictive model as informed by Alvarez and Babuska (1999)[10]:

\[
VAF = \left[ 1 - \frac{\text{var}(y - y')}{\text{var}(y)} \right] \times 100
\]

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y - y')^2}
\]

where $y = $ observed value , and $y' = $ predicted value

The accuracy of the predictive performance models measured by the mean absolute percentage error (MAPE [6]. The expresses MAPE as

| Dependent Variable | Independent Variables | R   | R²   | Std. Error of the Estimate | Regression Equation | Number of Samples |
|--------------------|-----------------------|-----|------|---------------------------|--------------------|------------------|
| $C_s$              | LL (%)                | 0.776 | 0.602 | 0.02750           | $C_s = 0.002LL\% + 0.008$ | 102              |
|                    | PI (%)                | 0.780 | 0.608 | 0.02728           | $C_s = 0.002PI\% + 0.036$ | 102              |
|                    | $C_c$                 | 0.927 | 0.860 | 0.01630           | $C_s = 0.281C_c - 0.002$ | 102              |
|                    | $e_o$                 | 0.875 | 0.766 | 0.02110           | $C_s = 0.327e_o - 0.113$ | 102              |

**Table 3.** Calculate, $C_s$ by linear regression models.

**Figure 3.** Swelling index relationships, $C_s$ with LL, PI, $e_o$ and $C_c$. 
Table 4. Calculate, $C_s$ by MLR models.

| Dependent Variable | Independent Variables | R   | $R^2$  | Std. Error of the Estimate | Regression Equation | Number of Samples |
|--------------------|-----------------------|-----|--------|----------------------------|---------------------|-------------------|
| $C_s$              | $C_c$, LL (%)         | 0.943 | 0.882  | 0.01460                    | $C_s=0.007 - 0.001LL\% + 0.395C_c$ | 102               |
| $C_s$              | $C_c$, PI (%)         | 0.936 | 0.871  | 0.01543                    | $C_s=-0.005 - 0.001PI\% + 0.36C_c$ | 102               |
| $C_s$              | $C_c$, $e_o$          | 0.931 | 0.859  | 0.01596                    | $C_s=0.044 - 0.123e_o+0.377C_c$ | 102               |

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{A_i - P_i}{A_i} \right| \times 100$$

where $A_i$ = The value of actual, and $P_i$ = The value of predicted.

Table 5 shown the values of RMSE, VAF, and MAPE which produced the high performances prediction.

3.2 The models of ANN

Artificial neural networks are a form of artificial intelligence that attempts to mimic the behavior of the human brain and neural system. The use of ANN's has increased in many areas of engineering in last few years [11]. As an alternative for nonlinear regression or cluster analysis techniques, neural networks are often used [12]. The confident network defined by three essential elements: transfer function, network production, law learning [13]. Multi- Layer Perception is one of the most commonly used patterns in the architecture of neural networks to solve problems as classification pattern [14]. It works with a strong generalization capability for inaccurate data entry. MLP approach is more distributed due to the use of the production output by linear combinations of the contract hidden layer output.

The data was divided into three sets as training of 70%, test of 15%, and check of 15% of data. Matlab 7.1 (2005) software was used in the neural network analyses with nutrition network forward three layers; one input layer (2 or 3 neurons), one hidden layer (9 neurons for MLP) and one output layer (Figure 5) [15]. It was selected as the number of neurons in the hidden layers of a series of test flights of the networks and the presence of 1 neuron to neuron 50 to get the number of neurons in the network with minimum error. The network parameters of the learning rate and momentum to 0.01 and 0.9, respectively were used in the analysis for all layers [6].

3.2. ANN-MLP Model

The most research applications such as engineering, mathematical modelling used Multi-layer perceptron (MLP) as network models. The feed forward neural network modules are arranged in the nutrition layers [6]. MLP model consist of one layer of input, one or more layers of hidden, and one layer of output. Through some nonlinear functions, the MLP transformed from $n$ inputs to $l$ outputs. The network output calculated by [6]:

$$X_o = f(\sum_h X_h W_{ho})$$  \hspace{1cm} (4)

where $f(\cdot)$ = The function of activation $X_h$ = The measured of $h$th node of hidden layer, and $W_{ho}$ = The connection between $h$th node and $o$th node of output layer.

The measured function of sigmoid as follows[6]:

$$X_o = \frac{1}{1+\exp(-\sum_h X_h W_{ho})}$$  \hspace{1cm} (5)
Figure 4. The observed and predicted values of $C_s$ relations by MLR models.

Table 5. The values of RMSE, VAF, MAPE, and $R^2$ for predictive models, $C_s$

| Type     | Predictive Model                         | RMSE  | MAPE% | VAF%  | $R^2$ |
|----------|------------------------------------------|-------|--------|-------|-------|
| MR(1)    | $C_s = 0.007-0.001LL\% + 0.395C_c$       | 0.0151| 12.37  | 88.14 | 0.882 |
| ANN-MLP(1)|                                          | 0.0112| 9.48   | 92.74 | 0.949 |
| MR(2)    | $C_s = -0.005-0.001PI\% + 0.36C_c$       | 0.0162| 13.43  | 87.06 | 0.871 |
| ANN-MLP(2)|                                          | 0.0123| 9.70   | 92.31 | 0.934 |
| MR(3)    | $C_s = 0.044-0.123 e_c+0.377 C_c$        | 0.0157| 12.79  | 85.92 | 0.859 |
| ANN-MLP(3)|                                          | 0.0117| 8.17   | 92.11 | 0.939 |

Figure 5. A model of MLP artificial neural network [6].
The error between calculated output and target value found by [6]:

\[ E = \frac{1}{2} \sum_{s} \sum_{o} \left( t_{o}^{(s)} - X_{o}^{(s)} \right)^{2} \]  

(6)

where

\( N \) = The patterns number for data set, and

\( L \) = The output nodes number.

The gradient descent backpropagation (BP) algorithm is used to adjust the weights. The training data of algorithm requires a set of patterned values input and target, \( t_{o} \). MLP starts the training process where a random set of initial weights and then continues to set of \( W_{ho} \) and that of \( W_{ih} \) are optimized to a specific error threshold is achieved in advance between the \( X_{o} \) and \( t_{o} \) [6]. The connection between the nodes are adopted to the BP algorithm by:

\[ \Delta W_{ho} = \eta \frac{\delta E}{\delta W_{ho}} = \eta \delta_{o} X_{h} \]  

(7)

\[ \Delta W_{ih} = \eta \frac{\delta E}{\delta W_{ih}} = \eta \delta_{h} X_{i} \]  

(8)

where

\( E \) = The cost function of error as equation 6,

\( \delta_{o} = X_{o}^{'} (t_{o} - X_{o}) \), and

\( \delta_{h} = X_{h}^{'} = \sum_{o} \delta_{o} W_{ho} \)

where \( X_{o}^{'} = X_{o} (1 - X_{o}) \) and \( X_{h}^{'} = X_{h} (1 - X_{h}) \) when the activation function of sigmoid used[6].

The measured and predicted values of \( C_{s} \) relations are shown in Figure 6. It can be seen that the MLP-ANN model is very acceptable to predict \( C_{s} \). The values of \( RMSE, VAF, MAPE \), and \( R^{2} \) is presented in Table 5.

Figure 6. \( C_{s} \) relations by the models of ANN-MLP.
4. Discussion of results

To predict the swelling index, $C_s$ of Ramadi cohesive soils, the models of MR and ANN-MLP were utilized. It can be observed that there are relations between $C_s$ and LL, PI, $e_a$ and $C_y$. The best relationship for $C_s$ was with $e_a$ and $C_y$ was obtained by using linear regression models.

To predict of $C_s$ using two or three inputs and one output expressed by the MR and the ANN-MLP models conducted as follow:

a. MR models have higher predictive performance compared to simple linear regression models
b. A more reliable prediction revealed by ANN-MLP compared with the MR model.

So, the indicators of the VAF, RMSE and MAPE and $R^2$ to predict $C_s$ values for ANN-MLP model are higher than MR model and the best performers.

5. Conclusions

Statistical parameters and correlations of soil properties determined from more than 102 consolidation tests on samples from undisturbed soil concluded in the development of useful and simple relationships, suitable for estimating $C_s$ of the soil where the various soil index properties. It was found that the swelling index, $C_s$ is greatly expressed for $e_a$ and $C_y$ by using the models of linear regression.

A high performance for ANN-MLP model compared with MR model to predict $C_s$ values for Ramadi cohesive soils was obtained that will provide a good way to minimize the potential inconsistency for correlations. So, there is a possibility to use the proposed empirical relations to calculate $C_s$ values of soils. In this study, the analyzed data is relatively specific area. Therefore, it can be used the practical outcome of proposed equations with acceptable accuracy at the stage of the preliminary design

References

[1] Jones D E and Jones K A 1987 Treating expansive soils Civil Engineering ASCE Vol. 57 No. 8
[2] Arya A and Shahaboddin S 2009 A micromechanical approach to swelling behavior of unsaturated expansive clays under controlled drainage conditions Applied Clay Science Vol 45, issues 1-2 pp 1-104
[3] Shi B, Jiang H, Liu Z and Fang H Y 2002 Engineering and geological characteristics of swelling soils in China Engineering Geology Journal Vol 67 pp 63–71
[4] Munthar A S 2000 Prediction and classification of swelling clay soil Swelling soils recent advances in characterization and treatment London pp 25-36
[5] Israr J, Farooq K and Mujtaba H 2014 Modelling of swelling parameters and associated characteristics based on index properties of expansive soils Pak. J. Engg. & Appl Sci Vol 15 July pp 1-9
[6] Yilmaz I and Kaynar O 2011 Multiple regression, ANN(RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils Expert systems with applications 38 pp 5958-5966 www.elsevier.com/locate/eswa
[7] Chen F H 1975 Foundation on swelling soils Development on Geotechnical Engineering 12 Elsevier Scientific Co Amsterdam pp 150-158
[8] Holtz W G and Gibbs H J 1956 Engineering properties of swelling clays Transaction ASCE 121 pp 641–677
[9] SPSS Inc 2011 PASW Statistics for Windows Version 21 Chicago: SPSS Inc
[10] Alvarez G M and Babuska R 1999 Fuzzy model for the prediction of unconfined compressive strength of rock samples International Journal of Rock Mechanics and Mining Sciences vol 36 pp 339–349
[11] Yoon G L, Kim B T and Joen S S 2004 Empirical correlations of compression index for marine clay from regression analysis Canadian Geotechnical Journal Vol 41 pp 1213-1221
[12] Shahin M A, Jaksa M B and Maier H R 2000 Predicting the settlement of shallow foundations on cohesionless soils using back-propagation neural networks Research Report No. R 167 The University of Adelaide, Adelaide
[13] Simpson P K 1990 Artificial neural system-foundation, paradigm, application and implementation New York Pergamon Press.
[14] Cohen S and Intrator N 2002 Automatic model selection in a hybrid perceptron/radial network Information Fusion: Special Issue on Multiple Experts 3(4) pp 259–266.
[15] Matlab 7.1 2005 Software for technical computing and Model-Based Design The MathWorks Inc.