Estimation of Soil Compression Coefficient Using Artificial Neural Network and Multiple Regressions

Farzaneh Namdarvand *1, Alireza Jafarnejadi 2, Gholamabbas Sayyad 3

1. Department of Soil Science, Science and Research Branch, Islamic Azad University, Khouzestan, Iran
2. Soil and Water Research Department of Agricultural and Natural Resources Research Center of Khouzestan, Iran
3. Department of Soil Science, Shahid Chamran University, Ahwaz, Iran

*Corresponding Author email: farzaneh.namdar@gmail.com

ABSTRACT: Measurement of some significant properties of soil might be difficult, costly and time-consuming. Thus, estimation of these characteristics using conveniently measurable soil properties may be useful. In this research, it is attempted to evaluate and examine the artificial neural network technique and multiple regression in order to measure the soil compression coefficient using conveniently measurable soil properties. A total of 100 soil samples were taken randomly from various areas of Ahwaz and the percentage of clay, silt, sand, wet bulk density, dry bulk density, friction coefficient, viscosity coefficient, plastic limit were determined as conveniently measurable properties (dependent variable) and the compression coefficient as costly measured properties (independent variable). In order to form Annual Neural Network, instructional algorithm Mark- Laurinburg and Perseptron stricter was use and the stepwise method was used in order to make regression transfer functions. The compression coefficients of soils were mean 0.16, at least 0.11 and at most 0.25 and depended on clay soil class. The results showed that the compression coefficient r=0.63 and MSE=0.006 was determined by neural network method. In the regression method, it was measured as r=0.47 and MSE=0.002. By comparing the values of correlation coefficient and error square mean by two using methods, it was revealed that artificial neural network has the least error and the most accuracy. Therefore in the study area the practice of this method is recommended for estimating the compression coefficient.

Keywords: Soil Physical Properties, Khuzestan, Compression Coefficient, Multivariate Linear Regression, Artificial neural Network

INTRODUCTION

Soil Sampling along with laboratory or field measurement is very costly and time-consuming. Further, due to temporal and spatial variations, an appropriate assessment of the properties involves a massive spending (Tahooni. 2002). Hence, there have been many attempts to obtain the possible relations between costly measured properties with the other properties of soil which in pedology are known as transfer functions. The transfer functions convert the basic data obtained from soil studies in to properties obtained from soil (it involves much cost and time). In fact, the transfer functions have been developed to estimate the chemical, biological and mechanical properties. There are various methods for fitting these functions which can be referred as multivariate regression and computational intelligence technique (artificial neural network) (Jemsi. 2011). The compression coefficient is one the costly measured properties of soil that its direct measurement in the laboratory is too costly and time-consuming. The estimation of this characteristic using the conveniently measurable soil properties more often is possible. Once the saturated soil is loaded, first all the loadings are tolerated by pore water pressure, leading to increase of pore water pressure. If there is any drainage, the soil volume is decreased over time that is followed by soil deposition. This deposition is induced by compression coefficient (Tahooni. 2002). A wide range of soil textures can be often determined based on the degree of compression coefficient. In sandy texture the range of compression coefficient is within 0.01 to 0.1 and in clay texture it ranges 0.1-2.5 (Behnia. 2012). Generally, the compression coefficient changes the agriculture soil texture and it also alters the order, size and shape of fine soil and empty space between them (Bouma. 1988). Compression coefficient of soil is created as a result of natural incidents such as impact caused by rainfall, swelling soil or as a result of man-made events such as travelling of
tractors and other agricultural machineries (Levine, 1996). Therefore, it is very necessary to find a suitable way for predicking the degree of this coefficient with the lowest cost and doing heavy tests. In this regard, one of the proper tools is the artificial neural network so that their development has led to use of this instrument in the different sciences such as modeling of dynamical, physical and chemical properties and optimization (Kia, 2010). Shelmani et al., (2010) have concluded that the neural network method has a higher accuracy over regression one and their coefficient of determination ($R^2$) and root mean square error (RMSE) are 0.89 AND 0.24 respectively (Shelmani, 2010). In other research, Bolouri et al., (2000) by using one neural network and laboratory data include liquid limit, plastic limit. Porosity ratio, the amount of clay soil density showed that the proposed model has an acceptable accuracy and efficiency with correlation coefficient(R) 0.92 and root mean square error (RMSE) 0.021. Daryaei and Kashef pour (2000) in their study on neural network and comparing of it with empirical relation by examining the correlation between fine soil physical properties and compression index have reported that the artificial neural network estimates the compression index with higher accuracy and lesser error percentage with coefficient of determination ($R^2$) 0.64 root mean square error (RMSE) 0.3 (Blouri Bazaz. 2010). In this study, the estimation of some mechanical and physical soil properties using artificial neural network, the predication of some physical soil properties (soil shear resistance, weighted mean of fine soil diameter) is obtained better than regression methods. The values of correlation root mean square error (RMSE) for estimating the soil shear resistance by neural network were 0.89 and 0.006 respectively (Basalatpoor. 2011). In this research it is attempted by estimating the compression coefficient using the available data from soil physical properties and using the artificial neural network and multivariate regression, the accurate results to be obtained.

**METHOD AND MATERIAL**

In this study, from the measured data of 100 soil samples which were randomly taken by Geotechnical Consultant Engineering CO were used. The physical soil properties used in this research are sand percentage, clay percentage, silt percentage, dry bulk specific gravity, wet bulk specific gravity, swelling indication, compression indication, friction angle, viscosity coefficient, liquid limit, plastic limit, plastic indication. The compression coefficient was measured by assisting the confirmation test (Behnia. 2012).

**Normalization of data**

The data obtained from laboratory experiment was first classified in EXCEL software and then the unrelated data was specified by BOXPLOT methods. In order to use the data in the neural network, first the data obtained from both conveniently measured and costly measured properties were normalized by Komaron formula in the EXCEL software (Jemsi S. 2011).

$$X_{norm} = 0.5 \left( \frac{X_0 - \bar{X}}{X_{max} - X_{min}} \right) + 0.5$$  \hspace{1cm} (1)

Where, the $X_{norm}$ is a normalized value of each conveniently measurable input ($X_0$). The ($X_0$) is the value of each conveniently measurable input, $\bar{X}$ is the data mean, $X_{max}$ is the maximum data and $X_{min}$ is the minimum data. The normalized properties that include clay percentage, silt, sand, wet bulk specific gravity, dry bulk specific gravity, friction angle and viscosity amount were transferred as input in to INPUT part in the MATLAB software and the normalized data transferred as network real output in to OUTPUT part.

**Neural Model**

In neural network modeling, the MATLAB software as well as multilayer perceptron (MLP) model was used. These networks have a potential to be evolved by input vectors and include a series of sensory units (base neuron) comprising an input layer, one or more latent layers and an output layer.

The input signals are released as layer by layer through network and follow a forward route (Rezaei Arshad. 2010). The Trainlm (lm algorithm) is used as a training function because of its high speed and efficiency (Kia. 2010). Each network has its own function for implementation that in the multilayer perceptron this function is Newff. In this network, two latent layers are applied which in each one the transfer functions Tansing and Pulin has been used. It is through trial and error method, that is, replacement of transfer functions and modification of the number of neurons that the results can be obtained at best, having the highest correlation coefficient ($R$) and the least mean square error (MSE).
In the MATLAB software, 70 percent of both normalized output and input data were entered into network as training and the remaining 30 percent entered as test (the training is an observer one). If the value of predicated output is close to the actual one entered in to OUTPUT part, indicating the ideal predication of the network and ensure that the MSE is low and r is high. Therefore, the predicted output of the network is close to the target.

**Regression Model**

In order to examine the data based on regression, storing and normalization of the data was done by SPSS 17 (Al-Busod . 2010) packages. The relation between compression coefficient and the other properties was carried out by stepwise regression. To investigate being interdependent of the factors, the Pearson correlation coefficient was measured and the properties which had a significant correlation were specified.

**RESULTS AND DISCUSSION**

The results obtained by descriptive statistics based on properties data of interest are shown in table 1. The compression coefficient was mean 16%, at least 11% and at most 25% and they fall into clay soil category (Mahnaj. 2010).

<table>
<thead>
<tr>
<th>variable</th>
<th>C</th>
<th>Q</th>
<th>Sand</th>
<th>Clay</th>
<th>silt</th>
<th>pbd</th>
<th>pbw</th>
<th>Cc</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
</tr>
<tr>
<td>mean</td>
<td>0.32</td>
<td>8.01</td>
<td>6.82</td>
<td>77.72</td>
<td>14.9</td>
<td>1.58</td>
<td>1.93</td>
<td>0.16</td>
<td>20.7</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.06</td>
<td>1.3</td>
<td>2.45</td>
<td>3.48</td>
<td>2.71</td>
<td>0.06</td>
<td>0.05</td>
<td>0.03</td>
<td>2.16</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.15</td>
<td>0.67</td>
<td>0.52</td>
<td>-0.31</td>
<td>0.52</td>
<td>-0.2</td>
<td>-0.31</td>
<td>0.71</td>
<td>0.11</td>
</tr>
<tr>
<td>Elongation</td>
<td>-0.17</td>
<td>0.3</td>
<td>-0.6</td>
<td>-0.8</td>
<td>-0.64</td>
<td>0.17</td>
<td>-0.33</td>
<td>0.06</td>
<td>-0.83</td>
</tr>
<tr>
<td>Max</td>
<td>0.19</td>
<td>6</td>
<td>3</td>
<td>70</td>
<td>8</td>
<td>1.4</td>
<td>1.83</td>
<td>0.11</td>
<td>16.61</td>
</tr>
<tr>
<td>Min</td>
<td>0.45</td>
<td>11</td>
<td>13</td>
<td>85</td>
<td>19</td>
<td>1.73</td>
<td>2.01</td>
<td>0.24</td>
<td>25.55</td>
</tr>
</tbody>
</table>

C(kg/cm\(^2\)), Q(degree), Sand & Clay & Silt (%), pbd & pbw (gr/cm\(^3\)), PL(%)  

**Regression Results**

The results derived from stepwise regression between the properties are shown in the table 2.

**Table 2. Statistical Summary of data used for the multivariate regression compression coefficient**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>correlation coefficient (R)</th>
<th>Coefficient of determination R(^2)</th>
<th>Standardized coefficients</th>
<th>the least mean square error (MSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pb(_d)</td>
<td>0.474</td>
<td>0.224</td>
<td>-0.474</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Cc = -0.474 pb\(_d\) \(^2\)

On the basis of this equation it was revealed that the dry bulk specific gravity was the only factor which has a significant impact on compression coefficient (Cc). For this reason, the efficiency of regression model was examined for the data of interest. The correlation coefficient (R) and mean square error (MSE) for dry bulk specific gravity were 0.474 and 0.002 respectively. The results of table 2 showed that about 22% of compression coefficient variation is interpretable by the value of dry bulk specific gravity.

Abbasi (2012) in his investigation on appraisal of empirical methods for determination of fine soil compression coefficient found that the soil compression coefficient has a significant correlation with two groups of soil physical specification (viscosity coefficient, friction angle) and the specification related to initial soil structure (bulk specific gravity) and dry bulk specific gravity has much significant correlation with compression coefficient and root mean square error (RSME) equal to 0.060 (Jemsi. 2011). Albosda and co-workers (2010) in their research indicated that the amount of soil compression coefficient has a significant correlation with other soil physical properties and their linear regression model is predictable by the value of compression coefficient with correlation coefficient 0.5 (Al-Busod. 2010).

**Table 3. Specs of Neural Network of compression coefficient**

<table>
<thead>
<tr>
<th>Training Network</th>
<th>Testing</th>
<th>Training</th>
<th>Transfer functions</th>
<th>MSE</th>
<th>Network architecture</th>
<th>Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trainlm</td>
<td>0.63</td>
<td>0.79</td>
<td>Tansig Purlin</td>
<td>0.006</td>
<td>1,1,10,8</td>
<td>MLP</td>
</tr>
</tbody>
</table>
The selected optimal architecture by using trial and error method for this network (table 3), have 8 neurons (wet bulk specific gravity, friction, angle, plastic limit, percentage of sand particles, clay and silt, wet and dry bulk specific gravities) in the input layer, 10 neurons in the latent layer with Sigmund tangent threshold function and a neuron in the external layer with linear threshold function. The values of measured compression coefficient against estimated values by neural network are shown in the tables (1) and (2).

Daryaei and co-workers (2011), in their study on compression coefficient and other physical specification of soil by neural network have reported the coefficient of determination and mean square error were 0.02 and 0.76 respectively.

Kalantari pour and co-workers (2012) in their investigation indicated that the compression coefficient in the artificial neural network for coefficient of determination ($R^2$) and mean square error were 0.033 and 0.98 respectively (Kalntary. 2012). Also in their studies, fani Komar and co-workers (2011) showed that the neural network is able to predicate the compression coefficient with coefficient of determination ($R^2$) and mean square error were .97 and 0.0002 respectively.

Figure 1. The relation between estimated compression coefficients by neural network in the test stage.

Figure 2. The relation between estimated compression coefficients by neural network in the training stage.

**Neural Network Results**
The table 3 shows the results of evaluation of neural network MLP for predicating the degree of compression coefficient. The results indicated that the model has a correlation coefficient (R) 0.63 in the test stage and 0.79 in the training stage and the mean square error (MSE) 0.006 was measure.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>multiple regression</th>
<th>Neural network in the training stage.</th>
</tr>
</thead>
<tbody>
<tr>
<td>compression coefficient (C&lt;sub&gt;c&lt;/sub&gt;)</td>
<td>0.47</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>0.002</td>
<td>0.006</td>
</tr>
</tbody>
</table>

The table (4) indicates the comparison of both artificial neural network and linear multivariate regression. The results showed that the values of correlation coefficient (R) and mean square error (MSE) for compression coefficient were determined as 0.47 and 0.002 respectively and artificial neural network were 0.63 and 0.006. On this basis with regard to the value of correlation coefficient in artificial neural network, compared to regression, the neural network has a greater capability for predicating of this property.

**CONCLUSION**

The results of this research showed that the neural network has more capability than multivariate in establishing a logical relation among various data. There is a much higher speed connection between data and probability of error between connection of logical data is less. Also according to the modeling could explain To increase the accuracy of neural networks, there is need for more data.

**REFERENCES**

Abbasi N. 2011. Experimental Methods to evaluate the Fine Fraction Compression Coefficient. 4th International Conference of Soil Mechanics and Geotechnical Engineering. Iran


Kia M. 2010 Calculation on MATLAB. Kian Rayaneh Sabz publisher


Mahm M. 2010. Foundation of Neural Networks. 7th edition. Amirkabir publisher


SPSS Inc. 2007. SPSS IG Manual.