Estimation of swell index of fine grained soils using regression equations and artificial neural networks

Nihat Sinan Işık

Department of Construction Education, Gazi University, 06500, Beşevler, Ankara, Turkey. E-mail: nihatsinan@gazi.edu.tr. Tel: +903122028879. Fax: +903122120059.

Accepted 27 August, 2009

The swelling index which is the slope of the rebound curve of void ratio versus the logarithm of the effective pressure curve is used to estimate the consolidation settlement of over-consolidated fine grained soils. Because determination of swelling index from oedometer tests takes a relatively long time, empirical equations involving index soil properties, are needed to estimate it for preliminary calculations and to control the validity of consolidation tests. Geotechnical engineering literature involves empirical equations for the estimation of compression and swelling indexes. In this study the performance of widely used empirical equations were assessed using a database consisting of 42 test data. In addition to this, new empirical relationships with single and multiple dependent variables were developed with better estimation capability. An artificial neural network (ANN) which has two input variables, one hidden layer and eight hidden layer nodes was also developed to estimate swelling index. It was concluded that the performance of the ANN is better than empirical equations.

Key words: Swell index, clay soils, regression analyses, artificial neural networks.

INTRODUCTION

The compression index \((C_c)\) represents the slope of the curve of void ratio versus logarithm of effective pressure beyond maximum past effective stress and the swell index \((C_s)\) represents the slope of the rebound curve of void ratio versus logarithm of effective pressure. Compression index and swell index are used for the calculation of consolidation settlement of overconsolidated fine grained soils and they are conventionally determined by laboratory oedometer tests. However, the duration of consolidation tests is very long compared to standard index tests. For this reason, it is important to estimate compression and swell indexes with reasonable accuracy for preliminary calculations and to control the validity of consolidation tests. A large number of empirical equations are present in the geotechnical literature for the estimation of compression index (Skempton, 1944; Helenelund, 1951; Cozzolino, 1961; Sowers, 1970; Wroth and Wood, 1978; Carrier, 1985; Nagaraj and Murthy, 1986; Nakase et al., 1988; Bowles, 1989; Yin, 1999; Sridharan and Nagaraj, 2000; Giasi et al., 2003; Yoon et al., 2004; Ozer et al., 2008) however only two widely used equations are present for the estimation of swell index.

The purpose of this study is to compare the performances of widely used empirical swell index equations and to develop new empirical equations and a neural network based estimation technique for swell index by using the results of conventional oedometer and index test results.

DATABASE COMPILATION

In order to build the database, 42 conventional oedometer tests were performed according to ASTM 2435 (ASTM, 1996) on various undisturbed clay samples which were taken from various locations in Turkey. In addition to oedometer tests, index tests were performed on each sample according to relevant ASTM standards. Soil parameters used in the database were natural water content \((w_n)\), natural unit weight \((γ_n)\), dry unit weight \((γ_d)\), percent of soil passing from No. 200 sieve, percent of soil passing from No. 4 sieve, liquid limit \((LL)\), plastic limit \((PL)\), initial void ratio \((e_0)\), specific gravity of soil particles \((G_s)\), saturation ratio \((S_r)\), compression index \((C_c)\) and swell index \((C_s)\).

In order to test and obtain an empirical equation which is valid for all clay soils, the database should include a wide range of adequate data. In order to assess the ade-
Table 1. Descriptive statistics of variables used in the database.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_n$ (%)</td>
<td>15.1</td>
<td>62.0</td>
<td>31.2</td>
<td>10.6</td>
</tr>
<tr>
<td>$\gamma_n$ (t/m$^3$)</td>
<td>1.705</td>
<td>2.167</td>
<td>1.936</td>
<td>0.124</td>
</tr>
<tr>
<td>$\gamma_d$ (t/m$^3$)</td>
<td>1.105</td>
<td>1.856</td>
<td>1.489</td>
<td>0.197</td>
</tr>
<tr>
<td>No. 200 p (%)</td>
<td>42.02</td>
<td>93.13</td>
<td>74.42</td>
<td>12.29</td>
</tr>
<tr>
<td>No. 4 p (%)</td>
<td>90.98</td>
<td>100.0</td>
<td>97.57</td>
<td>2.62</td>
</tr>
<tr>
<td>LL (%)</td>
<td>23.90</td>
<td>102.6</td>
<td>56.72</td>
<td>18.96</td>
</tr>
<tr>
<td>PL (%)</td>
<td>0 (N.P)</td>
<td>45.30</td>
<td>27.15</td>
<td>8.68</td>
</tr>
<tr>
<td>PI (%)</td>
<td>0 (N.P)</td>
<td>69.55</td>
<td>29.02</td>
<td>13.43</td>
</tr>
<tr>
<td>$e_0$</td>
<td>0.497</td>
<td>1.780</td>
<td>0.930</td>
<td>0.297</td>
</tr>
<tr>
<td>$G_s$</td>
<td>2.560</td>
<td>2.740</td>
<td>2.641</td>
<td>0.046</td>
</tr>
<tr>
<td>$S_r$ (%)</td>
<td>59.5</td>
<td>100.0</td>
<td>86.2</td>
<td>10.3</td>
</tr>
<tr>
<td>$C_c$</td>
<td>0.0922</td>
<td>0.5671</td>
<td>0.2189</td>
<td>0.0929</td>
</tr>
<tr>
<td>$C_s$</td>
<td>0.0147</td>
<td>0.1294</td>
<td>0.0459</td>
<td>0.0246</td>
</tr>
</tbody>
</table>

Figure 1. Histogram of swell index values.

According to Table 1, it can be concluded that the database consists of a wide range of data. Therefore, this database can be used for the comparison of the performance of existing empirical equations and for the development of new equations.

EXPERIMENTAL DETERMINATION OF SWELL INDEX AND PREVIOUS EMPIRICAL STUDIES

During undisturbed soil sampling, some degree of disturbance is unavoidable; this leads to a slightly remoulded state. Remoulded specimens will display some deviation on the $e$ - $\log p$ plot as observed in the laboratory from the actual behavior in the field. Schmertmann (1953) described a procedure to obtain field (virgin) consolidation curve from the laboratory consolidation curve. Figure 2 presents this procedure.

In this study compression index values were determined using the procedure described by Schmertmann (1953); therefore, they correspond to the virgin compression index and swell index values were calculated from the average slope of the rebound curve.

Settlement behavior of fine grained soils is dependent on the relative proportion of silt and clay fractions, with the highly colloidal clays having the larger equilibrium...
void ratios; the colloidal size particles (< 0.001 mm) with greater surface area per unit mass have the ability to attract larger amounts of water, since liquid limits are a measure of the water attracted to these clay particles, some correlation between liquid limit and the compression index would be expected (Al-Khafaji and Andersland, 1992). The amount of consolidation settlement of fine grained soils is dependent on the water absorption capacity of clay sized particles, the existing stress state, pre-consolidation pressure of the soil sample and to some extent the compressibility of soil grains. Therefore, any direct or indirect parameters which define these conditions should be related to swell index. Atterberg limits reflect the relative amount of clay sized particles and their mineralogy; initial void ratio of the soil is an indication of the existing stress state and the pre-consolidation pressure; natural water content is a measure of the water attracted to clay particles and free water present within the voids; percent of soil passing from No. 4 and No. 200 sieves reflects the grain size distribution of the soil and dry unit weight may be an indication of the compressibility of soil grains to some degree. The other weight volume relationship parameters, saturation ratio, specific gravity and natural unit weight are physically related to dry unit weight, natural water content and the void ratio of the soil.

There are very large numbers of empirical equations for the estimation of compression index. Published regression equations generally relate compression index to one variable, liquid limit, natural water content, or in-situ void ratio, the majority of these equations are linear in form (Al-Khafaji and Andersland, 1992). Herrero (1983) recommended multiple soil parameter models for the estimation of compression index. In contrast to compression index only two widely accepted empirical equations were developed for the estimation of swell index; these are Nagaraj and Murty (1985) (Equation 1) and Nakase et al. (1988) (Equation 2) equations.

\[
C_s = 0.0463 \left[ \frac{LL(\%)}{100} \right] G_s
\]  
(1)

\[
C_s = 0.00194(PI - 4.6)
\]  
(2)

In order to compare the performance of equations developed by Nagaraj and Murty (1985) and Nakase et al. (1988), root mean square error (RMSE) term was utilized (Equation 3). Compiled database was used to calculate estimated swell index values and the root mean square error (RMSE) term of Nagaraj and Murty (1985) and Nakase et al. (1988) were calculated. RMSE of
Multiple linear regression analysis was performed using equations to estimate swell index, good predictor variables. Therefore, for the development of multi-parameter equations regression analyses were performed between parameters and swell index. Figures 4 to 8 present the regression equations developed for the index parameters and swell index, regression analyses were performed between $w_n$, LL, $e_0$, $C_s$, and swell index. In order to determine the relationship between the highest regression coefficient between the index parameters and swell index, regression analyses were performed between $w_n$, LL, $e_0$, $\gamma_n$, $\gamma_d$, and swell index. Figures 4 to 8 present the regression equations between $w_n$, LL, $e_0$, $\gamma_n$, $\gamma_d$, and swell index.

According to Figures 4 to 8, $e_0$, $\gamma_d$, and $w_n$ seem to be good predictor variables. Therefore, for the development of multi-parameter equations to estimate swell index, multiple linear regression analysis was performed using $e_0$, $\gamma_d$, and $w_n$. Equation 4 was determined as a result of multiple linear regression analysis. Newly developed equations and their performance indices are presented in Table 3.

$$C_s = (-0.000319 \cdot w_n) - (0.027277 \cdot \gamma_d) + (0.064019 \cdot e_0) + 0.037$$  (4)

Figure 9 presents the performance comparison of the newly developed empirical equations listed in Table 3. Table 3 and Figure 9 suggest that performances of newly developed equations are almost similar but Equation 2 and multi linear equation (Equation 4) seem to have slightly better estimation capacity.

**ARTIFICIAL NEURAL NETWORK BASED SWELL INDEX ESTIMATION**

An artificial neural network (ANN) is a computational model which tries to simulate the functional aspects of biological neural networks. ANNs consist of connected artificial neurons and process information using a connectionist approach. The purpose of ANNs is to set a relationship between model inputs and outputs by continuously updating connection weights according to inputs - outputs. They can be used to model complex relationships between inputs and outputs or to find patterns in data. Complex relationships between inputs and outputs can be discovered by changing model structure and connection weights. In spite of these advantages ANNs have an important disadvantage; they are not transparent as a closed form equation. Neural network model development involves six main stages. These are: determination of input and output variables, grouping of database as training and validating datasets, determination of network structure, optimization of connection weights, stopping according to a predefined criteria and validation of the neural network.

In recent times, artificial neural networks (ANNs) have been applied to many geotechnical engineering tasks and have demonstrated some degree of success (Shahin et al., 2002). Ozer et al. (2008) were applied ANN for the estimation of compression index of clay bearing soils and they determined that ANN has a better estimation performance than regression equations. There are also applications of ANNs to civil engineering materials, for example Subasi (2009) applied ANN for the prediction of mechanical properties of cement containing class C fly ash.

For the development of ANN to estimate swell index, $w_n$ and $e_0$ were used as predictor variables. All input variables were scaled between 0 and 1 as recommended by Masters (1993). It is common practice to divide the available data into two subsets; a training set to construct the neural network model, and an independent validation set to estimate model performance in the deployed development (Twomey and Smith, 1997). Approximately 85% of the data was used for training and 15% was used for validation. Hornik et al. (1989) showed that a network with one hidden layer can approximate any continuous function.
Table 2. Correlation matrix of variables.

<table>
<thead>
<tr>
<th></th>
<th>( w_n )</th>
<th>( \gamma_n )</th>
<th>( \gamma_d )</th>
<th>No. 200 p</th>
<th>No. 4 p</th>
<th>LL</th>
<th>PL</th>
<th>PI</th>
<th>( e_0 )</th>
<th>( G_s )</th>
<th>( S_r )</th>
<th>( C_c )</th>
<th>( C_s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w_n )</td>
<td>1.00</td>
<td>-0.75</td>
<td>-0.94</td>
<td>0.30</td>
<td>0.13</td>
<td>0.52</td>
<td>0.53</td>
<td>0.45</td>
<td>0.93</td>
<td>0.13</td>
<td>0.08</td>
<td>0.73</td>
<td>0.78</td>
</tr>
<tr>
<td>( \gamma_n )</td>
<td>-0.75</td>
<td>1.00</td>
<td>0.92</td>
<td>-0.19</td>
<td>-0.14</td>
<td>-0.36</td>
<td>-0.37</td>
<td>-0.31</td>
<td>-0.73</td>
<td>-0.26</td>
<td>-0.07</td>
<td>-0.69</td>
<td>-0.71</td>
</tr>
<tr>
<td>( \gamma_d )</td>
<td>-0.94</td>
<td>0.92</td>
<td>1.00</td>
<td>-0.29</td>
<td>-0.15</td>
<td>-0.48</td>
<td>-0.51</td>
<td>-0.41</td>
<td>-0.89</td>
<td>-0.18</td>
<td>-0.11</td>
<td>-0.74</td>
<td>-0.77</td>
</tr>
<tr>
<td>No. 200 p</td>
<td>0.30</td>
<td>-0.19</td>
<td>-0.29</td>
<td>1.00</td>
<td>0.51</td>
<td>0.51</td>
<td>0.62</td>
<td>0.44</td>
<td>0.36</td>
<td>-0.13</td>
<td>-0.12</td>
<td>0.40</td>
<td>0.38</td>
</tr>
<tr>
<td>No. 4 p</td>
<td>0.13</td>
<td>-0.14</td>
<td>-0.15</td>
<td>0.51</td>
<td>1.00</td>
<td>0.01</td>
<td>0.04</td>
<td>0.06</td>
<td>0.21</td>
<td>0.12</td>
<td>-0.23</td>
<td>0.22</td>
<td>0.09</td>
</tr>
<tr>
<td>LL</td>
<td>0.52</td>
<td>-0.36</td>
<td>-0.48</td>
<td>0.51</td>
<td>0.01</td>
<td>1.00</td>
<td>0.83</td>
<td>0.95</td>
<td>0.58</td>
<td>-0.20</td>
<td>-0.14</td>
<td>0.48</td>
<td>0.54</td>
</tr>
<tr>
<td>PL</td>
<td>0.53</td>
<td>-0.37</td>
<td>-0.51</td>
<td>0.62</td>
<td>0.04</td>
<td>0.83</td>
<td>1.00</td>
<td>0.67</td>
<td>0.59</td>
<td>-0.21</td>
<td>-0.16</td>
<td>0.40</td>
<td>0.51</td>
</tr>
<tr>
<td>PI</td>
<td>0.45</td>
<td>-0.31</td>
<td>-0.41</td>
<td>0.44</td>
<td>0.06</td>
<td>0.95</td>
<td>0.67</td>
<td>1.00</td>
<td>0.51</td>
<td>-0.21</td>
<td>-0.13</td>
<td>0.47</td>
<td>0.49</td>
</tr>
<tr>
<td>( e_0 )</td>
<td>0.93</td>
<td>-0.73</td>
<td>-0.89</td>
<td>0.36</td>
<td>0.21</td>
<td>0.58</td>
<td>0.59</td>
<td>0.51</td>
<td>1.00</td>
<td>0.16</td>
<td>-0.23</td>
<td>0.82</td>
<td>0.84</td>
</tr>
<tr>
<td>( G_s )</td>
<td>0.13</td>
<td>-0.26</td>
<td>-0.18</td>
<td>-0.13</td>
<td>0.12</td>
<td>-0.20</td>
<td>-0.21</td>
<td>-0.21</td>
<td>0.16</td>
<td>1.00</td>
<td>-0.07</td>
<td>0.25</td>
<td>0.05</td>
</tr>
<tr>
<td>( S_r )</td>
<td>0.08</td>
<td>-0.07</td>
<td>-0.11</td>
<td>-0.12</td>
<td>-0.23</td>
<td>-0.14</td>
<td>-0.16</td>
<td>-0.13</td>
<td>-0.23</td>
<td>-0.07</td>
<td>1.00</td>
<td>-0.23</td>
<td>-0.21</td>
</tr>
<tr>
<td>( C_c )</td>
<td>0.73</td>
<td>-0.69</td>
<td>-0.74</td>
<td>0.40</td>
<td>0.22</td>
<td>0.48</td>
<td>0.40</td>
<td>0.47</td>
<td>0.82</td>
<td>0.25</td>
<td>-0.23</td>
<td>1.00</td>
<td>0.82</td>
</tr>
<tr>
<td>( C_s )</td>
<td>0.78</td>
<td>-0.71</td>
<td>-0.77</td>
<td>0.38</td>
<td>0.09</td>
<td>0.54</td>
<td>0.51</td>
<td>0.49</td>
<td>0.84</td>
<td>0.05</td>
<td>-0.21</td>
<td>0.82</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Figure 4. Regression equation between \( w_n \) and \( C_s \).

\[
C_s = 0.0133e^{0.033w_n}
\]

\( R^2 = 0.6187 \)

Figure 5. Linear regression equation between LL and \( C_s \).

\[
C_s = 0.0007LL + 0.0062
\]

\( R^2 = 0.2914 \)
Figure 6. Regression equation between $e_0$ and $C_s$.

$C_s = 0.0121e^{1.3131e_0}$

$R^2 = 0.6501$

Figure 7. Regression equation between $\gamma_n$ and $C_s$.

$C_s = 9.3158e^{-2.8048\gamma_n}$

$R^2 = 0.5228$
function provided that sufficient connection weights are used; therefore, in this study a network with one hidden layer was used and the number of hidden layer nodes was increased until a successful ANN was achieved. Neural networks with back propagation algorithm are the most widely used method (Rumelhart et al., 1986). Therefore, in this study back propagation algorithm is used during training with 0.6 momentum and 0.8 learning rate. The training process was stopped when the average error of training data was below 10%. Figure 10 depicts the architecture of the neural network for the prediction of swell index and the relative connection weights.

Figure 11 displays the laboratory determined scaled swell index values versus ANN estimated scaled swell index values of (a) training and (b) validating data. RMSE of the training and validation data was calculated as 0.0113, this is lower than the RMSE of best empirical equations. According to RMSE of the training and validation data and the Figure 12, it can be concluded that the ANN’s prediction performance is better than empirical equations.

**SUMMARY AND CONCLUSIONS**

In this study the performance of widely used empirical equations for the estimation of swell index was assessed using the database consisting of 42 laboratory test data. Results indicate that Nagaraj and Murty’s (1985) equation using liquid limit and specific gravity of soil particles ge-

**Table 3.** Newly developed equations and their performance indices.

<table>
<thead>
<tr>
<th>Number</th>
<th>Equation</th>
<th>RMSE</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$C_s = 0.0133e^{0.036w_n}$</td>
<td>0.0144</td>
<td>0.6187</td>
</tr>
<tr>
<td>2</td>
<td>$C_s = 0.0121e^{-3.313c_n}$</td>
<td>0.0115</td>
<td>0.6501</td>
</tr>
<tr>
<td>3</td>
<td>$C_s = 0.1257\gamma_d^{-2.8826}$</td>
<td>0.0134</td>
<td>0.6532</td>
</tr>
<tr>
<td>4</td>
<td>$C_s = (-0.000319.w_n) - (0.027277.\gamma_d) + (0.064019.e_0) + 0.037$</td>
<td>0.0131</td>
<td>0.6857</td>
</tr>
</tbody>
</table>

* Unit of $\gamma_n$ is t/m$^3$. 

![Figure 8. Regression equation between $\gamma_d$ and $C_s$.](image-url)

![Figure 9. Performance comparison of the new empirical equations listed on Table 3.](image-url)
generally overestimate the swell index.

Statistical analyses performed on the compiled database revealed that there is a strong positive relationship between $w_n$, $LL$, $e_0$, $C_c$ and swell index and there is strong negative relationship between $\gamma_n$, $\gamma_d$ and swell index. Single variable regression equations for the estimation of swell index were developed using $w_n$, $LL$, $e_0$, $\gamma_n$, $\gamma_d$ as predictor variables. Among these equations regression equation using $e_0$ has the best estimation performance. In addition to single variable regression equations one multi variable equation to estimate swell index was developed using multiple linear regression analysis.

Because artificial neural networks have significant advantages over traditional regression methods such as

Figure 10. The architecture of the neural network for the prediction of swell index.
being more flexible and being able to discover more complex relationships with less effort, an artificial neural network using $w_n$ and $e_0$ as predictor variables was developed for the estimation of swell index. This ANN has one hidden layer and eight hidden layer nodes. Back propagation algorithm was used to train the network. RMSE of the training and validation and data was calculated as 0.0113, which is lower than the RMSE’s of regression equations. Therefore, it can be concluded the ANN’s prediction performance is better than regression equations.

REFERENCES


